Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} = Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/~pulino/ALESA/

Janeiro de 2012

Conteúdo

1

Estruturas Algébricas

Conteúdo

1.1 Operação Binária. Grupos

Definição 1.1.1 (Operação) *Seja* $\mathbb{E}$ *um conjunto não vazio. Uma operação binária em* $\mathbb{E}$ *é uma aplicação, que denotamos por* $\star$*, que a cada par ordenado* $(x, y) \in \mathbb{E} \times \mathbb{E}$ *associa o elemento* $x \star y \in \mathbb{E}$.

Definição 1.1.2 (Fechamento) *Seja* $\star$ *uma operação binária sobre* $\mathbb{E}$*. Dizemos que o subconjunto* $\mathcal{A} \subseteq \mathbb{E}$*, não vazio, é* ***fechado*** *com relação a operação* $\star$ *se para todo par ordenado* $(x, y) \in \mathcal{A} \times \mathcal{A}$ *tem–se que* $x \star y \in \mathcal{A}$.

Exemplo 1.1.1 *Considere* $\mathbb{N} = \{ 1, 2, 3, 4, \cdots \}$*, o conjunto dos números naturais. Podemos verificar facilmente que* $\mathbb{N}$ *é fechado com relação a operação de adição* $+$ *e também com relação a operação de multiplicação* $\times$.

Exemplo 1.1.2 *Considere* $\mathbb{Z} = \{ \cdots, -2, -1, 0, 1, 2, \cdots \}$*, o conjunto dos números inteiros. Podemos verificar facilmente que* $\mathbb{Z}$ *é fechado com relação a operação de adição* $+$ *e também com relação a operação de multiplicação* $\times$.

Definição 1.1.3 *Uma operação* $\star$ *definida em* $\mathbb{E}$ *pode ter as seguintes* ***propriedades****:*

- *Dizemos que uma operação* $\star$ *definida em* $\mathbb{E}$ *é* **associativa** *se* $\forall\, x,\ y,\ z \in \mathbb{E}$ *tem–se que* $(x \star y) \star z = x \star (y \star z)$.
- *Dizemos que uma operação* $\star$ *definida em* $\mathbb{E}$ *é* **comutativa** *se* $\forall\, x,\ y \in \mathbb{E}$ *tem–se que* $x \star y = y \star x$.
- *Dizemos que o elemento* $e \in \mathbb{E}$ *é o* **elemento neutro** *da operação* $\star$ *se* $\forall\, x \in \mathbb{E}$ *tem–se que* $x \star e = e \star x = x$.
- *Dizemos que o elemento* $x \in \mathbb{E}$ *é* **simetrizável** *para uma operação* $\star$ *com o elemento neutro* e *se existe um elemento* $\overline{x} \in E$ *tal que* $x \star \overline{x} = \overline{x} \star x = e$.

Exemplo 1.1.3 *Considere* $\mathbb{E} = \mathbb{R}$ *o conjunto dos números reais. Vamos definir em* $\mathbb{E}$ *uma operação* $\star$ *da seguinte forma: para cada par ordenado* $(x, y) \in \mathbb{R} \times \mathbb{R}$ *associamos o elemento* $x \star y = (x + y) + (x \times y)$*. Podemos mostrar facilmente que a operação* $\star$ *é associativa, comutativa e possui elemento neutro.*

Exemplo 1.1.4 *Podemos verificar que o conjunto dos números inteiros* $\mathbb{Z}$ *possui uma estrutura de grupo com relação a operação usual de adição.*

Definição 1.1.4 (Grupo) *Seja* $\mathbb{G}$ *um conjunto não vazio e* $\star$ *uma operação definida sobre* $\mathbb{G}$*. Dizemos que* $\mathbb{G}$ *tem uma estrutura de grupo em relação a operação* $\star$ *se essa operação possui as seguintes propriedades:*

(a) *Associativa*

(b) *Elemento Neutro*

(c) *Todo elemento de* $\mathbb{G}$ *é simetrizável*

Utilizamos a notação $(\mathbb{G}, \star)$ para denotar que o conjunto $\mathbb{G}$ tem uma operação $\star$ definida nele. Se a operação definida no grupo $\mathbb{G}$ for a adição, dizemos que $(\mathbb{G}, +)$ é um **grupo aditivo**. Se a operação definida no grupo $\mathbb{G}$ for a multiplicação, dizemos que $(\mathbb{G}, \times)$ é um **grupo multiplicativo**. No entanto, existem grupos com outras operações em vários ramos da matemática.

Exemplo 1.1.5 *Podemos verificar que a Tabela 1.1 proporciona uma estrutura de grupo ao conjunto* $\mathbb{G} = \{ e, a, b, c \}$*, isto é,* $(\mathbb{G}, \star)$ *é um grupo. A regra de operação na tabela é definida de forma que o elemento* $a_{ij} = a_i \star a_j$*.*

Tabela 1.1: Grupo de Klein

$\star$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Definição 1.1.5 (Grupo Comutativo) *Seja* $(\mathbb{G}, \star)$ *um grupo. Dizemos que* $(\mathbb{G}, \star)$ *é um grupo comutativo, ou grupo abeliano, se a operação* $\star$ *for comutativa.*

Exemplo 1.1.6 *O conjunto* $\mathbb{Z}$ *dos números inteiros possui uma estrutura de grupo abeliano em relação a operação usual de adição. Assim, dizemos que* $(\mathbb{Z}, +)$ *é um grupo aditivo abeliano.*

Exemplo 1.1.7 *O conjunto das matrizes reais de ordem* n*, que denotamos por* $\mathbb{M}_n(\mathbb{R})$*, tem uma estrutura de grupo aditivo abeliano, isto é,* $(\mathbb{M}_n(\mathbb{R}), +)$ *tem uma estrutura de grupo abeliano, onde* $+$ *indica a operação usual de adição de matrizes.*

Exemplo 1.1.8 *O conjunto* $\mathbb{Q}(\sqrt{2}) = \{ a + b\sqrt{2} \ / \ a, b \in \mathbb{Q} \}$ *tem uma estrutura de grupo multiplicativo abeliano, com relação a operação usual de multiplicação dos números racionais.*

Exemplo 1.1.9 *Considere o subconjunto* $S \subset \mathbb{M}_n(\mathbb{R})$ *definido por:*

$$S = \{ D \in \mathbb{M}_n(\mathbb{R}) \ / \ D \text{ é uma matriz diagonal invertível} \}.$$

Mostre que $(S, \star)$ *tem uma estrutura de grupo multiplicativo abeliano, onde* $\star$ *é a operação usual de multiplicação de matrizes.*

Definição 1.1.6 (Subgrupo) *Seja* $(\mathbb{G}, \star)$ *um grupo. Dizemos que um subconjunto* $\mathbb{S} \subset \mathbb{G}$ *não vazio é um subgrupo de* $\mathbb{G}$ *se* $\mathbb{S}$ *for fechado com relação a operação* $\star$ *e* $(\mathbb{S}, \star)$ *tem uma estrutura de grupo.*

Exemplo 1.1.10 *Dado* $n \in \mathbb{N}$, *o subconjunto* $\mathbb{Z}\mathbf{n} \subset \mathbb{Z}$ *definido da forma:*

$$\mathbb{Z}\mathbf{n} = \{ x \in \mathbb{Z} \ / \ x = n \times m \quad ; \quad m \in \mathbb{Z} \}$$

é um subgrupo do grupo aditivo $(\mathbb{Z}, +)$, *onde* $\times$ *é a operação de multiplicação em* $\mathbb{Z}$.

Exercícios

Exercício 1.1 *Verifique se* $(E, \star)$ *tem uma estrutura de grupo abeliano. Em caso negativo, dizer quais propriedades não são satisfeitas.*

(a) $E = I\!N_0 = \{\, 0, 1, 2 \cdots \}$ *e* $x \star y = x + y$

(b) $E = \mathbb{Z}$ *e* $x \star y = x + y - 1$

(c) $E = \mathbb{Z}$ *e* $x \star y = x + y + 1$

(d) $E = \mathbb{Z}$ *e* $x \star y = 2 \times x + y$

(e) $E = \mathbb{Z}$ *e* $x \star y = x \times y$

onde $+$ *indica a operação usual de adição e* $\times$ *indica a operação usual de multiplicação.*

Exercício 1.2 *Considere o conjunto dos números reais* $I\!R$ *munido da operação* $\star$ *definida por* $x \star y = x + y + 4$*. Mostre que* $(I\!R, \star)$ *tem uma estrutura de grupo comutativo.*

Exercício 1.3 *Considere o conjunto dos números reais* $I\!R$ *munido da operação* $\star$ *definida por* $x \star y = x + 2 \times y - 4$*. Verifique se* $(I\!R, \star)$ *tem uma estrutura de grupo comutativo. Em caso negativo, dizer quais propriedades não são satisfeitas.*

Exercício 1.4 *Considere o conjunto dos números reais positivos* $I\!R^+$*, isto é,*

$$I\!R^+ = \{\, x \in I\!R \;/\; x > 0 \,\},$$

munido da operação $\star$ *definida por* $x \star y = x + y + 8$*. Verifique se* $(I\!R^+, \star)$ *possui uma estrutura de grupo comutativo.*

Exercício 1.5 *Considere o conjunto dos números reais positivos* $I\!R^+$ *munido da operação* $\star$ *definida por* $x \star y = x + y - 6$*. Verifique se* $(I\!R^+, \star)$ *possui uma estrutura de grupo comutativo.*

Exercício 1.6 *Considere o subconjunto* $I\!R^*$ *dos números reais definido por:*

$$I\!R^* = \{\, x \in I\!R \;/\; x \neq 0 \,\},$$

Mostre que $(I\!R^*, \times)$ *possui uma estrutura de grupo multiplicativo abeliano, com relação a operação usual de multiplicação.*

Exercício 1.7 *Verifique se* $(\mathbb{M}_n(\mathbb{R}), \star)$ *tem uma estrutura de grupo multiplicativo, onde* $\star$ *é a operação usual de multiplicação de matrizes.*

Exercício 1.8 *Seja* $S = \{ A \in \mathbb{M}_n(\mathbb{R}) \ / \ A$ *é invertível*$\}$. *Verifique se* $(S, \star)$ *tem uma estrutura de grupo multiplicativo, onde* $\star$ *é a operação usual de multiplicação de matrizes.*

Exercício 1.9 *Seja* $S = \{ A \in \mathbb{M}_n(\mathbb{R}) \ / \ A$ *é invertível*$\}$. *Verifique se* $(S, \star)$ *tem uma estrutura de grupo multiplicativo abeliano, onde* $\star$ *é a operação usual de multiplicação de matrizes.*

Exercício 1.10 *Considere o conjunto dos números reais* $\mathbb{R}$. *Vamos definir em* $\mathbb{R}$ *uma operação* $\star$ *da seguinte forma: para cada par ordenado* $(x, y) \in \mathbb{R} \times \mathbb{R}$ *associamos o elemento* $x \star y = x + (x \times y)$. *Faça um estudo sobre as propriedades da operação* $\star$, *onde* $+$ *indica a operação usual de adição e* $\times$ *indica a operação usual de multiplicação.*

Exercício 1.11 *Mostre que* $(\mathbb{Z}, \times)$, *onde* $\times$ *é a operação usual de multiplicação de números inteiros, não possui uma estrutura de grupo multiplicativo.*

Exercício 1.12 *Considere o conjunto* $G = \{ e, a, b, c \}$. *Se* $(G, \star)$ *tem uma estrutura de grupo, complete a tabela abaixo.*

$\star$	e	a	b	c
e	e	a	b	c
a	a			
b	b	c		
c	c	e	a	

A regra de operação na tabela é de forma que o elemento $a_{ij} = a_i \star a_j$.

Exercício 1.13 *Considere o conjunto dos números racionais*

$$\mathbb{Q} = \left\{ \frac{p}{q} \ / \ p, q \in \mathbb{Z} \quad , \quad q \neq 0 \right\}.$$

Mostre que $(\mathbb{Q}, +)$ *possui uma estrutura de grupo aditivo abeliano, com relação a operação usual de adição.*

1.2 Corpo Comutativo

Definição 1.2.1 *Um **corpo comutativo** é um conjunto não vazio $\mathbb{F}$ munido de duas operações, denominadas **adição** e **multiplicação**, que vamos denotar por $+$ e $\times$, respectivamente, que satisfazem os seguintes axiomas:*

Axiomas de Fechamento

(F_1) *$\mathbb{F}$ é fechado com relação a operação de adição. Para todos $x,\ y \in \mathbb{F}$ temos que $x + y \in \mathbb{F}$.*

(F_2) *$\mathbb{F}$ é fechado com relação a operação de multiplicação. Para todos $x,\ y \in \mathbb{F}$ temos que $x \times y \in \mathbb{F}$.*

Axiomas da Operação de Adição

(A_1) *Associatividade: para todos $x,\ y,\ z \in \mathbb{F}$ temos que $(x + y) + z = x + (y + z)$.*

(A_2) *Comutatividade: para todos $x,\ y \in \mathbb{F}$ temos que $x + y = y + x$.*

(A_3) *Elemento Neutro: existe um único elemento em $\mathbb{F}$, denotado por $0_{\mathbb{F}}$, tal que $x + 0_{\mathbb{F}} = x \quad ; \quad \forall\ x \in \mathbb{F}$.*

(A_4) *Elemento Simétrico: todo elemento $x \in \mathbb{F}$ possui um único elemento simétrico $(-x) \in \mathbb{F}$ tal que $x + (-x) = 0_{\mathbb{F}}$.*

Axiomas da Operação de Multiplicação

(M_1) *Associatividade: $\forall\ x,\ y,\ z \in \mathbb{F}$ temos que $(x \times y) \times z = x \times (y \times z)$.*

(M_2) *Comutatividade: $\forall\ x,\ y \in \mathbb{F}$ temos que $x \times y = y \times x$.*

(M_3) *Elemento Neutro: existe um único elemento em $\mathbb{F}$, denotado por $1_{\mathbb{F}}$, tal que $x \times 1_{\mathbb{F}} = x$ para todo $x \in \mathbb{F}$.*

(M_4) *Inverso Multiplicativo: todo elemento $x \in \mathbb{F}$ com $x \neq 0_{\mathbb{F}}$ possui um único elemento $x^{-1} \in \mathbb{F}$ tal que $x \times x^{-1} = 1_{\mathbb{F}}$.*

(D_1) *Distributividade: $\forall\ x,\ y,\ z \in \mathbb{F}$ temos que $x \times (y + z) = (x \times y) + (x \times z)$.*

É interessante observar que $(\mathbb{F}, +)$ tem uma estrutura de *grupo aditivo abeliano*. Seja o conjunto $\mathbb{F}^* = \{\ x \in \mathbb{F} \ / \ x \neq 0_{\mathbb{F}}\ \}$. Observamos que $(\mathbb{F}^*, \times)$ tem uma estrutura de *grupo multiplicativo abeliano*.

Definição 1.2.2 (subcorpo) *Seja* $\mathbb{F}$ *um corpo. Dizemos que um subconjunto* $\mathbb{S} \subset \mathbb{F}$*, não vazio, é um subcorpo de* $\mathbb{F}$ *se* $\mathbb{S}$ *possui uma estrutura de corpo com relação às operações de adição e multiplicação definidas em* $\mathbb{F}$*.*

Teorema 1.2.1 *Sejam* $a, b, c \in \mathbb{F}$ *com* $a + b = a + c$*. Então* $b = c$*.*

Demonstração – Pelo Axioma (A_4) de corpo, temos que existe um único elemento $d \in \mathbb{F}$ tal que $a + d = 0_{\mathbb{F}}$. Desse modo, obtemos

$$b = (a + d) + b = d + (a + b) = d + (a + c) = (a + d) + c = c \, ,$$

utilizando os Axiomas (A_1) e (A_3) de corpo, completando a demonstração. ■

Essa propriedade é denominada **Lei do Cancelamento para a Adição**. Em particular, essa propriedade mostra a unicidade do elemento neutro da adição.

Teorema 1.2.2 *Sejam* $a, b \in \mathbb{F}$*. Então, existe um único* $x \in \mathbb{F}$ *tal que* $a + x = b$*.*

Demonstração – Pelo Axioma (A_4) de corpo, temos que existe um único elemento $d \in \mathbb{F}$ tal que $a + d = 0_{\mathbb{F}}$. Seja o elemento $x = d + b$. Assim, temos que

$$a + x = a + (d + b) = (a + d) + b = 0_{\mathbb{F}} + b = b \, ,$$

utilizando os Axiomas (A_1) e (A_3) de corpo, completando a demonstração. ■

Denotamos o elemento $x = b - a$ para indicar a diferença entre os elementos a e b. Em particular, $0_{\mathbb{F}} - a$ é simplesmente escrito como $-a$ que é denominado negativo ou simétrico do elemento a. A operação que a cada par $(a, b) \in \mathbb{F} \times \mathbb{F} \longrightarrow a - b$ é denominada **subtração**.

Teorema 1.2.3 *Sejam* $a, b, c \in \mathbb{F}$ *com* $a \times b = a \times c$ *e* $a \neq 0_{\mathbb{F}}$*. Então* $b = c$*.*

Demonstração – Como $a \neq 0_{\mathbb{F}}$, pelo Axioma (M_4) de corpo, temos que existe um único elemento $d \in \mathbb{F}$ tal que $a \times d = 1_{\mathbb{F}}$. Desse modo, obtemos

$$b = (a \times d) \times b = d \times (a \times b) = d \times (a \times c) = (a \times d) \times c = c \, ,$$

utilizando os Axiomas (M_1) e (M_3) de corpo, completando a demonstração. ■

Essa propriedade é denominada **Lei do Cancelamento para a Multiplicação**. Em particular, essa propriedade mostra a unicidade do elemento neutro da multiplicação.

Teorema 1.2.4 *Sejam* $a, b \in I\!\!F$ *com* $a \neq 0_{I\!\!F}$. *Então, existe um único elemento* $x \in I\!\!F$ *tal que* $a \times x = b$.

Demonstração – Como $a \neq 0_{I\!\!F}$, pelo Axioma (M_4) de corpo, temos que existe um único elemento $d \in I\!\!F$ tal que $a \times d = 1_{I\!\!F}$. Seja $x = d \times b$. Assim, temos que

$$a \times x \;=\; a \times (d \times b) \;=\; (a \times d) \times b \;=\; b \,,$$

utilizando os Axiomas (M_1) e (M_3) de corpo, completando a demonstração. ■

Denotamos o elemento $x = \dfrac{b}{a}$ para indicar o quociente do elemento b pelo elemento a. Em particular, $\dfrac{1}{a}$ é simplesmente escrito como a^{-1} e é chamado recíproco do elemento a.

A operação que a cada par $(b, a) \longrightarrow \dfrac{b}{a}$, definida para todo elemento $b \in I\!\!F$ e $a \in I\!\!F$ não–nulo, é denominada **divisão**.

Teorema 1.2.5 *Sejam* $a, b, c \in I\!\!F$.

(a) $a \times 0_{I\!\!F} = 0_{I\!\!F}$.

(b) $(-a) \times b = a \times (-b) = -(a \times b)$.

(c) $(-a) \times (-b) = a \times b$.

Demonstração – *(a)* Utilizando o Axioma (D_1) de corpo, temos que

$$a \times 0_{I\!\!F} \;=\; 0_{I\!\!F} \;+\; a \times 0_{I\!\!F} \;=\; a \times (0_{I\!\!F} \;+\; 0_{I\!\!F}) \;=\; a \times 0_{I\!\!F} \;+\; a \times 0_{I\!\!F} \,.$$

Pela Lei do Cancelamento da Adição, obtemos $a \times 0_{I\!\!F} = 0_{I\!\!F}$.

(b) Pelo Axioma (A_4) de corpo, o elemento $-(a \times b)$ é o único elemento em $I\!\!F$ tal que $(a \times b) + (\,-(a \times b)\,) = 0_{I\!\!F}$. Além disso, o elemento $-a$ é o único elemento em $I\!\!F$ tal que $a + (-a) = 0_{I\!\!F}$. Desse modo, temos que

$$a \times b \;+\; (-a) \times b \;=\; (\,a \;+\; (-a)\,) \times b \;=\; 0_{I\!\!F} \times b \;=\; 0_{I\!\!F} \,.$$

Assim, mostramos que $(-a) \times b = -(a \times b)$. De modo análogo, podemos provar que $a \times (-b) = -(a \times b)$.

(c) Aplicando duas vezes o resultado do item *(b)*, obtemos

$$(-a) \times (-b) \;=\; -(\,a \times (-b)\,) \;=\; -(\,-(a \times b)\,) \;=\; a \times b \,,$$

que completa a demonstração. ■

Exemplo 1.2.1 *O conjunto dos números racionais*

$$\mathbb{Q} = \left\{ \frac{p}{q} \;/\; p, q \in \mathbb{Z} \quad , \quad q \neq 0 \right\}$$

tem uma estrutura de corpo, com as operações usuais de adição e multiplicação.

Exemplo 1.2.2 *O conjunto dos números inteiros* $\mathbb{Z} = \{ \cdots, -2, -1, 0, 1, 2, \cdots \}$ *não possui uma estrutura de corpo, pois o Axioma* (M_4)*, da definição de corpo, não é satisfeito, exceto para* $n = 1$ *ou* $n = -1$.

Exemplo 1.2.3 *O conjunto dos números reais, denotado por* $I\!R$*, tem uma estrutura de corpo, com as operações usuais de adição e multiplicação.*

1.3 Corpo com Valor Absoluto

Definição 1.3.1 *Seja* $I\!F$ *um corpo. Definimos o* ***valor absoluto*** *em* $I\!F$ *como sendo uma aplicação* $v(\cdot)$ *que associa a cada elemento* $x \in I\!F$ *um número real* $v(x)$*, que possui as seguintes propriedades:*

(a) $v(x) \geq 0$.

(b) $v(x) = 0$ *se, e somente se,* $x = 0_{I\!F}$.

(c) $v(x \times y) = v(x)v(y)$.

(d) $v(x + y) \leq v(x) + v(y)$.

Definição 1.3.2 *Seja* $I\!F$ *um corpo. A aplicação* $v(\cdot)$ *que associa a cada elemento* $x \in I\!F$ *um número real* $v(x)$ *definida por:*

$$v(x) = \begin{cases} 1 & se \quad x \neq 0_{I\!F} \\ 0 & se \quad x = 0_{I\!F} \end{cases}$$

é denominada ***valor absoluto trivial*** *em* $I\!F$.

Lema 1.3.1 *Sejam* $\mathbb{F}$ *um corpo com valor absoluto* $v(\cdot)$ *e* $x \in \mathbb{F}$ *tal que* $x^n = 1_{\mathbb{F}}$ *para todo inteiro positivo* n. *Então,* $v(x) = 1$.

Demonstração – Primeiramente vamos observar que $v(1_{\mathbb{F}}) = 1$. De fato,

$$(1_{\mathbb{F}})^2 = 1_{\mathbb{F}} \times 1_{\mathbb{F}} = 1_{\mathbb{F}} \quad \Longrightarrow \quad v((1_{\mathbb{F}})^2) = v(1_{\mathbb{F}})v(1_{\mathbb{F}}) = v(1_{\mathbb{F}})\,.$$

Assim, concluímos que $v(1_{\mathbb{F}}) = 0$ ou $v(1_{\mathbb{F}}) = 1$. Note que, se $v(1_{\mathbb{F}}) = 0$, então $v(a) = 0$ para todo $a \in \mathbb{F}$. Portanto, temos que $v(1_{\mathbb{F}}) = 1$.

Finalmente, tomando $x^n = 1_{\mathbb{F}}$, obtemos

$$v(x^n) = v(1_{\mathbb{F}}) \quad \Longrightarrow \quad (v(x))^n = 1 \quad \Longrightarrow \quad v(x) = 1\,,$$

o que completa a demonstração. ■

Lema 1.3.2 *Seja* $\mathbb{F}$ *um corpo com valor absoluto* $v(\cdot)$. *Então,*

$$v(-x) = v(x) \qquad \textit{para todo} \qquad x \in \mathbb{F}\,.$$

Demonstração – Sabemos que

$$(-1_{\mathbb{F}})^2 = -1_{\mathbb{F}} \times -1_{\mathbb{F}} = 1_{\mathbb{F}} \times 1_{\mathbb{F}} = 1_{\mathbb{F}}\,.$$

Portanto, obtemos $v(-1_{\mathbb{F}}) = 1$. Desse modo, temos que

$$v(-x) = v(-1_{\mathbb{F}} \times x) = v(-1_{\mathbb{F}})v(x) = v(x)\,,$$

o que completa a demonstração. ■

Lema 1.3.3 *Seja* $\mathbb{F}$ *um corpo com valor absoluto* $v(\cdot)$. *Então,*

$$v(x) - v(y) \leq v(x+y) \qquad \textit{para todos} \qquad x,\, y \in \mathbb{F}\,.$$

Demonstração – Considerando $x = x + y - y$ e a propriedade da desigualdade triangular, isto é,

$$v(x+y) \leq v(x) + v(y) \qquad \text{para todos} \qquad x,\, y \in \mathbb{F}\,,$$

obtemos

$$v(x) = v(x + y - y) \leq v(x + y) + v(-y) = v(x + y) + v(y)$$

Portanto, temos que $v(x) - v(y) \leq v(x+y)$. ■

1.4 Corpo Ordenado

Definição 1.4.1 *Um **corpo ordenado** é um corpo $I\!F$ com uma relação de ordem, que vamos denotar por $<$, e escrevemos $x < y$ para indicar que o elemento $x \in I\!F$ é **menor** que o elemento $y \in I\!F$, satisfazendo os seguintes axiomas:*

(O_1) Princípio da Comparação:
Se $x, y \in I\!F$, então uma e somente uma das seguintes relações é satisfeita:

$$x < y \ , \ y < x \ , \ x = y \ .$$

(O_2) Transitividade:
Se $x, y, z \in I\!F$, com $x < y$ e $y < z$, então $x < z$.

(O_3) Consistência da Adição com a Relação de Ordem:
se $x, y, z \in I\!F$ e $y < z$, então $x + y < x + z$.

(O_4) Consistência da Multiplicação com a Relação de Ordem:
Se $x, y \in I\!F$, com $0_{I\!F} < x$ e $0_{I\!F} < y$, então $0_{I\!F} < x \times y$.

Equivalentemente, podemos definir um corpo ordenado da forma a seguir.

Definição 1.4.2 *Seja $I\!F$ um corpo. Assumimos que existe um subconjunto $I\!F^+ \subset I\!F$, denominado **conjunto dos elementos positivos**, o qual satisfaz os seguinte axiomas:*

(O_1) Se $x, y \in I\!F^+$, então $x + y \in I\!F^+$ e $x \times y \in I\!F^+$.

(O_2) Para todo elemento $x \neq 0_{I\!F}$, temos que $x \in I\!F^+$ ou $-x \in I\!F^+$.

(O_3) O elemento neutro $0_{I\!F} \notin I\!F^+$.

Note que num corpo ordenado, se $x \neq 0_{I\!F}$, então $x^2 \in I\!F^+$. De fato, como o elemento $x \neq 0_{I\!F}$, temos que $x \in I\!F^+$ ou $-x \in I\!F^+$.

No primeiro caso, isto é, $x \in I\!F^+$, obtemos $x \times x = x^2 \in I\!F^+$.

No segundo caso, isto é, $-x \in I\!F^+$, obtemos

$$(-x) \times (-x) = -(\, x \times (-x)\,) = -(\, -(x \times x)\,) = x \times x = x^2 \in I\!F^+ \ .$$

Em particular, num corpo ordenado, o elemento $-1_{I\!F}$ não é o quadrado de nenhum elemento de $I\!F$.

Num corpo ordenado $\mathbb{F}$, podemos escrever $x < y$ para indicar que $y - x \in \mathbb{F}^+$, isto é, o elemento $y - x$ é **positivo**. De modo análogo, escrevemos $y > x$ para indicar que o elemento y é **maior** que o elemento x. Em particular, escrevemos $x > 0_{\mathbb{F}}$ para dizer que $x \in \mathbb{F}^+$, isto é, o elemento x é positivo. De mesmo modo, escrevemos $x < 0_{\mathbb{F}}$ para dizer que o elemento x é **negativo**, isto é, o elemento $-x \in \mathbb{F}^+$.

A partir dos axiomas da Definição 1.4.2 vamos mostrar os axiomas da Definição 1.4.1.

(O_1) Princípio da Comparação
Dados os elementos $x, y \in \mathbb{F}$. Pelo axioma *(O_2)* da Definição 1.4.2, temos as seguintes possibilidades:

$$(1)\ \ y - x \in \mathbb{F}^+ \qquad (2)\ \ -(y - x) \in \mathbb{F}^+ \qquad (3)\ \ y - x = 0_{\mathbb{F}}\ .$$

Assim, podemos concluir que ou $x < y$ ou $y < x$ ou $x = y$.

(O_2) Transitividade
Considere os elementos $x, y, z \in \mathbb{F}$ com $x < y$ e $y < z$. Assim, podemos afirmar que $y - x \in \mathbb{F}^+$ e $z - y \in \mathbb{F}^+$. Pelo axioma *(O_1)* da Definição 1.4.2, temos que o elemento $(y - x) + (z - y) \in \mathbb{F}^+$. Logo, o elemento $z - x \in \mathbb{F}^+$. Assim, podemos concluir que $x < z$.

(O_3) Consistência da Adição com a Relação de Ordem
Considere os elementos $x, y, z \in \mathbb{F}$ com $y < z$, isto é, o elemento $z - y \in \mathbb{F}^+$. Desse modo, temos que $z - y = (z + x) - (y + x) \in \mathbb{F}^+$. Assim, podemos concluir que $y + x < z + x$.

(O_4) Consistência da Multiplicação com a Relação de Ordem
Considere os elementos $x, y \in F$ com $0_{\mathbb{F}} < x$ e $0_{\mathbb{F}} < y$, isto é, $x \in \mathbb{F}^+$ e $y \in \mathbb{F}^+$. Logo, temos que $x \times y \in \mathbb{F}^+$, pelo axioma *($O_1$)* da Definição 1.4.2. Assim, podemos concluir que $x \times y > 0_{\mathbb{F}}$.

Portanto, tomando os elementos $x, y, z \in \mathbb{F}$ com $x < y$ e $0_{\mathbb{F}} < z$, isto é, $y - x \in \mathbb{F}^+$ e $z \in \mathbb{F}^+$. Pelo axioma *(O_1)* da Definição 1.4.2, temos que o elemento $(y - x) \times z \in \mathbb{F}^+$, isto é, o elemento $y \times z - x \times z \in \mathbb{F}^+$. Desse modo, podemos concluir que $x \times z < y \times z$.

De modo análogo, a partir dos axiomas da Definição 1.4.1 podemos obter os axiomas da Definição 1.4.2. Assim, essas definições são equivalentes. Para um estudo mais detalhado sobre corpos ordenados podemos consultar a referência [17].

Exemplo 1.4.1 *O conjunto dos números racionais*

$$\mathbb{Q} = \left\{ \frac{p}{q} \ / \ p, q \in \mathbb{Z} \quad , \quad q \neq 0 \right\}$$

é um corpo ordenado, com as operações usuais de adição e multiplicação. O conjunto dos números racionais positivos $\mathbb{Q}^+$ *é definido por:*

$$\mathbb{Q}^+ = \left\{ \frac{p}{q} \ / \ p, q \in I\!N \right\}.$$

Exemplo 1.4.2 *O conjunto*

$$\mathbb{Q}(\sqrt{2}) = \left\{ a + b\sqrt{2} \ / \ a, b \in \mathbb{Q} \right\}$$

é um corpo ordenado, com as operações usuais de adição e multiplicação.

Exemplo 1.4.3 *Considere o seguinte conjunto* $\mathbb{Z}_p = \{ 0, 1, 2, \cdots, (p-1) \}$, *onde* p *é um inteiro positivo, no qual definimos as operações:*

- ***Adição:***
  $a \oplus b = c$, *onde* c *é o resto da divisão da soma* $a + b$ *pelo inteiro* p, *isto é,*
  $$a + b = mp + c \qquad \textit{para algum} \qquad m \in I\!N \cup \{0\}.$$

- ***Multiplicação:***
  $a \otimes b = d$, *onde* d *é o resto da divisão do produto* ab *pelo inteiro* p, *isto é,*
  $$ab = mp + d \qquad \textit{para algum} \qquad m \in I\!N \cup \{0\}.$$

Como $\mathbb{Z}_p$ *deve ser fechado com relação as operações, temos que* $c, d \in \mathbb{Z}_p$.

Podemos mostrar que $\mathbb{Z}_p$ tem um estrutura de corpo quanto p é um número primo. Considere como exemplo $\mathbb{Z}_5 = \{ 0, 1, 2, 3, 4 \}$. Faça a verificação que $\mathbb{Z}_5$ satisfaz os axiomas de corpo.

Exemplo 1.4.4 *O corpo* $\mathbb{Z}_5 = \{ 0, 1, 2, 3, 4 \}$ *não é um corpo ordenado. De fato, tomando por exemplo* $2 + 3 = 0$ *em* $\mathbb{Z}_5$. *Entretanto, num corpo ordenado a soma de dois elementos positivos deve ser igual a um elemento positivo. Assim, mostramos que* $\mathbb{Z}_5$ *não comporta uma relação de ordem.*

1.5 Valor Absoluto num Corpo Ordenado

Definição 1.5.1 *Seja $\mathbb{F}$ um corpo ordenado. O* **valor absoluto** *em $\mathbb{F}$ é uma aplicação $|\cdot|$ que associa a cada elemento $x \in \mathbb{F}$ um número real $|x|$ definida por:*

$$|x| = \begin{cases} x & se \quad x > 0 \\ 0 & se \quad x = 0 \\ -x & se \quad x < 0 \end{cases}$$

Podemos observar que $|x|$ é escolhido o maior elemento entre x e $-x$. Logo, temos que $|x| \geq x$ e $|x| \geq -x$. Portanto, podemos concluir que

$$-|x| \leq x \leq |x|$$

para todo $x \in \mathbb{F}$.

Teorema 1.5.1 *Sejam $\mathbb{F}$ um corpo ordenado e os elementos $x, a \in \mathbb{F}$. As seguintes afirmações são equivalentes:*

(a) $-a \leq x \leq a$.

(b) $x \leq a$ *e* $-x \leq a$.

(c) $|x| \leq a$.

Demonstração – Temos que

$$\begin{aligned} -a \leq x \leq a \quad &\Longleftrightarrow \quad -a \leq x \ \text{ e } \ x \leq a \\ &\Longleftrightarrow \quad a \geq x \ \text{ e } \ a \geq -x \\ &\Longleftrightarrow \quad a \geq |x| \end{aligned}$$

a última equivalência vem do fato que $|x|$ é o maior elemento entre x e $-x$. ■

Corolário 1.5.1 *Sejam $a, b, x \in \mathbb{F}$. Então,*

$$|x - a| \leq b \quad \textit{se, e somente se,} \quad a - b \leq x \leq a + b.$$

Demonstração – Pelo Teorema 1.5.1, temos que

$$|x - a| \leq b \iff -b \leq x - a \leq b \iff -b + a \leq x \leq b + a,$$

o que completa a demonstração. ■

Teorema 1.5.2 *Sejam* $x, y \in \mathbb{F}$. *Então,* $|\,x \times y\,| = |\,x\,|\,|\,y\,|$.

Demonstração – Primeiramente vamos observar que $x^2 = |\,x\,|^2$, para todo $x \in \mathbb{R}$, pois $|\,x\,| = x$ ou $|\,x\,| = -x$ e vale $x^2 = (-x)^2$.

Desse modo, temos que

$$|\,x \times y\,|^2 = (x \times y)^2 = x^2 \times y^2 = |\,x\,|^2\,|\,y\,|^2 .$$

Assim, temos que $|\,x \times y\,| = \pm|\,x\,|\,|\,y\,|$. Entretanto, como $|\,x \times y\,|$ e $|\,x\,|\,|\,y\,|$ são ambos positivos, concluímos que $|\,x \times y\,| = |\,x\,|\,|\,y\,|$. ■

Teorema 1.5.3 *Sejam* $x, y \in \mathbb{F}$. *Então,* $|\,x + y\,| \leq |\,x\,| + |\,y\,|$.

Demonstração – Pela definição de valor absoluto em $\mathbb{F}$, temos que

$$-|\,x\,| \leq x \leq |\,x\,| \qquad \text{e} \qquad -|\,y\,| \leq y \leq |\,y\,| .$$

Somando as duas desigualdades acima, obtemos

$$-(\,|\,x\,| + |\,y\,|\,) \leq x + y \leq (\,|\,x\,| + |\,y\,|\,) .$$

Pelo Teorema 1.5.1, concluímos que $|\,x + y\,| \leq |\,x\,| + |\,y\,|$. ■

Teorema 1.5.4 *Sejam* $x, y \in \mathbb{F}$. *Então,* $|\,|\,x\,| - |\,y\,|\,| \leq |\,x - y\,|$.

Demonstração – Considerando $x = x - y + y$ e o Teorema 1.5.3, obtemos

$$|\,x\,| = |\,x - y + y\,| \leq |\,x - y\,| + |\,y\,|$$

Assim, temos que $|\,x\,| - |\,y\,| \leq |\,x - y\,|$.

De modo análogo, obtemos $|\,y\,| - |\,x\,| \leq |\,y - x\,|$. Podemos verificar facilmente que $|\,y - x\,| = |\,x - y\,|$. Portanto, mostramos que

$$|\,x\,| - |\,y\,| \leq |\,x - y\,| \qquad \text{e} \qquad -(|\,x\,| - |\,y\,|) \leq |\,x - y\,|$$

Pelo Teorema 1.5.1, obtemos

$$|\,|\,x\,| - |\,y\,|\,| \leq |\,x - y\,| .$$

1.6 Números Reais

O conjunto dos números reais, denotado por $I\!R$, tem uma estrutura de corpo com relação as operações usuais de adição e multiplicação. Assim, o conjunto dos números reais tem as propriedades apresentadas na Seção 1.2. Considerando que existe um subconjunto $I\!R^+ \subset I\!R$, denominado conjunto dos números positivos, que satisfaz os seguintes axiomas:

(O_1) Se $x, y \in I\!R^+$, então $x + y \in I\!R^+$ e $xy \in I\!R^+$.

(O_2) Para todo elemento $x \neq 0_{I\!R}$, temos que $x \in I\!R^+$ ou $-x \in I\!R^+$.

(O_3) O elemento neutro $0_{I\!R} \notin I\!R^+$.

temos que o conjunto dos números reais $I\!R$ tem uma estrutura de corpo ordenado. Desse modo, no corpo ordenado $I\!R$ valem as observações apresentadas na Seção 1.4. Com os axiomas de ordem, podemos obter importantes desigualdades que apresentamos no teorema a seguir.

Teorema 1.6.1 *Considere* $a, b, c \in I\!R$.

1. *Se* $ab = 0$, *então* $a = 0$ *ou* $b = 0$.
2. *Se* $a < b$ *e* $c > 0$, *então* $ac < bc$.
3. *Se* $a \neq 0$, *então* $a^2 > 0$.
4. *Se* $a < b$ *e* $c < 0$, *então* $ac > bc$.
5. *Se* $a < c$ *e* $b < d$, *então* $a + b < c + d$.
6. *Se* $ab > 0$, *então* a *e* b *são positivos ou ambos são negativos.*

Demonstração – A prova pode ficar a cargo do leitor. □

Definição 1.6.1 *Para* $x \in I\!R$, *definimos seu* ***valor absoluto***, *ou* ***módulo***, *que vamos denotar por* $|\,x\,|$, *como sendo o número real não negativo*

$$|\,x\,| = \begin{cases} x & se \quad x > 0 \\ 0 & se \quad x = 0 \\ -x & se \quad x < 0 \end{cases}$$

Podemos observar que $|\,x\,|$ é escolhido o maior número entre x e $-x$. Logo, temos que $|\,x\,| \geq x$ e $|\,x\,| \geq -x$. Portanto, podemos concluir que

$$-|\,x\,| \leq x \leq |\,x\,| \,.$$

Teorema 1.6.2 *Seja* $a \geq 0$. *Então,* $|\,x\,| \leq a$ *se, e somente se,* $-a \leq x \leq a$.

Demonstração

($\Longrightarrow$) Da definição de módulo de um número real, temos que

$$-|\,x\,| \leq x \leq |\,x\,| \,,$$

isto é, $|\,x\,| = x$ ou $|\,x\,| = -x$.

Tomando a hipótese que $|\,x\,| \leq a$, podemos escrever

$$-a \leq -|\,x\,| \leq x \leq |\,x\,| \leq a \,.$$

Assim, provamos que $-a \leq x \leq a$.

($\Longleftarrow$) Tomando por hipótese que $-a \leq x \leq a$. Desse modo, se $x \geq 0$, temos que $|\,x\,| = x \leq a$. Se $x < 0$, temos que $|\,x\,| = -x \leq a$. Portando, provamos que $|\,x\,| \leq a$, o que completa a demonstração. ■

Corolário 1.6.1 *Sejam* $a, b, x \in I\!R$. *Então,*

$$|\,x - a\,| \leq b \quad \textit{se, e somente se,} \quad a - b \leq x \leq a + b \,.$$

Teorema 1.6.3 *Sejam* $x, y \in I\!R$. *Então,* $|\,xy\,| = |\,x\,|\,|\,y\,|$.

Demonstração – Primeiramente vamos observar que $x^2 = |\,x\,|^2$, para todo $x \in I\!R$, pois $|\,x\,| = x$ ou $|\,x\,| = -x$ e vale $x^2 = (-x)^2$.

Desse modo, temos que

$$|\,xy\,|^2 = (xy)^2 = x^2y^2 = |\,x\,|^2\,|\,y\,|^2 \,.$$

Assim, concluímos que $|\,xy\,| = \pm|\,x\,|\,|\,y\,|$. Entretanto, como $|\,xy\,|$ e $|\,x\,|\,|\,y\,|$ são ambos positivos, obtemos que $|\,xy\,| = |\,x\,|\,|\,y\,|$. ■

Teorema 1.6.4 *Sejam* $x, y \in I\!R$. *Então,* $|x + y| \leq |x| + |y|$.

Demonstração – Pela definição de valor absoluto de um número real, temos que

$$-|x| \leq x \leq |x| \quad \text{e} \quad -|y| \leq y \leq |y| .$$

Somando as duas desigualdades acima, obtemos

$$-(|x| + |y|) \leq x + y \leq (|x| + |y|) .$$

Pelo Teorema 1.6.2, podemos concluir que $|x + y| \leq |x| + |y|$. ■

Essa propriedade é denominada **desigualdade triangular** para números reais.

A seguir, apresentamos a desigualdade triangular numa forma que é mais utilizada na prática.

Fazendo $x = a - c$ e $y = c - b$, temos que $x + y = a - b$. Agora, utilizando a desigualdade triangular $|x + y| \leq |x| + |y|$, obtemos

$$|a - b| \leq |a - c| + |b - c| .$$

Teorema 1.6.5 *Sejam* $a_1, a_2, \cdots, a_n$ *números reais quaisquer. Então,*

$$\left| \sum_{k=1}^{n} a_k \right| \leq \sum_{k=1}^{n} |a_k| .$$

Demonstração – A prova é feita por indução utilizando a desigualdade triangular. □

Teorema 1.6.6 *Sejam* $x, y \in I\!R$. *Então,* $|x| - |y| \leq |x - y|$.

Demonstração – Utilizando a desigualdade triangular para os números reais e o fato que $x = x - y + y$, obtemos

$$|x| = |x - y + y| \leq |x - y| + |y| .$$

Portanto, obtemos $|x| - |y| \leq |x - y|$, o que completa a demonstração. ■

Finalmente, podemos dizer que $I\!R$ é um corpo ordenado com um valor absoluto $|\cdot|$, que vamos fazer referência como sendo o valor absoluto usual.

1.7 Números Complexos

Definição 1.7.1 *Definimos o conjunto dos números complexos da seguinte forma:*

$$\mathbb{C} = \{ a + bi \;/\; a, b \in I\!R \}$$

onde $i = \sqrt{-1}$ *é a unidade imaginária.*

Considere os números complexos $z = a + bi$ *e* $w = c + di$*. A operação de* ***adição*** *de dois números complexos é definida por:*

$$z + w = (a + bi) + (c + di) = (a + c) + (b + d)i \,.$$

A operação de ***multiplicação*** *de dois números complexos é definida por:*

$$z \cdot w = (a + bi) \cdot (c + di) = (ac - bd) + (bc + ad)i \,.$$

Em particular, temos que $i^2 = i \cdot i = -1$*, que é compatível com a definição da unidade imaginária.*

Definição 1.7.2 *Considere o número complexo* $z = a + bi$*. Definimos a* ***parte real*** *como sendo o número real* $Re(z) = a$ *e a* ***parte imaginária*** *como sendo o número real* $Im(z) = b$*.*

Teorema 1.7.1 *O conjunto dos números complexos* $\mathbb{C}$ *com as operações de adição e multiplicação definidas acima tem uma estrutura de corpo.*

Demonstração – A prova pode ficar a cargo do leitor. □

Devemos observar que $\mathbb{C}$ não é um corpo ordenado. De fato, num corpo ordenado $I\!F$, para todo elemento $x \neq 0_{I\!F}$, tem–se que $x^2 > 0_{I\!F}$. Desse modo, considerando $z = i \in \mathbb{C}$, que é diferente do elemento neutro, temos que $z^2 = -1 < 0$. Logo, provamos que $\mathbb{C}$ não é um corpo ordenado.

Definição 1.7.3 *Definimos o* ***complexo conjugado*** *do número complexo* $z = a + bi$*, que denotamos por* $\overline{z}$*, como sendo o número complexo* $\overline{z} = a - bi$*.*

Definição 1.7.4 *O* ***valor absoluto****, ou* ***módulo****, do número complexo* $z = a + bi$*, que denotamos por* $|\,z\,|$*, é definido como sendo o número real não negativo*

$$|\,z\,| = \sqrt{a^2 + b^2} \,.$$

Definição 1.7.5 *Considere o número complexo* $z = a + bi$. *O* ***argumento principal*** *do número complexo* z *é definido da seguinte forma:*

$$\arg(z) = \begin{cases} \arctan\left(\dfrac{b}{a}\right) & \textit{para} \quad a \neq 0 \\[2ex] \dfrac{\pi}{2} & \textit{para} \quad a = 0, \ \ b > 0 \\[2ex] -\dfrac{\pi}{2} & \textit{para} \quad a = 0, \ \ b < 0 \end{cases}$$

onde $\theta = \arg(z)$ *indica o ângulo formado entre o eixo* $0x$ *e a reta que passa pela origem do plano* $\mathbb{R}^2$ *e pelo ponto* (a, b).

Definição 1.7.6 *Considere o número complexo* $z = a + bi$. *A* ***forma polar*** *do número complexo* z *é definida da seguinte forma:*

$$z = |\,z\,|\exp(i\theta) = |\,z\,|\cos(\theta) \ + \ i\,|\,z\,|\sin(\theta)$$

onde $\theta = \arg(z)$.

Teorema 1.7.2 *Considere os números complexos* $z = a + bi$ *e* $w = c + di$. *Então,*

(a) $\overline{\overline{z}} = z$.

(b) $z + \overline{z} = 2Re(z)$ *e* $z - \overline{z} = 2iIm(z)$.

(c) $z \cdot \overline{z} = |\,z\,|^2$ *e* $|\,\overline{z}\,| = |\,z\,|$.

(d) $\overline{z + w} = \overline{z} + \overline{w}$.

(e) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.

Demonstração – Vamos fazer a prova o item *(e)*. Temos que

$$\begin{aligned} \overline{z \cdot w} &= \overline{(a + bi) \cdot (c + di)} = (ac - bd) - (ad + bc)i \\ &= (a - bi) \cdot (c - di) = \overline{z} \cdot \overline{w}\,. \end{aligned}$$

Vamos fazer a prova o item *(d)*. Temos que

$$\begin{aligned} \overline{z + w} &= \overline{(a + bi) + (c + di)} = (a + c) - (b + d)i \\ &= (a - bi) + (c - di) = \overline{z} + \overline{w}\,. \end{aligned}$$

A prova dos outros ítens pode ficar a cargo do leitor. $\square$

Teorema 1.7.3 *Considere os números complexos* $z = a + bi$ *e* $w = c + di$. *Então,*

(a) $|\, z \cdot w \,| = |\, z \,|\, |\, w \,|$.

(b) $|\, Re(z) \,| \leq |\, z \,|$ *e* $|\, Im(z) \,| \leq |\, z \,|$.

(c) $z^{-1} = \dfrac{1}{z} = \dfrac{\overline{z}}{z \cdot \overline{z}} = \dfrac{\overline{z}}{|\, z \,|^2}$ *com* $z \neq 0$.

Demonstração – Vamos fazer a prova o item *(a)*. Pelo Teorema 1.7.2, temos que

$$|\, z \cdot w \,|^2 = (z \cdot w) \cdot \overline{(z \cdot w)} = (z \cdot \overline{z}) \cdot (w \cdot \overline{w}) = |\, z \,|^2 \,|\, w \,|^2 ,$$

o que completa a prova do item *(a)*.

A prova dos outros ítens pode ficar a cargo do leitor. □

Exemplo 1.7.1 *Considere o número complexo* $z = 3 + 4i$. *Como* $z \neq 0$, *temos que*

$$z^{-1} = \frac{\overline{z}}{|\, z \,|^2} = \frac{3 - 4i}{25} .$$

Exemplo 1.7.2 *Considere os números complexos* $z = 3 + 4i$ *e* $w = 1 + 2i$. *Como* $z \neq 0$, *temos que*

$$\frac{w}{z} = \frac{w \cdot \overline{z}}{z \cdot \overline{z}} = \frac{w \cdot \overline{z}}{|\, z \,|^2} = \frac{(1 + 2i) \cdot (3 - 4i)}{25} = \frac{11 + 2i}{25} .$$

Lema 1.7.1 *Se* $z, w \in \mathbb{C}$, *então*

$$|\, z + w \,|^2 = |\, z \,|^2 + |\, w \,|^2 + 2Re(z \cdot \overline{w}) .$$

Demonstração – Inicialmente, vamos escrever

$$\begin{aligned} |\, z + w \,|^2 &= (z + w) \cdot (\overline{z} + \overline{w}) = z \cdot \overline{z} + w \cdot \overline{w} + (z \cdot \overline{w} + \overline{z} \cdot w) \\ &= |\, z \,|^2 + |\, w \,|^2 + (z \cdot \overline{w} + \overline{z} \cdot w) \end{aligned}$$

Podemos observar que, $\overline{z} \cdot w$ é o conjugado de $z \cdot \overline{w}$. Portanto, temos que

$$z \cdot \overline{w} + \overline{z} \cdot w = 2Re(z \cdot \overline{w}) ,$$

o que completa a prova. ■

Teorema 1.7.4 *Se* $z, w \in \mathbb{C}$, *então* $|z + w| \leq |z| + |w|$.

Demonstração – Utilizando o resultado do Lema 1.7.1, temos que

$$|z + w|^2 - (|z| + |w|)^2 = -2|z||w| + 2Re(z \cdot \overline{w}) .$$

Sabemos que $Re(z \cdot \overline{w}) \leq |z \cdot \overline{w}| = |z||w|$. Logo,

$$|z + w|^2 - (|z| + |w|)^2 \leq 0 ,$$

o que completa a prova da **desigualdade triangular** para os números complexos. ■

Lema 1.7.2 *Sejam* z *e* w *dois números complexos. Então,* $|z| - |w| \leq |z + w|$.

Demonstração – Utilizando a desigualdade triangular para os números complexos e a propriedade $|z \cdot w| = |z||w|$, temos que

$$|z| = |(z + w) - w| \leq |z + w| + |-w| = |z + w| + |w| .$$

Portanto, temos que $|z| - |w| \leq |z + w|$. ■

Lema 1.7.3 *Se* $z, w \in \mathbb{C}$ *com* $w \neq 0$, *então*

$$\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}} \qquad e \qquad \left|\frac{z}{w}\right| = \frac{|z|}{|w|} .$$

Demonstração – A prova pode ficar a cargo do leitor. □

Assim, podemos dizer que $\mathbb{C}$ é um corpo com um valor absoluto $|\cdot|$, que vamos fazer referência como sendo o valor absoluto usual. Finalmente, podemos apresentar os seguintes exemplos de subcorpos de $\mathbb{C}$.

Exemplo 1.7.3 *O conjunto dos números reais* $\mathbb{R}$ *é um subcorpo do corpo dos números complexos* $\mathbb{C}$.

Exemplo 1.7.4 *O conjunto dos números racionais* $\mathbb{Q}$ *é um subcorpo do corpo dos números complexos* $\mathbb{C}$.

Exemplo 1.7.5 *O conjunto* $\mathbb{Q}(\sqrt{2})$ *é um subcorpo do corpo dos números complexos* $\mathbb{C}$.

Exercícios

Exercício 1.14 *Calcule o módulo e o argumento dos seguintes números complexos.*

(a) $z = (1 + 2i) \cdot (3 - i)$ *(b)* $z = \dfrac{2 + 3i}{2 + i}$ *(c)* $z = \dfrac{1}{(1 + i)^2}$

Exercício 1.15 *Determine a forma polar dos seguintes números complexos.*

(a) $z = (1 + i)^2$ *(b)* $z = 2 + 2i$ *(c)* $z = \dfrac{1}{(1 + i)}$

Exercício 1.16 *Verifique que os números complexos* $z = 1 + i$ *e* $w = 1 - i$ *satisfazem a equação* $z^2 - 2z + 2 = 0$.

Exercício 1.17 *Seja* z *um número complexo. Mostre que*

$$|\,z\,|\sqrt{2} \geq |\,Re(z)\,| + |\,Im(z)\,| \,.$$

Exercício 1.18 *Sejam* $z, w, u \in \mathbb{C}$ *com* $|\,z\,| \neq |\,w\,|$, *mostre que*

$$\left| \frac{u}{z + w} \right| \leq \frac{|\,u\,|}{|\,|\,z\,| - |\,w\,|\,|} \,.$$

Exercício 1.19 *Faça a representação gráfica no plano complexo dos subconjuntos.*

(a) $S = \{\, z \in \mathbb{C} \;/\; |\,z\,| = 1 \,\}$.

(b) $S = \{\, z \in \mathbb{C} \;/\; z + \overline{z} = 1 \,\}$.

(c) $S = \{\, z \in \mathbb{C} \;/\; z - \overline{z} = i \,\}$.

Exercício 1.20 *Determine os números reais* a *e* b *tais que*

$$\sum_{k=0}^{100} i^k = a + bi \,.$$

Exercício 1.21 *Expresse o número complexo*

$$z = \frac{1 + i}{2 - i} \,.$$

na forma $z = a + bi$.

1.8 Característica do Corpo

Definição 1.8.1 *Seja* $\mathbb{F}$ *um corpo. Definimos a* ***característica*** *do corpo* $\mathbb{F}$ *como sendo o menor inteiro positivo* p *tal que* $1_{\mathbb{F}} + 1_{\mathbb{F}} + \cdots + 1_{\mathbb{F}} = 0_{\mathbb{F}}$, *com* p *termos no somatório. O corpo* $\mathbb{F}$ *tem* ***característica zero*** *se* $1_{\mathbb{F}} + 1_{\mathbb{F}} + \cdots + 1_{\mathbb{F}} \neq 0_{\mathbb{F}}$, *para qualquer quantidade de termos no somatório.*

Podemos verificar que se o corpo $\mathbb{F}$ tem característica $p \neq 0$, então

$$x + x + \cdots + x = 0 ,$$

com p termos no somatório, para todo $x \in \mathbb{F}$.

Exemplo 1.8.1 *O corpo do números reais* $\mathbb{R}$ *tem característica zero.*

Exemplo 1.8.2 *O corpo do números racionais* $\mathbb{Q}$ *tem característica zero.*

Exemplo 1.8.3 *O corpo do números complexos* $\mathbb{C}$ *tem característica zero.*

Exemplo 1.8.4 *O corpo* $\mathbb{Z}_2 = \{ 0, 1 \}$, *definido no Exemplo 1.4.3, tem característica* $p = 2$. *De fato,* $1 + 1 = 0$ *no corpo* $\mathbb{Z}_2$.

Exemplo 1.8.5 *O corpo* $\mathbb{Z}_5 = \{ 0, 1, 2, 3, 4 \}$ *tem característica* $p = 5$. *De fato,* $1 + 1 + 1 + 1 + 1 = 0$ *no corpo* $\mathbb{Z}_5$.

Podemos verificar facilmente que se $\mathbb{F}$ é um corpo de característica zero, então $\mathbb{N} \subset \mathbb{F}$. De fato, fazendo a identificação $1_{\mathbb{F}} = 1$, temos que

$$1 + 1 = 2 \quad , \quad 1 + 1 + 1 = 3 \qquad \text{e} \qquad \sum_{i=1}^{n} 1 = 1 + 1 + 1 + \cdots + 1 = n .$$

Finalmente, os elementos simétricos $-n$ dos elementos $n \in \mathbb{N}$ também pertencem ao corpo $\mathbb{F}$. Desse modo, temos que $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{F}$. Em particular, todo corpo de característica zero é infinito.

Podemos verificar que todo corpo ordenado tem característica zero, tendo em vista que

$$1 + 1 + 1 + \cdots + 1 \neq 0$$

para qualquer quantidade de termos no somatório, pois num corpo ordenado a soma de elementos positivos é sempre um elemento positivo.

Definição 1.8.2 *Seja $\mathbb{F}$ um corpo de característica p. Definimos n em $\mathbb{F}$ da seguinte forma:*

$$n := 1_{\mathbb{F}} + 1_{\mathbb{F}} + \cdots + 1_{\mathbb{F}}$$

com n termos no somatório.

Podemos observar que vamos tomar n como sendo o resto da divisão do somatório

$$1_{\mathbb{F}} + 1_{\mathbb{F}} + \cdots + 1_{\mathbb{F}}$$

pelo inteiro positivo p, que é a característica do corpo $\mathbb{F}$.

1.9 Métricas

Definição 1.9.1 *Seja* $\mathbb{X}$ *um conjunto não vazio. Uma* ***métrica*** *ou uma* ***distância*** *em* $\mathbb{X}$ *é uma aplicação* $d(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \longrightarrow I\!R$ *satisfazendo as seguintes propriedades:*

1. ***Simetria:*** $d(x,y) \;=\; d(y,x)$ *para todos* $x,\, y \;\in\; \mathbb{X}$.

2. ***Positividade:*** $d(x,y) \;\geq\; 0$ *com* $d(x,y) \;=\; 0 \iff x \;=\; y$.

3. ***Desigualdade Triangular:*** $d(x,z) \;\leq\; d(x,y) + d(y,z)$ *para todos* $x,\, y,\, z \in \mathbb{X}$.

Utilizamos a notação $(\mathbb{X}, d)$ para denotar que o conjunto $\mathbb{X}$ está munido com a métrica $d(\cdot,\cdot)$ e dizemos que $(\mathbb{X}, d)$ é um **espaço métrico**.

Proposição 1.9.1 *Seja* $I\!F$ *um corpo com valor absoluto* $v(\cdot)$*. A aplicação*

$$d(x,y) \;=\; v(x-y) \qquad \textit{para todos} \qquad x,\, y \;\in\; I\!F\,,$$

define uma métrica no corpo $I\!F$.

Demonstração – Para todos $x,\, y \;\in\; I\!F$, temos que

$$d(x,y) \;=\; v(x-y) \;=\; v(-(y-x)) \;=\; v(y-x) \;=\; d(y,x)\,,$$

provando a propriedade de simetria.

Considerando a propriedade de positividade do valor absoluto $v(\cdot)$, temos que

$$d(x,y) \;=\; v(x-y) \;\geq\; 0$$

com

$$d(x,y) \;=\; v(x-y) \;=\; 0 \qquad \iff \qquad x \,-\, y \;=\; 0_{I\!F} \qquad \iff \qquad x \;=\; y\,,$$

provando a propriedade de positividade.

Para todos $x,\, y,\, z \;\in\; I\!F$, temos que

$$d(x,z) \;=\; v(x-z) \;=\; v(x-y \,+\, y-z) \;\leq\; v(x-y) \,+\, v(y-z) \;=\; d(x,y) \,+\, d(y,z)\,,$$

provando a desigualdade triangular, o que completa a demonstração. ∎

Definição 1.9.2 *O valor absoluto trivial* $v(\cdot)$ *em qualquer corpo* $\mathbb{F}$*, define uma* ***métrica discreta***

$$d(x,y) = v(x-y) = \begin{cases} 1 & se \quad x \neq y \\ 0 & se \quad x = y \end{cases}$$

para todos $x, y \in \mathbb{F}$.

Exemplo 1.9.1 *A aplicação* $d(x,y) = |x - y|$ *para todos* $x, y \in \mathbb{R}$*, define uma métrica no corpo dos números reais* $\mathbb{R}$*, onde* $|\cdot|$ *é o valor absoluto usual em* $\mathbb{R}$.

Exemplo 1.9.2 *A aplicação* $d(z,w) = |z - w|$ *para todos* $z, w \in \mathbb{C}$*, define uma métrica no corpo dos números complexos* $\mathbb{C}$*, onde* $|\cdot|$ *é o valor absoluto usual em* $\mathbb{C}$.

Exemplo 1.9.3 *Considere o seguinte conjunto*

$$\mathcal{C}([a,b]) = \{ f : [a,b] \longrightarrow \mathbb{R} \ / \ f \ \text{é uma função contínua} \}.$$

Podemos verificar facilmente que a aplicação

$$d_\infty(f,g) = \max\{ |f(x) - g(x)| \quad ; \quad x \in [a,b] \} \qquad ; \qquad \forall \ f, g \in \mathcal{C}([a,b])$$

define uma métrica no conjunto $\mathcal{C}([a,b])$*, onde* $|\cdot|$ *é o valor absoluto usual em* $\mathbb{R}$.

Exemplo 1.9.4 *Determine* $d_\infty(f,g)$*, com* $f(x) = x$ *e* $g(x) = x^2$ *para* $x \in [0,1]$.

Devemos calcular

$$d_\infty(f,g) = \max\{ \left| x - x^2 \right| \quad ; \quad x \in [0,1] \} = \frac{1}{4}.$$

Exemplo 1.9.5 *A aplicação*

$$d_1(f,g) = \int_a^b |f(x) - g(x)| \, dx$$

define uma métrica no conjunto $\mathcal{C}([a,b])$*, onde* $|\cdot|$ *é o valor absoluto usual em* $\mathbb{R}$.

Considerando as funções $f(x) = x$ *e* $g(x) = x^2$ *para* $x \in [0,1]$*, determine* $d_1(f,g)$.

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] LUENBERGER, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \quad = \quad Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q \quad = \quad \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações
Notas de Aula

Petronio Pulino
Departamento de Matemática Aplicada
Instituto de Matemática, Estatística e Computação Científica
Universidade Estadual de Campinas
E–mail: pulino@ime.unicamp.br
www.ime.unicamp.br/~pulino/ALESA/

Janeiro de 2012

Conteúdo

2

Matrizes e Sistemas Lineares

Conteúdo

2.1 Matrizes

Definição 2.1.1 *Denominamos* ***matriz*** *a um conjunto de números reais, ou a um conjunto de números complexos, dispostos em linhas e colunas, numa certa ordem, e colocados entre colchetes. Assim, uma matriz* ***real****, ou uma matriz* ***complexa****, que vamos denotar por A, com m linhas e n colunas é representada da forma:*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

com $a_{ij} \in I\!R$, ou $a_{ij} \in \mathbb{C}$. Os escalares a_{ij} são denominados ***elementos*** *da matriz, onde o primeiro índice indica a linha e o segundo índice indica a coluna às quais pertence o elemento. Neste caso, dizemos que a matriz A é de ordem $m \times n$. Por simplicidade, vamos utilizar a indicação $A = [a_{ij}]$ para denotar a matriz A e seus elementos.*

Definição 2.1.2 *Dizemos que uma matriz $A = [a_{ij}]$ de ordem $m \times n$ é* ***quadrada*** *se $m = n$, isto é, se possui o mesmo número de linhas e de colunas. Neste caso, dizemos simplesmente que A é uma matriz de ordem n.*

Definição 2.1.3 *Dizemos que uma matriz $A = [a_{ij}]$ de ordem $m \times n$ é a* ***matriz nula*** *se seus elementos a_{ij} são todos nulos. Neste caso, denotamos $A = 0$. Freqüentemente, indicamos $0_{m \times n}$ para denotar uma matriz nula de ordem $m \times n$, onde pode causar alguma dúvida sobre a ordem da matriz.*

Definição 2.1.4 *Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ duas matrizes de ordem $m \times n$. Dizemos que as matrizes A e B são* ***iguais*** *se, e somente se,*

$$a_{ij} = b_{ij} \quad ; \quad i = 1, \cdots, m \quad e \quad j = 1, \cdots, n\,.$$

Definição 2.1.5 *Dizemos que uma matriz $A = [a_{ij}]$ de ordem $m \times 1$ é uma* ***matriz coluna****, que representamos por:*

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

Definição 2.1.6 *Dizemos que uma matriz* $A = [a_{ij}]$ *de ordem* $1 \times n$ *é uma* ***matriz linha****, que representamos por:*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} .$$

Em geral, uma matriz coluna também é denominada **vetor coluna** e uma matriz linha também é denominada **vetor linha**. Em particular, podemos considerar um escalar $a \in \mathbb{R}$, ou $a \in \mathbb{C}$, como uma matriz de ordem 1×1.

Exemplo 2.1.1 *A seguir temos o exemplo de uma matriz real*

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 4 & 6 & 7 \end{bmatrix}$$

de ordem 2×3.

Exemplo 2.1.2 *Determine os valores de* a, b, c *e* d *de modo que* $A = B$*, onde*

$$A = \begin{bmatrix} 3 & -1 \\ c & 5 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 2a - b & a + 2b \\ 3c - d & c - 3d \end{bmatrix} .$$

Exemplo 2.1.3 *A seguir temos o exemplo de uma matriz complexa*

$$A = \begin{bmatrix} 1 + i & 2i \\ 2 & 6 - 3i \end{bmatrix}$$

de ordem 2×2.

Exemplo 2.1.4 *A seguir temos o exemplo de uma matriz coluna real* X*, de ordem* 3×1*, e de uma matriz linha* Y*, de ordem* 1×4*,*

$$X = \begin{bmatrix} 1 \\ 5 \\ 8 \end{bmatrix} \qquad e \qquad Y = \begin{bmatrix} 2 & -1 & 4 & 6 \end{bmatrix} .$$

De modo análogo, podemos considerar uma matriz coluna complexa e uma matriz linha complexa. Nos casos em que fica claro qual é a ordem da matriz podemos omitir essa especificação. Omitimos também se a matriz é real ou complexa nos casos que não causam dúvidas ou que o resultado é válido tanto para matriz real quanto para matriz complexa.

Definição 2.1.7 *Considere os seguintes subconjuntos de* $I\!N$

$$\mathcal{I}_m = \{1, 2, \cdots, m\} \qquad e \qquad \mathcal{I}_n = \{1, 2, \cdots, n\}.$$

Uma ***matriz sobre o corpo*** $\boldsymbol{I\!F}$ *de ordem* $m \times n$ *é uma* ***função***

$$A : \mathcal{I}_m \times \mathcal{I}_n \longrightarrow I\!F$$

que para cada para ordenado $(i, j) \in \mathcal{I}_m \times \mathcal{I}_n$ *está associado um único escalar*

$$a_{ij} = A(i, j) \in I\!F,$$

denominado ***elemento*** *da matriz* A.

Rigorosamente falando, a tabela retangular exibida na Definição 2.1.1, não é uma matriz, mas sim a representação de uma matriz.

Exemplo 2.1.5 *Considere o seguinte conjunto* $\mathcal{I}_3 = \{1, 2, 3\}$. *Vamos definir uma* ***matriz real*** $A : \mathcal{I}_3 \times \mathcal{I}_3 \longrightarrow I\!R$ *da seguinte forma:*

$$a_{ij} = A(i, j) = \frac{1}{i + j - 1},$$

que é denominada ***matriz de Hilbert*** *de ordem* 3×3.

De acordo com a Definição 2.1.1, representamos a matriz A *da seguinte forma:*

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}.$$

De modo análogo, definimos a matriz de Hilbert de ordem $n \times n$.

Exemplo 2.1.6 *Considere o seguinte conjunto* $\mathcal{I}_4 = \{1, 2, 3, 4\}$. *Vamos definir uma matriz real* $A : \mathcal{I}_4 \times \mathcal{I}_4 \longrightarrow I\!R$ *cuja regra funcional é dada por:*

$$a_{ij} = A(i, j) = |\, i - j \,|.$$

De acordo com a Definição 2.1.1, representamos a matriz A *da seguinte forma:*

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}.$$

Definição 2.1.8 *Sejam* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *duas matrizes de ordem* $m \times n$. *Definimos a* ***soma*** *das matrizes* A *e* B, *que denotamos por* $A + B$, *como sendo a matriz* $C = [c_{ij}]$, *de ordem* $m \times n$, *onde cada elemento é definido da seguinte forma:*

$$c_{ij} = a_{ij} + b_{ij} \quad ; \quad i = 1, \cdots, m \quad e \quad j = 1, \cdots, n.$$

Por simplicidade, indicamos $A + B = [a_{ij} + b_{ij}]$ para denotar a soma das matrizes A e B. De modo análogo, definimos a **diferença** das matrizes A e B, que denotamos por $A - B = [a_{ij} - b_{ij}]$.

Definição 2.1.9 *Sejam* $A = [a_{ij}]$ *uma matriz de ordem* $m \times n$ *e um escalar* λ. *Definimos a multiplicação da matriz* A *pelo escalar* λ, *e denotamos* λA, *como sendo a matriz* $C = [c_{ij}]$, *de ordem* $m \times n$, *onde cada elemento é definido da seguinte forma:*

$$c_{ij} = \lambda\, a_{ij} \quad ; \quad i = 1, \cdots, m \quad e \quad j = 1, \cdots, n.$$

Por simplicidade, indicamos $\lambda A = [\lambda a_{ij}]$ *para denotar a multiplicação da matriz* A *pelo escalar* λ.

Exemplo 2.1.7 *Considerando as matrizes* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *de ordem* 2×3,

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 5 & 1 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \end{bmatrix},$$

a matriz $C = A + 2B = [a_{ij} + 2b_{ij}]$ *é dada por:*

$$C = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 5 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 6 & 2 \\ 8 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 8 & 3 \\ 11 & 7 & 5 \end{bmatrix}.$$

Teorema 2.1.1 *Sejam* A, B *e* C *matrizes de mesma ordem. Então,*

(a) $A + B = B + A$.

(b) $A + (B + C) = (A + B) + C$.

(c) Existe uma matriz nula 0, *da mesma ordem da matriz* A, *tal que* $A + 0 = A$.

(d) Existe uma matriz D, *da mesma ordem da matriz* A, *tal que* $A + D = 0$.

Demonstração – A prova é feita utilizando as definições das operações de soma de matrizes e da multiplicação de uma matriz por escalar, juntamente com as propriedades das operações com números reais (complexos). □

Exemplo 2.1.8 *Dadas as matrizes* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *de ordem* 3×2*,*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 4 & 3 \end{bmatrix},$$

determine a matriz D *tal que* $A + B - D = 0$*.*

Definição 2.1.10 *Sejam* X *uma matriz linha de ordem* $1 \times m$ *e* Y *uma matriz coluna de ordem* $m \times 1$*,*

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \end{bmatrix} \qquad e \qquad Y = \begin{bmatrix} y_{11} \\ \vdots \\ y_{m1} \end{bmatrix},$$

o ***produto*** XY*, nesta ordem, é a matriz* Z *de ordem* 1×1 *dada por:*

$$Z = \begin{bmatrix} x_{11}y_{11} + x_{12}y_{21} + \cdots + x_{1j}y_{j1} + \cdots + x_{1m}y_{m1} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} x_{1j}y_{j1} \end{bmatrix}.$$

Exemplo 2.1.9 *Dada a matriz linha* X *de ordem* 1×3 *e a matriz coluna* Y *de ordem* 3×1*,*

$$X = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix} \qquad e \qquad Y = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix},$$

a matriz $Z = XY$ *de ordem* 1×1 *é dada por:*

$$Z = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 + 12 + 2 \end{bmatrix} = \begin{bmatrix} 16 \end{bmatrix}.$$

Definição 2.1.11 *Sejam* $A = [a_{ij}]$ *uma matriz de ordem* $m \times p$ *e* $B = [b_{ij}]$ *uma matriz de ordem* $p \times n$*. O* ***produto*** AB*, nesta ordem, é a matriz* $C = [c_{ij}]$ *de ordem* $m \times n$ *cujos elementos são definidos por:*

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} \qquad ; \qquad i = 1, \cdots, m \quad e \quad j = 1, \cdots, n,$$

isto é, o elemento c_{ij} *é o produto da* i*–ésima linha de* A *pela* j*–ésima coluna de* B*. Assim, podemos definir o produto* AB *somente quando o número de colunas de* A *é igual ao número de linhas de* B*.*

Exemplo 2.1.10 *Dada a matriz* A *de ordem* 3×2 *e a matriz* B *de ordem* 2×4,

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 4 & 2 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 2 & 1 & 1 & 3 \\ 0 & 1 & 2 & 1 \end{bmatrix},$$

a matriz $C = AB$ *de ordem* 3×4 *é dada por:*

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 3 \\ 0 & 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 5 & 5 \\ 6 & 4 & 5 & 10 \\ 8 & 6 & 8 & 14 \end{bmatrix}.$$

Exemplo 2.1.11 *Dada a matriz coluna* X, *de ordem* 3×1,

$$X = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix},$$

determine a matriz $Z = XX^t$ *de ordem* 3×3.

Exemplo 2.1.12 *Dada uma matriz coluna* X, *de ordem* $m \times 1$,

$$X = \begin{bmatrix} x_{11} \\ \vdots \\ x_{i1} \\ \vdots \\ x_{m1} \end{bmatrix},$$

deduza uma regra para a formação da matriz $Z = XX^t$ *de ordem* $m \times m$.

Exemplo 2.1.13 *Dada a matriz coluna* X, *de ordem* 3×1,

$$X = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix},$$

determine todas as matrizes Y, *de ordem* 3×1, *tais que* $Y^t X = 0$.

Exemplo 2.1.14 *Determine um escalar* λ *tal que* $AX = \lambda X$, *onde*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad e \qquad X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Teorema 2.1.2 *Sejam as matrizes* $A = [a_{ij}]$ *de ordem* $m \times n$, $B = [b_{ij}]$ *de ordem* $n \times p$ *e* $C = [c_{ij}]$ *de ordem* $n \times p$ *. Então,* $A(B + C) = AB + AC$.

Demonstração – Chamando $D = A(B + C) = [d_{ij}]$, sabemos que

$$d_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

para $i = 1, \cdots, m$ e $j = 1, \cdots, p$.

Logo, temos que a primeira parcela é o elemento da i–ésima linha e da j–ésima coluna do produto AB e a segunda parcela é o elemento da i–ésima linha e da j–ésima coluna do produto AC. Portanto, provamos que $A(B + C) = AB + AC$. ■

Teorema 2.1.3 *Sejam as matrizes* $A = [a_{ij}]$ *de ordem* $m \times n$, $B = [b_{ij}]$ *de ordem* $m \times n$ *e* $C = [c_{ij}]$ *de ordem* $n \times p$ *. Então,* $(A + B)C = AC + BC$.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 2.1.4 *Sejam as matrizes* $A = [a_{ij}]$ *de ordem* $m \times n$, $B = [b_{ij}]$ *de ordem* $n \times p$ *e* $C = [c_{ij}]$ *de ordem* $p \times q$ *. Então,* $A(BC) = (AB)C$.

Demonstração – A prova pode ficar a cargo do leitor. □

É importante observar que

(a) $AB \neq BA$, em geral.

(b) $AB = 0$ não implica necessariamente que $A = 0$ ou $B = 0$.

(c) $AB = AC$ não implica necessariamente que $B = C$.

onde a ordem das matrizes, A, B e C, são tais que as operações indicadas acima podem ser efetuadas.

Exemplo 2.1.15 *Dadas as matrizes*

$$A = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}, \ B = \begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}.$$

Mostre que $AB = BA = 0$, $AC = A$ *e* $CA = C$.

Exemplo 2.1.16 *Dadas as matrizes*

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & -3 \\ 4 & -3 & -1 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 4 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 2 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 2 & 1 & -1 & -2 \\ 3 & -2 & -1 & -1 \\ 2 & -5 & -1 & 0 \end{bmatrix}.$$

Verifique que $AB = AC$, *entretanto,* $B \neq C$.

Exemplo 2.1.17 *Dadas as matrizes*

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -3 & 2 & -1 \\ -2 & 1 & 0 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix}.$$

Verifique que $AB = 0$ *e que*

$$BA = \begin{bmatrix} 11 & 6 & -1 \\ -22 & 12 & -2 \\ -11 & 6 & -1 \end{bmatrix}.$$

Portanto, em geral, $AB \neq BA$.

Teorema 2.1.5 *Sejam* A *e* B *matrizes de mesma ordem e* α *e* β *escalares. Então,*

(a) $\alpha(\beta A) = (\alpha\beta)A$.

(b) $(\alpha + \beta)A = \alpha A + \beta A$.

(c) $\alpha(A + B) = \alpha A + \alpha B$.

Demonstração – A prova é feita utilizando as definições das operações de soma de matrizes e de multiplicação de uma matriz por escalar, juntamente com as propriedades das operações com números reais (complexos). $\square$

Teorema 2.1.6 *Sejam* A *uma matriz de ordem* $m \times n$, B *uma matriz de ordem* $n \times p$ *e* λ *um escalar. Então,* $A(\lambda B) = \lambda(AB) = (\lambda A)B$.

Demonstração – A prova é feita utilizando as definições de produto de matrizes e de multiplicação de uma matriz por escalar, juntamente com as propriedades das operações com números reais (complexos). $\square$

Teorema 2.1.7 *Considere* $A = [a_{ij}]$ *uma matriz de ordem* $m \times n$ *e* λ *um escalar. Se* $\lambda A = 0_{m\times n}$*, então* $\lambda = 0$ *ou* $A = 0_{m\times n}$*.*

Demonstração – Pela Definição 2.1.9, sabemos que a matriz λA é dada por:

$$\lambda A = [\lambda\, a_{ij}] \quad \text{para} \quad i = 1, \cdots m \quad \text{e} \quad j = 1, \cdots n\,.$$

Desse modo, pela hipótese, temos que

$$\lambda\, a_{ij} = 0 \quad \text{para} \quad i = 1, \cdots m \quad \text{e} \quad j = 1, \cdots n\,.$$

Sendo assim, pelo Teorema 1.2.5, temos que

$$\lambda = 0 \quad \text{ou} \quad a_{ij} = 0 \quad \text{para} \quad i = 1, \cdots m \quad \text{e} \quad j = 1, \cdots n\,,$$

o que completa a demonstração. ■

Teorema 2.1.8 *Seja* A *uma matriz de ordem* $m \times n$*. Então,* $AX = 0_{m\times 1}$ *para toda matriz coluna* X *de ordem* $n \times 1$ *se, e somente se,* $A = 0_{m\times n}$*.*

Demonstração – Considerando que $A = 0_{m\times n}$, o resultado segue trivialmente.

Considerando que $AX = 0_{m\times 1}$ para toda matriz coluna X de ordem $n \times 1$, e tomando a equação $A = AI_n$, obtemos

$$A = AI_n = [AE_{\cdot 1} \cdots AE_{\cdot j} \cdots AE_{\cdot n}] = [0_{m\times 1} \cdots 0_{m\times 1} \cdots 0_{m\times 1}]\,,$$

onde a matriz coluna $E_{\cdot j}$ de ordem $n \times 1$ é a j–ésima coluna da matriz identidade I_n, uma vez que $AE_{\cdot j} = 0_{m\times 1}$ para $j = 1, \cdots, n$.

Portanto, mostramos que $A = 0_{m\times n}$, o que completa a demonstração. ■

Teorema 2.1.9 *Sejam* A *e* B *matrizes de ordem* $m \times n$*. Então,* $A = B$ *se, e somente se,* $AX = BX$ *para toda matriz coluna* X *de ordem* $n \times 1$*.*

Demonstração – A prova segue imediata pelo resultado do Teorema 2.1.8. De fato,

$$AX = BX \iff (A - B)X = 0_{m\times 1} \iff A - B = 0_{m\times n}\,.$$

para toda matriz coluna X de ordem $n \times 1$.

Como $A - B = 0_{m\times n}$, tem–se $A = B$, o que completa a demonstração. ■

Exercícios

Exercício 2.1 *Considere o subconjunto* $\mathcal{I}_n = \{1, 2, \cdots, n\}$ *de* $\mathbb{N}$. *Determine a matriz* $A : \mathcal{I}_n \times \mathcal{I}_n \longrightarrow \mathbb{R}$ *definida pela seguinte regra funcional*

$$a_{ij} = A(i,j) = \begin{cases} \;\;1 & se \quad |\, i - j\,| > 1 \\ -1 & se \quad |\, i - j\,| \leq 1 \end{cases}$$

Exercício 2.2 *Considere o subconjunto* $\mathcal{I}_n = \{1, 2, \cdots, n\}$ *de* $\mathbb{N}$. *Determine a matriz* $A : \mathcal{I}_n \times \mathcal{I}_n \longrightarrow \mathbb{R}$ *definida pela seguinte regra funcional*

$$a_{ij} = A(i,j) = \begin{cases} 1 & se \quad |\, i - j\,| < 2 \\ 0 & se \quad |\, i - j\,| \geq 2 \end{cases}$$

Exercício 2.3 *Sejam* A *uma matriz de ordem* $m \times n$ *e* X *uma matriz coluna de ordem* $n \times 1$ *que são indicadas da seguinte forma:*

$$A = [Y_1 \cdots Y_j \cdots Y_n] \qquad e \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix},$$

onde a matriz coluna Y_j *de ordem* $m \times 1$ *é a* j*–ésima coluna da matriz* A. *Mostre que podemos escrever o produto* AX *da seguinte forma:*

$$AX = x_1 Y_1 + \cdots + x_j Y_j + \cdots + x_n Y_n .$$

Exercício 2.4 *Sejam* A *uma matriz de ordem* $m \times n$ *e* B *uma matriz de ordem* $n \times p$ *que vamos indicar da seguinte forma:*

$$B = [Y_1 \cdots Y_j \cdots Y_p] ,$$

onde a matriz coluna Y_j *de ordem* $n \times 1$ *é a* j*–ésima coluna da matriz* B. *Mostre que podemos escrever a matriz* $C = AB$ *da seguinte forma:*

$$C = AB = A[Y_1 \cdots Y_j \cdots Y_p] = [AY_1 \cdots AY_j \cdots AY_p] .$$

onde a matriz coluna $Z_j = AY_j$ *de ordem* $m \times 1$ *é a* j*–ésima coluna da matriz* C.

Exercício 2.5 *Dadas as matrizes*

$$A = \begin{bmatrix} a+2b & 2a-b \\ 2c+d & c-2d \end{bmatrix} \quad e \quad B = \begin{bmatrix} 9 & -2 \\ 4 & 7 \end{bmatrix}.$$

Determine os parâmetros a, b, c *e* d *de modo que* $A = B$.

Exercício 2.6 *Dadas as matrizes*

$$X = \begin{bmatrix} a \\ 2 \\ 1 \end{bmatrix} \quad , \quad Y = \begin{bmatrix} -1 & b & 2 \end{bmatrix} \quad e \quad Z = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

Determine os parâmetros a *e* b *tais que* $YX = 0$ *e* $YZ = 1$.

Exercício 2.7 *Determine todas as matrizes* X *tais que* $YX = 0$*, onde*

$$X = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad e \quad Y = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}.$$

Exercício 2.8 *Dadas as matrizes*

$$A = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \quad , \quad X = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad e \quad Y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Determine os valores do parâmetro $\theta \in \mathbb{R}$ *de modo que* $AX = Y$.

Exercício 2.9 *Dadas as matrizes*

$$A = \begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix} \quad ; \quad B = \begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix} \quad e \quad C = \begin{bmatrix} -4 & -3 \\ 0 & -4 \end{bmatrix}.$$

Verifique que $AB = AC$.

Exercício 2.10 *Dada a matriz*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Determine as matrizes B *de modo que* $AB - BA = 0_{2\times2}$*, se possível.*

2.2 Tipos Especiais de Matrizes

Definição 2.2.1 *Seja* $U = [u_{ij}]$ *uma matriz de ordem* $n \times n$. *Dizemos que* U *é uma matriz* ***triangular superior*** *se os elementos abaixo da diagonal principal são todos nulos, isto é,* $u_{ij} = 0$ *para* $j < i$.

Exemplo 2.2.1 *A matriz* U *dada por:*

$$U = \begin{bmatrix} 2 & 1 & 5 \\ 0 & 3 & 3 \\ 0 & 0 & 6 \end{bmatrix}$$

é uma matriz triangular superior.

Definição 2.2.2 *Seja* $L = [l_{ij}]$ *uma matriz de ordem* $n \times n$. *Dizemos que* L *é uma matriz* ***triangular inferior*** *se os elementos acima da diagonal principal são todos nulos, isto é,* $l_{ij} = 0$ *para* $j > i$.

Exemplo 2.2.2 *A matriz* L *dada por:*

$$L = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 2 & 7 & 4 \end{bmatrix}$$

é uma matriz triangular inferior.

Exemplo 2.2.3 *Mostre que o produto de duas matrizes triangulares superiores é uma matriz triangular superior.*

Exemplo 2.2.4 *Mostre que o produto de duas matrizes triangulares inferiores é uma matriz triangular inferior.*

Definição 2.2.3 *Seja* $D = [d_{ij}]$ *uma matriz de ordem* $n \times n$. *Dizemos que* D *é uma matriz* ***diagonal*** *se os elementos fora da diagonal principal são todos nulos, isto é,* $d_{ij} = 0$ *para* $j \neq i$. *Freqüentemente, indicamos*

$$D = diag(d_1, \cdots, d_n),$$

para dizer que D *é uma matriz diagonal de ordem* $n \times n$.

Exemplo 2.2.5 *A matriz* D *dada por:*

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

é uma matriz diagonal.

Definição 2.2.4 *O* ***traço*** *de uma matriz* $A = [a_{ij}]$, *de ordem* n*, que denotamos por* $tr(A)$*, é a soma dos elementos da diagonal principal, isto é,*

$$tr(A) = \sum_{i=1}^{n} a_{ii} \, .$$

Exemplo 2.2.6 *Dada a matriz real*

$$A = \begin{bmatrix} 1 & 2 & 7 \\ 3 & 4 & 8 \\ 0 & 1 & 3 \end{bmatrix} ,$$

temos que $tr(A) = 1 + 4 + 3 = 8$.

Exemplo 2.2.7 *Dada a matriz complexa*

$$A = \begin{bmatrix} 4i & 2-i & 7+i \\ 3+2i & 4+i & 8+2i \\ 0 & 1+3i & 3-i \end{bmatrix} ,$$

temos que $tr(A) = 4i + (4+i) + (3-i) = 7 + 4i$.

Teorema 2.2.1 *Sejam* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *matrizes de ordem* n*. Então,*

(a) $tr(A+B) = tr(A) + tr(B)$.

(b) $tr(\lambda A) = \lambda tr(A)$ *para qualquer escalar* λ.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 2.2.2 *Sejam* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *matrizes de ordem* n*. Então,* $tr(AB) = tr(BA)$.

Demonstração – A prova pode ficar a cargo do leitor. □

Definição 2.2.5 *Uma matriz diagonal* $D = diag(d_{11}, \cdots, d_{nn})$ *cujos elementos da diagonal principal são todos iguais, isto é,* $d_{ii} = \alpha$ *para* $i = 1, \cdots, n$*, é denominada* ***matriz escalar****.*

Exemplo 2.2.8 *A matriz* D *dada por:*

$$D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

é uma matriz escalar de ordem 3.

Definição 2.2.6 *Uma matriz diagonal cujos elementos da diagonal principal são todos iguais a* 1 *é denominada* ***identidade****. Freqüentemente, indicamos* I_n *para denotar uma matriz identidade de ordem* n.

Exemplo 2.2.9 *A matriz* I *dada por:*

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

é uma matriz identidade de ordem 3.

Exemplo 2.2.10 *Seja* A *uma matriz de ordem* $m \times n$*. Podemos verificar facilmente que* $I_m A = A$ *e* $A I_n = A$.

Definição 2.2.7 *Se* A *é uma matriz de ordem* $m \times n$*, denominamos* ***transposta*** *de* A *à matriz de ordem* $n \times m$ *obtida trocando–se as linhas pelas colunas. Denotamos a transposta da matriz* A *por* A^t.

Exemplo 2.2.11 *Temos o seguinte exemplo de uma matriz real* A *de ordem* 4×3 *e de sua respectiva transposta* A^t *de ordem* 3×4.

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 3 & 5 \\ 2 & 1 & 4 \\ 1 & 2 & 7 \end{bmatrix} \qquad A^t = \begin{bmatrix} 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 3 & 5 & 4 & 7 \end{bmatrix}$$

Exemplo 2.2.12 *Seja* A *uma matriz real de ordem* n*. Podemos verificar facilmente que* $tr(A^t) = tr(A)$.

Exemplo 2.2.13 *Temos o seguinte exemplo de uma matriz complexa* A *de ordem* 2×3 *e de sua respectiva transposta* A^t *de ordem* 3×2.

$$A = \begin{bmatrix} 2 & 1+i & i \\ 3+i & 2i & 1 \end{bmatrix} \qquad A^t = \begin{bmatrix} 2 & 3+i \\ 1+i & 2i \\ i & 1 \end{bmatrix}$$

Definição 2.2.8 *Seja* $A = [a_{ij}]$ *uma matriz quadrada. Dizemos que* A *é* ***simétrica*** *se* $A^t = A$, *isto é,* $a_{ij} = a_{ji}$ *para todos* i, j.

Exemplo 2.2.14 *As matrizes* A *e* B *dadas por:*

$$A = \begin{bmatrix} 5 & 1 & 2 \\ 1 & 6 & 3 \\ 2 & 3 & 8 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 1+2i & 2+i \\ 2+i & 3 \end{bmatrix}$$

são matrizes simétricas, isto é, $A^t = A$ *e* $B^t = B$.

Definição 2.2.9 *Seja* A *uma matriz quadrada. Dizemos que* A *é* ***anti–simétrica*** *se* $A^t = -A$, *isto é,* $a_{ij} = -a_{ji}$ *para todos* i, j.

Exemplo 2.2.15 *As matrizes* A *e* B *dadas por:*

$$A = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 0 & 2-i & -3 \\ -2+i & 0 & i \\ 3 & -i & 0 \end{bmatrix}$$

são matrizes anti–simétricas, isto é, $A^t = -A$ *e* $B^t = -B$.

Definição 2.2.10 *Considere* $A = [a_{ij}]$ *uma matriz complexa de ordem* $m \times n$. *A matriz obtida de* A *substituindo cada elemento por seu conjugado é denominada* ***matriz conjugada*** *da matriz* A, *que denotamos por* $\overline{A}$. *Assim,* $\overline{A} = [\overline{a}_{ij}]$.

Exemplo 2.2.16 *Dada a matriz complexa*

$$A = \begin{bmatrix} 1+2i & i \\ 3 & 2-3i \end{bmatrix}.$$

A ***matriz conjugada*** *de* A, *que denotamos por* $\overline{A}$, *é obtida da seguinte forma:*

$$\overline{A} = \begin{bmatrix} 1-2i & -i \\ 3 & 2+3i \end{bmatrix}.$$

Definição 2.2.11 *Seja* $A = [a_{ij}]$ *uma matriz complexa de ordem* $m \times n$. *Definimos a matriz* ***transposta Hermitiana*** *da matriz* A, *que indicamos por* A^*, *como sendo a matriz* $A^* = [\overline{a}_{ji}]$ *de ordem* $n \times m$, *isto é,* $A^* = (\overline{A})^t$.

Exemplo 2.2.17 *Dada a matriz complexa*

$$A = \begin{bmatrix} 1+2i & i \\ 3 & 2-3i \end{bmatrix}.$$

A ***transposta Hermitiana*** *de* A *é dada por:*

$$A^* = \begin{bmatrix} 1-2i & 3 \\ -i & 2+3i \end{bmatrix}.$$

Teorema 2.2.3 *Sejam* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *matrizes complexas, com ordens compatíveis com as operações. Então,*

(a) $\overline{(A + B)} = \overline{A} + \overline{B}$.

(b) $\overline{(AB)} = \overline{A}\,\overline{B}$.

(c) $\overline{(\lambda A)} = \overline{\lambda}\,\overline{A}$ *para qualquer escalar* $\lambda \in \mathbb{C}$.

(d) $(\overline{A})^t = \overline{(A^t)}$.

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 2.2.18 *Seja* A *uma matriz complexa de ordem* n. *Observamos facilmente que* $tr(A^*) = \overline{tr(A)}$.

Definição 2.2.12 *Dizemos que uma matriz* $A = [a_{ij}]$ *complexa de ordem* n *é uma matriz* ***Hermitiana*** *se* $(\overline{A})^t = A$, *isto é,* $a_{ij} = \overline{a}_{ji}$ *para todos* i, j. *Geralmente indicamos* $A^* = A$ *para denotar uma matriz Hermitiana.*

Exemplo 2.2.19 *A matriz complexa*

$$A = \begin{bmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{bmatrix}$$

é uma matriz Hermitiana, isto é, $(\overline{A})^t = A$.

Definição 2.2.13 *Dizemos que uma matriz* $A = [a_{ij}]$ *complexa de ordem* n *é uma matriz* ***anti–Hermitiana*** *se* $(\overline{A})^t = -A$, *isto é,* $a_{ij} = -\overline{a}_{ji}$ *para todos* i, j. *Geralmente indicamos* $A^* = -A$ *para denotar uma matriz anti–Hermitiana.*

Exemplo 2.2.20 *A matriz complexa*

$$A = \begin{bmatrix} i & 1-i & 2 \\ -1-i & 3i & i \\ -2 & i & 0 \end{bmatrix}$$

é uma matriz anti–Hermitiana, isto é, $(\overline{A})^t = -A$.

Teorema 2.2.4 *Sejam* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *matrizes de mesma ordem e* α *um escalar. Então,*

(a) $(A^t)^t = A$.

(b) $(A + B)^t = A^t + B^t$.

(c) $(\alpha A)^t = \alpha A^t$.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 2.2.5 *Sejam as matrizes* $A = [a_{ij}]$ *de ordem* $m \times n$ *e* $B = [b_{ij}]$ *de ordem* $n \times p$. *Então,* $(AB)^t = B^t A^t$.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 2.2.6 *Sejam* $A = [a_{ij}]$ *e* $B = [b_{ij}]$ *matrizes complexas de mesma ordem e* α *um escalar. Então,*

(a) $(A^*)^* = A$.

(b) $(A + B)^* = A^* + B^*$.

(c) $(\alpha A)^* = \overline{\alpha} A^*$.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 2.2.7 *Sejam as matrizes complexas* $A = [a_{ij}]$ *de ordem* $m \times n$ *e* $B = [b_{ij}]$ *de ordem* $n \times p$. *Então,* $(AB)^* = B^* A^*$.

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 2.2.21 *Seja* A *uma matriz real de ordem* $m \times n$*. Podemos verificar facilmente que as matrizes* AA^t *e* A^tA *são simétricas.*

Exemplo 2.2.22 *Seja* A *uma matriz complexa de ordem* $m \times n$*. Podemos verificar facilmente que as matrizes* AA^* *e* A^*A *são Hermitianas.*

Definição 2.2.14 *Seja* A *uma matriz quadrada. Define–se* ***potenciação*** *para expoentes naturais da seguinte forma:*

$$A^0 = I \,, \quad A^1 = A \,, \quad A^2 = AA \qquad e \qquad A^{k+1} = AA^k \,.$$

Exemplo 2.2.23 *O calculo da expressão* $A^2 - 2A + 3I_2$*, onde*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix},$$

é obtido da seguinte forma:

$$A^2 - 2A + 3I_2 = \begin{bmatrix} 7 & 4 \\ 6 & 7 \end{bmatrix} - \begin{bmatrix} 2 & 4 \\ 6 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix}.$$

Podemos definir a matriz $p(A) = A^2 - 2A + 3I_2$*, de mesma ordem da matriz* A*, que é o* ***polinômio matricial*** *em* A *associado ao polinômio* $p(x) = 3 - 2x + x^2$*.*

Definição 2.2.15 *Dizemos que a matriz quadrada* A *é* ***idempotente*** *se* $A^2 = A$*.*

Exemplo 2.2.24 *A matriz* A*, dada abaixo, é idempotente, isto é,* $A^2 = A$*.*

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Exemplo 2.2.25 *A matriz* A*, dada abaixo, é idempotente, isto é,* $A^2 = A$*.*

$$A = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Definição 2.2.16 *Seja* A *uma matriz quadrada. Dizemos que* A *é* ***periódica****, com período* k*, se* $A^{k+1} = A$*, onde* k *é o menor inteiro positivo com tal propriedade.*

Definição 2.2.17 *Seja A uma matriz quadrada de ordem $n \times n$. Dizemos que A é* ***nilpotente*** *se existe um $k \in I\!N^*$ tal que $A^k = 0_n$. Se k é o menor inteiro positivo tal que $A^k = 0_n$, dizemos que A é nilpotente de* ***índice*** *k.*

Exemplo 2.2.26 *A matriz dada por:*

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

é uma matriz nilpotente de índice $k = 3$, isto é, $A^3 = 0_3$.

Definição 2.2.18 *Dizemos que a matriz quadrada A é* ***auto–reflexiva*** *se $A^2 = I$.*

Exemplo 2.2.27 *A matriz A dada por:*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

é uma matriz auto–reflexiva, isto é, $A^2 = I$.

Definição 2.2.19 *Se A e B são matrizes quadradas tais que $AB = BA$, dizemos que as matrizes A e B são* ***comutativas***.

Exemplo 2.2.28 *Podemos verificar facilmente que as matrizes*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 5 & 4 \\ 6 & 11 \end{bmatrix}$$

são comutativas, isto é, $AB = BA$.

Definição 2.2.20 *Se A e B são matrizes quadradas tais que $AB = -BA$, dizemos que as matrizes A e B são* ***anti–comutativas***.

Exemplo 2.2.29 *Podemos verificar facilmente que as matrizes*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad e \qquad C = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

são anti–comutativas duas a duas.

Teorema 2.2.8 *Sejam A uma matriz de ordem n e $D = diag(d, \cdots, d)$ uma matriz escalar de mesma ordem da matriz A. Então, $DA = AD$.*

Demonstração – Podemos verificar facilmente que uma matriz escalar pode ser escrita como $D = dI$, Exercício 2.11. Assim, utilizando o Teorema 2.1.6, temos que

$$DA = (dI)A = d(IA) = dA \qquad \text{e} \qquad AD = A(dI) = d(AI) = dA\,,$$

o que completa a demonstração. ■

Definição 2.2.21 *Seja A uma matriz real de ordem n. Dizemos que A é uma matriz* ***normal*** *se $A^tA = AA^t$, isto é, as matrizes A e A^t são comutativas.*

Exemplo 2.2.30 *As matrizes reais*

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 3 & 0 \\ 4 & 0 & 1 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 3 & -1 & 5 \\ 1 & 3 & 2 \\ -5 & -2 & 3 \end{bmatrix}$$

são matrizes normais, isto é, $A^tA = AA^t$ e $B^tB = BB^t$.

Exemplo 2.2.31 *Podemos verificar facilmente que se A é uma matriz simétrica real, então A é uma matriz normal real.*

Exemplo 2.2.32 *Podemos verificar facilmente que se A é uma matriz anti–simétrica real, então A é uma matriz normal real.*

Exemplo 2.2.33 *Podemos verificar facilmente que se A é a soma de uma matriz escalar real e uma matriz anti–simétrica real, então A é uma matriz normal real.*

De fato, vamos escrever $A = D + B$, onde D é uma matriz escalar e B é uma matriz anti–simétrica, isto é, $B^t = -B$. Assim, pelo Teorema 2.2.8, temos que

$$\begin{aligned} (D + B)^t(D + B) &= (D - B)(D + B) = D^2 + DB - BD - B^2 = D^2 - B^2 \\ (D + B)(D + B)^t &= (D + B)(D - B) = D^2 - DB + BD - B^2 = D^2 - B^2 \end{aligned}$$

Portanto, mostramos que $A^tA = AA^t$, isto é, A é uma matriz normal real.

Exemplo 2.2.34 *A matriz real*

$$A = \begin{bmatrix} 2 & 1 & -3 & 2 \\ -1 & 2 & 4 & -1 \\ 3 & -4 & 2 & 0 \\ -2 & 1 & 0 & 2 \end{bmatrix}$$

é uma matriz normal, isto é, $A^t A = A A^t$*. De fato, podemos observar facilmente que a matriz* A *é a soma de uma matriz escalar e uma matriz anti–simétrica.*

Definição 2.2.22 *Seja* A *uma matriz complexa de ordem* n*. Dizemos que* A *é uma matriz* ***normal*** *se* $A^* A = A A^*$*, isto é, as matrizes* A *e* A^* *são comutativas..*

Exemplo 2.2.35 *A matriz complexa*

$$A = \begin{bmatrix} 2+3i & 1 \\ i & 1+2i \end{bmatrix}$$

é uma matriz normal, isto é, $A^* A = A A^*$*.*

Exemplo 2.2.36 *Podemos verificar facilmente que se* A *é uma matriz Hermitiana, então* A *é uma matriz normal.*

Exemplo 2.2.37 *Podemos verificar facilmente que se* A *é uma matriz anti–Hermitiana, então* A *é uma matriz normal.*

Exemplo 2.2.38 *Podemos verificar facilmente que se* A *é a soma de uma matriz escalar complexa e uma matriz anti–Hermitiana, então* A *é uma matriz normal.*

De fato, vamos escrever $A = D + B$, onde D é uma matriz escalar e B é uma matriz anti–Hermitiana, isto é, $B^* = -B$. Assim, pelo Teorema 2.2.8, temos que

$$\begin{aligned} (D+B)^*(D+B) &= (D^* - B)(D+B) = D^*D + D^*B - BD - B^2 \\ &= D^*D + D^*B - DB - B^2 \end{aligned}$$

$$\begin{aligned} (D+B)(D+B)^* &= (D+B)(D^* - B) = DD^* - DB + BD^* - B^2 \\ &= D^*D + D^*B - DB - B^2 \end{aligned}$$

Portanto, mostramos que $A^* A = A A^*$, isto é, A é uma matriz normal complexa.

Exemplo 2.2.39 *A matriz complexa* $C = A + D$, *onde*

$$A = \begin{bmatrix} i & 1-i & 2 & 3+i \\ -1-i & 3i & i & 2i \\ -2 & i & 0 & -3 \\ -3+i & 2i & 3 & 2i \end{bmatrix} \qquad e \qquad D = \begin{bmatrix} 1+i & & & \\ & 1+i & & \\ & & 1+i & \\ & & & 1+i \end{bmatrix},$$

é uma matriz normal, isto é, $C^*C = CC^*$. *De fato, podemos observar facilmente que* A *é uma matriz anti–Hermitiana e* D *é uma matriz escalar complexa.*

Exemplo 2.2.40 *Podemos observar facilmente que uma matriz simétrica complexa não necessariamente é uma matriz normal. Tome como exemplo as seguintes matrizes simétricas*

$$A = \begin{bmatrix} 1 & i \\ i & i \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} i & i \\ i & 1 \end{bmatrix}.$$

De fato, temos que

$$A^*A = AA^* = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Logo, A *é uma matriz normal. Entretanto,*

$$BB^* = \begin{bmatrix} 2 & 1+i \\ 1-i & 2 \end{bmatrix} \qquad e \qquad B^*B = \begin{bmatrix} 2 & 1-i \\ 1+i & 2 \end{bmatrix}.$$

Logo, B *não é uma matriz normal.*

Exemplo 2.2.41 *Podemos verificar facilmente que a matriz complexa*

$$A = \begin{bmatrix} 1 & 1-2i \\ 1+2i & 1 \end{bmatrix}$$

é uma matriz normal, pois A *é Hermitiana, isto é,* $A^* = A$. *Assim, temos que*

$$A^*A = AA^* = \begin{bmatrix} 6 & 2-4i \\ 2+4i & 6 \end{bmatrix}.$$

Exemplo 2.2.42 *Seja* A *uma matriz real de ordem* $m \times n$. *Podemos verificar facilmente que a matriz* $C = A^tA$, *de ordem* n, *é uma matriz normal.*

Exemplo 2.2.43 *Seja* A *uma matriz real de ordem* $m \times n$. *Podemos verificar facilmente que a matriz* $C = AA^t$, *de ordem* m, *é uma matriz normal.*

Teorema 2.2.9 *Seja A uma matriz normal real de ordem 2×2. Então, A ou é uma matriz simétrica ou é a soma de uma matriz escalar e uma matriz anti–simétrica.*

Demonstração – Vamos escrever a matriz A da seguinte forma:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Assim, temos que

$$AA^t = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix}$$

Como, por hipótese, temos que $AA^t = A^tA$, obtemos três equações

(1) $a^2 + b^2 = a^2 + c^2$.

(2) $c^2 + d^2 = b^2 + d^2$.

(3) $ac + bd = ab + cd$.

Desse modo, da primeira equação, ou da segunda equação, obtemos $b^2 = c^2$. Logo, temos duas possibilidades $b = c$ ou $b = -c$.

Primeiramente, considerando o caso $b = c$, o que inclui o caso $b = -c = 0$, obtemos que a matriz A é simétrica, isto é,

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

Finalmente, considerando a situação $b = -c \neq 0$, da terceira equação obtemos

$$c(a - d) = ac + bd = ab + cd = c(d - a).$$

Assim, temos que

$$c(a - d) = c(d - a) \qquad \Longleftrightarrow \qquad 2c(a - d) = 0$$

como $c \neq 0$, obtemos $a = d$. Portanto, a matriz A tem a seguinte forma:

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$$

que é a soma de uma matriz escalar e uma matriz anti–simétrica, o que completa a demonstração. ■

Exercícios

Exercício 2.11 *Mostre que se* $A = [a_{ij}]$ *é uma matriz escalar de ordem* n, *então* $A = cI_n$ *para qualquer escalar* c.

Exercício 2.12 *Sejam* $A\,,B$ *e* C *matrizes quadradas de mesma ordem. Mostre que*

$$(ABC)^t = C^t\,B^t\,A^t\,.$$

Exercício 2.13 *Seja* $A = [a_{ij}]$ *uma matriz anti–simétrica. Mostre que os elementos da diagonal principal são todos nulos, isto é,* $a_{ii} = 0$ *para* $i = 1, \cdots, n$.

Exercício 2.14 *Seja* $A = [a_{ij}]$ *uma matriz Hermitiana. Mostre que os elementos da diagonal principal são números reais, isto é,* $a_{ii} \in I\!R$ *para* $i = 1, \cdots, n$..

Exercício 2.15 *Seja* $A = [a_{ij}]$ *uma matriz anti–Hermitiana. Mostre que os elementos da diagonal principal são ou nulo ou imaginário puro.*

Exercício 2.16 *Seja* A *uma matriz de ordem* n. *Então, a matriz* $B = A + A^t$ *é simétrica e a matriz* $C = A - A^t$ *é anti–simétrica.*

Exercício 2.17 *Seja* A *uma matriz complexa de ordem* n. *Então,* $B = A + A^*$ *é uma matriz Hermitiana e* $C = A - A^*$ *é uma matriz anti–Hermitiana.*

Exercício 2.18 *Mostre que as matrizes*

$$A = \begin{bmatrix} a & b \\ b & a \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} c & d \\ d & c \end{bmatrix}$$

comutam para quaisquer valores de a, b, c *e* d.

Exercício 2.19 *Sejam* A *e* B *matrizes simétricas de mesma ordem. Então,* AB *é uma matriz simétrica se, e somente se,* A *e* B *comutam, isto é,* $AB = BA$.

Exercício 2.20 *Seja* A *uma matriz idempotente, de ordem* $n \times n$. *Então,*

$$B = I - A$$

é uma matriz idempotente. Além disso, temos que $AB = BA = 0_n$.

Exercício 2.21 *Sejam* A *e* B *matrizes quadradas de mesma ordem tais que*

$$AB = A \qquad e \qquad BA = B .$$

Então, A *e* B *são matrizes idempotentes.*

Exercício 2.22 *Seja* A *uma matriz nilpotente com* $k = 2$*. Então,* $A(I + A)^3 = A$*.*

Exercício 2.23 *Qual a relação entre uma matriz* A *ser periódica e* A *ser nilpotente?*

Exercício 2.24 *Seja* A *uma matriz de ordem* n*. Mostre que* A *pode ser decomposta, de maneira única, como* $A = B + C$*, onde* B *é uma matriz simétrica e* C *é uma matriz anti–simétrica.*

Exercício 2.25 *Seja* A *uma matriz complexa de ordem* n*. Mostre que* A *pode ser decomposta, de maneira única, como* $A = B + C$*, onde* B *é uma matriz Hermitiana e* C *é uma matriz anti–Hermitiana.*

Exercício 2.26 *Considere* A *e* B *matrizes quadradas de mesma ordem. Seja* A *uma matriz simétrica. Então,* B^tAB *é uma matriz simétrica.*

Exercício 2.27 *Considere* A *e* B *matrizes quadradas de mesma ordem. Seja* A *uma matriz Hermitiana. Então,* B^*AB *é uma matriz Hermitiana.*

Exercício 2.28 *Seja* A *uma matriz Hermitiana de ordem* n*. Mostre que* A *pode ser escrita como* $A = B + iC$*, onde* B *é uma matriz simétrica real e* C *é uma matriz anti–simétrica real.*

Exercício 2.29 *Seja* A *uma matriz anti–Hermitiana de ordem* n*. Mostre que* A *pode ser escrita como* $A = B + iC$*, onde* B *é uma matriz anti–simétrica real e* C *é uma matriz simétrica real.*

Exercício 2.30 *Considere* A *e* B *matrizes quadradas de mesma ordem. Seja* A *uma matriz anti–simétrica. Então,* B^tAB *é uma matriz anti–simétrica.*

Exercício 2.31 *Considere* A *e* B *matrizes quadradas de mesma ordem. Sejam* A *e* B *matrizes anti–simétricas. Então,* AB *é simétrica se, e somente se, as matrizes* A *e* B *comutam, isto é,* $AB = BA$*.*

Exercício 2.32 *Seja* A *uma matriz real de ordem* $m \times n$. *Mostre que* $C = A^t A$ *é uma matriz simétrica.*

Exercício 2.33 *Sejam* A *uma matriz quadrada e* $B = \lambda A + \alpha I$, *onde* $\lambda, \alpha \in I\!R$. *Então, as matrizes* A *e* B *comutam.*

Exercício 2.34 *Mostre que não existem matrizes* A *e* B, *de ordem* n, *tais que*

$$AB - BA = I \,,$$

utilizando as propriedades de traço.

Exercício 2.35 *Se* A *é uma matriz simétrica (anti–simétrica) de ordem* m *e* P *é uma matriz de ordem* $m \times n$, *então* $B = P^t A P$ *é uma matriz simétrica (anti–simétrica).*

Exercício 2.36 *Seja* A *uma matriz de ordem* n *tal que* $AB = BA$ *para toda matriz* B *de ordem* n. *Mostre que* $A = cI_n$, *onde* c *é um escalar qualquer.*

Exercício 2.37 *Seja* A *uma matriz de ordem* n. *Mostre que*

$$I - A^{k+1} = (I - A)(I + A + \ldots + A^k) = (I + A + \ldots + A^k)(I - A) \,.$$

Exercício 2.38 *Mostre que a matriz*

$$A = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

é idempotente, isto é, $A^2 = A$.

Exercício 2.39 *Mostre que a matriz*

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$$

é nilpotente de ordem 3, *isto é,* $A^3 = 0$.

Exercício 2.40 *Mostre que se* A *é nilpotente de ordem* 2, *isto é,* $A^2 = 0$, *então*

$$A(I + A)^n = A \,,$$

para qualquer inteiro positivo n.

Exercício 2.41 *Mostre que uma matriz A é auto–reflexiva se, e somente se,*

$$(I - A)(I + A) = 0 .$$

Exercício 2.42 *Mostre que se A e B são matrizes quadradas, então A e B comutam se, e somente se, $A - \lambda I$ e $B - \lambda I$ comutam para qualquer escalar λ.*

Exercício 2.43 *Mostre que se A é uma matriz idempotente, de ordem $n \times n$, então $B = I - A$ é uma matriz idempotente e $AB = BA = 0_n$.*

Exercício 2.44 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} .$$

Mostre que $A^2 - 4A - 5I = 0_3$, onde $0_3 \in \mathbb{M}_3(\mathbb{R})$ é a matriz nula..

Exercício 2.45 *Dada a matriz complexa*

$$A = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} .$$

Mostre que uma fórmula para as potências inteiras positivas da matriz A é dada por:

$$A^n = I, A, -I, -A$$

para $n = 4m, 4m+1, 4m+2, 4m+3 \quad ; \quad m \in \mathbb{N}$, *respectivamente.*

Exercício 2.46 *Mostre que a matriz*

$$A = \begin{bmatrix} 1 & -2 & -6 \\ -3 & 2 & 9 \\ 2 & 0 & -3 \end{bmatrix}$$

é periódica com período 2, isto é, $A^3 = A$.

Exercício 2.47 *Mostre que a matriz*

$$A = \begin{bmatrix} 1 & -3 & -4 \\ -1 & 3 & 4 \\ 1 & -3 & -4 \end{bmatrix}$$

é nilpotente, isto é, existe um $k \in \mathbb{N}^*$ tal que $A^k = 0_3$.

Exercício 2.48 *Mostre que as matrizes*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 0 \\ -1 & -1 & -1 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} -2 & -1 & -6 \\ 3 & 2 & 9 \\ -1 & -1 & -4 \end{bmatrix}$$

comutam, isto é, $AB = BA$.

Exercício 2.49 *Mostre que a matriz*

$$A = \begin{bmatrix} 4 & 3 & 3 \\ -1 & 0 & -1 \\ -4 & -4 & -3 \end{bmatrix}$$

é auto–reflexiva, isto é, $A^2 = I$.

Exercício 2.50 *Determine todas as matrizes reais de ordem* 2 *da forma*

$$A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

tal que $A^2 = I_2$, *isto é,* A *é auto–reflexiva.*

Exercício 2.51 *Sejam* A *uma matriz de ordem* $m \times n$ *e* $D = diag(d_1, \cdots, d_m)$ *uma matriz diagonal. Deduza uma regra para o produto* DA.

Exercício 2.52 *Sejam* A *uma matriz de ordem* $m \times n$ *e* $D = diag(d_1, \cdots, d_n)$ *uma matriz diagonal. Deduza uma regra para o produto* AD.

Exercício 2.53 *Mostre que se* A *é auto–reflexiva, então as matrizes*

$$\frac{1}{2}(I + A) \qquad e \qquad \frac{1}{2}(I - A)$$

são idempotentes.

Exercício 2.54 *Mostre que se* A *é uma matriz auto–reflexiva, de ordem* $n \times n$, *então*

$$(I + A)(I - A) = 0_n .$$

Exercício 2.55 *Mostre que as matrizes*

$$A = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 1 & 1 \\ 4 & -1 \end{bmatrix}$$

são anti–comutativas. Assim, temos que $(A + B)^2 = A^2 + B^2$.

Exercício 2.56 *Sejam A e B matrizes quadradas de mesma ordem. Qual a condição que devemos ter para que $(A + B)(A - B) = A^2 - B^2$?*

Exercício 2.57 *Sejam A e B matrizes quadradas de mesma ordem. Qual a condição que devemos ter para que $(A + B)^2 = A^2 + 2AB + B^2$?*

Exercício 2.58 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Deduzir uma fórmula para as potências inteiras positivas da matriz A.

Exercício 2.59 *Dada a matriz*

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix},$$

determine as matriz B, de ordem 2, tais que $AB = BA$.

Exercício 2.60 *Sejam $X = [x_{i1}]$ e $Y = [y_{i1}]$ matrizes coluna de ordem $n \times 1$. Mostre que $tr(XY^t) = X^tY$.*

Exercício 2.61 *Seja $A = [a_{ij}]$ uma matriz real de ordem $n \times n$. Mostre que*

(a) $tr(A^tA) \geq 0$.

(b) $tr(A^tA) = 0$ *se, e somente se,* $A = 0_n$.

Exercício 2.62 *Dada a matriz*

$$A = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \qquad para \qquad \theta \in \mathbb{R}.$$

(a) Determine A^2 e A^3.

(b) Faça a dedução de uma expressão para A^k, $k \in \mathbb{N}$, se possível.

2.3 Inversa de uma Matriz

Definição 2.3.1 *Se A e B são matrizes quadradas de mesma ordem tais que*

$$AB = BA = I \, ,$$

dizemos que B é a ***inversa*** *de A e escrevemos $B = A^{-1}$. De modo análogo, temos que a matriz A é a inversa da matriz B e podemos escrever $A = B^{-1}$. Uma matriz que possui inversa dizemos que é* ***invertível****. Caso contrário, dizemos que a matriz é* ***não–invertível****.*

Exemplo 2.3.1 *As matrizes A e B dadas por:*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

satisfazem $AB = BA = I$. Logo, uma é a inversa da outra.

Teorema 2.3.1 *Sejam A e B matrizes quadradas de mesma ordem com inversas A^{-1} e B^{-1}, respectivamente. Então, $(AB)^{-1} = B^{-1}A^{-1}$.*

Demonstração – Por definição, temos que

$$(AB)^{-1}\,(AB) = (AB)\,(AB)^{-1} = I \, .$$

Desse modo, podemos escrever

$$(B^{-1}A^{-1})\,(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I \, .$$

Por outro lado, temos que

$$(AB)\,(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I \, .$$

Portanto, provamos que $(AB)^{-1} = B^{-1}A^{-1}$. ■

Teorema 2.3.2 *Seja A uma matriz quadrada com inversa A^{-1}. Então,*

$$(A^{-1})^t = (A^t)^{-1} \, .$$

Demonstração – Sabemos que $AA^{-1} = I$ e $A^{-1}A = I$. Assim, calculando suas transpostas, obtemos

$$(AA^{-1})^t = (A^{-1})^t A^t = I \qquad \text{e} \qquad (A^{-1}A)^t = A^t(A^{-1})^t = I \, .$$

Desse modo, temos que $(A^{-1})^t = (A^t)^{-1}$, o que completa a demonstração. ■

Teorema 2.3.3 *Sejam* A, B *e* C *matrizes quadradas tais que*

$$AB = I \qquad e \qquad CA = I\,.$$

Então, $B = C = A^{-1}$ *é a **única** inversa da matriz* A.

Demonstração – Como $CA = I$ e $AB = I$, temos que

$$(CA)B = C(AB) \qquad \Longrightarrow \qquad B = C\,.$$

Portanto, pela Definição 2.3.1, temos que $B = C = A^{-1}$. Assim, mostramos que a inversa da matriz A é **única**. ■

Exemplo 2.3.2 *Dada a matriz*

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}.$$

Determine a matriz A^{-1}, *se possível.*

Sabendo que a inversa da matriz A é única, caso exista, vamos representar a matriz A^{-1} da seguinte forma:

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

para em seguida utilizar o fato que $AA^{-1} = I_2$, isto é,

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Assim, temos que obter a solução de dois sistemas lineares

$$\begin{cases} 2a + 3c = 1 \\ 3a + 4c = 0 \end{cases} \qquad e \qquad \begin{cases} 2b + 3d = 0 \\ 3b + 4d = 1 \end{cases}$$

que são equivalentes aos seguintes sistemas lineares, respectivamente,

$$\begin{cases} 6a + 9c = 3 \\ c = 3 \end{cases} \qquad e \qquad \begin{cases} 6b + 9d = 0 \\ d = -2 \end{cases}$$

que possuem solução única. Portanto, obtemos

$$A^{-1} = \begin{bmatrix} -4 & 3 \\ 3 & -2 \end{bmatrix},$$

mostrando também a sua unicidade.

Exercícios

Exercício 2.63 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 8 \end{bmatrix}.$$

Determine a matriz A^{-1}.

Exercício 2.64 *Considere a matriz real* A *dada por:*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad com \qquad ad - bc \neq 0 .$$

Mostre que

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Exercício 2.65 *Sejam* A, B *e* C *matrizes quadradas de mesma ordem com inversas* A^{-1}, B^{-1} *e* C^{-1}*, respectivamente. Mostre que* $(ABC)^{-1} = C^{-1} B^{-1} A^{-1}$.

Exercício 2.66 *Seja* A *uma matriz quadrada com inversa* A^{-1} *. Mostre que*

$$(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$$

para qualquer escalar λ *não–nulo.*

Exercício 2.67 *Seja* $D = diag(a_{11}, \cdots, a_{nn})$ *uma matriz diagonal, de ordem* n*, com os elementos* $a_{ii} \neq 0$ *para* $i = 1, \cdots, n$*. Mostre que*

$$D^{-1} = diag\left(\frac{1}{a_{11}}, \cdots, \frac{1}{a_{nn}} \right)$$

Exercício 2.68 *Determine a inversa da matriz* A *definida por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Exercício 2.69 *Sejam* A *e* B *matrizes quadradas de mesma ordem e* B *com inversa* B^{-1}*. Mostre que* $tr(B^{-1}AB) = tr(A)$*.*

Exercício 2.70 *Sejam A e B matrizes quadradas de mesma ordem tais que AB é uma matriz invertível. Mostre que as matrizes A e B são invertíveis.*

Exercício 2.71 *Sejam A e B matrizes quadradas não–nulas, de ordem n, tais que $AB = 0_n$. Mostre que as matrizes A e B são não–invertíveis.*

Exercício 2.72 *Seja A uma matriz quadrada complexa com inversa A^{-1} . Mostre que*

$$(\,\overline{A}\,)^{-1} = \overline{(A^{-1})}\,.$$

Exercício 2.73 *Seja A uma matriz de ordem n tal que $A^4 = 0_4$. Mostre que*

$$(I_4 - A)^{-1} = I_4 + A + A^2 + A^3\,.$$

onde $I_4 \in I\!M_4(I\!R)$ é a matriz identidade e $0_4 \in I\!M_4(I\!R)$ é a matriz nula.

Exercício 2.74 *Seja A uma matriz nilpotente de ordem n. Mostre que a matriz $(I_n - A)$ é invertível, exibindo sua matriz inversa.*

Exercício 2.75 *Sejam A e B matrizes de ordem n. Mostre que*

(a) Se $AB = I_n$, então $BA = I_n$.

(b) Se $BA = I_n$, então $AB = I_n$.

Exercício 2.76 *Determine a matriz A^{-1}, se possível, da matriz A dada por:*

$$A = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \qquad para \qquad \theta \in I\!R\,.$$

Exercício 2.77 *Seja X uma matriz coluna de ordem $n \times 1$ tal que $X^tX = 1$. A matriz H, de ordem n, definida por:*

$$H = I_n - 2XX^t$$

*é denominada **matriz de Householder**. Mostre que*

(a) H é uma matriz simétrica.

(b) $H^tH = I_n$.

(c) $H^{-1} = H^t$.

Dê um exemplo de uma matriz de Householder de ordem 3.

2.4 Matrizes em Blocos

Definição 2.4.1 *Dizemos que uma matriz* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *é uma* ***matriz em blocos*** *quando podemos particionar linhas e colunas da seguinte forma:*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qr} \end{bmatrix},$$

onde cada matriz $A_{\alpha\beta}$ *é de ordem* $m_\alpha \times n_\beta$*, com*

$$m_1 + \cdots + m_q = m \qquad e \qquad n_1 + \cdots + n_r = n\,.$$

Exemplo 2.4.1 *Considere a matriz em blocos* $A \in \mathbb{M}_{3 \times 5}(\mathbb{R})$ *definida na forma:*

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix},$$

onde as matrizes $A_{\alpha\beta}$ *são dadas por:*

$$A_{11} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \quad , \quad A_{12} = \begin{bmatrix} 0 & 3 \\ 1 & 2 \end{bmatrix} \quad , \quad A_{13} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} 3 & 1 \end{bmatrix} \quad , \quad A_{22} = \begin{bmatrix} 2 & 4 \end{bmatrix} \quad , \quad A_{23} = \begin{bmatrix} -8 \end{bmatrix}$$

com $m_1 = 2$*,* $m_2 = 1$*,* $n_1 = 2$*,* $n_2 = 2$ *e* $n_3 = 1$*. Assim, temos que*

$$m_1 + m_2 = 3 \qquad e \qquad n_1 + n_2 + n_3 = 5\,.$$

Portanto, a matriz $A \in \mathbb{M}_{3 \times 5}(\mathbb{R})$ *é dada por:*

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 2 & 1 & 2 & -3 \\ 3 & 1 & 2 & 4 & -8 \end{bmatrix}.$$

Finalmente, é importante observar que podemos particionar a matriz A *em blocos de diversas maneiras.*

Exemplo 2.4.2 *Considere a matriz em blocos* $A \in \mathbb{M}_4(\mathbb{R})$ *definida na forma:*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

onde as matrizes $A_{\alpha\beta}$ *são dadas por:*

$$A_{11} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \\ 2 & 4 & 0 \end{bmatrix} \quad , \quad A_{12} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad , \quad A_{21} = \begin{bmatrix} 2 & 1 & 4 \end{bmatrix} \quad , \quad A_{22} = \begin{bmatrix} 5 \end{bmatrix}$$

com $m_1 = 3$, $m_2 = 1$, $n_1 = 3$ *e* $n_2 = 1$. *Assim, temos que*

$$m_1 + m_2 = 4 \qquad e \qquad n_1 + n_2 = 4 .$$

Portanto, a matriz $A \in \mathbb{M}_4(\mathbb{R})$ *é dada por:*

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & 1 & 2 \\ 2 & 4 & 0 & 0 \\ 2 & 1 & 4 & 5 \end{bmatrix}.$$

Exemplo 2.4.3 *Considere a matriz em blocos* $A \in \mathbb{M}_4(\mathbb{R})$ *definida na forma:*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

onde as matrizes $A_{\alpha\beta}$ *são dadas por:*

$$A_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad , \quad A_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad A_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad A_{22} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$$

com $m_1 = 2$, $m_2 = 2$, $n_1 = 2$ *e* $n_2 = 2$. *Assim, temos que*

$$m_1 + m_2 = 4 \qquad e \qquad n_1 + n_2 = 4 .$$

Portanto, a matriz $A \in \mathbb{M}_4(\mathbb{R})$ *é dada por:*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

Definição 2.4.2 *Dizemos que uma matriz A é uma **matriz quadrada em blocos** se*

(a) A é uma matriz quadrada.

(b) Os blocos formam uma matriz quadrada.

(c) O blocos diagonais são matrizes quadradas.

Definição 2.4.3 *Dizemos que uma matriz quadrada em blocos $D \in \mathbb{M}_n(\mathbb{R})$ é uma **matriz diagonal em blocos** se os blocos não diagonais são matrizes nulas. Denotamos a matriz diagonal em blocos da seguinte forma:*

$$D = \begin{bmatrix} D_{11} & & & \\ & D_{22} & & \\ & & \ddots & \\ & & & D_{rr} \end{bmatrix},$$

onde cada matriz $D_{\alpha\alpha}$ é de ordem $n_\alpha \times n_\alpha$, com $n_1 + \cdots + n_r = n$.

Em geral, representamos a matriz diagonal em bloco D da forma:

$$D = D_{11} \oplus D_{22} \oplus \cdots \oplus D_{rr} = \oplus \sum_{i=1}^{r} D_{ii} ,$$

que também é denominada **soma direta** das matrizes $D_{11}, \cdots, D_{rr}$.

Exemplo 2.4.4 *A matriz do Exemplo 2.4.3 é uma matriz diagonal em blocos.*

Definição 2.4.4 *Dizemos que uma matriz quadrada em blocos $L \in \mathbb{M}_n(\mathbb{R})$ é uma **matriz triangular inferior em blocos** se os blocos acima da diagonal principal são matrizes nulas.*

Exemplo 2.4.5 *A matriz quadrada em blocos $L \in \mathbb{M}_4(\mathbb{R})$ definida na forma:*

$$L = \begin{bmatrix} L_{11} & 0_2 \\ L_{21} & L_{22} \end{bmatrix},$$

onde $0_2 \in \mathbb{M}_2(\mathbb{R})$ é a matriz nula, e as matrizes $L_{\alpha\beta}$ são dadas por:

$$L_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad , \quad L_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad e \quad L_{22} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix},$$

é uma matriz triangular inferior em blocos.

Portanto, a matriz $L \in I\!M_4(I\!R)$ *é dada por:*

$$L = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 5 \\ 0 & 1 & 1 & 3 \end{bmatrix}.$$

Definição 2.4.5 *Dizemos que uma matriz quadrada em blocos* $U \in I\!M_n(I\!R)$ *é uma* ***matriz triangular superior em blocos*** *se os blocos abaixo da diagonal principal são matrizes nulas.*

Exemplo 2.4.6 *A matriz quadrada em blocos* $U \in I\!M_4(I\!R)$ *definida na forma:*

$$U = \begin{bmatrix} U_{11} & U_{12} \\ 0_4 & U_{22} \end{bmatrix},$$

onde as matrizes $U_{\alpha\beta}$ *são dadas por:*

$$U_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \ , \ U_{12} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad e \quad U_{22} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix},$$

é uma matriz triangular superior em blocos.

Portanto, a matriz $U \in I\!M_4(I\!R)$ *é dada por:*

$$U = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

Definição 2.4.6 *Sejam* $A, B \in I\!M_{m \times n}(I\!R)$ *matrizes em blocos dadas por:*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qr} \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \vdots \\ B_{q1} & \cdots & B_{qr} \end{bmatrix},$$

onde as matrizes $A_{\alpha\beta}, B_{\alpha\beta}$ *são de ordem* $m_\alpha \times n_\beta$*, com*

$$m_1 + \cdots + m_q = m \qquad e \qquad n_1 + \cdots + n_r = n \,.$$

*Definimos a **soma*** $C = A + B$ *da seguinte forma:*

$$C = \begin{bmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & \ddots & \vdots \\ C_{q1} & \cdots & C_{qr} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & \cdots & A_{1r} + B_{1r} \\ \vdots & \ddots & \vdots \\ A_{q1} + B_{q1} & \cdots & A_{qr} + B_{qr} \end{bmatrix},$$

que é uma matriz em blocos, onde cada matriz $C_{\alpha\beta}$ *é de ordem* $m_\alpha \times n_\beta$.

Lema 2.4.1 *Sejam* $A \in \mathbb{M}_{m\times p}(\mathbb{R})$ *e* $B \in \mathbb{M}_{p\times n}(\mathbb{R})$ *matrizes em blocos dadas por:*

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_q \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} B_1 & \cdots & B_r \end{bmatrix},$$

onde cada matriz A_α *é de ordem* $m_\alpha \times p$, *com*

$$m_1 + \cdots + m_q = m,$$

e cada matriz B_β *é de ordem* $p \times n_\beta$, *com*

$$n_1 + \cdots + n_r = n.$$

*Então, o **produto*** $C = AB$, *que é uma matriz em blocos, é definido na forma:*

$$C = \begin{bmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & \ddots & \vdots \\ C_{q1} & \cdots & C_{qr} \end{bmatrix},$$

onde cada matriz $C_{\alpha\beta} = A_\alpha B_\beta$ *é de ordem* $m_\alpha \times n_\beta$.

Demonstração – Veja Lema 1.3.1, página 25, da referência [11]. □

Lema 2.4.2 *Sejam* $A \in I\!M_{m \times p}(I\!R)$ *e* $B \in I\!M_{p \times n}(I\!R)$ *matrizes em blocos dadas por:*

$$A = \begin{bmatrix} A_1 & \cdots & A_q \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} B_1 \\ \vdots \\ B_q \end{bmatrix},$$

onde as matrizes A_γ *são de ordem* $m \times n_\gamma$ *e as matrizes* B_γ *são de ordem* $n_\gamma \times n$, *com* $n_1 + \cdots + n_q = p$.

Então, o ***produto*** $C = AB$, *que é uma matriz em blocos, é definido na forma:*

$$C = \sum_{\gamma=1}^{q} A_\gamma B_\gamma ,$$

onde cada matriz $A_\gamma B_\gamma \in I\!M_{m \times n}(I\!R)$ *para* $\gamma = 1, \cdots, q$.

Demonstração – Veja Lema 1.3.2, página 26, da referência [11]. □

Lema 2.4.3 *Sejam* $A \in \mathbb{M}_{m \times p}(\mathbb{R})$ *e* $B \in \mathbb{M}_{p \times n}(\mathbb{R})$ *matrizes em blocos dadas por:*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qs} \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \vdots \\ B_{s1} & \cdots & B_{sr} \end{bmatrix},$$

onde cada matriz $A_{\alpha\gamma}$ *é de ordem* $m_\alpha \times l_\gamma$, *com*

$$m_1 + \cdots + m_q = m \qquad e \qquad l_1 + \cdots + l_s = p\,,$$

e cada matriz $B_{\gamma\beta}$ *é de ordem* $l_\gamma \times n_\beta$, *com*

$$n_1 + \cdots + n_r = n\,.$$

Então, o ***produto*** $C = AB$, *que é uma matriz em blocos, é definido na forma:*

$$C = \begin{bmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & \ddots & \vdots \\ C_{q1} & \cdots & C_{qr} \end{bmatrix},$$

que é uma matriz em blocos, onde cada matriz $C_{\alpha\beta}$ *é dada por:*

$$C_{\alpha\beta} = \sum_{\gamma=1}^{s} A_{\alpha\gamma} B_{\gamma\beta}\,,$$

que é de ordem $m_\alpha \times n_\beta$, *para*

$$\alpha = 1, \cdots, q \qquad e \qquad \beta = 1, \cdots, r\,.$$

Demonstração – Veja Teorema 1.3.3, página 26, da referência [11]. □

Exemplo 2.4.7 *Sejam a matriz em blocos A e o vetor coluna em blocos X dados por:*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad e \qquad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},$$

onde cada matriz $A_{\alpha\beta} \in I\!M_n(I\!R)$ e cada vetor coluna $X_\beta \in I\!M_{n\times 1}(I\!R)$.

Assim, o produto $Y = AX$ é escrito da seguinte forma:

$$Y = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} A_{11}X_1 + A_{12}X_2 \\ A_{21}X_1 + A_{22}X_2 \end{bmatrix},$$

de acordo com o Lema 2.4.3.

Para exemplificar, considere a matriz em blocos $A \in I\!M_4(I\!R)$ definida na forma:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

onde as matrizes $A_{\alpha\beta} \in I\!M_2(I\!R)$ são dadas por:

$$A_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} , \ A_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} , \ A_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} , \ A_{22} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix},$$

e o vetor coluna $X \in I\!M_{4\times 1}(I\!R)$ definido na forma:

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \qquad com \qquad X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad e \qquad X_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Assim, o produto $Y = AX$ é escrito da seguinte forma:

$$Y = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} A_{11}X_1 \\ A_{22}X_2 \end{bmatrix}.$$

Assim, o vetor coluna $Y \in I\!M_{4\times 1}(I\!R)$ é dado por:

$$Y = \begin{bmatrix} 2 \\ 3 \\ 3 \\ 2 \end{bmatrix}.$$

Exemplo 2.4.8 *Considere a matriz diagonal em blocos* A *definida na forma:*

$$A = \begin{bmatrix} A_{11} & 0_n \\ 0_n & A_{22} \end{bmatrix},$$

onde $0_n \in \mathbb{M}_n(\mathbb{R})$ *é a matriz nula, e as matrizes* $A_{\alpha\alpha} \in \mathbb{M}_n(\mathbb{R})$ *são invertíveis.*

Desse modo, a matriz em blocos B *definida na forma:*

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

tal que

$$AB = BA = \begin{bmatrix} I_n & 0_n \\ 0_n & I_n \end{bmatrix},$$

onde $I_n \in \mathbb{M}_n(\mathbb{R})$ *é a matriz identidade, é a inversa da matriz* A.

De acordo com o Lema 2.4.3, temos que o produto AB *é dado por:*

$$AB = \begin{bmatrix} A_{11}\,B_{11} & A_{11}\,B_{12} \\ A_{22}\,B_{21} & A_{22}\,B_{22} \end{bmatrix} = \begin{bmatrix} I_n & 0_n \\ 0_n & I_n \end{bmatrix}.$$

Portanto, temos que

$$\begin{aligned}
A_{11}\,B_{11} &= I_n & \iff && B_{11} &= A_{11}^{-1} \\
A_{11}\,B_{12} &= 0_n & \iff && B_{12} &= A_{11}^{-1}\,0_n = 0_n \\
A_{22}\,B_{21} &= 0_n & \iff && B_{21} &= A_{22}^{-1}\,0_n = 0_n \\
A_{22}\,B_{22} &= I_n & \iff && B_{22} &= A_{22}^{-1}
\end{aligned}$$

Assim, obtemos a matriz diagonal em blocos

$$B = \begin{bmatrix} A_{11}^{-1} & 0_n \\ 0_n & A_{22}^{-1} \end{bmatrix}$$

que é a inversa da matriz A.

Exemplo 2.4.9 *Considere a matriz diagonal em blocos* $A \in I\!M_4(I\!R)$ *definida na forma:*

$$A = \begin{bmatrix} A_{11} & 0_2 \\ 0_2 & A_{22} \end{bmatrix},$$

onde $0_2 \in I\!M_2(I\!R)$ *é a matriz nula, e as matrizes* $A_{\alpha\alpha} \in I\!M_2(I\!R)$ *são dadas por:*

$$A_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad e \qquad A_{22} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}.$$

Assim, a matriz diagonal em blocos $A^{-1} \in I\!M_4(I\!R)$ *definida na forma:*

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & 0_2 \\ 0_2 & A_{22}^{-1} \end{bmatrix},$$

é a inversa da matriz A*, onde as matrizes* $A_{\alpha\alpha}^{-1} \in I\!M_2(I\!R)$ *são dadas por:*

$$A_{11}^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \qquad e \qquad A_{22}^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}.$$

Portanto, temos que

$$AA^{-1} = A^{-1}A = \begin{bmatrix} I_2 & \\ & I_2 \end{bmatrix},$$

onde $I_2 \in I\!M_2(I\!R)$ *é a matriz identidade.*

Portanto, as matrizes $A\,,\, A^{-1} \in I\!M_4(I\!R)$ *são dadas por:*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 \end{bmatrix} \qquad e \qquad A^{-1} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 3 & -5 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Lema 2.4.4 *Considere a matriz em blocos* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *dada na seguinte forma:*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qr} \end{bmatrix},$$

onde cada matriz $A_{\alpha\beta}$ *é de ordem* $m_\alpha \times n_\beta$*, com*

$$m_1 + \cdots + m_q = m \qquad e \qquad n_1 + \cdots + n_r = n \,.$$

Então, a matriz em blocos $A^t \in \mathbb{M}_{n \times m}(\mathbb{R})$ *é definida na forma:*

$$A^t = \begin{bmatrix} A_{11}^t & \cdots & A_{q1}^t \\ \vdots & & \vdots \\ A_{1r}^t & \cdots & A_{qr}^t \end{bmatrix}.$$

Demonstração – A prova pode ficar a carga do leitor. $\square$

Exemplo 2.4.10 *Considere a matriz em blocos* $A \in I\!M_{3\times5}(I\!R)$ *definida na forma:*

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix},$$

onde as matrizes $A_{\alpha\beta}$ *são dadas por:*

$$A_{11} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \quad , \quad A_{12} = \begin{bmatrix} 0 & 3 \\ 1 & 2 \end{bmatrix} \quad , \quad A_{13} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} 3 & 1 \end{bmatrix} \quad , \quad A_{22} = \begin{bmatrix} 2 & 4 \end{bmatrix} \quad , \quad A_{23} = \begin{bmatrix} -8 \end{bmatrix}$$

Desse modo, a matriz $A^t \in I\!M_{5\times3}(I\!R)$ *é dada por:*

$$A^t = \begin{bmatrix} A^t_{11} & A^t_{21} \\ A^t_{12} & A^t_{22} \\ A^t_{13} & A^t_{23} \end{bmatrix}$$

onde as matrizes $A^t_{\alpha\beta}$ *são dadas por:*

$$A^t_{11} = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \quad , \quad A^t_{12} = \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix} \quad , \quad A^t_{13} = \begin{bmatrix} 1 & -3 \end{bmatrix}$$

$$A^t_{21} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad , \quad A^t_{22} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad , \quad A^t_{23} = \begin{bmatrix} -8 \end{bmatrix}$$

Assim, obtemos

$$A^t = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 2 & 1 \\ 0 & 1 & 2 \\ 3 & 2 & 4 \\ 1 & -3 & -8 \end{bmatrix}$$

Exemplo 2.4.11 *Sejam A uma matriz normal real de ordem $n \times n$ e B uma matriz normal real de ordem $m \times m$. Vamos mostrar que a matriz em blocos dada por:*

$$C = \begin{bmatrix} A & 0_{n\times m} \\ 0^t_{n\times m} & B \end{bmatrix}$$

é uma matriz normal de ordem $(n+m)$, onde $0_{n\times m}$ é a matriz nula de ordem $n \times m$.

Assim, temos que

$$CC^t = \begin{bmatrix} A & 0_{n\times m} \\ 0^t_{n\times m} & B \end{bmatrix} \begin{bmatrix} A^t & 0_{n\times m} \\ 0^t_{n\times m} & B^t \end{bmatrix} = \begin{bmatrix} AA^t & 0_{n\times m} \\ 0^t_{n\times m} & BB^t \end{bmatrix}$$

$$C^tC = \begin{bmatrix} A^t & 0_{n\times m} \\ 0^t_{n\times m} & B^t \end{bmatrix} \begin{bmatrix} A & 0_{n\times m} \\ 0^t_{n\times m} & B \end{bmatrix} = \begin{bmatrix} A^tA & 0_{n\times m} \\ 0^t_{n\times m} & B^tB \end{bmatrix}$$

Como, por hipótese, $A^tA = AA^t$ e $B^tB = BB^t$, obtemos $CC^t = C^tC$. Logo, mostramos que a matriz em blocos C é uma matriz normal.

Exemplo 2.4.12 *Podemos verificar facilmente que as matrizes*

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 4 \\ 0 & 4 & 5 \end{bmatrix} \quad e \quad B = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$$

são matrizes normais.

Portanto, a matriz em blocos dada por:

$$C = \begin{bmatrix} A & 0_{3\times 2} \\ 0_{2\times 3} & B \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 6 & 4 & 0 & 0 \\ 0 & 4 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 3 & 2 \end{bmatrix}$$

é também uma matriz normal.

2.5 Operações Elementares. Equivalência

Definição 2.5.1 *As **operações elementares** com matrizes, são operações que mantém tanto a ordem da matriz quanto a sua característica. Vamos definir dois tipos de operações elementares. As operações elementares de linhas, que vamos indicar por h, e as operações elementares de colunas, que vamos indicar por k.*

*São **operações elementares de linhas**:*

(a) Permutação da i–ésima linha com a j–ésima linha, que indicaremos por:

$$h : l_i \longleftrightarrow l_j \,.$$

(b) Multiplicação da i–ésima linha por um escalar r não–nulo, que indicaremos por:

$$h : l_i \longleftarrow rl_i \,.$$

(c) Substituição da i–ésima linha pela i–ésima linha mais a j–ésima linha multiplicada por um escalar r não–nulo, que indicaremos por:

$$h : l_i \longleftarrow l_i + rl_j \,.$$

*De modo análogo, definimos os mesmos tipos de operações elementares com as colunas da matriz, que são denominadas **operações elementares de colunas**.*

*São **operações elementares de colunas**:*

(a) Permutação da i–ésima coluna com a j–ésima coluna, que indicaremos por:

$$k : c_i \longleftrightarrow c_j \,.$$

(b) Multiplicação da i–ésima coluna por um escalar r não–nulo, que indicaremos por:

$$k : c_i \longleftarrow rc_i \,.$$

(c) Substituição da i–ésima coluna pela i–ésima coluna mais a j–ésima coluna multiplicada por um escalar r não–nulo, que indicaremos por:

$$k : c_i \longleftarrow c_i + rc_j \,.$$

Vamos nos dedicar mais às operações elementares de linhas, pois temos como objetivo central suas aplicações na análise de soluções de sistemas de equações lineares.

Exemplo 2.5.1 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 5 & 5 \\ 1 & 2 & 3 \end{bmatrix},$$

a operação elementar de linhas

$$h : l_2 \longleftarrow l_2 - 3l_1$$

e a operação elementar de colunas

$$k : c_2 \longleftarrow c_2 + c_3 \,.$$

Portanto, aplicando a seqüência $k(h(A))$ *obtemos a seguinte matriz resultante*

$$C = k(h(A)) = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & -1 \\ 1 & 5 & 3 \end{bmatrix}.$$

Podemos verificar facilmente que $C = h(k(A))$.

Exemplo 2.5.2 *Dada a matriz*

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 3 & 4 \\ 3 & 1 & 1 \end{bmatrix},$$

vamos aplicar a seguinte seqüência de operações elementares de linhas

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & 3 & 4 \\ 3 & 1 & 1 \end{bmatrix} \quad l_2 \longleftarrow l_2 - 2l_1 \quad \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & 0 \\ 3 & 1 & 1 \end{bmatrix} \quad l_3 \longleftarrow l_3 - 3l_1 \quad \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & 0 \\ 0 & 4 & -5 \end{bmatrix}$$

$$l_3 \longleftarrow 5l_3 \quad \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & 0 \\ 0 & 20 & -25 \end{bmatrix} \quad l_3 \longleftarrow l_3 - 4l_2 \quad \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & 0 \\ 0 & 0 & -25 \end{bmatrix}.$$

Assim, encontramos uma matriz triangular superior

$$U = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 5 & 0 \\ 0 & 0 & -25 \end{bmatrix},$$

obtida da matriz A *através de operações elementares de linhas.*

Definição 2.5.2 *A* ***operação elementar inversa*** *é uma operação que desfaz o efeito da operação elementar, isto é, depois de haver realizado uma operação elementar sobre uma matriz, aplicando sobre a matriz resultante a operação elementar inversa retornamos à matriz original.*

Exemplo 2.5.3 *Considere as seguintes operações elementares de linhas*

(a) $h : l_i \longleftarrow l_i + cl_j$ ***(b)*** $h : l_i \longleftarrow rl_i$ ***(c)*** $h : l_i \longleftrightarrow l_j$

onde os escalares c *e* r *são não–nulos.*

As respectivas operações elementares inversas são dadas por:

(a) $h_1 : l_i \longleftarrow l_i - cl_j$ ***(b)*** $h_1 : l_i \longleftarrow \dfrac{1}{r}l_i$ ***(c)*** $h_1 : l_i \longleftrightarrow l_j$

o que pode ser facilmente verificada.

Exemplo 2.5.4 *Considere a seguinte seqüência de operações elementares de linhas*

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - 3l_1 \qquad e \qquad l_2 \longleftarrow \frac{1}{5}l_2 \, .$$

Desse modo, a seqüência de operações elementares inversas é dada por:

$$l_2 \longleftarrow 5l_2 \quad , \quad l_3 \longleftarrow l_3 + 3l_1 \qquad e \qquad l_2 \longleftarrow l_2 + 2l_1 \, .$$

Exemplo 2.5.5 *Dada a matriz*

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 3 & 4 \\ 3 & 1 & 1 \end{bmatrix} .$$

Aplicando a seqüência de operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - 3l_1 \quad , \quad l_2 \longleftarrow \frac{1}{5}l_2 \qquad e \qquad l_3 \longleftarrow l_3 - 4l_2 \, ,$$

na matriz A*, obtemos a seguinte matriz resultante*

$$B = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix} .$$

Finalmente, aplicando a seqüência de operações elementares inversas

$$l_3 \longleftarrow l_3 + 4l_2 \quad , \quad l_2 \longleftarrow 5l_2 \quad , \quad l_3 \longleftarrow l_3 + 3l_1 \qquad e \qquad l_2 \longleftarrow l_2 + 2l_1 \, .$$

na matriz B*, obtemos novamente a matriz* A*.*

Assim, podemos verificar facilmente que a operação inversa de uma operação elementar de linhas é uma operação elementar de linhas do mesmo tipo. Desse modo, temos que

$$h_1(h(A)) \;=\; h(h_1(A)) \;=\; A\,.$$

De modo análogo, a operação inversa de uma operação elementar de colunas é uma operação elementar de colunas do mesmo tipo.

Definição 2.5.3 *Sejam A e B matrizes de mesma ordem. Dizemos que a matriz B é* ***linha equivalente*** *a matriz A, se a matriz B pode ser obtida da matriz A através de uma seqüência finita de operações elementares sobre as linhas de A.*

Exemplo 2.5.6 *Considere a matriz A, de ordem 3×4, dada por:*

$$A \;=\; \begin{bmatrix} 1 & 4 & 2 & 1 \\ 2 & 1 & -1 & 1 \\ 4 & -5 & -7 & 1 \end{bmatrix}.$$

Aplicando a seqüência de operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - 4l_1 \qquad e \qquad l_3 \longleftarrow l_3 - 3l_2$$

na matriz A, obtemos a matriz

$$B \;=\; \begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & -7 & -5 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

que é linha equivalente a matriz A.

Definição 2.5.4 *Sejam A e B matrizes de mesma ordem. Dizemos que a matriz B é* ***equivalente por coluna*** *a matriz A, se a matriz B pode ser obtida da matriz A através de uma seqüência finita de operações elementares sobre as colunas de A.*

Definição 2.5.5 *Sejam A e B matrizes de mesma ordem. Dizemos que a matriz B é* ***equivalente*** *a matriz A, se a matriz B pode ser obtida da matriz A através de uma seqüência finita de operações elementares sobre as linhas e sobre as colunas de A. Indicamos $B \sim A$ para denotar que a matriz B é equivalente a matriz A.*

Exercícios

Exercício 2.78 *Mostre que as matrizes*

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} \qquad e \qquad U = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -5 & -5 \\ 0 & 0 & 3 \end{bmatrix}$$

são equivalentes, indicando a seqüência de operações elementares de linhas utilizada para reduzir a matriz A a matriz triangular superior U.

Exercício 2.79 *Mostre que as matrizes*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -2 & 0 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 1 & 2 \\ 0 & -2 \\ 0 & 0 \end{bmatrix}$$

são equivalentes, indicando a seqüência de operações elementares de linhas utilizada para reduzir a matriz A a matriz B.

Exercício 2.80 *Mostre que as matrizes*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -2 & 0 \end{bmatrix} \qquad e \qquad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

são equivalentes, indicando a seqüência de operações elementares utilizada.

Exercício 2.81 *Mostre que as matrizes*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 4 \end{bmatrix} \qquad e \qquad R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

são equivalentes, indicando a seqüência de operações elementares utilizada.

2.6 Forma Escalonada. Forma Escada

Definição 2.6.1 *Uma matriz R, de ordem $m \times n$, está na **forma escalonada**, linha reduzida, se prevalecem as seguintes condições:*

(a) Todas as linhas nulas, se houver, aparecem nas últimas linhas da matriz.

*(b) O primeiro elemento não–nulo de uma linha, que é denominado **pivô**, está à direita do primeiro elemento não–nulo da linha anterior.*

Exemplo 2.6.1 *Nos Exemplos 2.5.5 e 2.5.6 efetuamos uma seqüência de operações elementares de linhas na matriz A com o objetivo de obter uma matriz B na forma escalonada, linha equivalente a matriz A.*

Definição 2.6.2 *Uma matriz R, de ordem $m \times n$, na **forma escalonada** está na **forma escada**, linha reduzida, se prevalecem mais as seguintes condições:*

(c) O primeiro elemento não–nulo de uma linha não–nula de R é igual a 1.

(d) Cada coluna de R que contém o primeiro elemento não–nulo tem todos os seus outros elementos nulos.

Exemplo 2.6.2 *Um exemplo de uma matriz de ordem n na forma escada é a matriz identidade I_n. De fato, podemos verificar facilmente que a matriz identidade satisfaz as propriedades exigidas. Para ilustrar, tome como exemplo a matriz*

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A matriz nula $0_{m \times n}$ é um outro exemplo de uma matriz na forma escada.

Exemplo 2.6.3 *Considerando as matrizes A e B do Exemplo 2.5.5. Aplicando a seqüência de operações elementares de linhas*

$$l_3 \longleftarrow -\frac{1}{5} l_3 \quad , \quad l_1 \longleftarrow l_1 + l_2 \qquad e \qquad l_1 \longleftarrow l_1 - 2 l_3$$

na matriz B na forma escalonada, linha equivalente a matriz A, obtemos a matriz $R = I_3$ na forma escada, que é linha equivalente a matriz A.

Exemplo 2.6.4 *Considerando novamente as matrizes A e B do Exemplo 2.5.6. Podemos realizar uma seqüência de operações elementares de linhas na matriz B, que está na forma escalonada, para obter uma matriz R na forma escada, linha equivalente a matriz A. De fato,*

$$B = \begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & -7 & -5 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad l_2 \longleftarrow -\frac{1}{7} l_2 \quad \begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & \frac{5}{7} & \frac{1}{7} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$l_1 \longleftarrow l_1 - 4 l_2 \quad \begin{bmatrix} 1 & 0 & -\frac{6}{7} & \frac{3}{7} \\ 0 & 1 & \frac{5}{7} & \frac{1}{7} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, obtemos a matriz na forma escada

$$R = \begin{bmatrix} 1 & 0 & -\frac{6}{7} & \frac{3}{7} \\ 0 & 1 & \frac{5}{7} & \frac{1}{7} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

que é linha equivalente a matriz A.

Exemplo 2.6.5 *Dada a matriz*

$$A = \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 2 & 2 & 5 & 2 & 14 \\ 6 & 6 & 0 & 20 & 19 \end{bmatrix}.$$

Encontre uma matriz R na forma escalonada, linha equivalente a matriz A, indicando a seqüência de operações elementares de linhas utilizada.

Exemplo 2.6.6 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{bmatrix}.$$

Encontre uma matriz R na forma escada, linha equivalente a matriz A, indicando a seqüência de operações elementares de linhas utilizada.

Definição 2.6.3 *Sejam A uma matriz de ordem $m \times n$ e R a matriz na forma escalonada linha equivalente a matriz A. Definimos o* ***posto linha*** *da matriz A, ou* ***posto de A****, como sendo o número de linhas não–nulas da matriz R, e denotamos esse número inteiro por posto(A).*

Exemplo 2.6.7 *Determine o posto linha da matriz A dada por:*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 4 \\ 1 & 4 & 2 \end{bmatrix},$$

e também o posto linha da matriz A^t.

Exemplo 2.6.8 *Determine o posto linha da matriz A dada por:*

$$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 8 & 4 & 1 \\ 1 & 4 & 2 & 1 \end{bmatrix},$$

e também o posto linha da matriz A^t.

Exemplo 2.6.9 *Determine o posto linha da matriz A dada por:*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 4 \\ 1 & 4 & 2 \\ 1 & 1 & 1 \end{bmatrix},$$

e também o posto linha da matriz A^t.

Na seção 8.10 apresentamos um estudo mais detalhado sobre os resultados envolvendo o **posto de A**, onde demonstraremos o fato observado nos exemplos anteriores.

2.7 Matrizes Elementares

Definição 2.7.1 *A matriz resultante da aplicação de uma única operação elementar de linhas à matriz identidade, é denominada **matriz elementar de linha**.*

Definição 2.7.2 *A matriz resultante da aplicação de uma única operação elementar de colunas à matriz identidade, é denominada **matriz elementar de coluna**.*

Exemplo 2.7.1 *Vamos considerar o seguinte exemplo de uma matriz elementar de linha obtida da matriz identidade* I_3*, que denotamos por* H*,*

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad l_2 \longleftarrow l_2 + 2l_1 \quad H = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definição 2.7.3 *A matriz resultante da aplicação de uma única operação elementar de permutação de linhas sobre a matriz identidade, é denominada **matriz de permutação de linhas**.*

Definição 2.7.4 *A matriz resultante da aplicação de uma única operação elementar de permutação de colunas sobre a matriz identidade, é denominada **matriz de permutação de colunas**.*

Exemplo 2.7.2 *Vamos considerar o seguinte exemplo de uma matriz de permutação de linhas obtida da matriz identidade* I_3*, que denotamos por* P*,*

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad l_1 \longleftrightarrow l_3 \quad P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Observamos facilmente que uma matriz de permutação também é uma matriz elementar, pois foi obtida da matriz identidade através de uma única operação elementar.

Seja h uma operação elementar de linhas, denotamos por $H = h(I_n)$ a matriz elementar de linha correspondente à operação elementar h. De modo análogo, se k é uma operação elementar de colunas, vamos denotar por $K = k(I_n)$ a matriz elementar de coluna correspondente à operação elementar k.

Lema 2.7.1 *Sejam* A *uma matriz de ordem* $m \times n$, B *uma matriz de ordem* $p \times m$ *e* h *uma operação elementar de linhas. Então,* $h(B)A = h(BA)$.

Demonstração – Seja $E_{i\cdot}$ a matriz linha de ordem $1 \times p$ dada por:

$$E_{i\cdot} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} ,$$

onde o valor 1 aparece na i–ésima coluna, que é a i–ésima linha da matriz identidade de ordem $p \times p$. Podemos verificar facilmente que

$$E_{i\cdot}B = \begin{bmatrix} b_{i1} & \cdots & b_{ij} & \cdots & b_{im} \end{bmatrix} = B_{i\cdot} ,$$

onde $B_{i\cdot}$ é a matriz linha de ordem $1 \times m$ que denota a i–ésima linha da matriz B.

Por simplicidade, vamos denotar as matrizes A e B, e a matriz identidade I_p, da seguinte forma:

$$A = \begin{bmatrix} A_{1\cdot} \\ \vdots \\ A_{i\cdot} \\ \vdots \\ A_{m\cdot} \end{bmatrix} \quad , \quad B = \begin{bmatrix} B_{1\cdot} \\ \vdots \\ B_{i\cdot} \\ \vdots \\ B_{p\cdot} \end{bmatrix} \quad \text{e} \quad I_p = \begin{bmatrix} E_{1\cdot} \\ \vdots \\ E_{i\cdot} \\ \vdots \\ E_{p\cdot} \end{bmatrix} ,$$

onde $A_{i\cdot}$ é a matriz linha de ordem $1 \times n$ que denota a i–ésima linha da matriz A.

De modo análogo, podemos verificar facilmente que

$$BA = \begin{bmatrix} B_{1\cdot}A \\ \vdots \\ B_{i\cdot}A \\ \vdots \\ B_{p\cdot}A \end{bmatrix} ,$$

utilizando a definição de multiplicação de matrizes. Note que $B_{i\cdot}A$ indica a multiplicação da i–ésima linha da matriz B pela matriz A, obtendo a i–ésima linha da matriz BA.

A seguir passamos para a demonstração, considerando cada uma das operações elementar de linhas, onde as observações acima serão de muita utilidade.

(1) Considere h como sendo a operação elementar de linhas que multiplica a i–ésima linha por um escalar r não–nulo, isto é, $h : l_i \longleftarrow rl_i$, cuja matriz elementar correspondente é dada por:

$$h(I_p) = H = \begin{bmatrix} E_{1\cdot} \\ \vdots \\ rE_{i\cdot} \\ \vdots \\ E_{p\cdot} \end{bmatrix} .$$

Desse modo, temos que

$$h(B) = \begin{bmatrix} B_{1\cdot} \\ \vdots \\ rB_{i\cdot} \\ \vdots \\ B_{p\cdot} \end{bmatrix} = \begin{bmatrix} E_{1\cdot}B \\ \vdots \\ (rE_{i\cdot})B \\ \vdots \\ E_{p\cdot}B \end{bmatrix} = h(I_p)B = HB .$$

Portanto, temos que $h(B)A = (HB)A = H(BA) = h(BA)$, fazendo uso do fato que multiplicação de matrizes e associativa.

(2) Considere h como sendo a operação elementar de linhas que substitui a i–ésima linha pela i–ésima mais a k–ésima linha multiplicada por um escalar r não–nulo, isto é, $h : l_i \longleftarrow l_i + rl_k$, cuja matriz elementar correspondente é dada por:

$$h(I_p) = H = \begin{bmatrix} E_{1\cdot} \\ \vdots \\ E_{i\cdot} + rE_{k\cdot} \\ \vdots \\ E_{p\cdot} \end{bmatrix} .$$

Desse modo, temos que

$$h(B) = \begin{bmatrix} B_{1\cdot} \\ \vdots \\ B_{i\cdot} + rB_{k\cdot} \\ \vdots \\ B_{p\cdot} \end{bmatrix} = \begin{bmatrix} E_{1\cdot}B \\ \vdots \\ (E_{i\cdot} + rE_{k\cdot})B \\ \vdots \\ E_{p\cdot}B \end{bmatrix} = h(I_p)B = HB .$$

Portanto, temos que $h(B)A = (HB)A = H(BA) = h(BA)$, fazendo uso do fato que multiplicação de matrizes e associativa.

(3) Considere h como sendo a operação elementar de linhas que permuta a i–ésima linha com a k–ésima linha, isto é, $h : l_i \longleftrightarrow l_k$, para $i < k$, cuja matriz elementar correspondente é dada por:

$$h(I_p) = H = \begin{bmatrix} E_{1\cdot} \\ \vdots \\ E_{k\cdot} \\ \vdots \\ E_{i\cdot} \\ \vdots \\ E_{p\cdot} \end{bmatrix} .$$

Desse modo, temos que

$$h(B) = \begin{bmatrix} B_{1\cdot} \\ \vdots \\ B_{k\cdot} \\ \vdots \\ B_{i\cdot} \\ \vdots \\ B_{p\cdot} \end{bmatrix} = \begin{bmatrix} E_{1\cdot}B \\ \vdots \\ E_{k\cdot}B \\ \vdots \\ E_{i\cdot}B \\ \vdots \\ E_{p\cdot}B \end{bmatrix} = h(I_p)B = HB .$$

Portanto, temos que $h(B)A = (HB)A = H(BA) = h(BA)$, fazendo uso do fato que multiplicação de matrizes e associativa. ■

Teorema 2.7.1 *Seja A uma matriz de ordem $m \times n$. A matriz C resultante da aplicação de uma única operação elementar com as linhas da matriz A, é a mesma matriz C resultante da multiplicação pela esquerda da matriz A pela matriz elementar H de ordem $m \times m$, correspondente à operação elementar efetuada com as linhas de A, isto é, $C = HA$.*

Demonstração – A prova segue do Lema 2.7.1, considerando a matriz $B = I_m$. De fato, Seja H a matriz elementar de linha correspondente à operação elementar de linhas h. Desse modo, temos que

$$C = h(A) = h(I_m)A = HA ,$$

o que completa a demonstração. ■

Lema 2.7.2 *Sejam A uma matriz de ordem $m \times n$, k uma operação elementar de colunas e h a operação elementar de linhas correspondente à operação k. Então,*

$$k(A) \;=\; (\,h(A^t)\,)^t\,.$$

Demonstração – A demonstração segue diretamente do fato que as colunas da matriz A são as linhas da matriz A^t, e vice–versa. ■

Corolário 2.7.1 *Sejam k uma operação elementar de colunas, sendo K a matriz elementar de coluna correspondente, e h a operação elementar de linhas análoga à operação k, com H a matriz elementar de linha correspondente à operação elementar de linhas h. Então, $K = H^t$.*

Exemplo 2.7.3 *Vamos considerar a seguinte operação elementar de colunas*

$$k : c_2 \;\longleftarrow\; c_2 + 2c_1$$

e a operação elementar de linhas h correspondente à operação k, isto é,

$$h : l_2 \;\longleftarrow\; l_2 + 2l_1\,.$$

Desse modo, temos as matrizes elementares correspondentes às operações k e h

$$K \;=\; k(I_3) \;=\; \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad e \qquad H \;=\; h(I_3) \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Assim, podemos verificar que $K = H^t$.

Teorema 2.7.2 *Seja A uma matriz de ordem $m \times n$. A matriz C resultante da aplicação de uma única operação elementar com as colunas da matriz A, é a mesma matriz C resultante da multiplicação pela direita da matriz A pela matriz elementar K, de ordem $n \times n$, correspondente à operação elementar efetuada com as colunas de A, isto é, $C = AK$.*

Demonstração – A prova segue do Lema 2.7.2 e do Teorema 2.7.1. De fato, Sejam K a matriz elementar de coluna correspondente à operação elementar de colunas k e H a matriz elementar de linha correspondente à operação elementar de linhas h análoga à operação k. Desse modo, obtemos

$$k(A) \;=\; (\,h(A^t)\,)^t \;=\; (\,HA^t\,)^t \;=\; AH^t \;=\; AK\,,$$

o que completa a demonstração. ■

Exemplo 2.7.4 *Considere a matriz*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 10 & 6 \\ 3 & 0 & 1 \end{bmatrix}.$$

Aplicando a operação elementar de linhas $l_2 \longleftarrow l_2 - 4l_1$ *na matriz* A, *obtemos a matriz resultante* C

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 10 & 6 \\ 3 & 0 & 1 \end{bmatrix} \quad l_2 \longleftarrow l_2 - 4l_1 \quad C = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 2 \\ 3 & 0 & 1 \end{bmatrix}.$$

Equivalentemente, considerando a matriz elementar de linha E *correspondente à operação elementar de linhas, definida acima,*

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad l_2 \longleftarrow l_2 - 4l_1 \quad E = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

obtemos a matriz C *da seguinte forma:*

$$C = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 4 & 10 & 6 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 2 \\ 3 & 0 & 1 \end{bmatrix},$$

isto é, $C = EA$.

Exemplo 2.7.5 *Considerando o Exemplo 2.5.1, vamos denotar por* H *a matriz elementar de linha correspondente a operação elementar de linhas* h *e por* K *a matriz elementar de coluna correspondente a operação elementar de colunas* k. *Desse modo, temos que a matriz* C *pode ser obtida da seguinte forma:*

$$C = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 3 & 5 & 5 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & -1 \\ 1 & 5 & 3 \end{bmatrix}.$$

Portanto, temos que $C = (HA)K = H(AK)$, *pois o produto de matrizes possui a propriedade associativa. Logo, provamos que* $k(h(A)) = h(k(A))$.

Exemplo 2.7.6 *Considere a matriz* A *de ordem* 3×2 *e a matriz* B *de ordem* 2×3,

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 4 & 6 \end{bmatrix} \qquad e \qquad B = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 4 & 8 \end{bmatrix},$$

e a operação elementar de linhas $h : l_2 \longleftarrow l_2 - 2l_1$. *Podemos verificar facilmente que*

$$h(B)A = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 15 & 22 \\ 12 & 18 \end{bmatrix} = h(BA)$$

onde as matrizes BA *e* $h(BA)$ *são dadas por:*

$$BA = \begin{bmatrix} 15 & 22 \\ 42 & 62 \end{bmatrix} \quad l_2 \longleftarrow l_2 - 2l_1 \quad h(BA) = \begin{bmatrix} 15 & 22 \\ 12 & 18 \end{bmatrix}.$$

Assim, vemos que $h(B)A = h(BA)$, *que é uma ilustração do Lema 2.7.1.*

Exemplo 2.7.7 *Considere a matriz* A *de ordem* 3×2

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 4 & 6 \end{bmatrix}$$

e a seguinte seqüência de operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - 4l_1 \qquad e \qquad l_3 \longleftarrow l_3 - 2l_2$$

com as correspondentes matrizes elementares E_1, E_2 *e* E_3, *todas de ordem* 3,

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \quad e \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}.$$

Desse modo, obtemos a matriz $B = E_3 E_2 E_1 A$, *que está na forma escalonada, que corresponde a aplicação da seqüência de operações elementares de linhas, definida acima, na matriz* A. *De fato,*

$$B = E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Portanto, temos que as matrizes A *e* B *são equivalentes,* $A \sim B$.

Teorema 2.7.3 *Sejam A e B matrizes de ordem $m \times n$. Então, a matriz B é linha equivalente a matriz A se, e somente se, $B = PA$, com $P = H_r \cdots H_2 H_1$, onde cada matriz H_i é uma matriz elementar de linha de ordem $m \times m$.*

Demonstração

($\Longrightarrow$) Considerando que B é linha equivalente a matriz A. Sejam $h_1, \cdots h_r$ uma seqüência de operações elementares com as linhas de A resultando na matriz B. Desse modo, tomando H_i a matriz elementar de linha correspondente à operação elementar de linhas h_i, temos que $B = (H_r \cdots H_2 H_1)A = PA$.

($\Longleftarrow$) Considerando que $B = PA$, com $P = H_r \cdots H_2 H_1$, onde cada matriz H_i é uma matriz elementar de linha de ordem $m \times m$. Temos que a matriz $H_1 A$ é linha equivalente a matriz A e $H_2 H_1 A$ é linha equivalente a matriz $H_1 A$. Assim, a matriz $H_2 H_1 A$ é linha equivalente a matriz A. Continuando o processo, vemos que a matriz $(H_r \cdots H_2 H_1)A = PA$ é linha equivalente a matriz A. ■

Teorema 2.7.4 *Uma matriz elementar de linha H é invertível e sua inversa é uma matriz elementar de linha H_1 que corresponde à operação elementar inversa da operação elementar de linhas efetuada por H.*

Demonstração – Seja H a matriz elementar de linha correspondente à operação elementar de linhas h. Se h_1 é a operação inversa de h e $H_1 = h_1(I)$, então

$$\begin{aligned} H H_1 &= h(H_1) = h(h_1(I)) = I \\ H_1 H &= h_1(H) = h_1(h(I)) = I \end{aligned}$$

Desse modo, temos que H é uma matriz invertível e $H_1 = H^{-1}$. Logo, da definição de inversa de uma matriz, temos que $H = H_1^{-1}$. ■

Teorema 2.7.5 *Uma matriz elementar de coluna K é invertível e sua inversa é uma matriz elementar de coluna K_1 que corresponde à operação elementar inversa da operação elementar de colunas efetuada por K.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 2.7.8 *Vamos considerar o seguinte exemplo de uma matriz elementar de linha obtida da matriz identidade* I_3

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad l_2 \longleftarrow l_2 + 2l_1 \quad E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Assim, temos que a operação elementar inversa é $l_2 \longleftarrow l_2 - 2l_1$ *e a inversa da matriz elementar* E_1 *é dada por:*

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad l_2 \longleftarrow l_2 - 2l_1 \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Portanto, temos que $E_1E_2 = E_2E_1 = I_3$*. Logo,* $E_2 = E_1^{-1}$ *e* $E_1 = E_2^{-1}$*, decorrente da definição de inversa de uma matriz.*

Exemplo 2.7.9 *Vamos considerar o seguinte exemplo de uma matriz de permutação de linhas obtida da matriz identidade* I_3

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad l_2 \longleftrightarrow l_3 \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Sabemos que a matriz de permutação P *é uma matriz elementar e podemos observar que* $P^{-1} = P$*, isto é,* $PP = P^2 = I_3$*. Logo, a matriz de permutação* P *é idempotente.*

Exemplo 2.7.10 *No Exemplo 2.7.7 temos que* $B = E_3E_2E_1A$*. Logo, como as matrizes elementares são invertíveis, obtemos que* $A = E_1^{-1}\,E_2^{-1}\,E_3^{-1}\,B$*.*

Assim, denotando $E = E_3E_2E_1$*, temos que* $E^{-1} = E_1^{-1}\,E_2^{-1}\,E_3^{-1}$*. Portanto, obtemos*

$$A = E_1^{-1}\,E_2^{-1}\,E_3^{-1}\,B = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 4 & 6 \end{bmatrix}.$$

Exemplo 2.7.11 *Considerando o Exemplo 2.7.10, calcule explicitamente as matrizes*

$$E = E_3E_2E_1 \qquad e \qquad E^{-1} = E_1^{-1}\,E_2^{-1}\,E_3^{-1}.$$

Teorema 2.7.6 *Sejam* $H_1, H_2, \cdots, H_r$ *matrizes elementares de linha e*

$$H = H_r \cdots H_2 H_1 .$$

Então, $H^{-1} = H_1^{-1} H_2^{-1} \cdots H_r^{-1}$.

Demonstração – Pelo Teorema 2.7.4, temos que cada matriz elementar de linha H_i é invertíveis. Assim, a prova segue da aplicação do Teorema 2.3.1. □

Teorema 2.7.7 *Sejam* $K_1, K_2, \cdots, K_r$ *matrizes elementares de coluna e*

$$K = K_r \cdots K_2 K_1 .$$

Então, $K^{-1} = K_1^{-1} K_2^{-1} \cdots K_r^{-1}$.

Demonstração – Pelo Teorema 2.7.5, temos que cada matriz elementar de coluna H_i é invertíveis. Assim, a prova segue da aplicação do Teorema 2.3.1. □

Exemplo 2.7.12 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -2 \\ 3 & 6 & -4 \end{bmatrix}$$

Encontre uma matriz R *na forma escalonada, linha equivalente a matriz* A*, indicando a seqüência de matrizes elementares de linha utilizada.*

Exemplo 2.7.13 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Determine uma seqüência de matrizes elementares $H_1, \cdots, H_r$*, onde cada matriz* H_i *é uma matriz elementar de linha correspondente à operação elementar de linhas* h_i*, de modo que* $H_r H_{r-1} \cdots H_2 H_1 A = I_3$*. Mostre que* $H_r H_{r-1} \cdots H_2 H_1 = A^{-1}$.

Exemplo 2.7.14 *Mostre que necessariamente uma matriz elementar de linha de ordem* 2×2 *é uma das seguintes matrizes:*

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} \quad , \quad \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \quad , \quad \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix} \quad e \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}$$

para o escalar $c \neq 0$.

Teorema 2.7.8 *Seja A uma matriz de ordem $n \times n$. Então, as seguintes afirmações são equivalentes:*

(a) A é invertível.

(b) A é linha equivalente a matriz identidade.

(c) A é um produto de matrizes elementares de linha.

Demonstração

Vamos mostrar que *(a)* $\Longrightarrow$ *(b)*. Considerando que A é invertível e linha equivalente a uma matriz B na forma escada. Sejam $H_1, H_2, \cdots, H_r$ matrizes elementares de linha, onde cada matriz H_i corresponde a uma operação elementar de linhas h_i, tais que

$$B = H_r \cdots H_2 H_1 A .$$

Como A é invertível e cada matriz elementar de linha H_i também é invertível, temos que B é invertível. Entretanto, se $B \neq I_n$, então B possui uma linha nula, o que é uma contradição, pois B é invertível. Logo, $B = I_n$.

Vamos mostrar que *(b)* $\Longrightarrow$ *(c)*. Considerando que A é linha equivalente a matriz I_n. Sejam $h_1, \cdots h_r$ uma seqüência de operações elementares com as linhas de A resultando na matriz I_n. Desse modo, tomando cada matriz H_i como sendo a matriz elementar de linha correspondente à operação elementar de linhas h_i, temos que

$$I_n = H_r \cdots H_2 H_1 A .$$

Pelo Teorema 2.7.6, temos que o produto de matrizes elementares de linha é invertível. Assim, temos que

$$A = H_1^{-1} H_2^{-1} \cdots H_r^{-1} .$$

Sabemos que H_i^{-1} também é uma matriz elementar de linha. Portanto, mostramos que A é um produto de matrizes elementares de linha.

Finalmente, mostraremos que *(c)* $\Longrightarrow$ *(a)*. Sejam $H_1, H_2, \cdots, H_r$ matrizes elementares de linha, onde cada matriz H_i corresponde a uma operação elementar de linhas h_i, tais que $A = H_r \cdots H_2 H_1$. Pelo Teorema 2.7.6, temos que o produto de matrizes elementares de linha é invertível. Desse modo, obtemos

$$A^{-1} = H_1^{-1} H_2^{-1} \cdots H_r^{-1} ,$$

provando que A é invertível, o que completa a demonstração. ■

Exemplo 2.7.15 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Vamos determinar a matriz A^{-1}, *se possível, através da aplicação de uma seqüência de operações elementares de linhas* $h_1, h_2, \cdots, h_r$ *na matriz* A *de modo que* $A \sim I_3$. *Pelo Teorema 2.7.8, sabemos que aplicando essa mesma seqüência de operações elementares de linhas na matriz identidade* I_3 *obtemos a matriz* A^{-1}.

Inicialmente, aplicamos as operações elementares de linhas

$$h_1 : l_2 \longleftarrow l_2 - l_1 \quad , \quad h_2 : l_3 \longleftarrow l_3 - l_1 \qquad \text{e} \qquad h_3 : l_2 \longleftrightarrow l_3$$

simultaneamente na matriz A, obtendo uma matriz R na forma escalonada, e na matriz identidade I_3, obtendo

$$R = H_3 H_2 H_1 A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix} \qquad \text{e} \qquad H_3 H_2 H_1 I_3 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}.$$

Agora, aplicamos as operações elementares de linhas

$$h_4 : l_2 \longleftarrow l_2 - l_3 \quad , \quad h_5 : l_1 \longleftarrow l_1 - 2l_3 \qquad \text{e} \qquad h_6 : l_1 \longleftrightarrow l_1 + l_2$$

simultaneamente na matriz R e na matriz $H_3 H_2 H_1 I_3$, obtendo

$$H_6 H_5 H_4 R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \text{e} \qquad H_6 H_5 H_4 H_3 H_2 H_1 I_3 = \begin{bmatrix} -1 & 1 & 1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{bmatrix}.$$

Finalmente, aplicamos as operações elementares de linhas

$$h_7 : l_2 \longleftarrow -l_2 \qquad \text{e} \qquad h_8 : l_3 \longleftarrow -l_3$$

nas matrizes acima, obtendo $I_3 = H_8 H_7 H_6 H_5 H_4 H_3 H_2 H_1 A$ e

$$A^{-1} = H_8 H_7 H_6 H_5 H_4 H_3 H_2 H_1 I_3 = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix},$$

o que completa a resolução do exemplo.

Exemplo 2.7.16 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 7 & 5 \\ 3 & 8 & 7 \end{bmatrix}.$$

Vamos determinar a matriz A^{-1}*, se possível, através da aplicação de uma seqüência de operações elementares de linhas* $h_1, h_2, \cdots, h_r$ *na matriz* A *de modo que* $A \sim I_3$*. Pelo Teorema 2.7.8, sabemos que aplicando essa mesma seqüência de operações elementares de linhas na matriz identidade* I_3 *obtemos a matriz* A^{-1}*.*

Para facilitar a aplicação da seqüência de operações elementares de linhas, vamos cria uma **matriz ampliada** M da seguinte forma:

$$M = \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 3 & 7 & 5 & 0 & 1 & 0 \\ 3 & 8 & 7 & 0 & 0 & 1 \end{array}\right].$$

Inicialmente, na primeira parte da matriz M temos a matriz A e na segunda parte da matriz M temos a matriz identidade I_3.

Agora, aplicando as operações elementares de linhas

$$h_1 : l_2 \longleftarrow l_2 - 3l_1 \quad , \quad h_2 : l_3 \longleftarrow l_3 - 3l_1 \qquad \text{e} \qquad h_3 : l_3 \longleftarrow l_3 - 2l_2$$

na matriz ampliada M, obtemos

$$H_3\,H_2\,H_1\,M = \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 0 & 3 & -2 & 1 \end{array}\right].$$

Temos que a matriz $R = H_3\,H_2\,H_1\,A$ na forma escalonada, linha equivalente a matriz A, possui uma linha nula. Assim, pelo Teorema 2.7.8, podemos concluir que a matriz A não possui inversa, pois não poderá ser reduzida linha equivalente à matriz identidade I_3.

Teorema 2.7.9 *Sejam* A *e* B *matrizes de ordem* $m \times n$. *Mostre que a matriz* B *é equivalente por coluna a matriz* A *se, e somente se,* $B = AQ$, *com* $Q = K_1 K_2 \cdots K_r$, *onde cada matriz* K_i *é uma matriz elementar de coluna de ordem* $n \times n$.

Demonstração – A prova é feita de modo análogo ao Teorema 2.7.3. $\square$

Corolário 2.7.2 *Sejam* A *e* B *matrizes de ordem* $m \times n$. *Então, a matriz* B *é equivalente a matriz* A, *que indicamos por* $B \sim A$, *se, e somente se, existe uma matriz invertível* P *de ordem* $m \times m$ *e uma matriz invertível* Q *de ordem* $n \times n$, *tais que* $B = PAQ$.

Demonstração – A prova segue imediata da Definição 2.5.5, do Teorema 2.7.3 e do Teorema 2.7.9. $\square$

Exemplo 2.7.17 *Vamos mostrar que a equivalência de matrizes, que indicamos por* $\sim$, *é uma relação de equivalência sobre o conjunto das matrizes de ordem* $m \times n$, *isto é, satisfaz as seguintes propriedades:*

(a) ***Propriedade Reflexiva:*** $A \sim A$.

(b) ***Propriedade Simétrica:*** *Se* $A \sim B$, *então* $B \sim A$.

(c) ***Propriedade Transitiva:*** *Se* $A \sim B$ *e* $B \sim C$, *então* $A \sim C$.

para A, B *e* C *matrizes de ordem* $m \times n$.

Podemos verificar facilmente que $A \sim A$. De fato, pois $A = I_m A I_n$, onde I_n é a matriz identidade de ordem $n \times n$, e I_m é a matriz identidade de ordem $m \times m$. Assim, mostramos que a equivalência de matrizes satisfaz a propriedade reflexiva.

Considerando que a matriz A é equivalente a matriz B, $A \sim B$, isto é, existe uma matriz invertível P de ordem $m \times m$ e uma matriz invertível Q de ordem $n \times n$, tais que $A = PBQ$.

Assim, temos que $B = P^{-1} A Q^{-1}$. Logo, tomando as matrizes $P_1 = P^{-1}$ e $Q_1 = Q^{-1}$, obtemos $B = P_1 A Q_1$. Desse modo, mostramos que a matriz B é equivalente a matriz A, $B \sim A$. Portanto, mostramos que a equivalência de matrizes satisfaz a propriedade simétrica.

Considerando que a matriz A é equivalente a matriz B, $A \sim B$, isto é, existe uma matriz invertível P_1 de ordem $m \times m$ e uma matriz invertível Q_1 de ordem $n \times n$, tais que $A = P_1 B Q_1$, e que a matriz B é equivalente a matriz C, $B \sim C$, isto é, existe uma matriz invertível P_2 de ordem $m \times m$ e uma matriz invertível Q_2 de ordem $n \times n$, tais que $B = P_2 C Q_2$. Desse modo, temos que

$$A = P_1 B Q_1 = P_1 (P_2 C Q_2) Q_1 = (P_1 P_2) C (Q_2 Q_1)$$

Sabemos que a matriz $P_1 P_2$ é invertível, pois as matrizes P_1 e P_2 são invertíveis, e que a matriz $Q_2 Q_1$ é invertível, pois as matrizes Q_1 e Q_2 são invertíveis. Desse modo, mostramos que a matriz A é equivalente a matriz C, $A \sim C$. Assim, mostramos que a equivalência de matrizes satisfaz a propriedade transitiva.

Portanto, mostramos que a equivalência de matrizes é uma relação de equivalência sobre o conjunto das matrizes de ordem $m \times n$.

Exercícios

Exercício 2.82 *Determine as matrizes elementares de linha* $H_1\,,\,H_2\,,\,H_3$ *e* $H_4\,,$ *de ordem* 3×3, *que correspondem às operações elementares de linhas*

$$h_1 : l_1 \longleftrightarrow l_3 \quad , \quad h_2 : l_2 \longleftarrow l_2 + l_1 \quad , \quad h_3 : l_3 \longleftarrow l_3 - 2l_1 \quad e \quad h_4 : l_3 \longleftarrow l_3 - l_2$$

e encontre a inversa da matriz $H = H_4\,H_3\,H_2\,H_1$.

Exercício 2.83 *Escreva a matriz*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$$

como um produto de matrizes elementares de linha, se possível.

Exercício 2.84 *Escreva a matriz*

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$$

como um produto de matrizes elementares de linha e determine sua inversa, se possível.

Exercício 2.85 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 2 & 1 \\ 1 & 4 & -7 \end{bmatrix}.$$

Determine a matriz A^{-1}, *se possível, através da aplicação de uma seqüência de operações elementares de linhas* $h_1,\, h_2,\, \cdots,\, h_r$ *na matriz* A *de modo que* $A \sim I_3$.

Exercício 2.86 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Determine a matriz A^{-1}, *se possível, através da aplicação de uma seqüência de operações elementares de linhas* $h_1,\, h_2,\, \cdots,\, h_r$ *na matriz* A *de modo que* $A \sim I_3$.

Exercício 2.87 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 1 & 3 & 6 \\ 1 & 4 & 3 & 1 \end{bmatrix}.$$

Determine uma matriz R na forma escalonada que seja linha equivalente a matriz A, e uma matriz P invertível de ordem 3×3 tal que $R = PA$.

Exercício 2.88 *Dada a matriz*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 6 \end{bmatrix}.$$

Determine uma matriz R na forma escalonada que seja linha equivalente a matriz A e uma matriz P invertível de ordem 3×3 tal que $R = PA$.

Exercício 2.89 *Considere a seguinte matriz simétrica*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 2 & 4 \\ 1 & 3 & 4 \end{bmatrix}.$$

Determine uma matriz L triangular inferior que seja equivalente por coluna a matriz A, indicando a seqüência de operações elementares de colunas utilizada e a respectiva seqüência de matrizes elementares de coluna, isto é, $L = AK_1 K_2 \cdots K_r$.

Exercício 2.90 *Considere a matriz L triangular inferior obtida no Exercício 2.89. Determine a matriz D linha equivalente a matriz L através da seqüência de operações elementares de linhas correspondente à seqüência de operações elementares de colunas utilizada para obter a matriz L, isto é, $D = H_r \cdots H_2 H_1 L$ onde $H_i = (K_i)^t$.*

Exercício 2.91 *Determine o posto linha da matriz A dada por:*

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 & 0 \\ 3 & 4 & 4 & 5 & 1 \\ 1 & 0 & 2 & 1 & 1 \end{bmatrix},$$

e também o posto linha da matriz A^t.

2.8 Matrizes Congruentes. Lei da Inércia

Definição 2.8.1 *Sejam* A *e* B *matrizes reais de ordem* $n \times n$. *Dizemos que a matriz* B *é* ***congruente*** *com a matriz* A *se existe uma matriz real invertível* P *de ordem* $n \times n$ *tal que* $B = PAP^t$.

Podemos verificar facilmente que se a matriz B é congruente a uma matriz simétrica A, então B é simétrica. De fato,

$$B^t = (PAP^t)^t = PA^t P^t = PAP^t = B \,,$$

e assim mostramos que B é uma matriz simétrica.

Como a congruência é uma relação de equivalência, Exemplo 2.8.3, temos que somente as matrizes simétricas podem ser mutuamente congruentes e, em particular, somente as matrizes simétricas são congruentes a matrizes diagonais.

Teorema 2.8.1 (Lei da Inércia) *Seja* A *uma matriz simétrica real. Então, existe uma matriz real invertível* P *tal que* $D = PAP^t$ *é uma matriz diagonal. Além disso, o número de elementos na diagonal de* D *que são positivos, negativos e nulos é sempre o mesmo, independente da matriz* P *que realiza a relação de congruência.*

Na seção 6.7 vamos apresentar a demonstração da *Lei de Inércia de Sylvester*, que é uma generalização do Teorema 2.8.1, utilizando os resultados de diagonalização de matrizes Hermitianas, tornando a prova mais simples e elegante.

A seguir, apresentamos um procedimento para determinar da matriz P que realiza a diagonalização da matriz simétrica A através da relação de congruência.

Sejam $h_1, \cdots, h_r$ a seqüência de operações elementares de linhas, sendo H_i a matriz elementar de linha correspondente à operação elementar de linhas h_i, tais que

$$R = H_r \cdots H_2 \, H_1 \, A = H A$$

é a matriz na forma escalonada linha equivalente a matriz A, que é uma matriz triangular superior, onde estamos indicando a matriz $H = H_r \cdots H_2 \, H_1$.

Sejam $k_1, \cdots, k_r$ a seqüência de operações elementares de colunas correspondente à seqüência de operações elementares de linhas $h_1, \cdots, h_r$. Indicamos por K_i a matriz elementar de coluna correspondente à operação elementar de colunas k_i. Sabemos que

$$K_i = (H_i)^t ,$$

pelo Corolário 2.7.1.

Aplicando a seqüência de operações elementares de colunas $K_1, \cdots, K_r$ na matriz $R = HA$, obtemos a matriz diagonal

$$D = H_r \cdots H_2 H_1 A K_1 \cdots K_r = HAH^t$$

Desse modo, a matriz P, que é triangular inferior, é dada por:

$$P = H_r \cdots H_2 H_1 = H$$

realiza a diagonalização da matriz simétrica A através da relação de congruência, isto é,

$$D = PAP^t$$

é uma matriz diagonal.

Portanto, a matriz P é invertível, pois é o produto de matrizes elementares de linhas. Assim, a matriz $L = P^{-1}$, que é triangular inferior, é dada por:

$$L = (H_1)^{-1} (H_2)^{-1} \cdots (H_r)^{-1} .$$

Desse modo, temos a decomposição da matriz simétrica A na forma:

$$A = LDL^t ,$$

que é bastante utilizada em várias aplicações. Essa forma de decomposição de matrizes simétricas será estudada na seção 8.5, onde apresentamos a Decomposição de Cholesky.

Exemplo 2.8.1 *Dada a matriz simétrica*

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 6 \\ 1 & 6 & 3 \end{bmatrix} .$$

Vamos determinar uma matriz invertível P *de modo que*

$$D = PAP^t$$

seja uma matriz diagonal.

Para facilitar a aplicação da seqüência de operações elementares, vamos cria uma **matriz ampliada** M da seguinte forma:

$$M = \begin{bmatrix} 1 & 2 & 1 & | & 1 & 0 & 0 \\ 2 & 6 & 6 & | & 0 & 1 & 0 \\ 1 & 6 & 3 & | & 0 & 0 & 1 \end{bmatrix} .$$

Inicialmente, na primeira parte da matriz M temos a matriz A e na segunda parte da matriz M temos a matriz identidade I_3.

Agora, aplicando as operações elementares de linhas

$$h_1 : l_2 \longleftarrow l_2 - 2l_1 \quad , \quad h_2 : l_3 \longleftarrow l_3 - l_1 \qquad \text{e} \qquad h_3 : l_3 \longleftarrow l_3 - 2l_2$$

na matriz ampliada M, obtemos

$$H_3 \, H_2 \, H_1 \, M = \begin{bmatrix} 1 & 2 & 1 & | & 1 & 0 & 0 \\ 0 & 2 & 4 & | & -2 & 1 & 0 \\ 0 & 0 & -6 & | & 3 & -2 & 1 \end{bmatrix} .$$

Desse modo, temos que as matrizes $P = H_3 \, H_2 \, H_1$ e PA são dadas por:

$$PA = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & -6 \end{bmatrix} \qquad \text{e} \qquad P = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix} .$$

Finalmente, aplicando as correspondentes operações elementares de colunas na matriz PA, isto é, $PA \, K_1 \, K_2 \, K_3$, onde $K_i = (H_i)^t$, obtemos a matriz diagonal

$$D = PAP^t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -6 \end{bmatrix} ,$$

que é a diagonalização da matriz A através da transformação de congruência. Sabemos que a matriz P é invertível, pois é um produto de matrizes elementares de linhas.

Podemos observar facilmente que

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad H_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}.$$

Desse modo, chamando a matriz triangular inferior $L = P^{-1}$, temos que

$$L = P^{-1} = (H_1)^{-1} (H_2)^{-1} (H_3)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix},$$

onde

$$(H_1)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad (H_2)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad (H_3)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}.$$

Portanto, temos que a matriz simétrica A pode ser decomposta da seguinte forma:

$$A = P^{-1} D P^{-t} = L D L^t ,$$

que também é uma forma bastante usual de decomposição de uma matriz simétrica, que possui várias aplicações interessantes.

Exemplo 2.8.2 *Considere a matriz simétrica A dada por:*

$$A = \begin{bmatrix} 4 & 2 & 12 & 2 \\ 2 & 28 & 6 & 1 \\ 12 & 6 & 72 & 6 \\ 2 & 1 & 6 & 3 \end{bmatrix}.$$

Determine uma matriz invertível P de modo que $D = PAP^t$ seja uma matriz diagonal, e a matriz $L = P^{-1}$ tal que $A = LDL^t$.

Exemplo 2.8.3 *Vamos mostrar que a relação de congruência, que indicaremos por* $\approx$*, é uma relação de equivalência sobre o conjunto das matrizes de ordem* $n \times n$*, isto é, satisfaz as seguintes propriedades:*

(a) ***Reflexiva:*** $A \approx A$.

(b) ***Simétrica:*** *Se* $B \approx A$*, então* $A \approx B$.

(c) ***Transitiva:*** *Se* $A \approx B$ *e* $B \approx C$*, então* $A \approx C$.

para A*,* B *e* C *matrizes de ordem* $n \times n$*.*

Podemos verificar facilmente que $A \approx A$. De fato, pois $A = IAI^t$, onde I é a matriz identidade de ordem $n \times n$. Assim, mostramos que a relação de congruência satisfaz a propriedade reflexiva.

Considerando que a matriz B é congruente com a matriz A, $B \approx A$, isto é, existe uma matriz invertível P tal que $B = PAP^t$. Sendo assim, temos que $A = P^{-1}BP^{-t}$. Portanto, tomando a matriz $Q = P^{-1}$, obtemos $A = QBQ^t$. Desse modo, mostramos que a matriz A é congruente com a matriz B, $A \approx B$. Assim, mostramos que a relação de congruência satisfaz a propriedade simétrica.

Considerando que a matriz A é congruente com a matriz B, $A \approx B$, isto é, existe uma matriz invertível P tal que $A = PBP^t$, e que a matriz B é congruente com a matriz C, $B \approx C$, isto é, existe uma matriz invertível Q tal que $B = QCQ^t$. Desse modo, temos que

$$A \;=\; PBP^t \;=\; P(\,QCQ^t\,)P^t \;=\; (PQ)C(PQ)^t\,.$$

Sabemos que a matriz PQ é invertível, pois P e Q são invertíveis. Desse modo, mostramos que a matriz A é congruente com a matriz C, $A \approx C$. Assim, mostramos que a relação de congruência satisfaz a propriedade transitiva.

Portanto, mostramos que a relação de congruência é uma relação de equivalência sobre o conjunto das matrizes de ordem $n \times n$.

Exercícios

Exercício 2.92 *Considere a matriz simétrica*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 8 \\ 3 & 8 & 15 \end{bmatrix}.$$

Determine uma matriz invertível P *de modo que* $D = PAP^t$ *seja uma matriz diagonal, e a matriz* $L = P^{-1}$ *tal que* $A = LDL^t$.

Exercício 2.93 *Considere a matriz simétrica*

$$A = \begin{bmatrix} 3 & 9 & 6 \\ 9 & 29 & 22 \\ 6 & 22 & 20 \end{bmatrix}.$$

Determine uma matriz invertível P *de modo que* $D = PAP^t$ *seja uma matriz diagonal, e a matriz* $L = P^{-1}$ *tal que* $A = LDL^t$.

2.9 Sistemas de Equações Lineares

Seja $A = [a_{ij}]$ uma matriz de ordem $m \times n$ definida sobre o corpo $\mathbb{F}$, isto é, seus elementos $a_{i,j} \in \mathbb{F}$ para $1 \leq i \leq m$ e $1 \leq j \leq n$. Consideremos o problema de encontrar escalares $x_1, \cdots, x_n \in \mathbb{F}$ satisfazendo simultaneamente o seguinte sistema de equações lineares

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{i1}x_1 & + & a_{i2}x_2 & + & \cdots & + & a_{in}x_n & = & y_i \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & y_m \end{array} \right.$$

conhecendo os escalares $y_1, \cdots, y_n \in \mathbb{F}$. Esse problema é denominado **sistema linear**, com m equações lineares e n incógnitas.

Por simplicidade, vamos representar o sistema linear acima na sua **forma matricial**

$$AX = Y$$

onde

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad , \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{bmatrix}$$

A matriz A de ordem $m \times n$ é denominada **matriz dos coeficientes** do sistema linear, o vetor coluna X de ordem $n \times 1$ é denominado **vetor de incógnitas** e o vetor coluna Y de ordem $m \times 1$ é denominado **vetor do lado direito** do sistema linear.

Toda n–upla $(x_1, \cdots, x_n)$ de elementos do corpo $\mathbb{F}$ que satisfaz cada uma das equações do sistema linear é denominada uma **solução** do sistema linear. O vetor coluna X, associado a essa n–upla, é denominado **vetor solução** do sistema linear. O conjunto de todas as soluções do sistema linear é chamado **conjunto solução**.

Quanto $y_1 = y_2 = \cdots = y_m = 0$ dizemos que o sistema linear é **homogêneo**, isto é, temos que cada equação do sistema linear é uma **equação homogênea**.

Teorema 2.9.1 *Considere a equação linear*

$$ax = b\,,$$

com $a, b \in \mathbb{R}$.

(a) Se $a \neq 0$, *então* $x = \dfrac{b}{a}$ *é a única solução da equação linear.*

(b) Se $a = 0$ *e* $b \neq 0$, *então a equação linear não possui solução.*

(c) Se $a = 0$ *e* $b = 0$, *então a equação linear possui infinitas solução.*

Demonstração – A prova pode ficar a cargo do leitor. $\square$

Definição 2.9.1 *Dizemos que a equação linear*

$$a_1x_1 + \cdots + a_nx_n = b\,,$$

nas incógnitas $x_1, \cdots, x_n$, *é* ***degenerada*** *se* $a_1 = a_2 = \cdots = a_n = 0$.

Teorema 2.9.2 *Considere a equação linear degenerada*

$$a_1x_1 + \cdots + a_nx_n = b\,,$$

$a_1 = a_2 = \cdots = a_n = 0$.

(a) Se $b \neq 0$, *então a equação linear degenerada não possui solução.*

(b) Se $b = 0$, *então a equação linear degenerada possui infinitas solução.*

Demonstração – A prova pode ficar a cargo do leitor. $\square$

Definição 2.9.2 *Dizemos que a equação linear*

$$a_1x_1 + \cdots + a_nx_n = b$$

é ***não–degenerada*** *se os coeficientes* $a_1, a_2, \cdots, a_n$ *não são todos nulos. Além disso, a primeira incógnita com coeficiente não–nulo é denominada* ***variável básica****. De modo análogo,* x_k *é a variável básica, se* $a_j = 0$ *para todo* $j < k$, *mas* $a_k \neq 0$.

Teorema 2.9.3 *Considere a equação linear não–degenerada*

$$a_1 x_1 + \cdots + a_n x_n = b \,,$$

com x_k *a variável básica.*

(a) Qualquer conjunto de valores x_j*, para* $j \neq k$*, fornece uma única solução para a equação linear. As incógnitas* x_j*, para* $j \neq k$*, são denominadas* ***variáveis livres****.*

(b) Toda solução da equação linear não–degenerada é obtida em (a).

Demonstração – Inicialmente atribuímos valores arbitrários às variáveis livres x_j, para $j \neq k$. Como $a_j = 0$, para $j < k$, obtemos

$$a_k x_k = b - (a_{k+1} x_{k+1} + \cdots + a_n x_n) \,,$$

com $a_k \neq 0$. Pelo Teorema 2.9.1, a incógnita x_k é determinada de modo único na forma:

$$x_k = \frac{b - (a_{k+1} x_{k+1} + \cdots + a_n x_n)}{a_k} \,,$$

o que prova o item *(a)*.

Finalmente, vamos supor que a n–upla $(x_1, \cdots, x_n)$ seja uma solução da equação linear. Desse modo, temos que

$$x_k = \frac{b - (a_{k+1} x_{k+1} + \cdots + a_n x_n)}{a_k} \,,$$

que é exatamente a solução obtida no item *(a)*, o que completa a demonstração. ■

Exemplo 2.9.1 *O conjunto solução da equação linear*

$$2x + 6y - 4z = 10 \,,$$

nas incógnitas x, y, *e* z, *é dado por:*

$$S = \{ \, (x, y, z) \in I\!R^3 \; / \; x = 5 - 3y + 2z \quad , \quad y, z \in I\!R \, \} \,,$$

onde x *é a variável básica, com* y *e* z *as variáveis livres.*

Exemplo 2.9.2 *Considere o sistema de equações lineares não–degeneradas dado por:*

$$\begin{cases} a_1 x & + & b_1 y & = & c_1 \\ \\ a_2 x & + & b_2 y & = & c_2 \end{cases}$$

nas incógnitas x e y.

Podemos observar facilmente que cada uma das equações do sistema linear representa a equação na forma canônica de uma reta contida no plano numérico $I\!R^2$. Assim, podemos dar uma interpretação geométrica ao conjunto solução do sistema linear.

Podemos descrever três situações geométricas para o conjunto solução, a saber:

(1) O gráfico das equações lineares são retas que se interceptam em um único ponto, isto é, são retas concorrentes. Assim, O sistema linear possui somente uma única solução.

(2) O gráfico das equações lineares são retas paralelas distintas. Assim, O sistema linear não possui solução.

(3) O gráfico das equações lineares são retas paralelas coincidentes. Assim, O sistema linear possui infinitas soluções.

A seguir vamos analisar separadamente cada um dos casos acima. As situações (2) e (3) ocorrerem quando as retas possuem coeficientes angulares iguais, isto é,

$$-\frac{a_1}{b_1} = -\frac{a_2}{b_2} \qquad \Longleftrightarrow \qquad \frac{a_1}{a_2} = \frac{b_1}{b_2} \qquad \Longleftrightarrow \qquad a_1\,b_2 \,-\, a_2\,b_1 \,=\, 0\,.$$

Note que a condição acima pode ser escrita da seguinte forma:

$$\det(A) \;=\; \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \;=\; 0\,,$$

onde A é a matriz do sistema linear. Desse modo, os casos (2) e (3) ocorrerem somente quando a matriz do sistema linear não possui inversa.

Assim, o caso (1) ocorre quando as retas possuem coeficientes angulares diferentes, isto é,

$$\det(A) \;=\; \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \;=\; a_1\,b_2 \,-\, a_2\,b_1 \;\neq\; 0\,.$$

Portanto, o caso (1) ocorre quando a matriz do sistema linear for invertível.

O caso (2) ocorre quando as retas são paralelas e cortam o eixo vertical OY em pontos distintos, isto é,

$$-\frac{c_1}{b_1} \neq -\frac{c_2}{b_2} \quad \Longleftrightarrow \quad \frac{b_1}{b_2} \neq \frac{c_1}{c_2} \quad \Longleftrightarrow \quad \frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2}\,.$$

O caso (3) ocorre quando as retas são paralelas e cortam o eixo vertical OY no mesmo pontos, isto é,

$$-\frac{c_1}{b_1} = -\frac{c_2}{b_2} \quad \Longleftrightarrow \quad \frac{b_1}{b_2} = \frac{c_1}{c_2} \quad \Longleftrightarrow \quad \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}\,.$$

Assim, analisamos o conjunto solução do sistema linear tanto do ponto de vista geométrico quanto do ponto de vista algébrico.

Exemplo 2.9.3 *Analisar o conjunto solução do sistema linear*

$$\begin{cases} 2x + y = 1 \\ 3x + 2y = 4 \end{cases}$$

apresentando uma interpretação geométrica.

Exemplo 2.9.4 *Analisar o conjunto solução do sistema linear*

$$\begin{cases} 3x + 2y = 1 \\ 3x + 2y = 4 \end{cases}$$

apresentando uma interpretação geométrica.

Exemplo 2.9.5 *Analisar o conjunto solução do sistema linear*

$$\begin{cases} 10x + 5y = 15 \\ 2x + 1y = 3 \end{cases}$$

apresentando uma interpretação geométrica.

Exemplo 2.9.6 *Analisar o conjunto solução do sistema linear*

$$\begin{cases} 10x + 5y = 15 \\ 2x + 1y = -3 \end{cases}$$

apresentando uma interpretação geométrica.

Exemplo 2.9.7 *Uma equação linear nas incógnitas* x, y, z *é representada na forma:*

$$ax + by + cz = d$$

conhecendo os escalares $a, b, c, d \in I\!R$. *Por exemplo,*

$$x - 4y + 3z = 6$$

é uma equação linear nas incógnitas x, y, z.

Exemplo 2.9.8 *Uma equação linear nas incógnitas* x, y, z *representada na forma:*

$$ax + by + cz = 0 \, ,$$

conhecendo os escalares $a, b, c \in I\!R$, *é denominada* ***equação linear homogênea****.*

Por exemplo,

$$x - 4y + 3z = 0$$

é uma equação linear homogênea nas incógnitas x, y, z.

Podemos verificar facilmente que toda solução da equação linear homogênea, dada acima, pode ser escrita como:

$$(x, y, z) = y(4, 1, 0) + z(-3, 0, 1) \qquad \text{para} \qquad y, z \in I\!R \, .$$

denominada **solução geral**.

Desse modo, pode representar o conjunto solução da seguinte forma:

$$S = \{ \, (x, y, z) \; / \; (x, y, z) = y(4, 1, 0) + z(-3, 0, 1) \quad \text{para} \quad y, z \in I\!R \, \} \, .$$

Note que as ternas $(4, 1, 0)$ e $(-3, 0, 1)$ são soluções da equação homogênea, utilizadas na representação da solução geral. Essas soluções são chamadas **soluções básicas**.

Uma técnica bastante simples e muito importante na obtenção de soluções de um sistema de equações lineares é o **método de eliminação** que utiliza as operações elementares de linhas. Vamos exemplificar esse método através de um sistema linear homogêneo.

Exemplo 2.9.9 *Considere o seguinte sistema linear homogêneo*

$$\begin{cases} x - 2y + z = 0 \\ 2x - 5y + z = 0 \end{cases}$$

Aplicando a operação elementar de linhas $l_2 \longleftarrow l_2 - 2l_1$ obtemos o seguinte sistema linear homogêneo

$$\begin{cases} x - 2y + z = 0 \\ \quad - y - z = 0 \end{cases}$$

Assim, da segunda equação temos que $y = -z$ para $z \in \mathbb{R}$. Substituindo na primeira equação obtemos $x = -3z$.

Portanto, temos que toda solução do sistema linear homogêneo é escrita da seguinte forma:

$$(x, y, z) = z(-3, -1, 1) \qquad \text{para} \qquad z \in \mathbb{R}.$$

Note que $(-3, -1, 1)$ é uma solução do sistema linear homogêneo, que é utilizada na representação da solução geral. Assim, essa solução é chamada solução básica. Neste caso, dizemos que o sistema linear homogêneo possui um **grau de liberdade**, que é a **variável livre** z. As varáveis x, y são denominadas **variáveis básicas**.

Podemos verificar que o sistema linear homogêneo obtido através da operação elementar de linhas, possui o mesmo conjunto solução do sistema linear homogêneo original. Desse modo, dizemos que os dois sistemas lineares são **equivalentes**. Vamos apresentar uma análise detalhada do processo de eliminação mais a frente. Claramente, o processo de eliminação é válido para sistema linear não homogêneo, como exemplificamos a seguir.

Exemplo 2.9.10 *Considere o seguinte sistema linear*

$$\begin{cases} x - 2y + z - t = 1 \\ 2x - 3y + z + 2t = 3 \end{cases}$$

Aplicando a operação elementar de linhas $l_2 \longleftarrow l_2 - 2l_1$ obtemos o sistema linear

$$\begin{cases} x - 2y + z - t = 1 \\ \quad y - z + 4t = 1 \end{cases}$$

Assim, da segunda equação temos que $y = z - 4t + 1$ para $z, t \in \mathbb{R}$. Substituindo na primeira equação obtemos $x = z - 7t + 3$.

Portanto, temos que a solução geral do sistema linear é escrita da seguinte forma:

$$(x, y, z, t) \;=\; z(1,1,1,0) \;+\; t(-7,-4,0,1) \;+\; (3,1,0,0) \qquad \text{para} \qquad z,\, t \,\in\, I\!R\,.$$

Neste exemplo, temos duas variáveis livres, que são z e t, e dizemos que o grau de liberdade do sistema linear é igual a dois.

Podemos verificar facilmente que as ternas $(1,1,1,0)$ e $(-7,-4,0,1)$ são duas soluções do sistema linear homogêneo associado, isto é,

$$\begin{cases} x - 2y + z - t = 0 \\ 2x - 3y + z + 2t = 0 \end{cases} \iff \begin{cases} x - 2y + z - t = 0 \\ \qquad y - z + 4t = 0 \end{cases}$$

utilizadas na representação da solução geral do sistema linear. Assim, essas soluções são as soluções básicas do sistema linear homogêneo associado. Note que a terna $(3,1,0,0)$ é uma solução do sistema linear original, denominada **solução particular**.

Exemplo 2.9.11 *Considere o seguinte sistema linear*

$$\begin{cases} x - 2y + z - t = 1 \\ 2x - 3y + z + 2t = 3 \\ 3x - 9y + 6z - 15t = 5 \end{cases}$$

Aplicando as seguintes operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \qquad \text{e} \qquad l_3 \longleftarrow l_3 - 3l_1$$

obtemos o sistema linear

$$\begin{cases} x - 2y + z - t = 1 \\ \qquad y - z + 4t = 1 \\ \quad - 3y + 3z - 12t = 2 \end{cases}$$

Aplicando a operação elementar de linhas $l_3 \longleftarrow l_3 + 3l_2$ obtemos o sistema linear

$$\begin{cases} x - 2y + z - t = 1 \\ \qquad y - z + 4t = 1 \\ \qquad\qquad 0 = 5 \end{cases}$$

Desse modo, temos que a terceira equação é **degenerada**, isto é, pode ser escrita como:

$$0x \,+\, 0y \,+\, 0z \,+\, 0t \,=\, 5\,.$$

Logo, o sistema linear é **inconsistente**, isto é, não possui solução.

Exemplo 2.9.12 *Considere o seguinte sistema linear*

$$\begin{cases} x - 2y + z = -1 \\ 2x - 3y + z = -3 \\ x + 4y + 2z = 7 \end{cases}$$

Aplicando as seguintes operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad \text{e} \quad l_3 \longleftarrow l_3 - l_1$$

obtemos o sistema linear

$$\begin{cases} x - 2y + z = 1 \\ y - z = -1 \\ 6y + z = 8 \end{cases}$$

Aplicando a operação elementar de linhas $l_3 \longleftarrow l_3 - 6l_2$ obtemos o sistema linear

$$\begin{cases} x - 2y + z = 1 \\ y - z = 1 \\ 7z = 14 \end{cases}$$

Assim, temos que o sistema linear possui uma única solução

$$z = 2 \,, \quad y = 1 \quad \text{e} \quad x = -1 \,,$$

isto é, a terna $(-1, 1, 2)$ é a única solução do sistema linear.

Definição 2.9.3 *Dizemos que um sistema de equações lineares é* ***consistente*** *se possui solução. Quanto não possui solução, dizemos que é* ***inconsistente****.*

$$\underline{\textbf{Sistema de Equações Lineares}} \begin{cases} \textit{Inconsistente} \longrightarrow \text{não possui solução} \\ \\ \textit{Consistente} \begin{cases} \longrightarrow \text{possui uma única solução} \\ \longrightarrow \text{possui infinitas soluções} \end{cases} \end{cases}$$

Teorema 2.9.4 *Sejam A e B matrizes de ordem $m \times n$ que são equivalentes por linha. Então, os sistemas lineares homogêneos $AX = 0$ e $BX = 0$ possuem as mesmas soluções.*

Demonstração – Considerando que a matriz B é linha equivalente a matriz A, pelo Teorema 2.7.3, existe uma seqüência de operações elementares com as linhas da matriz A resultando na matriz B, que vamos denotar por $h_1, \cdots h_r$. Desse modo, tomando H_i a matriz elementar de linha correspondente à operação elementar de linhas h_i, temos que $B = (H_r \cdots H_2 H_1)A = PA$. Pelo Teorema 2.7.6, sabemos que a matriz P é invertível. Assim, temos que $A = P^{-1}B$. Logo, a matriz A é linha equivalente a matriz B.

Desse modo, se o vetor coluna X^*, de ordem $n \times 1$, é uma solução do sistema linear homogêneo $AX = 0$, isto é, $AX^* = 0$, temos que

$$BX^* = (PA)X^* = P(AX^*) = 0 .$$

Logo, X^* é também uma solução do sistema linear homogêneo $BX = 0$.

De modo análogo, se o vetor coluna X^*, de ordem $n \times 1$, é uma solução do sistema linear homogêneo $BX = 0$, isto é, $BX^* = 0$, temos que

$$AX^* = (P^{-1}B)X^* = P^{-1}(BX^*) = 0 .$$

Logo, X^* é também uma solução do sistema linear homogêneo $AX = 0$.

Portanto, provamos que os sistemas lineares homogêneos

$$AX = 0 \qquad \text{e} \qquad BX = 0$$

são equivalentes, isto é, possuem o mesmo conjunto solução. ■

Definição 2.9.4 *Seja A uma matriz de ordem $m \times n$. Dizemos que A é uma matriz* ***não–singular*** *se $AX = 0$ somente para $X = 0$. Caso contrário, dizemos que A é uma matriz* ***singular****.*

No resultado que apresentamos a seguir, provamos que se A é uma matriz de ordem $m \times n$, com $m < n$, então necessariamente A é uma matriz singular, isto é, o sistema linear homogêneo $AX = 0$ possui solução não trivial. Para uma matriz A de ordem n, vamos mostrar que A é uma matriz invertível se, e somente se, A é uma matriz não–singular. Esses são resultados importantes em Álgebra Linear, que iremos utilizar em todo o texto.

Teorema 2.9.5 *Seja* A *uma matriz de ordem* $m \times n$, *com* $m < n$. *Então, o sistema linear homogêneo* $AX = 0$ *admite pelo menos uma solução não trivial.*

Demonstração – Seja $R = [r_{ij}]$ uma matriz na forma escalonada linha equivalente a matriz A. Então, pelo Teorema 2.9.4, os sistemas lineares homogêneos $AX = 0$ e $RX = 0$ são equivalentes, isto é, possui o mesmo conjunto solução.

Considerando que p é o número de linhas não–nulas na matriz R, certamente temos que $p < m < n$. Desse modo, o sistema linear homogêneo tem grau de liberdade igual a $(n - p)$. Vamos considerar que o primeiro elemento não–nulo da i–ésima linha ocorra na coluna k_i, com $k_1 < k_2 < \cdots < k_i < \cdots < k_p$.

Assim, o sistema linear $RX = 0$ possui p equações não–triviais, que podem ser escritas da seguinte forma:

$$r_{ik_i} x_{k_i} + \sum_{j = k_{i+1}}^{n} r_{ij} x_j = 0 \qquad \text{para} \qquad i = 1, 2, \cdots p \, .$$

Como $r_{ik_i} \neq 0$, temos que as incógnitas x_{k_i} são obtidas da seguinte forma:

$$x_{k_i} = -\frac{\displaystyle\sum_{j = k_{i+1}}^{n} r_{ij} x_j}{r_{ik_i}} \qquad \text{para} \qquad i = 1, 2, \cdots p \, .$$

Finalmente, atribuindo valores, não todos nulos, para as $(n - p)$ **variáveis livres**, que são diferentes das **variáveis básicas** $x_{k_1}, \cdots, x_{k_p}$, obtemos o conjunto solução do sistema linear homogêneo $RX = 0$. ■

Exemplo 2.9.13 *No sistema linear homogêneo associado do Exemplo 2.9.10*

$$\begin{cases} x - 2y + z - t = 0 \\ 2x - 3y + z + 2t = 0 \end{cases} \iff \begin{cases} x - 2y + z - t = 0 \\ y - z + 4t = 0 \end{cases}$$

temos as seguintes matrizes de coeficientes

$$A = \begin{bmatrix} 1 & -2 & 1 & -1 \\ 2 & -3 & 1 & 2 \end{bmatrix} \qquad e \qquad R = \begin{bmatrix} 1 & -2 & 1 & -1 \\ 0 & 1 & -1 & 4 \end{bmatrix}$$

que são equivalentes por linha. Neste exemplo, temos $k_1 = 1$ *e* $k_2 = 2$, *onde* x, y *são as variáveis básicas e* z, t *são as variáveis livres. Desse modo, obtemos*

$$y = z - 4t \qquad e \qquad x = z - 7t$$

para $z, t \in I\!R$ *livres.*

Teorema 2.9.6 *Seja A uma matriz de ordem n. Então, o sistema linear homogêneo $AX = 0$ possui somente a solução trivial se, e somente se, a matriz A é linha equivalente a matriz identidade I_n.*

Demonstração – Seja R uma matriz na forma escada linha equivalente a matriz A e p o número de linhas não–nulas de R. Como $AX = 0$ possui somente a solução trivial, pelo Teorema 2.9.4, sabemos que o sistema homogêneo $RX = 0$ também possui somente a solução trivial. Assim, temos que $p \geq n$. Entretanto, a matriz R possui somente n linhas. Logo, como $p \leq n$, temos que $p = n$.

Desse modo, temos que na matriz R o primeiro elemento não–nulo da i–ésima linha ocorre na i–ésima coluna, isto é, sempre na diagonal principal. Portanto, a matriz R é necessariamente a matriz identidade I_n.

Podemos verificar facilmente que se a matriz A é linha equivalente a matriz identidade, então o sistema linear homogêneo $AX = 0$ possui somente a solução trivial. ■

Teorema 2.9.7 *Seja A uma matriz de ordem n. Então, a matriz A é invertível se, e somente se, o sistema linear homogêneo $AX = 0$ possui somente a solução trivial.*

Demonstração – A prova segue imediatamente do Teorema 2.7.8 e do Teorema 2.9.6, o que completa a demonstração. ■

Exemplo 2.9.14 *Considere o seguinte sistema linear homogêneo*

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

que possui somente a solução trivial $x = y = z = 0$.

Desse modo, temos que a matriz do sistema linear homogêneo $AX = 0$ que é dada por:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 2 & 4 & 5 \end{bmatrix}$$

é uma matriz não–singular, isto é, $AX = 0$ somente para $X = 0$. Portanto, temos que A é uma matriz invertível, de acordo com o Teorema 2.9.7.

Corolário 2.9.1 *Seja A uma matriz de ordem n. Então, a matriz A é invertível se, e somente se, o sistema linear $AX = Y$ possui somente uma única solução.*

Demonstração – Inicialmente, tomando por hipótese que A é invertível, temos que $X^* = A^{-1}Y$ é uma solução do sistema linear $AX = Y$.

Vamos mostrar que X^* é a única solução. Para isso, supomos que X_1^* e X_2^* sejam duas soluções do sistema linear $AX = Y$, isto é, $AX_1^* = Y$ e $AX_2^* = Y$.

Desse modo, temos que

$$AX_1^* - AX_2^* = A(X_1^* - X_2^*) = 0 .$$

Pelo Teorema 2.9.7, temos que $(X_1^* - X_2^*) = 0$. Logo, obtemos que $X_1^* = X_2^*$.

Finalmente, tomando por hipótese que o sistema linear $AX = Y$ possui somente uma única solução e considerando $Y = 0$, pelo Teorema 2.9.7, temos que a matriz A é invertível, o que completa a demonstração. ■

Vamos fazer algumas considerações. Sejam A uma matriz de ordem $m \times n$ e Y um vetor coluna de ordem $m \times 1$. Queremos determinar o conjunto solução do sistema linear

$$AX = Y .$$

Seja R uma matriz na forma escalonada linha equivalente a matriz A, isto é, existe uma matriz P de ordem $m \times m$ invertível tal que $R = PA$. Sabemos que a matriz P é o produto de matrizes elementares de linha, isto é, $P = H_r \cdots H_2 H_1$, onde cada matriz H_i é uma matriz elementar de linha associada a uma operação elementar de linhas h_i, que foram utilizadas na redução da matriz A na forma escalonada. Desse modo, sabemos que o sistema linear $AX = Y$ possui o mesmo conjunto solução do sistema linear $RX = Z$, onde $Z = PY$.

Portanto, um procedimento eficiente para a resolução do sistema linear $AX = Y$ é a aplicação de uma seqüência de operações elementares de linhas na matriz ampliada $[A \,|\, Y]$ para obtermos a matriz $[R \,|\, Z]$

$$[R \,|\, Z] = H_r \cdots H_2 H_1 [A \,|\, Y] ,$$

sem a necessidade do cálculo da matriz P. Esse procedimento, inclusive o cálculo da matriz P, foi bastante discutido na seção 2.7. Em particular, se A é uma matriz quadrada, esse procedimento também vai determinar se A é uma matriz invertível.

Exemplo 2.9.15 *Considere o sistema linear não–homogêneo*

$$\begin{cases} x - 2y + z - t = 1 \\ 2x - 3y + z + 2t = 4 \\ 3x - 9y + 6z - 15t = -3 \end{cases}$$

Vamos fazer uma análise do seu conjunto solução.

Para facilitar a aplicação da seqüência de operações elementares de linhas, vamos cria uma matriz ampliada $[A \,|\, Y]$ da seguinte forma:

$$[A \,|\, Y] = \begin{bmatrix} 1 & -2 & 1 & -1 & | & 1 \\ 2 & -3 & 1 & 2 & | & 4 \\ 3 & -9 & 6 & -15 & | & -3 \end{bmatrix}$$

onde A é a matriz dos coeficientes do sistema linear e Y é o vetor do lado direito.

Aplicando a seqüência de operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - 3l_1 \qquad \text{e} \qquad l_3 \longleftarrow l_3 + 3l_2$$

na matriz ampliada $[A \,|\, Y]$, obtemos a matriz $[R \,|\, Z]$

$$[R \,|\, Z] = \begin{bmatrix} 1 & -2 & 1 & -1 & | & 1 \\ 0 & 1 & -1 & 4 & | & 2 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}$$

Assim, obtemos o sistema linear equivalente $RX = Z$

$$\begin{cases} x - 2y + z - t = 1 \\ \quad\;\; y - z + 4t = 2 \end{cases}$$

Portanto, a solução geral do sistema linear não–homogêneo $AX = Y$ é escrita como:

$$(x, y, z, t) = z(1, 1, 1, 0) + t(-7, -4, 0, 1) + (5, 2, 0, 0) \qquad \text{para} \qquad z,\, t \in \mathbb{R}\,.$$

Logo, o sistema linear $AX = Y$ é consistente e possui infinitas soluções.

Podemos verificar facilmente que as ternas $(1, 1, 1, 0)$ e $(-7, -4, 0, 1)$ são as soluções básicas do sistema linear homogêneo associado $AX = 0$, que é equivalente ao sistema linear homogêneo $RX = 0$. Portanto, a solução geral do sistema linear homogêneo associado $AX = 0$ é escrita como:

$$(x, y, z, t) = z(1, 1, 1, 0) + t(-7, -4, 0, 1) \qquad \text{para} \qquad z,\, t \in \mathbb{R}\,.$$

A terna $(5, 2, 0, 0)$ é uma solução particular do sistema linear $AX = Y$.

Exemplo 2.9.16 *Considere o sistema linear não–homogêneo*

$$\begin{cases} x - 2y + z - t = 1 \\ 2x - 3y + z + 2t = 4 \\ 3x - 9y + 6z - 15t = 5 \end{cases}$$

Vamos fazer uma análise do seu conjunto solução.

Para facilitar a aplicação da seqüência de operações elementares de linhas, vamos cria uma matriz ampliada $[A\,|\,Y]$ da seguinte forma:

$$[A\,|\,Y] = \begin{bmatrix} 1 & -2 & 1 & -1 & | & 1 \\ 2 & -3 & 1 & 2 & | & 4 \\ 3 & -9 & 6 & -15 & | & 5 \end{bmatrix}$$

onde A é a matriz dos coeficientes do sistema linear e Y é o vetor do lado direito.

Aplicando a seqüência de operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - 3l_1 \qquad \text{e} \qquad l_3 \longleftarrow l_3 + 3l_2$$

na matriz ampliada $[A\,|\,Y]$, obtemos a matriz $[R\,|\,Z]$

$$[R\,|\,Z] = \begin{bmatrix} 1 & -2 & 1 & -1 & | & 1 \\ 0 & 1 & -1 & 4 & | & 2 \\ 0 & 0 & 0 & 0 & | & 8 \end{bmatrix}$$

Assim, obtemos o sistema linear equivalente $RX = Z$

$$\begin{cases} x - 2y + z - t = 1 \\ \qquad y - z + 4t = 2 \\ \qquad\qquad\qquad 0 = 8 \end{cases}$$

Assim, temos que a terceira equação é degenerada, isto é, pode ser escrita como:

$$0x + 0y + 0z + 0t = 8\,.$$

Logo, o sistema linear $AX = Y$ é inconsistente, isto é, não possui solução.

Exemplo 2.9.17 *Considere o sistema linear não–homogêneo*

$$\begin{cases} x - 2y + z = -1 \\ 2x - 3y + z = -3 \\ x + 4y + 2z = 7 \end{cases}$$

Vamos fazer uma análise do seu conjunto solução.

Para facilitar a aplicação da seqüência de operações elementares de linhas, vamos cria uma matriz ampliada $[A\,|\,Y]$ da seguinte forma:

$$[A\,|\,Y] = \begin{bmatrix} 1 & -2 & 1 & | & -1 \\ 2 & -3 & 1 & | & -3 \\ 1 & 4 & 2 & | & 7 \end{bmatrix}$$

onde A é a matriz dos coeficientes do sistema linear e Y é o vetor do lado direito.

Aplicando a seqüência de operações elementares de linhas

$$l_2 \longleftarrow l_2 - 2l_1 \quad , \quad l_3 \longleftarrow l_3 - l_1 \qquad \text{e} \qquad l_3 \longleftarrow l_3 - -l_2$$

na matriz ampliada $[A\,|\,Y]$, obtemos a matriz $[R\,|\,Z]$

$$[R\,|\,Z] = \begin{bmatrix} 1 & -2 & 1 & | & 1 \\ 0 & 1 & -1 & | & 1 \\ 0 & 0 & 7 & | & 14 \end{bmatrix}$$

Assim, obtemos o sistema linear equivalente $RX = Z$

$$\begin{cases} x - 2y + z = 1 \\ \qquad y - z = 1 \\ \qquad\quad 7z = 14 \end{cases}$$

Assim, o sistema linear não–homogêneo $AX = Y$ possui uma única solução

$$z = 2 \quad , \quad y = 1 \qquad \text{e} \qquad x = -1.$$

Portanto, o sistema linear $AX = Y$ é consistente e a terna $(-1, 1, 2)$ é a única solução.

Observando os exemplos apresentados nessa seção, podemos fazer a seguinte afirmação:

$$\boxed{\text{Solução Geral do Sistema Linear} \quad \mathrm{AX} = \mathrm{Y}}$$

$$=$$

$$\boxed{\text{Solução Geral do Sistema Linear Homogêneo Associado} \quad \mathrm{AX} = 0}$$

$$+$$

$$\boxed{\text{Solução Particular do Sistema Linear} \quad \mathrm{AX} = \mathrm{Y}}$$

De fato, consideramos que X_h é uma solução do sistema linear homogêneo associado, isto é, $AX_h = 0$, e que X_p é uma solução particular do sistema linear original, isto é, $AX_p = Y$. Desse modo, temos que

$$A(X_h + X_p) = AX_h + AX_p = 0 + Y = Y .$$

Assim, mostramos que $X_h + X_p$ é uma solução do sistema linear $AX = Y$.

Por outro lado, considerando que X^* é uma solução do sistema linear $AX = Y$, que pode ser distinta da solução particular X_p, temos que

$$A(X^* - X_p) = AX^* - AX_p = Y - Y = 0 .$$

Desse modo, temos que $X^* - X_p$ é uma solução do sistema homogêneo $AX = 0$.

Entretanto, podemos escrever X^* da seguinte forma:

$$X^* = X_p + (X^* - X_p) .$$

Portanto, qualquer solução do sistema linear $AX = Y$ pode ser escrita como a soma de uma solução particular do sistema linear original com uma solução do sistema linear homogêneo associado.

Portanto, a solução geral do sistema linear $AX = Y$ pode ser escrita da seguinte forma:

$$X_g = X_h + X_p ,$$

o que prova a nossa afirmação.

Teorema 2.9.8 *Sejam A uma matriz de ordem $m \times n$, e $X_1, \cdots, X_n$ soluções do sistema linear homogêneo $AX = 0$. Então, toda* ***combinação linear***

$$X_c = \alpha_1 X_1 + \cdots + \alpha_1 X_n ,$$

onde $\alpha_1, \cdots, \alpha_n$ são escalares, é também solução do sistema linear $AX = 0$.

Demonstração – Sabemos que

$$AX_1 = 0, \ \cdots, \ AX_n = 0 .$$

Logo,

$$AX_c = \alpha_1 AX_1 + \cdots + \alpha_1 AX_n = 0 ,$$

o que completa a demonstração. ■

Teorema 2.9.9 *Sejam A uma matriz de ordem $m \times n$ e Y um vetor coluna de ordem $m \times 1$. Então, o sistema linear $AX = Y$ não possui solução, possui uma única solução, ou possui infinitas soluções.*

Demonstração – Basta mostrar que se o sistema linear $AX = Y$ possui mais de uma solução, então possui infinitas soluções.

Sejam X_1 e X_2 duas soluções distintas de $AX = Y$, isto é, $AX_1 = Y$ e $AX_2 = Y$. Desse modo, para todo escalar α temos que

$$X^* = X_1 + \alpha(X_1 - X_2)$$

é também uma solução do sistema linear $AX = Y$. De fato,

$$AX^* = AX_1 + \alpha(AX_1 - AX_2) = Y + \alpha(Y - Y) = Y .$$

Finalmente, devemos observar que para cada escalar α os vetores colunas

$$X^* = X_1 + \alpha(X_1 - X_2)$$

são distintos, o que completa a demonstração. ■

Teorema 2.9.10 *Sejam* A *uma matriz de ordem* $m \times n$ *e* Y *um vetor coluna de ordem* $m \times 1$. *Então,*

(a) O sistema linear $AX = Y$ *é consistente se, e somente se,*

$$posto(\,[A\,|\,Y]\,) \;=\; posto(A)\,.$$

(b) O sistema linear $AX = Y$ *possui uma única solução se, e somente se,*

$$posto(\,[A\,|\,Y]\,) \;=\; posto(A) \;=\; n\,.$$

(c) O sistema linear $AX = Y$ *possui infinitas solução se, e somente se,*

$$posto(\,[A\,|\,Y]\,) \;=\; posto(A) \;<\; n\,.$$

(d) O sistema linear $AX = Y$ *é inconsistente se, e somente se,*

$$posto(A) \;<\; posto(\,[A\,|\,Y]\,)\,.$$

Demonstração – Seja $[R\,|\,Z]$ uma matriz na forma escalonada, linha equivalente a matriz ampliada $[A\,|\,Y]$, isto é, existe uma matriz P invertível de ordem $m \times m$ tal que $R = PA$ e $Z = PY$. Logo, os sistemas lineares $AX = Y$ e $RX = Z$ possuem o mesmo conjunto solução.

(a) O $posto(\,[A\,|\,Y]\,) = posto(A)$, isto é,

$$posto(\,[R\,|\,Z]\,) \;=\; posto(R)$$

se, e somente se, o sistema linear reduzido $RX = Z$ não possui equações degeneradas. Desse modo, não existem condições sobre as componentes de Z para que o sistema linear reduzido $RX = Z$ tenha solução. Portanto, $AX = Y$ é um sistema linear consistente, isto é, possui solução.

(b) O $posto(\,[A\,|\,Y]\,) = posto(A) = n$, isto é,

$$posto(\,[R\,|\,Z]\,) \;=\; posto(R) \;=\; n$$

se, e somente se, o sistema linear reduzido $RX = Z$ não possui variáveis livres. Desse modo, cada uma das variáveis assume um valor fixo, que são obtidos resolvendo o sistema linear reduzido $RX = Z$ pelo processo de substituição atrasada. Portanto, o sistema linear $AX = Y$ possui uma única solução.

Exemplificando a situação descrita acima, podemos representar a matriz reduzida R, linha equivalente a matriz A, e o vetor coluna Z por:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & r_{nn} \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \qquad \text{e} \qquad Z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

onde os elementos da diagonal principal da matriz R são todos não–nulos.

(c) O $posto([A \,|\, Y]) = posto(A) < n$, isto é,

$$posto([R \,|\, Z]) = posto(R) < n$$

se, e somente se, o sistema linear $RX = Z$ possui pelo menos uma variável livre. Desse modo, para cada conjunto de valores atribuídos às variáveis livres, obtemos uma solução. Como cada variável livre pode assumir qualquer valor, o sistema linear reduzido $RX = Z$ possui infinitas soluções. Portanto, o sistema linear $AX = Y$ possui infinitas soluções.

(d) O $posto(A) < posto([A \,|\, Y])$, isto é,

$$posto(R) < posto([R \,|\, Z])$$

se, e somente se, o sistema linear reduzido $RX = Z$ possui pelo menos uma equação degenerada, com a correspondente componente de Z não–nula. Desse modo, o sistema linear reduzido $RX = Z$ não possui solução. Portanto, $AX = Y$ é um sistema linear inconsistente. ■

Sistemas Lineares em Forma Triangular

Definição 2.9.5 *Sejam* L *uma matriz triangular inferior de ordem* $n \times n$ *e* Y *um vetor coluna de ordem* $n \times 1$*. Dizemos que*

$$LX = Y$$

é um sistema linear ***triangular inferior****.*

Exemplo 2.9.18 *O sistema linear dado por:*

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 1 & 5 & 0 & 0 \\ 2 & 1 & 4 & 0 \\ 1 & 2 & 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \\ 8 \\ 26 \end{bmatrix}$$

é um sistema triangular inferior.

Teorema 2.9.11 *Sejam* L *uma matriz triangular inferior de ordem* $n \times n$*, com todos os elementos da diagonal principal não–nulos, e* Y *um vetor coluna de ordem* $n \times 1$*. Então, o sistema linear* ***triangular inferior***

$$LX = Y$$

possui uma única solução, que é obtida pelo processo de ***substituição avançada****.*

Demonstração – Resolvendo a primeira equação, obtemos o valor da primeira incógnita

$$l_{11}\, x_1 = y_1 \qquad \Longleftrightarrow \qquad x_1 = \frac{y_1}{l_{11}} \, .$$

Substituindo o valor de x_1 na segunda equação, obtemos o valor da segunda incógnita

$$x_2 = \frac{y_2 - l_{21}\, x_1}{l_{22}} \, .$$

Desse modo, sucessivamente determinamos o valor de x_k, levando os valores das incógnitas previamente obtidos na k–ésima equação, da forma:

$$l_{kk}\, x_k = y_k - \sum_{j=1}^{k-1} l_{kj}\, x_j \qquad \Longleftrightarrow \qquad x_k = \frac{y_k - \displaystyle\sum_{j=1}^{k-1} l_{kj}\, x_j}{l_{kk}}$$

para $k = 2, 3, \cdots, n$.

Note que a unicidade de x_k, para $k = 1, \cdots, n$, é dada pelo Teorema 2.9.1, o que completa a demonstração. ■

Definição 2.9.6 *Sejam* U *uma matriz triangular superior de ordem* $n \times n$ *e* Y *um vetor coluna de ordem* $n \times 1$. *Dizemos que*

$$UX = Y$$

é um sistema linear ***triangular superior***.

Exemplo 2.9.19 *O sistema linear dado por:*

$$\begin{bmatrix} 6 & 2 & 3 & 1 \\ 0 & 4 & 1 & 2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 16 \\ 15 \\ 8 \\ 12 \end{bmatrix}$$

é um sistema triangular superior.

Teorema 2.9.12 *Sejam* U *uma matriz triangular superior de ordem* $n \times n$, *com todos os elementos da diagonal principal não–nulos, e* Y *um vetor coluna de ordem* $n \times 1$. *Então, o sistema linear* ***triangular superior***

$$UX = Y$$

possui uma única solução, que é obtida pelo processo de ***substituição atrasada***.

Demonstração – Resolvendo a última equação, obtemos o valor da última incógnita

$$u_{nn}\, x_n = y_n \qquad \Longleftrightarrow \qquad x_n = \frac{y_n}{u_{nn}} \,.$$

Substituindo o valor de x_n na penúltima equação, obtemos o valor da penúltima incógnita

$$x_{n-1} = \frac{y_{n-1} - u_{n-1,n}\, x_n}{u_{n-1,n-1}} \,.$$

Desse modo, sucessivamente determinamos o valor de x_k, levando os valores das incógnitas previamente obtidos na k–ésima equação, da forma:

$$u_{kk}\, x_k = y_k - \sum_{j=k+1}^{n} u_{kj}\, x_j \qquad \Longleftrightarrow \qquad x_k = \frac{y_k - \displaystyle\sum_{j=k+1}^{n} u_{kj}\, x_j}{u_{kk}}$$

para $k = (n-1), \cdots, 1$.

Note que a unicidade de x_k, para $k = n, \cdots, 1$, é dada pelo Teorema 2.9.1, o que completa a demonstração. ■

A seguir apresentamos o algoritmo para o obter a solução de um sistema linear triangular inferior $LX = Y$, pelo processo de substituição avançada.

Algoritmo 2.9.1 (Processo de Substituição Avançada)

```
for i = 1, ... ,n
    soma = 0.0
    for j = 1, ... ,(i-1)
        soma = soma  +  L(i,j)*X(j)
    end
    X(i) = ( Y(i) - soma ) / L(i,i)
end
```

A seguir apresentamos o algoritmo para o obter a solução de um sistema linear triangular superior $UX = Y$, pelo processo de substituição atrasada.

Algoritmo 2.9.2 (Processo de Substituição Atrasada)

```
for i = n, ... ,1
    soma = 0.0
    for j = (i+1), ... ,n
        soma = soma + U(i,j)*X(j)
    end
    X(i) = ( Y(i) - soma ) / U(i,i)
end
```

Exemplo 2.9.20 *Determine a solução do sistema linear triangular inferior*

$$\begin{cases} 4x & & & & = & 4 \\ x & + & 5y & & & = & 11 \\ 2x & + & y & + & 4z & & = & 8 \\ x & + & 2y & + & 3z & + & 6t & = & 26 \end{cases}$$

pelo processo de substituição avançada.

Exemplo 2.9.21 *Determine a solução do sistema linear triangular superior*

$$\begin{cases} 6x + 2y + 3z + t = 16 \\ 4y + z + 2t = 15 \\ 5z + t = 8 \\ 4t = 12 \end{cases}$$

pelo processo de substituição atrasada.

Fatoração LU

Sejam A uma matriz de ordem $n \times n$ invertível e Y um vetor coluna de ordem $n \times 1$. Vamos supor que a matriz A possa ser decomposta na forma:

$$A = LU ,$$

onde L é uma matriz triangular inferior, com todos os elementos da diagonal principal iguais a 1, e U uma matriz triangular superior com todos os elementos da diagonal principal diferentes de zero. Assim, dizemos que a matriz A possui uma **fatoração LU**, ou uma **decomposição LU**.

A fatoração LU da matriz A pode ser utilizada para obter, de uma maneira eficiente, a solução do sistema linear

$$AX = Y ,$$

repetidamente para diferentes vetores Y.

Substituindo a fatoração $A = LU$, obtemos

$$AX = Y \qquad \Longleftrightarrow \qquad (LU)X = Y \qquad \Longleftrightarrow \qquad L(UX) = Y .$$

Chamando $UX = Z$, e substituindo no sistema linear acima, obtemos o seguinte sistema linear triangular inferior

$$LZ = Y ,$$

cuja solução Z^* pode ser obtida facilmente pelo processo de substituição avançada.

Finalmente, obtemos a solução do sistema linear triangular superior

$$UX = Z^*$$

pelo processo de substituição atrasada. Assim, determinamos a solução X^* do sistema linear $AX = Y$.

Portanto, caso a matriz A possua uma decomposição LU, podemos obter a solução do sistema linear $AX = Y$ através da resolução de dois sistemas triangulares, isto é,

$$AX = Y \qquad \Longleftrightarrow \qquad \begin{cases} LZ = Y \\ UX = Z^* \end{cases}$$

Exemplo 2.9.22 *Determine a solução do seguinte sistema linear*

$$\begin{cases} 4x_1 + x_2 + 2x_3 = 2 \\ 8x_1 + 4x_2 + 5x_3 = 6 \\ 12x_1 + 7x_2 + 10x_3 = 6 \end{cases}$$

através da fatoração LU da matriz de coeficientes do sistema.

A matriz de coeficientes do sistema linear é dada por:

$$A = \begin{bmatrix} 4 & 1 & 2 \\ 8 & 4 & 5 \\ 12 & 7 & 10 \end{bmatrix}$$

e possui uma fatoração $A = LU$, onde

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Primeiramente, resolvemos o sistema linear triangular inferior

$$\begin{cases} z_1 = 2 \\ 2z_1 + z_2 = 6 \\ 3z_1 + 2z_2 + z_3 = 6 \end{cases}$$

pelo processo de substituição avançada. Assim, obtemos a solução

$$Z = \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix}.$$

Finalmente, uma vez determinado Z, resolvemos o sistema linear triangular superior

$$\begin{cases} 4x_1 + x_2 + 2x_3 = 2 \\ 2x_2 + x_3 = 2 \\ 2x_3 = -4 \end{cases}$$

pelo processo de substituição atrasada. Assim, obtemos a solução

$$X = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix},$$

o que completa a resolução do sistema linear $AX = Y$.

Seja A uma matriz de ordem $n \times n$ invertível que pode ser reduzida à forma escalonada, através de operações elementares de linhas, sem qualquer permutação de linhas, isto é, a cada passo do processo de redução à forma escalonada o pivô é sempre não–nulo.

Vamos descrever o processo de redução da matriz A à forma escalonada, que também é conhecido como processo de **Eliminação Gaussiana**, através do seguinte algoritmo.

Algoritmo 2.9.3 (Fatoração LU)

$$
\begin{array}{llcl}
para & j & = & 1, \cdots, (n-1) \\
para & i & = & (j+1), \cdots, n \\
 & m_{ij} & = & \dfrac{a_{ij}}{a_{jj}} \\
 & A_{i\cdot} & \longleftarrow & A_{i\cdot} \, - \, m_{ij} \, A_{j\cdot} \\
 & a_{ij} & = & m_{ij} \\
fim & & & \\
fim & & &
\end{array}
$$

Os escalares m_{ij} são denominados **multiplicadores**. Por simplicidade, indicamos por $A_{i\cdot}$ para denotar a i–ésima linha da matriz A.

No final do procedimento, teremos a matriz U armazenada na parte triangular superior da matriz A, e a matriz L armazenada abaixo da diagonal principal da matriz A, sabendo que os elementos da diagonal principal da matriz L são todos iguais à 1. Desse modo, a matriz triangular inferior L é dada por:

$$
L = \begin{bmatrix}
1 & & & & & \\
m_{21} & 1 & & & & \\
m_{31} & m_{32} & 1 & & & \\
m_{41} & m_{42} & m_{43} & 1 & & \\
\vdots & \vdots & \vdots & & \ddots & \\
m_{n1} & m_{n2} & m_{n3} & \cdots & m_{n,n-1} & 1
\end{bmatrix} .
$$

Teorema 2.9.13 *Seja* A *uma matriz de ordem* $n \times n$ *invertível que pode ser reduzida à forma escalonada, através de operações elementares de linhas, sem qualquer permutação de linhas. Então, existe uma matriz triangular inferior* L*, com todos os elementos da diagonal principal iguais a* 1*, e uma matriz triangular superior* U*, com todos os elementos da diagonal principal não–nulos, tais que* $A = LU$*.*

Demonstração – Sejam $h_1, \cdots, h_r$ a seqüência de operações elementares de linhas, sendo H_i a matriz elementar de linha correspondente à operação elementar de linhas h_i, tais que

$$U = H_r \cdots H_2 H_1 A$$

é a matriz na forma escalonada linha equivalente a matriz A, que é uma matriz triangular superior. Sendo assim, temos que

$$A = (H_r \cdots H_2 H_1)^{-1} U = (H_1^{-1} H_2^{-1} \cdots H_r^{-1}) U = LU ,$$

onde a matriz triangular inferior L é dada por:

$$L = H_1^{-1} H_2^{-1} \cdots H_r^{-1} ,$$

com os elementos da diagonal principal todos iguais a 1.

Finalmente, os elementos da diagonal principal da matriz U são todos não–nulos. De fato, caso contrário existiria um vetor coluna X não–nulo, de ordem $n \times 1$, tal que

$$UX = 0 \iff AX = LUX = 0 ,$$

o que é uma contradição, pois A é invertível, o que completa a demonstração. ∎

De uma maneira geral, podemos resumir os resultados apresentados da seguinte forma.

Seja A uma matriz de ordem n. As seguintes afirmações são equivalentes:

(a) A é uma matriz não–singular.

(b) A é uma matriz invertível.

(c) A matriz triangular superior U na forma escalonada, linha equivalente a matriz A, possui todos os elementos da diagonal principal não–nulos.

(d) $posto(A) = n$.

(e) O sistema linear homogêneo $AX = 0$ possui somente a solução trivial.

(f) Para todo vetor coluna Y, de ordem $n \times 1$, o sistema linear $AX = Y$ possui uma única solução, que é dada por $X^* = A^{-1}Y$.

Exemplo 2.9.23 *Considere a matriz A de ordem 4×4 dada por:*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}.$$

Determine a fatoração $A = LU$, e a matriz inversa A^{-1}.

Exemplo 2.9.24 *Considere a matriz* A *de ordem* 3×3 *dada por:*

$$A = \begin{bmatrix} 4 & 1 & 2 \\ 8 & 4 & 5 \\ 12 & 7 & 10 \end{bmatrix}.$$

Vamos determinar a fatoração $A = LU$ *sem permutações de linhas, se possível.*

Considere a seguinte seqüência de operações elementares de linhas

$$h_1 : l_2 \longleftarrow l_2 - 2l_1 \quad , \quad h_2 : l_3 \longleftarrow l_3 - 3l_1 \qquad \text{e} \qquad h_3 : l_3 \longleftarrow l_3 - 2l_2$$

com as correspondentes matrizes elementares H_1, H_2 e H_3, todas de ordem 3,

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \quad \text{e} \quad H_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}.$$

Desse modo, obtemos a matriz $U = H_3 H_2 H_1 A$, que está na forma escalonada, que corresponde a aplicação da seqüência de operações elementares de linhas, definida acima, na matriz A. De fato,

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 8 & 4 & 5 \\ 12 & 7 & 10 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Portanto, a matriz triangular inferior $L = H_1^{-1} H_2^{-1} H_3^{-1}$ é dada por:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$$

Assim, obtemos a fatoração $A = LU$, sem qualquer permutação de linhas.

Exercícios

Exercício 2.94 *Determine a fatoração* $A = LU$*, onde a matriz* A *é dada por:*

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 6 & 9 & 4 \\ 6 & 9 & 8 \end{bmatrix},$$

e determine a solução do sistema linear $AX = Y$*, onde*

$$Y = \begin{bmatrix} 0 \\ 0 \\ 24 \end{bmatrix}.$$

Exercício 2.95 *Determine a solução do sistema linear*

$$\begin{cases} 2x_1 & + & x_2 & - & x_3 & = & 1 \\ 6x_1 & + & 7x_2 & - & 2x_3 & = & 3 \\ 8x_1 & + & 12x_2 & + & 0x_3 & = & -4 \end{cases}$$

através da fatoração LU *da matriz de coeficientes do sistema.*

Exercício 2.96 *Descreva o conjunto solução da equação linear*

$$2x + 3y - 6z = 8.$$

Exercício 2.97 *Descreva o conjunto solução do sistema liner*

$$\begin{cases} 2x & + & 3y & - & 6z & = & 8 \\ x & + & y & + & z & = & 1 \end{cases}$$

Qual é o lugar geométrico em $\mathbb{R}^3$ *definido pelo conjunto solução?*

Exercício 2.98 *Considere a matriz* A *dada por:*

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 3 & 2 & -1 \\ 1 & 6 & -7 \end{bmatrix}.$$

Determine todos os vetores colunas

$$X = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

tais que $AX = 0$*.*

Exercício 2.99 *Considere a matriz* A *dada por:*

$$A = \begin{bmatrix} 1 & 3 \\ 4 & -3 \end{bmatrix}.$$

Determine todos os vetores colunas

$$X = \begin{bmatrix} a \\ b \end{bmatrix}$$

tais que $AX = -5X$.

Exercício 2.100 *Seja* A *uma matriz simétrica decomposta na forma* $A = LDL^t$, *onde*

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \qquad e \qquad D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Determine a solução do sistema linear $AX = Y$, *onde*

$$Y = \begin{bmatrix} 10 \\ 29 \\ 32 \end{bmatrix}.$$

Exercício 2.101 *Considere o seguinte sistema linear*

$$\begin{cases} 2x_1 + x_2 - x_3 = 1 \\ 6x_1 + 7x_2 - 2x_3 = 3 \\ 4x_1 + 6x_2 - x_3 = -4 \end{cases}$$

Faça uma análise do conjunto solução.

Exercício 2.102 *Considere o seguinte sistema linear*

$$\begin{cases} 2x_1 + 5x_2 + 2x_3 = -5 \\ 6x_1 + 7x_2 - 2x_3 = 3 \\ 8x_1 + 12x_2 + 0x_3 = -4 \end{cases}$$

Faça uma análise do conjunto solução.

Exercício 2.103 *Considere a matriz* A *dada por:*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 4 \end{bmatrix}.$$

Determine a inversa da matriz A *utilizando a fatoração* $A = LU$, *de maneira eficiente.*

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] Luenberger, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \quad = \quad Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q \quad = \quad \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/∼pulino/ALESA/

Janeiro de 2012

Conteúdo

3

Espaços Vetoriais

Conteúdo

3.1 Espaço Vetorial. Propriedades

Em vários ramos da matemática, defrontamo-nos com conjuntos, nos quais são, ao mesmo tempo significativos e interessante trabalhar com *combinações lineares* de seus elementos. Por exemplo, no estudo de equações lineares, é bastante natural considerar combinações lineares das linhas ou colunas de uma matriz. Em cálculo diferencial trabalhamos com combinações lineares de funções, por exemplo, no estudo de equações diferenciais. De um modo geral, a primeira experiência com vetores é apresentada com o estudo dos espaços euclidianos bidimensional e tridimensional.

Em geral, a Álgebra Linear é o ramo da matemática que trata das propriedades comuns a sistemas algébricos constituídos por um conjunto no qual a noção de combinação linear de seus elementos possa ser definida. Nesta seção vamos definir o objeto matemático que, como a experiência mostrou, é a abstração mais útil e interessante deste tipo de sistema algébrico.

Definição 3.1.1 *Um* ***Espaço Vetorial*** *consiste do seguinte:*

(1) Um conjunto não vazio V *de objetos, denominados vetores.*

(2) Um corpo $\mathbb{F}$ $(\mathbb{R}$ *ou* $\mathbb{C})$ *de escalares.*

(3) uma ***operação de adição de vetores****, que associa a cada par de elementos* $u, v \in V$ *um elemento* $u + v \in V$*, isto é,* V *é* ***fechado*** *com relação à operação de adição. Esta operação tem as seguintes propriedades:*

(A_1) Comutatividade. $u + v = v + u \;;\; \forall\ u, v \in V$.

(A_2) Associatividade. $u + (v + w) = (u + v) + w \;;\; \forall\ u, v, w \in V$.

(A_3) Elemento Neutro. Existe um elemento $0_V \in V$ *tal que* $u + 0_V = u \;;\; \forall\, u \in V$.

(A_4) Elemento Simétrico. Para todo elemento $u \in V$ *existe o elemento* $-u \in V$ *tal que* $u + (-u) = 0_V \;;\; \forall\ u \in V$.

(4) uma ***operação de multiplicação por escalar****, que associa a cada elemento* $u \in V$ *e cada escalar* $\alpha \in \mathbb{F}$ *um elemento* $\alpha\, u \in V$*, isto é,* V *é* ***fechado*** *com relação à operação de multiplicação por escalar. Esta operação tem as seguintes propriedades:*

(M_1) Associatividade. $(\alpha\, \beta)\, u = \alpha\, (\beta\, u) \;;\; \forall\ u \in V$ *e* $\forall\ \alpha, \beta \in \mathbb{F}$.

(M_2) Distributividade para a Adição de Elementos.
$\alpha\, (u + v) = \alpha\, u + \alpha\, v \;;\; \forall\ u, v \in V$ *e* $\forall\ \alpha \in \mathbb{F}$.

(M_3) *Distributividade para a Multiplicação por Escalar.*
$(\alpha + \beta)\, u = \alpha\, u + \beta\, u \ ; \ \forall \ u \in V \quad e \quad \forall \ \alpha, \beta \in \mathbb{F}.$

(M_4) *Elemento Identidade.* $1_{\mathbb{F}}\, u = u \ ; \ \forall \ u \in V.$

Quando consideramos o corpo dos escalares como sendo $\mathbb{F} = \mathbb{R}$, dizemos que $(V, +, \cdot)$ é um espaço vetorial real. Quando consideramos o corpo dos escalares como sendo $\mathbb{F} = \mathbb{C}$, dizemos que $(V, +, \cdot)$ é um espaço vetorial complexo. Por simplicidade, em todo texto podemos pensar $\mathbb{F} = \mathbb{R}$ ou $\mathbb{F} = \mathbb{C}$, a menos nos casos em que o corpo é explicitado. De uma maneira mais geral, podemos considerar que $\mathbb{F}$ é um corpo de característica zero, veja Definição 1.8.1.

Exemplo 3.1.1 *O conjunto do números reais, $\mathbb{R}$, com as operações usuais de adição e multiplicação de números reais, é um espaço vetorial real.*

Exemplo 3.1.2 *O conjunto do números complexos, $\mathbb{C}$, com as operações usuais de adição e multiplicação de números complexos, é um espaço vetorial complexo, considerando o corpo dos escalares como sendo $\mathbb{F} = \mathbb{C}$. Entretanto, podemos considerar o corpo de escalares como sendo $\mathbb{F} = \mathbb{R}$. Desse modo, temos que $\mathbb{C}$ é um espaço vetorial real.*

Exemplo 3.1.3 *O conjunto $\mathbb{R}^n = \{\ u = (x_1, \cdots, x_n) \ / \ x_i \in \mathbb{R}\ \}$, conjunto de todas as n–uplas reais, com a operação de adição de elementos definida por:*

$$u + v = (x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$$

e a operação de multiplicação por escalar definida por:

$$\lambda u = (\lambda x_1, \cdots, \lambda x_n)$$

é um espaço vetorial real.

Para mostrar que a operação de adição de elementos e a operação de multiplicação por escalar definidas em $\mathbb{R}^n$ verificam os axiomas da definição de espaço vetorial, basta utilizar as propriedades da operação de adição e da operação de multiplicação de elementos do corpo $\mathbb{R}$.

Exemplo 3.1.4 *O conjunto* $\mathbb{C}^n = \{ (z_1, \cdots, z_n) \ / \ z_i \in \mathbb{C} \}$, *conjunto de todas as* n*–uplas complexas, com as operações usuais, é um espaço vetorial complexo, considerando o corpo dos escalares como sendo* $\mathbb{F} = \mathbb{C}$.

Para mostrar que a operação de adição de elementos e a operação de multiplicação por escalar definidas em $\mathbb{C}^n$ verificam os axiomas da definição de espaço vetorial, basta utilizar as propriedades da operação de adição e da operação de multiplicação de elementos do corpo $\mathbb{C}$. Entretanto, podemos considerar o corpo dos escalares como sendo $\mathbb{F} = \mathbb{R}$. desse modo, temos que $\mathbb{C}^n$ é um espaço vetorial real.

Exemplo 3.1.5 *O conjunto* $\mathcal{F}(\mathbb{R}) = \{ f : \mathbb{R} \longrightarrow \mathbb{R} \ / \ f$ *é uma função* $\}$, *com a operação de adição de elementos definida como:*

$$(f + g)(x) = f(x) + g(x) \quad ; \quad \forall\, f, g \in \mathcal{F}(\mathbb{R})$$

e a operação de multiplicação por escalar definida como:

$$(\lambda f)(x) = \lambda f(x) \quad ; \quad \forall\, f \in \mathcal{F}(\mathbb{R}) \ \ e \ \ \lambda \in \mathbb{R}$$

é um espaço vetorial real.

Exemplo 3.1.6 *O conjunto* $\mathcal{C}([a,b]) = \{ f : [a,b] \longrightarrow \mathbb{R} \ / \ f$ *é uma função contínua* $\}$, *com a operação de adição de elementos e como a operação de multiplicação por escalar definidas em* $\mathcal{F}(\mathbb{R})$, *é um espaço vetorial real.*

Exemplo 3.1.7 *Seja* $n \geq 0$ *um número natural. O conjunto dos polinômios reais de grau* $\leq n$, *com coeficientes reais, que denotamos por* $\mathcal{P}_n(\mathbb{R})$, *munido da operação de adição de elementos e da operação de multiplicação por escalar definidas de modo análogo ao Exemplo 3.1.5, é um espaço vetorial real. Assim, todo elemento* $p(x) \in \mathcal{P}_n(\mathbb{R})$ *é escrito na forma:*

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n ,$$

com os coeficientes $a_0, a_1, \cdots, a_n \in \mathbb{R}$, *para todo* $x \in \mathbb{R}$.

Exemplo 3.1.8 *O conjunto da matrizes reais de ordem* $m \times n$, *que vamos denotar por* $\mathbb{M}_{m\times n}(\mathbb{R})$, *é um espaço vetorial real, com as operações usuais de soma de matrizes e multiplicação de uma matriz por um escalar.*

Teorema 3.1.1 (Unicidade do Elemento Neutro)
Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Então, existe um único elemento neutro da operação de adição $0_V \in V$.

Demonstração – O axioma (A_3) afirma que existe pelo menos um elemento neutro 0_V em V. Vamos supor que existem dois elementos neutros 0_V e 0_1, isto é,

$$0_V = 0_V + 0_1 = 0_1 + 0_V = 0_1 ,$$

o que prova a unicidade do elemento neutro da operação de adição. ■

Exemplo 3.1.9 *Considere o espaço vetorial $\mathcal{P}_3(\mathbb{R})$. Assim, o elemento neutro da operação de adição é o polinômio $p_0(x) \in \mathcal{P}_3(\mathbb{R})$ definido por:*

$$p_0(x) = a + bx + cx^2 + dx^3 = 0$$

para todo $x \in \mathbb{R}$. Assim, temos que $a = b = c = d = 0$.

Teorema 3.1.2 (Unicidade do Elemento Simétrico)
Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Então, todo elemento $u \in V$ possui um único elemento simétrico.

Demonstração – O axioma (A_4) afirma que todo elemento $u \in V$ possui pelo menos um elemento simétrico $-u \in V$. Vamos supor que o elemento $u \in V$ possui dois elementos simétricos $-u$ e u_1, isto é,

$$u + (-u) = 0_V \qquad \text{e} \qquad u + u_1 = 0_V$$

Desse modo, temos que

$$(-u) = 0_V + (-u) = (u + u_1) + (-u) = 0_V + u_1 = u_1 ,$$

o que prova a unicidade do elemento simétrico. ■

Exemplo 3.1.10 *Considere o espaço vetorial real $\mathcal{C}([a,b])$. Assim, o elemento neutro da operação de adição é a função $f_0 \in \mathcal{C}([a,b])$ dada por:*

$$f_0(x) = 0 \qquad \textit{para todo} \qquad x \in [a,b] \, .$$

Além disso, dada uma função $f \in \mathcal{C}([a,b])$, o seu elemento simétrico é a função $(-f)$ definida por:

$$(-f)(x) = -f(x) \qquad \textit{para todo} \qquad x \in [a,b] \, .$$

Teorema 3.1.3 (Lei do Cancelamento) *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$, $u, v, w \in V$ *e* $u + v = u + w$. *Então,* $v = w$.

Demonstração – Somando $(-u)$ em ambos os lados na igualdade temos

$$v = u + (-u) + v = u + (-u) + w = w$$

o que completa a prova. ■

Teorema 3.1.4 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$, *e* $u, w \in V$. *Então, existe um único elemento* $v \in V$ *tal que* $u + v = w$.

Demonstração – Somando $(-u)$ em ambos os lados da equação tem–se que

$$v = u + (-u) + v = (-u) + w$$

o que completa a prova. ■

Teorema 3.1.5 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$, $u, v \in V$ *e* $\alpha, \beta \in \mathbb{F}$. *Então, temos as seguintes propriedades:*

(a) $0_{\mathbb{F}}\, u = 0_V$;

(b) $\alpha\, 0_V = 0_V$;

(c) $(-\alpha)\, u = -(\alpha\, u) = \alpha\, (-u)$;

(d) se $\alpha\, u = 0_V$, *então* $\alpha = 0_{\mathbb{F}}$ *ou* $u = 0_V$;

(e) se $\alpha\, u = \alpha\, v$ *e* $\alpha \neq 0_{\mathbb{F}}$, *então* $u = v$;

(f) se $\alpha\, u = \beta\, u$ *e* $u \neq 0_V$, *então* $\alpha = \beta$;

(g) $-(u + v) = (-u) + (-v) = -u - v$;

(h) $u + u = 2\, u$, $u + u + u = 3\, u$, *de um modo geral,* $\displaystyle\sum_{i=1}^{n} u = n\, u$.

Demonstração

(a) Seja $v = 0_{I\!F}\, u$. Queremos mostrar que $v = 0_V$. Fazendo

$$v + v = 0_{I\!F}\,(u + u) = v$$

e somando $(-v)$ em ambos os lados da igualdade, obtemos

$$v = v + 0_V = v + (\, v + (-v)\,) = v + (-v) = 0_V$$

Logo, $v = 0_V$. ■

(b) A prova é feita de modo análogo ao item *(a)*, e pode ficar a cargo do leitor. □

(c) Seja $v = (-\alpha)\,u$. Temos que

$$v + \alpha\,u = (-\alpha)\,u + \alpha\,u = (-\alpha + \alpha)\,u = 0_V$$

Assim, obtemos que $v = -(\alpha\,u)$. Vamos provar agora que $v = \alpha\,(-u)$. Fazendo

$$\alpha\,(-u) + \alpha\,u = \alpha(\,(-u) + u\,) = \alpha\,0_V = 0_V$$

provamos que $-(\alpha\,u) = \alpha\,(-u)$. ■

(d) Se $\alpha = 0_{I\!F}$ reduz–se ao caso do item *(a)*. Tomamos $\alpha\,u = 0_V$ com $\alpha \neq 0_F$. Sabemos que existe um único $\alpha^{-1} \in I\!F$ tal que $\alpha\,\alpha^{-1} = 1_{I\!F}$. Desse modo, tem–se que

$$u = 1_{I\!F}\,u = (\alpha^{-1}\,\alpha)\,u = \alpha^{-1}\,(\alpha\,u) = \alpha^{-1}\,0_V = 0_V\,,$$

pelo item *(b)*. Logo, $u = 0_V$. ■

(e) Como $\alpha \neq 0_{I\!F}$, sabemos que existe um único $\alpha^{-1} \in I\!F$ tal que $\alpha\,\alpha^{-1} = 1_{I\!F}$. Desse modo, temos que

$$u = (\alpha^{-1}\,\alpha)\,u = \alpha^{-1}\,(\alpha\,u) = \alpha^{-1}\,(\alpha\,v) = (\alpha^{-1}\,\alpha)\,v = v$$

Logo, $u = v$. ■

(f) Somando $-(\beta\,u)$ em ambos os lados da igualdade $\alpha\,u = \beta\,u$, obtemos

$$\alpha\,u + (-(\beta\,u)) = \alpha\,u + (-\beta)\,u = = (\alpha + (-\beta))\,u = (\alpha - \beta)\,u = 0_V$$

como $u \neq 0_V$, temos que $(\alpha - \beta) = 0_{I\!F}$. Logo, $\alpha = \beta$. ■

(g) A prova é feita de modo análogo ao item *(f)*, e pode ficar a cargo do leitor. □

(h) A prova é feita por indução a partir dos axiomas (A_2) e (M_3) da definição de espaço vetorial. □

Exercícios

Exercício 3.1 *Mostre que o conjunto* $\mathbb{R}^2 = \{ (x,y) \ / \ x, y \in \mathbb{R} \}$ *é um espaço vetorial real, com as operações usuais de adição de elementos e multiplicação por escalar.*

Exercício 3.2 *Mostre que o conjunto de todas as matrizes reais de ordem* n*, que denotamos por* $\mathbb{M}_n(\mathbb{R})$*, com a operação de adição de elementos,* $A = [a_{ij}]$ *e* $B = [b_{ij}]$*, definida por:* $A + B = [a_{ij} + b_{ij}]$ *e a operação de multiplicação por escalar definida por:* $\lambda A = [\lambda a_{ij}]$*, é um espaço vetorial real.*

Exercício 3.3 *Considere o espaço vetorial real* $V = \{ (x,y) \in \mathbb{R}^2 \ / \ x > 0 \}$ *com as operações:*

- ***adição de elementos:*** $(x_1, y_1) \oplus (x_2, y_2) = (x_1 x_2, y_1 + y_2)$.
- ***multiplicação por escalar:*** $\alpha \odot (x, y) = (x^\alpha, \alpha y) \quad , \quad \alpha \in \mathbb{R}$.

(a) *Exiba o elemento neutro da operação adição.*

(b) *Exiba o elemento simétrico aditivo do elemento* $(x, y) \in V$.

(c) *Mostre que* $\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v \quad , \quad u, v \in V \quad e \quad \alpha \in \mathbb{R}$.

Exercício 3.4 *Considere o conjunto* $V = \{ x \in \mathbb{R} \ / \ x > 0 \}$*. Definimos as seguintes operações em* V*:*

1. $x \oplus y = xy , \ \forall \, x, y \in V$;
2. $\alpha \odot x = x^\alpha , \ \forall \, x \in V, \ \forall \, \alpha \in \mathbb{R}$.

Verifique se $(V, \oplus, \odot)$ *é um espaço vetorial real.*

Exercício 3.5 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$ *. Mostre que*

$$Z = V \times W = \{ (v,u) \ / \ v \in V \quad e \quad w \in W \}$$

munido das seguintes operações:

$$(v_1 , w_1) + (v_2 , w_2) = (v_1 + v_2 , w_1 + w_2)$$

$$\lambda(v , w) = (\lambda v , \lambda w) \quad ; \quad \lambda \in \mathbb{F}$$

é um espaço vetorial sobre o corpo $\mathbb{F}$.

3.2 Subespaço Vetorial

Definição 3.2.1 *Seja V um espaço vetorial sobre o corpo $I\!F$. Um* ***subespaço vetorial*** *de V é um subconjunto U de V que é ele mesmo um espaço vetorial sobre o corpo $I\!F$ com as operações de adição de vetores e multiplicação por escalar definidas em V.*

Exemplo 3.2.1 *O subconjunto $S = \{ (x,y) \in I\!R^2 \ / \ y - ax = 0 \ ; \ a \in I\!R \}$ é um subespaço vetorial de $I\!R^2$. Dê uma interpretação geométrica para S.*

Exemplo 3.2.2 *o conjunto $\mathcal{C}_0([a,b]) = \{ f \in \mathcal{C}([a,b]) \ / \ f(a) = f(b) = 0 \}$ é um subespaço vetorial de $\mathcal{C}([a,b])$.*

Exemplo 3.2.3 *O subconjunto S do $I\!R^3$ definido da forma:*

$$S = \{ w \in I\!R^3 \ / \ w = a(1,-1,1) + b(2,1,-1) \quad ; \quad a, b \in I\!R \},$$

é um subespaço vetorial de $I\!R^3$. Dê uma interpretação geométrica para S.

Teorema 3.2.1 (Subespaço Vetorial) *Um subconjunto* ***não vazio*** *U de um espaço vetorial V é um subespaço vetorial de V se, e somente se, para quaisquer elementos $u, v \in U$ e para qualquer escalar $\alpha \in I\!F$, tem–se que $u + v \in U$ e $\alpha u \in U$, isto é, U é fechado com relação as operações de adição e multiplicação por escalar.*

Demonstração

($\Longrightarrow$) Se U é um subespaço vetorial de V, então satisfaz todos os axiomas de espaço vetorial, em particular satisfaz os axiomas de fechamento.

($\Longleftarrow$) Agora, vamos mostrar que se U satisfaz os axiomas de fechamento, então satisfaz os axiomas da adição de elementos e os axiomas da multiplicação por escalar. Como $U \subset V$, os axiomas (A_1) e (A_2) são automaticamente satisfeitos, pois são válidos para todos os elementos de V. De modo análogo, os axiomas (M_1), (M_2), (M_3) e (M_4) são satisfeitos automaticamente. Finalmente, devemos provar somente os axiomas:

(A_3) Elemento Neutro. Para quaisquer $u \in U$ e $\lambda \in I\!F$, temos que $\lambda u \in U$. Fazendo $\lambda = 0_{I\!F}$, obtemos $0_{I\!F}\, u = 0_V \in U$. Logo, U possui elemento neutro.

(A_4) Elemento Simétrico. Para quaisquer $u \in U$ e $\lambda \in I\!F$, temos que $\lambda u \in U$. Fazendo $\lambda = -1_{I\!F}$, obtemos $-1_{I\!F}\, u = 1_{I\!F}\,(-u) = -u \in U$. Logo, todo elemento de U possui o elemento simétrico. O que completa a demonstração. ■

Exemplo 3.2.4 *O subconjunto* $S = \{ (x,y) \in I\!R^2 \ / \ y - 2x = 1 \}$ *não é um subespaço vetorial de* $I\!R^2$.

De fato, o elemento neutro da operação de adição, $0_{I\!R^2} = (0,0)$, não pertence a S. Além disso, o subconjunto S não é fechado com relação às operações de adição de elementos e de multiplicação por escalar.

Exemplo 3.2.5 *o subconjunto* $U = \{ f \in \mathcal{C}([a,b]) \ / \ f(a) = 1 \}$ *não é um subespaço vetorial de* $\mathcal{C}([a,b])$.

De fato, o elemento neutro da operação de adição, $f \equiv 0$, não pertence a U. Além disso, o subconjunto U não é fechado com relação às operações de adição de elementos e de multiplicação por escalar.

Exemplo 3.2.6 *Considere o espaço vetorial real* $\mathcal{P}_3(I\!R)$. *O subconjunto*

$$S = \{ p(x) \in \mathcal{P}_3(I\!R) \ / \ p(-1) = 0 \quad e \quad p'(1) = 0 \}$$

é um subespaço vetorial de $\mathcal{P}_3(I\!R)$.

Para mostrar que S é um subespaço vetorial de $\mathcal{P}_3(I\!R)$, vamos verificar se o elemento neutro da adição pertence a S e se os axiomas de fechamento são satisfeitos. É fácil ver que o polinômio identicamente nulo satisfaz as condições $p(-1) = 0$ e $p'(1) = 0$.

Inicialmente, vamos verificar se o subconjunto S é fechado com relação à operação de adição de elementos, isto é, dados os elementos $p(x)$, $q(x) \in S$ temos que

$$(p + q)(-1) = p(-1) + q(-1) = 0 \qquad \text{e} \qquad (p + q)'(1) = p'(1) + q'(1) = 0$$

Logo, o elemento $(\, p(x) + q(x)\,) \in S$.

Finalmente, vamos verificar se o subconjunto S é fechado com relação à operação de multiplicação por escalar, isto é, dados os elementos $p(x) \in S$ e $\lambda \in I\!R$ temos que

$$(\lambda\, p)(-1) = \lambda\, p(-1) = 0 \qquad \text{e} \qquad (\lambda\, p)'(1) = \lambda\, p'(1) = 0$$

Logo, o elemento $\lambda\, p(x) \in S$. Portanto, o subconjunto S é um subespaço de $\mathcal{P}_3(I\!R)$.

Exemplo 3.2.7 *Considere o sistema linear homogêneo*

$$\begin{cases} -x + 2y + z = 0 \\ 2x - y + z = 0 \end{cases} .$$

Mostre que o conjunto solução é um subespaço do $\mathbb{R}^3$.

Vamos obter a solução do sistema linear utilizando o escalonamento

$$\begin{cases} -x + 2y + z = 0 \\ 2x - y + z = 0 \end{cases} \iff \begin{cases} -x + 2y + z = 0 \\ 3y + 3z = 0 \end{cases}$$

Portanto, temos que $x = -z$ e $y = -z$ com $z \in \mathbb{R}$.

Assim, o conjunto solução do sistema linear pode ser escrito da seguinte forma:

$$S = \{ (x,y,z) \in \mathbb{R}^3 \ / \ (x,y,z) = \alpha(-1,-1,1) \quad , \quad \alpha \in \mathbb{R} \}$$

onde $v = (-1,-1,1) \in \mathbb{R}^3$ é denominada solução básica. Agora podemos verificar facilmente que S é um subespaço do $\mathbb{R}^3$.

Por simplicidade, representamos o sistema linear homogêneo na sua forma matricial

$$AX = \begin{bmatrix} -1 & 2 & 1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

Assim, podemos definir o conjunto solução da seguinte forma:

$$S = \{ (x,y,z) \in \mathbb{R}^3 \ / \ AX = 0 \} \qquad \text{onde} \qquad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} .$$

Temos uma representação mais interessante com a qual podemos obter vários resultados sobre o conjunto solução. Apresentamos o mesmo problema de uma maneira mais geral no Exemplo 3.2.10.

Exemplo 3.2.8 *O subconjunto*

$$S = \left\{ f \in \mathcal{C}([0,1]) \ / \ \int_0^1 f(x)dx \geq 0 \right\}$$

não é um subespaço do espaço vetorial $\mathcal{C}([0,1])$.

De fato, o conjunto S não é fechado em relação à operação de multiplicação por escalar. Tomando um elemento $f \in S$ e um escalar $\lambda \in \mathbb{R}$ negativo, temos que o elemento $(\lambda f) \notin S$. Note que o elemento neutro da operação de adição, $f \equiv 0$, pertence ao conjunto S e o conjunto S é fechado com relação à operação de adição de elementos.

Exemplo 3.2.9 *O subconjunto* U*, do espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$*, definido por:*

$$U = \left\{ p(x) \in \mathcal{P}_2(\mathbb{R}) \ / \ \int_{-1}^{1} p(x)dx \ + \ p'(0) = 0 \right\}$$

é um subespaço vetorial de $\mathcal{P}_2(\mathbb{R})$*.*

Podemos verificar facilmente que U é um subconjunto não vazio de $\mathcal{P}_2(\mathbb{R})$. De fato, o polinômio identicamente nulo satisfaz a condição para que um elemento de $\mathcal{P}_2(\mathbb{R})$ pertença a U, isto é, $0_{\mathcal{P}_2(\mathbb{R})} \in U$.

Assim, devemos mostrar que U é fechado com relação à operação de adição e fechado com relação à operação de multiplicação por escalar.

Tomando os elementos $p(x), q(x) \in U$, isto é, satisfazendo a condição

$$\int_{-1}^{1} p(x)dx \ + \ p'(0) = 0 \qquad \text{e} \qquad \int_{-1}^{1} q(x)dx \ + \ q'(0) = 0 \,.$$

Logo, temos que

$$\begin{aligned}
\int_{1}^{1} (p+q)(x)dx \ + \ (p+q)'(0) &= \int_{1}^{1} (\, p(x) + q(x)\,)dx \ + \ p'(0) \ + \ q'(0) \\
&= \left\{ \int_{-1}^{1} p(x)dx + p'(0) \right\} + \left\{ \int_{-1}^{1} q(x)dx + q'(0) \right\} \\
&= 0
\end{aligned}$$

Assim, mostramos que $(\, p(x) + q(x)\,) \in U$.

Tomando $p(x) \in U$ e $\lambda \in \mathbb{R}$, temos que

$$\begin{aligned}
\int_{1}^{1} (\lambda p)(x)dx \ + \ (\lambda p)'(0) &= \int_{1}^{1} \lambda p(x)dx \ + \ \lambda p'(0) \\
&= \lambda \left\{ \int_{-1}^{1} p(x)dx + p'(0) \right\} = 0 \,.
\end{aligned}$$

Assim, mostramos que $\lambda p(x) \in U$.

Portanto, o subconjunto U é um subespaço vetorial de $\mathcal{P}_2(\mathbb{R})$.

Exemplo 3.2.10 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$. *O subconjunto* S *do* $\mathbb{R}^n$ *definido da forma:*

$$S = \{ x = (x_1, \cdots, x_n) \in \mathbb{R}^n \ / \ AX = 0_{\mathbb{R}^n} \} \qquad onde \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

que é o conjunto solução do sistema linear homogêneo $AX = 0_{\mathbb{R}^n}$, *é um subespaço vetorial de* $\mathbb{R}^n$.

Por simplicidade, vamos utilizar a seguinte representação

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e} \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

para os elementos $x = (x_1, \cdots, x_n)$ e $y = (y_1, \cdots, y_n)$. Assim, podemos fazer a representação matricial do sistema linear homogêneo $AX = 0_{\mathbb{R}^n}$.

Podemos verificar facilmente que o elemento neutro da adição $0_{\mathbb{R}^n} \in S$, isto é, o elemento neutro $0_{\mathbb{R}^n}$ é a solução trivial do sistema linear homogêneo.

Considerando $x, y \in S$ e $\lambda \in \mathbb{R}$, temos que

$$A(X + Y) = AX + AY = 0_{\mathbb{R}^n} \qquad \Longrightarrow \qquad (x + y) \in S$$

e que

$$A(\lambda X) = \lambda AX = 0_{\mathbb{R}^n} \qquad \Longrightarrow \qquad (\lambda x) \in S.$$

Portanto, o subconjunto S é um subespaço vetorial do $\mathbb{R}^n$.

Exemplo 3.2.11 *O conjunto* S *de todas as funções representadas da forma*

$$f(x) = ae^x + be^{-x} \qquad ; \qquad a, b \in \mathbb{R},$$

para $x \in \mathbb{R}$, *é um subespaço vetorial de* $\mathcal{F}(\mathbb{R})$.

De fato, o elemento neutro da operação de adição, $f \equiv 0$, pertence a S, bastando tomar $a = b = 0$. Além disso, o conjunto S é fechado com relação às operações de adição de elementos e de multiplicação por escalar.

Exercícios

Exercício 3.6 *Verifique se o subconjunto* S *de* $\mathbb{M}_n(\mathbb{R})$ *definido por:*

$$S \;=\; \{\; A \;\in\; \mathbb{M}_n(\mathbb{R}) \;\;/\;\; A^2 \;=\; A \;\}\,,$$

o conjunto das matrizes idempotentes, é um subespaço vetorial de $\mathbb{M}_n(\mathbb{R})$.

Exercício 3.7 *Mostre que o subconjunto de* $\mathbb{M}_2(\mathbb{R})$ *dado por:*

$$U \;=\; \left\{ \begin{bmatrix} x & y \\ z & t \end{bmatrix} \;\;/\;\; x - y - z = 0 \right\}$$

é um subespaço vetorial de $\mathbb{M}_2(\mathbb{R})$.

Exercício 3.8 *Considere o espaço vetorial real* $V = \{\,(x,y) \;\;/\;\; x,\, y \in \mathbb{R}\,\}$, *com as operações:*

- ***adição de elementos:*** $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2 + 5,\; y_1 + y_2)$
- ***multiplicação por escalar:*** $\alpha \odot (x, y) = (\alpha\, x + 5(\alpha - 1),\; \alpha\, y) \quad , \quad \alpha \in \mathbb{R}$.

(a) Exiba o elemento neutro da operação adição.

(b) Exiba o elemento simétrico aditivo do elemento $(x, y) \in V$.

(c) Verifique se $W = \{\,(x,y) \in V \;\;/\;\; x = -5\,\}$ *é um subespaço vetorial de* V.

Definição 3.2.2 *Dado um elemento* $c \;=\; (c_1,\,\ldots,\,c_n) \;\in\; \mathbb{R}^n$ *fixo, porém arbitrário, e um escalar* $d \in \mathbb{R}$. *O subconjunto* $H \;\subset\; \mathbb{R}^n$ *definido por:*

$$H \;=\; \{\; (x_1,\,\ldots,\,x_n) \;\in\; \mathbb{R}^n \;\;/\;\; c_1x_1 \;+\; \cdots \;+\; c_nx_n = d \;\}$$

é denominado **hiperplano**.

Exercício 3.9 *Considere um hiperplano* H *contido em* $\mathbb{R}^n$. *Mostre que* H *é um subespaço vetorial de* $\mathbb{R}^n$, *no caso em que* $d = 0$.

Exercício 3.10 *Considere o seguinte subconjunto* S *de* $\mathcal{C}([a,b])$ *definido por:*

$$S \;=\; \{\; f \;\in\; \mathcal{C}([a,b]) \;\;/\;\; f \;\; \textit{é uma função crescente} \;\}\;.$$

Verifique se S *é um subespaço vetorial de* $\mathcal{C}([a,b])$.

Exercício 3.11 *Mostre que o seguinte subconjunto*

$$S = \left\{ f \in \mathcal{C}([0,1]) \ / \ \int_0^1 f(x)dx = 0 \right\}$$

é um subespaço do espaço vetorial $\mathcal{C}([0,1])$.

Exercício 3.12 *Considere o espaço vetorial real* $\mathcal{P}_3(I\!R)$*. Mostre que o subconjunto*

$$U = \{ \, p(x) \in \mathcal{P}_3(I\!R) \ / \ p(-1) = p(1) = 0 \, \}$$

é um subespaço vetorial de $\mathcal{P}_3(I\!R)$.

Exercício 3.13 *Definimos o* ***traço*** *da matriz* $A \in I\!M_n(I\!R)$*, que denotamos por* $tr(A)$*, da seguinte forma:*

$$tr(A) = \sum_{i=1}^{n} a_{ii} \, .$$

Mostre que o subconjunto de $I\!M_n(I\!R)$ *dado por:*

$$S = \{ \, A \in I\!M_n(I\!R) \ / \ tr(A) = 0 \, \}$$

é um subespaço vetorial de $I\!M_n(I\!R)$.

Exercício 3.14 *Mostre que os seguintes subconjuntos de* $I\!M_n(I\!R)$ *definidos por:*

$$U = \{ \, A \in I\!M_n(I\!R) \ / \ A^t = A \, \}$$

$$W = \{ \, A \in I\!M_n(I\!R) \ / \ A^t = -A \, \}$$

são subespaços vetoriais de $I\!M_n(I\!R)$.

Exercício 3.15 *Considere o espaço vetorial real* $\mathcal{C}([-a,a])$ *com* $a \in I\!R_+$*. Mostre que os seguintes subconjuntos*

$$U = \{ \, f \in \mathcal{C}([-a,a]) \ / \ f(-x) = f(x) \ ; \ x \in [-a,a] \, \}$$

$$W = \{ \, f \in \mathcal{C}([-a,a]) \ / \ f(-x) = -f(x) \ ; \ x \in [-a,a] \, \}$$

são subespaços vetoriais de $\mathcal{C}([-a,a])$.

Exercício 3.16 *Considere o subconjunto* S *do espaço vetorial* $I\!R^2$ *definido por:*

$$S = \{ \, v \in I\!R^2 \ / \ v = \alpha(1, \ 2) + (3, \ 2) \ , \ \alpha \in I\!R \, \} \, .$$

Verifique se S *é um subespaço vetorial de* $I\!R^2$.

3.3 Combinação Linear. Subespaço Gerado

Definição 3.3.1 *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$*. Dizemos que o elemento* $u \in V$ *é uma* ***combinação linear*** *dos elementos* $v_1, \cdots, v_n \in V$ *se existem escalares* $c_1, \cdots, c_n \in \mathbb{F}$ *tais que*

$$u = c_1 v_1 + \cdots + c_n v_n .$$

Definição 3.3.2 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *e* S *um conjunto finito de elementos de* V*, isto é,* $S = \{ v_1, \cdots, v_n \}$*. O subconjunto* U *construído a partir dos elementos de* S *da seguinte forma:*

$$U = \left\{ u \in V \ / \ u = \sum_{i=1}^{n} \alpha_i v_i \quad ; \quad \alpha_i \in \mathbb{F} \right\}$$

é um subespaço vetorial de V*, que vamos denotar por*

$$U = [v_1, \cdots, v_n] \qquad \text{ou por} \qquad U = [S] ,$$

denominado ***subespaço gerado*** *pelos elementos de* S*. Dizemos que o conjunto* S *é um* ***sistema de geradores*** *para o subespaço* U*.*

Exemplo 3.3.1 *Considere o seguinte espaço vetorial real*

$$\mathcal{C}_0([-\pi, \pi]) = \{ f \in \mathcal{C}([-\pi, \pi]) \ / \ f(-\pi) = f(\pi) = 0 \}$$

Note que $\mathcal{C}_0([-\pi, \pi])$ *é também um subespaço vetorial de* $\mathcal{C}([-\pi, \pi])$*. Considere o subconjunto* S *de elementos de* $\mathcal{C}_0([-\pi, \pi])$ *dado por:*

$$S = \{ \sin(x), \sin(2x), \cdots, \sin(nx) \}$$

O subconjunto W *definido como:*

$$W = \left\{ f \in \mathcal{C}_0([-\pi, \pi]) \ / \ f(x) = \sum_{k=1}^{n} c_k \sin(kx) \quad ; \quad c_k \in \mathbb{R} \right\}$$

é o subespaço gerado pelos elementos de S*. Logo,* W *é um subespaço de* $\mathcal{C}_0([-\pi, \pi])$*.*

Exemplo 3.3.2 *Considere uma matriz* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$*, com* $m > n$*. Vamos denotar por* $v_1, \cdots, v_n \in \mathbb{R}^m$ *as colunas da matriz* A*. O subconjunto* $\mathcal{R}(A) \subset \mathbb{R}^m$ *definido por:*

$$\mathcal{R}(A) = \left\{ y \in \mathbb{R}^m \ / \ y = \sum_{k=1}^{n} c_k v_k \quad ; \quad c_k \in \mathbb{R} \right\}$$

é o subespaço gerado pelas colunas da matriz A*, denominado* ***espaço coluna de*** A*.*

Exemplo 3.3.3 *Dada a matriz* $A \in \mathbb{M}_{3\times 2}(\mathbb{R})$

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix},$$

mostre que o elemento $v = (-1, 0, 1) \in \mathbb{R}^3$ *pertence ao* ***espaço coluna de*** A.

Basta mostrar que o elemento v pode ser representado pela combinação linear

$$(-1, 0, 1) = a(1, 2, 3) + b(1, 1, 1) \quad ; \quad a, b \in \mathbb{R},$$

isto é, devemos encontrar a solução do sistema linear

$$\begin{array}{rcrcr} a & + & b & = & -1 \\ 2a & + & b & = & 0 \\ 3a & + & b & = & 1 \end{array}$$

Logo, $a = 1$ e $b = -2$ é a única solução do sistema linear acima. Assim, mostramos que o elemento $v \in \mathcal{R}(A)$.

Exemplo 3.3.4 *Considere o espaço vetorial real* $\mathcal{P}_3(\mathbb{R})$. *Dados os elementos*

$$p_1(x) = x^3 - 1$$

$$p_2(x) = x^2 + x - 1$$

$$p_3(x) = x + 2$$

existem escalares $\alpha, \beta \in \mathbb{R}$ *tais que* $p_1(x) = \alpha\, p_2(x) + \beta\, p_3(x)$ *?*

Exemplo 3.3.5 *Considere o espaço vetorial real* $\mathbb{R}^2$. *Dados os elementos*

$$v_1 = (1, 1)$$

$$v_2 = (3, -2)$$

$$v_3 = (1, -1)$$

determine o elemento $u \in \mathbb{R}^2$ *tal que*

$$\frac{u + v_1}{2} + \frac{v_2 + u}{3} = v_3\,.$$

Definição 3.3.3 *Dizemos que um espaço vetorial* V *é* ***finitamente gerado*** *se existe um subconjunto finito* $S \subset V$ *de maneira que* $V = [S]$.

Exemplo 3.3.6 *Considere o subespaço* $W = \{ A \in \mathbb{M}_2(\mathbb{R}) \ / \ A = A^t \}$ *de* $\mathbb{M}_2(\mathbb{R})$. *Mostre que* W *é gerado pelas matrizes*

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad e \quad A_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} .$$

Basta observar que qualquer elemento $A \in \mathbb{M}_2(\mathbb{R})$ é representado de modo único pela combinação linear

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} .$$

Exemplo 3.3.7 *Mostre que o espaço vetorial* $\mathbb{M}_2(\mathbb{R})$ *é gerado pelas matrizes*

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad , \quad A_3 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad e \quad A_4 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} .$$

Basta mostrar que qualquer elemento $A \in \mathbb{M}_2(\mathbb{R})$ é representado pela combinação linear

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \alpha_4 \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} .$$

para $\alpha_1, \ \alpha_2, \ \alpha_3, \ \alpha_4 \in \mathbb{R}$. Assim, devemos mostrar que o sistema linear

$$\begin{array}{ccccccccc} \alpha_1 & + & \alpha_2 & + & \alpha_3 & & & = & a \\ \alpha_1 & + & \alpha_2 & + & & & \alpha_4 & = & b \\ \alpha_1 & + & & & \alpha_3 & + & \alpha_4 & = & c \\ & & \alpha_2 & + & \alpha_3 & + & \alpha_4 & = & d \end{array}$$

possui solução. Escalonando a matriz do sistema linear

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad \longrightarrow \quad \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix} ,$$

obtemos que o $posto(A) = 4$, veja Definição 2.6.3. Assim, concluímos que o sistema linear possui uma única solução. Desse modo, mostramos que o conjunto

$$\{ A_1, A_2, A_3, A_4 \}$$

gera de modo único os elementos do espaço vetorial $\mathbb{M}_2(\mathbb{R})$. Note que para obter a solução do sistema linear, devemos realizar as mesmas operações elementares no lado direito do sistema.

Exercícios

Exercício 3.17 *Considere o subespaço vetorial de* $I\!M_2(I\!R)$ *dado por:*

$$U = \left\{ \begin{bmatrix} x & y \\ z & t \end{bmatrix} \ / \ x - y - z = 0 \right\}.$$

Determine um sistema de geradores para U.

Exercício 3.18 *Considere o subespaço vetorial de* $I\!R^4$ *dado por:*

$$U = \{ (x, y, z, t) \in I\!R^4 \ / \ x - y + z + t = 0 \quad e \quad -x + 2y + z - t = 0 \}.$$

Determine um sistema de geradores para U.

Exercício 3.19 *Seja* W *o subespaço de* $I\!M_{3\times 2}(I\!R)$ *gerado pelas matrizes*

$$A_1 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad e \qquad A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Verifique se a matriz A *dada por:*

$$A = \begin{bmatrix} 0 & 2 \\ 3 & 4 \\ 5 & 0 \end{bmatrix}$$

pertence ao subespaço W.

Exercício 3.20 *Considere o espaço vetorial real* $I\!R^3$. *Dada a matriz*

$$A = \begin{bmatrix} 4 & 2 \\ 1 & 5 \\ 0 & 2 \end{bmatrix},$$

verifique se os elementos $u, v \in I\!R^3$, *dados abaixo, pertencem ao subespaço gerado pelas colunas da matriz* A.

(i) $u = (1, 2, -8)$ *(ii)* $v = (6, -3, -2)$.

Exercício 3.21 *Mostre que as matrizes*

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad e \quad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

formam um sistema de geradores para o subespaço $W = \{ A \in I\!M_2(I\!R) \ / \ A = A^t \}$.

3.4 Soma e Intersecção. Soma Direta

Teorema 3.4.1 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, U e W subespaços vetoriais de V. Então, o subconjunto de V definido por:*

$$U \cap W = \{ v \in V \; / \; v \in U \quad e \quad v \in W \}$$

é um subespaço vetorial de V.

Demonstração – Temos que $U \cap W \neq \emptyset$, pois $0_V \in U$ e $0_V \in W$. Logo, $0_V \in U \cap W$. Agora, basta mostrar que o subconjunto $U \cap W$ satisfaz as condições do Teorema 3.2.1, isto é, que satisfaz os axiomas de fechamento.

Sejam $u, v \in U \cap W$. Logo, $u, v \in U$ e $u, v \in W$. Como U e W são subespaços vetoriais de V temos que $u + v \in U$ e $u + v \in W$. Portanto, mostramos que $U \cap W$ é fechado com relação à operação de adição de elementos.

De modo análogo, seja $u \in U \cap W$. Logo, $u \in U$ e $u \in W$. Como U e W são subespaços vetoriais de V temos que $\lambda u \in U$ e $\lambda u \in W$ para todo $\lambda \in \mathbb{F}$. Portanto, mostramos que $U \cap W$ é fechado com relação a operação de multiplicação por escalar. Desse modo, provamos que $U \cap W$ é um subespaço vetorial de V. ■

Corolário 3.4.1 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Então, a intersecção de uma coleção arbitrária de subespaços de V é um subespaço vetorial de V.*

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 3.4.2 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, U e W subespaços vetoriais de V. Então, o subconjunto de V definido por:*

$$U + W = \{ v \in V \; / \; v = u + w \quad com \quad u \in U \quad e \quad w \in W \}$$

é um subespaço vetorial de V.

Demonstração – Temos que $U + W \neq \emptyset$, pois $0_V \in U$ e $0_V \in W$. Logo, $0_V \in U + W$. Agora, basta mostrar que o subconjunto $U + W$ satisfaz as condições do Teorema 3.2.1, isto é, que satisfaz os axiomas de fechamento. □

Definição 3.4.1 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, U e W subespaços vetoriais de V tais que $U \cap W = \{ 0_V \}$. Dizemos que o subespaço $U + W$ é a* ***soma direta*** *dos subespaços U e W, e denotamos por $U \oplus W$.*

Exemplo 3.4.1 *Considere os seguintes subespaços de* $\mathbb{R}^3$

$$U = \{ (x, y, z) \in \mathbb{R}^3 \ / \ z = 0 \} \qquad e \qquad W = \{ (x, y, z) \in \mathbb{R}^3 \ / \ y = 0 \} .$$

Temos que $\mathbb{R}^3 = U + W$*, entretanto, não como soma direta dos subespaços* U *e* W*.*

Podemos verificar facilmente que

$$U = [(1,0,0), (0,1,0)] \qquad \text{e} \qquad W = [(1,0,0), (0,0,1)] .$$

Assim, temos que

$$U + W = [(1,0,0), (0,1,0), (0,0,1)] \qquad \text{e} \qquad U \cap W = [(1,0,0)] .$$

Portanto, temos que $\mathbb{R}^3 = U + W$, mas não como soma direta.

Exemplo 3.4.2 *Considere os seguintes subespaços de* $\mathbb{R}^2$

$$U = \{ (x, y) \in \mathbb{R}^2 \ / \ y = 0 \} \qquad e \qquad W = \{ (x, y) \in \mathbb{R}^2 \ / \ x = 0 \} .$$

Temos que $\mathbb{R}^2 = U + W$ *é uma soma direta dos subespaços* U *e* W*.*

Podemos verificar facilmente que

$$U = [(1,0)] \qquad \text{e} \qquad W = [(0,1)] .$$

Assim, temos que

$$U + W = [(1,0), (0,1)] \qquad \text{e} \qquad U \cap W = \{ 0_{\mathbb{R}^2} \} .$$

Portanto, temos que $\mathbb{R}^2 = U \oplus W$.

Exemplo 3.4.3 *Considere os seguintes subespaços de* $\mathbb{R}^2$

$$U = \{ (x, y) \in \mathbb{R}^2 \ / \ y = x \} \qquad e \qquad W = \{ (x, y) \in \mathbb{R}^2 \ / \ y = -2x \} .$$

Temos que $\mathbb{R}^2 = U + W$ *é uma soma direta dos subespaços* U *e* W*.*

Podemos verificar facilmente que

$$U = [(1,1)] \qquad \text{e} \qquad W = [(1,-2)] .$$

Assim, temos que

$$U + W = [(1,1), (1,-2)] \qquad \text{e} \qquad U \cap W = \{ 0_{\mathbb{R}^2} \} .$$

Portanto, temos que $\mathbb{R}^2 = U \oplus W$.

Definição 3.4.2 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, U e W subespaços vetoriais de V. Dizemos que o espaço vetorial V é a soma direta dos subespaços U e W, e denotamos $V = U \oplus W$, se*

1. $U \cap W = \{ 0_V \}$
2. $V = U + W$

Proposição 3.4.1 *Sejam U e W subespaços vetoriais de um espaço vetorial V. Então, $V = U \oplus W$ se, e somente se, cada elemento $v \in V$ possui uma única decomposição $v = u + w$, com $u \in U$ e $w \in W$.*

Demonstração – ($\Longrightarrow$) Considerando por hipótese $V = U \oplus W$, temos a existência da decomposição, então basta mostrar a unicidade. Para isso, supomos que

$$v = u + w = u_1 + w_1 \quad ; \quad u,\, u_1 \in U \quad \text{e} \quad w,\, w_1 \in W$$

Assim, obtemos $u - u_1 = w - w_1$. Como $u - u_1 \in U$, então,

$$w - w_1 \in W \cap U = \{ 0_V \}$$

Logo, $w - w_1 = 0_V$ o que implica em $w = w_1$. De mesmo modo, temos que $u - u_1 = 0_V$ implicando em $u = u_1$, mostrando a unicidade da decomposição.

($\Longleftarrow$) Tomando por hipótese a unidade da decomposição $v = (u + w) \in V$, com $u \in U$ e $w \in W$, vamos mostrar que $V = U \oplus W$. Para isso, supomos que $v \in U \cap W$. Desse modo, temos que

$$u + w = (u + v) + (w - v)$$

Pela unicidade da decomposição, podemos afirmar que

$$u = u + v \qquad \text{e} \qquad w = w - v$$

Logo, $v = 0_V$. Assim, provamos que $U \cap W = \{ 0_V \}$. Portanto, mostramos que $V = U \oplus W$, o que completa a demonstração. ■

Exemplo 3.4.4 *Considere os seguintes subespaços de $\mathbb{R}^2$*

$$U = \{ (x,y) \in \mathbb{R}^2 \ / \ y = x \} \qquad e \qquad W = \{ (x,y) \in \mathbb{R}^2 \ / \ y = -x \}.$$

Temos que todo elemento $v \in \mathbb{R}^2$ é escrito de modo único como $v = u + w$ onde $u \in U$ e $w \in W$, isto é, $\mathbb{R}^2 = U \oplus W$.

Podemos verificar facilmente que $U + W = [(1,1), (1,-1)]$. Consideramos um elemento genérico $v = (x,y) \in I\!R^2$. Escrevendo

$$(x,y) = a(1,1) + b(1,-1)$$

obtemos o seguinte sistema linear

$$\begin{cases} a + b = x \\ a - b = y \end{cases}$$

que tem por única solução

$$a = \frac{x+y}{2} \qquad \text{e} \qquad b = \frac{x-y}{2} .$$

Assim, mostramos que os coeficientes da combinação linear, a e b, são obtidos de modo único em função das componentes do elemento genérico $v = (x,y) \in I\!R^2$. Portanto, obtemos o resultado desejado.

Do Corolário 3.4.1 decorre que se S é uma coleção arbitrária de elementos de V, então existe um menor subespaço de V que contém S, isto é, um subespaço de V que contém S e que está contido em todos os outros subespaços que contém S. Desse modo, podemos apresentar o conceito de subespaço gerado da forma a seguir.

Definição 3.4.3 *Sejam V um espaço vetorial sobre o corpo $I\!F$ e S um conjunto de elementos de V. O **subespaço gerado** por S, que vamos denotar por $U = [S]$, é definido como sendo a intersecção de todos os subespaços de V que contém o conjunto S. Quando S é um conjunto finito de elementos de V, isto é, $S = \{ v_1 , \cdots , v_n \}$, dizemos que U é o **subespaço gerado** pelos elementos de S. O conjunto S também é chamado de **sistema de geradores** do subespaço U.*

Exemplo 3.4.5 *Considere os seguintes subespaços de* $\mathbb{R}^3$

$$U = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x = z \}$$

$$W = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x + y + z = 0 \}$$

Determine um sistema de geradores para o subespaço $U \cap W$.

Inicialmente, vamos determinar um sistema de geradores para o subespaço U. Para os elementos $u \in U$ temos que

$$u = a(1,0,1) + b(0,1,0) \qquad \text{para} \qquad a, b \in \mathbb{R}.$$

Logo, $\{ (1,0,1), (0,1,0) \}$ é um sistema de geradores para o subespaço U.

Agora vamos determinar um sistema de geradores para o subespaço W. Para os elementos $w \in W$ temos que

$$w = c(-1,1,0) + d(-1,0,1) \qquad \text{para} \qquad c, d \in \mathbb{R}$$

Logo, $\{ (-1,1,0), (-1,0,1) \}$ é um sistema de geradores para o subespaço W.

Vamos determinar um sistema de geradores para o subespaço $U \cap W$. Sabemos que, se $v \in U \cap W$, então $v \in U$ e $v \in W$. Assim, temos que

$$a(1,0,1) + b(0,1,0) = c(-1,1,0) + d(-1,0,1) \qquad \text{para} \qquad a, b, c, d \in \mathbb{R}.$$

Desse modo, obtemos o seguinte sistema linear

$$\begin{cases} a = -c - d \\ b = c \\ a = d \end{cases}$$

cuja solução é dada por $a = d$, $b = -2d$ e $c = -2d$ para $d \in \mathbb{R}$.

Portanto, os elementos $v \in U \cap W$ são escritos como $v = d(1,-2,1)$ para $d \in \mathbb{R}$. Logo, $\{ (1,-2,1) \}$ é um sistema de geradores para o subespaço $U \cap W$.

Exemplo 3.4.6 *Considere os seguintes subespaços de* $\mathbb{R}^3$

$$U = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x - 2y + 3z = 0 \}$$

$$W = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x + y + z = 0 \}$$

Determine um sistema de geradores para o subespaço $U \cap W$.

Os elementos $v = (x,y,z) \in U \cap W$ satisfazem as equações

$$\begin{cases} x - 2y + 3z = 0 \\ x + y + z = 0 \end{cases} \iff \begin{cases} x - 2y + 3z = 0 \\ 3y - 2z = 0 \end{cases}$$

Escolhendo x e y como variáveis básicas e z como variável livre, temos que

$$y = \frac{2}{3}z \quad \text{e} \quad x = -\frac{5}{3}z \quad , \quad z \in \mathbb{R}.$$

Desse modo, o conjunto solução do sistema linear homogêneo é escrito da seguinte forma:

$$(x,y,z) = \alpha(-5,2,3) \quad , \quad \alpha \in \mathbb{R}.$$

Portanto, temos que $U \cap W = [(-5,2,3)]$.

Exemplo 3.4.7 *Sejam* U *e* W *subespaços vetoriais do* $\mathbb{R}^3$ *dados por:*

$$U = \{ u \in \mathbb{R}^3 \ / \ u = \lambda\overline{u} \quad , \quad \lambda \in \mathbb{R} \}$$

$$W = \{ w \in \mathbb{R}^3 \ / \ w = \alpha\overline{w} \quad , \quad \alpha \in \mathbb{R} \}$$

com $\overline{u}, \overline{w} \in \mathbb{R}^3$ *não-nulos. Temos que o subespaço* $U + W = [\overline{u}, \overline{w}]$.

Exemplo 3.4.8 *Considere os seguintes subespaços do* $\mathbb{R}^3$

$$U = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x - y - z = 0 \} \quad e \quad W = [(1,2,1)]$$

Temos que o espaço vetorial $\mathbb{R}^3 = U \oplus W$.

Podemos verificar facilmente que $U = [(1,1,0), (1,0,1)]$. Vamos mostra que qualquer elemento $u = (x,y,z)$ de $\mathbb{R}^3$ é escrito de modo único pela combinação linear

$$u = a(1,1,0) + b(1,0,1) + c(1,2,1) \quad , \quad a, b, c \in \mathbb{R}.$$

Portanto, basta mostrar que a matriz A dada por:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

é não–singular. Sendo assim, obtemos de modo único os coeficientes da combinação linear, a, b e c, em função das componentes do elemento $u = (x,y,z) \in \mathbb{R}^3$.

Exemplo 3.4.9 *Considere os seguintes subespaços de* $I\!R^3$

$$U = \{ (x, y, z) \in I\!R^3 \ / \ x - 2y + 3z = 0 \}$$

$$W = \{ (x, y, z) \in I\!R^3 \ / \ x + y + z = 0 \}$$

Determine um sistema de geradores para o subespaço $U + W$.

Inicialmente, vamos determinar um sistema de geradores para o subespaço U. Para os elementos $u \in U$ temos que

$$u = a(2, 1, 0) + b(-3, 0, 1) \qquad \text{para} \qquad a, b \in I\!R$$

Logo, $\{ (2, 1, 0), (-3, 0, 1) \}$ é um sistema de geradores para o subespaço U.

Agora vamos determinar um sistema de geradores para o subespaço W. Para os elementos $w \in W$ temos que

$$w = c(-1, 1, 0) + d(-1, 0, 1) \qquad \text{para} \qquad c, d \in I\!R$$

Logo, $\{ (-1, 1, 0), (-1, 0, 1) \}$ é um sistema de geradores para o subespaço W.

Portanto, o subespaço $U + W$ tem como um sistema de geradores os seguintes elementos

$$v_1 = (2, 1, 0) \quad , \quad v_2 = (-3, 0, 1) \quad , \quad v_3 = (-1, 1, 0) \quad \text{e} \quad v_4 = (-1, 0, 1) \,,$$

isto é, $U + W = [v_1, v_2, v_3, v_4]$.

Exemplo 3.4.10 *Considere os seguintes subespaços do* $I\!R^3$

$$U = [(1, 2, 1), (-1, 1, -1)] \qquad e \qquad W = [(2, 2, 1), (1, 1, -1)] \,.$$

Encontre um sistema de geradores para o subespaço $U \cap W$.

Temos que se $v \in U \cap W$, então $v \in U$ e $v \in W$. Assim, temos que

$$a(1, 2, 1) + b(-1, 1, -1) = c(2, 2, 1) + d(1, 1, -1) \qquad \text{para} \qquad a, b, c, d \in I\!R \,.$$

Desse modo, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a - b = 2c + d \\ 2a + b = 2c + d \\ a - b = c - d \end{cases} \iff \begin{cases} a - b - 2c - d = 0 \\ 2a + b - 2c - d = 0 \\ a - b - c + d = 0 \end{cases}$$

cuja solução é dada por $a = -2d$, $b = d$ e $c = -2d$ para $d \in I\!R$. Portanto, temos que os elementos $v \in U \cap W$ são escritos como $v = d(1, 1, 1)$ para $d \in I\!R$. Logo, $\{ (1, 1, 1) \}$ é um sistemas de geradores para o subespaço $U \cap W$.

Exercícios

Exercício 3.22 *Considere o espaço vetorial real* $\mathbb{R}^3$ *e os subespaços gerados*

$$U = [(1,0,0),(1,1,1)] \qquad e \qquad W = [(0,1,0),(0,0,1)].$$

Determine um sistema de geradores para o subespaço $V = U \cap W$.

Exercício 3.23 *Considere os seguintes subespaços*

$$U = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x + y = 0 \ e \ z - t = 0 \}$$

$$W = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x - y - z + t = 0 \}$$

Pede–se:

(a) Determine um sistema de geradores para o subespaço $U \cap W$.

(b) Determine um sistema de geradores para o subespaço $U + W$.

(c) O subespaço $U + W$ *é uma soma direta ? Justifique sua resposta.*

Exercício 3.24 *Sejam* U *o subespaço do* $\mathbb{R}^3$ *gerado pelo elemento* $u_1 = (1,\,0,\,0)$ *e* W *o subespaço do* $\mathbb{R}^3$ *gerado pelos elementos* $w_1 = (1,\,1,\,0)$ *e* $w_2 = (0,\,1,\,1)$. *Mostre que o espaço vetorial* $\mathbb{R}^3 = U \oplus W$.

Exercício 3.25 *Considere os seguintes subespaços vetoriais de* $\mathbb{M}_n(\mathbb{R})$

$$U = \{ A \in \mathbb{M}_n(\mathbb{R}) \ / \ A^t = A \} \qquad e \qquad W = \{ A \in \mathbb{M}_n(\mathbb{R}) \ / \ A^t = -A \}.$$

Mostre que $\mathbb{M}_n(\mathbb{R}) = U \oplus W$.

Exercício 3.26 *Considere os seguintes subespaços vetoriais de* $\mathcal{C}([-a,a])$

$$U = \{ f \in \mathcal{C}([-a,a]) \ / \ f(-x) = f(x) \ ; \ x \in [-a,a] \}$$

$$W = \{ f \in \mathcal{C}([-a,a]) \ / \ f(-x) = -f(x) \ ; \ x \in [-a,a] \}$$

Mostre que $\mathcal{C}([-a,a]) = U \oplus W$.

Exercício 3.27 *Considere o subespaço* V *do espaço vetorial* $\mathbb{R}^3$ *dado por:*

$$V = \{ (x,\,y,\,z) \in \mathbb{R}^3 \ / \ x + 2y + z = 0 \quad e \quad -x + 3y + 2z = 0 \},$$

Determine um subespaço W *do* $\mathbb{R}^3$ *tal que* $\mathbb{R}^3 = V \oplus W$.

Exercício 3.28 *Considere o espaço vetorial real $I\!R^2$ e os seguinte subespaços*

$$U = \{ (x,y) \in I\!R^2 \ / \ y = 3x \} \quad e \quad W = \{ (x,y) \in I\!R^2 \ / \ y = -2x \} .$$

Verifique se o seguinte subconjunto

$$U \cup W = \{ (x,y) \in I\!R^2 \ / \ (x,y) \in U \quad ou \quad (x,y) \in W \}$$

é um subespaço vetorial de $I\!R^2$.

Exercício 3.29 *Encontre o conjunto solução $S \subset I\!R^3$ do sistema linear homogêneo*

$$\begin{cases} 2x + 4y + z = 0 \\ x + y + 2z = 0 \\ x + 3y - z = 0 \end{cases}$$

Mostre que S é um subespaço do $I\!R^3$. Dado o subespaço

$$U = \{ (x,y,z) \in I\!R^3 \ / \ x - y + z = 0 \} ,$$

determine um sistema de geradores para o subespaço $S \cap U$.

Exercício 3.30 *Considere o subespaço S do $I\!R^3$ definido por:*

$$S = \{ (x,y,z) \in I\!R^3 \ / \ x + 2y + z = 0 \} .$$

Determine um subespaço W do $I\!R^3$ tal que $R^3 = S \oplus W$.

Exercício 3.31 *Considere os seguintes subespaços do espaço vetorial real $I\!M_2(I\!R)$*

$$W = \left\{ \begin{bmatrix} a & b \\ c & a \end{bmatrix} \quad ; \quad a \, , \, b \, , \, c \in I\!R \right\}$$

$$U = \left\{ \begin{bmatrix} 0 & a \\ -a & b \end{bmatrix} \quad ; \quad a \, , \, b \in I\!R \right\}$$

Determine um sistema de geradores para os subespaços

$$W \quad , \quad U \quad , \quad W \cap U \qquad e \qquad W + U \, .$$

Exercício 3.32 *Sejam W o subespaço de $I\!R^3$ gerado pelo vetor $w = (1,0,0)$ e U o subespaço do $I\!R^3$ gerado pelos vetores $u_1 = (1,1,0)$ e $u_2 = (0,1,1)$. Mostre que $I\!R^3 = U \oplus W$.*

3.5 Dependência e Independência Linear

Definição 3.5.1 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $v_1, \cdots, v_n \in V$. Dizemos que o conjunto $S = \{ v_1, \cdots, v_n \} \subset V$ é **Linearmente Independente**(LI) se, e somente se, toda combinação linear nula*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V \quad ; \quad \alpha_i \in \mathbb{F}$$

implicar que $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Definição 3.5.2 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $v_1, \cdots, v_n \in V$. Dizemos que o conjunto $S = \{ v_1, \cdots, v_n \} \subset V$ é **Linearmente Dependente**(LD) se, e somente se, é possível uma combinação linear nula*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V \quad ; \quad \alpha_i \in \mathbb{F}$$

sem que os escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ sejam todos nulos.

De maneira equivalente, encontramos o conceito de dependência e independência linear apresentado da forma a seguir.

Definição 3.5.3 *Seja V um espaço vetorial sobre $\mathbb{F}$. Um subconjunto S de V é dito **Linearmente Dependente** (LD), se existirem elementos distintos $v_1 , \cdots , v_n$ em S e escalares $\alpha_1 , \cdots , \alpha_n$ em $\mathbb{F}$, não todos nulos, tais que*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V .$$

*Um conjunto que não é linearmente dependente é **Linearmente Independente**(LI). Decorrem facilmente da definição as seguintes conseqüências:*

(a) Todo conjunto que contém um subconjunto linearmente dependente é LD.

(b) Todo subconjunto de um conjunto linearmente independente é LI.

(c) Todo conjunto que contém o elemento neutro, 0_V, é linearmente dependente.

(d) Um conjunto S de vetores é linearmente independente se, e somente se, todo subconjunto finito de S é linearmente independente, isto é, se, e somente se, para quaisquer elementos distintos $v_1, \cdots , v_n$ em S, com

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V$$

implicar que $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

(e) Convencionaremos que o conjunto vazio, $\emptyset \subset V$, é linearmente independente.

Teorema 3.5.1 *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *e* $v_1, \cdots, v_n \in V$. *O conjunto* $S = \{ v_1, \cdots, v_n \} \subset V$ *é Linearmente Dependente (LD) se, e somente se, um de seus elementos for uma combinação linear dos outros elementos.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 3.5.1 *O conjunto* $S = \{ v_1, v_2, v_3 \}$ *onde*

$$v_1 = (1,1,0) \quad , \quad v_2 = (1,4,5) \quad e \quad v_3 = (3,6,5) \, ,$$

é linearmente dependente no espaço vetorial $\mathbb{R}^3$.

Considerando a combinação linear nula $x\,v_1 + y\,v_2 + z\,v_3 = 0_{\mathbb{R}^3}$, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} x + y + 3z = 0 \\ x + 4y + 6z = 0 \\ \quad 5y + 5z = 0 \end{cases} \iff \begin{cases} x + y + 3z = 0 \\ \quad y + z = 0 \end{cases}$$

Assim, obtemos que o sistema possui infinitas soluções não–nulas, provando que o conjunto S é linearmente dependente.

Podemos verificar facilmente que $v_3 = 2v_1 + v_2$. Assim, utilizando o resultado do Teorema 3.5.1, mostramos que o conjunto S é linearmente dependente.

Exemplo 3.5.2 *O conjunto* $S = \{ v_1, v_2, v_3 \}$ *onde*

$$v_1 = (1,2,3) \quad , \quad v_2 = (1,4,9) \quad e \quad v_3 = (1,8,27) \, ,$$

é linearmente independente no espaço vetorial $\mathbb{R}^3$.

Considerando a combinação linear nula $x\,v_1 + y\,v_2 + z\,v_3 = 0_{\mathbb{R}^3}$, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} x + y + z = 0 \\ 2x + 4y + 8z = 0 \\ 3x + 9y + 27z = 0 \end{cases} \iff \begin{cases} x + y + 3z = 0 \\ \quad 2y + 6z = 0 \\ \quad\quad 6z = 0 \end{cases}$$

Assim, o sistema linear homogêneo possui somente a solução nula, isto é, $x = y = z = 0$, provando que o conjunto S é linearmente independente.

Definição 3.5.4 *Considere o espaço vetorial real* $\mathcal{C}([a,b])$. *Dizemos que o conjunto de funções* $S = \{ f_1(x), \cdots, f_n(x) \} \subset \mathcal{C}([a,b])$ *é **Linearmente Dependente**, se existirem escalares* $c_1, \cdots, c_n$, *não todos nulos, tais que*

$$c_1 f_1(x) + \cdots + c_n f_n(x) = 0 \quad ; \quad \forall \; x \in [a,b] \, .$$

O conjunto S *é **Linearmente Independente** se não for Linearmente Dependente.*

Exemplo 3.5.3 *O conjunto* $S = \{ 1, \cos(x), \cos(2x) \}$ *é linearmente independente no espaço vetorial* $\mathcal{C}([-\pi,\pi])$.

Considere a combinação linear nula

$$\alpha + \beta \cos(x) + \lambda \cos(2x) = 0 \quad ; \quad \forall \, x \in [-\pi,\pi] \, .$$

Avaliando a equação acima nos pontos $x = -\pi$, $x = 0$ e $x = \dfrac{\pi}{2}$, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} \alpha - \beta + \lambda = 0 \\ \alpha + \beta + \lambda = 0 \\ \alpha \qquad\quad - \lambda = 0 \end{cases}$$

Analisando o conjunto solução do sistema linear homogêneo, através de escalonamento, obtemos $\alpha = \beta = \gamma = 0$. Desse modo, provamos que o conjunto S é linearmente independente em $\mathcal{C}([-\pi,\pi])$, de acordo com a Definição 3.5.4.

Exemplo 3.5.4 *O conjunto* $S = \{ 1, \; x, \; x^2, \; 2 - 3x + 2x^2 \}$ *é linearmente dependente no espaço vetorial* $\mathcal{P}_3(\mathbb{R})$.

Por simplicidade, vamos denotar

$$p_1(x) = 1 \quad , \quad p_2(x) = x \quad , \quad p_3(x) = x^2 \quad \text{e} \quad p_4(x) = 2 - 3x + 2x^2 \, .$$

Podemos verificar facilmente que $p_4(x) = 2p_1(x) - 3p_2(x) + 2p_3(x)$. Assim, utilizando o resultado do Teorema 3.5.1, mostramos que o conjunto S é linearmente dependente.

Exemplo 3.5.5 *O conjunto* $S = \{ \cos^2(x), \sin^2(x), 1 \}$ *é linearmente dependente no espaço vetorial* $\mathcal{F}(\mathbb{R})$. *De fato, fazendo uso da identidade trigonométrica*

$$\cos^2(x) + \sin^2(x) = 1 \quad ; \quad x \in \mathbb{R}$$

obtemos o resultado desejado.

A seguir, iniciamos a apresentação de uma caracterização para que um conjunto de funções

$$S = \{ f_1(x), f_2(x), \cdots, f_n(x) \}$$

seja linearmente dependente em um determinado intervalo.

Teorema 3.5.2 *Considere o espaço vetorial real* $\mathcal{C}^1([a,b])$ *e as funções* $f, g \in \mathcal{C}^1([a,b])$. *O conjunto* $S = \{ f(x), g(x) \}$ *é* ***linearmente dependente*** *se, e somente se,* $\det(A(x)) = 0$ *para todo* $x \in [a,b]$, *onde a matriz* $A(x)$ *é dada por:*

$$A(x) = \begin{bmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{bmatrix} \qquad ; \qquad x \in [a,b] .$$

O determinante da matriz A *é denominado* ***wronskiano*** *das funções* f *e* g, *que vamos denotar por* $\mathbf{W}(f,g)(x)$.

Demonstração – Vamos considerar a combinação linear nula

$$c_1 f(x) + c_2 g(x) = 0 \qquad ; \qquad \forall \; x \in [a,b] .$$

Considerando a equação acima e a sua derivada com relação à x, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} c_1 f(x) & + & c_2 g(x) & = & 0 \\ c_1 f'(x) & + & c_2 g'(x) & = & 0 \end{cases} .$$

Sabemos que o sistema linear homogêneo possui solução não–nula se, e somente se, o determinante da matriz do sistema linear for igual à zero, isto é,

$$\mathbf{W}(f,g)(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix} = 0 \qquad ; \qquad \forall \; x \in [a,b] .$$

Assim, completamos a demonstração. ■

Exemplo 3.5.6 *Vamos utilizar o resultado do Teorema 3.5.2 para mostrar que as funções* $f(x) = \exp(x)$ *e* $g(x) = x\exp(x)$ *são linearmente independentes para* $x \in I\!R$.

Podemos observar facilmente que as funções f e g são continuamente diferenciáveis para todo $x \in I\!R$. Vamos calcular o wronskiano das funções f e g

$$\mathbf{W}(f,g)(x) = \begin{vmatrix} e^x & xe^x \\ e^x & e^x(1 + x) \end{vmatrix} = e^{2x} \neq 0 \qquad \text{para todo} \qquad x \in I\!R .$$

Assim, obtemos o resultado desejado.

Exemplo 3.5.7 *As funções* $f(x) = \sin(x)$ *e* $g(x) = x\sin(x)$ *são linearmente independentes para* $x \in \mathbb{R}$. *De fato, o wronskiano* $\mathbf{W}(f,g)(x) = -(\sin(x))^2$ *não é sempre nulo para* $x \in \mathbb{R}$.

A seguir, vamos apresentar uma extensão do Teorema 3.5.2, que o leitor poderá facilmente generalizar para um conjunto com n funções que sejam $(n - 1)$ vezes continuamente diferenciáveis para $x \in [a,b]$.

Teorema 3.5.3 *Considere o espaço vetorial* $\mathcal{C}^2([a,b])$ *e as funções* $f,\ g,\ h \in \mathcal{C}^2([a,b])$. *O conjunto* $S = \{ f(x),\ g(x),\ h(x) \}$ *é* ***linearmente dependente*** *se, e somente se,* $\mathbf{W}(f,g,h)(x) = 0$ *para todo* $x \in [a,b]$, *onde o wronskiano* $\mathbf{W}(f,g,h)(x)$ *é dado por:*

$$\mathbf{W}(f,g,h)(x) = \begin{vmatrix} f(x) & g(x) & h(x) \\ f'(x) & g'(x) & h'(x) \\ f''(x) & g''(x) & h''(x) \end{vmatrix} \quad ; \quad x \in [a,b]\,.$$

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 3.5.8 *Vamos utilizar o resultado do Teorema 3.5.3 para mostrar que as funções* $f(x) = 1$, $g(x) = \sin(x)$ *e* $h(x) = \cos(x)$ *são linearmente independentes para* $x \in \mathbb{R}$.

Podemos observar facilmente que as funções f , g e h são duas vezes continuamente diferenciáveis para todo $x \in \mathbb{R}$. Vamos calcular o wronskiano $\mathbf{W}(f,g,h)(x)$

$$\mathbf{W}(f,g)(x) = \begin{vmatrix} 1 & \sin(x) & \cos(x) \\ 0 & \cos(x) & -\sin(x) \\ 0 & -\sin(x) & -\cos(x) \end{vmatrix} = -1 \quad \text{para todo} \quad x \in \mathbb{R}\,.$$

Assim, obtemos o resultado desejado.

Exemplo 3.5.9 *Podemos verificar facilmente que as funções*

$$f(x) = \sin(x) \qquad e \qquad g(x) = \cos\left(x - \frac{\pi}{2}\right)$$

são linearmente dependentes para $x \in \mathbb{R}$. *De fato, basta utilizar a Definição 3.5.4 e a identidade trigonométrica para o cosseno da diferença.*

Exercícios

Exercício 3.33 *Verifique quais dos subconjuntos*

(a) $\{ (1,0,0) , (0,1,0) , (0,0,1) , (2,2,5) \}$

(b) $\{ (1,1,1) , (1,2,1) , (3,2,-1) \}$

são linearmente independentes no espaço vetorial real $\mathbb{R}^3$.

Exercício 3.34 *Verifique quais dos subconjuntos*

(a) $\{ 1 , x-1 , x^2+2x+1 , x^2 \}$

(b) $\{ x(x-1) , x^3 , 2x^3-x^2 , x \}$

são linearmente independentes no espaço vetorial real $\mathcal{P}_4(\mathbb{R})$.

Exercício 3.35 *Mostre que o conjunto*

$$\gamma = \{ (1,0,0,-1), (0,1,0,1), (0,0,1,1), (1,0,0,1) \}$$

é linearmente independente no espaço vetorial real $\mathbb{R}^4$.

Exercício 3.36 *Considere as seguintes funções* $f(x) = x$ *e* $g(x) = |\,x\,|$. *Mostre que o conjunto* $S = \{ f(x), g(x) \}$ *é linearmente independente no espaço vetorial* $\mathcal{C}([-1,1])$.

Exercício 3.37 *Considere as seguintes funções* $f(x) = x^2$ *e* $g(x) = x\,|\,x\,|$. *Mostre que o conjunto* $S = \{ f(x), g(x) \}$ *é linearmente independente no espaço vetorial real* $\mathcal{C}([-1,1])$. *Entretanto, é linearmente dependente em* $\mathcal{C}([0,1])$ *e em* $\mathcal{C}([-1,0])$.

Exercício 3.38 *Determine três elementos de* $\mathbb{R}^3$ *que sejam linearmente dependentes e tais que dois quaisquer sejam linearmente independentes.*

Exercício 3.39 *Considere* V *um espaço vetorial sobre o corpo* $\mathbb{F}$. *Mostre que se dois elementos de* V *são linearmente dependentes, então um é múltiplo escalar do outro.*

Exercício 3.40 *Mostre que o conjunto* $S = \{ 1, e^x, xe^x \}$ *é linearmente independente no espaço vetorial* $\mathcal{F}(\mathbb{R})$.

Exercício 3.41 *Mostre que o conjunto* $S = \{ 1, e^x, e^{2x} \}$ *é linearmente independente no espaço vetorial* $\mathcal{F}(\mathbb{R})$.

3.6 Bases e Dimensão

Passamos agora à tarefa de atribuir uma dimensão a certos espaços vetoriais. Apesar de associarmos usualmente **dimensão** a algo geométrico, precisamos encontrar uma definição algébrica adequada da dimensão de um espaço vetorial. Isto será feito através do conceito de uma base para o espaço vetorial.

Definição 3.6.1 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Uma* ***base*** *de V é um conjunto linearmente independente de elementos de V que gera V.*

Exemplo 3.6.1 *Considere o espaço vetorial real $\mathbb{R}^3$. O conjunto*

$$\beta = \{ (1,0,0), (0,1,0), (0,0,1) \}$$

é linearmente independente em $\mathbb{R}^3$ e gera o espaço $\mathbb{R}^3$. Logo, β é uma base para $\mathbb{R}^3$, denominada base canônica.

Exemplo 3.6.2 *Considere o espaço vetorial real $\mathbb{R}^2$. O conjunto*

$$\Gamma = \{ (1,1), (-1,1) \}$$

é linearmente independente em $\mathbb{R}^2$ e gera o espaço $\mathbb{R}^2$. Logo, Γ é uma base para $\mathbb{R}^2$.

Teorema 3.6.1 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$ finitamente gerado pelos elementos do conjunto $S = \{ v_1, \cdots, v_n \}$. Então, podemos extrair do conjunto S uma base para V.*

Demonstração – Se $v_1, \cdots, v_n \in V$ são linearmente independentes, então eles cumprem as condições de base, e não temos nada a fazer. Se $v_1, \cdots, v_n \in V$ são linearmente dependentes, então existe uma combinação linear nula

$$c_1 v_1 + \cdots + c_n v_n = 0_V$$

com os coeficientes c_i não todos nulos. Digamos que $c_n \neq 0$. Desse modo, temos que

$$v_n = -\frac{c_1}{c_n} v_1 - \cdots - \frac{c_{n-1}}{c_n} v_{n-1}$$

Assim, os elementos $v_1, \cdots, v_{n-1}$ ainda geram V. Se $\{ v_1, \cdots, v_{n-1} \}$ for linearmente dependente, repetimos o processo anterior e extraímos o elemento, digamos v_{n-1}, que é uma combinação linear dos outros. Repetindo esse processo um número finito de vezes, obtemos um subconjunto de $\{ v_1, \cdots, v_n \}$ formado com m elementos linearmente independentes $\{ v_1, \cdots, v_m \}$ que ainda geram V, com $m < n$. Assim, obtemos uma base para o espaço vetorial V. ■

Teorema 3.6.2 *Seja V um espaço vetorial gerado por um conjunto finito de elementos $v_1\,, \cdots\,, v_n \in V$. Então, todo conjunto linearmente independente de V é finito e contém no máximo n elementos.*

Demonstração – Para provar o teorema, basta mostrar que todo subconjunto W de V que contém mais de n elementos é linearmente dependente. Seja W um tal conjunto. Em W existem elementos distintos $w_1, \cdots, w_m$, com $m > n$. Como os elementos $v_1, \cdots, v_n$ geram V, existem escalares c_{ij} tais que

$$w_j = \sum_{i=1}^{n} c_{ij}\, v_i \qquad ; \qquad j = 1, \cdots, m$$

Consideramos agora uma combinação linear dos elementos de W, isto é,

$$\alpha_1 w_1 + \cdots + \alpha_m w_m = \sum_{j=1}^{m} \alpha_j \sum_{i=1}^{n} c_{ij}\, v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} c_{ij}\, \alpha_j \right) v_i$$

Como $m > n$, podemos encontrar escalares $\alpha_1, \cdots, \alpha_m$, não todos nulos, solução do sistema linear homogêneo

$$\sum_{j=1}^{m} c_{ij}\, \alpha_j = 0 \qquad ; \qquad i = 1, \cdots, n\,.$$

Logo, $\alpha_1 w_1 + \cdots + \alpha_m w_m = 0_V$ com algum $\alpha_i \neq 0$. Portanto, mostramos que W é um conjunto linearmente dependente em V. ■

Definição 3.6.2 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Dizemos que V é um espaço vetorial de* ***dimensão finita*** *se V possui uma base finita.*

Exemplo 3.6.3 *Vamos denotar por $\mathcal{P}(\mathbb{R})$ o conjunto de todos os polinômios reais, com a operação usual de adição de polinômios e a operação usual de multiplicação de um polinômio por um escalar. Assim, podemos mostrar facilmente que $\mathcal{P}(\mathbb{R})$ é um espaço vetorial real. Entretanto, $\mathcal{P}(\mathbb{R})$ não possui uma base finita.*

De fato, dado um conjunto $S = \{\, p_1(x), \cdots, p_n(x) \,\}$, considerando que o elemento $p_n(x)$ seja o polinômio de maior grau em S. Desse modo, se $\mathcal{P}(\mathbb{R}) = [S]$, qualquer elemento $p(x) \in \mathcal{P}(\mathbb{R})$ pode ser escrito como

$$p(x) = c_1 p_1(x) + c_2 p_2(x) + \cdots + c_n p_n(x)\,,$$

e seu grau é sempre menor ou igual ao grau de $p_n(x)$. Entretanto, como $\mathcal{P}(\mathbb{R})$ contém todos os polinômios reais, sabemos que existem polinômios em $\mathcal{P}(\mathbb{R})$ de grau maior que o grau de $p_n(x)$. Portanto, $\mathcal{P}(\mathbb{R}) \neq [S]$ para todo conjunto finito $S \subset \mathcal{P}(\mathbb{R})$.

O resultado a seguir, será necessário para que possamos apresentar uma definição algébrica adequada de dimensão de um espaço vetorial.

Corolário 3.6.1 *Seja V um espaço vetorial de dimensão finita. Então, quaisquer duas bases de V têm o mesmo número (finito) de elementos.*

Demonstração – Vamos supor que

$$\beta = \{ v_1, \cdots, v_n \} \qquad \text{e} \qquad \gamma = \{ w_1, \cdots, w_m \}$$

sejam duas bases finitas para V.

Como $\beta = \{ v_1, \cdots, v_n \}$ gera V e $\gamma = \{ w_1, \cdots, w_m \}$ é linearmente independente em V, pelo Teorema 3.6.2 temos que $m \leq n$.

Por outro lado, como $\gamma = \{ w_1, \cdots, w_m \}$ gera V e $\beta = \{ v_1, \cdots, v_n \}$ é linearmente independente em V, pelo Teorema 3.6.2 temos que $n \leq m$. Portanto, mostramos que $m = n$, o que completa a demonstração. ■

Exemplo 3.6.4 *Considere o espaço vetorial $\mathbb{R}^2$. Podemos verificar facilmente que os conjuntos $\beta = \{ (1,0), (0,1) \}$ e $\Gamma = \{ (1,1), (1,-1) \}$ são duas bases para o $\mathbb{R}^2$.*

Definição 3.6.3 *Seja V um espaço vetorial de dimensão finita, que possui uma base com n elementos. A **dimensão** de V é definida como sendo o número de elementos de uma base de V. Indicaremos a dimensão do espaço vetorial V por $\dim(V) = n$. No caso em que $V = \{ 0_V \}$, temos que o conjunto vazio, $\emptyset \subset V$, é uma base de V e dizemos que o espaço vetorial V tem dimensão nula.*

Exemplo 3.6.5 *Considere o espaço vetorial real $\mathcal{P}_n(\mathbb{R})$. Temos que o conjunto*

$$\beta = \{ 1, x, x^2, \cdots, x^n \}$$

é uma base para $\mathcal{P}_n(\mathbb{R})$, denominada base canônica. Assim, $\dim(\mathcal{P}_n(\mathbb{R})) = (n + 1)$.

Exemplo 3.6.6 *Considere o espaço vetorial real $\mathbb{M}_2(\mathbb{R})$. Temos que o conjunto*

$$\beta = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é uma base para $\mathbb{M}_2(\mathbb{R})$. Desse modo, $\dim(\mathbb{M}_2(\mathbb{R})) = 4$.

Corolário 3.6.2 *Seja V um espaço vetorial de dimensão n. Então,*

(a) Todo conjunto de elementos em V que contém mais de n elementos é LD.

(b) Nenhum conjunto contendo menos de n elementos pode gerar V.

Demonstração – A prova pode ficar a cargo do leitor. □

Lema 3.6.1 *Seja S um subconjunto linearmente independente de um espaço vetorial V. Seja u um elemento em V que não esteja no subespaço gerado por S. Então, o conjunto obtido acrescendo-se o elemento u a S é linearmente independente.*

Demonstração – Suponhamos que $v_1, \cdots, v_n$ sejam elementos distintos de S e que

$$c_1 v_1 + \cdots + c_n v_n + \alpha u = 0_V .$$

Então, $\alpha = 0$. Caso contrário

$$u = -\frac{c_1}{\alpha} v_1 - \cdots - -\frac{c_n}{\alpha} v_n$$

e u estaria no subespaço gerado por S. Assim, tem–se que

$$c_1 v_1 + \cdots + c_n v_n = 0_V$$

e como S é um conjunto linearmente independente, cada escalar $c_i = 0$, o que completa a demonstração. ■

Teorema 3.6.3 (Completamento) *Seja S um subconjunto linearmente independente de elementos de um espaço vetorial V de dimensão finita. Então, S pode ser completado de modo a formar uma base para V.*

Demonstração – Considere $dim(V) = n$ e $S = \{ v_1, \cdots, v_r \}$ linearmente independente. Note que pelo Teorema 3.6.2, temos que $r \leq n$. Se $V = [v_1, \cdots, v_r]$, então S é uma base de V e não temos nada a fazer.

Se existe $v_{r+1} \in V$ tal que $v_{r+1} \notin [v_1, \cdots, v_r]$, então pelo Lema 3.6.1 temos que $[v_1, \cdots, v_r, v_{r+1}]$ é linearmente independente. Se $V = [v_1, \cdots, v_r, v_{r+1}]$, então $\{ v_1, \cdots, v_r, v_{r+1} \}$ é uma base de V e não temos nada a fazer. Caso contrário, existe $v_{r+2} \in V$ tal que $v_{r+2} \notin [v_1, \cdots, v_r, v_{r+2}]$ e, então $[v_1, \cdots, v_r, v_{r+1}, v_{r+2}]$ é linearmente independente. Se $V = [v_1, \cdots, v_r, v_{r+1}, v_{r+2}]$ nossa prova está concluída. Caso contrário, prosseguimos com o mesmo processo. Como não podemos ter mais do que n elementos linearmente independente em V, após um número finito de passos teremos encontrado uma base para o espaço vetorial V que contém o conjunto original S, o que completa a demonstração. ■

Teorema 3.6.4 *Seja U um subespaço vetorial de um espaço vetorial V de dimensão finita. Então, todo subconjunto de U que é linearmente independente, é finito e é parte de uma base (finita) de U.*

Demonstração – Suponhamos que S_0 seja um subconjunto de U linearmente independente . Se S é um subconjunto de U linearmente independente contendo S_0, então S também é um subconjunto de V linearmente independente. Como V tem dimensão finita, S contém no máximo $dim(V)$ elementos. Portanto, existe um subconjunto S de U linearmente independente que é maximal e contém S_0. Como S é um subconjunto de U linearmente independente e maximal contendo S_0, o Lema 3.6.1 mostra que U é o subespaço gerado por S. Logo, S é uma base de U e o conjunto original S_0 é parte de uma base de U, o que completa a demonstração. ■

Corolário 3.6.3 *Seja U um subespaço vetorial próprio de um espaço vetorial V de dimensão finita. Então, U é de dimensão finita e $dim(U) < dim(V)$.*

Demonstração – Suponhamos que U contém um elemento $u \neq 0_V$. Pelo Teorema 3.6.4 e sua demonstração, existe uma base para U que contém o elemento u e no máximo $dim(V)$ elementos. Logo, U é de dimensão finita e $dim(U) \leq dim(V)$. Como U é um subespaço próprio, existe um elemento u_1 em V que não está em U. Acrescentando o elemento u_1 a uma base de U obtemos um subconjunto de V linearmente independente. Portanto, provamos que $dim(U) < dim(V)$. ■

Corolário 3.6.4 *Num espaço vetorial V de dimensão finita todo conjunto linearmente independente, não vazio, é parte de uma base de V.*

Teorema 3.6.5 *Sejam U e W subespaços de dimensão finita de um espaço vetorial V. Então, o subespaço $U + W$ é de dimensão finita e tem–se que*

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W).$$

Demonstração – Pelo Teorema 3.6.4 e seus Corolários, temos que o subespaço $U \cap W$ possui uma base finita $\{ v_1, \cdots, v_r \}$ que é parte de uma base

$$\{ v_1, \cdots, v_r, u_1, \cdots, u_m \}$$

do subespaço U, e parte de uma base

$$\{ v_1, \cdots, v_r, w_1, \cdots, w_n \}$$

do subespaço W.

O subespaço $U + W$ é gerado pelos elementos do conjunto

$$\{ v_1, \cdots, v_r, u_1, \cdots, u_m, w_1, \cdots, w_n \}$$

que é um conjunto linearmente independente em V.

De fato, considere a combinação linear nula

$$\sum_{i=1}^{r} a_i v_i + \sum_{j=1}^{m} b_j u_j + \sum_{k=1}^{n} c_k w_k = 0_V$$

Então, tem–se que

$$-\sum_{k=1}^{n} c_k w_k = \sum_{i=1}^{r} a_i v_i + \sum_{j=1}^{m} b_j u_j$$

o que mostra que o elemento

$$\widehat{u} = \sum_{k=1}^{n} c_k w_k \in U$$

Como o elemento $\widehat{u}$ também pertence a W, isto é, $\widehat{u} \in U \cap W$, temos que

$$\sum_{k=1}^{n} c_k w_k = \sum_{i=1}^{r} \alpha_i v_i \iff \sum_{k=1}^{n} c_k w_k - \sum_{i=1}^{r} \alpha_i v_i = 0_V$$

para certos escalares $\alpha_1, \cdots, \alpha_r$. Como o conjunto

$$\{ v_1, \cdots, v_r, w_1, \cdots, w_n \}$$

é linearmente independente, cada um dos escalares $c_k = 0$ e $\alpha_i = 0$. Assim,

$$\sum_{i=1}^{r} a_i v_i + \sum_{j=1}^{m} b_j u_j = 0_V .$$

Como o conjunto $\{ v_1, \cdots, v_r, u_1, \cdots, u_m \}$ é linearmente independente, temos que cada um dos escalares $a_i = 0$ e $b_j = 0$. Assim, mostramos que o conjunto

$$\{ v_1, \cdots, v_r, u_1, \cdots, u_m, w_1, \cdots, w_n \}$$

é uma base para o subespaço vetorial $U + W$.

Finalmente, temos que

$$\begin{aligned} dim(U) + dim(W) &= (r + m) + (r + n) \\ &= r + (r + m + n) \\ &= dim(U \cap W) + dim(U + W) , \end{aligned}$$

o que completa a demonstração. ■

Proposição 3.6.1 *Sejam V um espaço vetorial de dimensão finita e W um subespaço de V. Então, existe um subespaço U de V tal que $V = U \oplus W$.*

Demonstração – Seja $\{ w_1 , \cdots , w_m \}$ uma base para o subespaço W. Pelo Teorema do completamento, podemos completar a base de W para obter uma base de V, isto é,

$$\{ w_1 , \cdots , w_m , w_{m+1} , \cdots , w_n \}$$

é uma base para V.

Assim, vamos considerar U o subespaço gerado pelo conjunto linearmente independente $\{ w_{m+1} , \cdots , w_n \}$, que satisfaz as propriedades desejadas.

De fato, é evidente que $V = U + W$, pois o conjunto

$$\{ w_1 , \cdots , w_m , w_{m+1} , \cdots , w_n \}$$

é linearmente independente e gera V. Por outro lado, como

$$W = [w_1 , \cdots , w_m] \qquad , \qquad U = [w_{m+1} , \cdots , w_n]$$

e o conjunto $\{ w_1 , \cdots , w_m , w_{m+1} , \cdots , w_n \}$ é linearmente independente, temos que $W \cap U = \{ 0_V \}$, o que completa a demonstração. ■

Proposição 3.6.2 *Sejam V um espaço vetorial de dimensão finita e W um subespaço de V. Se $dim(W) = dim(V)$, então $W = V$.*

Demonstração – Sabemos que W tem uma base. Toda base de W também é uma base de V devido ao fato que $dim(W) = dim(V)$. Logo, todo elemento de V também pertence a W. Assim, $V \subset W$ e, como $W \subset V$, segue–se que $V = W$. ■

Exemplo 3.6.7 *Sejam V um espaço vetorial real, com $dim(V) = 9$, U e W subespaços vetoriais de V tais que $dim(U) = 6$ e $dim(W) = 5$. Mostre que*

$$2 \leq dim(U \cap W) \leq 5 .$$

Considerando o resultado do Teorema 3.6.5, temos que

$$6 \leq dim(U + W) = 6 + 5 - dim(U \cap W) \leq 9$$

Assim, mostramos que $2 \leq dim(U \cap W) \leq 5$.

Exemplo 3.6.8 *O conjunto* $S = \{ v_1, v_2, v_3 \}$ *onde*

$$v_1 = (1, 0, -1) \quad , \quad v_2 = (1, 2, 1) \quad e \quad v_3 = (0, -3, 2) \, ,$$

é uma base para o espaço vetorial real $I\!R^3$.

Devemos mostrar que S é linearmente independente e que gera o espaço vetorial $I\!R^3$.

Inicialmente, vamos mostrar que o conjunto S é linearmente independente. Considerando a combinação linear nula

$$av_1 \; + \; bv_2 \; + \; cv_3 \; = \; 0_{I\!R^3} \; ,$$

obtemos o seguinte sistema linear homogêneo

$$\begin{cases} \;\; a + b = 0 \\ \;\; 2b - 3c = 0 \\ -a + b + 2c = 0 \end{cases} \iff \begin{cases} a + b = 0 \\ 2b - 3c = 0 \\ 5c = 0 \end{cases}$$

Assim, o sistema linear homogêneo possui somente a solução trivial, isto é,

$$a \; = \; b \; = \; c \; = \; 0 \; ,$$

provando que o conjunto S é linearmente independente em $I\!R^3$.

Finalmente, vamos mostrar que S gera o espaço vetorial $I\!R^3$, isto é, que todo elemento $(x, y, z) \in I\!R^3$ é escrito como uma combinação linear dos elemento do conjunto S

$$(x, y, z) \; = \; a(1, 0, -1) \; + \; b(1, 2, 1) \; + \; c(0, -3, 2) \, .$$

Desse modo, obtemos o seguinte sistema linear

$$\begin{cases} \;\; a + b = x \\ \;\; 2b - 3c = y \\ -a + b + 2c = z \end{cases} \iff \begin{cases} a + b = x \\ 2b - 3c = y \\ 5c = z + x - y \end{cases}$$

Portanto, podemos obter de modo único os coeficientes da combinação linear, a, b e c, em função das componentes do elemento $(x, y, z) \in I\!R^3$. Assim, provamos que $I\!R^3 \; = \; [S]$. Logo, o conjunto S é uma base para o espaço vetorial $I\!R^3$.

Exemplo 3.6.9 *Seja* $I\!M_2(I\!R)$ *o espaço vetorial das matrizes reais de ordem* 2. *Mostre que a* $dim(\,I\!M_2(I\!R)\,) = 4$ *exibindo uma base* $\beta = \{\, A_1,\, A_2,\, A_3,\, A_4 \,\}$.

Podemos mostrar facilmente que tomando as matrizes

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad , \quad A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{e} \quad A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

o conjunto $\beta = \{\, A_1,\, A_2,\, A_3,\, A_4 \,\}$ é uma base para o espaço vetorial $I\!M_2(I\!R)$. Assim, temos que $dim(\,I\!M_2(I\!R)\,) = 4$.

Exemplo 3.6.10 *Considerando* $\mathbb{C} = \{\, z = a + bi \;\;/\;\; a, b \in I\!R \,\}$ *como um espaço vetorial real, temos que o conjunto* $\beta = \{\, 1,\, i \,\}$ *é uma base de* $\mathbb{C}$.

Sabemos que todo elemento $z \in \mathbb{C}$ é escrito como uma combinação linear dos elementos do conjunto β, com coeficientes em $I\!R$. Além disso, o conjunto β é linearmente independente. De fato, considere a combinação linear nula $a + bi = 0 + 0i$ para $a, b \in I\!R$. Desse modo, temos que $a = b = 0$. Portanto, mostramos que o conjunto β é uma base para $\mathbb{C}$ como espaço vetorial real. Desse modo, temos que $dim(\,\mathbb{C}\,) = 2$.

Exemplo 3.6.11 *Considere* $\mathbb{C}$ *um espaço vetorial complexo. Qual é a dimensão de* $\mathbb{C}$*?*

Neste caso, podemos verificar facilmente que o conjunto $\beta = \{\, 1 \,\}$ é uma base para $\mathbb{C}$. Assim, temos que $dim(\,\mathbb{C}\,) = 1$.

Exemplo 3.6.12 *Considerando* $\mathbb{C}^2$ *como um espaço vetorial complexo, temos que o conjunto* $\beta = \{\, e_1,\, e_2 \,\}$, *onde* $e_1 = (1,0)$ *e* $e_2 = (0,1)$, *é uma base de* $\mathbb{C}^2$.

Podemos verificar facilmente que todo elemento de $\mathbb{C}^2$ é escrito como uma combinação linear dos elementos de β, com coeficientes em $\mathbb{C}$, isto é,

$$(z_1,\, z_2) = z_1 e_1 + z_2 e_2 \qquad ; \qquad z_1,\, z_2 \in \mathbb{C}\,.$$

Além disso, sabemos que β é linearmente independente. Assim, temos que β é uma base para $\mathbb{C}^2$ como espaço vetorial complexo. Logo, temos que $dim(\,\mathbb{C}^2\,) = 2$.

Exemplo 3.6.13 *Considerando* $\mathbb{C}^2$ *como um espaço vetorial real, exibir uma base.*

Neste caso, podemos verificar facilmente que o conjunto $\gamma = \{\, z_1,\, z_2,\, z_3,\, z_4 \,\}$, onde

$$z_1 = (1,0) \quad , \quad z_2 = (i,0) \quad , \quad z_3 = (0,1) \quad \text{e} \quad z_4 = (0,i)\;,$$

é linearmente independente e que todo elemento de $\mathbb{C}^2$ é escrito como uma combinação linear dos elementos de γ, com coeficientes em $I\!R$. Assim, γ é uma base para $\mathbb{C}^2$. Logo, temos que $dim(\,\mathbb{C}^2\,) = 4$.

Exemplo 3.6.14 *Considerando* $\mathbb{C}^2$ *como um espaço vetorial complexo, temos que o conjunto* $S = \{ (1-i,\ i),\ (2,\ -1+i) \}$ *é linearmente dependente.*

Considerando a combinação linear nula

$$z_1(1-i,\, i) \;+\; z_2(2,\, -1+i) \;=\; (0,\, 0) \qquad ; \qquad z_1,\, z_2 \in \mathbb{C}$$

obtemos o seguinte sistema linear homogêneo

$$\left\{ \begin{array}{rcrcl} (1-i)z_1 & + & 2z_2 & = & 0 \\ iz_1 & + & (-1+i)z_2 & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcrcl} z_1 & + & (i+1)z_2 & = & 0 \\ iz_1 & + & (-1+i)z_2 & = & 0 \end{array} \right.$$

Note que o segundo sistema linear homogêneo foi obtido somando a segunda equação à primeira equação.

Tomando o segundo sistema linear homogêneo, multiplicamos a primeira equação por $-i$ e somamos à segunda equação. Assim, ficamos somente com a primeira equação

$$z_1 \;+\; (i+1)z_2 \;=\; 0$$

que possui infinitas soluções não–nulas, $z_1 = -(i+1)z_2$, $z_2 \in \mathbb{C}$. Portanto, o conjunto S é linearmente dependente.

Exemplo 3.6.15 *Considerando* $\mathbb{C}^2$ *como um espaço vetorial real, temos que o conjunto* $S = \{ (1-i,\, i),\, (2,\, -1+i) \}$ *é linearmente independente.*

Considerando a combinação linear nula

$$a(1-i,\, i) \;+\; b(2,\, -1+i) \;=\; (0,\, 0) \qquad ; \qquad a,\, b \in I\!R$$

obtemos o seguinte sistema linear homogêneo

$$\left\{ \begin{array}{rcrcl} (1-i)a & + & 2b & = & 0 \\ ia & + & (-1+i)b & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcrcl} a & + & (i+1)b & = & 0 \\ ia & + & (-1+i)b & = & 0 \end{array} \right.$$

Como $a,\, b \in I\!R$, temos que o sistema linear homogêneo possui somente a solução trivial, $a = b = 0$. Desse modo, mostramos que o conjunto S é linearmente independente.

Exemplo 3.6.16 *Considere a seguinte Equação Diferencial Ordinária (EDO)*

$$u''(x) \,+\, u(x) \,=\, 0\,.$$

Sabemos que as funções $u_1(x) = \sin(x)$ *e* $u_2 = \cos(x)$ *são duas soluções linearmente independentes da EDO, isto é, as funções* $u_1(x)$ *e* $u_2(x)$ *satisfazem a EDO e o conjunto* $\Gamma = \{\, u_1(x),\, u_2(x)\,\}$ *é linearmente independente para* $x \in I\!R$. *Além disso, temos que toda solução da EDO é uma combinação linear dessas duas funções. Desse modo, o conjunto* Γ *é uma base para o espaço solução da EDO.*

Exemplo 3.6.17 *O conjunto* $\Gamma = \{\,(1,2),\,(1,-1)\,\}$ *é uma base para o* $I\!R^2$.

Podemos verificar facilmente que Γ é linearmente independente, pois cada um de seus elementos não é múltiplo escalar do outro. Dado um elemento $u = (x,y) \in I\!R^2$, vamos mostrar que u pode ser escrito de modo único como $u = a(1,2) + b(1,-1)$ para $a,\, b \in I\!R$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{cases} a + b = x \\ 2a - b = y \end{cases} \iff \begin{cases} a + b = x \\ 3a = x + y \end{cases}$$

Assim, temos que

$$a = \frac{x \,+\, y}{3} \qquad \text{e} \qquad b = \frac{2x \,-\, y}{3}\,,$$

obtidos de modo único, em função das componentes do elemento $u \in I\!R^2$.

Exemplo 3.6.18 *Considere o espaço vetorial real* $I\!R^5$. *Determine uma base para o subespaço vetorial de* $I\!R^5$ *dado por:*

$$W = \{\, (x_1,\, x_2,\, x_3,\, x_4,\, x_5) \in I\!R^5 \;/\; x_1 + x_3 + x_5 = 0 \quad e \quad x_2 = x_4 \,\}\,.$$

Temos que $x_1 = -x_3 - x_5$ e $x_2 = x_4$. Assim, temos que os elementos

$$\begin{aligned} w_1 &= (-1,0,1,0,0) \\ w_2 &= (-1,0,0,0,1) \\ w_3 &= (0,1,0,1,0) \end{aligned}$$

formam uma base para o subespaço W. Desse modo, $dim(W) = 3$.

Exemplo 3.6.19 *Considere o espaço vetorial real* $\mathcal{P}_3(\mathbb{R})$. *Determine uma base para o subespaço vetorial de* $\mathcal{P}_3(\mathbb{R})$ *dado por:*

$$S = \{ \, p(x) \in \mathcal{P}_3(\mathbb{R}) \; / \; p(-1) = 0 \quad e \quad p'(1) = 0 \, \} .$$

Vamos tomar um elemento $p(x) \in \mathcal{P}_3(\mathbb{R})$ escrito da seguinte forma:

$$p(x) = a + bx + cx^2 + dx^3$$

Impondo as restrições $p(-1) = p'(1) = 0$, obtemos as seguintes equações algébricas

$$\begin{aligned} p(-1) &= a - b + c - d = 0 \\ p'(1) &= b + 2c + 3d = 0 \end{aligned}$$

Podemos observar que o sistema linear homogêneo possui dois graus de liberdade. Assim, $dim(S) = 2$. Desse modo, temos que

$$\begin{aligned} a &= -3c - 2d \\ b &= -2c - 3d \end{aligned}$$

Assim, todo elemento $p(x) \in S$ é escrito da seguinte forma:

$$p(x) = c(-3 - 2x + x^2) + d(-2 - 3x + x^3) \qquad , \qquad c, d \in \mathbb{R}$$

Portanto, mostramos que o subespaço S é gerado pelos elementos do conjunto

$$\Gamma = \{ \, -3 - 2x + x^2, \, -2 - 3x + x^3 \, \} ,$$

que é linearmente independente em $\mathcal{P}_3(\mathbb{R})$, pois os elementos não são múltiplos escalares. Logo, como $dim(S) = 2$, o conjunto Γ é uma base para o subespaço S.

Note que os elementos da base Γ satisfazem as condições para que um elemento do espaço vetorial $\mathcal{P}_3(\mathbb{R})$ pertença ao subespaço S.

Exemplo 3.6.20 *Considere o espaço vetorial real* $\mathcal{P}_3(\mathbb{R})$. *Os polinômios*

$$\begin{aligned} p(x) &= 1 - 2x^2 + x^3 \\ q(x) &= 3 - x + 4x^2 \\ r(x) &= -2 + 3x \\ s(x) &= x - 3x^3 \end{aligned}$$

formam uma base para o espaço vetorial $\mathcal{P}_3(\mathbb{R})$ *? Justifique sua resposta.*

Considerando que $dim(\mathcal{P}_3(\mathbb{R})) = 4$, basta verificar se o conjunto

$$\{ p(x), q(x), r(x), s(x) \}$$

é linearmente independente em $\mathcal{P}_3(\mathbb{R})$. Tomando a combinação linear nula

$$a(1 - 2x^2 + x^3) + b(3 - x + 4x^2) + c(-2 + 3x) + d(x - 3x^3) = 0$$

e organizando os termos de mesma potência, obtemos

$$(a + 3b - 2c) + (-b + 3c + d)x + (-2a + 4b)x^2 + (a - 3d)x^3 = 0 .$$

Desse modo, temos um sistema linear homogêneo nas incógnitas a, b, c, d cuja matriz é dada por:

$$A = \begin{bmatrix} 1 & 3 & -2 & 0 \\ 0 & -1 & 3 & 1 \\ -2 & 4 & 0 & 0 \\ 1 & 0 & 0 & -3 \end{bmatrix}.$$

Agora temos que analisar o tipo do conjunto solução do sistema linear homogêneo através do posto da matriz A. Efetuando o escalonamento na matriz A encontramos uma matriz $\widehat{A}$ na forma escalonada, linha equivalente a matriz A, dada por:

$$\widehat{A} = \begin{bmatrix} 1 & 3 & -2 & 0 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & 26 & 10 \\ 0 & 0 & 0 & -86 \end{bmatrix}.$$

Como o $posto(A) = posto(\widehat{A}) = 4$, veja Definição 2.6.3, o sistema linear homogêneo possui somente a solução trivial $a = b = c = d = 0$. Logo, o conjunto

$$\{ p(x), q(x), r(x), s(x) \}$$

é uma base para o espaço vetorial $\mathcal{P}_3(\mathbb{R})$.

Exemplo 3.6.21 *Seja* V *um espaço vetorial de dimensão finita com* $dim(V) = n$. *Se* U *e* W *são subespaços vetoriais de* V *com* $dim(U) > \dfrac{n}{2}$ *e* $dim(W) > \dfrac{n}{2}$. *Mostre que* $U \cap W \neq \{ 0_V \}$.

A prova é feita por absurdo, isto é, vamos negar a tese e obter uma contradição. Supondo que $U \cap W = \{ 0_V \}$, isto é, $dim(U \cap W) = 0$. Pelo Teorema 3.6.5 temos que

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W) = dim(U) + dim(W) > n$$

que é uma contradição, pois $U + W$ é um subespaço vetorial de V e temos que $dim(V) = n$. Portanto, $U \cap W \neq \{ 0_V \}$.

Exemplo 3.6.22 *Considere o subespaço vetorial* $S = \{ p(x) \in \mathcal{P}_2(\mathbb{R}) \ / \ p(1) = 0 \}$ *do espaço vetorial* $\mathcal{P}_2(\mathbb{R})$. *Determine uma base para* S.

Considerando $\mathcal{P}_2(\mathbb{R})$ com a base canônica $\{ 1, x, x^2 \}$, temos que $p(x) \in \mathcal{P}_2(\mathbb{R})$ é escrito de modo único como $p(x) = a + bx + cx^2$. Impondo a condição $p(1) = 0$, para que $p \in S$, obtemos a seguinte equação algébrica

$$p(1) = a + b + c = 0 .$$

Assim, temos que $a = -b - c$. Podemos concluir que $dim(S) = 2$, pois temos dois graus de liberdade. Logo, $p(x) \in S$ é escrito da seguinte forma:

$$p(x) = (-b - c) + bx + cx^2 = b(x - 1) + c(x^2 - 1) .$$

Portanto, $\{ (x - 1), (x^2 - 1) \}$ é uma base para o subespaço S.

Exemplo 3.6.23 *Considere o seguinte subespaço*

$$S = \{ p(x) \in \mathcal{P}_2(\mathbb{R}) \ / \ p(-1) = 0 \quad e \quad p(1) = 0 \}$$

do espaço vetorial $\mathcal{P}_2(\mathbb{R})$. *Temos que* $dim(S) = 1$ *e o conjunto* $\Gamma = \{ x^2 - 1 \}$ *é uma base para o subespaço* S.

Note que o elemento da base Γ satisfaz as condições para que um elemento do espaço vetorial $\mathcal{P}_2(\mathbb{R})$ pertença ao subespaço S.

Exemplo 3.6.24 *Considere o seguinte subespaço*

$$S = \{ \, p(x) \in \mathcal{P}_3(I\!R) \;\; / \;\; p'(-1) = 0 \quad e \quad p(1) = 0 \, \}$$

do espaço vetorial $\mathcal{P}_3(I\!R)$. *Encontre uma base para* S.

Consideramos um elemento genérico $p(x) = a + bx + cx^2 + dx^3 \in \mathcal{P}_3(I\!R)$ e vamos impor as condições para que esse elemento pertença ao subespaço S, isto é,

$$\begin{array}{rcccccccccc} p'(-1) & = & & & b & - & 2c & + & 3d & = & 0 \\ p(1) & = & a & + & b & + & c & + & d & = & 0 \end{array} .$$

Assim, obtemos um sistema linear homogêneo com dois graus de liberdade. Desse modo, podemos concluir que o subespaço S tem dimensão dois. Logo, temos uma relação entre os coeficientes dos elementos $p(x) \in S$. Podemos verificar facilmente que

$$b = 2c - 3d \qquad \text{e} \qquad a = -3c + 2d$$

para $c, d \in I\!R$. Substituindo a e b no polinômio $p(x)$, obtemos que todo elemento do subespaço S é escrito como:

$$p(x) = c(-3 + 2x + x^2) + d(2 - 3x + x^3) \qquad ; \qquad c, d \in I\!R \, .$$

Portanto, mostramos que o subespaço S é gerando pelo elementos do conjunto

$$\Gamma = \{ -3 + 2x + x^2 \, , \, 2 - 3x + x^3 \} \, ,$$

que é linearmente independente em $\mathcal{P}_3(I\!R)$. Logo, o conjunto Γ é uma base para o subespaço S.

Note que os elementos da base Γ satisfazem as condições para que um elemento do espaço vetorial $\mathcal{P}_3(I\!R)$ pertença ao subespaço S.

Exemplo 3.6.25 *Considere o seguinte subespaço*

$$S = \left\{ \, p(x) \in \mathcal{P}_2(I\!R) \;\; / \;\; \int_0^1 p(x)dx = 0 \, \right\}$$

do espaço vetorial $\mathcal{P}_2(I\!R)$. *Temos que* $dim(\, S \,) = 2$ *e o conjunto*

$$\Gamma = \left\{ \, -\frac{1}{2} + x \, , \, -\frac{1}{3} + x^2 \, \right\}$$

é uma base para o subespaço S.

Note que os elementos da base Γ satisfazem as condições para que um elemento do espaço vetorial $\mathcal{P}_2(I\!R)$ pertença ao subespaço S.

Estratégia para Completamento de Base

Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $\beta = \{ v_1 , \cdots , v_n \}$ uma base de V. Sabemos que todo elemento $u \in V$ pode ser representado pela combinação linear

$$u = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n ,$$

onde os escalares $c_i \in \mathbb{F}$ para $i = 1, \cdots , n$, são as coordenadas do elemento u com relação à base β.

Pelo resultado do Teorema 3.6.3, sabemos que todo conjunto linearmente independente em V pode ser completado até formar uma base para V. Vamos mostrar uma maneira prática, e eficiente, para determinar quais elementos que serão necessários para completar esse conjunto para formar uma base de V. Os resultados que vamos apresentar dependem fortemente dos resultados das seções 2.6 e 2.9.

Para isso, vamos considerar um conjunto $S = \{ w_1 , \cdots , w_m \}$ de elemento de V, com $m \leq n$. Sabemos que cada elemento $w_j \in S$ pode ser representado pela combinação linear

$$w_i = a_{i1} v_1 + a_{i2} v_2 + \cdots + a_{in} v_n ,$$

para $i = 1, \cdots , m$. A seguir, construímos uma matriz $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ cujas linhas são as coordenadas dos elementos do conjunto S em relação à base β, que é dada por:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} .$$

Em seguida, obtemos a matriz $\widehat{A}$ na forma escalonada, linha equivalente a matriz A, que representamos por:

$$\widehat{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \alpha_{24} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & \alpha_{ii} & \cdots & \alpha_{in} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \alpha_{mn} \end{bmatrix} .$$

As linhas da matriz $\widehat{A}$ são combinações lineares das linhas da matriz A, pois foram obtidas através de operações elementares de linhas da matriz A. Assim, podemos verificar facilmente que os elementos $u_1, \cdots, u_m$ representados da forma:

$$u_i = \sum_{j=i}^{n} \alpha_{ij}\, v_j \ ,$$

para $i = 1, \cdots, m$, são combinações lineares dos elementos $w_1, \cdots, w_m$.

Portanto, os subespaços vetoriais W e U do espaço vetorial V, gerados pelos conjuntos

$$\{ w_1 , \cdots , w_m \} \qquad \text{e} \qquad \{ u_1 , \cdots , u_m \} \ ,$$

respectivamente, são os mesmos.

Finalmente, temos que fazer uma análise do $posto(A) = posto(\widehat{A})$, com as seguintes possibilidades.

(1) No caso em que $posto(A) = posto(\widehat{A}) = m = n$, todas as linhas da matriz $\widehat{A}$ são não–nulas. Assim, temos que os conjuntos

$$\{ w_1 , \cdots , w_m \} \qquad \text{e} \qquad \{ u_1 , \cdots , u_m \}$$

são linearmente independentes, e formam uma base para V.

(2) No caso em que $posto(A) = posto(\widehat{A}) = m < n$, temos que os conjuntos

$$\{ w_1 , \cdots , w_m \} \qquad \text{e} \qquad \{ u_1 , \cdots , u_m \}$$

são linearmente independentes.

Construímos uma matriz $M \in I\!M_n(I\!R)$ na forma escalonada a partir da matriz $\widehat{A}$, acrescentando linhas que são as coordenadas de elementos da base canônica de $I\!F^n$, escolhidos de modo conveniente. Portanto, os elementos de V cujas coordenadas em relação à base β são dadas pelas linhas da matriz M formam um conjunto linearmente independente em V, e conseqüentemente, uma base para V.

É importante observar que estamos completando o conjunto

$$\{ u_1 , \cdots , u_m \}$$

com elementos da base β, para formar uma nova base para V.

(3) No caso em que $posto(A) = posto(\widehat{A}) = r < m$, as $m - r$ últimas linhas da matriz $\widehat{A}$ são nulas. Assim, temos que o conjunto

$$\{ u_1 , \cdots , u_r \}$$

é linearmente independente, e forma uma base para o subespaço $W = [w_1, \cdots, w_m]$.

De maneira análoga, construímos uma matriz $M \in \mathbb{M}_n(\mathbb{R})$ na forma escalonada a partir das linhas não–nulas da matriz $\widehat{A}$, acrescentando linhas que são as coordenadas de elementos da base canônica de $\mathbb{F}^n$, escolhidos de modo conveniente.

Desse modo, os elementos de V cujas coordenadas em relação à base β são dadas pelas linhas da matriz M formam um conjunto linearmente independente em V, e portanto, uma base para V.

Para exemplificar, considere $n = 6$, $m = 4$ e $posto(A) = posto(\widehat{A}) = 3 < m$. Assim sendo, a matriz $\widehat{A} \in \mathbb{M}_{4\times6}(\mathbb{R})$ é dada por:

$$\widehat{A} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} & \alpha_{16} \\ 0 & \alpha_{22} & \alpha_{23} & \alpha_{24} & \alpha_{25} & \alpha_{26} \\ 0 & 0 & \alpha_{33} & \alpha_{34} & \alpha_{35} & \alpha_{36} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Desse modo, a matriz $M \in \mathbb{M}_6(\mathbb{R})$ será construída da seguinte forma:

$$M = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} & \alpha_{15} & \alpha_{16} \\ 0 & \alpha_{22} & \alpha_{23} & \alpha_{24} & \alpha_{25} & \alpha_{26} \\ 0 & 0 & \alpha_{33} & \alpha_{34} & \alpha_{35} & \alpha_{36} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} ,$$

e portanto, o conjunto $\{ u_1, u_2, u_3, v_4, v_5, v_6 \}$ é uma nova base para o espaço vetorial V, onde os elementos u_1, u_2, u_3 são dados por:

$$u_i = \sum_{j=i}^{6} \alpha_{ij}\, v_j \ ,$$

para $i = 1, \cdots, 3$, e formam uma base para o subespaço $W = [w_1, \cdots, w_4]$.

A seguir apresentamos vários exemplos, que mostram como verificar se um conjunto é uma base, bem como a aplicação da estratégia de completamento de base.

Exemplo 3.6.26 *Considere os seguintes subespaços de* $I\!R^3$

$$U = \{ (x,y,z) \in I\!R^3 \ / \ x - 2y + 3z = 0 \}$$

$$W = \{ (x,y,z) \in I\!R^3 \ / \ x + y + z = 0 \}$$

Determine uma base para o subespaço $U + W$.

Do Exemplo 3.4.9, sabemos que o subespaço $U + W$ tem como um sistema de geradores os seguintes elementos

$$v_1 = (-1,0,1) \quad , \quad v_2 = (-1,1,0) \quad , \quad v_3 = (2,1,0) \quad \text{e} \quad v_4 = (-3,0,1)$$

dentre os quais vamos escolher uma base para o subespaço. Para isso, construímos uma matriz cujas linhas são as coordenadas dos elementos do sistema de geradores, em relação à base canônica, e efetuamos o escalonamento

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \\ 0 & 0 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Como a matriz tem posto igual a três, temos que $dim(U + W) = 3$. Assim, podemos escolher uma base para o subespaço $U + W$ um dos seguintes conjuntos

$$\{ (-1,0,1), (-1,1,0), (2,1,0) \} \qquad \text{ou} \qquad \{ (-1,0,1), (0,1,1), (0,0,3) \}$$

Note que qualquer base de $I\!R^3$ é uma base para o subespaço $U + W$.

Do Exemplo 3.4.6, sabemos que o subespaço $U \cap W = [(-5,2,3)]$. Desse modo, temos que $I\!R^3 = U + W$, entretanto, não como soma direta.

Exemplo 3.6.27 *Considere os seguintes subespaços de* $I\!\!R^3$

$$U = \{ (x, y, z) \in I\!\!R^3 \ / \ x = z \}$$

$$W = \{ (x, y, z) \in I\!\!R^3 \ / \ x + y + z = 0 \}$$

Determine uma base para o subespaço $U + W$.

Do Exemplo 3.4.5, sabemos que o subespaço $U + W$ tem como um sistema de geradores os seguintes elementos

$$v_1 = (1, 0, 1) \quad , \quad v_2 = (0, 1, 0) \quad , \quad v_3 = (-1, 1, 0) \quad \text{e} \quad v_4 = (-1, 0, 1)$$

dentre os quais vamos escolher uma base para o subespaço. Para isso, construímos uma matriz cujas linhas são as coordenadas dos elementos do sistema de geradores, em relação à base canônica, e efetuamos o escalonamento

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Como a matriz tem posto igual a três, temos que $dim(U + W) = 3$. Assim, podemos escolher como base para o subespaço $U + W$ um dos seguintes conjuntos

$$\{ (1, 0, 1), (0, 1, 0), (-1, 1, 0) \} \qquad \text{ou} \qquad \{ (1, 0, 1), (0, 1, 0), (0, 0, 1) \}$$

Do Exemplo 3.4.5, sabemos que o subespaço $U \cap W = [(1, -2, 1)]$. Desse modo, temos que $I\!\!R^3 = U + W$, entretanto, não como soma direta.

Exemplo 3.6.28 *Considere o espaço vetorial real* $\mathcal{P}_4(\mathbb{R})$. *Determine uma base para* $\mathcal{P}_4(\mathbb{R})$ *contendo elementos do conjunto* $S = \{ p(x), q(x), r(x) \}$, *onde*

$$\begin{aligned} p(x) &= 1 + x + x^2 + 3x^3 + 2x^4 \\ q(x) &= 1 + 2x + x^2 + 2x^3 + x^4 \\ r(x) &= 1 + 3x + 2x^2 + 1x^3 + 2x^4 \end{aligned}$$

Os elementos $p(x), q(x), r(x)$ estão escritos em relação à base canônica

$$\beta = \{ 1, x, x^2, x^3, x^4 \}.$$

Inicialmente, vamos construir uma matriz A cujas linhas são formadas pelas coordenadas dos elementos $p(x)$, $q(x)$ e $r(x)$ com relação à base canônica, respectivamente,

$$A = \begin{bmatrix} 1 & 1 & 1 & 3 & 2 \\ 1 & 2 & 1 & 2 & 1 \\ 1 & 3 & 2 & 1 & 2 \end{bmatrix}$$

Em seguida, efetuando o escalonamento na matriz A encontramos uma matriz $\widehat{A}$ na forma escalonada, linha equivalente a matriz A, dada por:

$$\widehat{A} = \begin{bmatrix} 1 & 1 & 1 & 3 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 2 \end{bmatrix}$$

da qual concluímos que o conjunto S é linearmente independente em $\mathcal{P}_4(\mathbb{R})$.

Finalmente, construímos uma matriz M de ordem 5×5 a partir da matriz $\widehat{A}$ na forma escalonada acrescentando linhas na matriz $\widehat{A}$ que são as coordenadas de elementos da base canônica escolhidos de modo conveniente

$$M = \begin{bmatrix} 1 & 1 & 1 & 3 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Assim, temos que $\gamma = \{ p(x), q(x), r(x), x^3, x^4 \}$ é uma base para o espaço $\mathcal{P}_4(\mathbb{R})$, contendo os elementos do conjunto S.

Exemplo 3.6.29 *Considere o espaço vetorial real* $\mathbb{M}_{3\times 2}(\mathbb{R})$. *Determine uma base para* $\mathbb{M}_{3\times 2}(\mathbb{R})$ *contendo elementos do conjunto* $S = \{ A_1, A_2, A_3 \}$, *onde*

$$A_1 = \begin{bmatrix} 1 & 1 \\ -1 & 3 \\ 3 & 2 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 3 & 2 \end{bmatrix} \quad e \quad A_3 = \begin{bmatrix} -1 & -3 \\ 4 & -2 \\ -3 & -2 \end{bmatrix}$$

Considere uma base Γ para o espaço vetorial $\mathbb{M}_{3\times 2}(\mathbb{R})$ dada por:

$$\Gamma = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

Sabemos que $dim(\mathbb{M}_{3\times 2}(\mathbb{R})) = 6$.

Inicialmente, vamos construir uma matriz A cujas linhas são formadas pelas coordenadas dos elementos A_1, A_2 e A_3 com relação à base Γ, respectivamente,

$$A = \begin{bmatrix} 1 & 1 & -1 & 3 & 3 & 2 \\ 1 & -1 & 2 & 4 & 3 & 2 \\ -1 & -3 & 4 & -2 & -3 & -2 \end{bmatrix}$$

Em seguida, efetuando o escalonamento na matriz A encontramos uma matriz $\widehat{A}$ na forma escalonada, linha equivalente a matriz A, dada por:

$$\widehat{A} = \begin{bmatrix} 1 & 1 & -1 & 3 & 3 & 2 \\ 0 & -2 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

da qual concluímos que o conjunto $\{ A_1, A_2 \}$ é linearmente independente em $\mathbb{M}_{3\times 2}(\mathbb{R})$, pois a matriz A_3 é uma combinação linear das matrizes A_1 e A_2.

Finalmente, construímos uma matriz M de ordem 6×6 a partir da matriz $\widehat{A}$ na forma escalonada acrescentando linhas na matriz $\widehat{A}$ que são as coordenadas de elementos da base Γ escolhidos de modo conveniente

$$M = \begin{bmatrix} 1 & 1 & -1 & 3 & 3 & 2 \\ 0 & -2 & 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Assim, temos que o conjunto

$$\Gamma' = \left\{ A_1 \ , \ A_2 \ , \ \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \ , \ \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \ , \ \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \ , \ \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é uma base para o espaço vetorial $\mathbb{M}_{3\times2}(\mathbb{R})$, contendo dois elementos do conjunto S.

Exemplo 3.6.30 *Considere o espaço vetorial real* $\mathbb{R}^4$*. Determine uma base para* $\mathbb{R}^4$ *contendo elementos do conjunto* $S = \{ v_1, v_2 \}$*, onde*

$$v_1 = (1, 0, -2, 2) \qquad e \qquad v_2 = (1, 2, -2, 1)\,.$$

Como nos exemplos anteriores, obtemos a matriz M de ordem 4×4 dada por:

$$M = \begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Assim, temos que o conjunto

$$\gamma = \{ (1,0,-2,2), (1,2,-2,1), (0,0,1,0), (0,0,0,1) \}$$

é uma base para o espaço vetorial $\mathbb{R}^4$, contendo os elementos do conjunto S.

Exemplo 3.6.31 *Considere o subespaço* $U = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x - y + 2z = 0 \}$ *do espaço vetorial* $\mathbb{R}^3$*. Determine um subespaço* W *de modo que* $\mathbb{R}^3 = U \oplus W$*.*

Inicialmente, vamos encontrar uma base para o subespaço U. Podemos escrever os elementos $u \in U$ da seguinte forma:

$$u = \alpha(1,1,0) + \beta(-2,0,1) \quad , \quad \alpha, \beta \in \mathbb{R}$$

Portanto, $\{ (1,1,0), (-2,0,1) \}$ é uma base para o subespaço U. Finalmente, vamos completar a base de U, para obter uma base para $\mathbb{R}^3$, com o elemento $(0,0,1)$. Assim, obtemos o subespaço $W = [(0,0,1)]$. Note que esse problema possui infinitas soluções, pois podemos completar a base com um outro elemento.

Exemplo 3.6.32 *Considere os seguintes subespaços do espaço vetorial real* $I\!R^4$

$$\begin{aligned} U &= [(1,0,1,0),\ (0,1,0,0)] \\ W &= \{\ (x,y,z,t) \in I\!R^4 \ /\ x + y = 0\ \} \end{aligned}$$

Determine $dim(U+W)$ *e* $dim(U \cap W)$.

Podemos verificar facilmente que os elementos $(1,0,1,0)$, $(0,1,0,0)$ formam uma base para o subespaço U. Logo, $dim(U) = 2$. Para o subespaço W temos a seguinte base

$$\{\ (-1,1,0,0),\ (0,0,1,0),\ (0,0,0,1)\ \}$$

portanto, $dim(W) = 3$.

Vamos agora determinar a dimensão do subespaço $U + W$, para isso construímos uma matriz cujas linhas são as coordenadas dos elementos do sistema de geradores do subespaço $U + W$, em relação à base canônica, que é dada por:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \end{bmatrix}$$

Em seguida, efetuando o escalonamento na matriz A encontramos uma matriz $\widehat{A}$ na forma escalonada, linha equivalente a matriz A, dada por:

$$\widehat{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Como $posto(A) = posto(\widehat{A}) = 4$, podemos concluir que $dim(U+W) = 4$.

Considerando o resultado do Teorema 3.6.5, temos que

$$dim(U \cap W) = dim(U) + dim(W) - dim(U+W) = 2 + 3 - 4 = 1$$

Assim, obtemos que $dim(U \cap W) = 1$.

Exemplo 3.6.33 *Sejam* U *e* W *subespaços vetoriais de* $\mathbb{R}^4$, *com* $U \neq W$, *tais que* $dim(U) = 3$ *e o subespaço* $U \cap W = [v_1, v_2, v_3]$, *onde*

$$v_1 = (1, 2, 1, 0) \quad , \quad v_2 = (-1, 1, 0, 1) \quad e \quad v_3 = (1, 5, 2, 1) \,.$$

Determine $dim(U \cap W)$ *e as possíveis dimensões dos subespaços* W *e* $U + W$.

Vamos determinar a dimensão do subespaço $U \cap W$. De modo análogo, escalonando a matriz A, cujas linhas são as coordenadas dos elementos do sistema de geradores do subespaço $U \cap W$, dada por:

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 1 & 5 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 1 \\ 0 & 3 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

obtemos que $dim(U \cap W) = 2$. Considerando o resultado do Teorema 3.6.5, temos que

$$3 \leq dim(U + W) = 1 + dim(W) \leq 4 \,,$$

uma vez que $U \subseteq U + W$ e $dim(U) = 3$. Portanto, temos duas possibilidades:

(1) $dim(W) = 2$ e $dim(U + W) = 3$

(2) $dim(W) = 3$ e $dim(U + W) = 4$

Exemplo 3.6.34 *Determine os valores de* $a \in \mathbb{R}$ *de modo que o conjunto*

$$S = \{ (a, 1, 0), (1, a, 1), (0, 1, a) \}$$

seja uma base para $\mathbb{R}^3$.

Vamos encontrar os valores para a de modo que o conjunto acima seja linearmente independente. Assim, vamos construir uma matriz $A \in \mathbb{M}_3(\mathbb{R})$ cujas linhas são formadas pelas coordenadas dos elementos do conjunto S, em relação à base canônica, e procedemos com o escalonamento

$$A = \begin{bmatrix} 1 & a & 1 \\ 0 & 1 & a \\ a & 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & a & 1 \\ 0 & 1 & a \\ 0 & 1 - a^2 & -a \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & a & 1 \\ 0 & 1 & a \\ 0 & 0 & a(a^2 - 2) \end{bmatrix}$$

Portanto, devemos impor que $a(a^2 - 2) \neq 0$. Assim, obtemos $a \neq 0$ e $a \neq \pm\sqrt{2}$.

De modo análogo, consideramos uma combinação linear nula dos elementos do conjunto S, obtendo um sistema linear homogêneo, e impomos a condição na matriz do sistema de modo que os coeficientes da combinação linear sejam todos nulos. Assim, obtemos os valores para o parâmetro a.

Exemplo 3.6.35 *Considere os seguintes subespaços do espaço vetorial real* $I\!R^3$

$$\begin{aligned} U &= [(1,-1,2),\ (2,1,1)] \\ W &= [(1,2,1),\ (0,1,-1)] \end{aligned}$$

Determine uma base para os subespaços $U + W$ *e* $U \cap W$*.*

Podemos verificar facilmente que o conjunto $\{\ (1,-1,2),\ (2,1,1)\ \}$ é uma base para o subespaço U e que $\{\ (1,2,1),\ (0,1,-1)\ \}$ é uma base para o subespaço W. Agora vamos determinar uma base para o subespaço $U + W$. Para isso construímos uma matriz cujas linhas são as coordenadas dos elementos do sistema de geradores do subespaço $U + W$, em relação à base canônica, e efetuamos o escalonamento

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -3 \\ 0 & 3 & -1 \\ 0 & 1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Assim, podemos escolher uma base para o subespaço $U + W$ um dos seguintes conjuntos

$$\{\ (1,-1,2),\ (2,1,1),\ (1,2,1)\ \} \qquad \text{ou} \qquad \{\ (1,-1,2),\ (0,3,-3),\ (0,0,2)\ \}$$

Logo, temos $dim(U + W) = 3$. Considerando o resultado do Teorema 3.6.5, obtemos

$$dim(U \cap W) \;=\; dim(U) \;+\; dim(W) \;-\; dim(U + W) \;=\; 1\,.$$

Neste caso, do processo de escalonamento, podemos concluir que o elemento $(0,1,-1)$ pertence ao subespaço $U \cap W$, pois observamos que ele também pertence ao subespaço U. Logo, $\{\,(0,1,-1)\,\}$ é uma base para o subespaço $U \cap W$.

Exemplo 3.6.36 *Considere os seguintes subespaços do espaço vetorial real* $I\!R^4$

$$\begin{aligned} U &= [(1,-1,0,2),\,(-1,2,0,1)] \\ W &= [(2,1,-1,3),\,(3,-4,0,3),\,(4,5,-3,5)] \end{aligned}$$

Determine uma base para o subespaço $U + W$ *e* $dim(U \cap W)$*.*

Vamos determinar a dimensão do subespaço U, procedendo o escalonamento na matriz

$$A = \begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 2 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 3 \end{bmatrix}$$

Como o $posto(A) = 2$, obtemos $dim(U) = 2$.

De modo análogo, vamos determinar a dimensão do subespaço W

$$A = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 3 & -4 & 0 & 3 \\ 4 & 5 & -3 & 5 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & -7 & 3 & -3 \\ 0 & 6 & 1 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & -7 & 3 & -3 \\ 0 & 0 & 25 & -32 \end{bmatrix}$$

Como o $posto(A) = 3$, obtemos $dim(W) = 3$.

De modo análogo, vamos determinar uma base para o subespaço $U + W$

$$\begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & -1 & 3 \\ 3 & -4 & 0 & 3 \\ 4 & 5 & -3 & 5 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 3 & -1 & -1 \\ 0 & -1 & 0 & -3 \\ 0 & 9 & -3 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & -10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & -30 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & -10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, provamos que o conjunto $\{\, (1,-1,0,2),\,(-1,2,0,1),\,(2,1,-1,3)\,\}$ é uma base para o subespaço $U + W$. Logo, $dim(U + W) = 3$.

Considerando o resultado do Teorema 3.6.5

$$dim(U \cap W) = dim(U) + dim(W) - dim(U + W) = 2 + 3 - 3 = 2$$

obtemos $dim(U \cap W) = 2$.

Exemplo 3.6.37 *Considere o subespaço U, do espaço vetorial $\mathcal{P}_2(\mathbb{R})$, definido por:*

$$U = \left\{ p \in \mathcal{P}_2(\mathbb{R}) \ / \ \int_{-1}^{1} p(x)dx \ + \ p'(0) = 0 \right\}.$$

Determine uma base para o subespaço U.

É importante lembrar que no exemplo 3.2.9 mostramos que U é um subespaço vetorial de $\mathcal{P}_2(\mathbb{R})$. Vamos determinar uma base para o subespaço U.

Tomando um elemento genérico $p(x) \in \mathcal{P}_2(\mathbb{R})$, que é escrito da forma:

$$p(x) = a + bx + cx^2 ,$$

e impondo a condição para que $p(x)$ pertença à U, isto é,

$$\int_{-1}^{1} (a + bx + cx^2)dx \ + \ b \ = \ 0 ,$$

obtemos a equação algébrica homogênea, que representa o vínculo entre os coeficientes do elemento $p(x)$, dada por:

$$6a + 3b + 2c = 0 ,$$

que possui dois grau de liberdade, de onde concluímos que $dim(U) = 2$.

Assim, podemos escolher

$$c = -3a - \frac{3}{2}b \quad ; \quad a, b \in \mathbb{R} .$$

Logo, todo elemento $p(x) \in U$ é escrito da seguinte forma:

$$p(x) = (1 - 3x^2)a + \left(x - \frac{3}{2}x^2 \right) b \quad ; \quad a, b \in \mathbb{R} .$$

Portanto, temos que o conjunto

$$\gamma = \left\{ 1 - 3x^2 , x - \frac{3}{2}x^2 \right\}$$

é uma base para o subespaço U, uma vez que γ é linearmente independente em $\mathcal{P}_2(\mathbb{R})$.

Exercícios

Exercício 3.42 *Verifique se os elementos*

$$u_1 = (1,1,1,1) \quad , \quad u_2 = (0,1,1,1) \quad , \quad u_3 = (0,0,1,1) \quad e \quad u_4 = (0,0,0,1)$$

formam uma base para o espaço vetorial real $I\!R^4$.

Exercício 3.43 *Encontre uma base para o subespaço* W *de* $I\!M_3(I\!R)$ *definido por:*

$$W = \{ A \in I\!M_3(I\!R) \ / \ A^t = -A \} .$$

Exercício 3.44 *Encontre uma base para o subespaço* W *de* $I\!M_3(I\!R)$ *definido por:*

$$W = \{ A \in I\!M_3(I\!R) \ / \ A^t = A \} .$$

Exercício 3.45 *Mostre que o conjunto* $\gamma = \{ 1, 1-x, (1-x)^2, (1-x)^3 \}$ *é uma base para o espaço vetorial real* $\mathcal{P}_3(I\!R)$.

Exercício 3.46 *Determine uma base para o subespaço vetorial de* $I\!M_2(I\!R)$ *dado por:*

$$U = \left\{ \begin{bmatrix} x & y \\ z & t \end{bmatrix} \ / \ x - y - z = 0 \right\} .$$

Exercício 3.47 *Determine uma base para o espaço solução do sistema linear*

$$\begin{cases} x + y + z + 2t = 0 \\ 2x - y - 2z - t = 0 \end{cases}$$

que é um subespaço vetorial do $I\!R^4$.

Exercício 3.48 *Considere os seguintes subespaços vetoriais de* $\mathcal{P}_2(I\!R)$

$$U = \{ p(x) = a + bx + cx^2 \in \mathcal{P}_2(I\!R) \ / \ a - 2c = 0 \}$$

$$W = [1-x \, , \, x-x^2]$$

Determine uma base para o subespaço $U + W$.

Exercício 3.49 *Considere os seguintes subespaços vetoriais de* $\mathcal{P}_2(I\!R)$

$$U = \{ p(x) = a + bx + cx^2 \in \mathcal{P}_2(I\!R) \ / \ a - 2c = 0 \}$$

$$W = [1-x \, , \, x-x^2]$$

Determine uma base para o subespaço $U \cap W$.

Exercício 3.50 *Considere os seguintes subespaços vetoriais de* $\mathbb{M}_2(\mathbb{R})$

$$U = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad ; \quad a,\, b,\, c, \, \in \, \mathbb{R} \right\} \qquad e \qquad W = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \quad ; \quad a,\, b,\, c \, \in \, \mathbb{R} \right\}.$$

Determine uma base para os subespaços U, W *e* $U \cap W$.

Exercício 3.51 *Para quais valores de* $a \in \mathbb{R}$ *o conjunto*

$$\beta \, = \, \{ \, (a, 1, 0) \, , \, (1, a, 1) \, , \, (0, 1, a) \, \}$$

é uma base para o espaço vetorial $\mathbb{R}^3$ *?*

Exercício 3.52 *Considere a seguinte Equação Diferencial Ordinária (EDO)*

$$-u''(x) \, + \, u(x) \, = \, 0 \, .$$

Mostre que as funções $u_1(x) \, = \, \exp(x)$ *e* $u_2 \, = \, \exp(-x)$ *são duas soluções linearmente independentes da EDO e que o conjunto* $\Gamma \, = \, \{ \, u_1(x), \, u_2(x) \, \}$ *é uma base para o espaço solução da EDO.*

Exercício 3.53 *Mostre que uma base para o espaço vetorial real* U *definido por:*

$$U \, = \, \{ \, p(x) \, \in \, \mathcal{P}_3(\mathbb{R}) \, \, / \, \, p(-1) \, = \, p(1) \, = \, 0 \, \}$$

é dada pelo conjunto $\gamma \, = \, \{ \, 1 \, - \, x^2 \, , \, x \, - \, x^3 \, \}$.

Exercício 3.54 *Sejam* U *e* W *subespaços vetoriais de dimensão finita de um espaço vetorial* V*. Determine as condições necessária e suficiente sobre os subespaços* U *e* W *para que* $dim(U \cap W) \, = \, dim(W)$.

Exercício 3.55 *Sejam* U *e* W *subespaços vetoriais de dimensão finita de um espaço vetorial* V *com dimensões* m *e* n*, respectivamente, onde* $m \geq n$.

(a) Prove que $dim(U \cap W) \, \leq \, n$.

(b) Prove que $dim(U + W) \, \leq \, n \, + \, m$.

Exercício 3.56 *Determine uma base para o espaço vetorial* $\mathbb{R}^4$ *contendo os elementos*

$$v_1 \, = \, (1, 1, 1, 0) \qquad e \qquad v_2 = (1, 1, 2, 1) \, .$$

Exercício 3.57 *Mostre que os polinômios*

$$p_1(x) = 1 \ , \quad p_2(x) = 1 + x \ , \quad p_3(x) = 1 - x^2 \quad e \quad p_4(x) = 1 - x - x^2 - x^3$$

formam uma base para o espaço vetorial $\mathcal{P}_3(\mathbb{R})$.

Exercício 3.58 *Sejam* V *e* W *subespaços vetoriais do espaço vetorial* $\mathbb{R}^3$ *tais que* $dim(V) = 1$, $dim(W) = 2$ *e* V *não está contido em* W. *Mostre que* $\mathbb{R}^3 = V \oplus W$.

Exercício 3.59 *Determine uma base para o espaço solução do sistema linear*

$$\begin{cases} x - y - z - t = 0 \\ \qquad 2y + 5z + t = 0 \end{cases}$$

que é um subespaço vetorial do $\mathbb{R}^4$.

Exercício 3.60 *Sejam* U *e* W *subespaços vetoriais de dimensão* 3 *do espaço vetorial* $\mathbb{R}^4$. *Considerando que*

$$U \cap W = [\, (1,2,1,0) \, , \, (-1,1,0,1) \, , \, (1,5,2,1) \,] \, .$$

Qual é a dimensão do subespaço $U + W$ *?*

Exercício 3.61 *Sejam* W *o subespaço de* $\mathbb{R}^4$ *definido por:*

$$W = \{ \, (x,y,z,t) \in \mathbb{R}^4 \ / \ x - y = z \quad e \quad x - 3y + t = 0 \, \}$$

e U *o subespaço de* $\mathbb{R}^4$ *gerado pelos elementos* $u_1 = (1,2,1,3)$ *e* $u_2 = (3,1,-1,4)$. *Determine uma base para o subespaço* $U + W$ *e para o subespaço* $U \cap W$.

Exercício 3.62 *Considere os seguintes subespaços de* $\mathbb{R}^3$

$$\begin{aligned} U &= \{ \, (x,y,z) \in \mathbb{R}^3 \ / \ x = 0 \, \} \\ V &= \{ \, (x,y,z) \in \mathbb{R}^3 \ / \ y - 2z = 0 \, \} \\ W &= [\, (1,1,0) \, , \, (0,0,2) \,] \end{aligned}$$

Determine uma base para cada um dos seguintes subespaços

$$U \cap V \quad , \quad V + W \qquad e \qquad U + V + W \, .$$

3.7 Coordenadas

Uma das características úteis de uma base β de um espaço vetorial V de dimensão finita é essencialmente que ela nos permite introduzir coordenadas em V de maneira análoga às *coordenadas naturais* x_i de um elemento $u = (x_1, \cdots, x_x)$ do espaço vetorial $\mathbb{R}^n$, por exemplo. Assim, as coordenadas de um elemento u de V em relação a base β serão os escalares que servem para representar u como uma combinação linear dos elementos da base ordenada β.

Se β é uma base arbitrária do espaço vetorial V de dimensão n, não teremos nenhuma ordenação natural para os elementos de β e será portanto necessário impormos uma certa ordem sobre esses elementos antes de podermos definir as coordenadas de um elemento de V em relação a β.

Definição 3.7.1 *Seja S um conjunto de n elementos. Uma **ordenação** do conjunto S, é uma função do conjunto dos inteiros positivos $1, \cdots, n$ sobre o conjunto S.*

Desse modo, uma ordenação do conjunto é simplesmente uma regra para nos dizer que elemento deve ser considerado como o primeiro elemento de S, que elemento é o segundo, e assim sucessivamente.

Uma **base ordenada** de um espaço vetorial V de dimensão finita é uma base β de V, mais uma ordenação fixa dos elementos de β. Desse modo, temos o seguinte resultado.

Teorema 3.7.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $\beta = \{ v_1, \cdots, v_n \}$ uma base ordenada de V. Então, todo elemento de V é escrito de modo único como uma combinação linear dos elementos de β, isto é, dado o elemento $u \in V$ temos que existe uma única n–upla $(c_1, \cdots, c_i, \cdots, c_n) \in \mathbb{F}^n$ tal que*

$$u = \sum_{i=1}^{n} c_i v_i .$$

Dizemos que c_i é a i–ésima coordenada do elemento u com relação à base ordenada β.

Demonstração – Para mostrar a unicidade, vamos considerar que

$$u = \sum_{j=1}^{n} c_j v_j = \sum_{j=1}^{n} b_j v_j \qquad \Longrightarrow \qquad \sum_{j=1}^{n} (c_j - b_j) v_j = 0_V .$$

Como $\{ v_1, \cdots, v_n \}$ é linearmente independente em V, temos que $c_j - b_j = 0$ para todo j. Logo, $c_j = b_j$ para todo j, o que completa a demonstração. ■

Observamos que, cada base ordenada β do espaço vetorial V determina uma bijeção

$$u \in V \quad \longrightarrow \quad (c_1\,,\,\cdots\,,\,c_i\,,\,\cdots\,,\,c_n) \in \mathbb{F}^n$$

entre o conjunto de elementos de V e o conjunto das n–uplas de $\mathbb{F}^n$.

Esta associação tem a propriedade de que o correspondente do elemento $(u + v) \in V$ é a soma em $\mathbb{F}^n$ das correspondentes n–uplas de u e v. Além disso, o correspondente do elemento $(\lambda\,u) \in V$ é o produto em $\mathbb{F}^n$ do escalar λ pela correspondente n–upla do elemento u.

Neste ponto poderíamos perguntar por que não tomar simplesmente uma base ordenada no espaço vetorial V e descrever cada elemento de V pelo seu vetor de coordenadas, visto que teríamos então a conveniência de operar apenas com elementos de $\mathbb{F}^n$. Esta atitude faria malograr nosso objetivo, por duas razões. A primeira, como indica a nossa definição de espaço vetorial, estamos aprendendo a raciocinar com espaços vetoriais como sistemas algébricos. A segunda razão, mesmo nos casos em que usamos coordenadas, os resultados importantes decorrem de nossa habilidade de mudar o sistema de coordenadas, isto é, mudar a base ordenada do espaço vetorial V.

Desse modo, será mais conveniente utilizar a **matriz de coordenadas** do elemento u em relação à base ordenada β, que denotamos por $[u]_\beta$, dada por:

$$[u]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{F})\,.$$

Esta notação será particularmente útil quando passarmos a descrever o que ocorre com as coordenadas de um elemento $u \in V$ quando fazemos a mudança de uma base ordenada para uma outra base ordenada, que é o tema da próxima seção.

Teorema 3.7.2 *Considere* $A \in \mathbb{M}_n(\mathbb{R})$. *Então, os vetores colunas de* A *formam um conjunto linearmente independente em* $\mathbb{R}^n$ *se, e somente se,* A *é uma matriz invertível.*

Demonstração

($\Longrightarrow$) Sejam $v_1, \cdots, v_n \in \mathbb{R}^n$ os vetores colunas da matriz A e W o subespaço de $\mathbb{R}^n$ gerado pelos vetores colunas de A. Como os vetores $v_1, \cdots, v_n$ são linearmente independentes, temos que $dim(\, W\,) = n$. Pelo Corolário 3.6.3, sabemos que $W = \mathbb{R}^n$.

Desse modo, existem escalares $b_{ij} \in \mathbb{R}$, para $i, j = 1, \cdots, n$, de modo que cada elemento e_j pode ser escrito de modo único da forma:

$$e_j = \sum_{i=1}^{n} b_{ij}\, v_i \qquad ; \qquad j = 1, \cdots, n \,,$$

onde $\{ e_1, \cdots, e_j, \cdots, e_n \}$ é a base canônica do $\mathbb{R}^n$.

Portanto, a matriz $B = [b_{ij}]$ satisfaz $AB = I$.

($\Longleftarrow$) A prova segue do Teorema 2.9.7 e do conceito de independência linear. ■

Exemplo 3.7.1 *Considere o espaço vetorial real* $\mathbb{R}^n$ *e o elemento*

$$u = (x_1, \cdots, x_n) \in \mathbb{R}^n \,.$$

Se $\beta = \{ e_1, \cdots, e_n \}$ *é a base ordenada canônica de* $\mathbb{R}^n$*, a matriz de coordenadas do elemento* u *em relação à base* β *é dada por:*

$$[u]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} .$$

Para exemplificar, considere o elemento $u = (1, 2, 4, 7) \in \mathbb{R}^4$. Assim, a matriz de coordenadas de u com relação à base ordenada canônica de $\mathbb{R}^4$ é dada por:

$$[u]_\beta = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 7 \end{bmatrix} .$$

Exemplo 3.7.2 *Podemos verificar facilmente que* $\gamma = \{ 1, 1+x, 1+x^2 \}$ *é uma base ordenada para o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$. *Determine as coordenadas do elemento* $p(x) = 2 + 4x + x^2$ *em relação à base ordenada* γ.

Inicialmente, escrevemos o polinômio $p(x)$ como uma combinação linear dos elementos da base ordenada γ

$$p(x) = 2 + 4x + x^2 = a + b(1+x) + c(1+x^2) = (a+b+c) + bx + cx^2$$

Assim, obtemos o seguinte sistema linear nas incógnitas a, b e c que são as coordenadas de $p(x)$ com relação à base ordenada γ

$$\begin{aligned} a + b + c &= 2 \\ b &= 4 \\ c &= 1 \end{aligned}$$

Desse modo, temos que $a = -3$, $b = 4$ e $c = 1$. Portanto, o vetor de coordenadas do elemento p com relação à base ordenada γ é dado por:

$$[p]_\gamma = \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}$$

Exemplo 3.7.3 *Considere o espaço vetorial real* $\mathbb{R}^3$. *Determine as coordenadas do elemento* $u = (2,1,4) \in \mathbb{R}^3$ *com relação à base* $\gamma = \{ (1,1,1), (1,0,1), (1,0,-1) \}$.

Vamos escrever o elemento $u = (2,1,4) \in \mathbb{R}^3$ como uma combinação linear dos elementos da base ordenada γ

$$u = (2,1,4) = a(1,1,1) + b(1,0,1) + c(1,0,-1)$$

Assim, obtemos o seguinte sistema linear nas incógnitas a, b e c que são as coordenadas de u com relação à base ordenada γ

$$\begin{aligned} a + b + c &= 2 \\ a &= 1 \\ a + b - c &= 4 \end{aligned}$$

Desse modo, temos que $a = 1$, $b = 2$ e $c = -1$. Portanto, o vetor de coordenadas do elemento u com relação à base ordenada γ é dado por:

$$[u]_\gamma = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Exemplo 3.7.4 *Sejam* V *um espaço vetorial real e* $\gamma = \{ v_1, v_2, v_3, v_4 \}$ *uma base ordenada para* V. *Pede–se:*

(a) *Mostre que* $\beta = \{ v_1, v_1 + v_2, v_1 + v_2 + v_3, v_1 + v_2 + v_3 + v_4 \}$ *é uma base de* V.

(b) *Se o elemento* $v \in V$ *tem como vetor de coordenadas*

$$[v]_\gamma = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 2 \end{bmatrix}$$

determine seu vetor de coordenadas $[v]_\beta$.

Vamos mostrar que β é linearmente independente em V. Para isso, vamos considerar a combinação linear nula

$$av_1 + b(v_1 + v_2) + c(v_1 + v_2 + v_3) + d(v_1 + v_2 + v_3 + v_4) = 0_V .$$

Escrevendo a combinação linear acima da seguinte forma:

$$(a + b + c + d)v_1 + (b + c + d)v_2 + (c + d)v_3 + dv_4 = 0_V$$

e utilizando a hipótese que $\gamma = \{ v_1, v_2, v_3, v_4 \}$ é linearmente independente em V, obtemos o seguinte sistema linear triangular superior homogêneo

$$\begin{array}{rcrcrcrcl} a & + & b & + & c & + & d & = & 0 \\ & & b & + & c & + & d & = & 0 \\ & & & & c & + & d & = & 0 \\ & & & & & & d & = & 0 \end{array}$$

que tem por solução $a = b = c = d = 0$. Assim, provamos que β é linearmente independente em V. Logo, β é uma base para V.

Note que na primeira parte da resolução, o vetor de coordenadas de um elemento $v \in V$ na base β está representado por:

$$[v]_\beta = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

e o vetor de coordenadas do elemento $v \in V$ na base γ está representado por

$$[v]_\beta = \begin{bmatrix} a+b+c+d \\ b+c+d \\ c+d \\ d \end{bmatrix} .$$

Assim, para encontrar o vetor de coordenadas $[v]_\beta$, conhecendo o vetor de coordenadas $[v]_\gamma$, basta obter a solução do seguinte sistema linear triangular superior

$$\begin{array}{rcrcrcrcl} a & + & b & + & c & + & d & = & 4 \\ & & b & + & c & + & d & = & 3 \\ & & & & c & + & d & = & 1 \\ & & & & & & d & = & 2 \end{array}$$

Portanto, temos que

$$[v]_\beta = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \end{bmatrix}$$

Exemplo 3.7.5 *Considere o espaço vetorial real* $I\!M_2(I\!R)$ *com a base ordenada* $\gamma = \{ A_1, A_2, A_3, A_4 \}$ *onde*

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad , \quad A_3 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad e \quad A_4 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Determine o vetor de coordenadas $[A]_\gamma$ da matriz A dada por:

$$A = \begin{bmatrix} 4 & 6 \\ 5 & 6 \end{bmatrix} .$$

Resposta: $\boxed{[A]_\gamma = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix}}$

Exemplo 3.7.6 *Considere o espaço vetorial* $\mathcal{F}(\mathbb{R})$ *e o subespaço* $U = [u_1(x), u_2(x)]$, *onde* $u_1 = \exp(x)$ *e* $u_2 = \exp(-x)$. *Podemos verificar facilmente que o conjunto* $\Gamma = \{ \exp(x), \exp(-x) \}$ *é uma base ordenada para o subespaço* U. *Temos que as funções hiperbólicas*

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \qquad e \qquad \sinh(x) = \frac{e^x - e^{-x}}{2}$$

pertencem ao subespaço U *e cujos vetores de coordenadas em relação à base ordenada* Γ *são dados por*

$$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \qquad e \qquad \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix},$$

respectivamente.

Teorema 3.7.3 *Considere* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$ *e* $\beta = \{ v_1 , \cdots , v_n \}$ *um conjunto finito de elementos de* V. *Se todo elemento de* V *é escrito de modo único como uma combinação linear dos elementos de* β, *então* β *é uma base de* V.

Demonstração – Como todo elemento de V é escrito como uma combinação linear dos elementos de β, temos que V é gerado pelos elementos do conjunto β. Agora basta mostrar que β é linearmente independente em V.

Considere os escalares $c_1 , \cdots , c_n \in \mathbb{F}$ tais que

$$\sum_{i=1}^{n} c_i \, v_i = 0_V .$$

Temos também que

$$\sum_{i=1}^{n} 0_{\mathbb{F}} \, v_i = 0_V .$$

Desse modo, pela hipótese de unicidade, obtemos $c_i = 0_{\mathbb{F}}$ para $i = 1, \cdots , n$.

Portanto, β é uma base para o espaço vetorial V. ■

Exercícios

Exercício 3.63 *Considere o espaço vetorial* $\mathcal{P}_3(\mathbb{R})$ *com a base ordenada*

$$\gamma = \{ 1, 1 - x, (1-x)^2, (1-x)^3 \}.$$

Encontre o vetor de coordenadas $[p]_\gamma$ *do polinômio* $p(x) = 3 - 2x - x^2$*.*

Exercício 3.64 *Considere a base* $\beta = \{ 2, a + x, 1 + bx^2 \}$ *de* $\mathcal{P}_2(\mathbb{R})$*. Determine as constantes* $a, b \in \mathbb{R}$ *de modo que o vetor de coordenadas do polinômio* $p(x) = x + x^2$ *em relação à base* β *seja dado por:*

$$[p]_\beta = \begin{bmatrix} \frac{3}{2} \\ 1 \\ -1 \end{bmatrix}.$$

Exercício 3.65 *Mostre que o conjunto* $\gamma = \{ 1, x + a, (x + a)^2 \}$*, para* $a \in \mathbb{R}$ *fixo, é uma base para o espaço vetorial* $\mathcal{P}_2(\mathbb{R})$*.*

Exercício 3.66 *Considere a base ordenada* $\gamma = \{ 1, x + a, (x + a)^2 \}$*,* $a \in \mathbb{R}$ *fixo, do espaço vetorial* $\mathcal{P}_2(\mathbb{R})$*. Considere que um elemento* $p(x) \in \mathcal{P}_2(\mathbb{R})$ *tem por vetor de coordenadas, em relação à base ordenada canônica* $\beta = \{ 1, x, x^2 \}$*,*

$$[p(x)]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

Determine o vetor de coordenadas do elemento $p(x)$ *em relação à base ordenada* γ*.*

Exercício 3.67 *Seja* $\gamma = \{ v_1 , v_2 , v_3 \}$ *uma base ordenada para o espaço vetorial real* V*. Pede–se:*

(a) Mostre que $\beta = \{ v_1 , v_1 + v_2 , -v_1 + v_2 + v_3 \}$ *é também uma base para* V*.*

(b) Considere que o vetor de coordenadas do elemento $v \in V$*, em relação à base* γ*, é dado por:*

$$[v]_\gamma = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}.$$

Determine o vetor de coordenadas do elemento v *em relação à base* β*.*

3.8 Mudança de Base

Teorema 3.8.1 *Sejam* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$*,* $\beta = \{ v_1 , \cdots , v_n \}$ *e* $\gamma = \{ w_1 , \cdots , w_n \}$ *bases ordenadas para* V*. Então, existe uma única matriz* $P \in \mathbb{M}_n(\mathbb{F})$ *invertível tal que para todo* $u \in V$ *tem–se que*

(a) $[u]_\gamma = P\,[u]_\beta$

(b) $[u]_\beta = P^{-1}\,[u]_\gamma$

Demonstração – Inicialmente vamos representar cada elemento v_j da base ordenada β em relação à base ordenada γ. Pelo Teorema 3.7.1, existem escalares $p_{1j}, \cdots , p_{nj} \in \mathbb{F}$, bem definidos, tais que

$$v_j = \sum_{i=1}^{n} p_{ij}\, w_i \qquad ; \qquad j = 1, \cdots , n\,.$$

Dado um elemento $u \in V$, sejam $c_1, \cdots , c_n \in \mathbb{F}$ suas coordenadas com relação à base ordenada β, isto é,

$$\begin{aligned} u &= \sum_{j=1}^{n} c_j\, v_j = \sum_{j=1}^{n} c_j \sum_{i=1}^{n} p_{ij}\, w_i \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} (\, p_{ij}\, c_j \,)\, w_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij}\, c_j \right) w_i\,. \end{aligned}$$

Assim, obtemos a relação

$$u = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij}\, c_j \right) w_i\,.$$

Como as coordenadas $b_1, \cdots , b_n$ do elemento $u \in V$ com relação à base γ são determinadas de modo único, temos que

$$b_i = \sum_{j=1}^{n} p_{ij}\, c_j \qquad ; \qquad i = 1, \cdots , n\,.$$

Tomando a matriz $P = [p_{ij}]$, podemos escrever a relação acima na forma matricial $[u]_\gamma = P\,[u]_\beta$. Como β e γ são linearmente independentes em V, então $[u]_\gamma = 0$ se, e somente se, $[u]_\beta = 0$. Assim, temos que P é uma matriz invertível. Logo, temos que $[u]_\beta = P^{-1}\,[u]_\gamma$. A matriz $P = [p_{ij}]$ é denominada matriz de mudança da base ordenada β para a base ordenada γ, o que completa a demonstração. ■

Utilizaremos a notação $[I]_\gamma^\beta$ para indicar a matriz de mudança da base ordenada β para a base ordenada γ.

Teorema 3.8.2 *Sejam* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$*,* $\gamma = \{ w_1, \cdots, w_n \}$ *uma base ordenada para* V*, e* $P \in \mathbb{M}_n(\mathbb{F})$ *uma matriz invertível. Então, existe uma única base ordenada* $\beta = \{ v_1, \cdots, v_n \}$ *de* V *tal que*

(a) $[u]_\gamma = P\,[u]_\beta$

(b) $[u]_\beta = P^{-1}\,[u]_\gamma$

para todo elemento $u \in V$*.*

Demonstração – Se $\beta = \{ v_1, \cdots, v_n \}$ é uma base ordenada para V para qual *(a)* é válido, é claro que

$$v_j = \sum_{i=1}^{n} p_{ij}\, w_i \qquad ; \qquad j = 1, \cdots, n\,.$$

Assim, basta mostrar que os elementos v_j assim definidos formam uma base para V. Considerando a matriz $Q = [q_{ij}] = P^{-1}$, temos que

$$\begin{aligned}
\sum_{j=1}^{n} q_{jk}\, v_j &= \sum_{j=1}^{n} q_{jk} \sum_{i=1}^{n} p_{ij}\, w_i \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} p_{ij}\, q_{jk} \right) w_i \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij}\, q_{jk} \right) w_i \\
&= w_k
\end{aligned}$$

Portanto, o subespaço gerado pelos elementos do conjunto β contém γ. Logo, é igual ao espaço vetorial V. Assim, β é uma base e, de sua definição e do Teorema 3.8.1, é evidente que *(a)* é válido. Logo, *(b)* também o é, o que completa a demonstração. ■

Note que a j–ésima coluna da matriz P são as coordenadas do j–ésimo elemento da base ordenada β com relação à base ordenada γ, isto é,

$$v_j = \sum_{i=1}^{n} p_{ij}\, w_i \qquad ; \qquad j = 1, \cdots, n\,.$$

Portanto, P é a matriz de mudança da base ordenada β para a base ordenada γ, isto é, $P = [I]^{\beta}_{\gamma}$.

Exemplo 3.8.1 *Considere a matriz* $P \in I\!M_2(I\!R)$ *dada por:*

$$P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

onde θ *é um número real. A matriz* P *é a matriz de rotação de um ângulo* θ *no sentido anti–horário. A matriz* P *é invertível e sua inversa é dada por:*

$$P^{-1} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} .$$

Portanto, para cada $\theta \in I\!R$, *o conjunto* $\gamma = \{ v_1, v_2 \}$ *onde*

$$v_1 = (\cos(\theta) , \sin(\theta)) \qquad e \qquad v_2 = (-\sin(\theta) , \cos(\theta))$$

é uma base ordenada de $I\!R^2$.

Podemos observar que a base ordenada γ *pode ser descrita como sendo a base obtida pela rotação de um ângulo* θ *da base ordenada canônica* $\beta = \{ e_1, e_2 \}$. *Assim, dado um elemento* $u = (x, y) \in I\!R^2$, *temos que*

$$[u]_\gamma = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_\beta .$$

Logo, temos que $[u]_\gamma = P^{-1}[u]_\beta$ *e* $[u]_\beta = P[u]_\gamma$.

Esse exemplo é uma excelente ilustração do Teorema 3.8.2. Para exemplificar, vamos considerar $\theta = \dfrac{\pi}{4}$. Desse modo, temos a base ordenada canônica $\beta = \{ (1,0), (0,1) \}$ e a base ordenada $\gamma = \{ v_1, v_2 \}$ onde

$$v_1 = \frac{\sqrt{2}}{2}(1,1) \qquad \text{e} \qquad v_2 = \frac{\sqrt{2}}{2}(-1,1) .$$

Neste caso, as seguintes matrizes

$$P = \begin{bmatrix} \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{2}}{2} \\ \dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} \end{bmatrix}_\beta^\gamma \qquad \text{e} \qquad P^{-1} = \begin{bmatrix} \dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} \\ -\dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} \end{bmatrix}_\gamma^\beta$$

que são, respectivamente, a matriz de mudança da base ordenada γ para a base canônica β e a matriz de mudança da base canônica β para a base ordenada γ.

Exemplo 3.8.2 *Considere o espaço vetorial real* $\mathbb{R}^3$*. Determine a matriz de mudança de base,* $[I]^{\beta}_{\gamma}$ *, da base canônica* $\beta = \{ (1,0,0), (0,1,0), (0,0,1) \}$ *para a base ordenada* $\gamma = \{ (1,1,1), (1,0,1), (1,0,-1) \}$.

Vamos escrever cada elemento da base canônica β como uma combinação linear dos elementos da base γ

$$\begin{aligned} e_1 &= (1,0,0) = p_{11}(1,1,1) + p_{21}(1,0,1) + p_{31}(1,0,-1) \\ e_2 &= (0,1,0) = p_{12}(1,1,1) + p_{22}(1,0,1) + p_{32}(1,0,-1) \\ e_3 &= (0,0,1) = p_{13}(1,1,1) + p_{23}(1,0,1) + p_{33}(1,0,-1) \end{aligned}$$

obtendo três sistemas lineares, que podem ser resolvidos pelo processo de escalonamento. Assim, temos que

$$[I]^{\beta}_{\gamma} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$$

Podemos verificar facilmente que

$$[I]^{\gamma}_{\beta} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix}$$

Portanto, temos que $[I]^{\gamma}_{\beta}\,[I]^{\beta}_{\gamma} = I$.

Exemplo 3.8.3 *Considere o espaço vetorial real* $\mathbb{R}^2$*. Determine a matriz de mudança da base* $\alpha = \{ (-3,-1), (-1,3) \}$ *para a base* $\gamma = \{ (-1,1), (1,1) \}$.

Resposta: $$[I]^{\alpha}_{\gamma} = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \quad \text{e} \quad (\,[I]^{\alpha}_{\gamma}\,)^{-1} = [I]^{\gamma}_{\alpha} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}$$

Exemplo 3.8.4 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$. *A matriz de mudança da base* $\gamma = \{ 1 + t, 1 - t^2 \}$ *para uma base* α, *de um mesmo subespaço de* $\mathcal{P}_2(\mathbb{R})$, *é dada por*

$$[I]^{\gamma}_{\alpha} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$

Determine a base α.

Fazendo $\alpha = \{ p_1, p_2 \}$, temos que

$$\begin{aligned} p_1(t) &+ p_2(t) = 1 + t \\ 2p_1(t) &- p_2(t) = 1 - t^2 \end{aligned}$$

obtendo a seguinte relação entre os elementos da base α com os elementos da base γ

$$\begin{aligned} p_1(t) &= \frac{1}{3}(1 + t) + \frac{1}{3}(1 - t^2) \\ p_2(t) &= \frac{2}{3}(1 + t) - \frac{1}{3}(1 - t^2) \end{aligned}$$

De outro modo, sabemos que

$$[I]^{\alpha}_{\gamma} = ([I]^{\gamma}_{\alpha})^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

Assim, obtemos a relação entre os elementos da base α com os elementos da base γ.

Exemplo 3.8.5 *Considere o espaço vetorial* $\mathcal{P}_2(\mathbb{R})$. *Determine a matriz de mudança da base canônica* $\beta = \{ 1, x, x^2 \}$ *para a base ordenada* $\gamma = \{ 2, 1 - x, 1 - x^2 \}$.

Podemos verificar facilmente que

$$[I]^{\gamma}_{\beta} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Portanto, temos que

$$[I]^{\beta}_{\gamma} = ([I]^{\gamma}_{\beta})^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Exercícios

Exercício 3.68 *Considere a base ordenada* $\gamma = \{ v_1, v_2, v_3 \}$ *do* $I\!R^3$ *onde*

$$v_1 = (1, 0, -1) , \qquad v_2 = (1, 1, 1) \quad e \quad v_3 = (1, 0, 0) .$$

Encontre o vetor de coordenadas do elemento $u = (a, b, c) \in I\!R^3$ *com relação à base ordenada* γ.

Exercício 3.69 *Considere o seguinte subespaço vetorial de* $I\!M_2(I\!R)$

$$U = \left\{ \begin{bmatrix} x & y \\ z & t \end{bmatrix} \ / \ x - y - z = 0 \right\} .$$

Considere as seguintes bases do subespaço vetorial U

$$\beta = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} , \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} \qquad e \qquad \gamma = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} , \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\} .$$

Determine a matriz de mudança de base $[I]^{\gamma}_{\beta}$.

Exercício 3.70 *Seja* V *um espaço vetorial real de dimensão* n. *Sejam*

$$\beta = \{ v_1 , \cdots , v_n \} \ , \ \alpha = \{ u_1 , \cdots , u_n \} \quad e \quad \gamma = \{ w_1 , \cdots , w_n \}$$

três bases ordenadas para V. *Considerando a matriz* $P = [I]^{\beta}_{\alpha}$, *a matriz de mudança da base* β *para a base* α, *e* $Q = [I]^{\alpha}_{\gamma}$, *a matriz de mudança da base* α *para a base* γ. *Determine a matriz de mudança da base* β *para a base* γ, $[I]^{\beta}_{\gamma}$. *Utilizando esse resultado mostre que uma matriz de mudança de base é sempre invertível.*

Exercício 3.71 *Seja* V *um espaço vetorial real e* $\beta = \{ u_1, \cdots , u_n \}$ *uma base para* V. *Mostre que se* $A = [a_{ij}]$ *é uma matriz de ordem* n *invertível, então os elementos*

$$v_j = \sum_{i=1}^{n} a_{ij} \, u_i \qquad para \qquad j = 1, \cdots , n$$

formam uma base para V, *e que* A *é a matriz de mudança da base*

$$\gamma = \{ v_1, \cdots , v_n \}$$

para a base β, *isto é* $A = [I]^{\gamma}_{\beta}$.

Exercício 3.72 *Considere o espaço vetorial real* $I\!R^2$. *A matriz de mudança da base ordenada* $\gamma = \{ u_1, u_2 \}$, *onde* $u_1 = (1,1)$ *e* $u_2 = (-2,2)$, *para a base ordenada* $\alpha = \{ v_1, v_2 \}$ *é dada por:*

$$[I]^{\gamma}_{\alpha} = \begin{bmatrix} 1 & 0 \\ 4 & -2 \end{bmatrix}.$$

Determine a base ordenada α. *Determine o elemento* $u \in I\!R^2$ *tal que* $[u]_\alpha = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Exercício 3.73 *Considere as bases* $\beta = \{u_1, u_2, u_3\}$ *e* $\gamma = \{w_1, w_2, w_3\}$ *de* $I\!R^3$, *relacionadas da seguinte forma:*

$$\begin{cases} w_1 = u_1 - u_2 - u_3 \\ w_2 = 2u_2 + 3u_3 \\ w_3 = 3u_1 + u_3 \end{cases}$$

Pede–se:

(a) Determine as matrizes de mudança de base $[I]^{\beta}_{\gamma}$ *e* $[I]^{\gamma}_{\beta}$.

(b) Considere que o elemento $u \in I\!R^3$ *tem por vetor de coordenadas*

$$[u]_\beta = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Determine o vetor de coordenadas do elemento u *com relação à base* γ.

Exercício 3.74 *Considere a seguinte matriz de mudança de base*

$$[I]^{\beta'}_{\beta} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

Encontre:

(a) $[v]_\beta$ *onde* $[v]_{\beta'} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$

(b) $[v]_{\beta'}$ *onde* $[v]_\beta = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] LUENBERGER, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} = Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações
Notas de Aula

Petronio Pulino
Departamento de Matemática Aplicada
Instituto de Matemática, Estatística e Computação Científica
Universidade Estadual de Campinas
E–mail: pulino@ime.unicamp.br
www.ime.unicamp.br/∼pulino/ALESA/

Janeiro de 2012

Conteúdo

4

Transformações Lineares

Conteúdo

4.1 Transformações do Plano no Plano

Exemplo 4.1.1 *Considere o espaço vetorial real* $I\!R^2$*, e um escalar* $\lambda \in I\!R$ *fixo. A transformação*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) \;=\; \lambda\,(x,\,y) \end{array}$$

é uma contração para $0 < \lambda < 1$*. Quando* $\lambda > 1$*, dizemos que* T *é uma expansão.*

Exemplo 4.1.2 *Considere o espaço vetorial real* $I\!R^2$*. A transformação*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) \;=\; (x,\,-y) \end{array}$$

é a reflexão em torno do eixo–ox.

Exemplo 4.1.3 *Considere o espaço vetorial real* $I\!R^2$*. A transformação*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) \;=\; (-x,\,-y) \end{array}$$

é a reflexão em torno da origem.

Exemplo 4.1.4 *Considere o espaço vetorial real* $I\!R^2$*. Dado um elemento* $(a,b) \in I\!R^2$*, a transformação*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) \;=\; (x\,,\,y) \;+\; (a\,,\,b) \end{array}$$

é uma translação.

Exemplo 4.1.5 *Considere o espaço vetorial real* $I\!R^2$*. A transformação*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) \;=\; (\,x\cos(\theta) \;-\; y\sin(\theta)\,,\,x\sin(\theta) \;+\; y\cos(\theta)\,) \end{array}$$

é uma rotação de um ângulo θ *no sentido anti–horário.*

4.2 Transformação Linear

Definição 4.2.1 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e T uma aplicação de V em W. Dizemos que T é uma* ***Transformação Linear*** *se possui as seguintes propriedades:*

(a) $T(u + v) = T(u) + T(v)$ *para todo* $u, v \in V$.

(b) $T(\lambda u) = \lambda T(u)$ *para todo* $u \in V$, $\lambda \in \mathbb{F}$.

Das duas propriedades de transformação linear, obtemos facilmente que

$$T(au + bv) = a\,T(u) + b\,T(v)$$

para todo $u, v \in V$ e todos escalares $a, b \in \mathbb{F}$. Por indução, obtemos uma relação mais geral

$$T\left(\sum_{j=1}^{n} \alpha_j\,u_j \right) = \sum_{j=1}^{n} \alpha_j\,T(u_j)$$

para quaisquer elementos $u_1, \cdots, u_n \in V$ e quaisquer escalares $\alpha_j, \cdots, \alpha_n \in \mathbb{F}$.

Finalmente, fazendo $\lambda = 0$ na propriedade *(b)*, tem–se $T(0_V) = 0_W$.

Exemplo 4.2.1 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Vamos definir a seguinte transformação linear $T(v) = v$ para todo $v \in V$, que é a transformação identidade, denotada por I_V.*

Exemplo 4.2.2 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Vamos definir a seguinte transformação linear $T(v) = 0_V$ para todo $v \in V$, que é a transformação nula.*

Exemplo 4.2.3 *Dado um elemento $c = (c_1, \cdots, c_n) \in \mathbb{R}^n$ fixo, porem arbitrário, vamos definir a seguinte transformação linear*

$$\begin{aligned} T :\ & \mathbb{R}^n \longrightarrow \mathbb{R} \\ & x \longrightarrow T(x) = \sum_{j=1}^{n} c_j\,x_j \end{aligned}$$

A transformação linear T é o produto escalar entre o elemento x e o elemento c.

Exemplo 4.2.4 *Seja V um espaço vetorial real. Considerando um escalar $\lambda \in I\!R$ fixo, porem arbitrário, definimos a seguinte transformação linear*

$$\begin{aligned} T: \ V & \longrightarrow V \\ v & \longrightarrow T(v) = \lambda v \end{aligned}$$

*A transformação linear T é uma **contração** para $|\lambda| < 1$. Quando $|\lambda| > 1$, dizemos que a transformação linear T é uma **expansão**.*

Exemplo 4.2.5 *Considerando os espaços vetoriais reais $\mathcal{C}([a,b])$ e $\mathcal{C}^1([a,b])$, definimos a seguinte transformação linear*

$$\begin{aligned} T: \ \mathcal{C}^1([a,b]) & \longrightarrow \mathcal{C}([a,b]) \\ f & \longrightarrow T(f) = f' \end{aligned}$$

com $T(f)(x) = f'(x) \quad ; \quad x \in [a,b]$.

Exemplo 4.2.6 *Considerando os espaços vetoriais reais $\mathcal{C}([a,b])$ e $\mathcal{C}^1([a,b])$, definimos a seguinte transformação linear*

$$\begin{aligned} T: \ \mathcal{C}([a,b]) & \longrightarrow \mathcal{C}^1([a,b]) \\ f & \longrightarrow g = T(f) \end{aligned}$$

com $g(x) = T(f)(x) = \displaystyle\int_a^x f(t)dt \quad ; \quad x \in [a,b]$.

Exemplo 4.2.7 *Considere os espaços vetoriais reais $I\!R^m$ e $I\!R^n$. Dada uma matriz $A = [a_{ij}] \in I\!M_{m \times n}(I\!R)$, definimos a transformação linear associada a matriz A da seguinte forma:*

$$\begin{aligned} T_A: \ I\!R^n & \longrightarrow I\!R^m \\ x & \longrightarrow y = T_A(x) \end{aligned}$$

onde

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j \qquad \text{para} \qquad i = 1, \cdots, m$$

é a i–ésima componente do elemento $y = (y_1, \cdots, y_m)$.

Teorema 4.2.1 *Sejam V e W espaços vetoriais sobre o corpo $I\!F$, com $dim(V)$ igual a n, $\beta = \{ v_1 , \cdots , v_n \}$ uma base ordenada para V e $w_1 , \cdots , w_n$ elementos arbitrários de W. Então, existe uma única transformação linear $T : V \longrightarrow W$ tal que*

$$T(v_j) = w_j \qquad \textit{para} \qquad j = 1, \cdots , n .$$

Demonstração – Inicialmente vamos mostrar que existe pelo menos uma transformação linear T com $T(v_j) = w_j$. Dado um elemento $u \in V$, sabemos que u é escrito de modo único como:

$$u = \sum_{i=1}^{n} c_i \, v_i .$$

Para este elemento u, vamos definir uma aplicação $T : V \longrightarrow W$ da forma:

$$T(u) = \sum_{i=1}^{n} c_i \, w_i .$$

Temos que T é uma transformação bem definida. Pela definição, fica evidente que $T(v_i) = w_i$. Para mostrar que T é uma transformação linear, sejam $\lambda \in I\!F$ e um elemento $v \in V$ escrito de modo único como:

$$v = \sum_{i=1}^{n} b_i \, v_i .$$

Assim, temos que

$$T(u + \lambda v) = \sum_{i=1}^{n} (c_i + \lambda b_i) \, w_i = \sum_{i=1}^{n} c_i \, w_i + \lambda \sum_{i=1}^{n} b_i \, w_i = T(u) + \lambda T(v) .$$

mostrando que a aplicação T é linear.

Finalmente vamos mostrar a unicidade da transformação linear T. Para isso, supomos que existe uma outra transformação linear $P : V \longrightarrow W$ tal que

$$P(v_j) = w_j \qquad \text{para} \qquad j = 1, \cdots , n .$$

Desse modo, temos que

$$P(u) = P\left(\sum_{i=1}^{n} c_i \, v_i \right) = \sum_{i=1}^{n} c_i \, P(v_i) = \sum_{i=1}^{n} c_i \, w_i .$$

Logo, P é exatamente a regra da transformação linear T definida acima. Portanto, provamos a unicidade da transformação linear T, o que completa a demonstração. ■

Exemplo 4.2.8 *A aplicação* $T : I\!R^2 \longrightarrow \mathcal{P}_2(I\!R)$ *tal que*

$$T(1,0) = 1 - x \qquad e \qquad T(0,1) = 1 - x^2$$

define uma transformação linear de $I\!R^2$ *em* $\mathcal{P}_2(I\!R)$.

Estamos considerando o espaço vetorial $I\!R^2$ com a base canônica

$$\beta = \{ (1,0), (0,1) \} .$$

Assim, dado um elemento $(a,b) \in I\!R^2$, podemos representa–lo de modo único como:

$$(a,b) = a(1,0) + b(0,1) .$$

Desse modo, temos que

$$T(a,b) = aT(1,0) + bT(0,1) = a(1 - x) + b(1 - x^2) .$$

Portanto, obtemos explicitamente a transformação linear T

$$T(a,b) = (a + b) - ax - bx^2 \qquad \text{para todo} \qquad (a,b) \in I\!R^2 .$$

Exemplo 4.2.9 *A aplicação* $T : I\!R^2 \longrightarrow \mathcal{P}_3(I\!R)$ *tal que*

$$T(1,1) = x \qquad e \qquad T(-1,1) = x - x^3$$

define uma transformação linear de $I\!R^2$ *em* $\mathcal{P}_3(I\!R)$.

Estamos considerando o espaço vetorial $I\!R^2$ com a base ordenada

$$\gamma = \{ (1,1), (-1,1) \} .$$

Assim, dado um elemento $(a,b) \in I\!R^2$, podemos representa–lo de modo único como:

$$(a,b) = \frac{a+b}{2}(1,1) + \frac{b-a}{2}(-1,1) .$$

Desse modo, para todo $(a,b) \in I\!R^2$, temos que

$$T(a,b) = \frac{a+b}{2}T(1,1) + \frac{b-a}{2}T(-1,1) = \frac{a+b}{2}x + \frac{b-a}{2}(x - x^3) .$$

obtendo explicitamente a transformação linear T, dada por:

$$T(a,b) = bx + \frac{a-b}{2}x^3 .$$

Exercícios

Exercício 4.1 *Determine a transformação linear* T *do plano no plano que representa uma rotação anti–horária de* $\dfrac{\pi}{4}$ *seguida por uma dilatação de* $\sqrt{2}$.

Exercício 4.2 *Considere a transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *definida por:*

$$T(x,y) \;=\; (x, -y)\,.$$

Seja K *um triângulo de vértices* $A = (-1,4)$, $B = (3,1)$ *e* $C = (2,6)$*. Faça a representação gráfica da imagem de* K *pela transformação* T*.*

Exercício 4.3 *Considere a transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *associada a matriz*

$$A \;=\; \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Considere o círculo $S = \{\, (x,y) \in \mathbb{R}^2 \;/\; x^2 + y^2 = 1 \,\}$*. Faça a representação gráfica da imagem do círculo* S *pela transformação linear* T*.*

Exercício 4.4 *Considere a transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *definida por:*

$$T(x,y) \;=\; (2x - y\,,\, -x + 2y)\,.$$

Determine uma base para cada um dos seguintes subespaços:

$$W \;=\; \{\, (x,y) \in \mathbb{R}^2 \;/\; T(x,y) \;=\; 3(x,y) \,\}$$

$$U \;=\; \{\, (x,y) \in \mathbb{R}^2 \;/\; T(x,y) \;=\; (x,y) \,\}$$

Exercício 4.5 *Determine a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *tal que*

$$T(1,0,0) \;=\; (0,0,1) \quad , \quad T(1,0,1) \;=\; (1,1,1) \qquad e \qquad T(0,-1,1) \;=\; (1,1,0)\,.$$

Exercício 4.6 *Determine a transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_3(\mathbb{R})$ *tal que*

$$T(1,1) \;=\; x^2 - 1 \qquad e \qquad T(1,-1) \;=\; x^3 + 1\,.$$

Exercício 4.7 *Determine a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathcal{P}_2(\mathbb{R})$ *tal que*

$$T(1,0,0) \;=\; 1 - x\,, \quad T(0,1,0) \;=\; 1 + x \qquad e \qquad T(0,0,1) \;=\; 1 - x^2\,.$$

Exercício 4.8 *Determine a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *tal que*

$$T(1,0,0) \;=\; (1,0), \quad T(0,1,0) \;=\; (1,-1) \qquad e \qquad T(0,0,1) \;=\; (0,1)\,.$$

4.3 Núcleo e Imagem

Definição 4.3.1 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e T uma transformação linear de V em W. O conjunto*

$$Im(T) = \{ w \in W \ / \ w = T(v) \quad \text{para algum} \quad v \in V \}$$

*é denominado **imagem** da transformação T.*

Teorema 4.3.1 *O conjunto $Im(T) \subset W$ é um subespaço vetorial de W.*

Demonstração – Sabemos que $T(0_V) = 0_W$. Assim, $0_W \in Im(T)$.

Agora, tomando $T(u), T(v) \in Im(T)$, tem-se que

$$T(u) + T(v) = T(u + v)$$

como $u + v \in V$ e $T(u + v) \in W$, temos que $T(u) + T(v) \in Im(T)$.

Finalmente, tomando $T(u) \in Im(T)$ e $\lambda \in \mathbb{F}$, temos que

$$\lambda T(u) = T(\lambda u)$$

como $\lambda u \in V$ e $T(\lambda u) \in W$, obtemos que $\lambda T(u) \in Im(T)$. ■

Definição 4.3.2 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e T uma transformação linear de V em W. O conjunto*

$$Ker(T) = \{ v \in V \ / \ T(v) = 0_W \}$$

*é denominado **núcleo** da transformação T.*

Teorema 4.3.2 *O conjunto $Ker(T) \subset V$ é um subespaço vetorial de V.*

Demonstração – Sabemos que $T(0_V) = 0_W$. Assim, $0_V \in Ker(T)$.

Agora, tomando $u, v \in Ker(T)$, temos que

$$T(u) + T(v) = T(u + v) = 0_W .$$

Logo, $u + v \in Ker(T)$.

Finalmente, tomando $u \in Ker(T)$ e $\lambda \in \mathbb{F}$, temos que

$$T(\lambda u) = \lambda T(u) = 0_W .$$

Logo, $\lambda u \in Ker(T)$, o que completa a demonstração. ■

Exemplo 4.3.1 *Determinar o núcleo da transformação linear*

$$\begin{array}{rccl} T: & I\!\!R^2 & \longrightarrow & I\!\!R \\ & (x,y) & \longrightarrow & T(x,y) = 3x + 2y \end{array}$$

Resposta: $\boxed{T(x,y) = 3x + 2y = 0 \implies y = -\dfrac{3}{2}x}$

Exemplo 4.3.2 *Determinar o núcleo da transformação diferenciação*

$$\begin{array}{rccl} T: & \mathcal{C}^1([a,b]) & \longrightarrow & \mathcal{C}([a,b]) \\ & u & \longrightarrow & T(u) = u' \end{array}$$

Note que $T(u)(x) \ = \ u'(x) \quad para \quad x \in [a,b]$.

Resposta: $\boxed{T(u)(x) = u'(x) = 0 \implies Ker(T) \ = \ [1] \ = \ \mathcal{P}_0(I\!\!R)}$

Exemplo 4.3.3 *Determinar o núcleo da transformação linear*

$$\begin{array}{rccl} T: & \mathcal{C}^2([a,b]) & \longrightarrow & \mathcal{C}([a,b]) \\ & u & \longrightarrow & T(u) = u'' \end{array}$$

Note que $T(u)(x) \ = \ u''(x) \quad para \quad x \in [a,b]$.

Resposta: $\boxed{T(u)(x) = u''(x) = 0 \implies Ker(T) \ = \ [1, x] \ = \ \mathcal{P}_1(I\!\!R)}$

Exemplo 4.3.4 *Determinar o núcleo da transformação linear*

$$\begin{array}{rccl} T: & \mathcal{C}^2([a,b]) & \longrightarrow & \mathcal{C}([a,b]) \\ & u & \longrightarrow & T(u) = u'' \ + \ u \end{array}$$

Note que $T(u)(x) \ = \ u''(x) \ + \ u(x) \quad para \quad x \in [a,b]$.

Resposta: $\boxed{T(u)(x) = u''(x) \ + \ u(x) = 0 \implies Ker(T) = [\sin(x), \cos(x)]}$

Exemplo 4.3.5 *Determinar o núcleo da transformação linear*

$$\begin{array}{rccl} T: & \mathcal{C}^2([a,b]) & \longrightarrow & \mathcal{C}([a,b]) \\ & u & \longrightarrow & T(u) = -u'' + u \end{array}$$

Note que $T(u)(x) = -u''(x) + u(x)$ *para* $x \in [a,b]$.

Resposta: $\boxed{T(u)(x) = -u''(x) + u(x) = 0 \implies Ker(T) = [\exp(x), \exp(-x)]}$

Exemplo 4.3.6 *Considere a transformação linear*

$$\begin{array}{rccl} T: & I\!\!R^3 & \longrightarrow & I\!\!R^2 \\ & (x,y,z) & \longrightarrow & T(x,y,z) = (x - 3y + 5z\,,\, -x + 4y - z) \end{array}$$

Determine os subespaços $Im(T)$ *e* $Ker(T)$, *e* $dim(Im(T))$.

A transformação T pode ser escrita da seguinte forma:

$$T(x,y,z) = x(1,-1) + y(-3,4) + z(5,-1) \qquad , \qquad x,\, y,\, z \in I\!\!R\,.$$

Assim, temos que $Im(T) = [(1,-1)\,,\,(-3,4)\,,\,(5,-1)]$. Podemos verificar facilmente que $\{\,(1,-1),\,(-3,4)\,\}$ é uma base para $Im(T)$. Logo, $dim(Im(T)) = 2$.

O núcleo da transformação T é o conjunto solução do sistema linear homogêneo

$$\left\{ \begin{array}{rcrcrcl} x & - & 3y & + & 5z & = & 0 \\ -x & + & 4y & - & z & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcrcrcl} x & - & 3y & + & 5z & = & 0 \\ & & y & + & 4z & = & 0 \end{array} \right.$$

Assim, temos que $Ker(T) = [(-17,-4,1)]$. Logo, $dim(Ker(T)) = 1$.

Exemplo 4.3.7 *Dado o elemento* $c = (1,-1,2) \in I\!\!R^3$. *O núcleo da transformação linear* T *definida pelo produto escalar entre os elementos* $(x,y,z) \in I\!\!R^3$ *e o elemento fixo* c, *veja o Exemplo 4.2.3, é dado por:*

$$Ker(T) = \{\,(x,y,z) \in I\!\!R^3 \;/\; T(x,y,z) = x - y + 2z = 0\,\}\,,$$

isto é, $Ker(T)$ *é um plano contido em* $I\!\!R^3$. *Assim, temos que* $dim(\,Ker(T)\,) = 2$. *Note que a* $Im(T) = I\!\!R$. *Logo,* $dim(\,Im(T)\,) = 1$.

Exemplo 4.3.8 *Considere a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathcal{P}_3(\mathbb{R})$ *definida por:*

$$T(1,0,1) = 2 + x^2 + x^3 \quad , \quad T(0,1,0) = 1 + x^2 \quad e \quad T(0,0,1) = x^2 - x^3 \, .$$

(a) Calcule $T(a,b,c)$ *para a transformação linear* T.

(b) Determine uma base para o subespaço $Im(T)$.

Podemos verificar facilmente que o conjunto

$$\gamma = \{ (1,0,1) \, , (0,1,0) \, , (0,0,1) \}$$

é uma base para o espaço vetorial $\mathbb{R}^3$. Desse modo, tomando um elemento genérico $(a,b,c) \in \mathbb{R}^3$ vamos fazer sua representação em relação à base γ, isto é,

$$(a,b,c) = c_1(1,0,1) + c_2(0,1,0) + c_3(0,0,1) \, ,$$

obtendo $c_1 = a \quad , \quad c_2 = b \quad$ e $\quad c_3 = c - a$. Desse modo, temos que

$$(a,b,c) = a(1,0,1) + b(0,1,0) + (c-a)(0,0,1) \, .$$

(a) Portanto, podemos escrever a transformação linear T da seguinte forma:

$$\begin{aligned} T(a,b,c) &= a\,T(1,0,1) + b\,T(0,1,0) + (c-a)\,T(0,0,1) \\ &= a(\,2 + x^2 + x^3\,) + b(\,1 + x^2\,) + (c-a)(\,x^2 - x^3\,) \\ &= a(\,2 + 2x^3\,) + b(\,1 + x^2\,) + c(\,x^2 - x^3\,) \end{aligned}$$

Assim, determinamos explicitamente a expressão da transformação linear T dada por:

$$T(a,b,c) = a(\,2 + 2x^3\,) + b(\,1 + x^2\,) + c(\,x^2 - x^3\,)$$

para todo $(a,b,c) \in \mathbb{R}^3$.

(b) Considerando $T(a,b,c)$, sabemos que

$$Im(T) = [2 + 2x^3 \, , 1 + x^2 \, , x^2 - x^3] \, .$$

Assim, a partir do sistema de geradores vamos determinar uma base para o subespaço $Im(T)$. Para isso, construímos a matriz A cujas linhas são formadas pelas coordenadas dos elementos do sistema de geradores em relação à base canônica de $\mathcal{P}_3(\mathbb{R})$, dada por:

$$A = \begin{bmatrix} 2 & 0 & 0 & 2 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \, .$$

Em seguida, efetuamos o escalonamento da matriz A, obtendo uma matriz $\widehat{A}$ na forma escalonada, linha equivalente a matriz A, dada por:

$$\widehat{A} = \begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

Portanto, podemos concluir que

$$\{\, 2 + 2x^3 \,,\, 1 + x^2 \,\}$$

é uma base para o subespaço $Im(T)$.

Exercícios

Exercício 4.9 *Considere a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *dada por:*

$$T(x,y,z) = (x - y - z \, , \, 2z - x) \, .$$

Determine uma base para $Ker(T)$ *e uma base para* $Im(T)$.

Exercício 4.10 *Seja* $U \subset \mathbb{M}_3(\mathbb{R})$ *o subespaço das matrizes diagonais. Considere a transformação linear* $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow U$ *definida por:*

$$T(a + bx + cx^2) = \begin{bmatrix} a - b + 2c & 0 & 0 \\ 0 & 2a + b & 0 \\ 0 & 0 & -a - 2b + 2c \end{bmatrix} .$$

Determine uma base para $Ker(T)$ *e uma base para* $Im(T)$.

Exercício 4.11 *Considere a transformação linear* $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ *dada por:*

$$T(p)(x) = p'(x) + \int_0^x p(t)dt \, .$$

Determine $Ker(T)$ *e uma base para* $Im(T)$.

Exercício 4.12 *Considere a transformação linear* $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ *dada por:*

$$T(p)(x) = p'(x) + xp(x) \, .$$

Determine uma base para $Ker(T)$ *e uma base para* $Im(T)$.

Exercício 4.13 *Considere a transformação linear* $T : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ *dada por:*

$$T(p(x)) = -p''(x) + p(x) \, .$$

Determine uma base para $Ker(T)$ *e uma base para* $Im(T)$.

Exercício 4.14 *Considere a transformação linear* $T : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ *dada por:*

$$T(p(x)) = p''(x) + p(x) \, .$$

Determine uma base para $Ker(T)$ *e uma base para* $Im(T)$.

Exercício 4.15 *Considere a transformação linear* $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ *dada por:*

$$T(p(x)) = \int_{-1}^{1} p(x)dx + p'(0) \, .$$

Determine uma base para $Ker(T)$ *e uma base para* $Im(T)$.

4.4 Posto e Nulidade

Definição 4.4.1 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$ *e* T *uma transformação linear de* V *em* W. *Definimos*

(a) *O **Posto** de* T, *que denotamos por* $posto(T)$, *como sendo a dimensão da imagem de* T, *isto é,* $posto(T) = dim(\, Im(T)\,)$.

(b) *A **Nulidade** de* T, *que denotamos por* $Null(T)$, *como sendo a dimensão do núcleo de* T, *isto é,* $Null(T) = dim(\, Ker(T)\,)$.

Definição 4.4.2 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$ *e* T *uma aplicação de* V *em* W. *Dizemos que* T *é uma aplicação **injetora** se, e somente se, para* $u, v \in V$, *com* $u \neq v$ *tem–se que* $T(u) \neq T(v)$. *De modo equivalente,* T *é **injetora** se, e somente se, para* $u, v \in V$, *com* $T(u) = T(v)$ *implica que* $u = v$.

Definição 4.4.3 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$ *e* T *uma aplicação de* V *em* W. *Dizemos que* T *é uma aplicação **sobrejetora** se, e somente se,* $Im(T) = W$, *isto é, para todo* $w \in W$, *existe* $v \in V$ *tal que* $T(v) = w$.

Teorema 4.4.1 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$ *e* T *uma transformação linear de* V *em* W. *Então,* T *é uma aplicação injetora se, e somente se,* $Ker(T) = \{\, 0_V \,\}$.

Demonstração

($\Longrightarrow$) Por hipótese temos que T é injetora, isto é, $T(u) = T(v)$ implica que $u = v$. Vamos mostrar que $Ker(T) = \{\, 0_V \,\}$.

Seja $u \in Ker(T)$, isto é, $T(u) = 0_W$. Como $T(0_V) = 0_W$, temos que $T(u) = T(0_V)$. Pelo fato de T ser injetora, devemos ter $u = 0_V$. Logo, $Ker(T) = \{\, 0_V \,\}$.

($\Longleftarrow$) Por hipótese temos que $Ker(T) = \{\, 0_V \,\}$. Vamos mostrar que T é injetora. Para isso, tomamos $u, v \in V$ tais que $T(u) = T(v)$. Assim, temos que

$$T(u) - T(v) = T(u - v) = 0_W \,.$$

Como $Ker(T) = \{\, 0_V \,\}$, obtemos

$$(u - v) \in Ker(T) \implies u - v = 0_V \implies u = v \,.$$

Logo, T é injetora, o que completa a demonstração. ■

Teorema 4.4.2 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$, com $dim(V) = n$, e $T : V \longrightarrow W$ é uma transformação linear. Então,*

$$dim(\,Ker(T)\,) \;+\; dim(\,Im(T)\,) \;=\; dim(\,V\,)\,.$$

Demonstração – Seja $\{\, v_1\,, \cdots, v_m \,\}$ uma base para $Ker(T)$, com $m \leq n$. Assim, esse conjunto faz parte de uma base para V. Desse modo, existem elementos $v_{m+1}\,, \cdots, v_n$ em V tais que $\{\, v_1\,, \cdots, v_m\,, v_{m+1}\,, \cdots, v_n \,\}$ seja uma base para V. Vamos mostrar que $\{\, T(v_{m+1})\,, \cdots, T(v_n)\,\}$ é uma base para $Im(T)$.

Inicialmente vamos mostrar que $Im(T) = [\, T(v_{m+1})\,, \cdots, T(v_n)\,]$. Seja $w \in Im(T)$, isto é, $w = T(u)$ para algum $u \in V$. Assim, como $u \in V$, tem–se que

$$u \;=\; \sum_{i=1}^{m} c_i\, v_i \;+\; \sum_{i=m+1}^{n} c_i\, v_i\,.$$

Logo, $w = T(u)$ é representado na forma:

$$w \;=\; \sum_{i=1}^{m} c_i\, T(v_i) \;+\; \sum_{i=m+1}^{n} c_i\, T(v_i) \;=\; \sum_{i=m+1}^{n} c_i\, T(v_i)\,,$$

uma vez que $T(v_1) = T(v_2) = \cdots = T(v_m) = 0_W$.

Devemos mostrar que $\{\, T(v_{m+1})\,, \cdots, T(v_n)\,\}$ é linearmente independente em $Im(T)$.

Suponhamos que existam escalares $c_{m+1}\,, \cdots, c_n \in \mathbb{F}$ tais que

$$\sum_{i=m+1}^{n} c_i\, T(v_i) \;=\; 0_W \qquad \Longleftrightarrow \qquad T\left(\sum_{i=m+1}^{n} c_i\, v_i \right) \;=\; T(v) \;=\; 0_W\,.$$

Desse modo, o elemento $v \in Ker(T)$, que é definido por:

$$v \;=\; \sum_{i=m+1}^{n} c_i\, v_i\,, \tag{4.1}$$

pode também ser representado da seguinte forma:

$$v \;=\; \sum_{i=1}^{m} b_i\, v_i\,, \tag{4.2}$$

uma vez que $\{\, v_1\,, \cdots, v_m \,\}$ é uma base para $Ker(T)$. Subtraindo (4.1) de (4.2), obtemos

$$\sum_{i=1}^{m} b_i\, v_i \;-\; \sum_{i=m+1}^{n} c_i\, v_i \;=\; 0_V\,,$$

como $\{ v_1 , \cdots , v_m , v_{m+1} , \cdots , v_n \}$ é linearmente independente em V, isso implica que

$$b_1 = \cdots = b_m = c_{m+1} = \cdots = c_n = 0 .$$

Assim, provamos que $\{ T(v_{m+1}) , \cdots , T(v_n) \}$ é linearmente independente em $Im(T)$. Logo, temos que $dim(Im(T)) = n - m$ e $dim(Ker(T)) = m$.

Portanto, provamos que

$$dim(Ker(T)) + dim(Im(T)) = n = dim(V) ,$$

o que completa a demonstração. ■

Teorema 4.4.3 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$, com $dim(V) = n$, e $T : V \longrightarrow W$ é uma transformação linear. As seguintes afirmações são equivalentes:*

(a) T *é injetora.*

(b) Seja $\{ v_1 , \cdots , v_m \}$ linearmente independente em V. Então,

$$\{ T(v_1) , \cdots , T(v_m) \}$$

é linearmente independente em $Im(T)$.

(c) $dim(Im(T)) = n$.

(d) Seja $\{ v_1 , \cdots , v_n \}$ uma base para V. Então, $\{ T(v_1) , \cdots , T(v_n) \}$ é uma base para $Im(T)$.

Demonstração

Vamos mostrar que *(a)* $\Longrightarrow$ *(b)*. Tomando uma combinação linear nula

$$0_W = \sum_{i=1}^{m} c_i T(v_i) = T\left(\sum_{i=1}^{m} c_i v_i \right)$$

e considerando a hipótese que T é injetora, pelo Teorema 4.4.1, temos que

$$\sum_{i=1}^{m} c_i v_i = 0_V .$$

Como $\{ v_1 , \cdots , v_m \}$ é linearmente independente em V, implica que

$$c_1 = c_2 = \cdots = c_m = 0 .$$

Portanto, $\{ T(v_1) , \cdots , T(v_m) \}$ é linearmente independente em $Im(T)$.

Vamos mostrar que *(b)* $\Longrightarrow$ *(c)*. Seja $\{ v_1 , \cdots , v_n \}$ uma base para V. Assim, temos que $\{ T(v_1) , \cdots , T(v_n) \}$ é linearmente independente em $Im(T)$. Desse modo, temos que $dim(Im(T)) \geq n$. Pelo Teorema 4.4.2, temos que $dim(Im(T)) \leq n$. Portanto, obtemos $dim(Im(T)) = n$.

Vamos mostrar que *(c)* $\Longrightarrow$ *(d)*. Seja $\{ v_1 , \cdots , v_n \}$ uma base para V. Considerando um elemento $w \in Im(T)$, isto é, $w = T(u)$ para algum $u \in V$. Como $u \in V$, temos que

$$u = \sum_{i=1}^{n} c_i v_i \implies w = T\left(\sum_{i=1}^{n} c_i v_i \right) = \sum_{i=1}^{n} c_i T(v_i)$$

Como $dim(Im(T)) = n$, temos que $\{ T(v_1) , \cdots , T(v_n) \}$ é uma base para $Im(T)$.

Finalmente, vamos mostrar que *(d)* $\Longrightarrow$ *(a)*. Seja $\{ v_1 , \cdots , v_n \}$ uma base para V. Tomando $u \in V$, temos que

$$u = \sum_{i=1}^{n} c_i v_i \implies T(u) = \sum_{i=1}^{n} c_i T(v_i)$$

Se $T(u) = 0_W$, isto é, $u \in Ker(T)$, implica que

$$c_1 = c_2 = \cdots = c_n = 0 ,$$

pois $\{ T(v_1), \cdots , T(v_n) \}$ é linearmente independente. Logo, $Ker(T) = \{ 0_V \}$. Pelo Teorema 4.4.1, temos que T é injetora, o que completa a demonstração. ■

Corolário 4.4.1 *Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow W$ uma transformação linear. Se $dim(V) = dim(W)$, então T é injetora se, e somente se, T é sobrejetora.*

Demonstração – A prova pode ficar a cargo do leitor. □

Corolário 4.4.2 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow W$ uma transformação linear injetora. Se $dim(V) = dim(W)$, então T leva base em base.*

Demonstração – Seja $\{ v_1, \cdots , v_n \}$ uma base para V. Pelo Teorema 4.4.3, temos que $\{ T(v_1), \cdots , T(v_n) \}$ é linearmente independente em W. Como $dim(V) = dim(W)$, obtemos que $\{ T(v_1), \cdots , T(v_n) \}$ é uma base para W. ■

Definição 4.4.4 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e T uma aplicação de V em W. Dizemos que T é uma aplicação* ***bijetora*** *se, e somente se, T é injetora e sobrejetora ao mesmo tempo.*

Exemplo 4.4.1 *Considere a seguinte transformação linear*

$$\begin{array}{rccl} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & (x,y,z) & \longrightarrow & T(x,y,z) = -x + y + 2z\,. \end{array}$$

Determine $Ker(T)$, $Null(T) = dim(Ker(T))$, $Im(T)$ e $posto(T) = dim(Im(T))$.

O núcleo da transformação T são os elementos $(x,y,z) \in \mathbb{R}^3$ que satisfazem a equação

$$T(x,y,z) = -x + y + 2z = 0\,.$$

Assim, temos que $Ker(T) = [(1,1,0),(2,0,1)]$. Podemos verificar facilmente que $dim(Ker(T)) = 2$. Portanto, pelo Teorema 4.4.2, tem–se que $dim(Im(T)) = 1$. Logo, como $Im(T) \subset \mathbb{R}$, temos que $Im(T) = \mathbb{R}$.

Exemplo 4.4.2 *A transformação linear definida da forma:*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (2x - y,\, x + y) \end{array}$$

é uma transformação linear injetora.

O núcleo da transformação linear T é conjunto solução do sistema linear homogêneo

$$\left\{\begin{array}{rcrcl} 2x & - & y & = & 0 \\ x & + & y & = & 0 \end{array}\right. \iff \left\{\begin{array}{rcrcl} 2x & - & y & = & 0 \\ & & 3y & = & 0 \end{array}\right.$$

que possui somente a solução trivial $x = 0$ e $y = 0$.

Desse modo, temos que $Ker(T) = \{\, 0_{\mathbb{R}^2} \,\}$. Assim, provamos que a transformação linear T é injetora.

Exemplo 4.4.3 *Determine o núcleo e a imagem do operador linear*

$$\begin{array}{rccl} T: & \mathcal{P}_3(\mathbb{R}) & \longrightarrow & \mathcal{P}_3(\mathbb{R}) \\ & p(x) & \longrightarrow & T(p(x)) = x^2\, p''(x) \end{array}$$

Resposta: $\boxed{Ker(T) = \mathcal{P}_1(\mathbb{R}) \quad \text{e} \quad Im(T) = [x^2,\, x^3]}$

Exemplo 4.4.4 *Considere a transformação linear* $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ *definida por:*

$$T(x, y, z, t) = (x - 2y + t , 2x + y - z , 5y - z - 2t) .$$

(a) Determine uma base para o subespaço $Ker(T)$*.*

(b) Determine uma base para o subespaço $Im(T)$*.*

(c) Determine uma base γ *para o* $\mathbb{R}^4$ *contendo uma base de* $Ker(T)$*.*

Todo elemento $(x, y, z, t) \in Ker(T)$ satisfaz o seguinte sistema linear homogêneo

$$\begin{cases} x - 2y + t = 0 \\ 2x + y - z = 0 \\ 5y - z - 2t = 0 \end{cases}$$

Colocando na forma matricial e escalonando, obtemos

$$\begin{bmatrix} 1 & -2 & 0 & 1 \\ 2 & 1 & -1 & 0 \\ 0 & 5 & -1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 5 & -1 & -2 \\ 0 & 5 & -1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2/5 & 1/5 \\ 0 & 1 & -1/5 & -2/5 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

Assim, temos que a solução do sistema linear homogêneo é dado por:

$$x = \frac{2}{5}z - \frac{1}{5}t \qquad \text{e} \qquad y = \frac{1}{5}z + \frac{2}{5}t .$$

Dessa forma, temos que todo elemento $(x, y, z, t) \in Ker(T)$ é escrito da forma:

$$(x, y, z, t) = z\left(\frac{2}{5}, -\frac{1}{5}, 1, 0\right) + t\left(-\frac{1}{5}, \frac{2}{5}, 0, 1\right) \qquad \text{para} \qquad z, t \in \mathbb{R} .$$

Como os elementos

$$\left(\frac{2}{5}, -\frac{1}{5}, 1, 0\right) \qquad \text{e} \qquad \left(-\frac{1}{5}, \frac{2}{5}, 0, 1\right)$$

são claramente linearmente independentes, temos que

$$\{ (2, -1, 5, 0) , (-1, 2, 0, 5) \}$$

é uma base de $Ker(T)$. Assim, $dim(Ker(T)) = 2$. Logo, $dim(Im(T)) = 2$.

Sabemos que $Im(T)$ é gerado pela imagem dos elementos de uma base do domínio pela transformação linear T. Desse modo, considerando a base canônica para $\mathbb{R}^4$, temos que $Im(T)$ é gerado pelos seguintes elementos:

$$\begin{aligned} T(1, 0, 0, 0) &= (1, 2, 0) \\ T(0, 1, 0, 0) &= (-2, 1, 5) \\ T(0, 0, 1, 0) &= (0, -1, -1) \\ T(0, 0, 0, 1) &= (1, 0, -2) \end{aligned}$$

Como a dimensão da imagem é 2, podemos escolher quaisquer dois desses elementos que sejam linearmente independentes, isto é,

$$Im(T) = [(1,2,0),(-2,1,5),(0,-1,-1),(1,0,-2)] = [(1,2,0),(0,-1,-1)] \,.$$

Assim, temos $\{\,(1,2,0),\,(0,-1,-1)\,\}$ é uma base do subespaço $Im(T)$.

Para obter a base γ de $I\!R^4$ temos que acrescentar mais dois elementos a base de $Ker(T)$, de maneira a termos 4 elementos linearmente independentes no $I\!R^4$. Escolhemos então

$$\gamma = \{\,(2,-1,5,0)\,,\,(-1,2,0,5)\,,\,(1,0,0,0)\,,\,(0,1,0,0)\,\} \,.$$

Quando calculamos T na base canônica de $I\!R^4$, já vimos que os elementos

$$e_1 = (1,0,0,0) \qquad \text{e} \qquad e_2 = (0,1,0,0)$$

não estavam no subespaço $Ker(T)$, sendo portanto uma boa escolha. Além disso, a equação

$$a_1(2,-1,5,0) + a_2(-1,2,0,5) + a_3(1,0,0,0) + a_4(0,1,0,0) = (0,0,0,0)$$

da origem ao sistema linear homogêneo

$$\begin{cases} 2a_1 - a_2 + 5a_3 = 0 \\ -a_1 + 2a_2 + 5a_4 = 0 \\ a_1 = 0 \\ a_2 = 0 \end{cases}$$

cuja única solução é $a_1 = 0,\; a_2 = 0,\; a_3 = 0,\; a_4 = 0$.

Logo, γ é um conjunto de 4 elementos do $I\!R^4$ que é linearmente independente. Como $I\!R^4$ tem dimensão 4, resulta que γ é uma base do $I\!R^4$ e contém a base de $Ker(T)$, conforme foi pedido.

Exemplo 4.4.5 *Seja* T *um operador linear sobre* $\mathbb{R}^3$ *tal que*

$$T(1,0,0) = (1,1,1) \quad , \quad T(0,1,0) = (1,-2,1) \quad e \quad T(0,0,1) = (1,0,-1)$$

Mostre que T *é um operador linear bijetor.*

Para isso basta mostrar que $dim(Im(T)) = 3$, onde

$$Im(T) = [(1,1,1), (1,-2,1), (1,0,-1)] .$$

Assim, pelo Teorema do núcleo e da imagem, temos que $dim(Ker(T)) = 0$, isto é, $Ker(T) = \{ 0_{\mathbb{R}^3} \}$. Portanto, T é um operador injetor e sobrejetor, isto é, T é um operador bijetor.

Exemplo 4.4.6 *A transformação linear definida da forma:*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ & (x,y) & \longrightarrow & T(x,y) = (x - y,\ x,\ x + y) \end{array}$$

não é uma transformação linear bijetora.

Podemos verificar facilmente que T é injetora, isto é, $Ker(T) = \{ 0_{\mathbb{R}^2} \}$, e utilizando o Teorema do núcleo e da imagem, obtemos que $dim(Im(T)) = 2$. Logo, $Im(T) \neq \mathbb{R}^3$. Portanto, T não é sobrejetora. Assim, T não é bijetora.

Exemplo 4.4.7 *Mostre que não existem transformações lineares* $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$*, com* $m < n$*, que são bijetoras.*

Considerando $Ker(T) = \{ 0_{\mathbb{R}^m} \}$, isto é, $dim(Ker(T)) = 0$, e utilizando o Teorema do núcleo e da imagem, obtemos $dim(Im(T)) = m < n$. Logo, $Im(T) \neq \mathbb{R}^n$. Portanto, T não é sobrejetora. Assim, podemos concluir que não existem transformações lineares bijetoras nestas condições.

Exemplo 4.4.8 *Determine uma transformação linear*

$$T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$$

tal que $Im(T) = [(1,1,2,1),(2,1,0,1)]$.

Podemos verificar facilmente que $dim(\, Im(T)\,) = 2$. Logo, $dim(\, Ker(T)\,) = 1$. Assim, tomando a base canônica para o $\mathbb{R}^3$, podemos definir uma transformação linear T da seguinte forma:

$$\begin{aligned} T(1,0,0) &= (0,0,0,0) \\ T(0,1,0) &= (1,1,2,1) \\ T(0,0,1) &= (2,1,0,1) \end{aligned}$$

Desse modo, dado um elemento $(x,y,z) \in \mathbb{R}^3$ temos que

$$\begin{aligned} T(x,y,z) &= xT(1,0,0) + yT(0,1,0) + zT(0,0,1) \\ &= (y+2z\,,\, y+z\,,\, 2y\,,\, y+z) \end{aligned}$$

Note que esse problema possui infinitas soluções, pois sua resolução depende da escolha de uma base para o espaço vetorial $\mathbb{R}^3$, e também da escolha do subespaço $Ker(T)$.

Exemplo 4.4.9 *Determine uma transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathcal{P}_2(\mathbb{R})$ *de modo que* $dim(\, Ker(T)\,) = 1$.

Como $dim(Ker(T)) = 1$, devemos ter que $dim(Im(T)) = 2$. Assim, escolhendo a base canônica para o $\mathbb{R}^3$, podemos definir uma transformação linear T da seguinte forma:

$$T(1,0,0) = 0 \quad , \quad T(0,1,0) = x \qquad \text{e} \qquad T(0,0,1) = x^2 \,,$$

onde $Im(T) = [x\,,\, x^2]$. Logo, T não é sobrejetora, pois $Im(T) \neq \mathcal{P}_2(\mathbb{R})$.

Desse modo, dado um elemento $(a,b,c) \in \mathbb{R}^3$ temos que

$$T(a,b,c) = aT(1,0,0) + bT(0,1,0) + cT(0,0,1) = bx + cx^2 \,.$$

Note que esse problema possui infinitas soluções, pois sua resolução depende da escolha de uma base para o espaço vetorial $\mathbb{R}^3$, e também da escolha do subespaço $Im(T)$, através da escolha de uma base.

Exemplo 4.4.10 *Determine um operador linear*

$$T : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$$

tal que $Ker(T) = [(1,0,1,0),(0,1,0,1)]$.

Podemos verificar facilmente que $dim(Ker(T)) = 2$. Assim, pelo Teorema do núcleo e da imagem, temos que $dim(Im(T)) = 2$. Agora vamos escolher uma base para o $\mathbb{R}^4$, contendo a base do núcleo do operador T, da seguinte forma:

$$\gamma = \{ (1,0,1,0),\, (0,1,0,1),\, (0,1,0,0),\, (0,0,1,0) \} .$$

Desse modo, podemos definir um operador linear T da seguinte forma:

$$\begin{aligned} T(1,0,1,0) &= (0,0,0,0) \\ T(0,1,0,1) &= (0,0,0,0) \\ T(0,1,0,0) &= (1,0,0,0) \\ T(0,0,1,0) &= (0,1,0,0) \end{aligned}$$

onde $Im(T) = [(1,0,0,0),(0,1,0,0)]$.

Finalmente, considerando um elemento $(x,y,z,t) \in \mathbb{R}^4$ vamos fazer sua representação em relação à base γ

$$(x,y,z,t) = a(1,0,1,0) + b(0,1,0,1) + c(0,1,0,0) + d(0,0,1,0)$$

obtendo

$$a = x \quad , \quad b = t \quad , \quad c = y - t \quad \text{e} \quad d = z$$

Desse modo, temos que

$$\begin{aligned} T(x,y,z,t) &= xT(1,0,1,0) + tT(0,1,0,1) + (y-t)T(0,1,0,0) + zT(0,0,1,0) \\ &= (y-t)(1,0,0,0) + z(0,1,0,0) \\ &= (y-t,z,0,0) \end{aligned}$$

Note que esse problema possui infinitas soluções, pois sua resolução depende da maneira como completamos a base do subespaço $Ker(T)$ para obter uma base do espaço vetorial $\mathbb{R}^4$, e também de como escolhemos o subespaço $Im(T)$, através da escolha de uma base.

Exercícios

Exercício 4.16 *Determine uma transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *tal que*

$$Ker(T) = \{ (x, y, z) \in \mathbb{R}^3 \ / \ x + y + z = 0 \} .$$

Exercício 4.17 *Sejam* U *e* W *subespaços vetoriais de* $\mathbb{R}^3$ *definidos por:*

$$\begin{aligned} U &= \{ (x, y, z) \in \mathbb{R}^3 \ / \ x + y + z = 0 \} \\ W &= [(1, 0, 1), (0, -1, 1)] \end{aligned}$$

Determine um operador linear T *sobre* $\mathbb{R}^3$ *tal que* $Im(T) = U$ *e* $Ker(T) = U \cap W$.

Exercício 4.18 *Seja* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *a transformação linear definida por:*

$$T(2, 1) = (3, 0, 2) \qquad e \qquad T(1, 2) = (1, 1, 0) .$$

Determine uma transformação linear $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *tal que* $Ker(P) = Im(T)$.

Exercício 4.19 *Verifique se é Falsa ou Verdadeira cada uma das afirmações abaixo, justificando sua resposta.*

(a) Existe uma transformação linear $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ *que é injetora.*

(b) Existe uma transformação linear $T : \mathbb{R}^4 \longrightarrow \mathcal{P}_2(\mathbb{R})$ *que é sobrejetora.*

(c) Existe uma transformação linear $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_2(\mathbb{R})$ *que é bijetora.*

Exercício 4.20 *Sejam* $u, v \in \mathbb{R}^2$ *tais que* $\beta = \{ u, v \}$ *é uma base para* $\mathbb{R}^2$. *Considere uma transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^n$, *para* $n \geq 2$. *Mostre que somente uma das seguintes alternativas se verifica:*

(a) $\{ T(u), T(v) \}$ *é linearmente independente.*

(b) $dim(Im(T)) = 1$.

(c) $Im(T) = \{ 0_{\mathbb{R}^n} \}$.

Exercício 4.21 *Sejam* V *um espaço vetorial de dimensão finita, com* $dim(V) = n$, *e* $T : V \longrightarrow V$ *uma transformação linear tal que* $Im(T) = Ker(T)$. *Mostre que* n *é par. Considerando* $V = \mathbb{R}^4$, *dê exemplo de uma transformação linear com essas propriedades.*

Exercício 4.22 *Determine uma transformação linear* $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{M}_2(\mathbb{R})$ *que satisfaça simultaneamente as seguintes condições:*

(a) O elemento $p(x) = (1 + x^2) \in Ker(T)$.

(b) O elemento $q(x) = 1 \notin Ker(T)$.

(c) O elemento $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \in Im(T)$.

Exercício 4.23 *Determine explicitamente a expressão de uma transformação linear*

$$T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$$

satisfazendo simultaneamente as seguintes condições:

(a) O elemento $p(x) = (1 + x) \in Ker(T)$.

(b) O elemento $q(x) = x \notin Ker(T)$.

(c) $Im(T) = [(1, 1, 1)]$.

Exercício 4.24 *Sejam* V *um espaço vetorial de dimensão finita e* $T : V \longrightarrow \mathbb{R}^3$ *uma transformação linear não–nula. Se* $dim(Ker(T)) = 2$, *determine as possíveis dimensões de* V. *Se* T *é sobrejetora, qual a dimensão de* V ? *Considerando* $V = \mathbb{R}^3$, *dê exemplo de uma transformação linear com essas propriedades, se possível.*

Exercício 4.25 *Considere* U, V *e* W *espaços vetoriais de dimensão finita. Sejam* $T : U \longrightarrow V$ *e* $P : V \longrightarrow W$ *transformações Lineares. Mostre que*

(a) se T *e* P *são injetoras, então* $dim(U) \leq dim(V) \leq dim(W)$.

(b) se T *e* P *são sobrejetoras, então* $dim(U) \geq dim(V) \geq dim(W)$.

4.5 Espaços Vetoriais Isomorfos

Definição 4.5.1 *Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $I\!F$. Uma transformação linear $T : V \longrightarrow W$ bijetora, isto é, injetora e sobrejetora ao mesmo tempo, é denominada um* ***isomorfismo*** *de V em W. Quando existe um isomorfismo de V em W, dizemos que eles são* ***isomorfos****, ou que V é* ***isomorfo*** *a W. Um isomorfismo $T : V \longrightarrow V$ é denominado um* ***automorfismo*** *de V.*

Pelo resultado do Teorema 4.4.2, podemos observar que espaços isomorfos devem ter a mesma dimensão. Desse modo, pelo Corolário 4.4.2, um isomorfismo leva base em base. Assim, do ponto de vista da Álgebra Linear, espaços isomorfos são considerados idênticos, mesmo que os elementos e as operações definidas nesses espaços sejam bem diferentes.

Seja $T : V \longrightarrow W$ é um isomorfismo, isto é, T é uma transformação linear bijetora. Então, para cada elemento $w \in W$ podemos fazer a associação $T(v) \longrightarrow w$ para um único elemento $v \in V$. Desse modo, temos uma nova aplicação de W em V, tendo em vista que não teremos $T(v_1) = T(v_2)$ com $v_1 \neq v_2$, uma vez que T é injetora.

Essa nova aplicação, que vamos denotar por T^{-1}, $T^{-1} : W \longrightarrow V$ é denominada **aplicação inversa** de T. Desse modo, temos que

$$T^{-1}(T(v)) = v \quad \text{e} \quad T(T^{-1}(w)) = w \quad \text{para todo} \quad v \in V \quad \text{e} \quad w \in W .$$

Na seção 4.7 vamos estudar com mais detalhes as transformações inversas e teremos a oportunidade de mostrar que a aplicação inversa de uma transformação linear também é linear. Além disso, mostraremos que se T é um isomorfismo, então T^{-1} também é um isomorfismo denominado **isomorfismo inverso**.

Exemplo 4.5.1 *Considere V o subespaço de $I\!M_2(I\!R)$ definido da seguinte forma:*

$$V = \left\{ A \in I\!M_2(I\!R) \;/\; A = \begin{bmatrix} a & a+b \\ 0 & c \end{bmatrix} \;,\; a,\, b,\, c \in I\!R \right\} .$$

Construa um isomorfismo de V em $I\!R^3$.

Podemos verificar facilmente que $dim(V) = 3$ e que os elementos

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad A_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

formam uma base ordenada para V. Desse modo, a aplicação $T : V \longrightarrow I\!R^3$ definida por $T(A) = (a,\, a+b,\, c)$ é um isomorfismo. Note que T leva a base ordenada $\gamma = \{ A_1 , A_2 , A_3 \}$ de V na base ordenada $\beta = \{ (1,1,0) , (0,1,0) , (0,0,1) \}$ do $I\!R^3$.

Exemplo 4.5.2 *Seja* $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ *definida por* $T(a,b) = a + (a+b)x$. *Podemos verificar facilmente que* T *é um isomorfismo e que o isomorfismo inverso* T^{-1} *de* $\mathcal{P}_1(\mathbb{R})$ *em* $\mathbb{R}^2$ *é definido por* $T^{-1}(a + bx) = (a, b - a)$. *Assim, os espaços vetoriais* $\mathbb{R}^2$ *e* $\mathcal{P}_1(\mathbb{R})$ *são isomorfos.*

Teorema 4.5.1 *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$, *com* $dim(V) = n$. *Então,* V *é isomorfo ao espaço vetorial* $\mathbb{F}^n$.

Demonstração – Seja $\beta = \{ v_1, \cdots, v_n \}$ uma base ordenada para V. Vamos definir uma aplicação T de V em $\mathbb{F}^n$ da seguinte forma:

$$
\begin{aligned}
T : \ V &\longrightarrow \mathbb{F}^n \\
u &\longrightarrow T(u) = (c_1, \cdots, c_n)
\end{aligned}
$$

onde $c_1, \cdots, c_n \in \mathbb{F}$ são as coordenadas do elemento u com relação à base β.

Podemos verificar facilmente que T é uma transformação linear. Pelo Teorema 3.7.1, temos que T é uma transformação linear injetora e pelo Teorema do núcleo e da imagem obtemos que $Im(T) = \mathbb{F}^n$. Logo, T é sobrejetora. Portanto, T é uma transformação linear bijetora, o que completa a demonstração. ■

Teorema 4.5.2 *Sejam* V *e* W *espaços vetoriais de dimensão finita sobre o corpo* $\mathbb{F}$. *Então,* V *e* W *são isomorfos se, e somente se,* $dim(V) = dim(W)$.

Demonstração

($\Longrightarrow$) Seja $T : V \longrightarrow W$ um isomorfismo, isto é, $Ker(T) = \{ 0_V \}$ e $Im(T) = W$. Pelo Teorema do núcleo e da imagem, temos que

$$
dim(V) = dim(Ker(T)) + dim(Im(T)) \quad \Longrightarrow \quad dim(V) = dim(W)
$$

($\Longleftarrow$) Sejam $\beta = \{ v_1, \cdots, v_n \}$ uma base para V e $\gamma = \{ w_1, \cdots, w_n \}$ uma base para W. Vamos definir uma transformação linear $T : V \longrightarrow W$ da seguinte forma:

$$
T\left(\sum_{i=1}^{n} c_i \, v_i \right) = \sum_{i=1}^{n} c_i \, T(v_i) = \sum_{i=1}^{n} c_i \, w_i \, ,
$$

onde estamos definindo $T(v_i) = w_i$ para $i = 1, \cdots, n$.

Podemos verificar facilmente que $Ker(T) = \{ 0_V \}$. De fato, considere um elemento $v \in Ker(T)$ que é escrito de modo único como:

$$v = \sum_{i=1}^{n} c_i v_i \qquad \Longrightarrow \qquad T(v) = \sum_{i=1}^{n} c_i w_i = 0_W \,,$$

como $\{ w_1 , \cdots , w_n \}$ é linearmente independente, implica em

$$c_1 = \cdots = c_n = 0 \,.$$

Portanto, provamos que T é uma transformação injetora.

Finalmente, pelo Teorema do núcleo e da imagem, tem–se que

$$dim(V) = dim(Ker(T)) + dim(Im(T)) \quad \Longrightarrow \quad dim(Im(T)) = n \,.$$

Assim, $Im(T) = W$. Logo, T é uma transformação sobrejetora.

Portanto, mostramos que T é um isomorfismo de V em W. ∎

Exemplo 4.5.3 *Em muita situações identificamos o espaço vetorial $I\!R^n$ com o espaço vetorial das matrizes reais de ordem $n \times 1$. Esta identificação é feita, muitas vezes, sem que façamos alguma referência ao fato que esses espaços vetoriais sejam isomorfos. Uma ilustração muito simples dessa idéia, é quando estamos trabalhando com **espaço solução** de um **sistema linear homogêneo**, veja Teorema 2.9.8, considerando a representação matricial para sistemas lineares. Desse modo, existe um isomorfismo evidente entre o espaço vetorial real $I\!R^n$ e o espaço vetorial real $I\!M_{n\times1}(I\!R)$, definido da seguinte forma:*

$$(x_1 , \cdots , x_i , \cdots , x_n) \in I\!R^n \qquad \longrightarrow \qquad \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} \in I\!M_{n\times1}(I\!R) \,.$$

Assim, sempre que conveniente vamos fazer a identificação do espaço vetorial real $I\!R^n$ com o espaço vetorial real $I\!M_{n\times1}(I\!R)$, em geral, por simplicidade de notação.

Exemplo 4.5.4 *De modo análogo ao Exemplo 4.5.3, podemos construir um isomorfismo entre o espaço vetorial complexo $\mathbb{C}^n$ e o espaço vetorial complexo $I\!M_{n\times1}(\mathbb{C})$, uma vez que $dim(\mathbb{C}^n) = dim(I\!M_{n\times1}(\mathbb{C})) = n$. Assim, sempre que conveniente vamos fazer a identificação do espaço vetorial complexo $\mathbb{C}^n$ com o espaço vetorial complexo $I\!M_{n\times1}(\mathbb{C})$, em geral, por simplicidade de notação.*

Exemplo 4.5.5 *Considerando* $\mathbb{C}^2$ *como um espaço vetorial real, do Exemplo 3.6.13, sabemos que o conjunto* $\gamma = \{ z_1, z_2, z_3, z_4 \}$, *onde*

$$z_1 = (1,0) \quad , \quad z_2 = (i,0) \quad , \quad z_3 = (0,1) \qquad e \qquad z_4 = (0,i) \, ,$$

é uma base ordenada para $\mathbb{C}^2$. *Logo, temos que* $dim(\mathbb{C}^2) = 4$. *Considere agora o espaço vetorial real* $\mathbb{R}^4$ *com a base canônica* $\beta = \{ e_1, e_2, e_3, e_4 \}$, *onde*

$$\begin{aligned} e_1 &= (1,0,0,0) \\ e_2 &= (0,1,0,0) \\ e_3 &= (0,0,1,0) \\ e_4 &= (0,0,0,1) \end{aligned}$$

De acordo com o Teorema 4.5.2, podemos definir um isomorfismo T *de* $\mathbb{C}^2$ *em* $\mathbb{R}^4$ *da seguinte forma:*

$$T(z_i) = e_i \qquad para \qquad i = 1, \cdots, 4 \, .$$

Desse modo, a expressão explicita de $T(z,w)$ *é dada por:*

$$T(z,w) = (a\,,\,b\,,\,c\,,\,d) \, ,$$

onde $(z,w) \in \mathbb{C}^2$ *é escrito como*

$$(z,w) = (a+ib\,,\,c+id) = az_1 + bz_2 + cz_3 + dz_4 \, ,$$

para $a, b, c, d \in \mathbb{R}$, *pois todo elemento de* $\mathbb{C}^2$ *é escrito de modo único como uma combinação linear dos elementos de* γ.

De fato, note que $T(z,w)$ *pode ser escrito da forma:*

$$\begin{aligned} T(z,w) &= T(az_1 + bz_2 + cz_3 + dz_4) \\ &= aT(z_1) + bT(z_2) + cT(z_3) + dT(z_4) \\ &= ae_1 + be_2 + ce_3 + de_4 \\ &= (a\,,\,b\,,\,c\,,\,d) \end{aligned}$$

o que completa a construção do isomorfismo T.

Assim, mostramos que $\mathbb{C}^2$ *como um espaço vetorial real é isomorfo ao espaço vetorial real* $\mathbb{R}^4$.

Exercícios

Exercício 4.26 *Mostre que o operador linear* T *sobre* $\mathbb{R}^3$ *definido por:*

$$T(x,y,z) = (x - 2y\,,\,z\,,\,x+y)$$

é um isomorfismo de $\mathbb{R}^3$.

Exercício 4.27 *Mostre que a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathcal{P}_2(\mathbb{R})$ *definida por:*

$$T(a,b,c) = (a - b) + (c - a)x + (b + c)x^2$$

é um isomorfismo de $\mathbb{R}^3$ *em* $\mathcal{P}_2(\mathbb{R})$.

Exercício 4.28 *Mostre que o espaço vetorial real* $\mathbb{R}^2$ *é isomorfo ao subespaço* S *do espaço vetorial real* $\mathbb{R}^3$ *definido por:*

$$S = \{\,(x,y,z) \in \mathbb{R}^3 \;/\; x - y + 2z = 0\,\}\,,$$

exibindo um isomorfismo T *de* $\mathbb{R}^2$ *em* S.

Exercício 4.29 *Mostre que o espaço vetorial real* $\mathbb{R}^2$ *é isomorfo a qualquer subespaço* S *do espaço vetorial real* $\mathbb{R}^3$ *tal que* $dim(S) = 2$, *exibindo um isomorfismo* T *de* $\mathbb{R}^2$ *em* S.

Exercício 4.30 *Mostre que o espaço vetorial real* $\mathbb{R}^3$ *é isomorfo ao subespaço* S *do espaço vetorial real* $\mathbb{M}_2(\mathbb{R})$ *definido por:*

$$S = \left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \;/\; a,\,b,\,c \in \mathbb{R} \right\},$$

exibindo um isomorfismo T *de* $\mathbb{R}^3$ *em* S.

Exercício 4.31 *Seja* V *um espaço vetorial real, com* $dim(V) \geq 3$. *Mostre que o espaço vetorial real* $\mathbb{R}^3$ *é isomorfo a qualquer subespaço* S *do espaço vetorial* V *tal que* $dim(S) = 3$, *exibindo um isomorfismo* T *de* $\mathbb{R}^3$ *em* S.

Exercício 4.32 *Mostre que o espaço vetorial real* $\mathbb{R}^3$ *é isomorfo ao espaço vetorial real* $\mathbb{M}_{3\times1}(\mathbb{R})$, *exibindo um isomorfismo* T *de* $\mathbb{R}^3$ *em* $\mathbb{M}_{3\times1}(\mathbb{R})$.

Exercício 4.33 *Mostre que* $\mathbb{C}^2$ *como um espaço vetorial real é isomorfo ao espaço vetorial real* $\mathbb{M}_2(\mathbb{R})$, *exibindo um isomorfismo* T *de* $\mathbb{C}^2$ *em* $\mathbb{M}_2(\mathbb{R})$.

4.6 Álgebra das Transformações Lineares

Definição 4.6.1 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$*. Denotamos por* $L(V,W)$ *o conjunto de todas as transformações lineares de* V *em* W*, isto é,*

$$L(V,W) = \{ T : V \longrightarrow W \ / \ T \text{ é uma transformação linear} \} .$$

Definição 4.6.2 *Dadas as transformações lineares* $T, P \in L(V,W)$*. Definimos a* **adição de transformações** $T + P : V \longrightarrow W$ *da seguinte forma:*

$$(T + P)(v) = T(v) + P(v) \quad ; \quad \forall\, v \in V .$$

A aplicação assim definida é também uma transformação linear.

Definição 4.6.3 *Dada a transformação linear* $T \in L(V,W)$ *e o escalar* $\lambda \in \mathbb{F}$*. Definimos a* **multiplicação de uma transformação por um escalar** $\lambda T : V \longrightarrow W$ *da seguinte forma:*

$$(\lambda T)(v) = \lambda T(v) \quad ; \quad \forall\, v \in V .$$

A aplicação assim definida é também uma transformação linear.

Teorema 4.6.1 $L(V,W)$ *é um espaço vetorial sobre o corpo* $\mathbb{F}$ *com relação as operações de adição de transformações lineares e multiplicação por escalar definidas acima.*

Demonstração

Inicialmente, devemos mostrar que a operação de adição tem as seguintes propriedades:

(A_1) Comutatividade. $T + P = P + T \quad ; \quad \forall\ T,\ P \in L(V,W)$.

(A_2) Associatividade. $T + (P + S) = (T + P) + S \quad ; \quad \forall\ T,\ P,\ S \in L(V,W)$.

(A_3) Elemento Neutro. A transformação linear nula $O_L : V \longrightarrow W$ é tal que

$$T + O_L = T \quad ; \quad \forall\ T \in L(V,W) ,$$

isto é, $O_L(v) = 0_W$ para todo $v \in V$.

(A_4) Elemento Simétrico. Para toda transformação linear $T \in L(V,W)$ existe a transformação $(-T) \in L(V,W)$ tal que $T + (-T) = O_L$.

Finalmente, devemos mostrar que a operação de multiplicação por escalar tem as seguintes propriedades:

(M_1) Associatividade. $(\alpha\,\beta)\,T \;=\; \alpha\,(\beta\,T) \;\;;\;\; \forall\; T \in L(V,W) \;\;$ e $\;\; \forall\; \alpha,\, \beta \in I\!\!F$.

(M_2) Distributividade para a Adição de Elementos.
$\alpha\,(T \,+\, S) \;=\; \alpha\,T \,+\, \alpha\,S \;\;;\;\; \forall\; T,\, S \in L(V,W) \;\;$ e $\;\; \forall\; \alpha \in I\!\!F$.

(M_3) Distributividade para a Multiplicação por Escalar.
$(\alpha \,+\, \beta)\,T \;=\; \alpha\,T \,+\, \beta\,T \;\;;\;\; \forall\; T \in L(V,W) \;\;$ e $\;\; \forall\; \alpha,\, \beta \in I\!\!F$.

(M_4) Elemento Identidade. $1_{I\!\!F}\,T \;=\; T \;\;;\;\; \forall\; T \in L(V,W)$.

As provas das propriedades acima podem ficar a cargo do leitor. □

Teorema 4.6.2 *Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $I\!\!F$, com dimensões n e m, respectivamente. Então, o espaço vetorial $L(V,W)$ tem dimensão finita e $dim(\,L(V,W)\,) \,=\, nm$.*

Demonstração – A demonstração pode ser vista na referência [12]. □

Definição 4.6.4 *Sejam U, V e W espaços vetoriais sobre o corpo $I\!\!F$. Considere as transformações lineares $T : U \longrightarrow V$ e $P : V \longrightarrow W$. Definimos a composição das transformações P e T, que denotamos por $S \,=\, P \circ T : U \longrightarrow W$, da seguinte forma:*

$$S(u) \;=\; (\,P \circ T\,)(u) \;=\; P(\,T(u)\,) \,\in\, W \qquad ; \qquad \forall\; u \in U\,.$$

Teorema 4.6.3 *A aplicação $S \,=\, P \circ T$ é uma transformação linear de U em W.*

Demonstração – A prova pode ficar a cargo do leitor. □

Definição 4.6.5 *Seja V um espaço vetorial sobre o corpo $I\!\!F$. Um* ***operador linear*** *sobre V é uma transformação linear de V em V.*

Definição 4.6.6 *Seja V um espaço vetorial sobre o corpo $I\!\!F$. Denotamos por $L(V)$ o conjunto de todos os operadores lineares sobre V, isto é,*

$$L(V) \;=\; \{\, T : V \longrightarrow V \;\;/\;\; T \text{ é um operador linear} \,\}\,.$$

Pelo Teorema 4.6.1, temos que $L(V)$ é um espaço vetorial sobre o corpo $I\!\!F$ com as operações de adição de operadores e multiplicação por escalar definidas para transformações lineares.

Exemplo 4.6.1 *Considere as transformações lineares* $T : I\!R^2 \longrightarrow I\!R$ *definida por:* $T(x,y) = x - 3y$ *e* $P : I\!R \longrightarrow I\!R$ *definida por:* $P(x) = 4x$*. Determine, se possível, as seguintes aplicações* $T + P$*,* $T \circ P$ *e* $P \circ T$*.*

Note que as aplicações $T + P$ e $T \circ P$ não estão definidas. Assim, podemos definir somente a aplicação $P \circ T$ que é dada por:

$$(P \circ T)(x,y) = P(T(x,y)) = P(x - 3y) = 4x - 12y \, .$$

Exemplo 4.6.2 *Sejam* P *e* T *operadores lineares sobre* $I\!R^2$ *definidos por:*

$$T(x,y) = (2x, x - y) \qquad e \qquad P(x,y) = (x + y, 4x) \, .$$

Determine os seguintes operadores $P + T$*,* $P \circ T$ *e* $T \circ P$*.*

Neste caso, todas as aplicações estão definidas. Assim, temos que

- $(T + P)(x,y) = T(x,y) + P(x,y) = (3x + y, 5x - y) \, .$

- $(P \circ T)(x,y) = P(T(x,y)) = P(2x, x - y) = (3x - y, 8x) \, .$

- $(T \circ P)(x,y) = T(P(x,y)) = P(x + y, 4x) = (2x + 2y, y - 3x) \, .$

Definição 4.6.7 *No espaço vetorial* $L(U)$ *podemos definir a operação* ***potenciação*** *para expoentes naturais de um operador* $T \in L(U)$ *da seguinte forma:*

$$T^0 = I \quad , \quad T^1 = T \quad , \quad T^2 = T \circ T \quad e \quad T^n = T \circ T^{n-1} \qquad para \qquad n \in I\!N \, .$$

Definição 4.6.8 *Sejam* V *um espaço vetorial sobre o corpo* $I\!F$ *e* T *um operador linear sobre* V*. Dizemos que* T *é um* ***operador idempotente*** *se* $T^2 = T$*, isto é,*

$$(T \circ T)(v) = T(T(v)) = T(v) \qquad para\ todo \qquad v \in V \, .$$

Exemplo 4.6.3 *Considere o espaço vetorial real* $I\!R^3$*. O operador linear de projeção sobre o plano xy*

$$\begin{array}{rccl} T : & I\!R^3 & \longrightarrow & I\!R^3 \\ & (x,y,z) & \longrightarrow & T(x,y,z) = (x,y,0) \end{array}$$

é um operador idempotente.

Definição 4.6.9 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Dizemos que T é um* **operador auto–reflexivo** *se $T^2 = I$, isto é,*

$$(T \circ T)(v) = T(T(v)) = v \qquad \text{para todo} \qquad v \in V .$$

Exemplo 4.6.4 *Considere o espaço vetorial real $\mathbb{R}^2$. O operador linear de reflexão em torno do eixo–ox*

$$\begin{array}{rccl} T : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x, y) & \longrightarrow & T(x, y) = (x, -y) \end{array}$$

é um operador auto–reflexivo.

Definição 4.6.10 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Dizemos que T é um* **operador nilpotente** *se $T^n = 0$ para um certo $n \in \mathbb{N}$, isto é,*

$$(T \circ T^{n-1})(v) = T(T^{n-1}(v)) = 0_V \qquad \text{para todo} \qquad v \in V .$$

Exemplo 4.6.5 *Considere o espaço vetorial real $\mathcal{P}_3(\mathbb{R})$ e o operação de derivação D sobre $\mathcal{P}_3(\mathbb{R})$, isto é, $D(p(x)) = p'(x)$ para $p(x) \in \mathcal{P}_3(\mathbb{R})$. Podemos verificar facilmente que D é um operador nilpotente.*

4.7 Transformação Inversa

Definição 4.7.1 *Sejam* V *e* W *espaços vetoriais sobre o corpo* $\mathbb{F}$ *e* $T : V \longrightarrow W$ *uma transformação linear. A aplicação* $L : Im(T) \subset W \longrightarrow V$ *é denominada **inversa a esquerda** da transformação linear* T *se*

$$(L \circ T)(u) = u \quad ; \quad \forall\ u \in V\,,$$

isto é, $L \circ T$ *é a transformação identidade em* V*.*

A transformação $R : Im(T) \subset W \longrightarrow V$ *é denominada **inversa a direita** de* T *se*

$$(T \circ R)(w) = w \quad ; \quad \forall\ w \in Im(T)\,,$$

isto é, $T \circ R$ *é a transformação identidade em* $Im(T)$*.*

Teorema 4.7.1 *Considere os espaços vetoriais* V *e* W *de dimensão finita sobre o corpo* $\mathbb{F}$*. Seja* $T : V \longrightarrow W$ *uma transformação linear que possui inversa a esquerda* L*. Então,* L *é também inversa a direita de* T*. Além disso,* L *é única.*

Demonstração – Primeiramente devemos observar que, pela Definição 4.7.1, tanto a inversa a esquerda quanto a inversa a direita estão definidas em $Im(T) \subset W$.

Vamos mostrar que L é única. Suponhamos que T possui duas inversas a esquerda L_1 e L_2, isto é, para todo $u \in V$ tem–se que

$$(L_1 \circ T)(u) = u \quad \text{e} \quad (L_2 \circ T)(u) = u\,.$$

Temos que $L_1(w) = u$ e $L_2(w) = u$ para $w \in Im(T)$. Desse modo,

$$L_1(w) - L_2(w) = 0_V \implies (\,L_1 - L_2\,)(w) = 0_V \quad ; \quad w \in Im(T)\,.$$

Logo, $L_1 = L_2$, o que prova a unicidade de L.

Agora vamos mostrar que L é também a inversa a direita de T. Seja $w \in Im(T)$. Assim, basta mostrar que $(T \circ L)(w) = w$.

Como $w \in Im(T)$, temos que $w = T(u)$ para algum $u \in V$. Como L é a inversa a esquerda de T , tem–se que

$$u = L(T(u)) = L(w) \implies T(u) = T(L(w))\,.$$

Portanto, $T(L(w)) = (T \circ L)(w) = w$, uma vez que $w = T(u)$. Assim, provamos que L é a inversa a direita de T, o que completa a demonstração. ∎

Exemplo 4.7.1 *Considere a transformação linear* $T : I\!R \longrightarrow I\!R^2$ *definida por:*

$$T(x) \;=\; (x,\, 2x)\,.$$

Primeiramente vamos observar que o subespaço $Im(T) \;=\; [(1,2)]$*. Logo, todo elemento* $(x,y) \in Im(T)$ *é da forma* $(x, 2x)$*. Além disso, a transformação linear* T *é* ***injetora****.*

Temos que a ***inversa a esquerda*** $L : Im(T) \subset I\!R^2 \longrightarrow I\!R$ *é definida por:*

$$L(x,y) \;=\; \frac{x}{3} \;+\; \frac{y}{3}\,.$$

Assim, temos que $(L \circ T)(x) \;=\; x$ *para todo* $x \in I\!R$*.*

A ***inversa a direita*** $R : Im(T) \subset I\!R^2 \longrightarrow I\!R$ *é definida por:*

$$R(x,y) \;=\; \frac{x}{3} \;+\; \frac{y}{3}\,.$$

Assim, temos que $(T \circ R)(x,y) \;=\; (x,y)$ *para todo* $(x,y) \in Im(T) \;\subset\; I\!R^2$*.*

Desse modo, apresentamos um excelente exemplo para o Teorema 4.7.1, mostrando que a existência e unicidade da inversa a esquerda implica na existência e unicidade da inversa a direita e que são iguais.

Exemplo 4.7.2 *Considere a transformação linear* $T : I\!R^2 \longrightarrow I\!R$ *definida por:*

$$T(x,y) \;=\; x \;+\; y\,.$$

Primeiramente vamos observar que a transformação linear T *não é injetora, pois*

$$Ker(T) \;=\; \{\, (x,y) \in I\!R^2 \;\;/\;\; x \;=\; -y \,\}\,.$$

Logo, $dim(\, Ker(T)\,) \;=\; 1$*, o que implica na* $Im(T) \;=\; I\!R$*, pelo Teorema do núcleo e da imagem.*

A transformação T *não possui* ***inversa a esquerda****, entretanto, podemos apresentar vários exemplos de* ***inversa a direita****. Desse modo, podemos tomar como exemplos de* ***inversa a direita*** $R : Im(T) \;=\; I\!R \longrightarrow I\!R^2$ *as seguintes transformações:*

$$R(x) \;=\; \left(\frac{x}{2}, \frac{x}{2}\right) \qquad e \qquad R(x) \;=\; \left(\frac{2x}{3}, \frac{x}{3}\right)\,.$$

Assim, temos que $(T \circ R)(x) \;=\; x$ *para todo* $x \;\in\; Im(T) \;=\; I\!R$*.*

Desse modo, apresentamos um exemplo onde a não existência da inversa a esquerda implica na não unicidade da inversa a direita.

Exemplo 4.7.3 *Considere a transformação linear* $T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *definida por:*

$$T_A(x,y) \;=\; (x+y,\, 2x+y)\,,$$

associada à matriz $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$.

Primeiramente vamos observar que a transformação linear T_A *é* ***injetora****. Além disso, pelo Teorema do núcleo e da imagem, podemos concluir que* $Im(T_A) = \mathbb{R}^2$*. Logo,* T *é um isomorfismo.*

Temos que a ***inversa a esquerda*** $L : Im(T_A) \longrightarrow \mathbb{R}^2$ *é a transformação linear*

$$L(x,y) \;=\; (-x+y,\, 2x-y)$$

associada à matriz $A^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$.

Assim, temos que $(L \circ T)(x,y) = (T \circ L)(x,y) = (x,y)$ *para todo* $(x,y) \in \mathbb{R}^2$.

Exemplo 4.7.4 *Considere a transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *definida por:*

$$T(x,y) \;=\; (x+y,\, y-x,\, x+3y)\,.$$

Note que a transformação linear T *é* ***injetora*** *e que o subespaço* $Im(T)$ *tem como uma base o conjunto* $\{\, (1,-1,1),\, (1,1,3)\,\}$.

Temos que a ***inversa a esquerda*** $L : Im(T) \subset \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *é definida por:*

$$L(x,y,z) \;=\; \left(\frac{-3x-3y+2z}{2},\, \frac{x+y}{2} \right)\,.$$

Assim, temos que $(L \circ T)(x,y) = (x,y)$ *para todo* $(x,y) \in \mathbb{R}^2$*. Além disso, temos que* $(T \circ L)(x,y,z) = (x,y,z)$ *para todo* $(x,y,z) \in Im(T)$.

Desse modo, desenvolvemos dois exemplos como ilustração do Teorema 4.7.2, sobre a existência da inversa a esquerda, que vamos apresentar a seguir.

Teorema 4.7.2 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e T uma transformação linear de V em W. Então, T possui inversa a esquerda se, e somente se, T é injetora.*

Demonstração

($\Longrightarrow$) Seja $L : Im(T) \longrightarrow V$ a inversa a esquerda de T. Sejam $u, v \in V$ com $T(u) = T(v)$. Assim, temos que

$$u = L(T(u)) = L(T(v)) = v$$

Logo, T é injetora.

($\Longleftarrow$) Por hipótese temos T injetora. Seja $w \in Im(T)$, isto é, $w = T(u)$ para um único $u \in V$. Assim, definimos a transformação $L : Im(T) \longrightarrow V$ da seguinte forma $L(w) = u$ tal que $T(u) = w$. Logo, L é a inversa a esquerda de T, o que completa a demonstração. ■

Definição 4.7.2 *Seja $T : V \longrightarrow W$ uma transformação linear injetora. A única transformação inversa a esquerda de T, que também é a transformação inversa a direita, é denotada por T^{-1}. Dizemos que a transformação T é **invertível**, e chamamos a transformação $T^{-1} : Im(T) \subset W \longrightarrow V$ de **transformação inversa** de T.*

Teorema 4.7.3 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow W$ um isomorfismo. Então, $T^{-1} : W \longrightarrow V$ é uma transformação linear.*

Demonstração – Sejam $w_1, w_2 \in W$ e $\lambda \in \mathbb{F}$. Queremos mostrar que

$$T^{-1}(\lambda w_1 + w_2) = \lambda T^{-1}(w_1) + T^{-1}(w_2) .$$

Sejam $u_1 = T^{-1}(w_1)$ e $u_2 = T^{-1}(w_2)$, isto é, u_1 e u_2 são os únicos elementos em V tais que $T(u_1) = w_1$ e $T(u_2) = w_2$.

Como T é uma transformação linear, temos que

$$T(\lambda u_1 + u_2) = \lambda T(u_1) + T(u_2) = \lambda w_1 + w_2 .$$

Desse modo, $\lambda u_1 + u_2$ é o único elemento em V que é levado pela transformação linear T no elemento $\lambda w_1 + w_2$ em W. Portanto

$$T^{-1}(\lambda w_1 + w_2) = \lambda u_1 + u_2 = \lambda T^{-1}(w_1) + T^{-1}(w_2) ,$$

provando que T^{-1} é uma transformação linear. ■

Proposição 4.7.1 *Sejam V e W espaços vetoriais de dimensão finitas sobre o corpo F e T um isomorfismo de V em W. Então, $T^{-1} : W \longrightarrow V$ é também um isomorfismo.*

Demonstração – Pelo Teorema 4.7.3, sabemos que T^{-1} é uma transformação linear. Assim, temos devemos mostrar que T^{-1} é bijetora.

Sejam $w_1, w_2 \in W$ tais que $T^{-1}(w_1) = T^{-1}(w_2) = v$. Desse modo, temos $T(v) = w_1$ e $T(v) = w_2$. Logo, $w_1 = w_2$, pois T é uma aplicação. Portanto, T^{-1} é injetora.

Para mostrar que T^{-1} é sobrejetora basta observar que $Ker(T^{-1}) = \{ 0_W \}$, pois T^{-1} é injetora, e aplicar o Teorema do núcleo e da imagem,

$$dim(\, Ker(T^{-1})\,) \; + \; dim(\, Im(T^{-1})\,) \; = \; dim(W) \,,$$

obtendo que $dim(\, Im(T^{-1})\,) = dim(W) = dim(V)$, pois V e W são isomorfos. Assim, $Im(T^{-1}) = V$, o que completa a demonstração. ■

Sejam V e U espaços vetoriais de dimensão finita sobre o corpo $I\!F$. Podemos observar que V é isomorfo a V, pois o operador identidade I_V é um isomorfismo de V em V.

Além disso, se V é isomorfo a U por meio de um isomorfismo T, então U é isomorfo a V por meio do isomorfismo inverso T^{-1}.

Exemplo 4.7.5 *Considere V, U e W espaços vetoriais de dimensão finita sobre o corpo $I\!F$. Sejam $T : V \longrightarrow U$ um isomorfismo, isto é, V é isomorfo a U, e $P : U \longrightarrow W$ um isomorfismo, isto é, U é isomorfo a W. Podemos verificar facilmente que $P \circ T : V \longrightarrow W$ é um isomorfismo, isto é, V é isomorfo a W.*

Primeiramente observamos que $dim(V) = dim(U)$ e $dim(U) = dim(W)$, pois são isomorfos. Logo, $dim(V) = dim(W)$. Além disso, sabemos que $Ker(T) = \{ 0_V \}$ e que $Ker(P) = \{ 0_U \}$. Assim, basta mostrar que $Ker(P \circ T) = \{ 0_V \}$.

Tomando um elemento $v \in Ker(P \circ T)$, isto é,

$$(P \circ T)(v) = P(T(v)) = 0_W \quad \Longrightarrow \quad T(v) = 0_U \quad \Longrightarrow \quad v = 0_V \,.$$

Logo, mostramos que $Ker(P \circ T) = \{ 0_V \}$.

Finalmente, podemos concluir que o **isomorfismo** é uma **relação de equivalência** sobre a classe dos espaços vetoriais de mesma dimensão.

Exemplo 4.7.6 *Considere* $T \in L(\mathbb{R}^2)$ *definido por:*

$$T(x,y) = (x+y,\, x-y)\,.$$

Mostre que T *é um isomorfismo, e determine o isomorfismo inverso.*

Inicialmente vamos mostrar que T é um isomorfismo sobre $\mathbb{R}^2$. Para isso, basta mostrar que $Ker(T) = \{\, 0_{\mathbb{R}^2} \,\}$, e em seguida utilizar o Teorema do núcleo e da imagem para mostrar que $Im(T) = \mathbb{R}^2$, isto é, T é um operador sobrejetor. Assim, provamos que T é um operador bijetor.

Para determinar o núcleo do operador T, temos que encontrar os elementos $(x,y) \in \mathbb{R}^2$ tais que

$$T(x,y) = (x+y,\, x-y) = (0,0) \quad\Longrightarrow\quad \begin{cases} x + y = 0 \\ x - y = 0 \end{cases}.$$

Podemos verificar facilmente que o sistema linear homogêneo acima possui somente a solução trivial $x = y = 0$. Logo, $Ker(T) = \{\, 0_{\mathbb{R}^2} \,\}$.

Finalmente, vamos determinar o isomorfismo inverso. Dado um elemento $(x,y) \in \mathbb{R}^2$, supomos que $T^{-1}(x,y) = (a,b)$, então $(x,y) = T(a,b)$. Assim, obtemos o sistema linear

$$(x,y) = (a + b,\, a - b) \quad\Longrightarrow\quad \begin{cases} a + b = x \\ a - b = y \end{cases}$$

que possui uma única solução

$$a = \frac{x+y}{2} \qquad \text{e} \qquad b = \frac{x-y}{2}\,.$$

Desse modo, temos que o isomorfismo inverso T^{-1} é definido por:

$$T^{-1}(x,y) = \left(\frac{x+y}{2}, \frac{x-y}{2} \right)$$

para todo $(x,y) \in \mathbb{R}^2$.

Exemplo 4.7.7 *Considere* $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ *a transformação linear tal que*

$$T(1,-1) \;=\; 2 + x \qquad e \qquad T(0,1) \;=\; x - 1\,.$$

Mostre que T *é um isomorfismo de* $\mathbb{R}^2$ *em* $\mathcal{P}_1(\mathbb{R})$*, e determine o isomorfismo inverso* T^{-1} *de* $\mathcal{P}_1(\mathbb{R})$ *em* $\mathbb{R}^2$*.*

Podemos verificar facilmente que $\gamma = \{ (1,-1), (0,1) \}$ é uma base para o $\mathbb{R}^2$. Para isso, basta mostrar que γ é linearmente independente.

Considere a combinação linear nula

$$a(1,-1) \;+\; b(0,1) \;=\; (0,0) \qquad \iff \qquad \begin{cases} a \phantom{{}+b} = 0 \\ -a + b = 0 \end{cases}.$$

Assim, obtemos $a = b = 0$. Logo, γ é linearmente independente em $\mathbb{R}^2$.

Vamos tomar um elemento genérico $(a,b) \in \mathbb{R}^2$ e representa–lo com relação à base ordenada γ, isto é, vamos representa–lo através da combinação linear

$$(a,b) \;=\; c(1,-1) \;+\; d(0,1) \;=\; (c, -c+d)\,.$$

Assim, obtemos o seguinte sistema linear

$$\begin{cases} c \phantom{{}+d} = a \\ -c + d = b \end{cases}$$

que possui uma única solução $c = a$ e $d = a + b$.

Desse modo, temos que o elemento $(a,b) \in \mathbb{R}^2$ é escrito de modo único como:

$$(a,b) \;=\; a(1,-1) \;+\; (a+b)(0,1)\,.$$

Finalmente, fazendo

$$\begin{aligned} T(a,b) &= aT(1,-1) \;+\; (a+b)T(0,1) \\ &= a(2+x) \;+\; (a+b)(x-1) \\ &= (a - b) \;+\; (2a + b)x \end{aligned}$$

obtemos a transformação linear T dada por:

$$T(a,b) \;=\; (a - b) \;+\; (2a + b)x$$

para todo $(a,b) \in \mathbb{R}^2$.

Para mostrar que T é um isomorfismo, basta mostrar que $Ker(T) = \{ 0_{\mathbb{R}^2} \}$. Assim, considerando um elemento $(a, b) \in Ker(T)$, temos que

$$T(a,b) = (a - b) + (2a + b)x = 0_{\mathcal{P}_1(\mathbb{R})} \qquad \text{para todo} \qquad x \in \mathbb{R} .$$

Assim, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a - b = 0 \\ 2a + b = 0 \end{cases}$$

que possui somente a solução trivial $a = b = 0$. Logo, T é um isomorfismo.

Vamos encontrar o isomorfismo inverso. Dado um elemento $p(x) = a + bx \in \mathcal{P}_1(\mathbb{R})$, supomos que $T^{-1}(a + bx) = (c, d)$. Assim, temos que $T(c, d) = a + bx$, isto é,

$$(c - d) + (2c + d)x = a + bx ,$$

obtendo o seguinte sistema linear

$$\begin{cases} c - d = a \\ 2c + d = b \end{cases}$$

que possui uma única solução

$$c = \frac{a+b}{3} \qquad \text{e} \qquad d = \frac{b-2a}{3} .$$

Portanto, temos que o isomorfismo inverso T^{-1} é dado por:

$$T^{-1}(a + bx) = \left(\frac{a+b}{3}, \frac{b-2a}{3} \right) ,$$

para todo $p(x) = a + bx \in \mathcal{P}_1(\mathbb{R})$.

Exemplo 4.7.8 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T e P operadores lineares sobre V tais que $T \circ P = P \circ T$. Então,*

$$Ker(T) + Ker(P) \subset Ker(T \circ P) = Ker(P \circ T) .$$

Considerando $v \in Ker(T) + Ker(P)$, isto é, $v = u + w$ com $u \in Ker(T)$ e $w \in Ker(P)$. Vamos mostrar que $v \in Ker(T \circ P)$.

Para isso, vamos avaliar $(T \circ P)(u + w)$

$$\begin{aligned}
(T \circ P)(u + w) &= (T \circ P)(u) \; + \; (T \circ P)(w) \\
&= T(P(u)) \; + \; T(P(w)) \; = \; T(P(u)) \; + \; T(0_V) \; = \; (T \circ P)(u) \\
&= (P \circ T)(u) \; = \; P(T(u)) \; = \; P(0_V) \; = \; 0_V
\end{aligned}$$

Assim, mostramos que $v \in Ker(T \circ P) \; = \; Ker(P \circ T)$, o que completa a nossa prova.

Exemplo 4.7.9 *O operador linear* $T \in L(I\!R^3)$ *dado por:*

$$T(x, y, z) \; = \; (x - y, \, 2y, \, y + z)$$

é invertível e T^{-1} *é dado por:* $T^{-1}(x, y, z) \; = \; \dfrac{1}{2}(2x + y, \, y, \, 2z - y)$.

Para mostrar que T é um operador invertível, basta mostrar que T é um operador bijetor, isto é, T é um automorfismo de $I\!R^3$. Para isso, vamos determinar o núcleo do operador T, isto é, vamos encontrar os elementos $(x, y, z) \in I\!R^3$ tais que

$$T(x, y, z) \; = \; (x - y, \, 2y, \, y + z) \; = \; (0, 0, 0) \quad \Longrightarrow \quad Ker(T) \; = \; \{ \, (0, 0, 0) \, \} \, .$$

Portanto, pelo teorema do núcleo e da imagem temos que $Im(T) \; = \; I\!R^3$. Logo, T é um automorfismo de $I\!R^3$.

Dado $(x, y, z) \in I\!R^3$, supomos que $T^{-1}(x, y, z) \; = \; (a, b, c)$, então $(x, y, z) \; = \; T(a, b, c)$. Assim, obtemos o sistema linear

$$(x, y, z) \; = \; (a - b, \, 2b, \, b + c) \quad \Longrightarrow \quad \left\{ \begin{array}{ccccccc} a & - & b & & & = & x \\ & & 2b & & & = & y \\ & & b & + & c & = & z \end{array} \right.$$

que possui uma única solução

$$a \; = \; x + \frac{y}{2} \quad , \quad b \; = \; \frac{y}{2} \qquad \text{e} \qquad c \; = \; z - \frac{y}{2} \, .$$

Desse modo, temos que o automorfismo inverso T^{-1} é definido por:

$$T^{-1}(x, y, z) \; = \; \frac{1}{2}(2x + y \, , \, y \, , \, 2z - y)$$

para todo $(x, y, z) \in I\!R^3$.

Exemplo 4.7.10 *Seja* $T \in L(V)$ *um operador linear tal que* $T^2 - T + I = 0$*. Então,* T *é um operador invertível e* $T^{-1} = I - T$*.*

Tomando a hipótese, temos que

$$T^2 - T + I = 0 \iff (T - I)T + I = 0 \iff (I - T)T = I\,.$$

Assim, mostramos que T é invertível, e que o automorfismo inverso é $T^{-1} = I - T$.

Exemplo 4.7.11 *Considere a transformação linear de derivação*

$$\begin{aligned} D: \ \mathcal{P}_3(\mathbb{R}) &\longrightarrow \mathcal{P}_2(\mathbb{R}) \\ p(x) &\longrightarrow D(p(x)) = p'(x) \end{aligned}$$

e a transformação linear de integração

$$\begin{aligned} T: \ \mathcal{P}_2(\mathbb{R}) &\longrightarrow \mathcal{P}_3(\mathbb{R}) \\ p(x) &\longrightarrow T(p(x)) = \int_0^x p(t)dt\,. \end{aligned}$$

Mostre que $D \circ T$ *é o operador identidade sobre* $\mathcal{P}_2(\mathbb{R})$ *e que* $T \circ D$ *é diferente do operador identidade sobre* $\mathcal{P}_3(\mathbb{R})$*. Determine o núcleo e a imagem do operador* $T \circ D$*.*

Inicialmente vamos determinar o operador $T \circ D$ sobre $\mathcal{P}_3(\mathbb{R})$

$$(T \circ D)(p)(x) = T(D(p))(x) = \int_0^x p'(t)dt = p(x) - p(0)\,.$$

Logo, $(T \circ D)(p)(x) \neq p(x)$.

Podemos verificar facilmente que $Im(T \circ D) = [x,\, x^2,\, x^3]$ e $Ker(T \circ D) = [1]$.

Finalmente, vamos determinar o operador $D \circ T$ sobre $\mathcal{P}_2(\mathbb{R})$. Utilizando o Primeiro Teorema Fundamental do Cálculo ([2] página 202), obtemos

$$(D \circ T)(p)(x) = D(T(p))(x) = \frac{d}{dx}\int_0^x p(t)dt = p(x)\,.$$

Logo, $D \circ T$ é o operador identidade sobre $\mathcal{P}_2(\mathbb{R})$.

Exemplo 4.7.12 *Considere os seguintes operadores lineares sobre o espaço vetorial* $\mathcal{P}(I\!R)$

$$D(p(x)) = p'(x) \qquad e \qquad T(p(x)) = xp(x)\,.$$

Podemos verificar facilmente que o operador $D \circ T - T \circ D$ *é o operador identidade sobre* $\mathcal{P}(I\!R)$*, isto é,* $(D \circ T - T \circ D)(p(x)) = p(x)$.

Para uma simples verificação considere o polinômio $p(x) = 2 + 3x - 4x^2 + 6x^3$.

Exemplo 4.7.13 *A transformação linear* $T : \mathcal{P}_1(I\!R) \longrightarrow I\!R^2$ *dada por:*

$$T(a + bx) = (a,\, a + b)$$

é um isomorfismo e o isomorfismo inverso $T^{-1} : I\!R^2 \longrightarrow \mathcal{P}_1(I\!R)$ *é dado por:*

$$T^{-1}(c, d) = c + (d - c)x\,.$$

Exemplo 4.7.14 *Sejam* V *um espaço vetorial sobre o corpo* $I\!F$ *e* T *um operador idempotente sobre* V*. Então,* $V = Ker(T) \oplus Im(T)$.

Inicialmente vamos provar que

$$Ker(T) \cap Im(T) = \{\, 0_V \,\}\,.$$

Para isso, consideramos um elemento $v \in Ker(T) \cap Im(T)$, isto é, $T(v) = 0_V$ e $v \in Im(T)$. Assim, existe um elemento $u \in V$ tal que $v = T(u)$.

Tomando $T(v)$, obtemos

$$0_V = T(v) = T(T(u)) = T^2(u) = T(u)\,.$$

Como $v = T(u)$, temos que $v = 0_V$. Assim, provamos que

$$Ker(T) \cap Im(T) = \{\, 0_V \,\}\,.$$

Finalmente, vamos mostrar que $V = Ker(T) + Im(T)$. Dado um elemento $v \in V$, vamos mostrar que $v = u + w$ com $u \in Ker(T)$ e $w \in Im(T)$.

Tomando $w = T(v) \in Im(T)$ e $u = v - w$. Vamos mostrar que $u \in Ker(T)$. De fato, fazendo

$$T(u) = T(v - w) = T(v) - T(w) = T(v) - T^2(v) = T(v) - T(v) = 0_V\,.$$

Logo, temos que $u = (\, v - T(v)\,) \in Ker(T)$, o que completa a nossa prova.

Exemplo 4.7.15 *Considere* $T : I\!\!R^2 \longrightarrow \mathcal{P}_1(I\!\!R)$ *a transformação linear tal que*

$$T(1,1) \;=\; 1 \,-\, x \qquad e \qquad T(1,-1) \;=\; 1 \,+\, 3x\,.$$

Mostre que T *é um isomorfismo de* $I\!\!R^2$ *em* $\mathcal{P}_1(I\!\!R)$. *Determine explicitamente a expressão do isomorfismo inverso* $T^{-1}(a_0 \,+\, a_1 x)$.

Vamos mostrar que T é um isomorfismo mostrando que $\{\,(1,1)\,,\,(1,-1)\,\}$ é uma base de $I\!\!R^2$, e que $\{\,1-x\,,\,1+3x\,\}$ é uma base de $\mathcal{P}_1(I\!\!R)$.

Dessa forma teremos que T leva base em base, demonstrando que T é um isomorfismo. Como $I\!\!R^2$ e $\mathcal{P}_1(I\!\!R)$ tem dimensão 2, basta mostrar que esses dois conjuntos são linearmente independente, pois cada um deles tem 2 elementos.

Inicialmente, vamos mostrar que o conjunto

$$\gamma \;=\; \{\,(1,1)\,,\,(1,-1)\,\}$$

é linearmente independente. Para isso, consideramos a equação

$$a(1,1) \;+\; b(1,-1) \;=\; 0$$

que resulta no sistema linear homogêneo

$$\begin{cases} a \;+\; b \;=\; 0 \\ a \;-\; b \;=\; 0 \end{cases}$$

com matriz

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

que possui determinante não–nulo. Logo invertível. Assim, o sistema linear homogêneo possui somente a solução trivial $a \;=\; 0$ e $b \;=\; 0$, mostrando que os elementos do conjunto γ são linearmente independentes.

Para com conjunto $\alpha \;=\; \{\,1-x\,,\,1+3x\,\}$ procedemos da mesma maneira. A equação

$$a(1 \,-\, x) \;+\; b(1 \,+\, 3x) \;=\; 0$$

da origem ao sistema homogêneo

$$\begin{cases} a \;+\; b \;=\; 0 \\ -a \;+\; 3b \;=\; 0 \end{cases}$$

com matriz

$$\begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$$

que possui determinante não–nulo. Logo invertível. Assim, o sistema homogêneo possui somente a solução trivial $a = 0$ e $b = 0$, mostrando que os elementos do conjunto α são linearmente independentes. Portanto, mostramos que a transformação linear T é um isomorfismo.

Vamos agora determinar o isomorfismo inverso. Já sabemos que

$$T^{-1}(1-x) = (1,1) \qquad \text{e} \qquad T^{-1}(1+3x) = (1,-1) .$$

Inicialmente, vamos escrever um polinômio genérico $p(x) = a_0 + a_1 x$ na base

$$\alpha = \{ 1-x \, , \, 1+3x \} ,$$

isto é,

$$p(x) = a_0 + a_1 x = m(1-x) + n(1+3x) .$$

Assim, podemos escrever o isomorfismo inverso T^{-1} da seguinte maneira:

$$\begin{aligned} T^{-1}(a_0 + a_1 x) &= T^{-1}(m(1-x) + n(1+3x)) \\ &= mT^{-1}(1-x) + nT^{-1}(1+3x) \\ &= m(1,1) + n(1,-1) = (m+n \, , \, m-n) \end{aligned}$$

Como

$$m(1-x) + n(1+3x) = (m+n) + (-m+3n)x ,$$

obtemos o sistema linear

$$\begin{cases} m + n = a_0 \\ -m + 3n = a_1 \end{cases}$$

que possui uma única solução

$$m = \frac{3a_0 - a_1}{4} \qquad \text{e} \qquad n = \frac{a_0 + a_1}{4} .$$

Desse modo, obtemos

$$m + n = a_0 \qquad \text{e} \qquad m - n = \frac{a_0 - a_1}{2} ,$$

e podemos concluir

$$T^{-1}(a_0 + a_1 x) = \left(a_0 \, , \, \frac{a_0 - a_1}{2} \right) .$$

Exemplo 4.7.16 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Então, T é um operador idempotente se, e somente se, $I - T$ é um operador idempotente.*

Inicialmente fazendo uso da hipótese $T^2 = T$, obtemos

$$(I - T)^2 = I - 2T + T^2 = I - T$$

Logo, $I - T$ é um operador idempotente.

Finalmente, fazendo uso da hipótese que $I - T$ é um operador idempotente, obtemos

$$I - T = (I - T)^2 = I - 2T + T^2$$

Logo, $T^2 = T$, isto é, T é um operador idempotente, o que completa a nossa prova.

Uma aplicação direta dos resultados do Exemplo 4.7.14 e do Exemplo 4.7.16, será feita quando da apresentação de **projeção ortogonal** em subespaço de dimensão finita, que vamos estudar com todo detalhe na seção 5.15.

Exercícios

Exercício 4.34 *Determine o operador linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *que representa a reflexão em torno da reta* $y = -x$*, utilizando conceitos de geometria analítica. Mostre que* T *é um operador auto–reflexivo.*

Exercício 4.35 *Determine o operador linear* $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *que representa a projeção no plano* $S = \{ (x, y, z) \in \mathbb{R}^3 \ / \ x + y + z = 0 \}$*, utilizando conceitos de geometria analítica. Mostre que* P *é um operador idempotente.*

Exercício 4.36 *Seja* $T \in L(\mathbb{R}^2)$ *tal que* $T(1, 0) = (2, 1)$ *e* $T(0, 1) = (1, 4)$*.*

(a) Determine $T(2, 4)$*.*

(b) Determine o elemento $(x, y) \in \mathbb{R}^2$ *tal que* $T(x, y) = (2, 3)$*.*

(c) Mostre que T *é um automorfismo de* $\mathbb{R}^2$*.*

Exercício 4.37 *Seja* $T \in L(\mathbb{R}^3)$ *tal que*

$$T(1, 0, 0) = (1, 1, 1) \quad , \quad T(0, 1, 0) = (1, 0, 1) \quad e \quad T(0, 1, 2) = (0, 0, 4) \, .$$

T *é um automorfismo de* $\mathbb{R}^3$ *? Em caso afirmativo, determine o automorfismo inverso.*

Exercício 4.38 *Dado o elemento* $q(x) = 3 + x \in \mathcal{P}_1(\mathbb{R})$*. Considere o operador linear* T *sobre* $\mathcal{P}_2(\mathbb{R})$ *definido por:* $T(p)(x) = q(x)p'(x) + 2p(x)$ *e a transformação linear* $P : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$ *definida por:* $P(a + bx + cx^2) = (a + b, \, c, \, a - b)$*. Determine a transformação linear* $P \circ T$ *e verifique se é um isomorfismo de* $\mathcal{P}_2(\mathbb{R})$ *em* $\mathbb{R}^3$*.*

Exercício 4.39 *Considere a transformação linear* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *definida por:*

$$T(x, y) = (2x \, , \, x - y \, , \, y)$$

e a transformação linear $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *definida por:*

$$P(x, y, z) = (y - z \, , \, z - x) \, .$$

(a) Determine a transformação linear $P \circ T$ *e uma base para o subespaço* $Ker(\, P \circ T \,)$*.*

(b) Determine a transformação linear $T \circ P$ *e uma base para o subespaço* $Im(\, T \circ P \,)$*.*

(c) Verifique se $T \circ P$ *é um automorfismo de* $\mathbb{R}^3$*. Em caso afirmativo, determine o automorfismo inverso.*

4.8 Representação Matricial

Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{F}$. Vamos considerar $\beta = \{ v_1 , \cdots , v_n \}$ uma base ordenada para V e $\gamma = \{ w_1 , \cdots , w_m \}$ uma base ordenada para W. Seja $T : V \longrightarrow W$ uma transformação linear, pelo Teorema 4.2.1, sabemos que T fica bem determinada pelo seu efeito sobre os elementos da base β de V. Assim, cada elemento $T(v_j) \in W$ pode ser escrito de modo único da forma:

$$T(v_j) = \sum_{i=1}^{m} t_{ij} w_i \qquad \text{para} \qquad j = 1, \cdots , n ,$$

onde os escalares $t_{1j} , \cdots , t_{mj} \in \mathbb{F}$ são as coordenadas do elemento $T(v_j)$ com relação à base ordenada γ de W . Desse modo, a transformação linear T fica bem determinada pela matriz $m \times n$ cuja j–ésima coluna são as coordenadas do elemento $T(v_j)$ com relação à base ordenada γ de W. Vamos denotar essa matriz por $[T]_{\gamma}^{\beta} = [t_{ij}]$, que é a representação matricial da transformação linear T com relação à base ordenada β de V e a base ordenada γ de W.

Teorema 4.8.1 *Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$, β uma base ordenada para V, γ uma base ordenada para W e T uma transformação linear de V em W. Então, para todo $v \in V$ temos que*

$$[T(v)]_{\gamma} = [T]_{\gamma}^{\beta} [v]_{\beta} .$$

Demonstração – Considere o elemento $v \in V$, que é escrito de modo único na forma:

$$v = \sum_{j=1}^{n} c_j v_j .$$

Assim, temos que

$$T(v) = \sum_{j=1}^{n} c_j T(v_j) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} t_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} t_{ij} c_j \right) w_i ,$$

onde

$$\sum_{j=1}^{n} t_{ij} c_j$$

é a i–ésima coordenada do elemento $T(v) \in W$ com relação a base ordenada γ, que é o produto da i–ésima linha da matriz $[T]_{\gamma}^{\beta}$ pelo vetor coordenada $[v]_{\beta}$ do elemento $v \in V$ com relação a base ordenada β. Portanto, mostramos que

$$[T(v)]_{\gamma} = [T]_{\gamma}^{\beta} [v]_{\beta} ,$$

o que completa a demonstração. ■

Teorema 4.8.2 *Sejam V e W espaços vetoriais de dimensão finita sobre o corpo $\mathbb{F}$, β uma base ordenada para V, γ uma base ordenada para W, T e P transformações lineares de V em W. Então,*

(a) $[T + P]^{\beta}_{\gamma} = [T]^{\beta}_{\gamma} + [P]^{\beta}_{\gamma}$.

(b) $[\lambda T]^{\beta}_{\gamma} = \lambda [T]^{\beta}_{\gamma}$ *para todo* $\lambda \in \mathbb{F}$.

Demonstração – Sejam $\beta = \{ v_1 , \cdots , v_n \}$ e $\gamma = \{ w_1 , \cdots , w_m \}$ bases ordenadas para V e W, respectivamente.

Pelo Teorema 4.8.1, sabemos que existem escalares a_{ij} e b_{ij}, para $i = 1, \cdots , n$ e $j = 1, \cdots , m$, tais que

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{e} \quad P(v_j) = \sum_{i=1}^{m} b_{ij} w_i$$

para $j = 1, \cdots , n$. Assim, $[T]^{\beta}_{\gamma} = A = [a_{ij}]$ e $[P]^{\beta}_{\gamma} = B = [b_{ij}]$.

Desse modo, temos que

$$\begin{aligned} (T + P)(v_j) &= T(v_j) + P(v_j) \\ &= \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i \\ &= \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i \end{aligned}$$

Portanto, mostramos que $[T + P]^{\beta}_{\gamma} = [T]^{\beta}_{\gamma} + [P]^{\beta}_{\gamma}$.

Finalmente, temos que

$$(\lambda T)(v_j) = \lambda T(v_j) = \lambda \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} (\lambda a_{ij}) w_i \, .$$

Assim, provamos que $[\lambda T]^{\beta}_{\gamma} = \lambda [T]^{\beta}_{\gamma}$, o que completa a demonstração. ■

Exemplo 4.8.1 *Considerando o espaço vetorial* V *de dimensão finita sobre o corpo* $\mathbb{F}$*, e a transformação identidade* $I_V : V \longrightarrow V$*, onde o domínio está com a base ordenada* β *e contra–domínio está com a base ordenada* γ*, podemos verificar facilmente que*

$$[I_V]^{\beta}_{\gamma} = [I]^{\beta}_{\gamma} .$$

Assim, justifica a notação $[I]^{\beta}_{\gamma}$ *para a matriz de mudança da base* β *para a base* γ*.*

Teorema 4.8.3 *Considere* U*,* V *e* W *espaços vetoriais com dimensão finita sobre o corpo* $\mathbb{F}$*, com as respectivas bases* γ *,* β *e* α*. Sejam* $T : U \longrightarrow V$ *e* $P : V \longrightarrow W$ *transformações lineares. Então, a matriz da transformação linear* $S = P \circ T : U \longrightarrow W$ *é dada por:*

$$[P \circ T]^{\gamma}_{\alpha} = [P]^{\beta}_{\alpha} \, [T]^{\gamma}_{\beta}$$

Demonstração – A prova é feita aplicando o Teorema 4.8.1 na transformação linear $S = P \circ T$, e utilizando também o resultado do Teorema 2.1.9. □

Corolário 4.8.1 *Sejam* V *e* W *espaços vetoriais de dimensão finitas sobre o corpo* $\mathbb{F}$*,* β *uma base ordenada para* V*,* γ *uma base ordenada para* W *e* T *um isomorfismo de* V *em* W*. Então,* $[T^{-1}]^{\gamma}_{\beta} = (\, [T]^{\beta}_{\gamma} \,)^{-1}$*.*

Demonstração – Aplicando o resultado Teorema 4.8.3, e considerando a transformação identidade I_V sobre V e a transformação identidade I_W sobre W, obtemos

$$[I_V]^{\beta}_{\beta} = [T^{-1} \circ T]^{\beta}_{\beta} = [T^{-1}]^{\gamma}_{\beta} \, [T]^{\beta}_{\gamma}$$

$$[I_W]^{\gamma}_{\gamma} = [T \circ T^{-1}]^{\gamma}_{\gamma} = [T]^{\beta}_{\gamma} \, [T^{-1}]^{\gamma}_{\beta}$$

onde $[I_V]^{\beta}_{\beta} = [I_W]^{\gamma}_{\gamma} = I_n$ é a matriz identidade de ordem n, com

$$dim(V) = dim(W) = n ,$$

o que completa a demonstração. ■

Corolário 4.8.2 *Sejam* V *e* W *espaços vetoriais de mesma dimensão sobre o corpo* $\mathbb{F}$*,* β *uma base ordenada para* V*,* γ *uma base ordenada para* W *e* T *uma transformação linear de* V *em* W*. Então,* T *é um isomorfismo se, e somente se,* $[T]^{\beta}_{\gamma}$ *é uma matriz não–singular.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 4.8.2 *Considere os espaços vetoriais* $\mathcal{P}_1(\mathbb{R})$ *e* $\mathbb{R}^2$ *com as respectivas bases canônicas* $\gamma = \{1\,,\,x\}$ *e* $\beta = \{(1,0),\ (0,1)\}$. *Para a transformação linear* T *de* $\mathcal{P}_1(\mathbb{R})$ *em* $\mathbb{R}^2$ *definida por:*

$$T(a + bx) = (a\,,\,a+b)\;,$$

temos que

$$[T]^{\gamma}_{\beta} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad e \qquad [T^{-1}]^{\beta}_{\gamma} = (\,[T]^{\gamma}_{\beta}\,)^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

Podemos verificar facilmente que $T^{-1} : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ *é definido por:*

$$T^{-1}(c,d) = c + (d-c)x\,,$$

como vimos no Exemplo 4.7.13.

Exemplo 4.8.3 *Considere a transformação linear*

$$\begin{aligned} T:\quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longrightarrow T(x,y) = (3x+y,\,x+3y) \end{aligned}$$

Determine a representação matricial de T *com relação à base canônica do* $\mathbb{R}^2$.

Utilizando a notação $[T]^{\beta}_{\beta} = [t_{ij}]$, temos que

$$T(1,0) = (3,1) = t_{11}(1,0) + t_{21}(0,1)$$

$$T(0,1) = (1,3) = t_{12}(1,0) + t_{22}(0,1)$$

Portanto, obtemos

$$[T]^{\beta}_{\beta} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Exemplo 4.8.4 *Considere a transformação linear*

$$\begin{aligned} T:\quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longrightarrow T(x,y) = (3x+y,\,x+3y) \end{aligned}$$

Determine a representação matricial da transformação linear T *com relação à base ordenada* γ *do* $\mathbb{R}^2$ *dada por:*

$$\gamma = \{\,(1,\,1)\,,\,(-1,\,1)\,\}\,.$$

Utilizando a notação $[T]^{\gamma}_{\gamma} = [t_{ij}]$, temos que

$$\begin{aligned} T(1,1) &= (4,4) = t_{11}(1,1) + t_{21}(-1,1) \\ T(-1,1) &= (-2,2) = t_{12}(1,1) + t_{22}(-1,1) \end{aligned}$$

Assim, temos que obter a solução dos seguintes sistemas lineares

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} t_{11} \\ t_{21} \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} t_{12} \\ t_{22} \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

Portanto, obtemos

$$[T]^{\gamma}_{\gamma} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}.$$

Exemplo 4.8.5 *Considere a transformação linear*

$$\begin{aligned} T: \ \mathcal{P}_3(I\!R) &\longrightarrow \mathcal{P}_2(I\!R) \\ p(x) &\longrightarrow T(p(x)) = p'(x) \end{aligned}$$

Determine a matriz $[T]^{\beta}_{\gamma}$*, onde* $\beta = \{ 1, x, x^2, x^3 \}$ *é a base canônica de* $\mathcal{P}_3(I\!R)$ *e* $\gamma = \{ 1, x, x^2 \}$ *é a base canônica de* $\mathcal{P}_2(I\!R)$.

Resposta: $\boxed{[T]^{\beta}_{\gamma} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}}$

Exemplo 4.8.6 *Considere a transformação linear*

$$\begin{aligned} T: \quad I\!R^3 &\longrightarrow I\!R^2 \\ (x,y,z) &\longrightarrow T(x,y,z) = (2x + y - z,\, 3x - 2y + 4z) \end{aligned}$$

Determine $[T]^{\beta}_{\gamma}$*, onde* β *é a base canônica de* $I\!R^3$ *e* γ *é a base canônica de* $I\!R^2$.

Resposta: $\boxed{[T]^{\beta}_{\gamma} = \begin{bmatrix} 2 & 1 & -1 \\ 3 & -2 & 4 \end{bmatrix}}$

Exemplo 4.8.7 *Considere a transformação linear*

$$\begin{array}{rccl} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x,y,z) & \longrightarrow & T(x,y,z) = (2x + y - z\,,\, 3x - 2y + 4z) \end{array}$$

Determine a representação matricial de T *com relação às bases*

$$\beta \;=\; \{\,(1,\,1,\,1)\,,\,(1,\,1,\,0)\,,\,(1,\,0,\,0)\,\} \quad do \quad \mathbb{R}^3$$

$$\gamma \;=\; \{\,(1,\,3)\,,\,(1,\,4)\,\} \quad do \quad \mathbb{R}^2.$$

Resposta: $\boxed{[T]^{\beta}_{\gamma} \;=\; \begin{bmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{bmatrix}}$

Exemplo 4.8.8 *Considere a transformação linear*

$$\begin{array}{rccl} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longrightarrow & T(x,y,z) = (3x + z\,,\, -2x + y\,,\, -x + 2y + 4z) \end{array}$$

Determine a matriz $[T]^{\beta}_{\beta}$, *onde* β *é a base canônica do* $\mathbb{R}^3$.

Resposta: $\boxed{[T]^{\beta}_{\beta} \;=\; \begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}}$

Exemplo 4.8.9 *Considere a transformação linear*

$$\begin{array}{rccl} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longrightarrow & T(x,y,z) = (3x + z\,,\, -2x + y\,,\, -x + 2y + 4z) \end{array}$$

Determine a representação matricial da transformação linear T *com relação à base ordenada* γ *do* $\mathbb{R}^3$ *dada por:*

$$\gamma \;=\; \{\,(1,\,0,\,1)\,,\,(-1,\,2,\,1)\,,\,(2,\,1,\,1)\,\}\,.$$

Resposta: $\boxed{[T]^{\gamma}_{\gamma} \;=\; \frac{1}{4}\begin{bmatrix} 17 & 35 & 22 \\ -3 & 15 & -6 \\ -2 & -14 & 0 \end{bmatrix}}$

Exemplo 4.8.10 *Considere a transformação linear*

$$\begin{array}{rcccl} T: & \mathcal{P}_2(I\!R) & \longrightarrow & \mathcal{P}_3(I\!R) & \\ & p & \longrightarrow & q & = T(p) \end{array}$$

com

$$q(x) = T(p(x)) = \int_0^x p(t)dt\,.$$

Determine a matriz $[T]_\gamma^\beta$, *onde* $\beta = \{\,1,\, x,\, x^2\,\}$ *é a base canônica de* $\mathcal{P}_2(I\!R)$ *e* $\gamma = \{\,1,\, x,\, x^2,\, x^3\,\}$ *é a base canônica de* $\mathcal{P}_3(I\!R)$.

Dado o elemento $p(x) = 2\,x^2 + x$, *determine as coordenadas de* $q(x) = T(p(x))$ *com relação à base canônica* γ.

Resposta: $[T]_\gamma^\beta = \dfrac{1}{6}\begin{bmatrix} 0 & 0 & 0 \\ 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ *e* $[q]_\gamma = [T]_\gamma^\beta\,[p]_\beta = \dfrac{1}{6}\begin{bmatrix} 0 \\ 0 \\ 3 \\ 4 \end{bmatrix}$

Exemplo 4.8.11 *Considere o operador linear* $T : I\!M_2(I\!R) \longrightarrow I\!M_2(I\!R)$ *definido por:*

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a+b & 2b \\ 2c & 3d \end{bmatrix}.$$

Considerando $I\!M_2(I\!R)$ *com a base canônica* $\beta = \{\,A_1,\, A_2,\, A_3,\, A_4\,\}$, *onde*

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad , \quad A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad e \quad A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Determine a matriz $[T]_\beta^\beta$. *Mostre que* T *é um automorfismo de* $I\!M_2(I\!R)$.

Resposta: $[T]_\beta^\beta = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

Exemplo 4.8.12 *Sejam V um espaço vetorial real e $\beta = \{ v_1, v_2, v_3 \}$ uma base ordenada para V. Determine um operador linear T sobre V tal que $T(v_1) = v_2$ e que deixa fixos todos os elementos do subespaço $W \subset V$ definido da seguinte forma:*

$$W = \{ x\,v_1 + y\,v_2 + z\,v_3 \in V \ / \ x - y + z = 0 \quad para \quad x, y, z \in I\!R \} .$$

Determine a matriz $[T]^{\beta}_{\beta}$. T é um automorfismo (isomorfismo) de V ?

Inicialmente, vamos determinar uma base para o subespaço W. Utilizando a condição $x - y + z = 0$, obtemos $x = y - z$. Logo, todo elemento $w \in W$ é escrito como

$$w = y\,(v_1 + v_2) + z\,(v_3 - v_1) \qquad \text{para} \qquad y, z \in I\!R .$$

Portanto, $\{ (v_1 + v_2), (v_3 - v_1) \}$ é uma base para o subespaço W. Impondo a condição que o operador T deixa fixos os elemento de W e que $T(v_1) = v_2$, tem–se que

$$v_1 + v_2 = T(v_1 + v_2) = T(v_1) + T(v_2) = v_2 + T(v_2) \Longrightarrow T(v_2) = v_1$$

$$v_3 - v_1 = T(v_3 - v_1) = T(v_3) - T(v_1) = T(v_3) - v_2 \Longrightarrow T(v_3) = v_3 - v_1 + v_2$$

Portanto, obtemos que a matriz do operador T com relação à base β é dada por:

$$[T]^{\beta}_{\beta} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} .$$

Vamos verificar se T é um automorfismo de V. Para isso, determinamos

$$Ker(T) = \{ v = x\,v_1 + y\,v_2 + z\,v_3 \in V \ / \ T(v) = 0_V \} .$$

Podemos observar que o operador T é escrito da forma:

$$T(v) = x\,T(v_1) + y\,T(v_2) + z\,T(v_3) = x\,v_2 + y\,v_1 + z\,(v_3 - v_1 + v_2) .$$

Assim, impondo que o elemento $v = a\,v_1 + b\,v_2 + c\,v_3 \in Ker(T)$, isto é,

$$T(v) = (b - c)\,v_1 + (a + c)\,v_2 + c\,v_3 = 0_V ,$$

obtemos $a = b = c = 0$, uma vez que $\{ v_1, v_2, v_3 \}$ é linearmente independente.

Portanto, $Ker(T) = \{ 0_V \}$, isto é, T é um operador injetor. Utilizando o Teorema do núcleo e da imagem, obtemos $dim(Im(T)) = dim(V)$. Logo, $Im(T) = V$, isto é, T é um operador sobrejetor. Portanto, T é um automorfismo de V.

Exemplo 4.8.13 *Sejam* T *um operador linear sobre* $\mathbb{R}^4$, $\gamma = \{ v_1, v_2, v_3, v_4 \}$ *uma base ordenada para o espaço vetorial real* $\mathbb{R}^4$ *e o subespaço* $S = [v_1, v_2, v_3]$.

(a) Sabendo que $T(v) = v$ *para todo* $v \in S$ *e* $T(v_4) = v_1 + v_3$, *determine* $[T]^{\gamma}_{\gamma}$.

(b) Sabendo que

$$[I]^{\beta}_{\gamma} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

onde $\beta = \{ e_1, e_2, e_3, e_4 \}$ *é a base canônica de* $\mathbb{R}^4$, *determine* $[T(e_1)]_{\gamma}$.

Sabendo que $T(v) = v$ para todo $v \in S$ e que $T(v_4) = v_1 + v_3$, obtemos

$$\begin{aligned} T(v_1) &= v_1 &&= 1v_1 + 0v_2 + 0v_3 + 0v_4 \\ T(v_2) &= v_2 &&= 0v_1 + 1v_2 + 0v_3 + 0v_4 \\ T(v_3) &= v_3 &&= 0v_1 + 0v_2 + 1v_3 + 0v_4 \\ T(v_4) &= v_1 + v_3 &&= 1v_1 + 0v_2 + 1v_3 + 0v_4 \end{aligned}$$

Portanto, temos que

$$[T]^{\gamma}_{\gamma} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Conhecemos as matrizes $[I]^{\beta}_{\gamma}$ e $[T]^{\gamma}_{\gamma}$, para encontrar $[T(e_1)]_{\gamma}$, vamos determinar inicialmente $[e_1]_{\gamma}$ da seguinte forma:

$$[e_1]_{\gamma} = [I]^{\beta}_{\gamma}\,[e_1]_{\beta} \quad \Longleftrightarrow \quad [e_1]_{\gamma} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Finalmente, calculamos

$$[T(e_1)]_{\gamma} = [T]^{\gamma}_{\gamma}\,[e_1]_{\gamma} \quad \Longleftrightarrow \quad [T(e_1)]_{\gamma} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

o que completa a resolução.

Exemplo 4.8.14 *Considere o operador linear* $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ *definido por:*

$$T(p(x)) = p(x) + (1 + x)p'(x)\,.$$

Verifique se T *é um automorfismo de* $\mathcal{P}_3(\mathbb{R})$*, e determine a matriz* $[T]^{\beta}_{\beta}$*, onde* β *é a base canônica de* $\mathcal{P}_3(\mathbb{R})$*.*

Vamos verificar se o operador T é um automorfismo de $\mathcal{P}_3(\mathbb{R})$. Para isso, devemos determinar o subespaço $Ker(T)$, isto é,

$$Ker(T) = \{\, p(x) \in \mathcal{P}_3(\mathbb{R}) \;/\; T(p(x)) = 0_{\mathcal{P}_3(\mathbb{R})} \,\}\,.$$

Tomando um elemento genérico $p(x) = a + bx + cx^2 + dx^3$, vamos impor a condição que $p(x) \in Ker(T)$, isto é,

$$\begin{aligned} T(p(x)) &= (\, a + bx + cx^2 + dx^3 \,) + (\, 1 + x \,)(\, b + 2cx + 3dx^2 \,) = 0_{\mathcal{P}_3(\mathbb{R})} \\ &= (\, a + b \,) + (\, 2b + 2c \,)x + (\, 3c + 3d \,)x^2 + 4dx^3 = 0_{\mathcal{P}_3(\mathbb{R})} \end{aligned}$$

o que nos leva a um sistema linear homogêneo que possui somente a solução trivial

$$a = b = c = d = 0\,.$$

Assim, mostramos que $Ker(T) = \{\, 0_{\mathcal{P}_3(\mathbb{R})} \,\}$, isto é, T é um operador injetor.

Pelo Teorema do núcleo e da imagem, sabemos que $Im(T) = \mathcal{P}_3(\mathbb{R})$, isto é, T é um operador sobrejetor. Portanto, mostramos que T é um automorfismo de $\mathcal{P}_3(\mathbb{R})$.

Finalmente, vamos determinar a representação matricial do operador T com relação à base canônica $\beta = \{\, 1, x, x^2, x^3 \,\}$ de $\mathcal{P}_3(\mathbb{R})$. Para isso, vamos calcular

$$\begin{aligned} T(1) &= 1 \\ T(x) &= 1 + 2x \\ T(x^2) &= 2x + 3x^2 \\ T(x^3) &= 3x^2 + 4x^3 \end{aligned}$$

Desse modo, obtemos

$$[T]^{\beta}_{\beta} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix},$$

o que completa a resolução.

Exemplo 4.8.15 *Considere uma matriz* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Definimos a transformação linear* T_A *de* $\mathbb{R}^n$ *em* $\mathbb{R}^m$ *associada à matriz* $A = [a_{ij}]$ *da seguinte forma:*

$$y = T_A(x) \qquad \text{para todo} \qquad x = (x_1, \cdots, x_n) \in \mathbb{R}^n ,$$

onde a i–ésima componente do elemento $y = (y_1, \cdots, y_m) \in \mathbb{R}^m$ *é dada por:*

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j \qquad ; \qquad i = 1, \cdots, m .$$

Podemos verificar que a matriz $[T_A]_\gamma^\beta = A$, *onde* β *é a base canônica do* $\mathbb{R}^n$ *e* γ *é a base canônica do* $\mathbb{R}^m$. *Assim,* $[T_A(x)]_\gamma = [T_A]_\gamma^\beta\,[x]_\beta = A[x]_\beta$ *para todo* $x \in \mathbb{R}^n$.

Definição 4.8.1 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e a transformação linear* T_A *de* $\mathbb{R}^n$ *em* $\mathbb{R}^m$, *associada à matriz* A. *Definimos o* ***posto*** *da matriz* A, *que denotamos por* $posto(A)$, *como sendo a dimensão da imagem de* T_A, *isto é,* $posto(A) = dim(\,Im(T_A)\,)$.

Exemplo 4.8.16 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e a transformação linear* T_A *de* $\mathbb{R}^n$ *em* $\mathbb{R}^m$ *associada à matriz* A. *Podemos verificar facilmente que o* $posto(A)$ *é igual ao número de colunas de* A *que são linearmente independentes em* $\mathbb{R}^m$, *tendo em vista que a* $Im(T_A)$ *é o subespaço gerado pelas colunas da matriz* A. *Sendo assim, denotando a matriz* $A = [v_1 \cdots v_j \cdots v_n]$, *onde* $v_j \in \mathbb{R}^m$ *é a j–ésima coluna de* A, *temos que todo elemento* $y \in Im(T_A)$ *é escrito como:*

$$y = \sum_{j=1}^{n} c_j\, v_j .$$

Exemplo 4.8.17 *Considere a matriz* $A \in \mathbb{M}_{3\times 4}(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 3 & 4 & 1 \\ -1 & 5 & 2 & 1 \\ 3 & 1 & 6 & 1 \end{bmatrix} .$$

Determine a transformação linear T_A *associada à matriz* A *e o* $posto(A)$.

Exemplo 4.8.18 *Considere a matriz* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -1 & 5 & 2 \\ 3 & 1 & 6 \\ 1 & 1 & 1 \end{bmatrix} .$$

Determine a transformação linear T_A *associada à matriz* A *e o* $posto(A)$.

Na seção 8.10 apresentamos um estudo mais detalhado sobre os resultados envolvendo o **posto de A**, mostrando que a Definição 2.6.3 e a Definição 4.8.1 são compatíveis. A seguir apresentamos alguns resultados importantes sobre o posto de uma matriz.

Teorema 4.8.4 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *a transformação linear associada à matriz* A, $Q \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível,* $T_Q : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *a transformação linear associada à matriz* Q, *e* $T_{AQ} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *a transformação linear associada à matriz* AQ. *Então,*

$$Im(T_{AQ}) \,=\, Im(T_A \circ T_Q) \,=\, Im(T_A)\,.$$

Demonstração – Como Q é uma matriz invertível, do Teorema 3.7.2, provamos que $Im(T_Q) = \mathbb{R}^n$, isto é, todo elemento $y \in \mathbb{R}^n$ pode ser escrito como $y = T_Q(x)$ para algum $x \in \mathbb{R}^n$. Além disso, do Teorema 2.9.7, temos que $Ker(T_Q) = \{ 0_{\mathbb{R}^n} \}$. Portanto, mostramos que T_Q é bijetora. Desse modo, temos que todo elemento $z \in Im(T_{AQ})$ é representado da forma:

$$z \,=\, T_{AQ}(x) \,=\, (T_A \circ T_Q)(x) \,=\, T_A(T_Q(x)) \,=\, T_A(y)\,,$$

onde $y = T_Q(x)$, para algum $x \in \mathbb{R}^n$, o que completa a demonstração. ■

Corolário 4.8.3 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e* $Q \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível. Então,*

$$posto(AQ) \,=\, posto(A)\,.$$

Demonstração – A prova segue imediata da Definição 4.8.1 e do Teorema 4.8.4. □

Exemplo 4.8.19 *Considere a matriz* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *e* $Q \in \mathbb{M}_3(\mathbb{R})$ *uma matriz invertível dadas por:*

$$A \,=\, \begin{bmatrix} 1 & 3 & 4 \\ -1 & 5 & 2 \\ 3 & 1 & 6 \\ 1 & 1 & 1 \end{bmatrix} \qquad e \qquad Q \,=\, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

(a) *Determine as transformações lineares* T_A, T_Q *e* T_{AQ}.

(b) *Verifique que* $posto(AQ) = posto(A)$.

Teorema 4.8.5 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e* $P \in \mathbb{M}_m(\mathbb{R})$ *uma matriz invertível. Então,*

$$posto(PA) \,=\, posto(A)\,.$$

Demonstração – A prova pode ficar a cargo do leitor. □

Exercícios

Exercício 4.40 *Considere a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *definida por:*

$$T(x,y,z) \;=\; (x \,-\, y \,+\, z \,,\, -x \,+\, 2z)\;.$$

Determine $[T]^{\gamma}_{\beta}$*, onde* γ *é a base canônica de* $\mathbb{R}^3$ *e* β *é a base canônica de* $\mathbb{R}^2$*.*

Exercício 4.41 *Considere o operador linear* T *sobre* $\mathbb{R}^3$ *definida por:*

$$T(x,y,z) \;=\; (x \,-\, y \,+\, z \,,\, x \,+\, y \,+\, 2z \,,\, x \,+\, 2y \,+\, z)\;.$$

Determine $[T]^{\beta}_{\beta}$*, onde* β *é a base canônica de* $\mathbb{R}^3$*.*

Exercício 4.42 *Considere o operador linear* T *sobre* $\mathbb{R}^2$ *definido por:*

$$T(x,y) \;=\; (x \,+\, 2y \,,\, 2x \,+\, 4y)\;.$$

Determine $[T]^{\beta}_{\beta}$ *,* $[T]^{\beta}_{\alpha}$ *,* $[T]^{\alpha}_{\beta}$ *e* $[T]^{\gamma}_{\alpha}$ *, onde as base* β*,* α *e* γ *são dadas por:*

$$\beta \;=\; \{\, (1,0)\,,\,(0,1)\,\} \quad , \quad \alpha \;=\; \{\,(1,-1)\,,\,(0,1)\,\} \qquad e \qquad \gamma \;=\; \{\,(1,-1)\,,\,(1,1)\,\}\;.$$

Exercício 4.43 *Considere o operador linear* T *sobre* $\mathbb{R}^3$ *definido por:*

$$T(x,y,z) \;=\; (\; x + 2y + z \;,\; 2x - y \;,\; 2y + z \;)\;.$$

Determine as matrizes $[T]^{\beta}_{\beta}$ *,* $[T]^{\gamma}_{\beta}$ *e* $[T]^{\beta}_{\gamma}$*, onde as bases* β *e* γ *são dadas por:*

$$\beta \;=\; \{\,(1,0,0)\,,\,(0,1,0)\,,\,(0,0,1)\,\} \qquad e \qquad \gamma \;=\; \{\,(1,0,1)\,,\,(0,1,1)\,,\,(0,0,1)\,\}\;.$$

Exercício 4.44 *Considere a transformação linear* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *definida por:*

$$T(x,y,z) \;=\; (x \,+\, y \,-\, z \,,\, 2x \,-\, y \,+\, 2z)\;.$$

Determine a matriz $[T]^{\gamma}_{\beta}$*, onde as bases* β *e* γ *são dadas por:*

$$\gamma \;=\; \{\,(1,0,1)\,,\,(0,1,1)\,,\,(1,1,0)\,\} \qquad e \qquad \beta \;=\; \{\,(1,-1)\,,\,(0,1)\,\}\;.$$

Exercício 4.45 *Seja* U *um subespaço de* $\mathcal{P}_3(\mathbb{R})$. *Considere a transformação linear* $T : U \longrightarrow \mathcal{P}_2(\mathbb{R})$ *dada por:* $T(p(x)) = p'(x) + (x+1)p(0)$. *Seja*

$$[T]_\gamma^\beta = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad onde \qquad \beta = \{\, x - x^2 + x^3,\, 1 + x + x^2 \,\}.$$

(a) *Determine* $[p(x)]_\beta$ *sabendo que* $[T(p(x))]_\gamma = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix}$.

(b) *Se* $\gamma = \{\, x+2,\, p_1(x),\, p_2(x) \,\}$, *determine* $3p_1(x) + 3p_2(x)$.

Exercício 4.46 *Mostre que a transformação linear* $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$ *definida por:* $T(p(x)) = (p(-1), p(0), p(1))$ *é bijetora. Determine a matriz* $[T]_\beta^\gamma$, *onde* γ *é a base canônica de* $\mathcal{P}_2(\mathbb{R})$ *e* β *é a base canônica de* $\mathbb{R}^3$.

Exercício 4.47 *Mostre que o operador linear* T *sobre* $\mathbb{R}^3$ *definido por:*

$$T(x,y,z) = (x - y,\, 2y,\, y + z)$$

é invertível e determine o isomorfismo inverso T^{-1}, *utilizando a matriz* $[T]_\beta^\beta$, *onde* β *é a base canônica do* $\mathbb{R}^3$.

Exercício 4.48 *Considere o operador linear* T *sobre* $\mathbb{R}^2$ *definido por:*

$$T(x,y) = (x - 2y,\, -2x + y).$$

(a) *Determine a matriz* $[T]_\beta^\beta$, *onde* $\beta = \{\,(1,1),\, (1,-1)\,\}$.

(b) *Determine o isomorfismo inverso* T^{-1}.

Exercício 4.49 *Considere a matriz* $A \in \mathbb{M}_{3\times 2}(\mathbb{R})$ *e* $P \in \mathbb{M}_3(\mathbb{R})$ *uma matriz invertível dadas por:*

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \end{bmatrix} \qquad e \qquad P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Verifique que $posto(PA) = posto(A)$.

Exercício 4.50 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, $P \in \mathbb{M}_m(\mathbb{R})$ *e* $Q \in \mathbb{M}_n(\mathbb{R})$ *matrizes invertíveis. Mostre que* $posto(PAQ) = posto(A)$.

Exercício 4.51 *Sejam* $A \in I\!M_{m\times n}(I\!R)$ *e* $H \in I\!M_m(I\!R)$ *uma matriz elementar de linha. Mostre que* $posto(HA) = posto(A)$. ***Sugestão:*** *faça uso do Teorema 2.7.4.*

Exercício 4.52 *Sejam* $A \in I\!M_{m\times n}(I\!R)$ *e* $K \in I\!M_n(I\!R)$ *uma matriz elementar de coluna. Mostre que* $posto(AK) = posto(A)$. ***Sugestão:*** *faça uso do Teorema 2.7.5.*

Exercício 4.53 *Considere a transformação linear* $T : I\!R^2 \longrightarrow I\!R^3$ *tal que*

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -2 & 3 \end{bmatrix},$$

onde β *é a base canônica de* $I\!R^2$ *e* $\gamma = \{ (1,0,1), (-1,0,1), (0,1,0) \}$.

(a) Determine $T(1,0)$ *e* $T(0,1)$.

(b) Determine uma base para $Im(T)$.

(c) A transformação T *é injetora ?*

Exercício 4.54 *Seja* $T : I\!R^2 \longrightarrow I\!R^2$ *o operador linear definido por:*

$$T(x,y) = (3x - 2y, -2x + 3y).$$

(a) Determine uma base para cada um dos seguintes subespaços:

$$U_1 = \{ (x,y) \in I\!R^2 \ / \ T(x,y) = 5(x,y) \}$$
$$U_2 = \{ (x,y) \in I\!R^2 \ / \ T(x,y) = (x,y) \}$$

(b) Mostre que o conjunto $\beta = \beta_1 \cup \beta_2$, *onde* β_1 *é uma base para* U_1 *e* β_2 *é uma base para* U_2, *é uma base para* $I\!R^2$ *e determine* $[T]^{\beta}_{\beta}$.

Exercício 4.55 *Sejam* T *um operador linear sobre* $I\!R^4$, $\gamma = \{ v_1, v_2, v_3, v_4 \}$ *uma base ordenada para o espaço vetorial real* $I\!R^4$ *e o subespaço* $S = [v_1, v_2, v_3]$.

(a) Sabendo que $T(v) = v$ *para todo* $v \in S$ *e* $T(v_4) = v_1 + v_3$, *determine* $[T]^{\gamma}_{\gamma}$.

(b) Sabendo que

$$[I]^{\beta}_{\gamma} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

onde $\beta = \{ e_1, e_2, e_3, e_4 \}$ *é a base canônica de* $I\!R^4$, *determine* $[T(e_1)]_{\gamma}$.

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] Luenberger, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \quad = \quad Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q \quad = \quad \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/∼pulino/ALESA/

Janeiro de 2012

Conteúdo

5

Produto Interno

Conteúdo

5.1 Introdução

Na geometria Euclidiana as propriedades que nos possibilitam expressar o comprimento de vetor e o ângulo entre dois vetores são denominadas de propriedades métricas. No estudo do $\mathbb{R}^n$, em geometria analítica, definimos comprimento de vetores e ângulo entre vetores através do produto escalar

$$x \cdot y = \sum_{i=1}^{n} x_i \, y_i \qquad \text{para} \qquad x, \; y \in \mathbb{R}^n \, .$$

Nosso objetivo é de estender esses conceitos para os espaços vetoriais sobre um corpo $\mathbb{F}$. O conceito de produto interno em um espaço vetorial real (complexo) é uma generalização do conceito de produto escalar definido em $\mathbb{R}^n$. Faremos essa generalização através do estudo de certos tipos de aplicações que são definidas sobre pares de elementos de um espaço vetorial e tomando valores no corpo.

Denotamos o **produto interno** entre dois elementos u e v de um espaço vetorial da seguinte forma: $\langle \, u \, , \, v \, \rangle$. Neste capítulo apresentamos um estudamos das propriedades geométricas que são atribuídas a um espaço vetorial por meio de algum produto interno definido sobre ele. Mais especificamente, estabelecemos as propriedades básicas, e suas aplicações, dos conceitos de comprimento, ângulo e ortogonalidade determinadas ao espaço vetorial pelo produto interno.

5.2 Definição de Produto Interno

Definição 5.2.1 *Seja V um espaço vetorial sobre o corpo $\mathbb{R}$. Uma aplicação*

$$\langle \, \cdot \, , \, \cdot \, \rangle : V \times V \longrightarrow \mathbb{R}$$

que satisfaz as seguintes propriedades:

(1) Simetria: $\langle \, u \, , \, v \, \rangle = \langle \, v \, , \, u \, \rangle \quad ; \quad \forall \; u, \; v \in V$

(2) Positividade: $\langle \, u \, , \, u \, \rangle \geq 0 \quad ; \quad \forall \; u \in V, \quad \textit{com} \quad \langle \, u \, , \, u \, \rangle = 0 \iff u = 0_V$

(3) Distributividade: $\langle \, u + w \, , \, v \, \rangle = \langle \, u \, , \, v \, \rangle + \langle \, w \, , \, v \, \rangle \quad ; \quad \forall \; u, \; v, \; w \in V$

(4) Homogeneidade: $\langle \, \lambda u \, , \, v \, \rangle = \lambda \langle \, u \, , \, v \, \rangle \quad ; \quad \forall \; u, \; v \in V \; \textit{e} \; \lambda \in \mathbb{R}$

define um ***produto interno*** *no espaço vetorial real V.*

Utilizando as propriedades de simetria, distributividade e homogeneidade tem–se

- $\langle u , v + w \rangle = \langle u , v \rangle + \langle u , w \rangle$ para todos $u, v, w \in V$.
- $\langle u , \lambda v \rangle = \lambda \langle u , v \rangle$ para todos $u, v \in V$ e $\lambda \in \mathbb{R}$.

Assim, dizemos que o produto interno, em um espaço vetorial real, é uma **aplicação bilinear**, isto é, é uma aplicação linear nas duas variáveis.

Definição 5.2.2 *Seja V um espaço vetorial sobre o corpo $\mathbb{C}$. Uma aplicação*

$$\langle \cdot , \cdot \rangle : V \times V \longrightarrow \mathbb{C}$$

que satisfaz as seguintes propriedades:

(1) Simetria Hermitiana: $\langle u , v \rangle = \overline{\langle v , u \rangle}$; $\forall\ u, v \in V$

(2) Positividade: $\langle u , u \rangle \geq 0$; $\forall\ u \in V$, *com* $\langle u , u \rangle = 0 \iff u = 0_V$

(3) Distributividade: $\langle u + w , v \rangle = \langle u , v \rangle + \langle w , v \rangle$; $\forall\ u, v, w \in V$

(4) Homogeneidade: $\langle \lambda u , v \rangle = \lambda \langle u , v \rangle$; $\forall\ u, v \in V$ *e* $\lambda \in \mathbb{C}$

*define um **produto interno** no espaço vetorial complexo V.*

Podemos verificar que com as propriedades de simetria Hermitiana, distributividade e homogeneidade temos que:

- $\langle u , v + w \rangle = \langle u , v \rangle + \langle u , w \rangle$ para todos $u, v, w \in V$.
- $\langle u , \lambda v \rangle = \overline{\lambda} \langle u , v \rangle$ para todos $u, v \in V$ e $\lambda \in \mathbb{C}$.

É importante observar que em um espaço vetorial complexo o produto interno possui a propriedade de simetria Hermitiana, que é necessária para garantir a propriedade de positividade. De fato, considere um elemento $u \in V$ não–nulo, como V é um espaço vetorial complexo, tem–se que o elemento $iu \in V$. Logo, obtemos

$$\langle iu , iu \rangle = i\, i \langle u , u \rangle = -1 \langle u , u \rangle < 0$$

que é uma contradição, proveniente da não utilização da simetria Hermitiana.

Considerando agora a propriedade de simetria Hermitiana, tem–se que

$$\langle iu , iu \rangle = i \bar{i} \langle u , u \rangle = \langle u , u \rangle > 0 ,$$

o que mostra a necessidade da propriedade de simetria Hermitiana.

Definição 5.2.3 *Um **espaço vetorial com produto interno**, que denotamos por* $(V, \langle \cdot , \cdot \rangle)$ *é um espaço vetorial* V *sobre o corpo* $\mathbb{F}$ *com produto interno* $\langle \cdot , \cdot \rangle$. *Um espaço vetorial real com produto interno é denominado **espaço Euclidiano**. Um espaço vetorial complexo com produto interno é denominado **espaço unitário**.*

Exemplo 5.2.1 *Seja* $\beta = \{ e_1, \cdots , e_n \}$ *a base canônica do* $\mathbb{R}^n$. *Todo elemento* $x = (x_1, \cdots , x_n) \in \mathbb{R}^n$ *é escrito de modo único da seguinte forma:*

$$x = \sum_{i=1}^{n} x_i \, e_i \, .$$

Em muitas situações, por simplicidade de notação, associamos o elemento $x \in \mathbb{R}^n$ *a matriz coluna* $X \in \mathbb{M}_{n\times 1}(\mathbb{R})$, *tendo em vista que os espaços vetoriais são isomorfos,*

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} .$$

Desse modo, o produto interno usual do $\mathbb{R}^n$ *que vamos denotar por* $\langle \cdot , \cdot \rangle$, *denominado produto interno Euclidiano, pode ser escrito como:*

$$\langle x , y \rangle = \sum_{i=1}^{n} x_i \, y_i = Y^t X = Y^t I_n X \qquad \text{para todos} \qquad x, \ y \in \mathbb{R}^n ,$$

onde $I_n \in \mathbb{M}_n(\mathbb{R})$ *é a matriz identidade de ordem* n.

De modo análogo, no espaço vetorial complexo $\mathbb{C}^n$ *o produto interno usual, denominado produto interno Hermitiano, é escrito da seguinte forma:*

$$\langle x , y \rangle = \sum_{i=1}^{n} x_i \, \overline{y}_i = Y^* X = Y^* I_n X \qquad \text{para todos} \qquad x, \ y \in \mathbb{C}^n .$$

onde Y^* *é a transposta Hermitiana da matriz coluna* Y.

Exemplo 5.2.2 *Considere o espaço vetorial real* $\mathcal{C}([a,b])$. *O produto interno usual é definido da seguinte forma:*

$$\langle f , g \rangle = \int_a^b f(x)g(x)dx \qquad ; \qquad \forall \ f, \ g \in \mathcal{C}([a,b]) .$$

Exemplo 5.2.3 *No espaço vetorial real* $\mathbb{M}_n(\mathbb{R})$ *o produto interno usual é definido da seguinte forma:*

$$\langle A , B \rangle = tr(B^t A) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \, b_{ij} \qquad ; \qquad \forall \ A, \ B \in \mathbb{M}_n(\mathbb{R}) .$$

Exemplo 5.2.4 *Considere o espaço vetorial real* $\mathcal{C}([a,b])$ *e o elemento* $w \in \mathcal{C}([a,b])$ *estritamente positivo. A aplicação* $\langle \cdot , \cdot \rangle_w$ *definida por:*

$$\langle f , g \rangle_w = \int_a^b w(x)f(x)g(x)dx \quad ; \quad \forall\, f,\ g \in \mathcal{C}([a,b]) ,$$

define um produto interno com peso no espaço vetorial $\mathcal{C}([a,b])$.

Exemplo 5.2.5 *Considere o espaço vetorial real* $\mathbb{R}^n$. *Verifique se as aplicações abaixo definem um produto interno em* $\mathbb{R}^n$.

1. $\langle x , y \rangle = \displaystyle\sum_{i=1}^{n} x_i \,|\, y_i \,|$

2. $\langle x , y \rangle = \left| \displaystyle\sum_{i=1}^{n} x_i y_i \right|$

3. $\langle x , y \rangle = \left(\displaystyle\sum_{i=1}^{n} x_i \right) \left(\displaystyle\sum_{j=1}^{n} y_j \right)$

Exemplo 5.2.6 *Seja* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um automorfismo de* V. *Mostre que a aplicação*

$$\begin{aligned} f(\cdot,\cdot) :\ V \times V &\longrightarrow \mathbb{R} \\ (u,v) &\longrightarrow f(u,v) = \langle T(u) , T(v) \rangle \end{aligned}$$

define um produto interno em V.

Para mostrar que a aplicação $f(\cdot,\cdot)$ satisfaz as propriedades de *simetria*, *distributividade* e *homogeneidade*, basta utilizar a hipótese que T é um operador linear e em seguida a hipótese que $\langle \cdot , \cdot \rangle$ é um produto interno em V.

Para mostrar que a aplicação $f(\cdot,\cdot)$ é *positiva* vamos utilizar a hipótese que T é um isomorfismo, isto é, T é um operador injetor ($Ker(T) = \{ 0_V \}$), e a hipótese que $\langle \cdot , \cdot \rangle$ é um produto interno em V. De fato,

$$f(u,u) = \langle T(u) , T(u) \rangle \geq 0 ,$$

pois $\langle \cdot , \cdot \rangle$ é um produto interno em V, e temos que

$$f(u,u) = \langle T(u) , T(u) \rangle = 0 \iff T(u) = 0_V \iff u = 0_V ,$$

pois T é um operador injetor, o que completa a prova.

Dada uma matriz $A \in \mathbb{M}_n(\mathbb{R})$ vamos definir o operador linear T_A sobre o $\mathbb{R}^n$ associado à matriz $A = [a_{ij}]$ da seguinte forma: $y = T_A(x)$ para $x \in \mathbb{R}^n$, onde a i–ésima componente do elemento $y = (y_1, \cdots, y_n) \in \mathbb{R}^n$ é dada por:

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j \qquad ; \qquad i = 1, \cdots, n\,.$$

Desse modo, podemos representar o elemento $y = T_A(x)$ na forma de matriz coluna como $Y = AX$, onde

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e} \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

A notação utilização para representar o operador T_A será muito útil no exemplo a seguir.

Exemplo 5.2.7 *Considere o espaço vetorial $\mathbb{R}^n$ munido do produto interno usual $\langle \cdot\,,\cdot \rangle$. Mostre que se $A \in \mathbb{M}_n(\mathbb{R})$ é uma matriz simétrica, então $\langle T_A(x)\,, y \rangle = \langle x\,, T_A(y) \rangle$ para todo $x, y \in \mathbb{R}^n$, onde T_A é operador linear associado à matriz simétrica A.*

Inicialmente, considerando o espaço vetorial $\mathbb{R}^n$ com a base canônica, vamos representar os elementos $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n) \in \mathbb{R}^n$ na forma de matriz coluna

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e} \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Finalmente, escrevendo o produto interno usual do $\mathbb{R}^n$ na forma matricial e utilizando a hipótese que A é uma matriz simétrica, temos que

$$\langle T_A(x)\,, y \rangle = Y^t A X = (A^t Y)^t X = (AY)^t X = \langle x\,, T_A(y) \rangle\,,$$

mostrando o resultado desejado.

Exemplo 5.2.8 *Seja V um espaço vetorial real munido do produto interno $\langle \cdot\,,\cdot \rangle$ e T um operador linear sobre V. Se $\langle T(u)\,, T(v) \rangle = \langle u\,, v \rangle$ para todos $u, v \in V$, então T é um operador injetor.*

Sabemos que T é um operador injetor se, e somente se, $Ker(T) = \{ 0_V \}$. Tomando um elemento $u \in Ker(T)$, isto é, $T(u) = 0_V$, e utilizando a hipótese, temos que

$$0_{\mathbb{R}} = \langle T(u)\,, T(u) \rangle = \langle u\,, u \rangle \qquad \Longleftrightarrow \qquad u = 0_V\,.$$

Portanto, $Ker(T) = \{ 0_V \}$, o que mostra o resultado desejado.

Matriz do Produto Interno

Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, que pode ser $\mathbb{R}$ ou $\mathbb{C}$, munido do produto interno $\langle \cdot , \cdot \rangle$ e $\beta = \{ v_1 , \cdots , v_n \}$ uma base ordenada de V. Vamos mostrar que o produto interno pode ser completamente descrito em termos de uma dada base por meio de uma determinada matriz.

Considere os elementos $u, v \in V$, que podem ser representados de modo único da forma:

$$u = \sum_{j=1}^{n} x_j v_j \qquad \text{e} \qquad v = \sum_{i=1}^{n} y_i v_i \, .$$

Desse modo, temos que

$$\begin{aligned}
\langle u , v \rangle &= \langle \sum_{j=1}^{n} x_j v_j , v \rangle \\
&= \sum_{j=1}^{n} x_j \langle v_j , v \rangle \\
&= \sum_{j=1}^{n} x_j \langle v_j , v \rangle \sum_{i=1}^{n} \overline{y}_i \langle v_j , v_i \rangle \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \langle v_j , v_i \rangle x_j \, \overline{y}_i \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} \, x_j \, \overline{y}_i \\
&= Y^* A X
\end{aligned}$$

onde $X , Y \in \mathbb{M}_{n \times 1}(\mathbb{F})$ são as matrizes de coordenadas dos elementos u e v em relação à base ordenada β, respectivamente, isto é,

$$X = [u]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e} \qquad Y = [v]_\beta = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} ,$$

e a matriz $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{F})$ com os elementos dados por:

$$a_{ij} = \langle v_j , v_i \rangle \qquad \text{para} \qquad i, j = 1, \cdots , n \, ,$$

é denominada **matriz do produto interno em relação à base ordenada** β.

Podemos verificar facilmente que A é uma matriz Hermitiana. De fato,

$$a_{ij} \;=\; \langle\, v_j \,,\, v_i \,\rangle \;=\; \overline{\langle\, v_i \,,\, v_j \,\rangle} \;=\; \overline{a}_{ji}$$

para $i, j \,=\, 1, \cdots, n$. Portanto, mostramos que $A^* \,=\, A$.

Da propriedade de positividade do produto interno, temos que

$$\langle\, u \,,\, u \,\rangle \;=\; \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}\, x_j\, \overline{x}_i \;=\; X^* A X \;>\; 0$$

para todo $u \neq 0_V$. Desse modo, temos que a matriz de coordenadas

$$X \;=\; [u]_\beta \;=\; \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;\neq\; 0_{\mathbb{M}_{n\times 1}(\mathbb{F})} \;.$$

Portanto, temos que

$$X^* A X \;>\; 0 \qquad \text{para todo} \qquad X \,\neq\, 0_{\mathbb{M}_{n\times 1}(\mathbb{F})} \;. \tag{5.1}$$

Além disso, sabemos também que

$$\langle\, u \,,\, u \,\rangle \,=\, 0 \qquad \Longleftrightarrow \qquad u \,=\, 0_V \;.$$

Assim, temos que

$$X^* A X \;=\; 0 \qquad \Longleftrightarrow \qquad X \,=\, 0_{\mathbb{M}_{n\times 1}(\mathbb{F})} \;. \tag{5.2}$$

Como A é Hermitiana e satisfaz a propriedade (5.1), dizemos que A é uma matriz **positiva–definida**. As matrizes positiva–definidas serão estudadas com mais detalhes nas seções 6.4 e 6.7, onde iremos fazer uma caracterização, facilitando sua identificação.

Finalmente, podemos observar que a matriz do produto interno A deve ser invertível. De fato, caso contrário, existiria um elemento $X \,\neq\, 0_{\mathbb{M}_{n\times 1}(\mathbb{F})}$ tal que $AX \,=\, 0_{\mathbb{F}^n}$, o que leva a uma contradição com a propriedade (5.1), pois teríamos

$$X^t A X \,=\, 0$$

para $X \,\neq\, 0_{\mathbb{M}_{n\times 1}(\mathbb{F})}$.

Exemplo 5.2.9 *Considere o espaço vetorial real* $I\!R^n$ *munido com o produto interno usual* $\langle \cdot , \cdot \rangle$ *e com a base canônica* $\beta = \{ e_1, \cdots, e_n \}$. *Podemos verificar facilmente que* I_n *é a matriz do produto interno usual com relação à base canônica* β.

Exemplo 5.2.10 *Considere o espaço vetorial real* $I\!R^3$ *munido com o produto interno usual* $\langle \cdot , \cdot \rangle$ *e com a base ordenada* $\Gamma = \{ v_1, v_2, v_3 \}$ *dada por:*

$$v_1 = (1, 0, -1) \quad , \quad v_2 = (1, 2, 1) \quad e \quad v_3 = (0, -3, 2) .$$

Podemos verificar facilmente que a matriz $A = [a_{ij}]$ *dada por:*

$$a_{ij} = \langle v_j , v_i \rangle = \langle v_i , v_j \rangle = a_{ji} \quad \Longrightarrow \quad A = \begin{bmatrix} 2 & 0 & -2 \\ 0 & 6 & -4 \\ -2 & -4 & 13 \end{bmatrix}$$

é a matriz do produto interno usual com relação à base ordenada Γ.

Exemplo 5.2.11 *Considere o espaço vetorial real* $I\!R^2$ *munido com o produto interno usual* $\langle \cdot , \cdot \rangle$ *e com a base ordenada* $\Gamma = \{ v_1, v_2 \}$ *dada por:*

$$v_1 = (1, -1) \quad e \quad v_2 = (1, 1) .$$

Podemos verificar facilmente que a matriz $A = [a_{ij}]$ *dada por:*

$$a_{ij} = \langle v_j , v_i \rangle = \langle v_i , v_j \rangle = a_{ji} \quad \Longrightarrow \quad A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

é a matriz do produto interno usual com relação à base ordenada Γ.

Exemplo 5.2.12 *Considere o espaço vetorial real* $\mathcal{P}_2(I\!R)$ *com o produto interno*

$$\langle p , q \rangle = \int_0^1 p(x)q(x)dx \quad ; \quad \forall \ p, q \in \mathcal{P}_2(I\!R) .$$

Podemos verificar facilmente que a matriz $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

é a matriz do produto interno $\langle \cdot , \cdot \rangle$ *com relação à base canônica*

$$\beta = \{ p_1(x) = 1 \ , \ p_2(x) = x \ , \ p_3(x) = x^2 \} ,$$

que é a matriz de Hilbert de ordem 3.

Exemplo 5.2.13 *Considere o espaço vetorial real* $\mathcal{P}_3(I\!R)$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx \qquad ; \qquad \forall\; p,\, q \,\in\, \mathcal{P}_3(I\!R)\,.$$

Podemos verificar facilmente que a matriz $A \in I\!M_4(I\!R)$ *dada por:*

$$A \;=\; \begin{bmatrix} 2 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{2}{5} \\ \frac{2}{3} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & 0 & \frac{2}{7} \end{bmatrix}$$

é a matriz do produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$ *com relação à base canônica*

$$\beta \;=\; \{\, p_1(x) \,=\, 1 \;,\; p_2(x) \,=\, x \;,\; p_3(x) \,=\, x^2 \;,\; p_4(x) \,=\, x^3 \,\}\,.$$

Para exemplificar a utilização da matriz do produto interno, considere os polinômios $p,\, q \in \mathcal{P}_3(I\!R)$ dados por:

$$p(x) \;=\; 1 \,+\, 2x \,+\, x^3 \qquad \text{e} \qquad q(x) \;=\; 3 \,+\, x \,-\, x^2\,.$$

Desse modo, o produto interno entre os elementos $p,\, q \in \mathcal{P}_3(I\!R)$, isto é,

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx,$$

pode ser calculado da seguinte forma:

$$\langle\, p \,,\, q \,\rangle \;=\; Y^t A X \;=\; \frac{106}{15}\,,$$

onde $X\,,\, Y \,\in\, I\!M_{4\times 1}(I\!R)$ são as matrizes de coordenadas dos elementos p e q em relação à base canônica de $\mathcal{P}_3(I\!R)$, respectivamente, isto é,

$$X \;=\; [p]_\beta \;=\; \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \qquad \text{e} \qquad Y \;=\; [q]_\beta \;=\; \begin{bmatrix} 3 \\ 1 \\ -1 \\ 0 \end{bmatrix}\,.$$

É importante observar que a utilização da matriz do produto interno torna o cálculo do produto interno em $\mathcal{P}_3(I\!R)$ mais simples, envolvendo somente produto de matrizes.

Exemplo 5.2.14 *Considere o espaço vetorial real* $I\!R^n$ *com a base canônica*

$$\beta = \{ e_1, \cdots, e_n \}.$$

Podemos descrever todos os produtos internos sobre $I\!R^n$ *através de suas matrizes em relação à base canônica.*

De fato, considere uma matriz $A = [a_{ij}] \in I\!M_n(I\!R)$ satisfazendo as propriedades:

$$A^t = A \quad \text{e} \quad X^t A X > 0 \quad \text{para todo} \quad X \neq 0_{I\!M_{n\times 1}(I\!R)},$$

e definimos o produto interno entre os elementos $u, v \in I\!R^n$,

$$u = (x_1, \cdots, x_n) \quad \text{e} \quad v = (y_1, \cdots, y_n),$$

da seguinte forma:

$$\langle u, v \rangle = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}\, x_j\, y_i = Y^t A X$$

onde $X, Y \in I\!M_{n\times 1}(I\!R)$ são as matrizes de coordenadas dos elementos u e v em relação à base canônica β, respectivamente, isto é,

$$X = [u]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad Y = [v]_\beta = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Portanto, fixando uma base ordenada β para um espaço vetorial V de dimensão finita sobre o corpo $I\!F$, digamos $dim(V) = n$, podemos obter uma descrição de todos os produtos internos possíveis sobre V. De fato, se $A \in I\!M_n(I\!F)$ é uma matriz com as propriedades:

$$A^* = A \quad \text{e} \quad X^* A X > 0 \quad \text{para todo} \quad X \neq 0_{I\!M_{n\times 1}(I\!F)},$$

então A é a matriz de algum produto interno sobre V em relação à base ordenada β. Sendo assim, esse produto interno é definido da seguinte forma:

$$\langle u, v \rangle = Y^* A X,$$

onde $X, Y \in I\!M_{n\times 1}(I\!F)$ são as matrizes de coordenadas dos elementos u e v em relação à base ordenada β, respectivamente.

Exercícios

Exercício 5.1 *Seja* V *um espaço vetorial real. Considere que* $\langle \cdot , \cdot \rangle_1$ *e* $\langle \cdot , \cdot \rangle_2$ *são dois produtos internos em* V. *Prove que a aplicação* $\langle \cdot , \cdot \rangle$ *dada por:*

$$\langle u , v \rangle = \langle u , v \rangle_1 + \langle u , v \rangle_2 \qquad ; \qquad u, v \in V$$

define um novo produto interno em V.

Exercício 5.2 *Sejam* V *um espaço vetorial real de dimensão finita munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $\beta = \{ v_1, \cdots , v_n \}$ *uma base de* V. *Dados os escalares* $c_1, \cdots , c_n$, *mostre que existe um único elemento* $u \in V$ *tal que*

$$\langle u , v_i \rangle = c_i \qquad \textit{para} \qquad i = 1, \cdots , n .$$

Exercício 5.3 *Considere o espaço vetorial real* U *definido por:*

$$U = \{ p(x) \in \mathcal{P}_3(I\!\!R) \ / \ p(-1) = p(1) = 0 \} .$$

Mostre que a aplicação

$$\langle p , q \rangle = \int_{-1}^{1} p'(x)q'(x)dx \qquad ; \qquad \forall \ p, q \in U .$$

define um produto interno em U.

Exercício 5.4 *Considere o espaço vetorial real* V *definido por:*

$$V = \{ f \in \mathcal{C}^1([a,b]) \ / \ f(a) = f(b) = 0 \} .$$

Mostre que a aplicação definida por:

$$F(f,g) = \int_a^b f^{'}(x)g^{'}(x)dx \qquad ; \qquad \forall \ f,g \in V$$

define um produto interno em V.

Exercício 5.5 *Verifique se a aplicação definida por:*

$$F(f,g) = \int_0^{\frac{1}{2}} f(x)g(x)dx \qquad ; \qquad \forall \ f, \ g \in \mathcal{C}([0,1])$$

define um produto interno em $\mathcal{C}([0,1])$.

Exercício 5.6 *Mostre que a aplicação* $\langle \cdot , \cdot \rangle : \mathcal{P}_2(I\!\!R) \times \mathcal{P}_2(I\!\!R) \longrightarrow I\!\!R$ *dada por:*

$$\langle p , q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

define um produto interno no espaço vetorial real $\mathcal{P}_2(I\!\!R)$.

Exercício 5.7 *Considere o espaço vetorial real* $\mathbb{M}_{m\times n}(\mathbb{R})$*, isto é, o espaço vetorial das matrizes de ordem* $m \times n$ *sobre o corpo* $\mathbb{R}$*. Mostre que a aplicação*

$$\langle A \,,\, B \rangle \;=\; \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, b_{ij} \;=\; tr(B^t A) \qquad ; \qquad \forall\;\, A,\, B \,\in\, \mathbb{M}_{m\times n}(\mathbb{R})\;,$$

define um produto interno sobre $\mathbb{M}_{m\times n}(\mathbb{R})$*.*

Exercício 5.8 *Considere o espaço vetorial* $\mathbb{R}^2$ *e a matriz diagonal* D *dada por:*

$$D \;=\; \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

Mostre que a aplicação

$$\langle x \,,\, y \rangle_D \;=\; y^t D x$$

define um produto interno em $\mathbb{R}^2$*, considere os elementos de* $\mathbb{R}^2$ *representados na forma de vetor coluna.*

Exercício 5.9 *Considere o espaço vetorial* $\mathbb{R}^n$ *e a matriz diagonal* D *dada por:*

$$D \;=\; diag(d_1, \,\cdots\,, d_n)$$

com $d_i > 0$ *para* $i = 1, \cdots, n$*. Mostre que a aplicação*

$$\langle x \,,\, y \rangle_D \;=\; y^t\, D\, x$$

define um produto interno em $\mathbb{R}^n$*, considere os elementos de* $\mathbb{R}^n$ *representados na forma de vetor coluna.*

Exercício 5.10 *Sejam o espaço vetorial real* $\mathbb{M}_{2\times 1}(\mathbb{R})$ *e uma matriz* $A \in \mathbb{M}_2(\mathbb{R})$*. Mostre que a aplicação*

$$f_A(X,Y) \;=\; Y^t A X \qquad ; \qquad \forall\;\, X,\, Y \,\in\, \mathbb{M}_{2\times 1}(\mathbb{R})\;,$$

define um produto interno sobre $\mathbb{M}_{2\times 1}(\mathbb{R})$ *se, e somente se,* $A^t = A$, $a_{11} > 0$, $a_{22} > 0$ *e* $\det(A) > 0$*.*

Exercício 5.11 *Mostre que a aplicação*

$$\langle\, u \,,\, v \,\rangle \;=\; x_1\, y_1 \;-\; x_1\, y_2 \;-\; y_1\, x_2 \;+\; 4x_2\, y_2$$

define um produto interno em $\mathbb{R}^2$*, onde* $u = (x_1, x_2)$ *e* $v = (y_1, y_2)$*. Determine a matriz do produto interno* $\langle\, \cdot\,,\, \cdot\, \rangle$ *com relação à base canônica do* $\mathbb{R}^2$*.*

Exercício 5.12 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *munido com o produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; p(-1)q(-1) \;+\; p(0)q(0) \;+\; p(1)q(1)$$

para todos $p,\, q \,\in\, \mathcal{P}_2(\mathbb{R})$.

(a) Determine a matriz do produto interno $\langle\, \cdot \,,\, \cdot \,\rangle$ *em relação à base canônica*

$$\beta \;=\; \{\, p_1(x) \,=\, 1 \;,\; p_2(x) \,=\, x \;,\; p_3(x) \,=\, x^2 \,\}\,.$$

(b) Considere os polinômios $p,\, q \in \mathcal{P}_2(\mathbb{R})$ *dados por:*

$$p(x) \;=\; -1 \;+\; 3x \;+\; x^2 \qquad e \qquad q(x) \;=\; 4 \;+\; 2x \;-\; x^2\,.$$

Determine o produto interno entre os elementos $p,\, q \in \mathcal{P}_2(\mathbb{R})$ *utilizando a matriz do produto interno, isto é,*

$$\langle\, p \,,\, q \,\rangle \;=\; Y^t A X \;,$$

onde A *é a matriz do produto interno e* $X, Y \,\in\, \mathbb{M}_{2\times 1}(\mathbb{R})$ *são as matriz de coordenadas dos polinômios* $p,\, q$ *com relação à base canônica, respectivamente.*

(c) Determine todos os polinômios $q \in \mathcal{P}_2(\mathbb{R})$ *tais que*

$$\langle\, p \,,\, q \,\rangle \;=\; 0$$

utilizando a matriz do produto interno.

Exercício 5.13 *Considere o espaço vetorial real* $\mathcal{P}_3(\mathbb{R})$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)x^2 dx \qquad ; \qquad \forall\; p,\, q \,\in\, \mathcal{P}_3(\mathbb{R})\,.$$

Determine a matriz do produto interno $\langle\, \cdot \,,\, \cdot \,\rangle$ *em relação à base canônica*

$$\beta \;=\; \{\, p_1(x) \,=\, 1 \;,\; p_2(x) \,=\, x \;,\; p_3(x) \,=\, x^2 \;,\; p_4(x) \,=\, x^3 \,\}\,.$$

Considere o polinômio $p \in \mathcal{P}_3(\mathbb{R})$ *dado por:*

$$p(x) \;=\; 2 \;-\; x \;+\; 4x^2 \;+\; x^3\,.$$

Determine o produto interno $\langle\, p \,,\, p \,\rangle$ *utilizando a matriz do produto interno, isto é,*

$$\langle\, p \,,\, p \,\rangle \;=\; X^t A X \;,$$

onde A *é a matriz do produto interno e* $X \,\in\, \mathbb{M}_{3\times 1}(\mathbb{R})$ *é a matriz de coordenadas do polinômio* p *com relação à base canônica.*

5.3 Desigualdade de Cauchy–Schwarz

Teorema 5.3.1 *Seja* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *. Então, para todos* $u , v \in V$ *temos que*

$$\langle u , v \rangle^2 \leq \langle u , u \rangle \langle v , v \rangle .$$

Além disso, a igualdade é válida se, e somente se, os elementos u *e* v *são linearmente dependentes.*

Demonstração – No caso em que os elementos u e v são linearmente dependentes, a igualdade é obtida trivialmente. Vamos considerar u e v linearmente independentes, isto é, $u + \lambda v \neq 0_V$ para todo $\lambda \in I\!R$. Desse modo, temos que

$$\begin{aligned} \langle u + \lambda v , u + \lambda v \rangle &= \langle u , u \rangle + \langle u , \lambda v \rangle + \langle \lambda v , u \rangle + \lambda^2 \langle v , v \rangle \\ &= \langle u , u \rangle + 2 \lambda \langle u , v \rangle + \lambda^2 \langle v , v \rangle > 0 \end{aligned}$$

é uma inequação de segundo grau na variável λ.

Note que a equação do segundo grau

$$\langle u , u \rangle + 2 \lambda \langle u , v \rangle + \lambda^2 \langle v , v \rangle = 0$$

não possui raízes reais. Assim, devemos ter

$$4 \langle u , v \rangle^2 - 4 \langle u , u \rangle \langle v , v \rangle < 0 \implies \langle u , v \rangle^2 < \langle u , u \rangle \langle v , v \rangle$$

o que completa da demonstração. ■

Exemplo 5.3.1 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual. Verifique a desigualdade de Cauchy–Schwarz para* $u = (1, -2, 1)$ *e* $v = (3, -1, 1)$.

Exemplo 5.3.2 *Considere o espaço vetorial* $\mathcal{C}([-1, 1])$ *com o produto interno usual*

$$\langle f , g \rangle = \int_{-1}^{1} f(x)g(x)dx \quad ; \quad \forall \ f, g \in \mathcal{C}([-\pi, \pi]) \ .$$

Verifique a desigualdade de Cauchy–Schwarz para os elementos $f(x) = x$ *e* $g(x) = x^3$.

Teorema 5.3.2 *Seja* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$. *Então, para todo* $u, v \in V$ *temos que*

$$| \langle u , v \rangle |^2 \leq \langle u , u \rangle \langle v , v \rangle .$$

Além disso, a igualdade é válida se, e somente se, os elementos u *e* v *são linearmente dependentes.*

Demonstração – No caso em que os elementos u e v são linearmente dependentes, a igualdade é obtida trivialmente. Vamos considerar u e v linearmente independentes, isto é, $u + \lambda v \neq 0_V$ para todo $\lambda \in \mathbb{C}$. Desse modo, temos que

$$\begin{aligned} 0 < \langle u + \lambda v , u + \lambda v \rangle &= \langle u , u \rangle + \langle u , \lambda v \rangle + \langle \lambda v , u \rangle + | \lambda |^2 \langle v , v \rangle \\ &= \langle u , u \rangle + \overline{\lambda} \langle u , v \rangle + \lambda \langle v , u \rangle + | \lambda |^2 \langle v , v \rangle \end{aligned}$$

Para o caso complexo, vamos escrever $\langle v , u \rangle \in \mathbb{C}$ da seguinte forma

$$\langle v , u \rangle = \exp(i \theta) \, | \langle v , u \rangle | \qquad ; \qquad \theta \in [0, 2\pi)$$

assim, temos que

$$\overline{\langle v , u \rangle} = \langle u , v \rangle = \exp(-i \theta) \, | \langle v , u \rangle |$$

Desse modo, temos que

$$\langle u , u \rangle + 2 Re(\lambda \exp(i \theta)) \, | \langle u , v \rangle | + | \lambda |^2 \langle v , v \rangle > 0$$

chamando $\beta = \lambda \exp(i \theta) \in \mathbb{C}$, note que $| \lambda |^2 = | \beta |^2$, e observando que $| Re(\beta) | \leq | \beta |$. Assim, encontramos

$$\langle u , u \rangle + 2 | \beta | \, | \langle u , v \rangle | + | \beta |^2 \langle v , v \rangle > 0$$

podemos concluir que a função quadrática em $| \beta |$ não possui raízes reais. Desse modo, temos que

$$4 | \langle u , v \rangle |^2 - 4 \langle u , u \rangle \langle v , v \rangle < 0 \implies | \langle u , v \rangle |^2 < \langle u , u \rangle \langle v , v \rangle$$

o que completa a demonstração. ■

5.4 Definição de Norma. Norma Euclidiana

Definição 5.4.1 (Norma) *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$. *Uma norma, ou comprimento, em* V *é uma aplicação* $\| \cdot \|$ *que para cada elemento* $u \in V$ *associa um número real* $\| u \|$, *que possui as seguintes propriedades:*

(a) Positividade: $\| u \| > 0$ *para* $u \neq 0_V$, *com* $\| u \| = 0 \iff u = 0_V$.

(b) Homogeneidade: $\| \lambda u \| = |\lambda| \, \| u \|$ *para todo* $u \in V$, $\lambda \in \mathbb{F}$.

(c) Desigualdade Triangular: $\| u + v \| \leq \| u \| + \| v \|$ *para todos* $u, v \in V$.

Um espaço vetorial V munido de uma norma $\| \cdot \|$ é denominado **espaço normado**, que denotamos por $(V, \| \cdot \|)$.

Exemplo 5.4.1 *No espaço vetorial real* $\mathbb{R}^n$ *temos as seguintes normas*

(a) Norma do Máximo: $\| x \|_\infty = \max \{ \, |x_i| \quad ; \quad 1 \leq i \leq n \, \}$

(b) Norma–1 ou Norma do Táxi: $\| x \|_1 = \displaystyle\sum_{i=1}^{n} |x_i|$

Podemos verificar facilmente que as aplicações $\| \cdot \|_\infty$ e $\| \cdot \|_1$ satisfazem as propriedades de norma utilizando as propriedades de módulo de um número real.

Exemplo 5.4.2 *Considere o espaço vetorial real* $\mathbb{M}_n(\mathbb{R})$. *A aplicação*

$$\| A \|_\infty = \max \left\{ \sum_{j=1}^{n} |a_{ij}| \quad ; \quad 1 \leq i \leq n \right\}$$

define uma norma em $\mathbb{M}_n(\mathbb{R})$.

Exemplo 5.4.3 *Considere o espaço vetorial real* $\mathbb{M}_n(\mathbb{R})$. *A aplicação*

$$\| A \|_1 = \max \left\{ \sum_{i=1}^{n} |a_{ij}| \quad ; \quad 1 \leq j \leq n \right\}$$

define uma norma em $\mathbb{M}_n(\mathbb{R})$.

Teorema 5.4.1 *Seja V um espaço vetorial sobre o corpo $I\!F$ munido do produto interno $\langle \cdot , \cdot \rangle$. Então, a aplicação $q(\cdot) : V \longrightarrow I\!R$ definida da seguinte forma:*

$$q(u) \;=\; \sqrt{\langle u \,,\, u \rangle} \qquad ; \qquad \forall\, u \,\in\, V \;,$$

satisfaz as propriedades de norma:

(a) Positividade: $q(u) > 0$ *para* $u \neq 0_V$, *com* $q(u) = 0 \iff u = 0_V$

(b) Homogeneidade: $q(\lambda u) = |\,\lambda\,|\, q(u)$ *para todo* $u \in V$, $\lambda \in I\!F$

(c) Desigualdade Triangular: $q(u + v) \leq q(u) + q(v)$ *para todos* u , $v \in V$

Demonstração – Vamos provar que a aplicação $q(\cdot)$ define uma norma em V com relação ao produto interno $\langle \cdot , \cdot \rangle$, que denotamos por $\| \cdot \|_2$, denominada **Norma Euclidiana**. As propriedades **(a)** e **(b)** seguem das propriedades de produto interno.

Para mostrar que a aplicação $\| \cdot \|_2$ satisfaz a propriedade da desigualdade triangular, utilizamos a desigualdade de Cauchy–Schwarz escrita da forma:

$$|\,\langle u \,,\, v \rangle\,| \;\leq\; \|\, u \,\|_2 \,\|\, v \,\|_2 \qquad \text{para todos} \qquad u \,,\, v \,\in\, V \;.$$

Temos que

$$\|\, u + v \,\|_2^2 \;=\; \langle\, u + v \,,\, u + v \,\rangle \;=\; \langle\, u \,,\, u \,\rangle \;+\; \langle\, u \,,\, v \,\rangle \;+\; \langle\, v \,,\, u \,\rangle \;+\; \langle\, v \,,\, v \,\rangle$$

Inicialmente considerando um espaço vetorial real, tem–se que

$$\|\, u + v \,\|_2^2 \;=\; \langle\, u \,,\, u \,\rangle \;+\; 2\,\langle\, u \,,\, v \,\rangle \;+\; \langle\, v \,,\, v \,\rangle \;\leq\; \langle\, u \,,\, u \,\rangle \;+\; 2\;|\,\langle\, u \,,\, v \,\rangle\,| \;+\; \langle\, v \,,\, v \,\rangle$$

Utilizando a desigualdade de Cauchy–Schwarz, obtemos

$$\|\, u + v \,\|_2^2 \;\leq\; \langle\, u \,,\, u \,\rangle \;+\; 2\,\|\, u \,\|_2\,\|\, v \,\|_2 \;+\; \langle\, v \,,\, v \,\rangle \;=\; \|\, u \,\|_2^2 \;+\; 2\,\|\, u \,\|_2\,\|\, v \,\|_2 \;+\; \|\, v \,\|_2^2$$

$$\|\, u + v \,\|_2^2 \;\leq\; (\;\|\, u \,\|_2 \;+\; \|\, v \,\|_2\,)^2 \quad\Longrightarrow\quad \|\, u + v \,\|_2 \;\leq\; \|\, u \,\|_2 \;+\; \|\, v \,\|_2$$

o que completa a prova para o caso de um espaço vetorial real.

Finalmente, para um espaço vetorial complexo, temos que

$$\begin{aligned} \| u + v \|_2^2 &= \langle u , u \rangle + \langle u , v \rangle + \overline{\langle u , v \rangle} + \langle v , v \rangle \\ &= \langle u , u \rangle + 2\, Re(\langle u , v \rangle) + \langle v , v \rangle \\ &\leq \langle u , u \rangle + 2\, | Re(\langle u , v \rangle) | + \langle v , v \rangle \\ &\leq \langle u , u \rangle + 2\, | \langle u , v \rangle | + \langle v , v \rangle \end{aligned}$$

Utilizando a desigualdade de Cauchy–Schwarz, obtemos

$$\| u + v \|_2^2 \leq \langle u , u \rangle + 2 \| u \|_2 \| v \|_2 + \langle v , v \rangle = \| u \|_2^2 + 2 \| u \|_2 \| v \|_2 + \| v \|_2^2$$

Portanto, temos que

$$\| u + v \|_2^2 \leq (\| u \|_2 + \| v \|_2)^2 \qquad \Longrightarrow \qquad \| u + v \|_2 \leq \| u \|_2 + \| v \|_2$$

o que completa a demonstração. ■

Definição 5.4.2 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Uma aplicação*

$$\begin{aligned} d(\cdot, \cdot) : \ V \times V &\longrightarrow \mathbb{R} \\ (u, v) &\longrightarrow d(u, v) \end{aligned}$$

com as propriedades:

(a) Positividade: $d(u,v) \geq 0$ *, com* $d(u,v) = 0 \iff u = v$

(b) Simetria: $d(u,v) = d(v,u) \quad ; \quad \forall\ u, v \in V$

(c) Desigualdade Triangular: $d(u,v) \leq d(u,w) + d(v,w) \quad ; \quad \forall\ u, v, w \in V$

*define uma **métrica**, ou **distância**, no espaço vetorial V.*

Um espaço vetorial V munido de uma métrica $d(\cdot, \cdot)$ é denominado **espaço métrico**, que denotamos por $(V, d(\cdot, \cdot))$.

Teorema 5.4.2 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$ com uma norma $\| \cdot \|$. A aplicação*

$$\begin{aligned} d(\cdot, \cdot) : \ V \times V &\longrightarrow \mathbb{R} \\ (u, v) &\longrightarrow d(u, v) = \| u - v \| \end{aligned}$$

*define uma **métrica** no espaço vetorial V.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 5.4.4 *Considere o espaço vetorial real* $\mathcal{C}([0,1])$. *A aplicação*

$$\| f \|_\infty = \max\{ \, | f(x) | \quad , \quad x \in [0,1] \, \}$$

define uma norma em $\mathcal{C}([0,1])$. *Dada a função* $f(x) = 1 - \exp(-x)$, *calcular* $\| f \|_\infty$.

Exemplo 5.4.5 *Considere o espaço vetorial real* $\mathcal{C}([0,1])$. *A aplicação*

$$\| f \|_1 = \int_0^1 | f(x) | \, dx \qquad ; \qquad \forall \, f \in \mathcal{C}([0,1])$$

define uma norma em $\mathcal{C}([0,1])$.

Exemplo 5.4.6 *Considere o espaço vetorial* $\mathcal{C}([0,1])$ *munido do produto interno usual*

$$\langle \, f \, , \, g \, \rangle = \int_0^1 f(x) g(x) dx \qquad ; \qquad \forall \, f, \, g \in \mathcal{C}([0,1])$$

Sabemos que a aplicação $\| f \|_2 = \sqrt{\langle \, f \, , \, f \, \rangle}$ *define uma norma em* $\mathcal{C}([0,1])$. *Dada a função* $f(x) = \cos(\pi \, x)$, *calcular* $\| f \|_2$.

Exemplo 5.4.7 *Considere o espaço vetorial* $\mathcal{C}([0,1])$ *munido do produto interno usual. Verifique a desigualdade de Cauchy–Schwarz para as funções* $f(x) = x$ *e* $g(x) = x^2$.

Exemplo 5.4.8 *Considere o espaço vetorial* $\mathcal{C}([0,1])$ *munido do produto interno usual e* $\| \cdot \|_2$ *a norma Euclidiana. Dadas as funções* $f(x) = x$ *e* $g(x) = x^2$, *determine* $d(f,g) = \| f - g \|_2$.

Utilizando a definição da métrica Euclidiana, temos que

$$d(f,g) = \| f - g \|_2 = \sqrt{\int_0^1 (f(x) - g(x))^2 dx} = \sqrt{\int_0^1 (x - x^2)^2 dx}$$

$$= \sqrt{\int_0^1 (\, x^2 - 2x^3 + x^4 \,) dx} = \frac{\sqrt{30}}{30} \, .$$

5.5 Definição de Ângulo. Ortogonalidade

Seja V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$. Observe que utilizando a desigualdade de Cauchy–Schwarz mostramos que para quaisquer elementos não–nulos $u, v \in V$ o quociente

$$\frac{\langle u , v \rangle}{\| u \|_2 \, \| v \|_2}$$

está no intervalo $[-1, 1]$. Desse modo, existe um número real $\theta \in [0, 2\pi]$ tal que

$$\frac{\langle u , v \rangle}{\| u \|_2 \, \| v \|_2} = \cos(\theta) .$$

Além disso, existe um único valor $\theta \in [0, \pi]$ satisfazendo a igualdade. Assim, podemos ter a noção de ângulo entre dois elementos de um espaço vetorial munido com um produto interno, que será compatível com a definição de ortogonalidade que apresentamos a seguir.

Definição 5.5.1 (Ângulo) *Seja V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$. O Ângulo entre dois elementos não–nulos $u, v \in V$ é definido como sendo o valor $\theta \in [0, \pi]$ que satisfaz a equação*

$$\cos(\theta) = \frac{\langle u , v \rangle}{\| u \|_2 \, \| v \|_2} .$$

Definição 5.5.2 (Ortogonalidade) *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$. Dizemos que os elementos $u, v \in V$ são ortogonais se, e somente se, $\langle u , v \rangle = 0$, e denotamos por $u \perp v$.*

Podemos observar facilmente que

$$\langle u , v \rangle = 0 \iff \cos(\theta) = 0 \iff \theta = \frac{\pi}{2} \quad \text{para} \quad 0 \leq \theta \leq \pi ,$$

mostrando a compatibilidade entre os conceitos de ângulo e ortogonalidade.

Teorema 5.5.1 *Num espaço vetorial V munido do produto interno $\langle \cdot , \cdot \rangle$, temos as seguintes propriedades:*

1. $0_V \perp v$ *para todo* $v \in V$;
2. $u \perp v$ *implica* $v \perp u$;
3. *Se* $v \perp u$ *para todo* $u \in V$, *então* $v = 0_V$;
4. *Se* $v \perp w$ *e* $u \perp w$, *então* $(v + u) \perp w$;
5. *Se* $v \perp u$ *e* λ *é um escalar, então* $\lambda v \perp u$.

Demonstração – A prova pode ficar a cargo do leitor. $\square$

Exemplo 5.5.1 *Considere o espaço vetorial* $\mathcal{C}([0,1])$ *munido do produto interno usual. Determine o ângulo entre as funções* $f(x) = x$ *e* $g(x) = x^2$*, de acordo com a geometria gerada no espaço vetorial pelo produto interno usual.*

Definição 5.5.3 *Considere* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *munido do produto interno* $\langle \cdot , \cdot \rangle$. *Seja* $S = \{ v_1 , \cdots , v_n \}$ *um conjunto de elementos de* V *com* $\langle v_i , v_j \rangle = 0$ *para* $i \neq j$. *Então, dizemos que* S *é um* ***conjunto ortogonal*** *em* V *com relação ao produto interno* $\langle \cdot , \cdot \rangle$. *Além disso, se* $\| v_j \|_2 = 1$ *para* $j = 1 , \cdots , n$ *, dizemos que* S *é um* ***conjunto ortonormal*** *em* V.

Teorema 5.5.2 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $S = \{ v_1 , \cdots , v_n \}$ *um conjunto ortogonal em* V *com elementos* $v_j \neq 0_V$ *para* $j = 1 , \cdots , n$. *Então,* S *é linearmente independente em* V.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 5.5.3 (Teorema de Pitágoras) *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma proveniente do produto interno. Então, os elementos* $u , v \in V$ *são ortogonais se, e somente se,*

$$\| v + u \|_2^2 = \| u \|_2^2 + \| v \|_2^2 .$$

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 5.5.4 (Lei do Paralelogramo) *Sejam* V *um espaço vetorial complexo com o produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma proveniente do produto interno. Então, para todos* $u, v \in V$ *tem–se que*

$$\| v + u \|_2^2 + \| u - v \|_2^2 = 2 \| u \|_2^2 + 2 \| v \|_2^2 .$$

Demonstração – A prova pode ficar a cargo do leitor. □

Proposição 5.5.1 (Lei dos Cossenos) *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$, $\| \cdot \|_2$ *a norma proveniente do produto interno e os elementos* $u, v \in V$ *não–nulos. Se* θ *é o ângulo entre os elementos* u *e* v*, então*

$$\| u \pm v \|_2^2 = \| u \|_2^2 + \| v \|_2^2 \pm 2 \| u \|_2 \| v \|_2 \cos(\theta) .$$

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 5.5.2 *Considere o espaço vetorial* $\mathcal{C}([-\pi, \pi])$ *com o produto interno usual*

$$\langle f \, , \, g \rangle \;=\; \int_{-\pi}^{\pi} f(x)g(x)dx \qquad ; \qquad \forall \;\, f, g \,\in\, \mathcal{C}([-\pi, \pi]) \;\; .$$

O conjunto $\{\, \sin(x), \; \sin(2x), \; \cdots, \; \sin(nx), \; \cdots \,\}$ *é ortogonal com relação ao produto interno usual de* $\mathcal{C}([-\pi, \pi])$.

Basta utilizar a seguinte identidade trigonométrica

$$\sin(nx) \, \sin(mx) \;=\; \frac{\cos((n-m)x) \;-\; \cos((n+m)x)}{2} \, ,$$

simplificando o cálculo das integrais, para obter o resultado desejado.

Exemplo 5.5.3 *Considere o espaço vetorial* $\mathcal{C}([-\pi, \pi])$ *com o produto interno usual*

$$\langle f \, , \, g \rangle \;=\; \int_{-\pi}^{\pi} f(x)g(x)dx \qquad ; \qquad \forall \;\, f, g \,\in\, \mathcal{C}([-\pi, \pi]) \;\; .$$

O conjunto $\{\, 1, \; \cos(x), \; \cos(2x), \; \cdots, \; \cos(nx), \; \cdots \,\}$ *é ortogonal com relação ao produto interno usual de* $\mathcal{C}([-\pi, \pi])$.

Basta utilizar a seguinte identidade trigonométrica

$$\cos(nx) \, \cos(mx) \;=\; \frac{\cos((n-m)x) \;+\; \cos((n+m)x)}{2} \, ,$$

simplificando o cálculo das integrais, para obter o resultado desejado.

De fato, por simplicidade vamos utilizar a seguinte notação

$$\varphi_k(x) \;=\; \cos(kx) \qquad \text{para} \qquad k \;=\; 0, \, 1, \, 2, \, \cdots, \, n, \, \cdots \, .$$

Para $n \,=\, m \,=\, 0$ temos

$$\langle \, \varphi_0 \, , \, \varphi_0 \, \rangle \;=\; \int_{-\pi}^{\pi} 1 \, dx \;=\; 2\pi \, .$$

Para $n \,=\, m \,\neq\, 0$, utilizando a identidade trigonométrica acima, temos que

$$\begin{aligned} \langle \, \varphi_n \, , \, \varphi_n \, \rangle \;&=\; \frac{1}{2} \int_{-\pi}^{\pi} 1 \, dx \;+\; \frac{1}{2} \int_{-\pi}^{\pi} \cos(2nx)dx \\ &=\; \pi \;+\; \frac{1}{4n} \sin(2nx) \Big|_{-\pi}^{\pi} \;=\; \pi \, . \end{aligned}$$

Para $n \neq m$, utilizando a identidade trigonométrica acima, temos que

$$\begin{aligned} \langle\, \varphi_n \,,\, \varphi_m \,\rangle &= \frac{1}{2}\int_{-\pi}^{\pi} \cos((n-m)x)dx \;+\; \frac{1}{2}\int_{-\pi}^{\pi} \cos((n+m)x)dx \\ &= \frac{1}{2(n-m)}\sin((n-m)x)\Big|_{-\pi}^{\pi} \;+\; \frac{1}{2(n+m)}\sin((n+m)x)\Big|_{-\pi}^{\pi} \;=\; 0\,. \end{aligned}$$

Assim, obtemos o resultado desejado.

Exemplo 5.5.4 *Sejam V um espaço vetorial complexo com o produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$ e $\|\cdot\|_2$ a norma proveniente do produto interno. Vamos mostrar que o produto interno fica completamente determinado pela sua "parte real", isto é, o produto interno pode ser representado da seguinte forma:*

$$\langle\, u \,,\, v \,\rangle \;=\; Re(\langle\, u \,,\, v \,\rangle) \;+\; iRe(\langle\, u \,,\, iv \,\rangle)$$

para todos $u, v \in V$.

De fato, se $z \in \mathbb{C}$, então $Im(z) = Re(-iz)$. Desse modo, temos que

$$\begin{aligned} \langle\, u \,,\, v \,\rangle &= Re(\langle\, u \,,\, v \,\rangle) \;+\; iIm(\langle\, u \,,\, v \,\rangle) \\ &= Re(\langle\, u \,,\, v \,\rangle) \;+\; iRe(-i\langle\, u \,,\, v \,\rangle) \\ &= Re(\langle\, u \,,\, v \,\rangle) \;+\; iRe(\langle\, u \,,\, iv \,\rangle) \end{aligned}$$

o que completa a prova da afirmação.

Exemplo 5.5.5 *Considere o espaço vetorial complexo $\mathbb{C}^3$ munido do produto interno usual, e os elementos $x, y \in \mathbb{C}^3$, representados na forma de vetor coluna, dados por:*

$$X \;=\; \begin{bmatrix} i \\ 2+3i \\ 1+2i \end{bmatrix} \qquad e \qquad Y \;=\; \begin{bmatrix} 3-i \\ 4-2i \\ 5+3i \end{bmatrix}.$$

Desse modo, temos que

$$\langle\, x \,,\, y \,\rangle \;=\; Y^*X \;=\; 12 \;+\; 26i\,.$$

Podemos observar facilmente que

$$\begin{aligned} \langle\, x \,,\, y \,\rangle &= Re(\langle\, u \,,\, v \,\rangle) \;+\; iRe(-i\langle\, u \,,\, v \,\rangle) \\ &= 12 \;+\; iRe(\,26 \;-\; 12i\,) \\ &= 12 \;+\; 26i \end{aligned}$$

exemplificando a propriedade descrita acima.

Exercícios

Exercício 5.14 *Sejam* $a_1,\ a_2,\ \cdots,\ a_n$ *reais quaisquer. Mostre que*

$$\left(\frac{a_1 + a_2 + \cdots + a_n}{n} \right)^2 \leq \frac{(a_1^2 + a_2^2 + \cdots + a_n^2)}{n},$$

utilizando a desigualdade de Cauchy–Schwarz em $\mathbb{R}^n$.

Exercício 5.15 *Sejam* $a_1,\ \cdots,\ a_n$ *reais estritamente positivos. Mostre que*

$$(a_1 + \cdots + a_n) \left(\frac{1}{a_1} + \cdots + \frac{1}{a_n} \right) \geq n^2,$$

utilizando a desigualdade de Cauchy–Schwarz em $\mathbb{R}^n$.

Exercício 5.16 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *com o produto interno*

$$\langle p, q \rangle = \int_0^1 p(x)q(x)dx \qquad ; \qquad \forall\ p, q \in \mathcal{P}_2(\mathbb{R}).$$

Dados os polinômios

$$p(x) = x + 2 \quad , \quad q(x) = 3x - 2 \qquad e \qquad h(x) = x^2 - 3,$$

determine

$$\langle p, q \rangle \quad , \quad \langle p+q, q \rangle \quad , \quad \| p \|_2 \quad , \quad \| q+h \|_2 \quad , \quad d(p,h) \qquad e \qquad \cos(\theta),$$

onde θ *é o ângulo entre os polinômios* $p(x)$ *e* $h(x)$.

Exercício 5.17 *Considere o espaço vetorial real* $\mathbb{M}_{2\times 3}(\mathbb{R})$ *com o produto interno usual*

$$\langle A, B \rangle = tr(B^t A).$$

Dadas as matrizes

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 2 & 1 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \qquad e \qquad C = \begin{bmatrix} 2 & -3 & 2 \\ 1 & 0 & -1 \end{bmatrix},$$

determine

$$\langle A, B \rangle \quad , \quad \langle A+B, C \rangle \quad , \quad \| A \|_2 \quad , \quad \| B \|_2 \qquad e \qquad \cos(\theta),$$

onde θ *é o ângulo entre as matrizes* A *e* B.

Exercício 5.18 *Considere o espaço vetorial real* $\mathbb{R}^4$ *munido do produto interno usual. Determine os valores do parâmetro* α *de modo que os elementos* $u = (1, 2, \alpha, 3)$ *e* $v = (\alpha, 2, \alpha, -2)$ *sejam ortogonais, isto é,* $\langle u, v \rangle = 0$.

Exercício 5.19 *Mostre que a aplicação* $\langle\cdot,\cdot\rangle : \mathcal{P}_1(\mathbb{R}) \times \mathcal{P}_1(\mathbb{R}) \longrightarrow \mathbb{R}$ *dada por*

$$\langle p \,,\, q \rangle \;=\; p(0)q(0) \;+\; p(1)q(1)$$

define um produto interno no espaço vetorial $\mathcal{P}_1(\mathbb{R})$. *Determine todos os polinômios* $q(x) = a + bx \in \mathcal{P}_1(\mathbb{R})$ *que são ortogonais ao polinômio* $p(x) = 1 + x$ *com relação ao produto interno* $\langle\cdot,\cdot\rangle$.

Resposta: $\boxed{q(x) \;=\; a\left(1 \;-\; \dfrac{3}{2}x\right) \quad , \quad a \in \mathbb{R}}$

Exercício 5.20 *Considere o espaço vetorial* $\mathcal{P}_1(\mathbb{R})$ *munido do produto interno*

$$\langle f \,,\, g \rangle \;=\; \int_0^1 f(t)g(t)dt\,.$$

Determine todos os polinômios $q(x) = a + bx \in \mathcal{P}_1(\mathbb{R})$ *que são ortogonais ao polinômio* $p(x) = 1 + x$ *com relação ao produto interno* $\langle\cdot,\cdot\rangle$.

Resposta: $\boxed{q(x) \;=\; a\left(1 \;-\; \dfrac{9}{5}x\right) \quad , \quad a \in \mathbb{R}}$

Exercício 5.21 *Seja* V *um espaço vetorial real munido do produto interno* $\langle\cdot,\cdot\rangle$ *e* $\|\cdot\|_2$ *a norma Euclidiana. Sejam* $u, v \in V$ *não–nulos. Prove que* $\langle u \,,\, v \rangle = 0$ *se, e somente se,* $\|u + \alpha v\|_2 \geq \|u\|_2$ *para todo* $\alpha \in \mathbb{R}$.

Exercício 5.22 *Seja* V *um espaço vetorial real munido do produto interno* $\langle\cdot,\cdot\rangle$. *Mostre que*

$$\|u\|_2 \;=\; \|v\|_2 \qquad \Longleftrightarrow \qquad \langle u+v \,,\, u-v \rangle \;=\; 0\,,$$

para todos $u\,,\, v \in V$.

Exercício 5.23 *Considere o espaço vetorial real* $\mathbb{R}^3$ *munido do produto interno usual. Dados o elementos* $u = (1, 1, 0)$ *e* $v = (0, 1, 1)$, *determine o elemento* $w \in \mathbb{R}^3$ *de modo que* $\|w\|_2 = 1$ *e* $\langle u \,,\, w \rangle = \langle v \,,\, w \rangle = 0$. *Dê uma interpretação geométrica.*

Exercício 5.24 *Considere o espaço vetorial real* $\mathcal{C}([0,1])$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *. Dadas as funções* $g(x) = \exp(-x)$ *e* $f(x) = x$*. Determine*

(a) $\langle f \, , \, g \rangle$.

(b) $\| \, f \, \|_\infty$, $\| \, g \, \|_\infty$, $\| \, f \, \|_2$ *e* $\| \, g \, \|_2$.

(c) $d(f,g) = \| \, f - g \, \|_1$.

(d) $d(f,g) = \| \, f - g \, \|_\infty$.

(e) Verificar a desigualdade de Cauchy-Schwarz aplicada às funções f *e* g .

(f) Verificar a desigualdade triangular $\| \, f + g \, \|_1 \leq \| \, f \, \|_1 + \| \, g \, \|_1$.

Exercício 5.25 *Considere* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma proveniente do produto interno. Mostre que se os elementos* $u \, , \, v \in V$ *são ortogonais, então*

$$\| \, v + u \, \|_2^2 = \| \, u \, \|_2^2 + \| \, v \, \|_2^2 \, .$$

A recíproca é verdadeira ? Justifique sua resposta.

Exercício 5.26 *Considere* V *um espaço vetorial complexo com o produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma proveniente do produto interno. Mostre que*

$$\| \, u \pm v \, \|_2^2 = \| \, u \, \|_2^2 \pm 2Re(\langle \, u \, , \, v \, \rangle) + \| \, v \, \|_2^2$$

para todos $u, \, v \in V$.

Exercício 5.27 *Considere* V *um espaço vetorial real com o produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma proveniente do produto interno. Mostre que*

$$\langle \, u \, , \, v \, \rangle = \frac{1}{4} \| \, u + v \, \|_2^2 - \frac{1}{4} \| \, u - v \, \|_2^2$$

para todos $u, \, v \in V$.

Exercício 5.28 *Considere* V *um espaço vetorial complexo com o produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma proveniente do produto interno. Mostre que*

$$\begin{aligned} \langle \, u \, , \, v \, \rangle &= \frac{1}{4} \| \, u + v \, \|_2^2 - \frac{1}{4} \| \, u - v \, \|_2^2 \\ &+ \frac{i}{4} \| \, u + iv \, \|_2^2 - \frac{i}{4} \| \, u - iv \, \|_2^2 \end{aligned}$$

para todos $u, \, v \in V$.

Exercício 5.29 *Seja* V *um espaço vetorial complexo munido com o produto interno* $\langle \cdot , \cdot \rangle$. *Mostre que o ângulo entre dois elementos não–nulos* $u, v \in V$ *pode ser definido, sem ambiguidade, como sendo o valor* $\theta \in \left[0 , \frac{\pi}{2} \right]$ *que satisfaz a equação*

$$\cos(\theta) \;=\; \frac{|\langle u , v \rangle|}{\| u \|_2 \, \| v \|_2} \,.$$

Exercício 5.30 *Sejam* V *um espaço vetorial real e* $\| \cdot \|$ *uma norma em* V *que satisfaz a Lei do Paralelogramo, isto é,*

$$\| u + v \|^2 \;+\; \| u - v \|^2 \;=\; 2\| u \|^2 \;+\; 2\| v \|^2$$

para todos $u, v \in V$. *Mostre que a aplicação* $\langle \cdot , \cdot \rangle$ *definida por:*

$$\langle u , v \rangle \;=\; \frac{1}{4} \left(\| u + v \|^2 \,-\, \| u - v \|^2 \right)$$

para todos $u, v \in V$, *define um produto interno em* V *tal que*

$$\| u \|^2 \;=\; \langle u , u \rangle$$

para todo $u \in V$.

Exercício 5.31 *Considere o espaço vetorial* $\mathcal{C}([0,1])$ *munido do produto interno usual. Determine uma função afim* $f(x) \;=\; ax + b$, *com* $a, b \in I\!R$, *que seja ortogonal com a função* $g(x) \;=\; x^2$, *de acordo com a geometria gerada no espaço vetorial pelo produto interno usual.*

5.6 Base Ortogonal. Coeficientes de Fourier

Definição 5.6.1 *Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$. Dizemos que uma base $\beta = \{ q_1 , \cdots , q_n \}$ de V é uma* ***base ortogonal*** *se β é um conjunto ortogonal em V. No caso em que o conjunto β é ortonormal, dizemos que β é uma* ***base ortonormal*** *de V.*

Teorema 5.6.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$ e $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortogonal de V. Então, todo elemento $u \in V$ é escrito de modo único da seguinte forma:*

$$u = \sum_{i=1}^{n} \alpha_i \, q_i \qquad com \qquad \alpha_i = \frac{\langle u , q_i \rangle}{\langle q_i , q_i \rangle} .$$

Neste caso, as coordenadas de u com relação à base ortogonal β são denominadas ***coeficientes de Fourier*** *de u com relação à base ortogonal β.*

Demonstração – Dado um elemento $u \in V$, como $\beta = \{ q_1 , \cdots , q_n \}$ é uma base para V, pelo Teorema 3.7.1, existem escalares $\alpha_1, \cdots , \alpha_n$ tais que o elemento u é escrito de modo único como:

$$u = \sum_{j=1}^{n} \alpha_j \, q_j .$$

Fazendo o produto interno entre o elemento u e um elemento q_i da base ortogonal β, obtemos

$$\langle u , q_i \rangle = \sum_{j=1}^{n} \alpha_j \langle q_j , q_i \rangle = \alpha_i \langle q_i , q_i \rangle \qquad \text{para} \qquad i = 1, \cdots , n .$$

Assim, temos que as coordenadas, coeficientes de Fourier, do elemento u em relação à base ortogonal β são dadas por:

$$\alpha_i = \frac{\langle u , q_i \rangle}{\langle q_i , q_i \rangle} \qquad \text{para} \qquad i = 1, \cdots , n .$$

No caso em que β é uma base ortonormal, temos que os coeficientes de Fourier do elemento u são dados por:

$$\alpha_i = \langle u , q_i \rangle \qquad \text{para} \qquad i = 1, \cdots , n ,$$

o que completa a demonstração. ∎

Definição 5.6.2 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $\beta = \{ q_1, \cdots , q_n \}$ *um conjunto ortogonal em* V *com elementos* $q_j \neq 0_V$ *para* $j = 1 , \cdots , n$*. Os* ***coeficientes de Fourier*** *do elemento* $u \in V$ *relativos ao conjunto ortogonal* β *são definidos como:*

$$\alpha_i = \frac{\langle u , q_i \rangle}{\langle q_i , q_i \rangle} \qquad para \qquad i = 1, \cdots , n ,$$

em homenagem ao matemático Francês ***Jean Baptiste Fourier****.*

Exemplo 5.6.1 *Considere o espaço vetorial complexo* $\mathbb{C}([0, 2\pi])$ *definido por:*

$$\mathbb{C}([0, 2\pi]) = \{ f : [0, 2\pi] \longrightarrow \mathbb{C} \ / \ f \ \text{é uma função contínua} \} ,$$

isto é, o espaço vetorial das funções complexas contínuas definidas em $[0, 2\pi]$*, onde a operação de adição e a operação de multiplicação por escalar são as mesmas definidas no espaço das funções reais. Uma função* $f \in \mathbb{C}([0, 2\pi])$ *é escrita da seguinte forma:*

$$f(x) = f_1(x) + i\, f_2(x) \qquad para \qquad x \in [0, 2\pi] ,$$

onde f_1 *e* f_2 *são funções reais contínuas em* $[0, 2\pi]$*.*

Definimos em $\mathbb{C}([0, 2\pi])$ *o seguinte produto interno*

$$\langle f , g \rangle = \int_0^{2\pi} f(x)\overline{g(x)}dx .$$

Podemos verificar facilmente que o subconjunto S *de* $\mathbb{C}([0, 2\pi])$ *definido por:*

$$S = \{ f_k(x) = \exp(ikx) \ ; \ x \in [0, 2\pi] \quad e \quad k = -r, \cdots , -1, 0, 1, \cdots , r \}$$

é um ***conjunto ortogonal*** *em* $\mathbb{C}([0, 2\pi])$*. Além disso, temos que* $\langle f_k , f_k \rangle = 2\pi$*.*

De fato, para $m \neq n$, e lembrando que $\exp(ikx) = \cos(kx) + i \sin(kx)$, temos que

$$\langle f_n , f_m \rangle = \int_0^{2\pi} \exp(inx) \exp(-imx)dx = \frac{1}{n - m} \exp(i(n - m)x) \Big|_0^{2\pi} = 0 ,$$

Além disso, temos que

$$\langle f_k , f_k \rangle = \int_0^{2\pi} \exp(ikx) \exp(-ikx)dx = \int_0^{2\pi} dx = 2\pi ,$$

o que completa a demonstração.

Exemplo 5.6.2 *Considere o espaço vetorial complexo* $\mathbb{C}([0, 2\pi])$ *com o produto interno*

$$\langle\, f \,,\, g \,\rangle \;=\; \int_0^{2\pi} f(x)\overline{g(x)}dx$$

e o conjunto ortogonal

$$S \;=\; \{\; f_k(x) \;=\; \exp(ikx) \;\;;\;\; x \in [0, 2\pi] \quad e \quad k \;=\; -m, \cdots, -1, 0, 1, \cdots, m \}\,.$$

Os ***coeficientes de Fourier*** *de uma função* $f \in \mathbb{C}([0, 2\pi])$ *relativos ao conjunto ortogonal* S *são dados por:*

$$\alpha_k \;=\; \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\overline{\exp(ikx)}dx \;=\; \frac{1}{2\pi}\int_0^{2\pi} f(x)\exp(-ikx)dx$$

para $k \,=\, -m, \cdots, -1, 0, 1, \cdots, m.$

Para exemplificar, os coeficientes de Fourier da função $f(x) \,=\, x$ são dados por:

$$\alpha_k \;=\; \frac{1}{2\pi}\int_0^{2\pi} x\,\overline{\exp(ikx)}dx \;=\; \frac{1}{2\pi}\int_0^{2\pi} x\,\exp(-ikx)dx \;=\; -\frac{1}{ik}$$

para $k \neq 0$. De fato, usando a técnica de integração por partes, obtemos

$$\int_0^{2\pi} x\,\exp(-ikx)dx \;=\; -\frac{x}{ik}\exp(-ikx)\Big|_0^{2\pi} \;+\; \int_0^{2\pi}\exp(-ikx)dx\,.$$

Agora, temos que

$$-\frac{x}{ik}\exp(-ikx)\Big|_0^{2\pi} \;=-\frac{2\pi}{ik}\exp(-i2k\pi) \;=\; -\frac{2\pi}{ik}\,(\,\cos(2k\pi) \;-\; i\sin(2k\pi)\,) \;=\; -\frac{2\pi}{ik}\,.$$

$$\int_0^{2\pi}\exp(-ikx)dx \;=\; -\frac{1}{ik}\exp(-ikx)\Big|_0^{2\pi} \;=\; -\frac{1}{ik}\,(\,\cos(2k\pi) \;-\; 1\,) \;=\; 0\,.$$

Assim, obtemos os coeficientes α_k para $k \neq 0$.

Para $k \,=\, 0$, temos que

$$\alpha_0 \;=\; \frac{1}{2\pi}\int_0^{2\pi} x\,dx \;=\; \pi\,.$$

Exemplo 5.6.3 *Considere o espaço vetorial real* $\mathcal{C}([-\pi,\pi])$ *com o produto interno usual*

$$\langle\, f \,,\, g \,\rangle \;=\; \int_{-\pi}^{\pi} f(x)g(x)dx \qquad ; \qquad \forall\; f, g \;\in\; \mathcal{C}([-\pi,\pi])\,.$$

Considere o subespaço vetorial

$$W \;=\; \{\; f \;\in\; \mathcal{C}([-\pi,\pi]) \;\;/\;\; f(-x) \;=\; -f(x) \;\}\,.$$

Sabemos que $\beta \;=\; \{\, \sin(x)\,,\, \sin(2x)\,,\, \cdots,\, \sin(nx) \,\}$ *é um conjunto ortogonal em* W*.*

Os coeficientes de Fourier da função $f(x) \;=\; x$*, para* $x \;\in\; [-\pi,\pi]$*, com relação ao conjunto ortogonal* β*, são dados por:*

$$\alpha_k \;=\; \begin{cases} \dfrac{2}{k} & , \quad k \;\; impar \\[2ex] -\dfrac{2}{k} & , \quad k \;\; par \end{cases}$$

para $k \;=\; 1,\; \cdots\;,\; n$*.*

Exemplo 5.6.4 *Considere o espaço vetorial real* $I\!R^2$ *munido do produto interno usual. Seja* $\beta \;=\; \{\; (1,1),\; (-1,1) \;\}$ *uma base ortogonal para* $I\!R^2$*. Calcular as coordenadas do elemento* $v \;=\; (3,4) \;\in\; I\!R^2$ *em relação à base ortogonal* β*.*

Por simplicidade, chamando $q_1 \;=\; (1,1)$ e $q_2 \;=\; (-1,1)$, temos que o elemento $v \in I\!R^2$ é escrito de modo único da seguinte forma:

$$v \;=\; \frac{\langle\, v\,,\, q_1 \,\rangle}{\langle\, q_1\,,\, q_1 \,\rangle}\, q_1 \;+\; \frac{\langle\, v\,,\, q_2 \,\rangle}{\langle\, q_2\,,\, q_2 \,\rangle}\, q_2 \;=\; \frac{7}{2}\, q_1 \;+\; \frac{1}{2}\, q_2\,.$$

Assim, o vetor de coordenadas de $v \;\in\; I\!R^2$ em relação à base ortogonal β é dado por:

$$[v]_\beta \;=\; \frac{1}{2}\begin{bmatrix} 7 \\ 1 \end{bmatrix}\,,$$

onde os coeficientes de Fourier do elemento v em relação à base ortogonal β são

$$\alpha_1 \;=\; \frac{7}{2} \qquad \text{e} \qquad \alpha_2 \;=\; \frac{1}{2}\,.$$

Exemplo 5.6.5 *Considere o espaço vetorial real* $\mathcal{C}([-\pi,\pi])$ *com o produto interno usual*

$$\langle\, f \,,\, g \,\rangle \;=\; \int_{-\pi}^{\pi} f(x)g(x)dx \qquad ; \qquad \forall\; f,g \,\in\, \mathcal{C}([-\pi,\pi])\,.$$

Considere o subespaço vetorial

$$U \,=\, \{\, f \,\in\, \mathcal{C}([-\pi,\pi]) \;/\; f(-x) \,=\, f(x) \,\}\,.$$

Sabemos que $\gamma \,=\, \{\, 1\,,\, \cos(x)\,,\, \cos(2x)\,,\, \cdots,\, \cos(nx) \,\}$ *é um conjunto ortogonal em* U*.*

Vamos calcular os coeficientes de Fourier da função $f(x) \,=\, x^2$*, para* $x \in [-\pi,\pi]$*, com relação ao conjunto ortogonal* γ*.*

Por simplicidade vamos utilizar a seguinte notação

$$\varphi_0(x) \,=\, 1\,, \qquad \varphi_k(x) \,=\, \cos(kx) \qquad \text{para} \qquad k \,=\, 1, \cdots, n\,.$$

Do Exemplo 5.5.3, sabemos que

$$\langle\, \varphi_0\,,\, \varphi_0 \,\rangle \,=\, 2\pi \qquad \text{e} \qquad \langle\, \varphi_k\,,\, \varphi_k \,\rangle \,=\, \pi \qquad \text{para} \qquad k \,=\, 1, \cdots, n\,.$$

Podemos verificar facilmente que

$$\langle\, f\,,\, \varphi_k \,\rangle \;=\; \frac{4\pi}{k^2}\cos(k\pi) \qquad \text{para} \qquad k \,=\, 1, \cdots, n\,.$$

Desse modo, obtemos

$$\alpha_0 \;=\; \frac{\langle\, f\,,\, \varphi_0 \,\rangle}{\langle\, \varphi_0\,,\, \varphi_0 \,\rangle} \;=\; \frac{\pi^2}{3}$$

$$\alpha_k \;=\; \frac{\langle\, f\,,\, \varphi_k \,\rangle}{\langle\, \varphi_k\,,\, \varphi_k \,\rangle} \;=\; \begin{cases} \dfrac{4}{k^2} & , \quad k \;\; \text{par} \\[2ex] -\dfrac{4}{k^2} & , \quad k \;\; \text{impar} \end{cases} \qquad \text{para} \qquad k \,=\, 1, \cdots, n$$

que são os coeficientes de Fourier da função f com relação ao conjunto ortogonal γ.

5.7 Processo de Gram–Schmidt

Teorema 5.7.1 *Considere V um espaço vetorial sobre o corpo $\mathbb{F}$ munido do produto interno $\langle \cdot , \cdot \rangle$. Sejam $v_1 , v_2 , \cdots , v_n , \cdots$ uma seqüência finita ou infinita de elementos de V e $S_k = [\, v_1 , \cdots , v_k \,]$ o subespaço gerado pelos k primeiros elementos. Então, existe uma seqüência correspondente de elementos $q_1 , q_2 , \cdots , q_n , \cdots$ em V a qual possui as seguintes propriedades:*

(a) O elemento q_k é ortogonal a todo elemento do subespaço $[\, q_1 , \cdots , q_{k-1} \,]$.

(b) O subespaço $S_k = [\, v_1 , \cdots , v_k \,]$ é igual ao subespaço $W_k = [\, q_1 , \cdots , q_k \,]$.

(c) A seqüência $q_1 , q_2 , \cdots , q_n , \cdots$ é única, a menos de uma constante multiplicativa, isto é, se existir uma outra seqüência $q'_1 , q'_2 , \cdots , q'_n , \cdots$, de elementos de V satisfazendo as propriedades (a) e (b), então existem escalares $c_k \in \mathbb{F}$ tais que $q'_k = c_k \, q_k$ para $k = 1, 2, \cdots , n, \cdots$.

Demonstração – Vamos construir os elementos $q_1 , q_2 , \cdots , q_n , \cdots$ por um processo de indução sobre k. Inicialmente, escolhemos $q_1 = v_1$. Agora vamos assumir que já construímos os elementos $q_1 , \cdots , q_r$ tais que *(a)* e *(b)* são satisfeitas quando $k = r$.

Desse modo, definimos o elemento q_{r+1} pela equação

$$q_{r+1} = v_{r+1} - \sum_{i=1}^{r} \alpha_i \, q_i \, ,$$

onde os escalares $\alpha_i \in \mathbb{F}$ são escolhidos de modo conveniente. Para $j \leq r$, calculamos

$$\langle \, q_{r+1} \, , \, q_j \, \rangle = \langle \, v_{r+1} \, , \, q_j \, \rangle - \sum_{i=1}^{r} \alpha_i \, \langle \, q_i \, , \, q_j \, \rangle = \langle \, v_{r+1} \, , \, q_j \, \rangle - \alpha_j \, \langle \, q_j \, , \, q_j \, \rangle \, ,$$

pois $\langle \, q_i \, , \, q_j \, \rangle = 0$ para $i \neq j$.

Se $q_j \neq 0_V$, construímos q_{r+1} ortogonal a q_j escolhendo

$$\alpha_j = \frac{\langle \, v_{r+1} \, , \, q_j \, \rangle}{\langle \, q_j \, , \, q_j \, \rangle} \, .$$

Caso $q_j = 0_V$, temos que q_{r+1} é ortogonal a q_j para qualquer escolha de α_j. Assim, escolhemos $\alpha_j = 0$. Desse modo, o elemento q_{r+1} fica bem definido e é ortogonal aos elementos $q_1 , \cdots , q_r$. Portanto, o elemento q_{r+1} é ortogonal a todo elemento do subespaço $[\, q_1 , \cdots , q_r \,]$. Portanto, provamos a propriedade *(a)* quando $k = r + 1$.

Para provar a propriedade *(b)*, para $k = r+1$, devemos mostrar que

$$S_{r+1} = [\, v_1 \,, \cdots , v_{r+1} \,] = W_{r+1} = [\, q_1 \,, \cdots , q_{r+1} \,]$$

tomando por hipótese que

$$S_r = [\, v_1 \,, \cdots , v_r \,] = W_r = [\, q_1 \,, \cdots , q_r \,] \,.$$

Os elementos $q_1 \,, \cdots , q_r$ pertencem ao subespaço S_r e também ao subespaço S_{r+1}.

Sabemos que o novo elemento q_{r+1} é escrito da seguinte forma:

$$q_{r+1} = v_{r+1} - \sum_{i=1}^{r} \alpha_i \, q_i \,.$$

Assim, o elemento q_{r+1} é escrito como a diferença de dois elementos que pertencem ao subespaço S_{r+1}. Desse modo, o elemento $q_{r+1} \in S_{r+1}$. Logo, provamos que

$$[\, q_1 \,, \cdots , q_{r+1} \,] \subseteq [\, v_1 \,, \cdots , v_{r+1} \,] \,.$$

De modo análogo, temos que o elemento v_{r+1} é escrito como:

$$v_{r+1} = q_{r+1} + \sum_{i=1}^{r} \alpha_i \, q_i \,.$$

Assim, o elemento v_{r+1} é escrito como a soma de dois elementos que pertencem ao subespaço W_{r+1}. Desse modo, o elemento $v_{r+1} \in W_{r+1}$. Logo, provamos que

$$[\, v_1 \,, \cdots , v_{r+1} \,] \subseteq [\, q_1 \,, \cdots , q_{r+1} \,] \,.$$

Portanto, provamos a propriedade *(b)* quanto $k = r+1$.

Finalmente, vamos provar a propriedade *(c)* por um processo de indução sobre k. Para $k = 1$, o resultado segue trivialmente. Vamos assumir que a propriedade *(c)* é válida para $k = r$ e considerar o elemento q'_{r+1}. Pela propriedade *(b)*, temos que o elemento $q'_{r+1} \in W_{r+1}$. Assim, podemos escrever

$$q'_{r+1} = \sum_{i=1}^{r+1} c_i \, q_i = w_r + c_{r+1} \, q_{r+1} \,,$$

onde o elemento $w_r \in W_r$. Agora, basta provar que $w_r = 0_V$. Pela propriedade *(a)*, sabemos que os elementos q'_{r+1} e $c_{r+1} \, q_{r+1}$ são ortogonais ao elemento w_r. Desse modo, obtemos

$$\langle \, q'_{r+1} \,, w_r \, \rangle = \langle \, w_r \,, w_r \, \rangle + \langle \, c_{r+1} \, q_{r+1} \,, w_r \, \rangle \implies \langle \, w_r \,, w_r \, \rangle = 0 \,.$$

Logo, $w_r = 0_V$, o que completa a demonstração do **processo de ortogonalização**. ■

Do processo de ortogonalização, sabemos que o elemento q_{r+1} é escrito da forma:

$$q_{r+1} = v_{r+1} - \sum_{i=1}^{r} \alpha_i \, q_i \,.$$

Desse modo, considerando que $q_{r+1} = 0_V$ para algum r, obtemos que o elemento v_{r+1} é uma combinação linear dos elementos $q_1 , \cdots , q_r$, e também dos elementos $v_1 , \cdots , v_r$. Logo, os elementos $v_1 , \cdots , v_r , v_{r+1}$ são linearmente dependentes em V.

Portanto, se os elementos $v_1 , \cdots , v_k$ são linearmente independentes, então os elementos $q_1 , \cdots , q_k$ são não–nulos. Assim, o **processo de ortogonalização de Gram–Schmidt** pode ser descrito da seguinte forma:

$$q_1 = v_1 \qquad \text{e} \qquad q_{r+1} = v_{r+1} - \sum_{i=1}^{r} \frac{\langle\, v_{r+1} \,,\, q_i \,\rangle}{\langle\, q_i \,,\, q_i \,\rangle} \, q_i$$

para $r = 1, \cdots , k-1$.

Exemplo 5.7.1 *Considere o espaço vetorial real $I\!R^2$ munido do produto interno usual e $\gamma = \{ (2,1), (1,1) \}$ uma base ordenada do $I\!R^2$. Obter a partir de γ uma base ordenada ortogonal para $I\!R^2$ com relação ao produto interno usual.*

Vamos utilizar o processo de ortogonalização de Gram–Schmidt. Inicialmente, escolhemos

$$q_1 = v_1 = (2,1)\,.$$

Em seguida, construímos o elemento q_2 da seguinte forma:

$$q_2 = v_2 - \alpha_{12} \, q_1 \,,$$

ortogonal ao subespaço $W_1 = [q_1]$. Assim, temos que

$$\alpha_{12} = \frac{\langle\, v_2 \,,\, q_1 \,\rangle}{\langle\, q_1 \,,\, q_1 \,\rangle} = \frac{3}{5}\,,$$

obtendo

$$q_2 = (1,1) - \frac{3}{5}(2,1) = \left(-\frac{1}{5} \,,\, \frac{2}{5} \right).$$

Assim, fazendo uso da propriedade *(c)* do Teorema 5.7.1, obtemos

$$\beta = \left\{ (2,1)\,,\, \left(-\frac{1}{5} \,,\, \frac{2}{5} \right) \right\} \qquad \text{ou} \qquad \beta = \{\, (2,1)\,,\, (-1,2) \,\}$$

uma base ortogonal para $I\!R^2$.

Teorema 5.7.2 *Todo espaço vetorial de dimensão finita munido de um produto interno possui uma base ortonormal.*

Demonstração – Sejam V um espaço vetorial munido de um produto interno $\langle \cdot , \cdot \rangle$ e $\beta = \{ v_1 , \cdots , v_n \}$ uma base ordenada para V. A partir da base ordenada β, vamos obter uma base ortogonal, através do processo de ortogonalização de Gram–Schmidt.

Em primeiro lugar, seja $q_1 = v_1$. Note que o subespaço $S_1 = [v_1]$ é igual ao subespaço $W_1 = [q_1]$. Agora vamos construir um vetor q_2 que seja ortogonal ao subespaço W_1 e que o subespaço $S_2 = [v_1 , v_2]$ seja igual ao subespaço $W_2 = [q_1 , q_2]$. Então,

$$q_2 = v_2 - \alpha_{12} q_1 \qquad \Longrightarrow \qquad \alpha_{12} = \frac{\langle v_2 , q_1 \rangle}{\| q_1 \|_2^2}$$

Como v_1 e v_2 são linearmente independentes, temos que $q_2 \neq 0_V$.

Agora vamos construir um vetor q_3 que seja ortogonal ao subespaço W_2 e que o subespaço $S_3 = [v_1 , v_2 , v_3]$ seja igual ao subespaço $W_3 = [q_1 , q_2 , q_3]$. Então,

$$q_3 = v_3 - \alpha_{13} q_1 - \alpha_{23} q_2 \quad \Longrightarrow \quad \alpha_{13} = \frac{\langle v_3 , q_1 \rangle}{\| q_1 \|_2^2} \qquad \text{e} \qquad \alpha_{23} = \frac{\langle v_3 , q_2 \rangle}{\| q_2 \|_2^2}$$

Como v_1 , v_2 , v_3 são linearmente independente, temos que $q_3 \neq 0_V$.

Repetindo o processo para $j = 2, \cdots n$, temos que

$$q_j = v_j - \sum_{i=1}^{j-1} \alpha_{ij} q_i \quad \Longrightarrow \quad \alpha_{ij} = \frac{\langle v_j , q_i \rangle}{\| q_i \|_2^2} \qquad \text{para} \qquad i = 1, \cdots , j-1 \,.$$

Como $v_1 , \cdots , v_j$ são linearmente independentes, temos que $q_j \neq 0_V$. Além disso, temos que o subespaço $S_j = [v_1 , \cdots , v_j]$ é igual ao subespaço $W_j = [q_1 , \cdots , q_j]$.

Assim, obtemos uma base ortogonal $\{ q_1 , \cdots , q_n \}$. Finalmente, fazendo

$$q_j^* = \frac{q_j}{\| q_j \|_2} \qquad \text{para} \qquad j = 1, \cdots , n$$

obtemos uma base ortonormal $\beta^* = \{ q_1^* , \cdots , q_n^* \}$, o que completa a prova. ∎

Exemplo 5.7.2 *Considere o espaço vetorial* $\mathcal{P}_3(I\!\!R)$ *munido do produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx\,.$$

Obter a partir da base $\beta \,=\, \{\, 1,\, x,\, x^2,\, x^3 \,\}$ *uma base ortogonal* $\gamma \,=\, \{\, P_0,\, \cdots\,,\, P_3 \,\}$*.*

Utilizando o processo de ortogonalização de Gram–Schmidt, obtemos

$$P_0(x) \,=\, 1 \quad , \quad P_1(x) \,=\, x \quad , \quad P_2(x) \,=\, x^2 \,-\, \frac{1}{3} \qquad \text{e} \qquad P_3(x) \,=\, x^3 \,-\, \frac{3}{5}x\,.$$

Os polinômios $P_0,\, \cdots\,,\, P_3$ são denominados **polinômios ortogonais de Legendre**. Esta denominação é em homenagem ao matemático Frances A. M. Legendre (1752–1833) que encontrou tais polinômios em seus estudos sobre a Teoria do Potencial.

Exemplo 5.7.3 *Considere o espaço vetorial* $\mathcal{P}_3(I\!\!R)$ *munido do produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{0}^{\infty} \exp(-x)p(x)q(x)dx\,.$$

Obter a partir da base $\beta \,=\, \{\, 1,\, x,\, x^2,\, x^3 \,\}$ *uma base ortogonal* $\gamma \,=\, \{\, L_0,\, \cdots\,,\, L_3 \,\}$*.*

Utilizando o processo de ortogonalização de Gram–Schmidt, obtemos os polinômios

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x \,-\, 1 \\ L_2(x) &= x^2 \,-\, 4x \,+\, 2 \\ L_3(x) &= x^3 \,-\, 9x^2 \,+\, 18x \,-\, 6 \end{aligned}$$

que são denominados **polinômios ortogonais de Laguerre**.

Na construção dos polinômios de Laguerre utilizamos o seguinte resultado

$$\int_{0}^{\infty} \exp(-x)\, x^n dx \;=\; n! \qquad \text{para} \qquad n \,\in\, I\!\!N\,,$$

que é facilmente obtido através do processo de indução sobre n, juntamente com a técnica de integração por partes.

Exemplo 5.7.4 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual e* $\beta = \{ v_1 = (1,1,1), v_2 = (0,2,1), v_3 = (0,0,1) \}$ *uma base do* $I\!R^3$*. Obter a partir de* β *uma base ortogonal* $\beta^* = \{ q_1, q_2, q_3 \}$ *para o* $I\!R^3$ *com relação ao produto interno usual.*

Vamos utilizar o processo de ortogonalização de Gram–Schmidt. Inicialmente, escolhemos

$$q_1 = v_1 = (1,1,1) .$$

Em seguida, construímos q_2 da seguinte forma:

$$q_2 = v_2 - \alpha_{12} \, q_1 ,$$

ortogonal ao subespaço $W_1 = [q_1]$. Assim, temos que

$$\alpha_{12} = \frac{\langle v_2 , q_1 \rangle}{\langle q_1 , q_1 \rangle} = \frac{3}{3} = 1$$

obtendo $q_2 = (-1,1,0)$.

Finalmente, construímos q_3 da seguinte forma:

$$q_3 = v_3 - \alpha_{13} \, q_1 - \alpha_{23} \, q_2 ,$$

ortogonal ao subespaço $W_2 = [q_1, q_2]$. Assim, temos que

$$\alpha_{13} = \frac{\langle v_3 , q_1 \rangle}{\langle q_1 , q_1 \rangle} = \frac{1}{3} \qquad \text{e} \qquad \alpha_{23} = \frac{\langle v_3 , q_2 \rangle}{\langle q_2 , q_2 \rangle} = 0 ,$$

obtendo $q_3 = \dfrac{1}{3}(-1,-1,2)$, ou $q_3' = (-1,-1,2)$.

Portanto, temos a seguinte base ortogonal para o $I\!R^3$

$$\beta^* = \{ (1,1,1), (-1,1,0), (-1,-1,2) \}$$

de acordo com o Teorema 5.7.1.

Exercícios

Exercício 5.32 *Considere o conjunto* Γ *formado pelos seguintes elementos de* $I\!R^3$

$$u_1 = (1, 1, 1)\,, \qquad u_2 = (1, 2, -3) \qquad e \qquad u_3 = (5, -4, -1)\,.$$

(a) Mostre que Γ *é uma base ortogonal para o* $I\!R^3$ *.*

(b) Encontre as coordenadas do elemento $v = (1, 5, -7)$ *com relação à base* Γ*.*

Exercício 5.33 *Considere o espaço vetorial real* $I\!R^4$ *munido do produto interno usual. Determine uma base ortogonal para o subespaço* S *de* $I\!R^4$ *gerado pelos elementos*

$$v_1 = (1, 1, 1, 1)\,, \qquad v_2 = (1, 1, 2, 4) \qquad e \qquad v_3 = (1, 2, -4, -3)\,.$$

Exercício 5.34 *Considere o espaço vetorial real* $\mathcal{C}([-\pi, \pi])$ *com o produto interno usual e o subespaço vetorial* $U = \{\, f \in \mathcal{C}([-\pi, \pi]) \ / \ f(-x) = f(x) \,\}$*. Seja*

$$\gamma = \{\, 1,\ \cos(x),\ \cos(2x),\ \cdots,\ \cos(nx) \,\}$$

um conjunto ortogonal em U*. Calcular os coeficientes de Fourier da função* $f(x) = |\, x \,|$*, para* $x \in [-\pi, \pi]$*, com relação ao conjunto ortogonal* γ*.*

Exercício 5.35 *Considere o espaço vetorial real* $\mathcal{P}_2(I\!R)$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle = \int_0^1 p(t)q(t)dt\,.$$

Aplique o processo de ortogonalização de Gram-Schmidt na base canônica $\{\, 1,\, t,\, t^2 \,\}$ *para obter uma base ortonormal* $\{\, P_0,\, P_1,\, P_2 \,\}$*.*

Resposta: $\boxed{P_0(t) = 1 \quad , \quad P_1(t) = \sqrt{3}(\,2t - 1\,) \quad \text{e} \quad P_2(t) = \sqrt{5}(\,6t^2 - 6t + 1\,)}$

Exercício 5.36 *Considere o espaço vetorial real* $\mathcal{P}_1(I\!R)$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle = \int_{-1}^1 p(x)q(x)dx \qquad ; \qquad \forall\ p, q \in \mathcal{P}_1(I\!R)\,.$$

Seja $T : \mathcal{P}_1(I\!R) \longrightarrow I\!R$ *a transformação linear definida por:* $T(p)(x) = p(1)$*. Determine um elemento* $q \in \mathcal{P}_1(I\!R)$ *tal que* $T(p) = \langle\, p \,,\, q \,\rangle$ *para* $p \in \mathcal{P}_1(I\!R)$*.*

Exercício 5.37 *Considere o espaço vetorial* $\mathbb{M}_2(\mathbb{R})$ *munido do produto interno usual. Determine uma base ortogonal para o subespaço* W *definido por:*

$$W = \{ A \in \mathbb{M}_2(\mathbb{R}) \ / \ A \ \textit{é triangular inferior} \}$$

a partir da base $\beta = \{ A_1, A_2, A_3 \}$ *formada pelos elementos*

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad e \quad A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} .$$

Exercício 5.38 *Considere o espaço vetorial* $\mathbb{R}^4$ *munido do produto interno usual. Seja* $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ *a transformação linear definida por:*

$$T(x,y,z,t) = (x - y - z + t, -x + z + t) .$$

Determine uma base ortogonal para o subespaço $Ker(T)$.

Exercício 5.39 *Considere o espaço vetorial real* $\mathbb{R}^3$ *munido do produto interno usual. Determine uma base ortogonal para o subespaço* W *definido por:*

$$W = \{ (x,y,z) \in \mathbb{R}^3 \ / \ x - y - z = 0 \} .$$

5.8 Complemento Ortogonal

Definição 5.8.1 *Sejam V um espaço vetorial munido do produto interno $\langle \cdot , \cdot \rangle$ e S um conjunto não vazio de elementos de V. O conjunto $S^{\perp}$ definido por:*

$$S^{\perp} = \{ u \in V \ / \ \langle u , v \rangle = 0 \quad \textit{para todo} \quad v \in S \} ,$$

*é denominado "S perpendicular". No caso em que S é um subespaço vetorial de V, o conjunto $S^{\perp}$ é denominado **complemento ortogonal** de S em V.*

Teorema 5.8.1 *O conjunto $S^{\perp}$ é um subespaço de V, mesmo que S não o seja. Além disso, tem–se que $S \cap S^{\perp} = \{ 0_V \}$ no caso em que S é um subespaço de V.*

Demonstração – Temos que $S^{\perp} \neq \emptyset$, pois $\langle 0_V , v \rangle = 0$ para todo $v \in S$. Desse modo, temos que $O_V \in S^{\perp}$. Agora, basta mostrar que $S^{\perp}$ satisfaz as condições do Teorema 3.2.1. Sejam $w_1, w_2 \in S^{\perp}$ e $v \in S$. Então, tem–se que

$$\langle w_1 , v \rangle = 0 \quad \text{e} \quad \langle w_2 , v \rangle = 0 \quad \Longrightarrow \quad \langle w_1 + w_2 , v \rangle = 0 .$$

Logo, $w_1 + w_2 \in S^{\perp}$. De modo análogo, temos que $\lambda w_1 \in S^{\perp}$ para todo $\lambda \in I\!\!F$.

Considerando agora S um subespaço de V, vamos mostrar que $S \cap S^{\perp} = \{ 0_V \}$. Tomando $w \in S^{\perp} \cap S$, isto é, $w \in S^{\perp}$ e $w \in S$. Como $w \in S^{\perp}$, temos que $\langle w , v \rangle = 0$ para todo $v \in S$. Em particular para $v = w$, pois $w \in S$, obtemos $\langle w , w \rangle = 0$. Logo, $w = 0_V$, o que completa a demonstração. ■

Teorema 5.8.2 *Sejam V um espaço vetorial de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$, U e W subespaços vetoriais de V. Então, $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$.*

Demonstração – Inicialmente, tomamos $v \in (U + W)^{\perp}$, isto é, v é ortogonal a todo elemento $u + w$ pertencente ao subespaço $U + W$. Como $U \subset U + W$ e $W \subset U + W$, temos que v é ortogonal a todo elemento de U e a todo elemento de W, isto é, $v \in U^{\perp}$ e $v \in W^{\perp}$. Logo, $v \in U^{\perp} \cap W^{\perp}$. Assim, mostramos que $(U + W)^{\perp} \subset U^{\perp} \cap W^{\perp}$.

Finalmente, seja $v \in U^{\perp} \cap W^{\perp}$, isto é, v é ortogonal a todo elemento de U e a todo elemento de W. Desse modo, dado um elemento $u + w \in U + W$, temos que

$$\langle v , u + w \rangle = \langle v , u \rangle + \langle v , w \rangle = 0 .$$

Logo, $v \in (U + W)^{\perp}$. Assim, mostramos que $U^{\perp} \cap W^{\perp} \subset (U + W)^{\perp}$. Portanto, provamos que $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$. ■

Proposição 5.8.1 *Sejam* V *um espaço vetorial munido do produto interno* $\langle \cdot , \cdot \rangle$, W *um subespaço de* V *e* $\beta = \{ w_1 , \cdots , w_n \}$ *uma base para* W. *Então,* $v \in W^{\perp}$ *se, e somente se,* $\langle w_i , v \rangle = 0$ *para todo* $i = 1, \cdots , n$.

Demonstração

($\Longrightarrow$) Se $v \in W^{\perp}$, isto é, $\langle w , v \rangle = 0$ para todo $w \in W$. Em particular, temos que $\langle w_i , v \rangle = 0$ para todo $i = 1, \cdots , n$.

($\Longleftarrow$) Seja $w \in W$, isto é, w é escrito de modo único como:

$$w = \alpha_1 w_1 + \cdots + \alpha_n w_n .$$

Desse modo, para todo $v \in V$, temos que

$$\langle w , v \rangle = \alpha_1 \langle w_1 , v \rangle + \cdots + \alpha_n \langle w_n , v \rangle .$$

Considerando $\langle w_i , v \rangle = 0$, para $1 \leq i \leq n$, temos que $\langle w , v \rangle = 0$. Logo, $v \in W^{\perp}$, o que completa a demonstração. ■

Exemplo 5.8.1 *Considere o espaço vetorial real* $\mathbb{R}^4$ *munido do produto interno usual. Seja* W *o subespaço de* $\mathbb{R}^4$ *dado por:* $W = [(1,0,1,1), (1,1,0,1)]$. *Determine uma base para o subespaço* $W^{\perp}$.

Temos que todo elemento $(x, y, z, t) \in W^{\perp}$ deve ser ortogonal aos elementos geradores de W. Assim, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} x + z + t = 0 \\ x + y + t = 0 \end{cases} \iff \begin{cases} x + z + t = 0 \\ y - z = 0 \end{cases}$$

Obtemos $x = -z - t$ e $y = z$ para $z, t \in \mathbb{R}$. Desse modo, todo elemento $(x, y, z, t) \in W^{\perp}$ é escrito da seguinte forma:

$$(x, y, z, t) = \alpha_1(-1, 1, 1, 0) + \alpha_2(-1, 0, 0, 1) \quad \text{para} \quad \alpha_1, \alpha_2 \in \mathbb{R} .$$

Portanto, temos que $W^{\perp} = [(-1, 1, 1, 0), (-1, 0, 0, 1)]$.

Exemplo 5.8.2 *Considere o espaço vetorial* $I\!R^2$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Determine* $U^\perp$ *do seguinte subespaço*

$$U = \{ (x,y) \in I\!R^2 \ / \ y - 2x = 0 \} .$$

Note que U é uma reta no plano passando pelo origem e que tem por vetor diretor $u = (1,2)$. Assim, todo elemento $v = (x,y) \in U^\perp$ satisfaz

$$\langle v , u \rangle = 0 \quad \Longrightarrow \quad x + 2y = 0 .$$

Portanto, todo elemento $(x,y) \in U^\perp$ satisfaz a equação da reta $y = -\dfrac{x}{2}$.

Exemplo 5.8.3 *Considere o espaço vetorial* $I\!R^3$ *com o produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* W *o subespaço gerado pelo elemento* $w = (1,-1,2) \in I\!R^3$. *Determine* $W^\perp$.

Tomando $v = (x,y,z) \in W^\perp$ e como W é gerado pelo elemento w, temos que

$$\langle v , w \rangle = 0 \quad \Longrightarrow \quad x - y + 2z = 0 .$$

Portanto, todo elemento $(x,y,z) \in W^\perp$ satisfaz a equação do plano $x - y + 2z = 0$.

Exemplo 5.8.4 *Considere o espaço vetorial* $\mathcal{C}([-\pi,\pi])$ *com o produto interno usual*

$$\langle f , g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx \qquad ; \qquad \forall \ f,g \in \mathcal{C}([-\pi,\pi]) .$$

Considere o subespaço vetorial das funções pares

$$S = \{ f \in \mathcal{C}([-\pi,\pi]) \ / \ f(-x) = f(x) \ , \ x \in [-\pi,\pi] \} .$$

Mostre que o complemento ortogonal de S *em* $\mathcal{C}([-\pi,\pi])$ *é o subespaço vetorial das funções ímpares, isto é,*

$$S^\perp = \{ g \in \mathcal{C}([-\pi,\pi]) \ / \ g(-x) = -g(x) \ , \ x \in [-\pi,\pi] \} .$$

Exemplo 5.8.5 *Considere o espaço vetorial real* $I\!M_n(I\!R)$ *com o produto interno usual*

$$\langle A , B \rangle = tr(B^t A) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij} \qquad ; \qquad \forall \, A, \ B \in I\!M_n(I\!R) .$$

Considere o subespaço vetorial das matrizes diagonais

$$S = \{ D \in I\!M_n(I\!R) \ / \ D \quad \text{é uma matriz diagonal} \} .$$

Determine o complemento ortogonal de S *em* $I\!M_n(I\!R)$.

Exemplo 5.8.6 *Considere o espaço vetorial real*

$$U = \{ p(x) \in \mathcal{P}_3(I\!R) \ / \ p(-1) = p(1) = 0 \}$$

com o produto interno

$$\langle p , q \rangle = \int_{-1}^{1} p'(x)q'(x)dx \quad ; \quad \forall \ p, q \in U .$$

Determine uma base para o complemento ortogonal do subespaço $S = [1 - x^2]$ *em* U *com relação ao produto interno* $\langle \cdot , \cdot \rangle$ *definido acima.*

Como $dim(U) = 2$ e $dim(S) = 1$, podemos concluir que $dim(S^{\perp}) = 1$.

Chamando $p(x) = 1 - x^2$, temos que o subespaço $S = [p(x)]$. O subespaço $S^{\perp}$ é definido por:

$$S^{\perp} = \{ q(x) \in \mathcal{P}_3(I\!R) \ / \ \langle r , q \rangle = 0 \quad ; \quad \forall \ r(x) \in S \} .$$

Tomando um elemento genérico $q(x) = a + bx + cx^2 + dx^3 \in S^{\perp}$, sabemos que $\langle p , q \rangle = 0$. Além disso, $q(-1) = q(1) = 0$. Assim, temos que

$$\begin{aligned} \langle p , q \rangle &= \int_{-1}^{1} (-2x)(b + 2cx + 3dx^2)dx = \int_{-1}^{1} (-2bx - 4cx^2 - 6dx^3)dx \\ &= -4c\int_{-1}^{1} x^2dx = -\frac{8}{3}c = 0 \end{aligned}$$

Assim, obtemos $c = 0$. Logo, $q(x) = a + bx + dx^3$, que impondo as condições

$$q(-1) = a - b - d = 0 \quad \text{e} \quad q(1) = a + b + d = 0 ,$$

temos um sistema linear homogêneo cuja solução é $a = 0$ e $d = -b$ para $b \in I\!R$.

Desse modo, todo elemento $q(x) \in S^{\perp}$ é escrito da seguinte forma:

$$q(x) = b(x - x^3) \quad ; \quad b \in I\!R .$$

Portanto, uma base para o subespaço $S^{\perp}$ é dada pelo conjunto

$$\gamma = \{ x - x^3 \} .$$

Exemplo 5.8.7 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *munido do produto interno*

$$\langle\, p\,,\, q\,\rangle \;=\; \int_{-1}^{1} x^2 p(x) q(x) dx \qquad ; \qquad \forall\ p,\, q \,\in\, \mathcal{P}_2(\mathbb{R})\,.$$

Determine uma base para o complemento ortogonal do subespaço $S = [1 + x]$ *em* $\mathcal{P}_2(\mathbb{R})$ *com relação ao produto interno* $\langle\,\cdot\,,\,\cdot\,\rangle$ *definido acima.*

Chamando $p(x) = 1 + x$, temos que o subespaço $S = [p(x)] \subset \mathcal{P}_2(\mathbb{R})$.

O subespaço $S^{\perp}$ é definido por:

$$S^{\perp} \;=\; \{\, q \,\in\, \mathcal{P}_2(\mathbb{R}) \;/\; \langle\, r\,,\, q\,\rangle = 0 \quad ; \quad \forall\, r \,\in\, S \,\}\,.$$

Tomando um elemento genérico $q(x) = a + bx + cx^2 \in S^{\perp}$, sabemos que $\langle\, p\,,\, q\,\rangle = 0$. Assim, temos que

$$\begin{aligned}
\langle\, p\,,\, q\,\rangle &= \int_{-1}^{1} x^2(1 + x)(a + bx + cx^2)dx \\
&= \int_{-1}^{1} (x^2 + x^3)(a + bx + cx^2)dx \\
&= \int_{-1}^{1} (\, ax^2 + bx^3 + cx^4 + ax^3 + bx^4 + cx^5\,)dx \;=\; 0 \\
&= \int_{-1}^{1} (\, ax^2 + cx^4 + bx^4\,)dx \;=\; 0
\end{aligned}$$

Calculando a integral, resulta a seguinte equação

$$\frac{2}{3}a + \frac{2}{5}c + \frac{2}{5}b \;=\; 0$$

Resolvendo a equação acima para a incógnita c, temos que

$$c \;=\; -\frac{5}{3}a \;-\; b\,.$$

Portanto, todo elemento $q(x) \in S^{\perp}$ é escrito como:

$$\begin{aligned}
q(x) &= a + bx + \left(-\frac{5}{3}a - b\right)x^2 \\
&= \left(1 - \frac{5}{3}x^2\right)a \;+\; (\,x - x^2\,)\,b \qquad \text{para} \qquad a,\, b \,\in\, \mathbb{R}\,.
\end{aligned}$$

Desse modo, uma base para o subespaço $S^{\perp}$ é formada pelos elementos

$$q_1(x) \;=\; 1 - \frac{5}{3}x^2 \qquad \text{e} \qquad q_2(x) \;=\; x - x^2\,.$$

5.9 Decomposição Ortogonal

Teorema 5.9.1 *Sejam V um espaço vetorial munido do produto interno $\langle \cdot , \cdot \rangle$ e S um subespaço de dimensão finita de V. Então, $V = S \oplus S^{\perp}$, isto é, todo elemento $u \in V$ pode ser escrito de modo único da seguinte forma:*

$$u = v + w \qquad \textit{com} \qquad v \in S \quad \textit{e} \quad w \in S^{\perp} .$$

Além disso, a norma do elemento u é dada pela ***Fórmula de Pitágoras***

$$\| u \|_2^2 = \| v \|_2^2 + \| w \|_2^2 .$$

Demonstração – Seja $\beta = \{ q_1, \cdots , q_n \}$ uma base ortonormal para o subespaço S. Desse modo, todo elemento $v \in S$ pode ser escrito de modo único da seguinte forma:

$$v = \sum_{j=1}^{n} c_j \, q_j ,$$

onde c_j são os coeficiente de Fourier de v em relação à base ortonormal β.

Dado um elemento $u \in V$ vamos escreve–lo da seguinte forma:

$$u = \sum_{j=1}^{n} \langle u , q_j \rangle q_j + w \qquad \Longrightarrow \qquad w = u - \sum_{j=1}^{n} \langle u , q_j \rangle q_j$$

para em seguida mostrar que o elemento $w \in S^{\perp}$. De fato, fazendo o produto interno entre o elemento w e os elementos da base β, obtemos

$$\langle w , q_i \rangle = \langle u , q_i \rangle - \sum_{j=1}^{n} \langle u , q_j \rangle \langle q_j , q_i \rangle = 0 \qquad , \qquad i = 1, \cdots , n$$

Como $\langle w , q_i \rangle = 0$ para $i = 1, \cdots , n$, temos que $w \in S^{\perp}$. Portanto, obtemos que $u = v + w$ com $v \in S$ e $w \in S^{\perp}$.

Finalmente, vamos mostrar a unicidade dos elementos $v \in S$ e $w \in S^{\perp}$. Para isso, vamos supor que

$$u = v_1 + w_1 \qquad \text{com} \qquad v_1 \in S \quad \text{e} \quad w_1 \in S^{\perp}$$

$$u = v_2 + w_2 \qquad \text{com} \qquad v_2 \in S \quad \text{e} \quad w_2 \in S^{\perp} .$$

Desse modo, temos que

$$0_V = (v_1 - v_2) + (w_1 - w_2) \qquad \Longrightarrow \qquad (v_1 - v_2) = (w_2 - w_1) .$$

Assim, podemos concluir que $(v_1 - v_2) \in S \cap S^{\perp}$ e que $(w_2 - w_1) \in S \cap S^{\perp}$. Como $S \cap S^{\perp} = \{ 0_V \}$, temos que $v_1 - v_2 = 0_V$ e $w_2 - w_1 = 0_V$. Portanto, segue a unicidade dos elementos v e w.

Como $u \in V$ é escrito como $u = v + w$ com $v \in S$ e $w \in S^{\perp}$, obtemos

$$\begin{aligned} \| u \|_2^2 &= \langle v + w , v + w \rangle = \langle v , v \rangle + \langle v , w \rangle + \langle w , v \rangle + \langle w , w \rangle \\ &= \| v \|_2^2 + \| w \|_2^2 , \end{aligned}$$

que é a **Fórmula de Pitágoras**, completando a demonstração. ■

Corolário 5.9.1 *Sejam V um espaço vetorial de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$ e S um subespaço de V. Então, $dim(V) = dim(S) + dim(S^{\perp})$.*

Demonstração – Combinando os resultados do Teorema 3.6.5, sobre a dimensão da soma de dois subespaços, e do Teorema 5.9.1, obtemos o resultado desejado. □

Teorema 5.9.2 *Sejam V um espaço vetorial de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$ e U um subespaço vetorial de V. Então $U = (U^{\perp})^{\perp}$.*

Demonstração – Tomando um elemento $u \in U$, temos que $\langle u , v \rangle = 0$ para todo $v \in U^{\perp}$. Logo, $u \in (U^{\perp})^{\perp}$. Assim, mostramos que $U \subset (U^{\perp})^{\perp}$.

Finalmente, tomando $w \in (U^{\perp})^{\perp}$, isto é, $\langle w , v \rangle = 0$ para todo $v \in U^{\perp}$.

Pelo Teorema 5.9.1, sabemos que $V = U \oplus U^{\perp}$. Desse modo, podemos escrever o elemento $w \in (U^{\perp})^{\perp}$ da seguinte forma:

$$w = u + v$$

com $u \in U$ e $v \in U^{\perp}$. Assim, temos que

$$\langle w , v \rangle = \langle u , v \rangle + \langle v , v \rangle = 0 .$$

Como $\langle u , v \rangle = 0$, obtemos $\langle v , v \rangle = 0$. Logo, temos que $v = 0_V$.

Desse modo, concluímos que $w \in U$. Assim, mostramos que $(U^{\perp})^{\perp} \subset U$. Portanto, provamos que $U = (U^{\perp})^{\perp}$, completando a demonstração. ■

Teorema 5.9.3 *Sejam* V *um espaço vetorial de dimensão finita munido do produto interno* $\langle \cdot , \cdot \rangle$, U *e* W *subespaços vetoriais de* V. *Então,* $(U \cap W)^{\perp} = U^{\perp} + W^{\perp}$.

Demonstração – Inicialmente, tomamos $v \in U^{\perp} + W^{\perp}$. Assim, podemos escrever o elemento v da seguinte forma $v = \overline{u} + \overline{w}$, com $\overline{u} \in U^{\perp}$ e $\overline{w} \in W^{\perp}$.

Desse modo, dado um elemento $\widehat{v} \in U \cap W$, temos que

$$\langle v , \widehat{v} \rangle = \langle \overline{u} + \overline{w} , \widehat{v} \rangle = \langle \overline{u} , \widehat{v} \rangle + \langle \overline{w} , \widehat{v} \rangle = 0 \,,$$

pois $\widehat{v} \in U$ e $\widehat{v} \in W$. Logo, tem–se $v \in (U \cap W)^{\perp}$.

Assim, mostramos que $U^{\perp} + W^{\perp} \subset (U \cap W)^{\perp}$.

Finalmente, tomamos $v \in (U \cap W)^{\perp}$, isto é, $\langle v , w \rangle = 0$ para todo $w \in U \cap W$.

Pelo Teorema 5.9.1, sabemos que

$$V = (U^{\perp} + W^{\perp}) \oplus (U^{\perp} + W^{\perp})^{\perp}$$

Utilizando o resultado do Teorema 5.8.2 e o resultado do Teorema 5.9.2, obtemos

$$(U^{\perp} + W^{\perp})^{\perp} = (U \cap W) \,.$$

Assim, temos que

$$V = (U^{\perp} + W^{\perp}) \oplus (U \cap W)$$

Desse modo, podemos escrever o elemento $v \in (U \cap W)^{\perp}$ da seguinte forma:

$$v = v_1 + w \qquad \text{com} \qquad v_1 \in (U^{\perp} + W^{\perp}) \qquad \text{e} \qquad w \in (U \cap W) \,.$$

Assim sendo, temos que

$$\langle v , w \rangle = \langle v_1 , w \rangle + \langle w , w \rangle = 0 \,.$$

Como $\langle v_1 , w \rangle = 0$, obtemos $\langle w , w \rangle = 0$. Logo, temos que $w = 0_V$.

Desse modo, concluímos que $v \in (U^{\perp} + W^{\perp})$.

Assim, temos que $(U \cap W)^{\perp} \subset U^{\perp} + W^{\perp}$.

Portanto, provamos que $U^{\perp} + W^{\perp} = (U \cap W)^{\perp}$, o que completa a demonstração. ■

Exemplo 5.9.1 *Considere o espaço vetorial* $\mathbb{R}^2$ *com o produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* S *o subespaço de* $\mathbb{R}^2$ *definido por:* $S = \{ (x,y) \in \mathbb{R}^2 \ / \ x - y = 0 \}$. *Dado o elemento* $u = (-1,3) \in \mathbb{R}^2$, *determine sua decomposição* $u = v + w$ *com* $v \in S$ *e* $w \in S^{\perp}$.

Sabemos que $\left\{ q = \dfrac{\sqrt{2}}{2}(1,1) \right\}$ é uma base para o subespaço S. Assim, temos que o elemento $v \in S$ e o elemento $w \in S^{\perp}$ são dados por:

$$v = \langle u , q \rangle q = (1,1) \qquad \text{e} \qquad w = u - v = (-2,2)\,.$$

de acordo com o Teorema 5.9.1.

Exemplo 5.9.2 *Considere o espaço vetorial real* $\mathbb{R}^4$ *munido do produto interno usual e* W *o subespaço definido por:*

$$W = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x - 2y + z - 3t = 0 \}\,,$$

isto é, W *é um hiperplano em* $\mathbb{R}^4$. *Temos que,* $dim(W) = 3$.

Podemos verificar facilmente que $W^{\perp} = [(1,-2,1,-3)]$. Assim, utilizando o resultado do Corolário 5.9.1 com $V = \mathbb{R}^4$, obtemos

$$dim(W) = 4 - dim(W^{\perp}) = 3\,.$$

De um modo geral, um **hiperplano** H contido no espaço vetorial $\mathbb{R}^n$ é definido como:

$$H = \{ (x_1, \cdots, x_n) \in \mathbb{R}^n \ / \ c_1x_1 + \cdots + c_nx_n = 0 \}\,,$$

conhecendo os escalares $c_1, \cdots, c_n$.

Note que considerando o elemento $v = (c_1, \cdots, c_n) \in \mathbb{R}^n$, fixo, temos que

$$H = \{ u \in \mathbb{R}^n \ / \ \langle u , v \rangle = 0 \}\,,$$

Desse modo, o subespaço $H^{\perp} = [v]$. Logo, pelo Corolário 5.9.1, obtemos que $dim(H)$ é igual a $(n - 1)$.

Exemplo 5.9.3 *Considere o espaço vetorial real* $\mathbb{R}^4$ *munido do produto interno usual e* W *o subespaço definido por:*

$$W = \{ (x, y, z, t) \in \mathbb{R}^4 \ / \ x - 2y + z + t = 0 \} .$$

Determine uma base para o subespaço W *e uma base para o subespaço* $W^\perp$*.*

Temos que todo elemento $w \in W$ é escrito como

$$w = \alpha_1(2, 1, 0, 0) + \alpha_2(-1, 0, 1, 0) + \alpha_3(-1, 0, 0, 1) \quad \text{para} \quad \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R} .$$

Portanto, os elementos $v_1 = (2, 1, 0, 0)$, $v_2 = (-1, 0, 1, 0)$ e $v_3 = (-1, 0, 0, 1)$ formam uma base para o subespaço W.

Temos que todo elemento $u = (x, y, z, t) \in W^\perp$ é ortogonal aos elementos de W. Assim, $u \in W^\perp$ deve ser ortogonal aos elementos da base de W, isto é,

$$\begin{array}{rcl} \langle u \, , \, v_1 \rangle & = & 0 \\ \langle u \, , \, v_2 \rangle & = & 0 \\ \langle u \, , \, v_3 \rangle & = & 0 \end{array} \quad \Longrightarrow \quad \left\{ \begin{array}{rcrcrcrcl} 2x & + & y & & & & & = & 0 \\ -x & & & + & z & & & = & 0 \\ -x & & & & & + & t & = & 0 \end{array} \right.$$

A solução do sistema linear homogêneo é dada por:

$$x = t \quad , \quad y = -2t \quad \text{e} \quad z = t \quad \text{para} \quad t \in \mathbb{R} .$$

Desse modo, todo elemento $u = (x, y, z, t) \in W^\perp$ é escrito da forma:

$$u = (x, y, z, t) = t \, (1, -2, 1, 1) \quad \text{para} \quad t \in \mathbb{R} .$$

Portanto, o subespaço $W^\perp = [(1, -2, 1, 1)]$.

Exemplo 5.9.4 *Considere o espaço vetorial real* $\mathcal{P}_3(\mathbb{R})$ *munido do produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_0^1 p(x)q(x)dx \,.$$

Determine uma base para o complemento ortogonal do subespaço $W = [1, x]$.

Por simplicidade, vamos chamar de $p_1(x) = 1$ e $p_2(x) = x$ os elementos da base do subespaço W. Assim, todo elemento $q(x) = a + bx + cx^2 + dx^3 \in W^{\perp}$ deve ser ortogonal aos elementos da base de W, isto é,

$$\begin{array}{rcl} \langle\, q \,,\, p_1 \,\rangle & = & 0 \\ \langle\, q \,,\, p_2 \,\rangle & = & 0 \end{array} \qquad \Longrightarrow \qquad \left\{ \begin{array}{rcrcrcrcl} a & + & \frac{1}{2}b & + & \frac{1}{3}c & + & \frac{1}{4}d & = & 0 \\[2ex] \frac{1}{2}a & + & \frac{1}{3}b & + & \frac{1}{4}c & + & \frac{1}{5}d & = & 0 \end{array} \right.$$

A solução do sistema linear homogêneo é dada por:

$$\begin{array}{rcl} a & = & \frac{1}{6}c + \frac{1}{5}d \\[2ex] b & = & -c - \frac{9}{10}d \end{array}$$

para $c, d \in \mathbb{R}$.

Portanto, todo elemento $q(x) = a + bx + cx^2 + dx^3 \in W^{\perp}$ é escrito da seguinte forma:

$$q(x) \;=\; \left(\frac{1}{6} \,-\, x \,+\, x^2 \right) c \;+\; \left(\frac{1}{5} \,-\, \frac{9}{10}x \,+\, x^3 \right) d$$

para $c, d \in \mathbb{R}$. Portanto, os elementos

$$q_1(x) \;=\; \frac{1}{6} \,-\, x \,+\, x^2 \qquad \text{e} \qquad q_2(x) \;=\; \frac{1}{5} \,-\, \frac{9}{10}x \,+\, x^3$$

formam uma base para o subespaço $W^{\perp}$.

Exercícios

Exercício 5.40 *Considere o espaço vetorial real* $\mathcal{C}([1,e])$ *munido do produto interno*

$$\langle\, f \,,\, g \,\rangle \;=\; \int_1^e \ln(t) f(t) g(t) dt\,.$$

Determine as funções $g(x) \;=\; a \;+\; bx$, $a, b \in I\!R$, *que são ortogonais à função* $f(x) \;=\; 1$. *Dê uma interpretação geométrica.*

Resposta: $\boxed{g(x) \;=\; b\left(x \;-\; \dfrac{e^2 \;+\; 1}{4}\right),\; b \;\in\; I\!R}$

Exercício 5.41 *Considere o espaço vetorial* $I\!R^3$ *com o produto interno usual. Seja* $T : I\!R^3 \longrightarrow I\!R^2$ *a transformação linear definida por:*

$$T(x,y,z) \;=\; (x - y - z,\, 2z - x)\,.$$

Determine uma base ortogonal para o complemento ortogonal do subespaço $Ker(T)$.

Exercício 5.42 *Considere o espaço vetorial real*

$$U \;=\; \{\, p(x) \;\in\; \mathcal{P}_3(I\!R) \;\;/\;\; p(1) \;=\; 0 \,\}$$

com o produto interno

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p'(x)q'(x)dx \qquad ; \qquad \forall\; p, q \;\in\; U\,.$$

Determine uma base para o complemento ortogonal do subespaço $S = [1 \;-\; x]$ *em* U *com relação ao produto interno* $\langle\, \cdot \,,\, \cdot \,\rangle$ *definido acima.*

Exercício 5.43 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual. Dados os elementos* $u \;=\; (1,1,1)$ *e* $v \;=\; (2,-1,1)$. *Determine os elementos* w_1 *e* w_2 *tais que* $v \;=\; w_1 + w_2$, *de modo que* w_1 *seja ortogonal ao elemento* u *e que o conjunto* $\{\, w_2 \,,\, u \,\}$ *seja linearmente dependente. Dê uma interpretação geométrica.*

Exercício 5.44 *Considere o espaço vetorial real* $\mathcal{P}_3(I\!R)$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx \qquad ; \qquad \forall\; p, q \;\in\; \mathcal{P}_3(I\!R)\,,$$

e o subespaço vetorial $S \;=\; \{\; p(x) \;\in\; \mathcal{P}_3(I\!R) \;\;/\;\; p(-1) \;=\; p(1) \;=\; 0 \;\}$. *Dado o polinômio* $q(x) \;=\; 1 \;+\; 2x \;-\; x^2$, *determine sua decomposição* $q(x) \;=\; p(x) \;+\; r(x)$ *com* $p(x) \;\in\; S$ *e* $r(x) \;\in\; S^{\perp}$.

Exercício 5.45 *Considere o espaço vetorial real* $\mathcal{P}_2(I\!R)$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx \qquad ; \qquad \forall \;\, p,\, q \,\in\, \mathcal{P}_2(I\!R)\,.$$

Determine uma base para o complemento ortogonal do subespaço $S = [1 \,+\, x,\, 1 \,-\, x^2]$ *em* $\mathcal{P}_2(I\!R)$ *com relação ao produto interno* $\langle\, \cdot\,,\, \cdot\,\rangle$ *definido acima.*

Exercício 5.46 *Considere o espaço vetorial real* $\mathcal{P}_2(I\!R)$ *com o produto interno*

$$\langle\, p \,,\, q \,\rangle \;=\; p(-1)q(-1) \;+\; p(0)q(0) \;+\; p(1)q(1)$$

Determine uma base para o complemento ortogonal do subespaço $U = [2 - x]$ *em* $\mathcal{P}_2(I\!R)$ *com relação ao produto interno* $\langle\, \cdot\,,\, \cdot\,\rangle$ *definido acima.*

Exercício 5.47 *Seja* U *o subespaço gerado pelo elemento* $u \,=\, (0,\, 1,\, -2,\, 1) \,\in\, I\!R^4$*. Encontre uma base para o complemento ortogonal do subespaço* U *em* $I\!R^4$ *com relação ao produto interno usual.*

Exercício 5.48 *Seja* W *o subespaço de* $I\!R^5$ *gerado pelos vetores* $w_1 \,=\, (1,\, 2,\, 3,\, -1,\, 2)$ *e* $w_2 \,=\, (2,\, 1,\, 0,\, 2,\, -1)$ *. Encontre uma base para o complemento ortogonal do subespaço* W *em* $I\!R^5$ *com relação ao produto interno usual.*

Exercício 5.49 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual. Seja* S *o subespaço de* $I\!R^3$ *definido por:*

$$S \;=\; \{\; (x, y, z) \;\in\; I\!R^3 \;\;/\;\; 2x \,-\, y \,+\, z \,=\, 0 \;\}\;.$$

Determine uma base para o complemento ortogonal do subespaço S *em* $I\!R^3$ *com relação ao produto interno usual.*

Exercício 5.50 *Considere o espaço vetorial real* $I\!R^4$ *munido do produto interno usual. Seja* U *o subespaço gerado pelo elemento* $u \,=\, (-1,\, 1,\, 1,\, -1)$*. Determine uma base ortogonal para o subespaço* $U^{\perp}$*.*

Exercício 5.51 *Considere o espaço vetorial* $I\!R^3$ *com o produto interno usual. Seja* $T : I\!R^3 \longrightarrow I\!R^3$ *o operador linear definido por:*

$$T(x, y, z) \;=\; (x - y - z,\, -x + y + 2z,\, x - y)\,.$$

Determine uma base ortogonal para o subespaço $Im(T)$*.*

5.10 Identidade de Parseval

Teorema 5.10.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$ e $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal de V. Então, para todos $u, v \in V$ temos que*

$$\langle u , v \rangle = \sum_{j=1}^{n} \langle u , q_j \rangle \overline{\langle v , q_j \rangle} .$$

Em particular para $u = v$, tem–se que

$$\| u \|_2^2 = \langle u , u \rangle = \sum_{j=1}^{n} | \langle u , q_j \rangle |^2 .$$

*que é uma generalização do **Teorema de Pitágoras**.*

Demonstração – Pelo Teorema 5.6.1, para todos $u, v \in V$, temos que

$$u = \sum_{j=1}^{n} \langle u , q_j \rangle q_j \qquad \text{e} \qquad v = \sum_{i=1}^{n} \langle v , q_i \rangle q_i$$

Calculando o produto interno entre os elementos u e v

$$\langle u , v \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle u , q_j \rangle \overline{\langle v , q_i \rangle} \langle q_j , q_i \rangle = \sum_{j=1}^{n} \langle u , q_j \rangle \overline{\langle v , q_j \rangle}$$

obtemos o resultado desejado.

No caso em que $v = u$, tem–se que

$$\langle u , u \rangle = \sum_{j=1}^{n} \langle u , q_j \rangle \overline{\langle u , q_j \rangle} = \sum_{j=1}^{n} | \langle u , q_j \rangle |^2 ,$$

o que completa a demonstração. ■

Podemos observar que se V é um espaço vetorial real, a identidade de Parseval fica escrita como:

$$\langle u , v \rangle = \sum_{j=1}^{n} \langle u , q_j \rangle \langle v , q_j \rangle \qquad \text{para todo} \qquad u, v \in V .$$

Exemplo 5.10.1 *Considere o espaço vetorial* $I\!R^2$ *munido do produto interno usual e* $\beta = \{ (1,1), (-1,1) \}$ *uma base ortogonal. Dados os elementos* $u = (2,4)$ *e* $v = (1,2)$, *verifique a identidade de Parseval.*

Por simplicidade, vamos denotar $q_1 = (1,1)$ e $q_2 = (-1,1)$. Inicialmente, vamos obter uma base ortonormal $\beta' = \{ q_1', q_2' \}$ a partir da base ortogonal $\beta = \{ q_1, q_2 \}$. Assim, temos que

$$q_1' = \frac{1}{\sqrt{2}}(1,1) \qquad \text{e} \qquad q_2' = \frac{1}{\sqrt{2}}(-1,1)\,.$$

Finalmente, aplicando a identidade de Parseval

$$10 = \langle u \,,\, v \rangle = \langle u \,,\, q_1' \rangle \langle v \,,\, q_1' \rangle + \langle u \,,\, q_2' \rangle \langle v \,,\, q_2' \rangle = 9 + 1 = 10$$

Podemos também verificar que

$$20 = \| u \|_2^2 = \langle u \,,\, u \rangle = \langle u \,,\, q_1' \rangle^2 + \langle u \,,\, q_2' \rangle^2 = 18 + 2 = 20$$

o que completa a verificação da identidade de Parseval.

Utilizando a notação $\langle \cdot \,,\, \cdot \rangle_V$ para o produto interno do espaço vetorial V sobre o corpo $I\!F$ e a notação $\langle \cdot \,,\, \cdot \rangle_{I\!F^n}$ para o produto interno do espaço vetorial $I\!F^n$ sobre o corpo $I\!F$, observamos facilmente que a **Identidade de Parseval** pode ser escrita da seguinte forma:

$$\langle u \,,\, v \rangle_V = \sum_{j=1}^{n} \langle u \,,\, q_j \rangle \overline{\langle v \,,\, q_j \rangle} = \langle [u]_\beta \,,\, [v]_\beta \rangle_{I\!F^n} \qquad \text{para todos} \qquad u, v \in V\,,$$

onde $[u]_\beta$ e $[v]_\beta$ são os vetores de coordenadas dos elementos u e v em relação à base ortonormal $\beta = \{ q_1 \,, \cdots \,, q_n \}$ de V, respectivamente.

5.11 Desigualdade de Bessel

Teorema 5.11.1 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $\beta = \{ q_1 , \cdots , q_n \}$ *um conjunto ortonormal em* V. *Então, para todo* $u \in V$ *temos que*

$$\| u \|_2^2 \ \geq \ \sum_{i=1}^{n} | \langle u , q_i \rangle |^2 .$$

Além disso, vale a igualdade se, e somente se, $u \in S = [q_1, \cdots , q_n]$, *isto é, o elemento* $u \in S$ *é escrito de modo único como:*

$$u \ = \ \sum_{i=1}^{n} \langle u , q_i \rangle \, q_i .$$

Demonstração – Vamos considerar o subespaço S gerado pelos elementos do conjunto ortonormal β, isto é, $S = [q_1, \cdots , q_n]$. Pelo Teorema 5.9.1, da Decomposição Ortogonal, sabemos que todo elemento $u \in V$ é escrito de modo único como:

$$u \ = \ v \ + \ w \qquad \text{para} \qquad v \ \in \ S \quad \text{e} \quad w \ \in \ S^{\perp} ,$$

onde o elemento $v \in S$ é dado por:

$$v \ = \ \sum_{i=1}^{n} \langle u , q_i \rangle \, q_i .$$

Além disso, temos a Fórmula de Pitágoras

$$\| u \|_2^2 \ = \ \| v \|_2^2 \ + \ \| w \|_2^2 .$$

Portanto, temos que

$$\| u \|_2^2 \ \geq \ \| v \|_2^2 \ = \ \langle v , v \rangle \ = \ \sum_{i=1}^{n} | \langle u , q_i \rangle |^2 .$$

Da Identidade de Parseval, podemos concluir que o elemento $u \in S$ se, e somente se,

$$\| u \|_2^2 \ = \ \sum_{i=1}^{n} | \langle u , q_i \rangle |^2 ,$$

o que completa a demonstração. ■

Exemplo 5.11.1 *Considere o espaço vetorial real* $I\!R^4$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *e o conjunto ortogonal* $\beta = \{ (1,1,-1,1), (-1,1,1,1) \}$. *Dado o elemento* $u = (2,4,1,-1) \in I\!R^4$, *verifique a desigualdade de Bessel.*

Por simplicidade, denotamos $q_1 = (1,1,-1,1)$ e $q_2 = (-1,1,1,1)$. Inicialmente, vamos obter um conjunto ortonormal $\beta' = \{ q_1', q_2' \}$ a partir do conjunto ortogonal $\beta = \{ q_1, q_2 \}$. Assim, temos que

$$q_1' = \frac{1}{2}(1,1,-1,1) \qquad \text{e} \qquad q_2' = \frac{1}{2}(-1,1,1,1)\,.$$

Os coeficientes de Fourier do elemento $u \in I\!R^4$ com relação ao conjunto ortonormal $\beta' = \{ q_1', q_2' \}$ são dados por:

$$\alpha_1 = \langle u , q_1' \rangle = 2 \qquad \text{e} \qquad \alpha_2 = \langle u , q_2' \rangle = 1\,.$$

Finalmente, obtemos que

$$22 = \| u \|_2^2 > \langle u , q_1' \rangle^2 + \langle u , q_2' \rangle^2 = 5\,.$$

Exemplo 5.11.2 *Considere o espaço vetorial real* $I\!R^4$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *e o conjunto ortogonal* $\beta = \{ (1,1,-1,1), (-1,1,1,1) \}$. *Dado o elemento* $u = (5,1,-5,1) \in I\!R^4$, *verifique a desigualdade de Bessel. O que podemos concluir ?*

Por simplicidade, denotamos $q_1 = (1,1,-1,1)$ e $q_2 = (-1,1,1,1)$. Inicialmente, vamos obter um conjunto ortonormal $\beta' = \{ q_1', q_2' \}$ a partir do conjunto ortogonal $\beta = \{ q_1, q_2 \}$. Assim, temos que

$$q_1' = \frac{1}{2}(1,1,-1,1) \qquad \text{e} \qquad q_2' = \frac{1}{2}(-1,1,1,1)\,.$$

Os coeficientes de Fourier do elemento $u \in I\!R^4$ com relação ao conjunto ortonormal $\beta' = \{ q_1', q_2' \}$ são dados por:

$$\alpha_1 = \langle u , q_1' \rangle = 6 \qquad \text{e} \qquad \alpha_2 = \langle u , q_2' \rangle = -4\,.$$

Finalmente, obtemos que

$$52 = \| u \|_2^2 = \langle u , q_1' \rangle^2 + \langle u , q_2' \rangle^2 = 52\,.$$

Assim, podemos concluir que o elemento u pertence ao subespaço gerado pelos elementos do conjunto ortogonal β.

5.12 Operadores Simétricos

Definição 5.12.1 *Sejam V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$, W um subespaço de V e $T : W \longrightarrow V$ um operador linear. Dizemos que T é um operador **simétrico** em W se*

$$\langle T(u) , v \rangle = \langle u , T(v) \rangle$$

para todos $u, v \in W$.

Exemplo 5.12.1 *Considere o espaço vetorial real $\mathcal{C}([0,1])$ munido do produto interno*

$$\langle f , g \rangle = \int_0^1 f(x)g(x)dx$$

Seja W o subespaço de $\mathcal{C}([0,1])$ definido da seguinte forma

$$W = \{ u \in \mathcal{C}^2([0,1]) \ / \ u(0) = u(1) = 0 \}$$

O operador linear $T : W \longrightarrow \mathcal{C}([0,1])$ definido por $T(u(x)) = -u''(x) + u(x)$ é um operador simétrico em W.

Vamos mostrar que T é simétrico em W. Para $u, v \in W$, temos que

$$\begin{aligned}
\langle T(u) , v \rangle &= \int_0^1 T(u(x))v(x)dx = \int_0^1 (-u''(x) + u(x))v(x)dx \\
&= -\int_0^1 u''(x)v(x)dx + \int_0^1 u(x)v(x)dx \\
&= \int_0^1 u'(x)v'(x)dx + \int_0^1 u(x)v(x)dx = \langle u , T(v) \rangle
\end{aligned}$$

Note que o resultado foi obtido fazendo uma integração por partes, na primeira integral da segunda linha, e utilizando o fato que a função $v \in W$ se anula nos extremos do intervalo de integração, isto é,

$$-\int_0^1 u''(x)v(x)dx = (-u'(1)v(1) + u'(0)v(0)) + \int_0^1 u'(x)v'(x)dx$$

Assim, provamos que T é um operador simétrico em W.

Teorema 5.12.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$, $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal para V e T um operador linear sobre V. Então, a matriz $A = [T]_\beta^\beta$ do operador linear T com relação à base ortonormal β é dada por $a_{ij} = \langle T(q_j) , q_i \rangle$.*

Demonstração – Como $\beta = \{ q_1, \cdots , q_n \}$ é uma base ortonormal para V, pelo Teorema 5.6.1, temos que todo elemento $u \in V$ é escrito de modo único como:

$$u = \sum_{i=1}^{n} \langle u , q_i \rangle q_i .$$

Desse modo, temos que o elemento $T(q_j) \in V$ é escrito de modo único como:

$$T(q_j) = \sum_{i=1}^{n} \langle T(q_j) , q_i \rangle q_i \qquad \text{para} \qquad j = 1, \cdots , n .$$

Portanto, os elemento da matriz $A = [a_{ij}]$, que é a matriz do operador linear T com relação à base ortonormal β, são dados como:

$$a_{ij} = \langle T(q_j) , q_i \rangle \qquad \text{para} \qquad i, j = 1, \cdots , n ,$$

o que completa a demonstração. ■

Teorema 5.12.2 *Sejam V um espaço vetorial real de dimensão finita com o produto interno $\langle \cdot , \cdot \rangle$, $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal para V, T um operador linear sobre V e $A = [T]_\beta^\beta$ a matriz do operador T com relação à base ortonormal β. Então, T é um operador simétrico se, e somente se, A é uma matriz simétrica.*

Demonstração – ($\Longrightarrow$) Vamos denotar por $A = [a_{ij}]$ a matriz do operador T com relação à base ortonormal β. Utilizando o resultado do Teorema 5.12.1 e a hipótese que T é um operador simétrico, temos que

$$a_{ij} = \langle T(q_j) , q_i \rangle = \langle q_j , T(q_i) \rangle = a_{ji} .$$

Logo, $A = [T]_\beta^\beta$ é uma matriz simétrica.

($\Longleftarrow$) Utilizando o resultado do Teorema 5.12.1 e a hipótese que $A = [T]_\beta^\beta$ é uma matriz simétrica, temos que

$$a_{ij} = \langle T(q_j) , q_i \rangle = a_{ji} = \langle q_j , T(q_i) \rangle .$$

Logo, T é um operador simétrico. ■

Exemplo 5.12.2 *Considere o espaço vetorial real* $I\!R^3$ *com o produto interno usual e o operador linear* T *sobre o* $I\!R^3$ *definido por* $T(x,y,z) = (x+2y, 2x+3y-z, -y+2z)$. *Mostre que* T *é um operador simétrico.*

Basta encontrar a matriz $[T]^{\beta}_{\beta}$ em relação à base canônica β. De fato,

$$T(1,0,0) \ = \ (1,2,0) \quad , \quad T(0,1,0) \ = \ (2,3,-1) \quad \text{e} \quad T(0,0,1) \ = \ (0,-1,2)$$

Portanto, obtemos

$$[T]^{\beta}_{\beta} \ = \ \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

que é uma matriz simétrica.

Definição 5.12.2 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$, W *um subespaço de* V *e* $T : W \longrightarrow V$ *um operador linear. Dizemos que* T *é um operador* ***anti–simétrico*** *em* W *se*

$$\langle\, T(u) \,,\, v \,\rangle \ = \ -\langle\, u \,,\, T(v) \,\rangle$$

para todos $u,\ v \in W$.

Exemplo 5.12.3 *Considere o espaço vetorial* $\mathcal{C}([a,b])$ *munido do produto interno*

$$\langle\, f \,,\, g \,\rangle \ = \ \int_a^b f(x)g(x)dx\,.$$

Seja W *o subespaço de* $\mathcal{C}([a,b])$ *definido por:*

$$W \ = \ \{\ f \in \mathcal{C}^1([a,b]) \ / \ f(a) \ = \ f(b)\ \}\,.$$

O operador linear $T : W \longrightarrow \mathcal{C}([a,b])$ *definido por:* $T(f)(x) \ = \ f'(x)$ *é um operador anti–simétrico em* W.

Proposição 5.12.1 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador anti–simétrico sobre* V. *Então,* $\langle\, T(u) \,,\, u \,\rangle = 0$ *para todo* $u \in V$.

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 5.12.4 *Vamos utilizar o Exemplo 5.12.3 para ilustrar a Proposição 5.12.1.*

Devemos mostrar que

$$\langle\, T(f) \,,\, f \,\rangle \ = \ \int_a^b f'(x)f(x)dx \ = \ 0 \qquad \text{para toda} \qquad f \in W\,.$$

Teorema 5.12.3 *Sejam V um espaço vetorial real de dimensão finita com o produto interno $\langle \cdot , \cdot \rangle$, $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal para V, T um operador linear sobre V e $A = [T]_\beta^\beta$ a matriz do operador linear T com relação à base ortonormal β. Então, T é um operador anti–simétrico se, e somente se, A é uma matriz anti–simétrica.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 5.12.5 *Considere o espaço vetorial real $I\!R^3$ com o produto interno usual e o operador linear T sobre o $I\!R^3$ definido por $T(x, y, z) = (-2y + z, 2x + 3z, -x - 3y)$. Mostre que T é um operador anti–simétrico.*

Basta encontrar a matriz $[T]_\beta^\beta$ em relação à base canônica β. Assim, temos que

$$[T]_\beta^\beta = \begin{bmatrix} 0 & 2 & -1 \\ -2 & 0 & -3 \\ 1 & 3 & 0 \end{bmatrix}$$

que é uma matriz anti–simétrica.

Exemplo 5.12.6 *Considere V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$. Sejam T_1 e T_2 operadores simétricos sobre V. Então, $aT_1 + bT_2$, para $a, b \in I\!R$, é um operador simétrico sobre V.*

Considerando o fato que T_1 e T_2 são operadores simétricos sobre V, temos que

$$\begin{aligned} \langle (aT_1 + bT_2)(u) , v \rangle &= \langle aT_1(u) + bT_2(u) , v \rangle \\ &= \langle aT_1(u) , v \rangle + \langle bT_2(u) , v \rangle \\ &= a\langle T_1(u) , v \rangle + b\langle T_2(u) , v \rangle \\ &= a\langle u , T_1(v) \rangle + b\langle u , T_2(v) \rangle \\ &= \langle u , aT_1(v) \rangle + \langle u , bT_2(v) \rangle \\ &= \langle u , aT_1(v) + bT_2(v) \rangle \\ &= \langle u , (aT_1 + bT_2)(v) \rangle \end{aligned}$$

o que mostra o resultado desejado.

5.13 Operadores Hermitianos

Definição 5.13.1 *Sejam V um espaço vetorial complexo munido do produto interno $\langle \cdot , \cdot \rangle$, W um subespaço de V e $T : W \longrightarrow V$ um operador linear. Dizemos que T é um operador **Hermitiano** em W se*

$$\langle T(u) , v \rangle = \langle u , T(v) \rangle$$

para todos $u, v \in W$.

Nesta seção é importante recordar o conceito de transposta Hermitiana de uma matriz $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$, que denotamos por A^*, que é definida da forma $A^* = [\overline{a}_{ji}]$. Assim, dizemos que $A \in \mathbb{M}_n(\mathbb{C})$ é uma matriz Hermitiana se $A^* = A$.

Teorema 5.13.1 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$, $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal para V, T um operador linear sobre V e $A = [T]_\beta^\beta$ a matriz do operador T com relação à base ortonormal β. Então, T é um operador Hermitiano se, e somente se, A é uma matriz Hermitiana.*

Demonstração – ($\Longrightarrow$) Vamos denotar por $A = [a_{ij}]$ a matriz do operador T com relação à base ortonormal β. Utilizando o resultado do Teorema 5.12.1 e a hipótese que T é um operador Hermitiano, temos que

$$a_{ij} = \langle T(q_j) , q_i \rangle = \langle q_j , T(q_i) \rangle = \overline{\langle T(q_i) , q_j \rangle} = \overline{a}_{ji} .$$

Logo, $A = [T]_\beta^\beta$ é uma matriz Hermitiana.

($\Longleftarrow$) Utilizando o resultado do Teorema 5.12.1 e a hipótese que $A = [T]_\beta^\beta$ é uma matriz Hermitiana, temos que

$$a_{ij} = \langle T(q_j) , q_i \rangle = \overline{a}_{ji} = \overline{\langle T(q_i) , q_j \rangle} = \langle q_j , T(q_i) \rangle .$$

Logo, T é um operador Hermitiano. $\blacksquare$

Teorema 5.13.2 *Considere* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador linear sobre* V*. Então,* T *é Hermitiano se, e somente se,* $\langle T(u) , u \rangle \in \mathbb{R}$ *para todo* $u \in V$*.*

Demonstração – Tomando a hipótese que T é um operador Hermitiano. Para todo $u \in V$, temos que

$$\overline{\langle u , T(u) \rangle} = \langle T(u) , u \rangle = \langle u , T(u) \rangle \quad \Longrightarrow \quad \langle T(u) , u \rangle \in \mathbb{R} .$$

Considerando a hipótese de que $\langle T(u) , u \rangle \in \mathbb{R}$, temos que

$$\langle T(u) , u \rangle = \overline{\langle u , T(u) \rangle} = \langle u , T(u) \rangle \qquad \text{para todo} \qquad u \in V .$$

Portanto, temos que T é um operador Hermitiano, o que completa a demonstração. ■

Definição 5.13.2 *Sejam* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$*,* W *um subespaço de* V *e* $T : W \longrightarrow V$ *um operador linear. Dizemos que* T *é um operador* ***anti–Hermitiano*** *em* W *se*

$$\langle T(u) , v \rangle = -\langle u , T(v) \rangle$$

para todos $u, v \in W$*.*

Teorema 5.13.3 *Sejam* V *um espaço vetorial complexo de dimensão finita munido do produto interno* $\langle \cdot , \cdot \rangle$*,* $\beta = \{ q_1 , \cdots , q_n \}$ *uma base ortonormal para* V*,* T *um operador linear sobre* V *e* $A = [T]^{\beta}_{\beta}$ *a matriz do operador* T *com relação à base ortonormal* β*. Então,* T *é um operador anti–Hermitiano se, e somente se,* A *é uma matriz anti–Hermitiana (*$A^* = -A$*).*

Demonstração – A prova pode ficar a cargo do leitor. □

5.14 Operadores Ortogonais

Definição 5.14.1 *Sejam* V *um espaço vetorial real com o produto interno* $\langle \cdot , \cdot \rangle$ *e* W *um subespaço de* V. *Seja* $T : W \longrightarrow V$ *um operador linear. Dizemos que* T *é um operador* ***ortogonal*** *em* W *se*

$$\langle T(u) , T(v) \rangle = \langle u , v \rangle$$

para todos $u, v \in W$.

Podemos verificar facilmente que se T é um operador ortogonal em V, então T preserva a norma Euclidiana, isto é, $\| T(u) \|_2 = \| u \|_2$ para todo $u \in V$. Assim, dizemos que T é uma **isometria** sobre V.

Proposição 5.14.1 *Sejam* V *um espaço vetorial real de dimensão finitas com o produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador ortogonal sobre* V. *Então,* T *é um automorfismo.*

Demonstração – Basta provar que T é um operador injetor, isto é, $Ker(T) = \{ 0_V \}$, e pelo Teorema do núcleo e da imagem, temos que $Im(T) = V$.

Tomando um elemento $u \in Ker(T)$, temos que

$$T(u) = 0_V \implies \| T(u) \|_2 = 0 \implies \| u \|_2 = 0 \implies u = 0_V .$$

Portanto, $Ker(T) = \{ 0_V \}$, o que completa a demonstração. ■

Proposição 5.14.2 *Sejam* V *um espaço vetorial real com o produto interno* $\langle \cdot , \cdot \rangle$ *e* T *uma isometria sobre* V. *Então,* T^{-1} *é uma isometria sobre* V.

Demonstração – Sabemos que T é um isomorfismo sobre V, pois T é uma isometria sobre V. Logo, T^{-1} existe. Desse modo,

$$\langle T^{-1}(u) , T^{-1}(v) \rangle = \langle T(T^{-1}(u)) , T(T^{-1}(v)) \rangle = \langle u , v \rangle .$$

Portanto, mostramos que T^{-1} é uma isometria sobre V. ■

Proposição 5.14.3 *Sejam V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$ e T um operador linear sobre V. Então, T é uma isometria sobre V se, e somente se, T é um operador ortogonal em V.*

Demonstração

($\Longrightarrow$) Tomando a hipótese que T é uma isometria sobre V, obtemos

$$\langle T(u-v) , T(u-v) \rangle = \langle u-v , u-v \rangle = \langle u , u \rangle - 2\langle u , v \rangle + \langle v , v \rangle .$$

para todos $u, v \in V$. Por outro lado, temos que

$$\begin{aligned} \langle T(u-v) , T(u-v) \rangle &= \langle T(u) , T(u) \rangle - 2\langle T(u) , T(v) \rangle + \langle T(v) , T(v) \rangle \\ &= \langle u , u \rangle - 2\langle T(u) , T(v) \rangle + \langle v , v \rangle \end{aligned}$$

Portanto, comparando as duas expressões, obtemos

$$\langle T(u) , T(v) \rangle = \langle u , v \rangle$$

para todos $u, v \in V$. Logo, mostramos que T é um operador ortogonal em V.

($\Longleftarrow$) Tomando a hipótese que T é um operador ortogonal em V, isto é,

$$\langle T(u) , T(v) \rangle = \langle u , v \rangle \qquad \text{para todos} \quad u, v \in V ,$$

obtemos $\langle T(v) , T(v) \rangle = \langle v , v \rangle$ para todo $v \in V$. Logo, $\| T(v) \|_2 = \| v \|_2$ para todo $v \in V$. Portanto, provamos que T é uma isometria sobre V. ■

Proposição 5.14.4 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot , \cdot \rangle$, T e P isometrias sobre V. Então, $T \circ P$ é uma isometria sobre V.*

Demonstração – Tomando a hipótese que T e P são isometrias sobre V, isto é,

$$\| T(v) \|_2 = \| v \|_2 \qquad \text{e} \qquad \| P(v) \|_2 = \| v \|_2$$

para todo $v \in V$, obtemos

$$\| (T \circ P)(v) \|_2 = \| (T(P(v)) \|_2 = \| P(v) \|_2 = \| v \|_2$$

para todo $v \in V$. Portanto, temos que $T \circ P$ é uma isometria sobre V. ■

Exemplo 5.14.1 *Considere o espaço vetorial real* $\mathbb{R}^2$ *munido do produto interno usual. Os operadores lineares* $T(x,y) = (x, -y)$ *e* $P(x,y) = (-x, -y)$ *, que representam uma reflexão em torno do eixo–ox e uma reflexão em torno da origem, respectivamente, são isometrias sobre o* $\mathbb{R}^2$*. Assim, o operador linear* $T \circ P$ *sobre o* $\mathbb{R}^2$ *que é dado por:*

$$(T \circ P)(x,y) \ = \ T(P(x,y)) \ = \ T(-x,-y) \ = \ (-x,y) \ ,$$

que representa uma reflexão em torno do eixo–oy, é uma isometria sobre o $\mathbb{R}^2$*.*

Teorema 5.14.1 *Sejam* V *um espaço vetorial real de dimensão finita com o produto interno* $\langle \cdot , \cdot \rangle$ *,* $\beta = \{ q_1 , \cdots , q_n \}$ *uma base ortonormal para* V *e* T *um operador linear sobre* V*. Então,* T *é um operador ortogonal em* V *se, e somente se,* T *leva a base ortonormal* β *na base ortonormal* $\{ T(q_1) , \cdots , T(q_n) \}$ *de* V*.*

Demonstração – Para todo $u, \, v \in V$ temos que

$$u \ = \ \sum_{i=1}^{n} b_i \, q_i \qquad \text{e} \qquad v \ = \ \sum_{j=1}^{n} c_j \, q_j$$

onde b_i e c_i , $i = 1, \cdots , n$, são os coeficientes de Fourier de u e de v com relação à base ortonormal β, respectivamente. Pela identidade de Parseval, temos que

$$\langle \, u \, , \, v \, \rangle \ = \ \sum_{i=1}^{n} b_i \, c_i$$

Desse modo, temos que

$$\langle \, T(u) \, , \, T(v) \, \rangle \ = \ \sum_{i=1}^{n} \sum_{j=1}^{n} b_i \, c_j \, \langle \, T(q_i) \, , \, T(q_j) \, \rangle$$

Portanto, $\langle \, T(u) \, , \, T(v) \, \rangle \ = \ \langle \, u \, , \, v \, \rangle$ se, e somente se, $\{ \, T(q_1) \, , \, \cdots \, , \, T(q_n) \, \}$ é um conjunto ortonormal em V, o que completa a demonstração. ■

Definição 5.14.2 *Dizemos que* $Q \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz* **ortogonal** *se* $Q^t Q = I$*. Assim, temos que* $Q Q^t = I$*. Desse modo, tem–se que* $Q^{-1} = Q^t$*.*

Teorema 5.14.2 *Uma matriz* $Q \in \mathbb{M}_n(\mathbb{R})$ *é ortogonal se, e somente se, suas colunas (suas linhas) formam um conjunto ortonormal em* $\mathbb{R}^n$*.*

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 5.14.3 *Sejam* V *um espaço vetorial real de dimensão finita com o produto interno* $\langle \cdot , \cdot \rangle$, $\beta = \{ q_1 , \cdots , q_n \}$ *uma base ortonormal para* V, T *um operador linear sobre* V *e* $A = [T]_\beta^\beta$ *a matriz do operador* T *com relação à base ortonormal* β. *Então,* T *é um operador ortogonal se, e somente se,* A *é uma matriz ortogonal.*

Demonstração – Seja $A = [a_{ij}]$ a representação matricial do operator T com relação à base ortonormal β, isto é,

$$T(q_i) = \sum_{k=1}^{n} a_{ki}\, q_k \qquad \text{e} \qquad T(q_j) = \sum_{r=1}^{n} a_{rj}\, q_r$$

Desse modo, temos que

$$\langle\, T(q_i) \,,\, T(q_j) \,\rangle = \sum_{k=1}^{n} \sum_{r=1}^{n} a_{ki}\, a_{rj} \,\langle\, q_k \,,\, q_r \,\rangle = \sum_{k=1}^{n} a_{ki}\, a_{kj}$$

Portanto,

$$\langle\, T(q_i) \,,\, T(q_j) \,\rangle = \delta_{ij} \qquad \Longleftrightarrow \qquad \sum_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij} \,,$$

onde δ_{ij} é o delta de Kronecker, o que completa a demonstração. ■

Teorema 5.14.4 *Seja* V *é um espaço vetorial real de dimensão finita com o produto interno* $\langle \cdot , \cdot \rangle$ *com* β *e* γ *duas bases ortonormais de* V. *Então, a matriz de mudança de base* $[I]_\beta^\gamma$ *é uma matriz ortogonal.*

Demonstração – Digamos que $dim(V) = n$. Sejam $\beta = \{ q_1, \cdots , q_n \}$ e $\gamma = \{ v_1, \cdots , v_n \}$. Vamos denotar por $C = [c_{ij}] = [I]_\beta^\gamma$. Assim, temos que

$$v_j = \sum_{l=1}^{n} c_{lj}\, q_l \qquad \text{e} \qquad v_i = \sum_{k=1}^{n} c_{ki}\, q_k$$

Pela identidade de Parseval tem–se que

$$\langle\, v_i \,,\, v_j \,\rangle = \sum_{k=1}^{n} c_{ki}\, c_{kj} \qquad \Longrightarrow \qquad C^t\, C = I$$

pois $\langle\, v_i \,,\, v_j \,\rangle = 1$ para $i = j$ e $\langle\, v_i \,,\, v_j \,\rangle = 0$ para $i \neq j$, o que completa a demonstração. ■

Exercícios

Exercício 5.52 *Considere* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$. *Sejam* T_1 *e* T_2 *operadores simétricos sobre* V. *Então,* $T_1 \circ T_2$ *é um operador simétrico sobre* V *se, e somente se,* $T_1 \circ T_2 = T_2 \circ T_1$.

Exercício 5.53 *Considere* V *um espaço vetorial real de dimensão finita munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador linear simétrico sobre* V. *Mostre que* $Ker(T) = (\, Im(T)\,)^{\perp}$.

Exercício 5.54 *Considere* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$. *Sejam* T_1 *e* T_2 *operadores Hermitianos sobre* V. *Então,* $aT_1 + bT_2$, *para* $a, b \in \mathbb{R}$, *é um operador Hermitiano sobre* V.

Exercício 5.55 *Considere* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$. *Sejam* T_1 *e* T_2 *operadores Hermitianos sobre* V. *Então,* $T_1 \circ T_2$ *é um operador Hermitiano sobre* V *se, e somente se,* $T_1 \circ T_2 = T_2 \circ T_1$.

Exercício 5.56 *Considere* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$. *Seja* T *um operador Hermitiano sobre* V. *Então,*

$$\| T(u) \pm iu \|_2^2 = \| T(u) \|_2^2 + \| u \|_2^2 \qquad \text{para todo} \qquad u \in V .$$

Exercício 5.57 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $T \in L(V)$. *Mostre que duas quaisquer das propriedades implicam a outra:*

(a) T *é simétrico.*

(b) T *é uma isometria sobre* V.

(c) $T^2 = I$.

Exercício 5.58 *Seja* $Q \in \mathbb{M}_2(\mathbb{R})$ *a matriz que representa uma rotação de um ângulo* θ *no sentido anti–horário, isto é,*

$$Q = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Mostre que Q *é uma matriz ortogonal.*

Exercício 5.59 *Considere o espaço vetorial* $\mathbb{R}^3$ *com o produto interno usual e seja* T *o operador linear sobre* $\mathbb{R}^3$ *definido por:*

$$T(x,y,z) = (x\cos(\theta) - y\sin(\theta),\, x\sin(\theta) + y\cos(\theta),\, z)\,,$$

onde θ *é um ângulo fixo. Mostre que* T *é um operador ortogonal.*

Exercício 5.60 *Considere* V *um espaço vetorial real munido do produto interno* $\langle\cdot,\cdot\rangle$ *e* W *um subespaço de dimensão finita de* V*. Sabemos que* $V = W \oplus W^{\perp}$*, isto é, todo elemento* $v \in V$ *é escrito de modo único da forma* $v = w + u$ *com* $w \in W$ *e* $u \in W^{\perp}$*. Seja* T *o operador linear sobre* V *definido da seguinte forma:* $T(v) = w - u$ *para todo* $v \in V$*.*

(a) Prove que T *é um operador linear simétrico e ortogonal em* V*.*

(b) Considerando $V = \mathbb{R}^3$ *com o produto interno usual e* W *o subespaço gerado pelo elemento* $w = (1,1,1)$*, encontre a matriz do operador linear* T *com relação à base canônica do* $\mathbb{R}^3$*.*

Exercício 5.61 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle\cdot,\cdot\rangle$ *e* G *o conjunto de todos os operadores ortogonais sobre* V*, isto é,*

$$G = \{\, T : V \longrightarrow V \ / \ T \text{ é um operador ortogonal} \,\}\,.$$

Mostre que G *tem uma* ***estrutura de grupo*** *em relação à operação de composição, isto é,* $(G, \circ)$ *é um* ***grupo****.*

Exercício 5.62 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle\cdot,\cdot\rangle$ *e* T *uma isometria sobre* V*. Mostre que* T *preserva o cosseno do ângulo entre dois elementos não–nulos de* V*.*

Exercício 5.63 *Considere a matriz* $A \in \mathbb{M}_2(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}.$$

Mostre que $B = (I_2 - A)(I_2 + A)^{-1}$ *é uma matriz ortogonal.*

Exercício 5.64 *Considere o espaço vetorial real* $\mathbb{M}_n(\mathbb{R})$ *munido do produto interno usual. Dada uma matriz* $Q \in \mathbb{M}_n(\mathbb{R})$*, definimos o operador linear* T_Q *sobre* $\mathbb{M}_n(\mathbb{R})$ *da seguinte forma:* $T_Q(X) = QX$*. Mostre que* T_Q *é uma isometria sobre* $\mathbb{M}_n(\mathbb{R})$ *se, e somente se,* Q *é uma matriz ortogonal.*

5.15 Projeção Ortogonal

A partir do Teorema da Decomposição Ortogonal temos a definição de projeção ortogonal, que é extremamente útil para a solução e interpretação geométrica de certas aplicações da Álgebra Linear.

Definição 5.15.1 *Sejam V um espaço vetorial munido do produto interno $\langle \cdot , \cdot \rangle$ e S um subespaço de dimensão finita de V, com $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal de S. Dado um elemento $u \in V$, o elemento $\widetilde{u} \in S$ definido por:*

$$\widetilde{u} \;=\; \sum_{j=1}^{n} \langle \, u \, , \, q_j \, \rangle \, q_j$$

é a projeção ortogonal do elemento u sobre o subespaço S e o elemento $w = u - \widetilde{u}$ é a projeção ortogonal do elemento u sobre o subespaço $S^{\perp}$.

Podemos observar que se $\beta = \{ q_1 , \cdots , q_n \}$ é uma **base ortogonal** para S, então a projeção ortogonal do elemento $u \in V$ sobre o subespaço S é dado por

$$\widetilde{u} \;=\; \sum_{j=1}^{n} \frac{\langle \, u \, , \, q_j \, \rangle}{\langle \, q_j \, , \, q_j \, \rangle} \, q_j \, .$$

Exemplo 5.15.1 *Considere o espaço vetorial real $I\!R^3$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Seja W o subespaço gerado pelo elemento $w = (1, -1, 2) \in I\!R^3$. Calcule a projeção ortogonal do elemento $u = (2, -1, 4) \in I\!R^3$ sobre o subespaço W e sobre o subespaço $W^{\perp}$.*

Pela Definição 5.15.1, temos que a projeção ortogonal de u sobre W é dada por:

$$\widetilde{u} \;=\; \frac{\langle \, u \, , \, w \, \rangle}{\langle \, w \, , \, w \, \rangle} w$$

Pelo Teorema 5.9.1 (Decomposição Ortogonal), temos que o elemento $v \in W^{\perp}$ dado por:

$$v \;=\; u \;-\; \widetilde{u} \;=\; u \;-\; \frac{\langle \, u \, , \, w \, \rangle}{\langle \, w \, , \, w \, \rangle} w$$

é a projeção ortogonal de u sobre o subespaço $W^{\perp}$. Assim, obtemos

$$\widetilde{u} \;=\; \frac{11}{6}(1, -1, 2) \qquad \text{e} \qquad v \;=\; (2, -1, 4) \;-\; \frac{11}{6}(1, -1, 2) \;=\; \frac{1}{6}(1, 5, 2) \, .$$

Exemplo 5.15.2 *Considere o espaço vetorial* $I\!R^n$ *com o produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* W *o subespaço gerado pelo elemento* $w \in I\!R^n$ *não–nulo. Dado um elemento* $u \in I\!R^n$, *determinar sua projeção ortogonal sobre* W. *Utilizando o Teorema 5.9.1, da decomposição ortogonal, determine a projeção ortogonal de* u *sobre o subespaço* $W^{\perp}$.

Pela definição 5.15.1, temos que a projeção ortogonal de u sobre W é dada por

$$\widetilde{u} = \frac{\langle u , w \rangle}{\langle w , w \rangle} w$$

Pelo teorema da decomposição ortogonal, temos que

$$v = u - \widetilde{u} = u - \frac{\langle u , w \rangle}{\langle w , w \rangle} w$$

é a projeção ortogonal de u no subespaço $W^{\perp}$.

Exemplo 5.15.3 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $\| \cdot \|_2$ *a norma Euclidiana. Considerando os elementos* $u, v \in V$, *com* $v \neq 0_V$, *determine o elemento* w^* *do conjunto* $S = \{ w \in V \ / \ w = u - tv \ , \ t \in I\!R \}$ *que possui a menor norma Euclidiana. Dê uma interpretação geométrica para* w^*.

Temos que encontrar um elemento $w^* \in S$ tal que

$$\| w^* \|_2 = \min\{ \| w \|_2 \ ; \ w \in S \} \iff \| w^* \|_2^2 = \min\{ \| u - tv \|_2^2 \ ; \ t \in I\!R \}$$

Portanto, temos que encontrar o mínimo da função $g(t)$ dada por

$$g(t) = \| u - tv \|_2^2 = \langle u - tv , u - tv \rangle = \langle u , u \rangle - 2t\langle u , v \rangle + t^2\langle v , v \rangle \quad ; \quad t \in I\!R$$

Inicialmente, vamos calcular os pontos críticos, fazendo $g'(t) = 0$, obtemos

$$t^* = \frac{\langle u , v \rangle}{\langle v , v \rangle} \qquad \Longrightarrow \qquad w^* = u - t^*v = u - \frac{\langle u , v \rangle}{\langle v , v \rangle} v$$

Temos que classificar o único ponto crítico da função g, calculando $g''(t)$, obtemos

$$g''(t) = 2 \langle v , v \rangle > 0 \quad ; \quad v \neq 0_V$$

Assim, t^* é um ponto de mínimo global da função g e o elemento $w^* = u - t^*v \in S$ é o elemento de menor norma Euclidiana.

Portanto, o elemento w^* é a projeção ortogonal do elemento u sobre o subespaço $U^{\perp}$ e o elemento $u^* = t^*v$ é a projeção ortogonal do elemento u sobre o subespaço $U = [v]$. Mais a frente vamos dar uma nova interpretação para a projeção ortogonal.

Vamos representar as projeções ortogonais através de operadores lineares. Considerando a definição de **projeção ortogonal**, definimos um operador P sobre V da forma:

$$P(u) \;=\; \sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, q_j \qquad \text{para todo} \qquad u \,\in\, V\,.$$

Podemos verificar facilmente que P é um operador linear sobre V. Desse modo, vamos denominar P como sendo o **operador de projeção ortogonal** sobre o subespaço S. Portanto, temos que $Im(P) \;=\; S$. Para o operador de projeção ortogonal podemos apresentar os seguintes resultados.

Teorema 5.15.1 *Sejam V um espaço vetorial real com o produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$, S um subespaço de dimensão finita de V e P o operador de projeção ortogonal sobre S. Então, P é um operador* ***simétrico****, isto é,*

$$\langle\, P(u)\,,\, v\,\rangle \;=\; \langle\, u\,,\, P(v)\,\rangle \qquad ; \qquad \forall\; u, v \,\in\, V$$

Demonstração – Seja $\beta \;=\; \{\, q_1\,,\,\cdots\,,\, q_n \,\}$ uma base ortonormal para S. Dado um elemento $u \,\in\, V$ sabemos que sua projeção ortogonal sobre S é dado por

$$P(u) \;=\; \sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, q_j$$

Desse modo, temos que

$$\langle\, P(u)\,,\, v\,\rangle \;=\; \sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, \langle\, q_j\,,\, v\,\rangle \;=\; \langle\, u\,,\, \sum_{j=1}^{n} \langle\, v\,,\, q_j \,\rangle\, q_j \,\rangle \;=\; \langle\, u\,,\, P(v)\,\rangle$$

o que completa a prova. ■

Teorema 5.15.2 *Sejam V um espaço vetorial real com o produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$, S um subespaço de dimensão finita de V e P o operador de projeção ortogonal sobre S. Então, $P^2 \;=\; P$ (**idempotente**), isto é, $P^2(u) \;=\; P(u)$ para todo $u \in V$.*

Demonstração – Seja $\beta \;=\; \{\, q_1\,,\,\cdots\,,\, q_n \,\}$ uma base ortonormal para S. Dado um elemento $u \,\in\, V$ sabemos que sua projeção ortogonal sobre S é dado por

$$P(u) \;=\; \sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, q_j$$

Desse modo, temos que

$$P^2(u) \;=\; P\left(\sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, q_j \right) \;=\; \sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, P(q_j) \;=\; \sum_{j=1}^{n} \langle\, u\,,\, q_j \,\rangle\, q_j \;=\; P(u)$$

desde que $P(q_j) \;=\; q_j$ para $j = 1, \cdots, n$. O que completa a prova. ■

Podemos observar que se $P^2 = P$ então, $P(u) = u$, para todo $u \in Im(P)$. De fato,

$$u \in Im(P) \implies u = P(v) \implies P(u) = P(P(v)) = P(v) = u\,.$$

Definição 5.15.2 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot, \cdot \rangle$ e P um operador linear sobre V. Dizemos que P é um operador de projeção ortogonal se P for simétrico e idempotente.*

Teorema 5.15.3 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot, \cdot \rangle$, S um subespaço de dimensão finita de V e P o operador de projeção ortogonal sobre S. Então, $Ker(P) = S^{\perp}$.*

Demonstração – Inicialmente, tomamos $u \in Ker(P)$, isto é, $P(u) = 0_V$. Desse modo, para todo $v \in V$, temos que

$$0 = \langle P(u) , v \rangle = \langle u , P(v) \rangle\,.$$

como $P(v) \in Im(P) = S$, obtemos que $u \in S^{\perp}$. Assim, $Ker(P) \subset S^{\perp}$.

Finalmente, tomamos $v \in S^{\perp}$, isto é, $\langle v , u \rangle = 0$ para todo $u \in S$. Como $u = P(w)$ para $w \in V$. Desse modo, obtemos

$$\langle v , u \rangle = \langle v , P(w) \rangle = \langle P(v) , w \rangle = 0 \qquad \text{para todo} \qquad w \in V\,.$$

Logo, $P(v) = 0_V$ o que implica em $v \in Ker(P)$.

Assim, mostramos que $S^{\perp} \subset Ker(P)$. Portanto, provamos que $Ker(P) = S^{\perp}$. ■

O Teorema 5.15.3 apresenta um resultado muito interessante para o estudo de autovalores e autovetores, que iremos ver no Capítulo 6. Podemos apresentar tal resultado da forma

$$P(u) = 0_V \qquad \text{para todo} \qquad u \in S^{\perp}$$

para que possamos obter as conclusões desejadas.

Corolário 5.15.1 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot, \cdot \rangle$, S um subespaço de dimensão finita de V e P o operador de projeção ortogonal sobre S. Então, $V = Ker(P) \oplus Im(P)$.*

Teorema 5.15.4 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot , \cdot \rangle$, S um subespaço de dimensão finita de V e P o operador de projeção ortogonal sobre S. Então, $I - P$ é o operador de projeção ortogonal sobre o subespaço $S^{\perp}$.*

Demonstração – Dado um elemento $u \in V$, temos que $P(u) \in S$. Vamos mostrar que o elemento $u - P(u)$ é ortogonal ao elemento $P(u)$. Pelo Teorema 5.15.1, sabemos que P é simétrico. Assim, temos que

$$\langle\, P(u)\,,\, u - P(u)\,\rangle \;=\; \langle\, u\,,\, P(u - P(u))\,\rangle \;=\; \langle\, u\,,\, P(u) - P^2(u)\,\rangle$$

Pelo Teorema 5.15.2, sabemos que $P^2 = P$. Logo, obtemos

$$\langle\, P(u)\,,\, u - P(u)\,\rangle \;=\; \langle\, u\,,\, P(u) - P(u)\,\rangle \;=\; \langle\, u\,,\, 0_V\,\rangle \;=\; 0 \qquad ; \qquad \forall\; u \in V$$

Portanto, $u - P(u) \in S^{\perp}$ para todo $u \in V$. Assim, provamos que $Im(I - P) = S^{\perp}$.

Podemos verificar facilmente que $I - P$ é um operador idempotente. De fato,

$$(I - P)^2 \;=\; (I - P)(I - P) \;=\; I - 2P + P^2 \;=\; I - P$$

Temos também que $I - P$ é um operador simétrico. De fato,

$$\langle\, (I - P)(u)\,,\, v\,\rangle = \langle\, u\,,\, v\,\rangle \;-\; \langle\, P(u)\,,\, v\,\rangle = \langle\, u\,,\, v\,\rangle \;-\; \langle\, u\,,\, P(v)\,\rangle = \langle\, u\,,\, (I - P)(v)\,\rangle$$

o que completa a prova. ■

Corolário 5.15.2 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot , \cdot \rangle$ e P um operador linear sobre V. Então, o operador P é simétrico e idempotente se, e somente se, P projeta ortogonalmente todo elemento $v \in V$ sobre o subespaço $Im(P)$.*

Exemplo 5.15.4 *Determine a projeção ortogonal do elemento $u = (2, 1, 2, 1) \in I\!R^4$ no subespaço $S = [q_1, q_2]$ onde $q_1 = (1, -1, 1, -1)$ e $q_2 = (-2, 1, 4, 1)$.*

Note que os geradores do subespaço S são ortogonais entre si. Logo, é uma base ortogonal para S. Desse modo, temos que a projeção ortogonal do elemento u sobre o subespaço S é dado por

$$\widetilde{u} \;=\; \frac{\langle\, u\,,\, q_1\,\rangle}{\langle\, q_1\,,\, q_1\,\rangle}\, q_1 \;+\; \frac{\langle\, u\,,\, q_2\,\rangle}{\langle\, q_2\,,\, q_2\,\rangle}\, q_2 \;=\; \frac{1}{2}\, q_1 \;+\; \frac{3}{11}\, q_2\,.$$

Assim, temos também que o elemento $w \in I\!R^4$ dado por:

$$w \;=\; u \;-\; \widetilde{u} \;=\; u \;-\; \left(\frac{1}{2}\, q_1 \;+\; \frac{3}{11}\, q_2 \right)$$

é a projeção ortogonal do elemento u sobre o subespaço $S^{\perp}$.

Teorema 5.15.5 *Sejam V um espaço vetorial real de dimensão finita com o produto interno $\langle \cdot , \cdot \rangle$, β uma base para V, P um operador linear sobre V e $[P]_\beta^\beta$ a matriz do operador P com relação à base β. Então, P é um operador idempotente se, e somente se, $[P]_\beta^\beta$ é uma matriz idempotente.*

Demonstração – Para todo $v \in V$, temos que $[P(v)]_\beta = [P]_\beta^\beta [v]_\beta$.

Pelo Teorema 4.8.3, temos que

$$[P^2(v)]_\beta = [(P \circ P)(v)]_\beta = ([P]_\beta^\beta)^2 [v]_\beta \qquad \text{para todo} \qquad v \in V .$$

Tomando por hipótese que P é um operador idempotente, temos que

$$[P^2(v)]_\beta = ([P]_\beta^\beta)^2 [v]_\beta = [P(v)]_\beta = [P]_\beta^\beta [v]_\beta \qquad \text{para todo} \qquad v \in V .$$

Portanto, provamos que $([P]_\beta^\beta)^2 = [P]_\beta^\beta$, isto é, $[P]_\beta^\beta$ é uma matriz idempotente.

Tomando por hipótese que $[P]_\beta^\beta$ é uma matriz idempotente, temos que

$$[P^2(v)]_\beta = ([P]_\beta^\beta)^2 [v]_\beta = [P]_\beta^\beta [v]_\beta = [P(v)]_\beta \qquad \text{para todo} \qquad v \in V .$$

Portanto, provamos que $P^2(v) = P(v)$ para todo $v \in V$. Logo, P é um operador idempotente, o que completa a demonstração. ■

Corolário 5.15.3 *Sejam V um espaço vetorial real de dimensão finita com o produto interno $\langle \cdot , \cdot \rangle$, β uma base ortonormal para V, P um operador linear sobre V e $[P]_\beta^\beta$ a matriz do operador P com relação à base ortonormal β. Então, P é um operador de projeção ortogonal se, e somente se, $[P]_\beta^\beta$ é uma matriz simétrica e idempotente.*

Exemplo 5.15.5 *Considere o espaço vetorial real* $I\!R^2$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* W *o subespaço gerado pelo elemento* $w = (1, 2) \in I\!R^2$. *Seja* P *o operador de projeção ortogonal sobre o subespaço* W. *Determine o operador* $P(x, y)$.

Dado um elemento genérico $u = (x, y) \in I\!R^2$, pela definição 5.15.1, temos que o operador de projeção ortogonal sobre W é dado por:

$$P(x,y) \;=\; \frac{\langle\, u \,,\, w \,\rangle}{\langle\, w \,,\, w \,\rangle} w \;=\; \frac{1}{5}(x + 2y,\, 2x + 4y)\,.$$

Para exemplificar que β não precisa ser uma base ortonormal no Teorema 5.15.5, vamos determinar a matriz $[P]_\beta^\beta$ com relação à base $\beta = \{\ (1, 1), (0, 1)\,\}$, para o caso do Exemplo 5.15.5. Desse modo, obtemos que

$$[P]_\beta^\beta \;=\; \frac{1}{5}\begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix}\,.$$

Podemos verificar facilmente que $[P]_\beta^\beta$ é uma matriz idempotente, mas não é simétrica, pois β não é ortonormal.

Exemplo 5.15.6 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* W *o subespaço gerado pelo elemento* $w = (1, -1, 2) \in I\!R^3$. *Seja* P *o operador de projeção ortogonal sobre o subespaço* W. *Determine o operador* $P(x, y, z)$ *e a matriz* $[P]_\beta^\beta$, *em relação à base canônica do* $I\!R^3$.

Dado um elemento genérico $u = (x, y, z) \in I\!R^3$, pela definição 5.15.1, temos que o operador de projeção ortogonal sobre W é dado por:

$$P(x,y,z) \;=\; \frac{\langle\, u \,,\, w \,\rangle}{\langle\, w \,,\, w \,\rangle} w \;=\; \frac{1}{6}(x - y + 2z,\, -x + y - 2z,\, 2x - 2y + 4z)\,.$$

Portanto, a matriz $[P]_\beta^\beta$ é dada por:

$$[P]_\beta^\beta \;=\; \frac{1}{6}\begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

onde β é a base canônica do $I\!R^3$. Podemos verificar facilmente que a matriz $[P]_\beta^\beta$ é simétrica e idempotente, pois β é uma base ortonormal, de acordo com o Corolário 5.15.3.

Exercícios

Exercício 5.65 *Considere o espaço vetorial* $\mathbb{R}^n$ *com o produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* W *o subespaço gerado pelo elemento* $w \in \mathbb{R}^n$, *não–nulo, e* P *o operador de projeção ortogonal sobre* W. *Mostre que a matriz* $[P]^{\beta}_{\beta}$, *onde* β *é a base canônica do* $\mathbb{R}^n$, *é dada por:*

$$[P]^{\beta}_{\beta} = \frac{ww^t}{w^t w},$$

considerando que o elemento $w \in \mathbb{R}^n$ *está representado na forma de vetor coluna.*

Exercício 5.66 *Considere o espaço vetorial* $\mathbb{R}^3$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* S *o subespaço de* $\mathbb{R}^3$ *dado por:* $S = [(1,1,1)]$ *e* $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *o operador de projeção ortogonal sobre* S. *Determine o operador* $P(x,y,z)$.

Exercício 5.67 *Seja* $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *o operador linear tal que* $u = P(v)$ *é a projeção ortogonal do elemento* $v \in \mathbb{R}^3$ *no plano* $3x + 2y + z = 0$. *Pede-se:*

(a) Encontre o operador $P(x,y,z)$.

(b) Determine a imagem do operador P.

(c) Determine o núcleo do operador P.

Exercício 5.68 *Seja* $H = \{ x \in \mathbb{R}^n \ / \ \langle x , c \rangle = 0 \}$ *para* $c \in \mathbb{R}^n$, *fixo. Pede–se:*

(a) Determine o subespaço $H^{\perp}$.

(b) Dado um elemento $u \in \mathbb{R}^n$, *determine sua projeção ortogonal no subespaço* $H^{\perp}$.

(c) Dado um elemento $u \in \mathbb{R}^n$, *determine sua projeção ortogonal no subespaço* H.

Exercício 5.69 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *com o produto interno*

$$\langle p , q \rangle = \int_0^1 p(x)q(x)dx \qquad ; \qquad \forall \ p, q \in \mathcal{P}_2(\mathbb{R}).$$

Determine a projeção ortogonal do elemento $q(x) = x^2$ *sobre o subespaço* $\mathcal{P}_1(\mathbb{R})$.

Resposta: $\boxed{\widetilde{q}(x) = -\frac{1}{6} + x}$

5.16 Reflexão sobre um Subespaço

Definição 5.16.1 *Sejam* V *um espaço vetorial real com o produto interno* $\langle \cdot , \cdot \rangle$, S *um subespaço de dimensão finita de* V *e* $P : V \longrightarrow V$ *o operador de projeção ortogonal sobre* S. *O operador* $R : V \longrightarrow V$ *definido por:* $R = 2P - I$, *isto é,*

$$R(v) = 2P(v) - v \qquad \text{para todo} \qquad v \in V,$$

é a ***reflexão*** *sobre o subespaço* S, *paralelamente ao subespaço* $S^{\perp}$.

A seguir vamos apresentam dois exemplos que são muito interessantes para o estudo de autovalores e autovetores, que iremos ver no Capítulo 6.

Exemplo 5.16.1 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* S *um subespaço de dimensão finita de* V. *Seja* R *o operador de reflexão sobre* S. *Podemos verificar facilmente que* $R(w) = -w$ *para todo* $w \in S^{\perp}$.

Exemplo 5.16.2 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* S *um subespaço de dimensão finita de* V. *Seja* R *o operador de reflexão sobre* S. *Podemos verificar facilmente que* $R(u) = u$ *para todo* $u \in S$.

Teorema 5.16.1 *Sejam* V *um espaço vetorial real com o produto interno* $\langle \cdot , \cdot \rangle$, S *um subespaço de dimensão finita de* V *e* R *o operador de reflexão sobre* S, *paralelamente ao subespaço* $S^{\perp}$. *Então,* R *é um operador simétrico, isto é,*

$$\langle R(u) , v \rangle = \langle u , R(v) \rangle \qquad \text{para todo} \qquad u, v \in V.$$

Demonstração – Fazendo uso da definição de operador simétrico e do fato que o operador de projeção ortogonal P é simétrico, temos que

$$\begin{aligned}
\langle R(u) , v \rangle &= \langle 2P(u) - u , v \rangle \\
&= 2\langle P(u) , v \rangle - \langle u , v \rangle \\
&= \langle u , 2P(v) \rangle - \langle u , v \rangle \\
&= \langle u , 2P(v) - v \rangle \\
&= \langle u , R(v) \rangle
\end{aligned}$$

para todo $u, v \in V$. Portanto, R é um operador simétrico. ■

Teorema 5.16.2 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot , \cdot \rangle$, S um subespaço de dimensão finita de V e R o operador de reflexão sobre S, paralelamente ao subespaço $S^{\perp}$. Então, R é um operador auto–reflexivo ($R^2 = I$), isto é, $(R \circ R)(v) = R(R(v)) = v$ para todo $v \in V$.*

Demonstração – Fazendo uso do fato que o operador de projeção ortogonal P é idempotente, isto é $P^2 = P$, temos que

$$R^2 = (2P - I)(2P - I) = 4P^2 - 4P + I = I .$$

Portanto, R é um operador auto–reflexivo. ■

Teorema 5.16.3 *Sejam V um espaço vetorial real com o produto interno $\langle \cdot , \cdot \rangle$, S um subespaço de dimensão finita de V e R o operador de reflexão sobre S, paralelamente ao subespaço $S^{\perp}$. Então, R é um operador ortogonal, isto é,*

$$\langle R(u) , R(v) \rangle = \langle u , v \rangle \qquad \textit{para todos} \qquad u, v \in V .$$

Demonstração – A prova é feito utilizando o fato que R é um operador simétrico e auto–reflexivo. De fato,

$$\langle R(u) , R(v) \rangle = \langle u , R(R(v)) \rangle = \langle u , R^2(v) \rangle = \langle u , v \rangle$$

para todos $u, v \in V$. Portanto, R é um operador ortogonal. ■

Podemos observar que, como R é um operador ortogonal em V, temos que R preserva a norma Euclidiana, isto é, $\| R(u) \|_2 = \| u \|_2$ para todo $u \in V$. Assim, R é uma **isometria** sobre V.

Exemplo 5.16.3 *Considere o espaço vetorial $\mathbb{R}^2$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Seja S o subespaço de $\mathbb{R}^2$ dado por: $S = \{ (x,y) \in \mathbb{R}^2 \ / \ x - 2y = 0 \}$. Dado o elemento $u = (-1, 2) \in \mathbb{R}^2$, determine sua reflexão sobre o subespaço S.*

Temos que o subespaço S é gerado pelo elemento $v = (1, 2)$. Assim, a projeção ortogonal de u sobre S é dada por:

$$\widetilde{u} = P(u) = \frac{\langle u , v \rangle}{\langle v , v \rangle} v = \frac{3}{5}(1, 2)$$

Portanto, a reflexão do elemento u sobre o subespaço S é dada por:

$$w = R(u) = 2\widetilde{u} - u = \frac{6}{5}(1, 2) - (-1, 2) = \frac{1}{5}(11, 2) .$$

Note que $\| v \|_2 = \| w \|_2 = \sqrt{5}$.

Exemplo 5.16.4 *Considere o espaço vetorial* $I\!R^2$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* S *o subespaço de* $I\!R^2$ *dado por:* $S = \{ (x,y) \in I\!R^2 \ / \ 2x - y = 0 \}$ *e* $R : I\!R^2 \longrightarrow I\!R^2$ *o operador de reflexão sobre* S. *Determine o operador* $R(x,y)$.

Temos que o subespaço S é gerado pelo elemento $v = (1,2)$. Dado um elemento genérico $u = (x,y) \in I\!R^2$, temos que

$$
\begin{aligned}
R(x,y) \ = \ 2P(u) - (u) \ &= \ 2\frac{\langle \, u \, , \, v \, \rangle}{\langle \, v \, , \, v \, \rangle} v \ - \ u \\
&= \ \frac{2x \, + \, 4y}{5}(1,2) \ - \ (x,y) \\
&= \ \frac{1}{5}(-3x + 4y, \, 4x + 3y)
\end{aligned}
$$

Portanto, o operador R é definido por:

$$
R(x,y) \ = \ \frac{1}{5}(-3x + 4y, \, 4x + 3y) \qquad \text{para todo} \qquad (x,y) \ \in \ I\!R^2 \, .
$$

Facilmente obtemos a matriz $[R]^{\beta}_{\beta}$, onde β é a base canônica do $I\!R^2$, que é dada por:

$$
[R]^{\beta}_{\beta} \ = \ \frac{1}{5} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}
$$

Exemplo 5.16.5 *Considere o espaço vetorial* $I\!R^2$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* S *o subespaço de* $I\!R^2$ *dado por:* $S = \{ (x,y) \in I\!R^2 \ / \ y - ax = 0 \}$ *para* $a \in I\!R$. *Seja* $R : I\!R^2 \longrightarrow I\!R^2$ *o operador de reflexão sobre* S, *paralelamente ao subespaço* $S^{\perp}$. *Determine a matriz do operador* R *com relação à base canônica do* $I\!R^2$.

Resposta: $\boxed{[R]^{\beta}_{\beta} \ = \ \dfrac{1}{1 + a^2} \begin{bmatrix} 1 - a^2 & 2a \\ 2a & a^2 - 1 \end{bmatrix}}$

Utilizando o resultado do Teorema 5.12.2 e o resultado do Teorema 5.14.3, temos que a matriz $[R]^{\beta}_{\beta}$ é simétrica e ortogonal. O que podemos verificar facilmente para a matriz $[R]^{\beta}_{\beta}$ do Exemplo 5.16.4 e para a matriz $[R]^{\beta}_{\beta}$ do Exemplo 5.16.5.

Exercícios

Exercício 5.70 *Considere o espaço vetorial* $I\!R^3$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Seja* S *o subespaço de* $I\!R^3$ *dado por:* $S = [(3,2,1)]$ *e* $R : I\!R^3 \longrightarrow I\!R^3$ *o operador de reflexão sobre* S. *Determine o operador* $R(x,y,z)$.

Exercício 5.71 *Seja* $R : I\!R^3 \longrightarrow I\!R^3$ *o operador linear tal que* $R(x,y,z)$ *é a reflexão do elemento* $(x,y,z) \in I\!R^3$ *sobre o plano* $3x + 2y + z = 0$. *Pede–se:*

(a) Determine o operador $R(x,y,z)$.

(b) Determine o núcleo do operador R.

(c) Determine a imagem do operador R.

Exercício 5.72 *Considere o espaço vetorial real* $\mathcal{P}_2(I\!R)$ *com o produto interno*

$$\langle p , q \rangle = \int_0^1 p(x)q(x)dx \qquad \text{para todo} \qquad p,\, q \in \mathcal{P}_3(I\!R)\,.$$

Determine a reflexão do elemento $q(x) = x^2$ *sobre o subespaço* $\mathcal{P}_1(I\!R)$.

Resposta: $\boxed{R(q)(x) = -\frac{1}{3} + 2x - x^2}$

5.17 Melhor Aproximação em Subespaços

Teorema 5.17.1 (Melhor Aproximação) *Considere V um espaço vetorial com o produto interno $\langle \cdot , \cdot \rangle$ e S um subespaço de dimensão finita de V. Se $u \in V$, então a projeção ortogonal $\widetilde{u} \in S$ é a **melhor aproximação** do elemento u no subespaço S com relação à norma proveniente do produto interno, isto é,*

$$\| \, u \, - \, \widetilde{u} \, \|_2 \;\leq\; \| \, u \, - \, v \, \|_2 \qquad \textit{para todo} \qquad v \in S \, .$$

Demonstração – Pelo Teorema 5.9.1, sabemos que $u = \widetilde{u} - w$, onde $\widetilde{u} \in S$ e $w \in S^{\perp}$. Desse modo, para todo $v \in S$, temos que

$$u - v = (\, u - \widetilde{u} \,) + (\, \widetilde{u} - v \,)$$

desde que $(\, \widetilde{u} - v \,) \in S$ e $(\, u - \widetilde{u} \,) \in S^{\perp}$, que é a decomposição ortogonal do elemento $u - v$. Pelo Teorema de Pitágoras 5.5.3, temos que

$$\| \, u - v \, \|_2^2 = \| \, u - \widetilde{u} \, \|_2^2 + \| \, \widetilde{u} - v \, \|_2^2 \quad \Longrightarrow \quad \| \, u - \widetilde{u} \, \|_2^2 \leq \| \, u - v \, \|_2^2$$

Portanto

$$\| \, u - \widetilde{u} \, \|_2 \;\leq\; \| \, u - v \, \|_2 \qquad \text{para todo} \qquad v \in S \, ,$$

o que completa da demonstração. ∎

Exemplo 5.17.1 *Considere o seguinte espaço vetorial real*

$$\mathcal{J}([0,\infty)) = \left\{ f : [0, +\infty) \longrightarrow I\!R \;\; \textit{função contínua} \;\; / \;\; \int_0^{\infty} \exp(-x) \, f^2(x) dx \, < \, +\infty \right\} .$$

Definimos em $\mathcal{J}([0,\infty))$ o seguinte produto interno

$$\langle \, f \, , \, g \, \rangle \;=\; \int_0^{\infty} \exp(-x) \, f(x) g(x) dx \, .$$

Dada a função $f(x) = \exp(-x)$, determine o polinômio $p(x) = a + b\,x$, $a, b \in I\!R$, que melhor aproxima a função f com relação à norma Euclidiana.

Resposta: $\boxed{p(x) \;=\; \dfrac{3}{4} - \dfrac{1}{4}x \quad , \quad x \geq 0}$

Exemplo 5.17.2 *Considere o espaço vetorial* $\mathcal{C}([-1,1])$ *munido do produto interno usual. Determine o polinômio* $p(x) = a + bx$, $a, b \in \mathbb{R}$, *mais próximo da função* $f(x) = \exp(x)$, $x \in [-1,1]$, *com relação à norma Euclidiana. Dê uma interpretação geométrica para o polinômio* $p(x)$.

Resposta: $\boxed{p(x) \;=\; \frac{1}{2}(\, e \,-\, e^{-1}\,) \,+\, \frac{3}{e}x \quad , \quad x \in [-1,1]}$

Exemplo 5.17.3 *Seja* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$. *Dados os elementos* $u\,,\, v \in V$ *ortogonais e um elemento qualquer* $w \in V$. *Considere a seguinte função*

$$J(\alpha,\beta) \;=\; \|\, w \,-\, (\, \alpha u \,+\, \beta v\,)\,\|_2^2 \qquad ; \qquad (\,\alpha\,,\,\beta\,) \in \mathbb{R}^2\,,$$

onde $\|\cdot\|_2$ *é a norma Euclidiana. Pede–se:*

(a) Determine o único ponto crítico, $(\,\alpha^*\,,\,\beta^*\,)$, *da função* J.

(b) Dê uma interpretação geométrica para o elemento $w^* \;=\; \alpha^* u \,+\, \beta^* v$.

(c) Classifique o ponto crítico $(\,\alpha^*\,,\,\beta^*\,)$.

Vamos escrever a função J da seguinte forma:

$$\begin{aligned}
J(\alpha,\beta) &= \langle\, w - (\,\alpha u + \beta v\,)\,,\, w - (\,\alpha u + \beta v\,)\,\rangle \\[2ex]
&= \langle\, w\,,\, w\,\rangle \,-\, 2\,\langle\, w\,,\, \alpha u + \beta v\,\rangle \,+\, \langle\, \alpha u + \beta v\,,\, \alpha u + \beta v\,\rangle \\[2ex]
&= \langle\, w\,,\, w\,\rangle \,-\, 2\alpha\,\langle\, w\,,\, u\,\rangle \,-\, 2\beta\,\langle\, w\,,\, v\,\rangle \,+\, \alpha^2\,\langle\, u\,,\, u\,\rangle \,+\, \beta^2\,\langle\, v\,,\, v\,\rangle
\end{aligned}$$

Como queremos encontrar os pontos críticos da função J, vamos calcular o seu gradiente

$$\nabla J(\alpha,\beta) \;=\; \begin{bmatrix} -2\,\langle\, w\,,\, u\,\rangle \,+\, 2\alpha\,\langle\, u\,,\, u\,\rangle \\[2ex] -2\,\langle\, w\,,\, v\,\rangle \,+\, 2\beta\,\langle\, v\,,\, v\,\rangle \end{bmatrix}$$

Finalmente, fazendo $\nabla J(\alpha,\beta) = (0,0)$ obtemos

$$\alpha^* \;=\; \frac{\langle\, w\,,\, u\,\rangle}{\langle\, u\,,\, u\,\rangle} \qquad \text{e} \qquad \beta^* \;=\; \frac{\langle\, w\,,\, v\,\rangle}{\langle\, v\,,\, v\,\rangle}$$

Portanto, temos que o elemento $w^* = \alpha^* u \,+\, \beta^* v$ é a projeção ortogonal do elemento w sobre o subespaço $S = [u,v]$. Pelo Teorema 5.17.1, podemos classificar o único ponto crítico $(\,\alpha^*\,,\,\beta^*\,)$ como sendo um ponto de mínimo da função J.

Exemplo 5.17.4 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *com o produto interno*

$$\langle\, p\,,\, q\,\rangle \;=\; \int_0^1 p(x)q(x)dx \qquad ; \qquad \forall\; p,\, q \;\in\; \mathcal{P}_2(\mathbb{R})\,.$$

Determine a melhor aproximação do polinômio $q(x) \,=\, 1 \,-\, x^2$ *no subespaço* $\mathcal{P}_1(\mathbb{R})$.

A melhor aproximação do elemento $q(x) \,=\, 1 \,-\, x^2$ no subespaço $\mathcal{P}_1(\mathbb{R}) \subset \mathcal{P}_2(\mathbb{R})$ é dada pela projeção ortogonal do elemento $q(x)$ sobre o subespaço $\mathcal{P}_1(\mathbb{R})$.

Inicialmente, vamos obter uma base ortogonal $\beta^* \,=\, \{\, q_1(x),\, q_2(x)\,\}$ para o subespaço $\mathcal{P}_1(\mathbb{R})$ a partir da base canônica $\beta \,=\, \{\, p_1(x) \,=\, 1,\, p_2(x) \,=\, x\,\}$, através do **Processo de Ortogonalização de Gram-Schmidt**.

Desse modo, escolhemos $q_1(x) \,=\, p_1(x) \,=\, 1$. Agora, vamos construir o elemento $q_2(x)$ da seguinte forma:

$$q_2(x) \;=\; p_2(x) \;-\; \alpha_{12}\, q_1(x)$$

ortogonal ao subespaço gerado pelo elemento $q_1(x)$. Assim, temos que

$$\alpha_{12} \;=\; \frac{\langle\, p_2\,,\, q_1\,\rangle}{\langle\, p_2\,,\, q_1\,\rangle} \;=\; \frac{1}{2}\,.$$

Logo, o elemento $q_2(x) \;=\; x \,-\, \dfrac{1}{2}$, completando a base ortogonal $\beta^* \,=\, \{\, q_1(x),\, q_2(x)\,\}$.

Finalmente, vamos determinar a projeção ortogonal, $\widetilde{q}(x)$, do elemento $q(x) \,=\, 1 \,-\, x^2$ no subespaço $\mathcal{P}_1(\mathbb{R})$ que é dada por:

$$\widetilde{q}(x) \;=\; \frac{\langle\, q\,,\, q_1\,\rangle}{\langle\, q_1\,,\, q_1\,\rangle}\, q_1(x) \;+\; \frac{\langle\, q\,,\, q_2\,\rangle}{\langle\, q_2\,,\, q_2\,\rangle}\, q_2(x)$$

onde

$$\langle\, q_1\,,\, q_1\,\rangle \;=\; \int_0^1 dx \;=\; 1 \qquad \text{e} \qquad \langle\, q_2\,,\, q_2\,\rangle \;=\; \int_0^1 \left(x \,-\, \frac{1}{2}\right)^2 dx \;=\; \frac{1}{12}$$

$$\langle\, q\,,\, q_1\,\rangle \;=\; \int_0^1 (\,1 \,-\, x^2\,)dx \;=\; \frac{2}{3}$$

$$\langle\, q\,,\, q_2\,\rangle \;=\; \int_0^1 (\,1 \,-\, x^2\,) \left(x \,-\, \frac{1}{2}\right) dx \;=\; -\frac{1}{12}$$

Portanto, temos que

$$\widetilde{q}(x) \;=\; \frac{2}{3} \;-\; \left(x \,-\, \frac{1}{2}\right) \;=\; \frac{7}{6} \;-\; x\,.$$

Exercícios

Exercício 5.73 *Considere o espaço vetorial real* $I\!R^4$ *munido do produto interno usual. Seja* S *o subespaço gerado pelos elementos* $q_1 = (1,1,1,1)$ *e* $q_2 = (-1,1,-1,1)$. *Encontre a melhor aproximação do elemento* $u = (2,1,3,1) \in I\!R^4$ *no subespaço* S.

Exercício 5.74 *Considere o espaço vetorial* $\mathcal{C}([-\pi,\pi])$ *com o produto interno usual*

$$\langle f \,,\, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx \qquad ; \qquad \forall \; f, g \in \mathcal{C}([-\pi,\pi])$$

e o subespaço vetorial

$$U = \{ \, f \in \mathcal{C}([-\pi,\pi]) \; / \; f(-x) = f(x) \, \} \, .$$

Seja $\beta = \{ 1, \cos(x), \cos(2x), \cdots, \cos(nx) \}$ *um conjunto ortogonal em* U. *Calcular a projeção ortogonal da função* $f(x) = |\, x \,|$, *pertencente ao subespaço* U, *sobre o subespaço de dimensão finita* S *gerado pelos elementos do conjunto ortogonal* β.

Exercício 5.75 *Considere o espaço vetorial real* $\mathcal{P}_3(I\!R)$ *com o produto interno*

$$\langle p \,,\, q \rangle = \int_{0}^{1} p(x)q(x)dx \qquad \textit{para todos} \qquad p, \, q \in \mathcal{P}_3(I\!R) \, .$$

Determine a melhor aproximação do elemento $q(x) = x^3$ *no subespaço* $\mathcal{P}_2(I\!R)$.

Exercício 5.76 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot \,,\, \cdot \rangle$, S *um subespaço de dimensão finita com* $\beta = \{ q_1, \cdots, q_n \}$ *uma base ortonormal para* S. *Dado um elemento* $u \in V$, *considere a função* $J : V \longrightarrow I\!R$ *definida por:*

$$J(v) = \| \, u - v \, \|_2^2 \qquad ; \qquad v \in S \, ,$$

onde $\| \cdot \|_2$ *é a norma Euclidiana. Mostre que o elemento* $v^* \in S$ *satisfazendo*

$$J(v^*) = \min\{ \, J(v) \quad ; \quad v \in S \, \}$$

é a ***projeção ortogonal*** *do elemento* u *sobre o subespaço* S.

Exercício 5.77 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual. Seja* S *o subespaço definido por:*

$$S = \{ \, (x,y,z) \in I\!R^3 \; / \; x + y - z = 0 \, \} \, .$$

Determine a melhor aproximação do elemento $v = (1,2,1)$ *no subespaço* $S^{\perp}$.

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] LUENBERGER, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} = Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/~pulino/ALESA/

Janeiro de 2012

Conteúdo

6

Autovalores e Autovetores

Conteúdo

6.1 Autovalor e Autovetor de um Operador Linear

Sejam V um espaço vetorial real e T um operador linear sobre V. Podemos fazer a colocação do seguinte problema:

$$\boxed{\text{Quais são os elementos } v \in V \text{ tais que } T(v) = -v\ ?}$$

Exemplo 6.1.1 *Considere o espaço vetorial real* $\mathbb{R}^2$. *O operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (-x\,,\,-y) \end{array}$$

é a reflexão em torno da origem, isto é, uma rotação de 180^o *no sentido anti–horário.*

Podemos verificar facilmente que

$$T(x,y) = (-x\,,\,-y) = -1(x\,,\,y)\,.$$

Portanto, todo elemento $v = (x,y) \in \mathbb{R}^2$ satisfaz a condição acima.

Exemplo 6.1.2 *Considere o espaço vetorial real* $\mathbb{R}^2$ *e o operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (x+2y\,,\,-y) \end{array}$$

Podemos verificar facilmente que

$$T(x,-x) = (-x\,,\,x) = -1(x\,,\,-x)\,.$$

Portanto, todo elemento $v = (x,-x) \in \mathbb{R}^2$ satisfaz a condição acima.

Sejam V um espaço vetorial real e T um operador linear sobre V. Podemos também fazer a colocação do seguinte problema:

Quais são os elementos $v \in V$, não–nulos, que são levados pelo operador T em um múltiplo de si mesmo, isto é, estamos procurando elementos $v \in V$, não–nulos, e escalares $\lambda \in \mathbb{R}$ tais que $T(v) = \lambda v$?

Definição 6.1.1 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *e* $T: V \longrightarrow V$ *um operador linear. Se existirem* $v \in V$, *diferentes do elemento neutro, e* $\lambda \in \mathbb{F}$ *tais que* $T(v) = \lambda\, v$, *então o escalar* $\lambda \in \mathbb{F}$ *é um* ***autovalor*** *de* T *e o elemento* v *é um* ***autovetor*** *de* T *associado ao autovalor* λ.

Exemplo 6.1.3 *Considere o espaço vetorial real* $I\!R^2$. *O operador linear*

$$T: \quad I\!R^2 \longrightarrow I\!R^2$$

$$(x, y) \longrightarrow T(x, y) = (y\,,\,x)$$

é a reflexão em torno da reta r *dada pela equação* $y = x$

Assim, para qualquer elemento $v = (x, y) \in r$, não nulo, temos que

$$T(x, y) = T(x, x) = 1\,(x\,,\,x)\,.$$

Portanto, qualquer elemento $v = (x, y) \in r$ não–nulo é um autovetor de T associado ao autovalor $\lambda = 1$.

De modo análogo, qualquer elemento $v = (x, y) \in s$, não nulo, onde s é a reta dada pela equação $y = -x$, é um autovetor de T associado ao autovalor $\lambda = -1$. De fato,

$$T(x, y) = T(x, -x) = (-x\,,\,x) = -1\,(x\,,\,y)\,.$$

Teorema 6.1.1 *Sejam* V *um espaço vetorial sobre o corpo* $I\!F$, T *um operador linear sobre* V *e* v *um autovetor associado ao autovalor* λ. *Então, qualquer elemento* $w = \alpha v$, *com* $\alpha \in I\!F$ *não–nulo, também é um autovetor de* T *associado a* λ.

Demonstração – Considerando que $(v\,,\,\lambda)$ é um autopar do operador linear T, isto é, $T(v) = \lambda v$, e que $w = \alpha v$, temos que

$$T(w) = T(\alpha v) = \alpha T(v) = \alpha(\lambda v) = \lambda(\alpha v) = \lambda w\,.$$

Logo, o elemento $w = \alpha v$, com $\alpha \in I\!F$ não–nulo, é um autovetor de T associado ao autovalor λ. ■

Podemos observar que o autovalor λ é unicamente determinado pelo operador T e pelo autovetor v. De fato, considere que λ e λ' são autovalores do operador T associados ao autovetor v, isto é,

$$T(v) = \lambda\,v \qquad \text{e} \qquad T(v) = \lambda'\,v\,.$$

Assim, temos que

$$\lambda\,v - \lambda'\,v = 0_V \implies (\,\lambda - \lambda'\,)\,v = 0_V \implies (\,\lambda - \lambda'\,) = 0 \implies \lambda = \lambda'\,,$$

pois $v \neq 0_V$. Assim, temos somente um autovalor λ associado ao autovetor v.

Nos casos em que o autovalor $\lambda \in I\!R$, podemos dar uma interpretação geométrica para os autovetores associados como sendo os elementos de V que tem suas direções preservadas pelo operador T.

Definição 6.1.2 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow V$ um operador linear. Fixando um autovalor λ do operador T, o subconjunto*

$$V_\lambda = \{ v \in V \ / \ T(v) = \lambda v \}$$

*é denominado **subespaço associado ao autovalor** λ.*

Podemos observar facilmente que o subconjunto V_λ é igual ao subespaço $Ker(T - \lambda I_V)$. De fato, tomando um elemento $v \in V_\lambda$ temos que

$$T(v) = \lambda v \iff (T - \lambda I_V)(v) = 0_V \iff v \in Ker(T - \lambda I_V).$$

Logo, temos que $V_\lambda = Ker(T - \lambda I_V)$. Assim, provamos que V_λ é um subespaço de V, pois sabemos que o núcleo de um operador linear é um subespaço de V.

Exemplo 6.1.4 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um operador linear sobre V, λ_1 e λ_2 autovalores distintos do operador T. Podemos verificar facilmente que $V_{\lambda_1} \cap V_{\lambda_2} = \{ 0_V \}$.*

De fato, tomando um elemento $v \in V_{\lambda_1} \cap V_{\lambda_2}$ temos que

$$T(v) = \lambda_1 v \quad \text{e} \quad T(v) = \lambda_2 v.$$

Assim, obtemos

$$\lambda_1 v - \lambda_2 v = 0_V \implies (\lambda_1 - \lambda_2)v = 0_V.$$

Como $\lambda_1 \neq \lambda_2$, temos que $v = 0_V$. Portanto, mostramos que $V_{\lambda_1} \cap V_{\lambda_2} = \{ 0_V \}$.

Exemplo 6.1.5 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um operador linear sobre V e λ um autovalor do operador T. Podemos verificar facilmente que o subespaço V_λ é invariante sob T, isto é $T(v) \in V_\lambda$ para todo $v \in V_\lambda$.*

Exemplo 6.1.6 *Considere o espaço vetorial real* $\mathbb{R}^2$*. O operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (x\,,\,-y) \end{array}$$

é a reflexão em torno do eixo–ox.

Assim, podemos observar que para os elementos do tipo $v = (0,y) \in \mathbb{R}^2$ temos que

$$T(0,y) = (0\,,\,-y) = -1(0\,,\,y)\,.$$

Portanto, os elementos $v = (0,y) \in \mathbb{R}^2$ são autovetores de T com autovalor $\lambda = -1$.

De modo análogo, temos que os elementos $v = (x,0) \in \mathbb{R}^2$ são autovetores de T associados ao autovalor $\lambda = 1$. De fato,

$$T(x,0) = (x\,,\,0) = 1(x\,,\,0)\,.$$

Exemplo 6.1.7 *Considere o espaço vetorial real* $\mathbb{R}^2$*. O operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (-x\,,\,-y) \end{array}$$

é a reflexão em torno da origem.

Assim, para qualquer elemento $v = (x,y) \in \mathbb{R}^2$, não nulo, temos que

$$T(x,y) = (-x\,,\,-y) = -1(x\,,\,y)\,.$$

Portanto, qualquer elemento $v = (x,y) \in \mathbb{R}^2$, não–nulo, é um autovetor de T associado ao único autovalor $\lambda = -1$.

Exemplo 6.1.8 *Considere o espaço vetorial real* $\mathbb{R}^2$*. O operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (x\,,\,0) \end{array}$$

é a projeção no eixo–ox, isto é, um operador de projeção de coordenadas.

Podemos verificar facilmente que qualquer elemento $v = (x,0) \in \mathbb{R}^2$, não–nulo, é um autovetor de T associado ao autovalor $\lambda = 1$. Além disso, qualquer elemento $v = (0,y) \in \mathbb{R}^2$, não–nulo, é um autovetor de T associado ao autovalor $\lambda = 0$.

Exemplo 6.1.9 *De um modo geral todo operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) \,=\, \lambda\,(x\,,\,y) \end{array}$$

com $\lambda \neq 0$, *tem* λ *como único autovalor e qualquer elemento* $v = (x,y) \in \mathbb{R}^2$*, não nulo, como autovetor associado.*

Exemplo 6.1.10 *Considere o espaço vetorial real* $\mathbb{R}^2$*. O operador linear*

$$\begin{array}{rccl} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longrightarrow & T(x,y) \,=\, (-y\,,\,x) \end{array}$$

é uma rotação de um ângulo $\theta = \dfrac{\pi}{2}$ *no sentido anti–horário.*

Note que nenhum vetor $v \in \mathbb{R}^2$ não–nulo é levado por T em um múltiplo de si mesmo. Logo, T não tem nem autovalores e nem autovetores. Este é um exemplo de que nem todo operador linear sobre um espaço vetorial real possui autovalores e autovetores. Mais a frente vamos fazer uma melhor colocação desse fato.

Exemplo 6.1.11 *Considere o espaço vetorial real* $\mathbb{R}^3$ *e* P *o operador linear sobre* $\mathbb{R}^3$ *definido da seguinte forma:* $P(x,y,z) = (x,y,0)$ *para* $(x,y,z) \in \mathbb{R}^3$*, que representa a projeção sobre o plano* xy*.*

Neste exemplo, temos que todo elemento $v = (0,0,z) \in \mathbb{R}^3$, elementos sobre o eixo–oz, é um autovetor de P associado ao autovalor $\lambda_1 = 0$. De fato, $P(0,0,z) = (0,0,0)$. Todo elemento $v = (x,y,0) \in \mathbb{R}^3$, elementos do plano xy, é um autovetor de P associado ao autovalor $\lambda_2 = 1$. De fato, $P(x,y,0) = (x,y,0)$.

Exemplo 6.1.12 *Seja* V *o espaço vetorial real das funções contínuas* f*, definidas em* (a,b)*, que possuem derivadas contínuas de todas as ordens, que denotamos por* $\mathcal{C}^\infty((a,b))$*. Considere o operador linear* D *sobre* V *definido da seguinte forma:* $D(f) = f'$*. Os autovetores do operador* D *são todas as funções contínuas não nulas* f *satisfazendo a equação da forma:* $f' = \lambda f$ *para algum* $\lambda \in \mathbb{R}$*. Assim, os autovetores são as funções* $f(x) = c\exp(\lambda x)$*, onde* $c \in \mathbb{R}$ *é uma constante não nula, associados aos autovalores* $\lambda \in \mathbb{R}$*. Note que para* $\lambda = 0$*, os autovetores associados são as funções constantes não nulas, isto é,* $f(x) = c$*, para* $c \in \mathbb{R}$ *não nula.*

Exemplo 6.1.13 *Considere o espaço vetorial real $I\!R^3$ munido do produto interno usual $\langle \cdot , \cdot \rangle$ e o subespaço $S = [(1,-1,2)]$. Seja P o operador linear sobre $I\!R^3$ onde $w = P(u)$, para $u \in I\!R^3$, é a projeção ortogonal do elemento u sobre o subespaço S. Vamos determinar os autovalores e autovetores de P.*

Fazendo $v = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, temos que $w = P(u) = \alpha^* v \in S$ com

$$\alpha^* = \frac{\langle u , v \rangle}{\langle v , v \rangle} = \frac{v^t u}{v^t v}$$

Assim, temos que w pode ser escrito da seguinte forma:

$$w = P(u) = \frac{v^t u}{v^t v} v = \frac{v v^t}{v^t v} u$$

Considerando o $I\!R^3$ com a base canônica β, temos que

$$[P]^\beta_\beta = \frac{v v^t}{v^t v} = \frac{1}{6} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix}$$

Sabemos que, para todo $z \in S$ temos $P(z) = z$. Portanto, $\lambda_1 = 1$ é um autovalor de P com $v_1 = (1,-1,2)$ o autovetor associado. Logo, o subespaço S é o subespaço associado ao autovalor $\lambda_1 = 1$.

O complemento ortogonal, $S^\perp$, do subespaço S em $I\!R^3$ é o hiperplano dado por:

$$S^\perp = H = \{ u \in I\!R^3 \ / \ \langle u , v \rangle = 0 \}$$

Note que $S^\perp$ é um plano em $I\!R^3$ dado pela equação $x - y + 2z = 0$. Temos também que, $P(u) = (0,0,0)$ para todo $u \in S^\perp$. Observamos também que $Ker(P) = S^\perp$.

Desse modo, como $P(u) = 0\,u$ para todo $u \in S^\perp$, podemos concluir que $\lambda_2 = 0$ é um autovalor de P e $S^\perp$ é o subespaço associado ao autovalor λ_2. Assim, quaisquer dois vetores v_2 e v_3 linearmente independentes em $S^\perp$ são autovetores associados ao autovalor $\lambda_2 = 0$.

Finalmente, escolhemos $v_2 = (1,1,0)$ e $v_3 = (0,2,1)$ como sendo os autovetores do operador P associados ao autovalor $\lambda_2 = 0$.

Exemplo 6.1.14 *Considere o espaço vetorial real* $\mathbb{R}^3$ *munido do produto interno usual* $\langle \cdot, \cdot \rangle$ *e o subespaço* $S = [(1,-1,2)]$. *Seja* R *o operador linear sobre* $\mathbb{R}^3$ *onde* $w = R(u)$, *para* $u \in \mathbb{R}^3$, *é a reflexão do elemento* u *em torno do subespaço* $S^\perp$. *Vamos determinar os autovalores e autovetores de* R.

Do Exemplo 6.1.13, sabemos que o operador P de projeção ortogonal sobre o subespaço S é dado por:

$$P(u) = \frac{v^t u}{v^t v} v = \frac{v v^t}{v^t v} u \qquad \text{para todo} \qquad u \in \mathbb{R}^3$$

Desse modo, o operador T de projeção ortogonal sobre o subespaço $S^\perp$ é dado por:

$$T(u) = u - P(u) = u - \frac{v^t u}{v^t v} v = \left(I - \frac{v v^t}{v^t v} \right) u$$

Temos que o operador R de reflexão em torno do subespaço $S^\perp$ é dado por:

$$R(u) = T(u) - P(u) = u - 2P(u) = = \left(I - 2\frac{v v^t}{v^t v} \right) u$$

Desse modo, temos que $R(u) = u$ para todo $u \in S^\perp$, concluindo que $\lambda_1 = 1$ é um autovalor de R e $S^\perp$ é o subespaço associado ao autovalor λ_1. Assim, quaisquer dois vetores v_1 e v_2 linearmente independentes em $S^\perp$ são autovetores associados ao autovalor $\lambda_1 = 1$.

Portanto, podemos escolher $v_1 = (1,1,0)$ e $v_2 = (0,2,1)$ como sendo os autovetores de R associados ao autovalor $\lambda_1 = 1$.

Sabemos que, para todo $w \in S$ temos $R(w) = -w$. Portanto, $\lambda_2 = -1$ é um autovalor de R com $v_3 = (1,-1,2)$ o autovetor associado. Logo, o subespaço S é o subespaço associado ao autovalor $\lambda_2 = -1$.

Note que podemos generalizar os dois últimos exemplos para o caso em que $S = [v]$, com $v \in \mathbb{R}^n$ não–nulo.

Exercícios

Exercício 6.1 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um operador linear sobre V, v um autovetor de T associado a um autovalor λ e α um escalar não–nulo. Mostre que $\alpha\lambda$ é um autovalor do operador linear αT com v o autovetor associado.*

Exercício 6.2 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Mostre que $\lambda = 0$ é um autovalor de T se, e somente se, T não é um operador injetor.*

Exercício 6.3 *Sejam V um espaço vetorial real e T um operador linear sobre V tal que $T^2 = T$, isto é, $T(T(v)) = T(v)$ para todo $v \in V$ (**operador idempotente**). Mostre que os autovalores de T são $\lambda_1 = 0$ e $\lambda_2 = 1$.*

Exercício 6.4 *Sejam V um espaço vetorial real e T um operador linear sobre V tal que $T^2 = I_V$, isto é, $T(T(v)) = v$ para todo $v \in V$ (**operador auto–reflexivo**). Mostre que os autovalores de T são $\lambda_1 = 1$ e $\lambda_2 = -1$.*

Exercício 6.5 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um operador linear sobre V e λ é um autovalor de T com v o autovetor associado. Mostre que $a\lambda + b$ é um autovalor do operador $aT + bI_V$, para $a, b \in \mathbb{F}$, com v o autovetor associado.*

Exercício 6.6 *Determine o operador linear T sobre o $\mathbb{R}^2$ satisfazendo as seguintes propriedades simultaneamente:*

(a) *$\lambda_1 = 1$ é um autovalor de T com os autovetores associados do tipo $v_1 = (y\,,\,-y)$ para $y \in \mathbb{R}$ não–nulo.*

(b) *$\lambda_2 = 3$ é um autovalor de T com os autovetores associados do tipo $v_2 = (0\,,\,y)$ para $y \in \mathbb{R}$ não nulo.*

Exercício 6.7 *Considere o espaço vetorial $\mathbb{R}^4$ munido do produto interno usual e W o subespaço vetorial gerado pelos elementos $w_1 = (1, -1, 0, 1)$ e $w_2 = (-1, 0, 1, 1)$. Sejam P o operador de projeção ortogonal sobre o subespaço W e R o operador de reflexão sobre o subespaço W. Pede–se:*

(a) *Determine os autovalores e os autovetores do operador P.*

(b) *Determine os autovalores e os autovetores do operador R.*

Exercício 6.8 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um operador linear sobre V e v um autovetor de T associado a um autovalor λ. Mostre que v é um autovetor do operador T^n associado ao autovalor λ^n para qualquer $n \in \mathbb{N}$.*

Exercício 6.9 *Sejam V um espaço vetorial real e T um operador linear sobre V de modo que existe um número inteiro n tal que $T^n = 0$, isto é, $T^n(v) = 0_V$ para todo $v \in V$ (**operador nilpotente**). Mostre que o único autovalor de T é $\lambda = 0$.*

Exercício 6.10 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um isomorfismo de V e v um autovetor de T associado a um autovalor λ. Mostre que v é um autovetor do isomorfismo inverso T^{-1} associado ao autovalor $\dfrac{1}{\lambda}$.*

Exercício 6.11 *Seja T um operador linear sobre o espaço vetorial real $\mathbb{M}_n(\mathbb{R})$ definido por: $T(A) = A^t$. Mostre que os autovalores de T são $\lambda_1 = 1$ e $\lambda_2 = -1$, descrevendo os subespaços associados a cada um dos autovalores.*

Exercício 6.12 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, T um operador linear sobre V e λ um autovalor de T. Mostre que o subconjunto definido por:*

$$V_\lambda = \{ \, v \in V \ / \ T(v) = \lambda v \, \}$$

é um subespaço vetorial de V.

Exercício 6.13 *Sejam V um espaço vetorial complexo e T um operador linear sobre V tal que $T^2 = -I_V$, isto é, $T(T(v)) = -v$ para todo $v \in V$. Mostre que T é um automorfismo de V e que os autovalores de T são $\lambda_1 = i$ e $\lambda_2 = -i$.*

Exercício 6.14 *Considere o espaço vetorial real $\mathbb{R}^4$ e o operador T sobre o $\mathbb{R}^4$ definido da seguinte forma: $T(x,y,z,t) = (-y\,,\,x\,,\,-t\,,\,z)$. Mostre que T satisfaz $T^2(v) = -v$ para todo $v = (x,y,z,t) \in \mathbb{R}^4$. Determine a matriz $[T]^\beta_\beta$, onde β é a base canônica de $\mathbb{R}^4$. O operador linear T possui autovalores e autovetores ?*

Exercício 6.15 *Considere o espaço vetorial complexo $\mathbb{C}^4$ e o operador T sobre o $\mathbb{C}^4$ definido da seguinte forma: $T(x,y,z,t) = (-t\,,\,z\,,\,-y\,,\,x)$. Mostre que T satisfaz $T^2(v) = -v$ para todo $v = (x,y,z,t) \in \mathbb{C}^4$. Determine a matriz $[T]^\beta_\beta$, onde β é a base canônica de $\mathbb{C}^4$, como espaço vetorial complexo. O operador linear T possui autovalores e autovetores ?*

6.2 Autovalor e Autovetor de uma Matriz

Sejam V um espaço vetorial de dimensão finita sobre um corpo $\mathbb{F}$, digamos que $dim(V) = n$, e T um operador linear sobre V. O problema de encontrar os autovalores do operador T será resolvido através do cálculo de determinantes. Queremos encontrar escalares $\lambda \in \mathbb{F}$ de modo que a equação $T(v) = \lambda v$ tenha solução $v \in V$, não nula. A equação $T(v) = \lambda v$ pode ser escrita na forma: $(T - \lambda I_V)(v) = 0_V$.

A equação acima terá solução v não nula se, e somente se, $Ker(T - \lambda I_V) \neq \{ 0_V \}$. Assim, se $A = [T]_\beta^\beta$ é a representação matricial do operador T, com relação a alguma base ordenada de V, então a matriz $A - \lambda I_n$ é a representação matricial para o operador $T - \lambda I_V$. Desse modo, a matriz $A - \lambda I_n$ deve ser singular, isto é, $\det(A - \lambda I_n) = 0$.

Portanto, $\lambda \in \mathbb{F}$ é um autovalor do operador T se, e somente se, satisfaz a equação

$$\det(A - \lambda I_n) = 0 .$$

Desse modo, dada uma matriz A de ordem n sobre um corpo $\mathbb{F}$, vamos definir um autovalor de A como sendo um autovalor do operador linear T_A sobre $\mathbb{F}^n$ associado à matriz A, isto é, $A = [T_A]_\beta^\beta$, onde β é a base canônica de $\mathbb{F}^n$. Portanto, os autovetores da matriz A, associados ao autovalor λ, são soluções não nulas da equação $T_A(v) = \lambda v$, representadas como matriz coluna. Assim, se $u = (x_1, \cdots, x_n) \in \mathbb{F}^n$ é um autovetor de T_A associado ao autovalor $\lambda \in \mathbb{F}$, isto é, $T_A(u) = \lambda u$, temos que

$$AX = \lambda X \quad , \quad \text{onde} \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{M}_{n \times 1}(\mathbb{F}) ,$$

isto é, (λ , X) é um autopar da matriz A. Note que $[u]_\beta = X$.

Definição 6.2.1 *Seja A uma matriz de ordem n sobre um corpo $\mathbb{F}$. Um* ***autovalor*** *da matriz A é um escalar $\lambda \in \mathbb{F}$ tal que a matriz $(A - \lambda I_n)$ seja singular.*

Equivalentemente, λ é um autovalor de A se, e somente se, $det(A - \lambda I_n) = 0$. Evidentemente, os autovalores de A são exatamente os escalares $\lambda \in \mathbb{F}$ que são raízes do polinômio $p(\lambda) = det(A - \lambda I_n)$. O polinômio $p(\lambda)$ é denominado **polinômio característico** da matriz A, que é um polinômio de grau n.

Definição 6.2.2 *Sejam $A, B \in \mathbb{M}_n(\mathbb{F})$. Dizemos que a matriz B é* ***similar*** *ou* ***semelhante*** *a matriz A, se existe uma matriz invertível $P \in \mathbb{M}_n(\mathbb{F})$ de maneira que $B = P^{-1}AP$.*

Note que matrizes similares possuem a seguinte propriedade:

$$\det(B) \;=\; \det(P^{-1}AP) \;=\; \det(P^{-1})\,\det(A)\,\det(P) \;=\; \det(A)\,.$$

Esta propriedade nos leva ao seguinte resultado, que é muito importante no estudo de autovalores.

Teorema 6.2.1 *Matrizes similares possuem o mesmo polinômio característico.*

Demonstração – Considerando que a matriz B é similar à matriz A, isto é, existe uma matriz P invertível tal que $B = P^{-1}AP$. Consideramos inicialmente o polinômio característico da matriz B, obtemos

$$\begin{aligned} p(\lambda) &= det(\,B - \lambda\,I_n\,) \\ &= det(\,P^{-1}AP - \lambda\,P^{-1}\,P\,) \\ &= det(\,P^{-1}(\,A - \lambda\,I_n\,)P\,) \\ &= det(P^{-1})\,det(\,A - \lambda\,I_n\,)\,det(P) \\ &= det(\,A - \lambda\,I_n\,)\,, \end{aligned}$$

o que completa a demonstração. ■

O Teorema 6.2.1 nos permite definir o polinômio característico do operador linear T como sendo o polinômio característico da matriz $A = [T]^{\beta}_{\beta}$, que é a representação matricial do operador T em relação a qualquer base ordenada β de V. Para isso, vamos precisar do seguinte resultado.

Teorema 6.2.2 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, T um operador linear sobre V, β e α bases ordenadas de V. Então,*

$$[T]^{\beta}_{\beta} \;=\; [I]^{\alpha}_{\beta}\,[T]^{\alpha}_{\alpha}\,[I]^{\beta}_{\alpha}\,.$$

Demonstração – Seja $P = [I]^{\beta}_{\alpha}$ a matriz mudança da base β para a base α, e lembrando que $[I]^{\alpha}_{\beta} = P^{-1}$. Inicialmente, vamos calcular

$$[T(u)]_{\alpha} \;=\; [T]^{\alpha}_{\alpha}\,[u]_{\alpha} \;=\; [T]^{\alpha}_{\alpha}\,[I]^{\beta}_{\alpha}\,[u]_{\beta} \qquad \text{para todo} \qquad u \in V\,.$$

Assim, podemos escrever $[T(u)]_{\beta}$ da seguinte forma:

$$[T(u)]_{\beta} \;=\; [I]^{\alpha}_{\beta}\,[T(u)]_{\alpha} \;=\; [I]^{\alpha}_{\beta}\,[T]^{\alpha}_{\alpha}\,[I]^{\beta}_{\alpha}\,[u]_{\beta} \qquad \Longrightarrow \qquad [T]^{\beta}_{\beta} \;=\; [I]^{\alpha}_{\beta}\,[T]^{\alpha}_{\alpha}\,[I]^{\beta}_{\alpha}\,.$$

Portanto, mostramos que $[T]^{\beta}_{\beta} = P^{-1}\,[T]^{\alpha}_{\alpha}\,P$, isto é, as matrizes $[T]^{\beta}_{\beta}$ e $[T]^{\alpha}_{\alpha}$ são similares, o que completa a demonstração. ■

Corolário 6.2.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, T um operador linear sobre V, β e α bases ordenadas de V. Então,*

$$\det([T]_\alpha^\alpha) = \det([T]_\beta^\beta) .$$

Demonstração – A prova é feita utilizando o resultado do Teorema 6.2.2. □

Definição 6.2.3 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, T um operador linear sobre V e β uma base ordenada de V. Definimos o* ***determinante*** *do operador T da seguinte forma: $\det(T) = \det([T]_\beta^\beta)$.*

Teorema 6.2.3 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Então, T é invertível se, e somente se, $\det(T) \neq 0$.*

Demonstração – A prova segue da definição de determinante e do Corolário 4.8.2. □

Exemplo 6.2.1 *Considere o espaço vetorial $\mathcal{P}_2(\mathbb{R})$ e o operador T sobre $\mathcal{P}_2(\mathbb{R})$ definido por: $T(p(x)) = p(x) + xp'(x)$. Considerando a base canônica $\beta = \{ 1, x, x^2 \}$, temos que*

$$[T]_\beta^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} .$$

Assim, $\det(T) = \det([T]_\beta^\beta) = 6$. Logo, o operador T é invertível, pois $\det(T) \neq 0$. Podemos observar facilmente que o determinante de um operador linear, assim definido, fica bem estabelecido devido ao Corolário 6.2.1.

Definição 6.2.4 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Definimos o* ***polinômio característico*** *do operador T como sendo o polinômio característico da matriz $[T]_\beta^\beta$ em relação a qualquer base ordenada β de V.*

Considerando o Exemplo 6.2.1, temos que o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (1 - \lambda)(2 - \lambda)(3 - \lambda) ,$$

com $A = [T]_\beta^\beta$, onde β é a base canônica de $\mathcal{P}_2(\mathbb{R})$.

Proposição 6.2.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, digamos $dim(v) = n$, e T um operador linear sobre V. Então, os autovalores do operador linear T são os escalares $\lambda \in \mathbb{F}$ que são raízes do polinômio característico da matriz $A = [T]_\beta^\beta$ em relação a qualquer base ordenada β de V.*

Demonstração – Por definição, um escalar $\lambda \in \mathbb{F}$ é um autovalor do operador T, se a equação $T(v) = \lambda v$ tem solução não nula. A equação $T(v) = \lambda v$ pode ser escrita na forma: $(T - \lambda I_V)(v) = 0_V$.

Assim, a equação acima terá solução não nula se, e somente se, $Ker(T - \lambda I_V) \neq \{ 0_V \}$. Desse modo, se $A = [T]_\beta^\beta$ é a representação matricial do operador linear T, com relação a alguma base ordenada β de V, então $A - \lambda I_n$ é a matriz do operador $T - \lambda I_V$. Desse modo, a matriz $A - \lambda I_n$ deve ser singular, isto é, $\det(A - \lambda I_n) = 0$.

Portanto, um escalar $\lambda \in \mathbb{F}$ é um autovalor do operador T se, e somente se, satisfaz a equação $\det(A - \lambda I_n) = 0$, o que completa a demonstração. ■

Finalmente, para determinar os autovetores do operador T associados ao autovalor λ, temos que encontrar os elementos não–nulos do núcleo do operador $T - \lambda I_V$, isto é, temos que encontrar as soluções não nulas da equação $T(v) = \lambda v$.

De uma maneira geral, podemos simplificar os cálculos para determinar os autovetores do operador T associados ao autovalor λ, fazendo a seguinte observação.

Considerando que $u \in V$ é um autovetor do operador linear T associado ao autovalor λ, isto é, $T(u) = \lambda u$, obtemos

$$[T(u)]_\beta = \lambda [u]_\beta \qquad \Longrightarrow \qquad [T]_\beta^\beta [u]_\beta = \lambda [u]_\beta .$$

Portanto, podemos observar facilmente que $[u]_\beta = X$, onde $X \in \mathbb{M}_{n\times 1}(\mathbb{F})$ é um autovetor da matriz $A = [T]_\beta^\beta$ associado ao autovalor λ.

Sabemos que os autovalores do operador T são os escalares $\lambda \in \mathbb{F}$ que são raízes do polinômio característico da matriz $A = [T]_\beta^\beta$ em relação a qualquer base ordenada β de V. Desse modo, podemos também simplificar os cálculos para encontrar os autovalores de T, escolhendo a base canônica de V para determinar a representação matricial do operador linear T.

Exemplo 6.2.2 *Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:*

$$T(p(x)) = (1 + x)p'(x) + p''(x) .$$

Determine os autovalores do operador linear T.

Temos que A $=$ $[T]_\beta^\beta$, onde $\beta = \{ 1, x, x^2 \}$ é a base canônica de $\mathcal{P}_2(\mathbb{R})$, é dada por:

$$[T]_\beta^\beta = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} .$$

Desse modo, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = -\lambda(1 - \lambda)(2 - \lambda)$$

Portanto, os autovalores de T são $\lambda_1 = 0$, $\lambda_2 = 1$ e $\lambda_3 = 2$.

Como $\lambda_1 = 0$ é um autovalor do operador linear T, podemos observar que

$$V_{\lambda_1} = Ker(T) .$$

Assim, o operador linear T não é um operador injetor.

Exemplo 6.2.3 *Considere o espaço vetorial real $\mathbb{R}^2$ com a base canônica β e T o operador linear definido por:*

$$\begin{array}{rccl} T : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x, y) & \longrightarrow & T(x, y) = (2x + 2y \, , \, y) \end{array}$$

Determine o polinômio característico do operador T.

Temos que a matriz $A = [T]_\beta^\beta$ é dada por:

$$A = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} .$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (2 - \lambda)(1 - \lambda) = \lambda^2 - 3\lambda + 2 .$$

Desse modo, temos que $\lambda_1 = 2$ e $\lambda_2 = 1$ são os autovalores do operador T.

Exemplo 6.2.4 *No Exemplo 6.2.3 considere o espaço vetorial real* $\mathbb{R}^2$ *com a base ordenada* $\gamma = \{ (1,1),\ (-1,1) \}$.

Temos que a matriz $A = [T]_\gamma^\gamma$ é dada por:

$$A = \frac{1}{2} \begin{bmatrix} 5 & 1 \\ -3 & 1 \end{bmatrix}.$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = \left(\frac{5}{2} - \lambda \right) \left(\frac{1}{2} - \lambda \right) + \frac{3}{4} = \lambda^2 - 3\lambda + 2 .$$

Assim, obtemos o resultado esperado, de acordo com o Teorema 6.2.2.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 2$, temos que determinar os elementos não–nulos $(x,y) \in \mathbb{R}^2$ tais que $T(x,y) = 2(x,y)$. Equivalentemente, temos que encontrar os elementos não nulos do núcleo do operador $(T - 2I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2\,x \\ 2\,y \end{bmatrix} \quad \Longrightarrow \quad \begin{cases} 0x + 2y = 0 \\ y = 0 \end{cases}$$

Portanto, os autovetores associados a $\lambda_1 = 2$ são do tipo $v_1 = (x,0)$, com $x \neq 0$. Desse modo, podemos escolher $v_1 = (1,0)$ o autovetor associado ao autovalor $\lambda_1 = 2$.

Para determinar os autovetores associados ao autovalor $\lambda_2 = 1$, temos que determinar os elementos não–nulos $(x,y) \in \mathbb{R}^2$ tais que $T(x,y) = (x,y)$. Equivalentemente, temos que encontrar os elementos não–nulos do núcleo do operador $(T - I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \Longrightarrow \quad x + 2y = 0$$

Portanto, os autovetores associados ao autovalor $\lambda_2 = 1$ são do tipo $v_2 = t\,(-2,1)$, para $t \in \mathbb{R}$ não–nulo. Assim, podemos escolher $v_2 = (-2,1)$ o autovetor associado ao autovalor $\lambda_2 = 1$.

Exemplo 6.2.5 *Considere a matriz $A \in I\!M_3(I\!R)$ dada por:*

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

Determine os autovalores e os autovetores da matriz A.

Seja T_A o operador linear sobre $I\!R^3$ associado a matriz A, isto é,

$$T_A(x, y, z) = (\, 2x + y + z \,,\, 2x + 3y + 4z \,,\, -x - y - 2z \,)$$

Assim, $A = [T_A]_\beta^\beta$, onde β é a base canônica do $I\!R^3$. Desse modo, os autovalores da matriz A são os autovalores do operador linear T_A, e os autovetores são os autovetores do operador T_A, representados como matriz coluna.

Temos que o polinômio característico da matriz $A = [T_A]_\beta^\beta$ é dado por:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2 - \lambda & 1 & 1 \\ 2 & 3 - \lambda & 4 \\ -1 & -1 & -2 - \lambda \end{vmatrix} = -(\lambda - 1)(\lambda + 1)(\lambda - 3)$$

Os autovalores da matriz A são $\lambda_1 = 1$, $\lambda_2 = -1$ e $\lambda_3 = 3$.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 1$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A - I)$. Assim, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \Longrightarrow \quad \left\{ \begin{array}{rcrcrcl} x & + & y & + & z & = & 0 \\ 2x & + & 2y & + & 4z & = & 0 \\ -x & - & y & - & 3z & = & 0 \end{array} \right.$$

Adicionando a primeira equação e a terceira equação encontramos $z = 0$, as duas primeiras equação ficam reduzidas a equação $x + y = 0$.

Portanto, os autovetores associados a $\lambda_1 = 1$ são do tipo $v_1 = (x, -x, 0)$, com $x \neq 0$. Assim, podemos escolher $v_1 = (1, -1, 0)$ o autovetor associado ao autovalor $\lambda_1 = 1$. De modo análogo, obtemos os autovetores associados ao autovalor $\lambda_2 = -1$ que são do tipo $v_2 = t\,(0, 1, -1)$, e os autovetores associados ao autovalor $\lambda_3 = 3$ que são do tipo $v_3 = t\,(2, 3, -1)$, para $t \in I\!R$ não–nulo.

Finalmente, os autovetores da matriz A são representados da seguinte forma:

$$X_1 = \begin{bmatrix} x \\ -x \\ 0 \end{bmatrix}$$

para $x \in I\!R$ não–nulo, são os autovetores associados ao autovalor $\lambda_1 = 1$. Desse modo, temos que $AX_1 = \lambda_1 X_1$.

$$X_2 = \begin{bmatrix} 0 \\ y \\ -y \end{bmatrix}$$

para $y \in I\!R$ não–nulo, são os autovetores associados ao autovalor $\lambda_2 = -1$. Desse modo, temos que $AX_2 = \lambda_2 X_2$.

$$X_3 = \begin{bmatrix} -2z \\ -3z \\ z \end{bmatrix}$$

para $z \in I\!R$ não–nulo, são os autovetores associados ao autovalor $\lambda_3 = 3$. Desse modo, temos que $AX_3 = \lambda_3 X_3$.

Portanto, podemos escolher os seguintes autovetores para a matriz A

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \quad \text{e} \quad X_3 = \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 1$, $\lambda_2 = -1$ e $\lambda_3 = 3$, respectivamente.

É importante observar que como A é uma matriz quadrada de ordem 3, seus autovetores são matrizes coluna de ordem 3×1.

Sejam V é um espaço vetorial complexo de dimensão finita, digamos $dim(V) = n$, e T um operador linear sobre V. Então, o polinômio característico do operador linear T é um polinômio complexo que possui n raízes em $\mathbb{C}$, levando em conta a multiplicidade, veja **Teorema Fundamental da Álgebra**. Neste caso, um operador linear T tem n autovalores. Entretanto, se V é um espaço vetorial real o número de autovalores do operador T é menor ou igual à dimensão de V. Para ilustrar este fato vamos considerar os seguintes exemplos.

Exemplo 6.2.6 *Seja* $\mathbb{C}^2$ *um espaço vetorial complexo com a base* $\beta = \{ (1,0), (0,1) \}$. *O operador linear*

$$\begin{array}{rccl} T: & \mathbb{C}^2 & \longrightarrow & \mathbb{C}^2 \\ & (x,y) & \longrightarrow & T(x,y) = (-y\,,\,x) \end{array}$$

é uma rotação de um ângulo $\theta = \dfrac{\pi}{2}$ *no sentido anti–horário.*

Temos que o polinômio característico da matriz $[T]_{\beta}^{\beta}$ é dado por $p(\lambda) = \lambda^2 + 1$ para $\lambda \in \mathbb{C}$. Assim, o operador T possui os autovalores $\lambda_1 = -i$ e $\lambda_2 = i$.

Desse modo, o autovetor $v_1 \in \mathbb{C}^2$ associado ao autovalor $\lambda_1 = -i$ é a solução do seguinte sistema linear homogêneo

$$\begin{cases} i\,x - y = 0 \\ x + i\,y = 0 \end{cases} \iff i\,x - y = 0 \implies y = i\,x$$

Portanto, todo elemento $v_1 = (a, i\,a) \in \mathbb{C}^2$, para $a \in \mathbb{C}$ não–nulo, é um autovetor do operador T associado ao autovalor $\lambda_1 = -i$. Assim, podemos escolher $v_1 = (1, i)$ o autovetor associado ao autovalor $\lambda_1 = -i$.

De modo análogo, o autovetor $v_2 \in \mathbb{C}^2$ associado ao autovalor $\lambda_2 = i$ é a solução do seguinte sistema linear homogêneo

$$\begin{cases} -i\,x - y = 0 \\ x - i\,y = 0 \end{cases} \iff i\,x + y = 0 \implies y = -i\,x$$

Portanto, todo elemento $v_2 = (a, -i\,a) \in \mathbb{C}^2$, para $a \in \mathbb{C}$ não–nulo, é um autovetor do operador T associado ao autovalor $\lambda_2 = i$. Assim, podemos escolher $v_2 = (1, -i)$ o autovetor associado ao autovalor $\lambda_2 = i$.

É importante observar que, neste caso, não temos a interpretação geométrica para o autovetor como sendo o elemento que tem sua direção preservada pelo operador T.

Exemplo 6.2.7 *Considere o operador linear T sobre o $I\!R^3$ definido por:*

$$\begin{aligned} T: \quad I\!R^3 &\longrightarrow I\!R^3 \\ (x,y,z) &\longrightarrow T(x,y,z) = (\, x \,,\, -z \,,\, y \,) \end{aligned}$$

que representa uma rotação de um ângulo $\theta = \dfrac{\pi}{2}$ no sentido anti–horário no plano yz.

A matriz $A = [T]^{\beta}_{\beta}$, onde β é a base canônica do $I\!R^3$ é dada por:

$$A = [T]^{\beta}_{\beta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Assim, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(\, A - \lambda\, I \,) = (1 - \lambda)(1 + \lambda^2)\,,$$

que possui as seguintes raízes $\lambda_1 = 1$, $\lambda_2 = i$ e $\lambda_3 = -i$.

Desse modo, como estamos considerando o operador linear T sobre o espaço vetorial real $I\!R^3$, temos que $\lambda_1 = 1$ é o único autovalor de T.

Podemos verificar facilmente que os autovetores associados ao autovalor $\lambda_1 = 1$ são do tipo $v = (x, 0, 0)$ para $x \in I\!R$ não–nulo. Assim, podemos escolher o autovetor $v_1 = (1, 0, 0)$ associado ao autovalor $\lambda_1 = 1$.

Entretanto, considerando o operador linear T sobre o espaço vetorial complexo $\mathbb{C}^3$, temos que $\lambda_1 = 1$, $\lambda_2 = i$ e $\lambda_3 = -i$ são os autovalores de T.

Neste caso, podemos verificar facilmente que $v_2 = (0, 1, -i)$ é um autovetor associado ao autovalor $\lambda_2 = i$. De modo análogo, temos que $v_3 = (0, 1, i)$ é um autovetor associado ao autovalor $\lambda_2 = -i$.

Exemplo 6.2.8 *Considere o operador linear* T *sobre* $\mathcal{P}_2(\mathbb{R})$ *definido por:*

$$T(p(x)) = p(0) + p(1)(x + x^2) .$$

Determine os autovalores e os autovetores do operador T*.*

Vamos determinar a matriz $[T]_\beta^\beta$, onde $\beta = \{ 1, x, x^2 \}$ é a base canônica de $\mathcal{P}_2(\mathbb{R})$. Desse modo, temos que

$$T(1) = 1 + x + x^2 \quad , \quad T(x) = x + x^2 \quad \text{e} \quad T(x^2) = x + x^2 .$$

Logo, a matriz $A = [T]_\beta^\beta$ é dada por:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} .$$

Assim, o polinômio característico do operador linear T é dado por:

$$p(\lambda) = \det(A - \lambda I) = \lambda(\lambda - 1)(\lambda - 2) .$$

Portanto, os autovalores do operador linear T são $\lambda_1 = 2$, $\lambda_2 = 1$ e $\lambda_3 = 0$, que são os autovalores da matriz A.

Podemos verificar facilmente que os autovetores da matriz A são

$$X_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \quad \text{e} \quad X_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 2$, $\lambda_2 = 1$ e $\lambda_3 = 0$, respectivamente.

Portanto, sabemos que

$$[p_1(x)]_\beta = X_1 \quad , \quad [p_2(x)]_\beta = X_2 \quad \text{e} \quad [p_3(x)]_\beta = X_3 ,$$

onde $p_1(x)$, $p_2(x)$ e $p_3(x)$ são os autovetores do operador linear T associados aos autovalores $\lambda_1 = 2$, $\lambda_2 = 1$ e $\lambda_3 = 0$, respectivamente. Logo, obtemos

$$p_1(x) = x + x^2 \quad , \quad p_2(x) = 1 - x - x^2 \quad \text{e} \quad p_3(x) = x - x^2 .$$

Podemos observar facilmente que

$$V_{\lambda_1} = [x + x^2] \quad , \quad V_{\lambda_2} = [1 - x - x^2] \quad \text{e} \quad V_{\lambda_3} = [x - x^2] .$$

Exercícios

Exercício 6.16 *Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:*

$$T(p(x)) = p(x) + (x+1)p'(x) .$$

Determine os autovalores e os autovetores do operador T.

Exercício 6.17 *Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:*

$$T(p(x)) = xp'(x) + p''(x) .$$

Determine os autovalores e os autovetores do operador T.

Exercício 6.18 *Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:*

$$T(p(x)) = p(x) + xp'(x) .$$

Determine os autovalores e os autovetores do operador T.

Exercício 6.19 *Considere o operador linear T sobre $\mathcal{P}_3(\mathbb{R})$ definido por:*

$$T(p(x)) = p(x) + xp''(x) .$$

Determine os autovalores e os autovetores do operador T. O operador linear T é um automorfismo de $\mathcal{P}_3(\mathbb{R})$?

Exercício 6.20 *Considere o operador T sobre $\mathbb{R}^3$ definido por:*

$$T(x,y,z) = (x + 2y - z\,,\, 3y + z\,,\, 4z) .$$

Determine os autovalores e os autovetores do operador T.

Exercício 6.21 *Considere o operador linear T sobre $\mathbb{R}^3$ definido por:*

$$T(x,y,z) = (x + y\,,\, x - y + 2z\,,\, 2x + y - z) .$$

Determine os autovalores e os autovetores do operador T.

Exercício 6.22 *Considere o operador linear T sobre $\mathbb{M}_2(\mathbb{R})$ definido por:*

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a+b & 2b \\ 2c & 3d \end{bmatrix} .$$

Determine os autovalores e os autovetores do operador T.

Exercício 6.23 *Sejam V um espaço vetorial real de dimensão finita e T um operador linear sobre V definido em uma base ordenada $\gamma = \{ v_1 , v_2 , \cdots , v_n \}$ de V da seguinte forma: $T(v_i) = \lambda_i v_i$, $\lambda_i \in I\!R$. Determine o polinômio característico e os autovalores do operador linear T.*

Exercício 6.24 *Seja $D \in I\!M_n(I\!R)$ uma matriz Diagonal. Mostre que D possui um conjunto de n autovetores linearmente independentes.*

Exercício 6.25 *Seja $A \in I\!M_n(I\!R)$. Mostre que as matrizes A e A^t possuem os mesmos autovalores.* ***Sugestão:*** *utilize o polinômio característico.*

Exercício 6.26 *Considere o operador linear T sobre o $I\!R^4$ cuja matriz em relação à base canônica β de $I\!R^4$ é dada por:*

$$[T]^{\beta}_{\beta} = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Determine os autovalores e os autovetores do operador linear T. O operador linear T é um automorfismo de $I\!R^4$?

Exercício 6.27 *Sejam $A \in I\!M_n(I\!R)$ uma matriz invertível e λ um autovalor de A. Mostre que $\dfrac{1}{\lambda}$ é um autovalor de A^{-1}.* ***Sugestão:*** *utilize o polinômio característico.*

Exercício 6.28 *Determine os autovalores da matriz A dada por:*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Exercício 6.29 *Sejam T um operador linear sobre $I\!R^3$, $\gamma = \{ v_1, v_2, v_3 \}$ uma base ordenada para o espaço vetorial real $I\!R^3$ e o subespaço $S = [v_1, v_3]$. Sabendo que $T(v) = v$ para todo $v \in S$ e $T(v_2) = v_1 + 2v_2 + 3v_3$. Determine os autovalores e os autovetores do operador linear T.*

Exercício 6.30 *Mostre que se λ é um autovalor de uma matriz A com X o autovetor associado, então $\alpha\lambda + \beta$ é um autovalor da matriz $\alpha A + \beta I_n$ com X o autovetor associado.*

Exercício 6.31 *Seja V o subespaço vetorial de $\mathbb{M}_2(\mathbb{R})$ das matrizes triangulares superiores. Pede–se:*

(a) *Exiba uma base ordenada para V.*

(b) *Seja $T : V \longrightarrow V$ o operador linear definido por:*

$$T\left(\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}\right) = \begin{bmatrix} a+b & b \\ 0 & c-a-b \end{bmatrix}.$$

Mostre que T é um automorfismo de V.

(c) *Determine os autovalores e os autovetores de T.*

Exercício 6.32 *Considere o operador linear T sobre $\mathcal{P}_3(\mathbb{R})$ definido por:*

$$T(p(x)) = p(x) + p'(x) + x^2 p''(x) .$$

Determine os autovalores e os autovetores do operador linear T, descrevendo para cada autovalor o subespaço associado.

Exercício 6.33 *Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:*

$$T(a + bx + cx^2) = (2b + c) + (2b - c)x + 2cx^2 .$$

Determine os autovalores e os autovetores do operador linear T, descrevendo para cada autovalor o subespaço associado.

Exercício 6.34 *Seja A uma matriz de ordem n triangular superior (inferior) ou uma matriz diagonal. Mostre que os autovalores de A são os elementos da diagonal principal da matriz A.*

Exercício 6.35 *Seja λ um autovalor de A com X o autovetor associado. Mostre que λ^n é um autovalor de A^n com X o autovetor associado, para $n \in \mathbb{N}$.*

Exercício 6.36 *Sejam A uma matriz invertível e λ um autovalor de A com X o autovetor associado. Mostre que $\dfrac{1}{\lambda}$ é um autovalor da matriz A^{-1} com X o autovetor associado.*

Exercício 6.37 *Determine os autovalores e autovetores das matrizes A e A^{-1}, onde*

$$A = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}.$$

Exercício 6.38 *Considere a matriz diagonal em blocos* $T \in \mathbb{M}_4(\mathbb{R})$ *dada por:*

$$T = \begin{bmatrix} U & 0_2 \\ 0_2 & D \end{bmatrix},$$

onde $U \in \mathbb{M}_2(\mathbb{R})$ *é uma matriz triangular superior e* $D \in \mathbb{M}_2(\mathbb{R})$ *é uma matriz diagonal, representadas por:*

$$U = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \quad e \quad D = \begin{bmatrix} d & 0 \\ 0 & e \end{bmatrix},$$

com $a, b, c, d, e \in R$. *Mostre que*

$$\lambda_1 = a \quad , \quad \lambda_2 = c \quad , \quad \lambda_3 = d \quad e \quad \lambda_4 = e$$

são os autovalores da matriz T, *com autovetores associados*

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} \dfrac{b}{c-a} \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad X_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad e \quad X_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

respectivamente, para $c \neq a$. *Considerando* $a = c$, *determine os autovalores e os autovetores da matriz* T.

Exercício 6.39 *Determine os autovalores e os autovetores da matriz* T *dada por:*

$$T = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

Exercício 6.40 *Determine os autovalores e os autovetores da matriz* T *dada por:*

$$T = \begin{bmatrix} 6 & 2 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

Exercício 6.41 *Sejam* $A, B \in \mathbb{M}_n(\mathbb{R})$ *matrizes similares, isto é, existe uma matriz invertível* $P \in \mathbb{M}_n(\mathbb{R})$ *tal que* $B = P^{-1}AP$. *Mostre que se* A *é invertível, então* B *é invertível e as matrizes* A^{-1} *e* B^{-1} *são similares.*

6.3 Multiplicidade Algébrica e Geométrica

Definição 6.3.1 *Definimos a* ***multiplicidade algébrica*** *de um* ***autovalor*** λ *como sendo a quantidade de vezes que ele aparece como raiz do polinômio característico.*

Definição 6.3.2 *Definimos a* ***multiplicidade geométrica*** *de um* ***autovalor*** λ *como sendo a dimensão do subespaço* V_λ *associado ao autovalor* λ.

Exemplo 6.3.1 *Considere a matriz* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}.$$

Determine os autovalores e os autovetores da matriz A.

Seja T_A o operador linear sobre $I\!R^3$ associado a matriz A, isto é, , isto é,

$$T_A(x,y,z) = (\, 2x - y + z \,,\, 3y - z \,,\, 2x + y + 3z \,) .$$

Assim, $A = [T_A]^\beta_\beta$, onde β é a base canônica do $I\!R^3$. Desse modo, os autovalores da matriz A são os autovalores do operador linear T_A, e os autovetores são os autovetores do operador T_A, representados como matriz coluna.

Temos que o polinômio característico da matriz $A = [T_A]^\beta_\beta$ é dado por:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -1 & 1 \\ 0 & 3-\lambda & -1 \\ 2 & 1 & 3-\lambda \end{vmatrix} = -(\lambda - 2)(\lambda - 2)(\lambda - 4)$$

Os autovalores do operador T_A são $\lambda_1 = 2$ com multiplicidade algébrica igual a 2, e $\lambda_2 = 4$ com multiplicidade algébrica igual a 1.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 2$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A - 2I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} \qquad \Longrightarrow \qquad \left\{ \begin{array}{rcrcrcl} & - & y & + & z & = & 0 \\ & & y & - & z & = & 0 \\ 2x & + & y & + & z & = & 0 \end{array} \right.$$

Assim, obtemos a solução $z = y = -x$. Portanto, os autovetores associados ao autovalor $\lambda_1 = 2$ são do tipo $v_1 = (x, -x, -x)$, com $x \neq 0$. Desse modo, o autovalor $\lambda_1 = 2$ tem multiplicidade geométrica igual a 1. De modo análogo, obtemos que os autovetores associados ao autovalor $\lambda_2 = 4$ são do tipo $v_2 = (x, -x, x)$, com $x \neq 0$. Note que o autovalor λ_2 tem multiplicidade geométrica igual a 1.

Finalmente, os autovetores da matriz A são representados da seguinte forma:

$$X_1 = \begin{bmatrix} x \\ -x \\ -x \end{bmatrix}$$

para $x \in I\!R$ não–nulo, são os autovetores associados ao autovalor $\lambda_1 = 2$, que possui multiplicidade algébrica igual a 2 e multiplicidade geométrica igual a 1. Desse modo, temos que $AX_1 = \lambda_1 X_1$.

$$X_2 = \begin{bmatrix} x \\ -x \\ x \end{bmatrix}$$

para $x \in I\!R$ não–nulo, são os autovetores associados ao autovalor $\lambda_2 = 4$, que possui multiplicidade algébrica igual a 1 e multiplicidade geométrica igual a 1. Desse modo, temos que $AX_2 = \lambda_2 X_2$.

Portanto, podemos escolher os seguintes autovetores para a matriz A

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \qquad \text{e} \qquad X_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 2$ e $\lambda_2 = 4$, respectivamente.

Exemplo 6.3.2 *Considere a matriz* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} .$$

Determine os autovalores e os autovetores da matriz A.

Seja T_A o operador linear sobre $I\!R^3$ associado a matriz A, isto é, , isto é,

$$T_A(x, y, z) = (\, 2x + y + z \,,\, 2x + 3y + 2z \,,\, 3x + 3y + 4z \,) .$$

Assim, $A = [T_A]_\beta^\beta$, onde β é a base canônica do $I\!R^3$. Desse modo, os autovalores da matriz A são os autovalores do operador linear T_A, e os autovetores são os autovetores do operador T_A, representados como matriz coluna.

Temos que o polinômio característico da matriz $A = [T_A]_\beta^\beta$ é dado por:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2 - \lambda & 1 & 1 \\ 2 & 3 - \lambda & 2 \\ 3 & 3 & 4 - \lambda \end{vmatrix} = -(\lambda - 1)(\lambda - 1)(\lambda - 7)$$

Os autovalores do operador T_A são $\lambda_1 = 7$ com multiplicidade algébrica igual a 1, e $\lambda_2 = 1$ com multiplicidade algébrica igual a 2.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 7$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A - 7I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7x \\ 7y \\ 7z \end{bmatrix} \quad \Longrightarrow \quad \begin{cases} 5x - y - z = 0 \\ -2x + 4y - 2z = 0 \\ -3x - 3y + 3z = 0 \end{cases}$$

Assim, obtemos a solução $y = 2x$ e $z = 3x$. Portanto, os autovetores associados ao autovalor $\lambda_1 = 7$ são do tipo $v_1 = t\,(1, 2, 3)$, com $t \neq 0$. Desse modo, o autovalor $\lambda_1 = 7$ tem multiplicidade geométrica igual a 1. De maneira análoga, obtemos que os autovetores associados ao autovalor $\lambda_2 = 1$ são do tipo $v_2 = a\,(1, 0, -1) + b\,(0, 1, -1)$, com $a, b \neq 0$. Note que o autovalor λ_2 tem multiplicidade geométrica igual a 2.

Finalmente, os autovetores da matriz A são representados da seguinte forma:

$$X_1 = \begin{bmatrix} x \\ 2x \\ 3x \end{bmatrix}$$

para $x \in I\!R$ não–nulo, são os autovetores associados ao autovalor $\lambda_1 = 7$, que possui multiplicidade algébrica igual a 1 e multiplicidade geométrica igual a 1. Desse modo, temos que $AX_1 = \lambda_1 X_1$.

$$X_2 = \begin{bmatrix} x \\ 0 \\ -x \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ -y \end{bmatrix}$$

para $x\,,\,y \in I\!R$ não–nulos, são os autovetores associados ao autovalor $\lambda_2 = 1$, que possui multiplicidade algébrica igual a 2 e multiplicidade geométrica igual a 2. Desse modo, temos que $AX_2 = \lambda_2 X_2$.

Portanto, podemos escolher os seguintes autovetores para a matriz A

$$X_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \text{e} \quad X_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

onde X_1 é o autovetor associado ao autovalor $\lambda_1 = 7$, X_2 e X_3 são os autovetores associados ao autovalor $\lambda_2 = 1$.

Exercícios

Exercício 6.42 *Considere a matriz A dada por:*

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Determine a multiplicidade algébrica e a multiplicidade geométrica dos autovalores da matriz A.

Exercício 6.43 *Sejam $A,\ B \in \mathbb{M}_n(\mathbb{R})$ matrizes similares, isto é, existe uma matriz invertível $P \in \mathbb{M}_n(\mathbb{R})$ tal que $A = P^{-1} B P$. Estabeleça a relação entre os autovalores e autovetores das matrizes A e B.*

Exercício 6.44 *Sejam A e B matrizes invertíveis de mesma ordem. Mostre que as matrizes AB^{-1} e $B^{-1}A$ possuem os mesmos autovalores.*

Exercício 6.45 *Sejam A e B matrizes invertíveis de mesma ordem. Pede–se:*

(a) Mostre que as matrizes AB e BA possuem os mesmos autovalores.

(b) Mostre que se λ é um autovalor da matriz AB com X o autovetor associado, então λ é um autovalor da matriz BA com BX autovetor associado.

(c) Mostre que se λ é um autovalor da matriz BA com Y o autovetor associado, então λ é um autovalor da matriz BA com AY o autovetor associado.

Exercício 6.46 *Seja A uma matriz de ordem n. Mostre que a transformação de similaridade preserva tanto a multiplicidade algébrica quanto a multiplicidade geométrica dos autovalores da matriz A.*

Exercício 6.47 *Sejam A, B, C matrizes quadradas de mesma ordem. Mostre que a transformação de similaridade é uma relação de equivalência, isto é,*

(a) A é similar a A.

(b) Se A é similar a B, então B é similar a A.

(c) Se A é similar a B e B é similar a C, então A é similar a C.

6.4 Matrizes Especiais

Com o objetivo de simplificar a notação e facilitar as demonstrações que apresentamos nesta seção, sempre que necessário, vamos considerar os elementos do espaço vetorial real $\mathbb{R}^n$ representados na forma de matriz coluna, elementos do espaço vetorial real $\mathbb{M}_{n\times 1}(\mathbb{R})$, tendo em vista que os espaços vetoriais reais $\mathbb{R}^n$ e $\mathbb{M}_{n\times 1}(\mathbb{R})$ são isomorfos. Sabemos também que os espaços vetoriais complexos $\mathbb{C}^n$ e $\mathbb{M}_{n\times 1}(\mathbb{C})$ são isomorfos.

Um conceito que utilizaremos com muita freqüência é o de transposta Hermitiana de uma matriz $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{C})$, que denotamos por A^*, que é definido da forma: $A^* = [\overline{a}_{ji}]$. Como ilustração, considere a matriz $A \in \mathbb{M}_2(\mathbb{C})$ e sua respectiva transposta Hermitiana $A^* \in \mathbb{M}_2(\mathbb{C})$, dadas por:

$$A = \begin{bmatrix} 2-3i & 1+i \\ 4i & 1+2i \end{bmatrix} \quad \text{e} \quad A^* = \begin{bmatrix} 2+3i & -4i \\ 1-i & 1-2i \end{bmatrix}.$$

De mesmo modo, esse conceito é aplicado aos elementos de $\mathbb{M}_{n\times 1}(\mathbb{C})$, resultando em um elemento de $\mathbb{M}_{1\times n}(\mathbb{C})$.

Considere o espaço vetorial $\mathbb{R}^n$ com a base canônica $\beta = \{ e_1, \cdots , e_n \}$. Sabemos que todo elemento $x = (x_1, \cdots , x_n) \in \mathbb{R}^n$ é escrito de modo único da forma:

$$x = \sum_{i=1}^{n} x_i \, e_i \,,$$

que vamos representar pela matriz coluna

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{M}_{n\times 1}(\mathbb{R}) .$$

Assim, o produto interno usual do $\mathbb{R}^n$, que denotamos por $\langle \cdot , \cdot \rangle$, pode ser escrito como:

$$\langle x , y \rangle = \sum_{i=1}^{n} x_i \, y_i = Y^t X$$

para todos $x, y \in \mathbb{R}^n$.

De modo análogo, considere o espaço vetorial complexo $\mathbb{C}^n$ com a base canônica. Desse modo, podemos escrever o produto interno usual da seguinte forma:

$$\langle x , y \rangle = \sum_{i=1}^{n} x_i \, \overline{y}_i = Y^* X$$

para todos $x, y \in \mathbb{C}^n$.

Teorema 6.4.1 *Seja* $A \in I\!M_n(I\!R)$. *Então, para todos* $x, y \in I\!R^n$ *temos que*

$$\langle Ax\,,\,y\rangle = \langle x\,,\,A^t y\rangle\,.$$

Demonstração – Para $x, y \in I\!R^n$, representados na forma de matriz coluna, tem–se

$$\langle Ax\,,\,y\rangle = y^t A x = (A^t y)^t x = \langle x\,,\,A^t y\rangle\,,$$

o que completa a demonstração. ■

Corolário 6.4.1 *Seja* $A \in I\!M_n(I\!R)$ *simétrica. Então, para todos* $x, y \in I\!R^n$ *tem–se*

$$\langle Ax\,,\,y\rangle = \langle x\,,\,Ay\rangle\,.$$

Demonstração – A prova é imediata utilizando o resultado do Teorema 6.4.1. □

Teorema 6.4.2 *Seja* $A \in I\!M_n(\mathbb{C})$. *Então, para todos* $x, y \in \mathbb{C}^n$ *temos que*

$$\langle Ax\,,\,y\rangle = \langle x\,,\,A^* y\rangle\,.$$

Demonstração – Para $x, y \in \mathbb{C}^n$, representados na forma de matriz coluna, tem–se

$$\langle Ax\,,\,y\rangle = y^* A x = (A^* y)^* x = \langle x\,,\,A^* y\rangle\,,$$

o que completa a demonstração. ■

Corolário 6.4.2 *Seja* $A \in I\!M_n(\mathbb{C})$ *Hermitiana. Então, para todos* $x, y \in \mathbb{C}^n$ *tem–se*

$$\langle Ax\,,\,y\rangle = \langle x\,,\,Ay\rangle\,.$$

Demonstração – A prova é imediata utilizando o resultado do Teorema 6.4.2. □

Os resultados do Teorema 6.4.1 e do Teorema 6.4.2, e os respectivos corolários, serão muito utilizados nas demonstrações que se seguem.

Teorema 6.4.3 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana. Então, seus autovalores são todos reais. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – Considere o espaço vetorial complexo $\mathbb{C}^n$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Seja T o operador linear sobre $\mathbb{C}^n$ associado a matriz A, isto é,

$$T_A(v) \;=\; Av$$

para todo $v \in \mathbb{C}^n$, representado na forma de matriz coluna.

Como a matriz $A = [T_A]^{\beta}_{\beta}$, onde β é a base canônica do $\mathbb{C}^n$, temos que um autovalor de A é um autovalor do operador T_A.

Como, por hipótese, a matriz $A = [T_A]^{\beta}_{\beta}$ é Hermitiana, sabemos que o operador T_A é Hermitiano.

Tomando λ um autovalor de T_A e v o autovetor associado, isto é, $T_A(v) = \lambda v$, obtemos

$$\lambda\langle v , v \rangle \;=\; \langle T_A(v) , v \rangle \;=\; \langle v , T_A(v) \rangle \;=\; \langle v , \lambda v \rangle \;=\; \overline{\lambda}\langle v , v \rangle \,.$$

Desse modo, obtemos a equação

$$(\lambda \,-\, \overline{\lambda})\langle v , v \rangle \;=\; 0 \,.$$

Como v é não–nulo, temos que $\lambda - \overline{\lambda} = 0$. Portanto, o autovalor λ é real, completando a demonstração da primeira parte.

Para a prova da segunda parte, sejam λ_1 e λ_2 autovalores distintos de T_A, com v_1 e v_2 os autovetores associados, respectivamente. Desse modo, temos que

$$\lambda_1\langle v_1 , v_2 \rangle \;=\; \langle T_A(v_1) , v_2 \rangle \;=\; \langle v_1 , T_A(v_2) \rangle \;=\; \langle v_1 , \lambda_2 v_2 \rangle \;=\; \lambda_2\langle v_1 , v_2 \rangle \,.$$

Desse modo, obtemos a equação

$$(\lambda_1 \,-\, \lambda_2)\langle v_1 , v_2 \rangle \;=\; 0 \,.$$

Portanto, tem–se

$$\langle v_1 , v_2 \rangle \;=\; 0 \,,$$

pois os autovalores λ_1 e λ_2 são distintos, o que completa a demonstração. ■

Corolário 6.4.3 *Seja $A \in \mathbb{M}_n(\mathbb{C})$ uma matriz simétrica. Então, seus autovalores são todos reais. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – A prova segue do Teorema 6.4.3, considerando que a matriz simétrica real é um caso particular de uma matriz Hermitiana. □

Exemplo 6.4.1 *Considere a matriz simétrica A dada por:*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Determine os autovalores e os autovetores de A.

Seja T_A o operador linear sobre $\mathbb{R}^2$ associado a matriz A. Sabemos que $A = [T_A]_\beta^\beta$, onde β é a base canônica do $\mathbb{R}^2$. Desse modo, os autovalores da matriz A são os autovalores do operador linear T_A, e os autovetores são os autovetores do operador T_A, representados na forma de matriz coluna.

O polinômio característico da matriz $A = [T_A]_\beta^\beta$ é dado por:

$$p(\lambda) = \det(A - \lambda I) = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3.$$

Portanto, os autovalores do operador T_A são $\lambda_1 = 3$ e $\lambda_2 = -1$.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 3$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A - 3I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix} \iff -x + y = 0.$$

Portanto, os autovetores associados a $\lambda_1 = 3$ são do tipo $v_1 = (x, x)$, com $x \neq 0$. Desse modo, podemos escolher $v_1 = (1, 1)$ o autovetor associado ao autovalor $\lambda_1 = 3$, do operador linear T_A.

Assim, podemos escolher

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

um autovetor da matriz A associado ao autovalor $\lambda_1 = 3$, isto é, $AX_1 = \lambda_1 X_1$.

De modo análogo, para determinar os autovetores associados ao autovalor $\lambda_2 = -1$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A + I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix} \iff x + y = 0 .$$

Portanto, os autovetores associados ao autovalor $\lambda_2 = -1$ são do tipo $v_2 = (x, -x)$, para $x \in \mathbb{R}$ não nulo. Assim, podemos escolher $v_2 = (1, -1)$ o autovetor associado ao autovalor $\lambda_2 = -1$, do operador linear T_A.

Assim, podemos escolher

$$X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

um autovetor da matriz A associado ao autovalor $\lambda_2 = -1$, isto é, $AX_2 = \lambda_2 X_2$. Note que os autovetores X_1 e X_2 são ortogonais.

Definição 6.4.1 *Seja $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz simétrica. Dizemos que A é uma matriz **positiva–definida** se*

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_j x_i = x^t A x = \langle Ax , x \rangle > 0$$

para todo $x \in \mathbb{R}^n$ não–nulo, representado na forma de matriz coluna.

No caso em que existe um elemento não–nulo $x \in \mathbb{R}^n$ tal que $x^t A x = 0$, dizemos que A é uma matriz **semipositiva–definida**.

Exemplo 6.4.2 *Considere a matriz simétrica $A \in \mathbb{M}_2(\mathbb{R})$ dada por:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} .$$

Mostre que a matriz A é positiva–definida.

Fazendo uso da Definição 6.4.1, temos que

$$\langle Ax , x \rangle = x^t A x = 2x_1^2 + 2x_2^2 + 2x_1 x_2 = x_1^2 + x_2^2 + (x_1 + x_2)^2 > 0$$

para todo $x = (x_1 , x_2) \in \mathbb{R}^2$ não–nulo, representado na forma de matriz coluna. Desse modo, mostramos que a matriz A é positiva–definida.

Definição 6.4.2 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana. Dizemos que* A *é uma matriz* ***positiva–definida*** *se*

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_j \overline{x}_i = x^* A x = \langle Ax, x \rangle > 0$$

para todo $x \in \mathbb{C}^n$ *não–nulo, representado na forma de matriz coluna.*

No caso em que existe um elemento não–nulo $x \in \mathbb{C}^n$ tal que $x^* A x = 0$, dizemos que A é uma matriz **semipositiva–definida**.

Teorema 6.4.4 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *positiva–definida. Então, seus autovalores são todos positivos. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – Como A é uma matriz Hermitiana, do Teorema 6.4.3, sabemos que seus autovalores são reais. Tomando λ uma autovalor da matriz A com v o autovetor associado, isto é, $Av = \lambda v$, e utilizando a hipótese que U é uma matriz positiva–definida, temos que

$$\langle Av, v \rangle = \lambda \langle v, v \rangle > 0 .$$

Como $\langle v, v \rangle > 0$, pois v é não-nulo, provamos que o autovalor $\lambda > 0$.

Como A é uma matriz Hermitiana, do Teorema 6.4.3, sabemos que autovetores associados a autovalores distintos são ortogonais, o que completa a demonstração. ■

Corolário 6.4.4 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *positiva–definida. Então, seus autovalores são todos positivos. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – A prova segue do Teorema 6.4.4, e do fato que a matriz simétrica real é um caso particular de uma matriz Hermitiana. □

Exemplo 6.4.3 *Considere a matriz positiva–definida* A *dada por:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} .$$

Determine os autovalores e os autovetores de A.

Seja T_A o operador linear sobre $\mathbb{R}^2$ associado a matriz A. Sabemos que $A = [T_A]_\beta^\beta$, onde β é a base canônica do $\mathbb{R}^2$. Desse modo, os autovalores da matriz A são os autovalores do operador linear T_A, e os autovetores são os autovetores do operador T_A, representados na forma de matriz coluna.

O polinômio característico da matriz $A = [T_A]_\beta^\beta$ é dado por:

$$p(\lambda) = \det(A - \lambda I) = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3.$$

Portanto, os autovalores do operador T_A são $\lambda_1 = 3$ e $\lambda_2 = 1$.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 3$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A - 3I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix} \iff -x + y = 0.$$

Portanto, os autovetores associados a $\lambda_1 = 3$ são do tipo $v_1 = (x, x)$, com $x \neq 0$. Desse modo, podemos escolher $v_1 = (1, 1)$ o autovetor associado ao autovalor $\lambda_1 = 3$, do operador linear T_A.

De modo análogo, para determinar os autovetores associados ao autovalor $\lambda_2 = 1$, temos que encontrar os elementos não–nulos do núcleo do operador $(T_A - I)$. Desse modo, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \iff x + y = 0.$$

Portanto, os autovetores associados ao autovalor $\lambda_2 = 1$ são do tipo $v_2 = (x, -x)$, para $x \in I\!R$ não nulo. Assim, podemos escolher $v_2 = (1, -1)$ o autovetor associado ao autovalor $\lambda_2 = 1$, do operador linear T_A.

Desse modo, temos que os autovetores da matriz A são dados por:

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{e} \quad X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 3$ e $\lambda_2 = 1$, respectivamente. Note que os autovetores X_1 e X_2 são ortogonais.

Na seção 6.7 vamos provar o resultado abaixo, que é muito importante na teoria de autovalores e autovetores, e nas suas aplicações, que é a caracterização de uma matriz positiva–definida.

Teorema 6.4.5 *Seja* $A \in I\!M_n(\mathbb{C})$ *uma matriz Hermitiana. Então,* A *é uma matriz positiva–definida se, e somente se, seus autovalores são todos positivos.*

Corolário 6.4.5 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz simétrica. Então,* A *é uma matriz positiva–definida se, e somente se, seus autovalores são todos positivos.*

Exemplo 6.4.4 *Fazendo uso do Corolário 6.4.5, podemos verificar que a matriz simétrica do Exemplo 6.4.1 não é positiva–definida, pois seus autovalores são* $\lambda_1 = 3$ *e* $\lambda_2 = -1$.

Definição 6.4.3 *Dizemos que* $U \in I\!M_n(\mathbb{C})$ *é uma* ***matriz unitária*** *se* $U^*U = I$. *Assim, temos que* $UU^* = I$. *Desse modo, tem–se que* $U^{-1} = U^*$.

Teorema 6.4.6 *Seja* $Q \in I\!M_n(I\!R)$ *uma matriz ortogonal. Então,* $\det(Q) = \pm 1$.

Demonstração – A prova segue da utilização da definição de matriz ortogonal e das propriedades de determinante de uma matriz. Como Q é ortogonal, tem–se que

$$\det(Q^tQ) = \det(I) = 1 .$$

Desse modo, obtemos

$$\det(Q^tQ) = \det(Q^t)\det(Q) = (\det(Q))^2 = 1 .$$

Portanto, mostramos que $\det(Q) = \pm 1$. ■

Teorema 6.4.7 *Seja* $Q \in I\!M_n(I\!R)$ *ortogonal. Então, para todos* $x, y \in I\!R^n$ *temos que*

1. $\langle Qx , Qy \rangle = \langle x , y \rangle$.
2. $\| Qx \|_2 = \| x \|_2$.

Demonstração – A prova do primeiro item segue do Teorema 6.4.1 e da definição de matriz ortogonal. De fato,

$$\langle Qx , Qy \rangle = \langle x , Q^tQy \rangle = \langle x , y \rangle .$$

A prova do segundo item segue de imediato do primeiro item a da definição de norma Euclidiana, o que completa a demonstração. ■

Teorema 6.4.8 *Sejam* $U \in \mathbb{M}_n(\mathbb{C})$ *unitária e* λ *um autovalor. Então,* $|\lambda| = 1$.

Demonstração – Considere o espaço vetorial complexo $\mathbb{C}^n$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Seja T o operador linear sobre $\mathbb{C}^n$ associado a matriz U, isto é, $T_U(v) = Uv$ para $v \in \mathbb{C}^n$, na forma de matriz coluna. Como $U = [T_U]_\beta^\beta$, onde β é a base canônica do $\mathbb{C}^n$, temos que um autovalor de U é um autovalor do operador T_U. Como $A = [T_U]_\beta^\beta$ é unitária, temos que o operador T_U é unitário.

Assim, tomando λ um autovalor de T_U e v o autovetor associado, isto é, $T_U(v) = \lambda v$, obtemos

$$|\lambda|^2 \langle v , v \rangle = \lambda\overline{\lambda} \langle v , v \rangle = \langle T_U(v) , T_U(v) \rangle = \langle Uv , Uv \rangle = \langle v , v \rangle .$$

Portanto, temos que $(1 - |\lambda|^2)\langle v , v \rangle = 0$. Como $v \in \mathbb{C}^n$ é não–nulo, obtemos $|\lambda| = 1$, o que completa a demonstração. ■

Corolário 6.4.6 *Sejam* $Q \in \mathbb{M}_n(\mathbb{R})$ *ortogonal e* λ *um autovalor. Então,* $|\lambda| = 1$.

Demonstração – A prova segue do Teorema 6.4.8, considerando que a matriz ortogonal é um caso particular de uma matriz unitária. □

Exemplo 6.4.5 *A matriz* $Q \in \mathbb{M}_2(\mathbb{R})$ *que representa uma rotação de um ângulo* θ *no sentido anti–horário*

$$Q = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

é uma matriz ortogonal. Determine os autovalores da matriz Q*, em função do ângulo* θ.

Podemos verificar facilmente que o polinômio característico da matriz Q é dado por:

$$p(\lambda) = \lambda^2 - 2\cos(\theta)\,\lambda + 1 .$$

Portanto, os autovalores da matriz Q, em função do ângulo θ, são dados por:

$$\lambda(\theta) = \cos(\theta) \pm \sqrt{\cos^2(\theta) - 1} = \cos(\theta) \pm \left(\sqrt{|\cos^2(\theta) - 1|}\right) i .$$

Note que $|\lambda(\theta)| = 1$, para todo $\theta \in \mathbb{R}$.

Definição 6.4.4 *As matrizes* $A, B \in \mathbb{M}_n(\mathbb{R})$ *são* ***ortogonalmente similares*** *se existe uma matriz ortogonal* $Q \in \mathbb{M}_n(\mathbb{R})$ *tal que* $B = Q^t A Q$.

Exemplo 6.4.6 *A matriz* $U \in \mathbb{M}_2(\mathbb{C})$ *dada por:*

$$U = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} .$$

é uma matriz unitária.

Definição 6.4.5 *As matrizes* $A, B \in \mathbb{M}_n(\mathbb{C})$ *são* ***unitáriamente similares*** *se existe uma matriz unitária* $U \in \mathbb{M}_n(\mathbb{C})$ *tal que* $B = U^* A U$.

Teorema 6.4.9 *Seja* $U \in \mathbb{M}_n(\mathbb{C})$ *é uma matriz unitária. Então,* $|\det(U)| = 1$.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 6.4.10 *Seja* $U \in \mathbb{M}_n(\mathbb{C})$ *uma matriz unitária. Então, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – Sejam λ_1 e λ_2 autovalores distintos de U com v_1 e v_2 os autovetores associados, respectivamente. Tomando a hipótese que U é uma matriz unitária, temos que

$$\lambda_1 \overline{\lambda}_2 \langle v_1 , v_2 \rangle = \langle \lambda_1 v_1 , \lambda_2 v_2 \rangle = \langle U v_1 , U v_2 \rangle = \langle v_1 , v_2 \rangle .$$

Desse modo, obtemos a equação

$$(1 - \lambda_1 \overline{\lambda}_2) \langle v_1 , v_2 \rangle = 0 .$$

Como os autovalores λ_1 e λ_2 são distintos, temos que $1 - \lambda_1 \overline{\lambda}_2 \neq 0$.

Portanto, obtemos

$$\langle v_1 , v_2 \rangle = 0 ,$$

mostrando que v_1 e v_2 são ortogonais, o que completa a demonstração. ■

Exemplo 6.4.7 *Considere a matriz unitária* $U \in \mathbb{M}_2(\mathbb{C})$ *dada por:*

$$U = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}.$$

Determine os autovalores e autovetores da matriz U.

O polinômio característico da matriz U é dado por: $p(\lambda) = \lambda^2 - 2\lambda + 2$. Assim, os autovalores da matriz U são dados por:

$$\lambda_1 = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i \qquad \text{e} \qquad \lambda_2 = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\,.$$

Temos que os autovetores associados aos autovalores λ_1 e λ_2 são dados por:

$$v_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{e} \qquad v_2 = \frac{\sqrt{2}}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

respectivamente.

Recordamos que $A \in \mathbb{M}_n(\mathbb{R})$ é uma matriz **idempotente** se $A^2 = A$, isto é,

$$A(Ax) = Ax$$

para todo $x \in \mathbb{R}^n$, representado na forma de matriz coluna.

Teorema 6.4.11 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz idempotente. Então, seus autovalores são* $\lambda_1 = 1$ *e* $\lambda_2 = 0$.

Demonstração – Tomando λ um autovalor de A e v o autovetor associado, isto é, $Av = \lambda v$, temos que

$$A(Av) = A(\lambda v) \quad \Longleftrightarrow \quad \lambda v = \lambda^2 v \quad \Longleftrightarrow \quad \lambda(1 - \lambda)v = 0\,.$$

Como v é não–nulo, obtemos a equação

$$\lambda(1 - \lambda) = 0\,,$$

que tem como soluções $\lambda_1 = 1$ e $\lambda_2 = 0$, o que completa a demonstração. ■

Recordamos que $A \in \mathbb{M}_n(\mathbb{R})$ é uma matriz **auto–reflexiva** se $A^2 = I$, isto é,

$$A(Ax) = x$$

para todo $x \in \mathbb{R}^n$, representado na forma de matriz coluna.

Teorema 6.4.12 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz auto–reflexiva. Então, seus autovalores são* $\lambda_1 = 1$ *e* $\lambda_2 = -1$.

Demonstração – Tomando λ um autovalor de A e v o autovetor associado, isto é, $Av = \lambda v$, temos que

$$A(Av) = A(\lambda v) \iff v = \lambda^2 v \iff (1 - \lambda^2)v = 0 .$$

Como v é não–nulo, obtemos a equação

$$1 - \lambda^2 = 0 ,$$

que tem como soluções $\lambda_1 = 1$ e $\lambda_2 = -1$, o que completa a demonstração. ■

Exemplo 6.4.8 *Considere a matriz auto–reflexiva* $A \in I\!M_2(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} .$$

Determine os autovalores da matriz A.

O polinômio característico da matriz A é dado por:

$$p(\lambda) = \lambda^2 - 1 ,$$

que possui como raízes $\lambda_1 = 1$ e $\lambda_2 = -1$, que são os autovalores da matriz A.

Exemplo 6.4.9 *Considere a matriz idempotente* $A \in I\!M_2(I\!R)$ *dada por:*

$$A = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} .$$

Determine os autovalores da matriz A.

O polinômio característico da matriz A é dado por:

$$p(\lambda) = \left(\frac{1}{2} - \lambda\right)^2 - \frac{1}{4} = \lambda^2 - \lambda = \lambda(1 - \lambda) ,$$

que possui como raízes $\lambda_1 = 1$ e $\lambda_2 = 0$, que são os autovalores da matriz A.

6.5 Aplicação. Classificação de Pontos Críticos

Exemplo 6.5.1 *Considere a função* $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$*, com derivadas contínuas de segunda ordem, definida da seguinte forma:*

$$F(x,y) \;=\; 6\,x \;-\; 4\,y \;-\; x^2 \;-\; 2\,y^2 \qquad , \qquad (x,y) \;\in\; \mathbb{R}^2 \;.$$

Pede–se:

1. *Determinar os pontos críticos da função* F*.*
2. *Classificar os pontos críticos.*
3. *Fazer um esboço do gráfico da função* F*.*

Por simplicidade, dado um elemento $(x,y) \in \mathbb{R}^2$, vamos representa–lo na forma de vetor coluna

$$X \;=\; \begin{bmatrix} x \\ y \end{bmatrix} \;.$$

Sabemos que os pontos críticos de F são os ponto $(\overline{x},\overline{y}) \in \Omega$ que satisfazem a equação

$$\nabla F(x,y) \;=\; 0$$

Para fazer a classificação dos pontos críticos, vamos fazer uso da *Fórmula de Taylor de Segunda–Ordem* da função F numa vizinhança do ponto crítico $\overline{X}$ dada por:

$$\begin{aligned} F(X) \;&=\; F(\overline{X}) \;+\; Y^t\,\nabla F(\overline{X}) \;+\; \frac{1}{2}\,Y^t\,H(\overline{X})\,Y \\ &=\; F(\overline{X}) \;+\; \langle\, \nabla F(\overline{X})\,,\, Y\,\rangle \;+\; \frac{1}{2}\,\langle\, H(\overline{X})\,Y\,,\, Y\,\rangle \\ &=\; F(\overline{X}) \;+\; \frac{1}{2}\,\langle\, H(\overline{X})\,Y\,,\, Y\,\rangle \end{aligned}$$

para todo $X \in B_r(\overline{X})$, onde $Y = X - \overline{X}$, $\nabla F(X)$ é o gradiente da função F e $H(X)$ é a matriz **Hessiana** da função F.

Portanto, se a matriz Hessiana $H(\overline{x},\overline{y})$ for positiva–definida, isto é, seus autovalores são positivos, temos que o ponto crítico $(\overline{x},\overline{y})$ é um ponto de mínimo relativo. No caso em que a matriz $H(\overline{x},\overline{y})$ for negativa–definida, isto é, seus autovalores são negativos, temos que o ponto crítico $(\overline{x},\overline{y})$ é um ponto de máximo relativo. No caso em que a matriz Hessiana $H(\overline{x},\overline{y})$ possuir autovalores negativo e positivo, isto é, $H(\overline{x},\overline{y})$ é uma **matriz indefinida**, temos que o ponto crítico $(\overline{x},\overline{y})$ é um ponto de sela.

Voltando ao exemplo, temos que

$$\nabla F(x,y) = \begin{pmatrix} 6 & - & 2\,x \\ -4 & - & 4\,y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Temos um ponto crítico $(\overline{x},\overline{y}) = (3,-1)$. Vamos calcular a matriz Hessiana $H(x,y)$

$$H(x,y) = \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix}$$

que não depende do ponto crítico $(\overline{x},\overline{y})$, e possui os autovalores $\lambda_1 = -2$ e $\lambda_2 = -4$. Logo, o ponto crítico $(\overline{x},\overline{y}) = (3,-1)$ é um ponto de **máximo global** da função F.

Exemplo 6.5.2 *Considere a função $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$, com derivadas contínuas de segunda ordem, definida da seguinte forma:*

$$F(x,y) = y^2 - x^2 \qquad , \qquad (x,y) \in \mathbb{R}^2 \,.$$

Pede–se:

1. *Determinar os pontos críticos da função F.*

2. *Classificar os pontos críticos.*

3. *Fazer um esboço do gráfico da função F.*

Neste caso, temos que

$$\nabla F(x,y) = \begin{pmatrix} -2\,x \\ 2\,y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Temos um ponto crítico $(\overline{x},\overline{y}) = (0,0)$. Vamos calcular a matriz Hessiana $H(x,y)$

$$H(x,y) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$

que não depende do ponto crítico $(\overline{x},\overline{y})$, e possui os autovalores $\lambda_1 = -2$ e $\lambda_2 = 2$.

Temos que os autovetores associados ao autovalor $\lambda_2 = 2$ são do tipo $v_2 = (0,y)$ com $y \neq 0$. Assim, tem–se que

$$\langle\, H(\overline{x},\overline{y})\, v_2 \,,\, v_2 \,\rangle = \lambda_2 \,\|\, v_2 \,\|_2^2 > 0 \qquad \text{para todo} \qquad v_2 \neq 0 \,.$$

Temos que os autovetores associados ao autovalor $\lambda_1 = -2$ são do tipo $v_1 = (x, 0)$ com $x \neq 0$. Assim, tem–se que

$$\langle\, H(\overline{x}, \overline{y})\, v_1 \,,\, v_1 \,\rangle \;=\; \lambda_1 \,\|\, v_1 \,\|_2^2 \;<\; 0 \qquad \text{para todo} \qquad v_1 \neq 0 \,.$$

Logo, o ponto crítico $(\overline{x}, \overline{y}) = (0, 0)$ é um ponto de **sela** da função F.

Exemplo 6.5.3 *Considere a função* $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$*, com derivadas contínuas de segunda ordem, definida da seguinte forma:*

$$F(x, y) \;=\; 2\,x^4 \;+\; y^2 \;-\; x^2 \;-\; 2\,y \qquad , \qquad (x, y) \in \mathbb{R}^2 \,.$$

Pede–se:

1. *Determinar os pontos críticos da função* F.
2. *Classificar os pontos críticos.*
3. *Fazer um esboço do gráfico da função* F.

Neste caso, temos que

$$\nabla F(x, y) \;=\; \begin{pmatrix} 8\,x^3 \;-\; 2\,x \\ 2\,y \;-\; 2 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Temos os seguintes pontos críticos $(0, 1)$, $(0.5, 1)$ e $(-0.5, 1)$. Vamos calcular a matriz Hessiana $H(\overline{x}, \overline{y})$ para o ponto crítico $(\overline{x}, \overline{y}) = (0, 1)$

$$H(x, y) \;=\; \begin{pmatrix} -2 \;+\; 24\,x^2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \Longrightarrow \qquad H(\overline{x}, \overline{y}) \;=\; \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$$

que possui os autovalores $\lambda_1 = -2$ e $\lambda_2 = 2$. Logo, o ponto crítico $(\overline{x}, \overline{y}) = (0, 1)$ é um ponto de **sela** da função F.

Agora, vamos calcular a matriz Hessiana $H(\overline{x}, \overline{y})$ para o ponto crítico $(\overline{x}, \overline{y}) = (0.5, 1)$

$$H(\overline{x}, \overline{y}) \;=\; \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

que possui os autovalores $\lambda_1 = 4$ e $\lambda_2 = 2$. Logo, o ponto crítico $(\overline{x}, \overline{y}) = (0.5, 1)$ é um ponto de **mínimo relativo** da função F. De modo análogo, temos que o ponto crítico $(\overline{x}, \overline{y}) = (-0.5, 1)$, também é um ponto de **mínimo relativo** da função F.

Exercícios

Exercício 6.48 *Determine quais das seguintes classes de matrizes são matrizes normais:*

(a) Matrizes Hermitianas.

(b) Matrizes anti–simétricas.

(c) Matrizes anti–Hermitianas.

(d) Matrizes unitárias.

(e) Matrizes simétricas.

(f) Matrizes ortogonais.

Exercício 6.49 *Mostre que toda matriz ortogonal em* $I\!M_2(I\!R)$ *pode ser escrita na forma:*

$$Q = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \qquad ou \qquad Q = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

com $a^2 + b^2 = 1$ *para* $a, b \in I\!R$.

Exercício 6.50 *Mostre que toda matriz unitária no espaço vetorial complexo* $I\!M_2(\mathbb{C})$ *pode ser escrita da seguinte forma:*

$$U = \begin{bmatrix} a & b \\ -e^{i\theta}\,\overline{b} & e^{i\theta}\,\overline{a} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}$$

onde $\theta \in I\!R$ *e* $|\,a\,|^2 + |\,b\,|^2 = 1$ *para* $a, b \in \mathbb{C}$.

Exercício 6.51 *Sejam* $Q \in I\!M_n(I\!R)$ *ortogonal e* λ *um autovalor real. Então,* $\lambda = 1$ *ou* $\lambda = -1$.

Exercício 6.52 *Sejam* $Q \in I\!M_n(I\!R)$ *uma matriz ortogonal e* λ *um autovalor complexo de* Q. *Então,* $\overline{\lambda}$ *é também um autovalor de* Q.

Exercício 6.53 *Seja* $Q \in I\!M_n(I\!R)$ *uma matriz ortogonal com* n *ímpar. Então,* Q *possui pelo menos um autovalor real.*

Exercício 6.54 *Seja* $A \in I\!M_n(I\!R)$ *idempotente. Se* $B \in I\!M_n(I\!R)$ *é similar a matriz* A, *então* B *é uma matriz idempotente.*

Exercício 6.55 *Seja* $A \in I\!M_n(I\!R)$ *auto–reflexiva. Se* $B \in I\!M_n(I\!R)$ *é similar a matriz* A*, então* B *é uma matriz auto–reflexiva.*

Exercício 6.56 *Seja* $A \in I\!M_n(I\!R)$ *simétrica. Se* $B \in I\!M_n(I\!R)$ *é ortogonalmente similar a matriz* A*, então* B *é simétrica.*

Exercício 6.57 *Seja* $A \in I\!M_n(\mathbb{C})$ *Hermitiana. Se* $B \in I\!M_n(I\!R)$ *é unitáriamente similar a matriz* A*, então* B *é Hermitiana.*

Exercício 6.58 *Sejam* $A \in I\!M_n(\mathbb{C})$ *anti–Hermitiana e* $B \in I\!M_n(I\!R)$ *unitáriamente similar a matriz* A*. Então,* B *é anti–Hermitiana.*

Exercício 6.59 *Sejam* $A \in I\!M_n(I\!R)$ *positiva–definida e* $B \in I\!M_n(I\!R)$ *ortogonalmente similar a* A*. Então,* B *é positiva–definida.*

Exercício 6.60 *Mostre que se* A *é positiva–definida, então* A *é uma matriz invertível e* A^{-1} *também é uma matriz positiva–definida.*

Exercício 6.61 *Mostre que se* A *é semipositiva–definida, então* A *é uma matriz singular e* $\lambda = 0$ *é um autovalor de* A*.*

Exercício 6.62 *Classifique os pontos críticos da função* $F : I\!R^2 \longrightarrow I\!R$ *dada por:*

$$F(x,y) \;=\; 2x^2 \;-\; xy \;-\; 3y^2 \;-\; 3x \;+\; 7y\,.$$

Exercício 6.63 *Determine todos os valores extremos, absoluto e relativo, e os pontos de sela da função* $F(x,y) \;=\; xy(\,1 \;-\; x^2 \;-\; y^2\,)$ *no quadrado definido por:*

$$Q \;=\; \{\,(x,y) \in I\!R^2 \;/\; 0 \;\leq\; x \;\leq\; 1 \quad e \quad 0 \;\leq\; y \;\leq\; 1\,\}\,.$$

Exercício 6.64 *Classifique os pontos críticos da função* $F : I\!R^2 \longrightarrow I\!R$ *dada por:*

$$F(x,y) \;=\; x^3 \;+\; y^3 \;-\; 3xy\,.$$

6.6 Diagonalização de Operadores Lineares

Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Nosso objetivo é determinar sob que condições V possui uma base ordenada com relação a qual a matriz do operador T seja uma matriz diagonal, que é a forma mais simples de se representar um operador linear. A solução para o problema de diagonalização de operadores lineares também nos leva naturalmente ao conceito de autovalores e autovetores do operador T.

Teorema 6.6.1 *Sejam $A \in \mathbb{M}_n(\mathbb{F})$, γ uma base ordenada para o espaço vetorial $\mathbb{F}^n$ e T_A o operador linear sobre $\mathbb{F}^n$ associado a matriz A. Então,*

$$[T_A]^{\gamma}_{\gamma} = P^{-1} A P ,$$

onde P é a matriz de mudança da base γ para a base canônica β de $\mathbb{F}^n$.

Demonstração – Como β é a base canônica para $\mathbb{F}^n$, temos que $A = [T_A]^{\beta}_{\beta}$.

Como $P = [I]^{\gamma}_{\beta}$, pelo Teorema 6.2.2, temos que

$$[T_A]^{\gamma}_{\gamma} = (\,[I]^{\gamma}_{\beta}\,)^{-1}\,[T_A]^{\beta}_{\beta}\,[I]^{\gamma}_{\beta} = P^{-1} A P ,$$

o que completa a demonstração. ■

Exemplo 6.6.1 *Sejam $\gamma = \{\,(1,2),\,(1,1)\,\}$ uma base ordenada para $\mathbb{R}^2$ e a matriz $A \in \mathbb{M}_2(\mathbb{R})$ dada por:*

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix} ,$$

para uma ilustração do Teorema 6.6.1.

Assim, temos que

$$P = [I]^{\gamma}_{\beta} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{e} \quad P^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} .$$

Desse modo, obtemos

$$[T_A]^{\gamma}_{\gamma} = P^{-1} A P = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -3 & -3 \\ 8 & 7 \end{bmatrix} ,$$

o que completa a ilustração do Teorema 6.6.1.

Teorema 6.6.2 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, β uma base ordenada para V, T um operador linear sobre V e B uma matriz similar a matriz $[T]_\beta^\beta$. Então, existe uma base ordenada γ para V tal que $B = [T]_\gamma^\gamma$.*

Demonstração – Considerando que a matriz B é similar a matriz $[T]_\beta^\beta$. Então, existe uma matriz $P = [p_{ij}]$ invertível tal que $B = P^{-1}\,[T]_\beta^\beta\,P$.

Seja $\beta = \{ v_1, \cdots, v_n \}$, e definimos

$$w_j = \sum_{i=1}^{n} p_{ij}\,v_i \qquad \text{para} \qquad j = 1, \cdots, n\,.$$

Então, $\gamma = \{ w_1, \cdots, w_n \}$ é uma base ordenada para V tal que P é a matriz de mudança da base γ para a base β, isto é, $P = [I]_\beta^\gamma$ (Exercício 3.71).

Portanto, temos que

$$[T]_\gamma^\gamma = P^{-1}\,[T]_\beta^\beta\,P = [I]_\gamma^\beta\,[T]_\beta^\beta\,[I]_\beta^\gamma$$

de acordo com o Teorema 6.2.2, o que completa a demonstração. ■

O conceito de matrizes similares é muito importante para o estudo de diagonalização de operadores lineares, tendo em vista que este problema pode ser reformulado no contexto de matrizes.

Definição 6.6.1 *Sejam V um espaço vetorial de dimensão finita sobre $\mathbb{F}$ e T um operador linear sobre V. Dizemos que T é um operador* ***diagonalizável*** *se existe uma base ordenada β para V tal que $[T]_\beta^\beta$ é uma matriz diagonal.*

Definição 6.6.2 *Seja $A \in \mathbb{M}_n(\mathbb{F})$. Dizemos que A é uma matriz* ***diagonalizável*** *se A é similar a uma matriz diagonal.*

Exemplo 6.6.2 *Considere a matriz simétrica $A \in \mathbb{M}_2(\mathbb{R})$ dada por:*

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

Pelo Exemplo 6.4.1, sabemos que seus autovalores são $\lambda_1 = 3$ e $\lambda_2 = -1$ com os autovetores associados

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad e \qquad X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

respectivamente.

Tomamos a matriz P e a matriz diagonal D dadas por:

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad e \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} .$$

Note que a matriz P foi construída a partir dos autovetores da matriz A e a matriz diagonal D foi construída com os autovalores da matriz A.

Assim, temos que a matriz A é similar a matriz diagonal D, onde P é a matriz que realiza a transformação de similaridade, isto é, $A = PDP^{-1}$ ou $D = P^{-1}AP$.

De fato, podemos verificar facilmente que

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \left(\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) .$$

Portanto, a matriz A é diagonalizável.

Teorema 6.6.3 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, β uma base ordenada para V e T um operador linear sobre V. Então, T é um operador* ***diagonalizável*** *se, e somente se, $[T]_\beta^\beta$ é uma matriz diagonalizável.*

Demonstração – Considerando que T é um operador diagonalizável. Então, existe uma base ordenada γ para V tal que $[T]_\gamma^\gamma$ é uma matriz diagonal. Portanto, pelo Teorema 6.2.2, temos que a matriz $[T]_\beta^\beta$ é similar a matriz $[T]_\gamma^\gamma$. Logo, $[T]_\beta^\beta$ é uma matriz diagonalizável.

Considerando agora que $[T]_\beta^\beta$ é uma matriz diagonalizável. Então, a matriz $[T]_\beta^\beta$ é similar a uma matriz diagonal B. Pelo Teorema 6.6.2, existe uma base ordenada γ para V tal que $B = [T]_\gamma^\gamma$. Logo, T é um operador diagonalizável. ∎

Corolário 6.6.1 *Sejam $A \in \mathbb{M}_n(\mathbb{F})$ e T_A o operador linear sobre $\mathbb{F}^n$ associado a matriz A. Então, A é uma matriz diagonalizável se, e somente se, T_A é um operador diagonalizável.*

Definição 6.6.3 *Seja $A \in \mathbb{M}_n(\mathbb{F})$. Dizemos que A é uma* ***matriz simples*** *se possui um conjunto de n autovetores linearmente independentes.*

Teorema 6.6.4 *Seja $A \in \mathbb{M}_n(\mathbb{F})$. Então, A é uma matriz simples se, e somente se, A é uma matriz diagonalizável.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 6.6.3 *Considere o espaço vetorial real* $\mathbb{R}^3$. *O operador linear* T *sobre o* $\mathbb{R}^3$ *definido por:* $T(x,y,z) = (3x - 4z,\ 3y + 5z,\ -z)$ *é um operador diagonalizável.*

Temos que a matriz $A = [T]_\beta^\beta$, onde β é a base canônica do $\mathbb{R}^3$, é dada por:

$$A = \begin{bmatrix} 3 & 0 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{bmatrix}$$

Assim, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = -(3 - \lambda)^2 (1 + \lambda).$$

Portanto, os autovalores de T são $\lambda_1 = 3$, com multiplicidade algébrica igual a 2, e $\lambda_2 = -1$, com multiplicidade algébrica igual a 1.

Podemos verificar facilmente que os autovetores associados ao autovalor $\lambda_1 = 3$ são do tipo $v = (x,y,0)$ para $x,\ y \in \mathbb{R}$ não–nulos. Assim, podemos escolher os autovetores $v_1 = (1,0,0)$ e $v_2 = (0,1,0)$ associados ao autovalor $\lambda_1 = 3$. Logo, o autovalor $\lambda_1 = 3$ tem multiplicidade geométrica igual a 2.

Os autovetores associados ao autovalor $\lambda_2 = -1$ são do tipo $v = (-4y, 5y, -4y)$ para $y \in \mathbb{R}$ não–nulo. Assim, podemos escolher o autovetor $v_3 = (-4,5,-4)$ associado ao autovalor $\lambda_2 = -1$.

Portanto, temos que a matriz $A = [T]_\beta^\beta$ é diagonalizável. Podemos verificar que

$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 5 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 5 \\ 0 & 0 & -4 \end{bmatrix} \right)^{-1}.$$

Assim, mostramos que T é um operador diagonalizável.

De modo análogo, sabemos que T é um operador diagonalizável, pois $\gamma = \{ v_1,\ v_2,\ v_3 \}$ é uma base para $\mathbb{R}^3$ de modo que $[T]_\gamma^\gamma$ é uma matriz diagonal. De fato, podemos verificar facilmente que

$$[T]_\gamma^\gamma = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Teorema 6.6.5 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e T um operador linear sobre V que possui autovalores distintos $\lambda_1, \cdots, \lambda_k$ com $v_1, \cdots, v_k$ os autovetores associados, respectivamente. Então, $\{ v_1, \cdots, v_k \}$ é linearmente independente em V.*

Demonstração – A prova é feita por indução matemática sobre k. Para $k = 1$ o resultado é obtido trivialmente. De fato, como $v_1 \neq 0_V$, pois v_1 é um autovetor, temos que $\{ v_1 \}$ é linearmente independente.

Agora supomos que o resultado seja válido para $k - 1$ autovalores distintos, onde $k - 1 \geq 1$. Finalmente, vamos mostrar que o resultado é válido para k autovalores distintos. Para isso, consideramos a combinação linear nula

$$\sum_{i=1}^{k} c_i \, v_i = 0_V .$$

Aplicando o operador T na equação acima e usando o fato $T(v_i) = \lambda \, v_i$, obtemos

$$\sum_{i=1}^{k} c_i \, \lambda_i \, v_i = 0_V .$$

Subtraindo da segunda equação o resultado da multiplicando a primeira equação por λ_k, tem–se a seguinte equação

$$\sum_{i=1}^{k-1} c_i \, (\lambda_i - \lambda_k) \, v_i = 0_V .$$

Pela hipótese de indução, temos que $v_1, \cdots, v_{k-1}$ são linearmente independentes, o que devemos ter $c_i \, (\lambda_i - \lambda_k) = 0$ para $i = 1, \cdots, k-1$. Como os autovalores são distintos, isto é, $\lambda_i \neq \lambda_k$ para $i \neq k$, temos que $c_i = 0$ para $i = 1, \cdots, k-1$. Assim da primeira equação, temos também que $c_k = 0$, provando que os autovetores $v_1, \cdots, v_k$ são linearmente independentes em V. ■

Corolário 6.6.2 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, digamos que $dim(V) = n$, e T um operador linear sobre V que possui n autovalores distintos. Então, T é um operador diagonalizável.*

Demonstração – Considerando que T possui n autovalores distintos $\lambda_1, \cdots, \lambda_n$. Sejam $v_1, \cdots, v_n$ os respectivos autovetores associados. Pelo Teorema 6.6.5, temos que $\beta = \{ v_1, \cdots, v_n \}$ é linearmente independente em V. Como $dim(V) = n$, temos que o conjunto β de autovetores é uma base para V. Como $T(v_i) = \lambda_i \, v_i$, temos que a matriz do operador T com relação à base ordenada de autovetores é a matriz diagonal $D = diag(\lambda_1, \cdots, \lambda_n)$, o que completa a demonstração. ■

Note que o operador linear T pode possuir n autovetores linearmente independentes que não estão associados a autovalores distintos, isto é, algum autovalor pode possuir multiplicidade algébrica r maior do que 1, mas possui uma multiplicidade geométrica igual a r, como já vimos em exemplos anteriores. Desse modo, podemos introduzir o conceito de diagonalização para operadores lineares.

Teorema 6.6.6 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e T um operador linear sobre V. Então, T é um operador* ***diagonalizável*** *se, e somente se, existe uma base ordenada β para V cujos elementos são autovetores de T.*

Demonstração – Inicialmente, supomos que T seja diagonalizável. Então, existe uma base ordenada $\beta = \{ v_1, \cdots, v_n \}$ para V tal que $[T]_\beta^\beta$ é uma matriz diagonal $\Lambda = diag(\lambda_1, \cdots, \lambda_n)$, onde os escalares $\lambda_1, \cdots, \lambda_n$ não necessariamente são distintos dois a dois. Assim, temos que

$$T(v_j) = \sum_{i=1}^{n} d_{ij}\, v_i = \lambda_j\, v_j \qquad \text{para} \qquad j = 1, \cdots, n\,.$$

Portanto, mostramos que cada elemento v_j da base ordenada β é um autovetor do operador T associado ao autovalor λ_j, isto é,

$$T(v_j) = \lambda_j\, v_j \qquad \text{para} \qquad j = 1, \cdots, n\,.$$

Desse modo, temos que $\beta = \{ v_1, \cdots, v_n \}$ é uma base de autovetores para V.

Finalmente, considerando que $\beta = \{ v_1, \cdots, v_n \}$ seja uma base ordenada para V formada de autovetores do operador T, isto é,

$$T(v_i) = \lambda_i\, v_i \qquad \text{para} \qquad i = 1, \cdots n\,,$$

onde $\lambda_1, \cdots, \lambda_n$ são os respectivos autovalores do operador T.

Podemos verificar facilmente que a matriz $[T]_\beta^\beta$ é dada por:

$$[T]_\beta^\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

o que completa a demonstração. ■

Exemplo 6.6.4 *Considere o operador linear* T *sobre* $I\!R^2$ *definido por:*

$$T(x,y) = (x+3y\,,\,4x+2y)\,.$$

Mostre que T *é um operador diagonalizável.*

Temos que a matriz $A = [T]^\beta_\beta$, com relação à base canônica β do $I\!R^2$, é dada por:

$$[T]^\beta_\beta = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (1-\lambda)(2-\lambda) - 12 = \lambda^2 - 3\lambda - 10\,.$$

Desse modo, temos que $\lambda_1 = 5$ e $\lambda_2 = -2$ são os autovalores de T. Podemos verificar facilmente que $v_1 = (3,4)$ e $v_2 = (1,-1)$ são os autovetores associados, respectivamente. Evidentemente, $\gamma = \{ v_1,\ v_2 \}$ é uma base de autovetores para $I\!R^2$. Logo, temos que T é um operador diagonalizável.

Exemplo 6.6.5 *Considere o operador linear* T *sobre* $I\!R^3$ *definido por:*

$$T(x,y,z) = (2y,\ 2x,\ 2z)\,.$$

Mostre que T *é um operador diagonalizável.*

Temos que a matriz $A = [T]^\beta_\beta$, com relação à base canônica β do $I\!R^3$, é dada por:

$$[T]^\beta_\beta = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (2-\lambda)(\lambda^2 - 4) = -(2-\lambda)^2(\lambda + 2)\,.$$

Desse modo, temos que $\lambda_1 = 2$ é um autovalor com multiplicidade algébrica igual 2 e $\lambda_2 = -2$ é um autovalor com multiplicidade algébrica igual a 1.

Podemos verificar facilmente que os autovalores associados ao autovalor λ_1 são do tipo $v = (x,x,z)$ para $x,\, z \in I\!R$ não–nulos. Assim, podemos escolher $v_1 = (1,1,0)$ e $v_2 = (0,0,1)$ os autovetores associados ao autovalor $\lambda_1 = 2$, que tem multiplicidade geométrica igual a 2. Para o autovalor $\lambda_2 = -2$ temos que os autovetores associados são do tipo $v = (0,0,z)$ para $z \in I\!R$ não–nulo. Assim, podemos escolher $v_3 = (1,-1,0)$ o autovetor associado ao autovalor $\lambda_2 = -2$, que tem multiplicidade geométrica igual a 1. Evidentemente, $\gamma = \{ v_1\,,\, v_2\,,\, v_3 \}$ é uma base de autovetores para $I\!R^3$. Logo, temos que T é um operador diagonalizável.

Exemplo 6.6.6 *Considere o espaço vetorial real* $I\!R^2$ *e* T *o operador linear dado por*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) = (-3x+4y,\, -x+2y) \end{array}$$

Mostre que T *é um operador linear diagonalizável.*

Seja β a base canônica do $I\!R^2$. Temos que a matriz $A = [T]_\beta^\beta$ é dada por:

$$A = \begin{bmatrix} -3 & 4 \\ -1 & 2 \end{bmatrix}.$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (-3 - \lambda)(2 - \lambda) + 4 = \lambda^2 + \lambda - 2.$$

Assim, $\lambda_1 = 1$ e $\lambda_2 = -2$ são os autovalores do operador T. Como os autovalores são distintos, podemos garantir que existe uma base de autovetores γ para $I\!R^2$ de modo que a matriz $[T]_\gamma^\gamma = \Lambda = diag(\lambda_1,\, \lambda_2)$.

Os autovetores associados a λ_1 e λ_2 são $v_1 = (1,1)$ e $v_2 = (4,1)$, respectivamente. Assim, temos que $\gamma = \{ v_1,\, v_2 \}$ é uma base de autovetores para $I\!R^2$. Sabemos que

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{e} \quad X_2 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

são os autovetores da matriz A associados aos autovalores $\lambda_1 = 1$ e $\lambda_2 = -2$, respectivamente. Podemos observar facilmente que $AP = P\Lambda$, onde

$$P = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \quad , \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{e} \quad P^{-1} = \frac{1}{3}\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix}.$$

Desse modo, a matriz $A = [T]_\beta^\beta$ pode ser representada da seguinte forma:

$$A = P\Lambda P^{-1} \quad \text{ou} \quad \Lambda = P^{-1} A P,$$

com a matriz P realizando a diagonalização da matriz $A = [T]_\beta^\beta$.

Exemplo 6.6.7 *Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:*

$$T(p(x)) = (1 + x)p'(x) + p''(x).$$

Determine uma base ordenada γ para $\mathcal{P}_2(\mathbb{R})$ tal que $[T]_\gamma^\gamma$ seja uma matriz diagonal.

Temos que $A = [T]_\beta^\beta$, onde $\beta = \{1, x, x^2\}$ é a base canônica de $\mathcal{P}_2(\mathbb{R})$, é dada por:

$$[T]_\beta^\beta = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}.$$

Desse modo, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = -\lambda(1 - \lambda)(2 - \lambda).$$

Portanto, os autovalores de T são $\lambda_1 = 0$, $\lambda_2 = 1$ e $\lambda_3 = 2$. Como o operador T possui três autovalores distintos, pelo Corolário 6.6.2, sabemos que T é um operador linear diagonalizável e que a base ordenada $\gamma = \{ p_1(x), p_2(x), p_3(x) \}$, formada pelos autovetores de T, é tal que $[T]_\gamma^\gamma = \Lambda = diag(\lambda_1, \lambda_2, \lambda_3)$.

Podemos verificar facilmente que os autovetores da matriz A são

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad X_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 0$, $\lambda_2 = 1$ e $\lambda_3 = 2$, respectivamente. Portanto, sabemos que

$$[p_1(x)]_\beta = X_1 \quad , \quad [p_2(x)]_\beta = X_2 \quad \text{e} \quad [p_3(x)]_\beta = X_3,$$

onde $p_1(x)$, $p_2(x)$ e $p_3(x)$ são os autovetores do operador linear T associados aos autovalores $\lambda_1 = 0$, $\lambda_2 = 1$ e $\lambda_3 = 2$, respectivamente. Logo, obtemos

$$p_1(x) = 1 \quad , \quad p_2(x) = 1 + x \quad \text{e} \quad p_3(x) = 2 + 2x + x^2.$$

É importante observar que $A = P \Lambda P^{-1}$ ou $\Lambda = P^{-1} A P$, onde

$$P = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad P^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Exemplo 6.6.8 *Considere o espaço vetorial real* $I\!R^2$ *e* T *o operador linear dado por*

$$\begin{array}{rccl} T: & I\!R^2 & \longrightarrow & I\!R^2 \\ & (x,y) & \longrightarrow & T(x,y) = (2x+2y,\, 2x+5y) \end{array}$$

Mostre que T *é um operador linear diagonalizável.*

Seja β a base canônica do $I\!R^2$. Temos que a matriz $A = [T]^{\beta}_{\beta}$ é dada por

$$A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}.$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (2 - \lambda)(5 - \lambda) - 4 = \lambda^2 - 7\lambda + 6.$$

Assim, $\lambda_1 = 1$ e $\lambda_2 = 6$ são os autovalores do operador T. Como os autovalores são distintos, podemos garantir que existe uma base de autovetores γ para $I\!R^2$ de modo que a matriz $[T]^{\gamma}_{\gamma} = \Lambda = diag(\lambda_1,\, \lambda_2)$.

Os autovetores associados aos autovalores λ_1 e λ_2 são $v_1 = (-2,1)$ e $v_2 = (1,2)$, respectivamente. Assim, temos que $\gamma = \{ v_1,\, v_2 \}$ é uma base ortogonal de autovetores para o espaço vetorial $I\!R^2$. Podemos obter também uma base ortonormal de autovetores $\gamma^* = \{ q_1 , q_2 \}$ para $I\!R^2$, obtida a partir da base ortogonal γ. Sabemos que

$$X_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \text{e} \quad X_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

são os autovetores da matriz A associados aos autovalores $\lambda_1 = 1$ e $\lambda_2 = 6$, respectivamente. Podemos observar facilmente que $AQ = Q\Lambda$, onde

$$Q = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}.$$

Desse modo, a matriz $A = [T]^{\beta}_{\beta}$ pode ser representada da seguinte forma:

$$A = Q\Lambda Q^t \quad \text{ou} \quad \Lambda = Q^t A Q.$$

Note que a matriz Q é uma matriz ortogonal, isto é, $QQ^t = Q^t Q = I$, e realiza a diagonalização da matriz $A = [T]^{\beta}_{\beta}$.

Exemplo 6.6.9 *Considere o espaço vetorial real* $I\!R^3$ *e* T *o operador linear dado por*

$$\begin{array}{rccl} T: & I\!R^3 & \longrightarrow & I\!R^3 \\ & (x,y,z) & \longrightarrow & T(x,y,z) = (y+z,\, x+z,\, x+y) \end{array}$$

Mostre que T *é um operador linear diagonalizável.*

Seja β a base canônica do $I\!R^3$. Temos que a matriz $A = [T]_\beta^\beta$ é dada por

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Portanto, o polinômio característico do operador T é dado por:

$$p(\lambda) = \det(A - \lambda I) = -\lambda^3 + 3\lambda + 2 = -(\lambda + 1)^2(\lambda - 2).$$

Os autovalores do operador T são $\lambda_1 = 2$, com multiplicidade algébrica igual a 1, e $\lambda_2 = -1$, com multiplicidade algébrica igual a 2.

Para determinar os autovetores associados ao autovalor $\lambda_1 = 2$, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} \iff \begin{cases} -2x + y + z = 0 \\ \quad 3y - 3z = 0 \end{cases}$$

que tem como solução $z = y = x$. Portanto, os autovetores associados ao autovalor $\lambda_1 = 2$ são do tipo $v_1 = (x,x,x)$, com $x \neq 0$. Desse modo, podemos escolher $v_1 = (1,1,1)$ o autovetor associado ao autovalor $\lambda_1 = 2$. Assim, temos que o autovalor λ_1 tem multiplicidade geométrica igual a 1.

Para determinar os autovetores associados ao autovalor $\lambda_2 = -1$, temos que obter a solução do seguinte sistema linear

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = -\begin{bmatrix} x \\ y \\ z \end{bmatrix} \iff x + y + z = 0$$

que tem como solução $z = -y - x$.

Os autovetores associados ao autovalor $\lambda_2 = -1$ são do tipo $v = (x, y, -x-y)$, com $x,\, y \neq 0$. Desse modo, podemos escolher $v_2 = (0,1,-1)$ e $v_3 = (1,0,-1)$ os autovetores associados ao autovalor $\lambda_2 = -1$. Assim, temos que o autovalor λ_2 tem multiplicidade geométrica igual a 2.

Desse modo, temos que $\gamma = \{ v_1 \,,\, v_2 \,,\, v_3 \}$ é uma base de autovetores para o espaço vetorial real $I\!R^3$ que realiza a diagonalização do operador T, isto é, a representação matricial $[T]^{\gamma}_{\gamma} = \Lambda = diag(\lambda_1, \lambda_2, \lambda_3)$.

A partir da base de autovetores γ podemos obter uma base de autovetores ortonormais $\gamma^* = \{ q_1 \,,\, q_2 \,,\, q_3 \}$ dada por:

$$q_1 = \frac{\sqrt{3}}{3}(1,1,1) \quad , \quad q_2 = \frac{\sqrt{2}}{2}(0,1,-1) \qquad \text{e} \qquad q_3 = \frac{\sqrt{6}}{6}(2,-1,-1) \,.$$

Assim, temos que

$$X_1 = \begin{bmatrix} \dfrac{\sqrt{3}}{3} \\[2mm] \dfrac{\sqrt{3}}{3} \\[2mm] \dfrac{\sqrt{3}}{3} \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 0 \\[2mm] \dfrac{\sqrt{2}}{2} \\[2mm] -\dfrac{\sqrt{2}}{2} \end{bmatrix} \qquad \text{e} \qquad X_3 = \begin{bmatrix} \dfrac{\sqrt{6}}{3} \\[2mm] -\dfrac{\sqrt{6}}{6} \\[2mm] -\dfrac{\sqrt{6}}{6} \end{bmatrix}$$

são os autovetores da matriz A associados aos autovalores $\lambda_1 = 2\,,\ \lambda_2 = -1$ e $\lambda_3 = -1$, respectivamente. Podemos observar facilmente que $A\,Q = Q\,\Lambda$, onde

$$Q = \begin{bmatrix} \dfrac{\sqrt{3}}{3} & 0 & \dfrac{\sqrt{6}}{3} \\[2mm] \dfrac{\sqrt{3}}{3} & \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{6}}{6} \\[2mm] \dfrac{\sqrt{3}}{3} & -\dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{6}}{6} \end{bmatrix} \qquad \text{e} \qquad \Lambda = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} .$$

Desse modo, a matriz $A = [T]^{\beta}_{\beta}$ pode ser representada da seguinte forma:

$$A = Q\,\Lambda\,Q^t \qquad \text{ou} \qquad \Lambda = Q^t\,A\,Q \,.$$

Note que a matriz Q é uma matriz ortogonal, isto é, $Q\,Q^t = Q^t\,Q = I$, e realiza a diagonalização da matriz $A = [T]^{\beta}_{\beta}$.

Exemplo 6.6.10 *Seja* A *uma matriz de ordem* n *diagonalizável, isto é, existe uma matriz* P *invertível tal que* $\Lambda = P^{-1} A P$ *ou* $A = P \Lambda P^{-1}$*, onde* Λ *é uma matriz diagonal dada por:*

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

com $\lambda_1 , \cdots , \lambda_n$ *os autovalores da matriz* A*, não necessariamente distintos dois a dois.*

Podemos verificar facilmente que $A^m = P \Lambda^m P^{-1}$*, onde* Λ^m *é dada por:*

$$\Lambda^m = \begin{bmatrix} (\lambda_1)^m & 0 & \cdots & 0 \\ 0 & (\lambda_2)^m & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (\lambda_n)^m \end{bmatrix},$$

para todo $m \in \mathbb{N}$*. Desse modo, quando* A *é uma matriz diagonalizável, podemos calcular com uma certa eficiência qualquer potência de* A*.*

Exemplo 6.6.11 *Seja* A *uma matriz de ordem* n*. Sabemos que a série*

$$I_n + A + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!} + \cdots$$

converge para a matriz $\exp(A)$*, isto é,*

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} .$$

Considerando que A *é uma matriz diagonalizável, temos que*

$$\exp(A) = \sum_{k=0}^{\infty} \frac{P \Lambda^k P^{-1}}{k!} = P \left(\sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} \right) P^{-1} = P \exp(\Lambda) P^{-1} ,$$

onde P *é a matriz que realiza a diagonalização de* A *e a matriz* $\exp(\Lambda)$ *é dada por:*

$$\exp(\Lambda) = \begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{bmatrix}.$$

Desse modo, quando A *é uma matriz diagonalizável, podemos calcular com uma certa eficiência a matriz* $\exp(A)$*.*

Exemplo 6.6.12 *Como uma aplicação direta do Exemplo 6.6.11, vamos considerar o seguinte Problema de Valor Inicial representado pelo sistema dinâmico*

$$\begin{cases} X'(t) &= A\,X(t) \\ X(0) &= X_0 \end{cases}$$

onde A *é uma matriz de ordem* n*, e os vetores coluna de ordem* $n \times 1$*,* $X(t)$ *e* X_0 *são dados por:*

$$X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_i(t) \\ \vdots \\ x_n(t) \end{bmatrix} \qquad e \qquad X_0 = \begin{bmatrix} x_1(0) \\ \vdots \\ x_i(0) \\ \vdots \\ x_n(0) \end{bmatrix} .$$

Estamos considerando que cada componente da função vetorial $X(t)$*, isto é,* $x_i(t)$*, é uma função continuamente diferenciável para todo* $t \geq 0$*. O vetor coluna* X_0 *é a condição inicial do sistema dinâmico. Os vetores* $X(t)$ *e* X_0 *são geralmente denominados vetor de estado e vetor de estado inicial, respectivamente.*

Vamos apresentar o desenvolvimento para obtenção da solução do sistema dinâmico acima, considerando que a matriz A seja diagonalizável. Desse modo, sabemos que existe uma matriz invertível P de ordem n tal que $A = P\,\Lambda\,P^{-1}$, onde $\Lambda = diag(\lambda_1\,, \cdots, \lambda_n)$.

Assim, substituindo $A = P\,\Lambda\,P^{-1}$ no sistema dinâmico e fazendo a mudança de variável $Y(t) = P^{-1}\,X(t)$ e $Y(0) = P^{-1}\,X(0)$, obtemos o sistema dinâmico equivalente

$$\begin{cases} Y'(t) &= \Lambda\,Y(t) \\ Y(0) &= Y_0 \end{cases}$$

Note que como Λ é uma matriz diagonal, ficamos com n Problemas de Valor Inicial de primeira ordem sem acoplamentos, isto é,

$$\begin{cases} y_i'(t) &= \lambda_i\,y_i(t) \\ y_i(0) &= c_i \end{cases}$$

onde c_i é a condição inicial de cada um dos problemas, para $i = 1, \cdots, n$.

Sabemos que a solução de cada um dos Problemas de Valor Inicial, sem acoplamentos, é

$$y_i(t) \;=\; c_i \exp(\lambda_i t) \qquad \text{e} \qquad i \;=\; 1, \cdots, n\,.$$

Podemos verificar facilmente que as n soluções podem ser escritas da seguinte forma:

$$Y(t) \;=\; \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} y_1(0) \\ \vdots \\ y_i(0) \\ \vdots \\ y_n(0) \end{bmatrix} \;=\; \exp(\Lambda t) Y(0)\,.$$

Utilizando novamente a mudança de variável, obtemos a solução do sistema dinâmico original que é dada por:

$$X(t) \;=\; P \exp(\Lambda t) P^{-1} X(0) \qquad \text{para todo} \qquad t \;\geq\; 0\,.$$

Assim, pelo Exemplo 6.6.11, temos que $X(t) \;=\; \exp(At)\, X_0$ para todo $t \;\geq\; 0$.

Para exemplificar, vamos considerar o seguinte sistema dinâmico

$$\begin{cases} x'(t) \;=\; -2x(t) \;+\; y(t) \;+\; z(t) \\ y'(t) \;=\; x(t) \;-\; 2y(t) \;+\; z(t) \\ z'(t) \;=\; x(t) \;+\; y(t) \;-\; 2z(t) \end{cases}$$

com a condição inicial

$$\begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix} \;=\; \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}.$$

Neste caso, a matriz do sistema dinâmico é dada por:

$$A \;=\; \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

Vamos determinar os autovalores e os autovetores da matriz A. O polinômio característico da matriz A é dada por:

$$p(\lambda) \;=\; \det(\, A \;-\; \lambda\, I\,) \;=\; -\lambda(\lambda \;+\; 3)^2$$

Assim, os autovalores da matriz A são $\lambda_1 \;=\; 0\,,$ $\lambda_2 \;=\; -3$ e $\lambda_3 \;=\; -3$.

Podemos verificar facilmente que os autovetores da matriz A são

$$X_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad X_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 0$, $\lambda_2 = -3$ e $\lambda_3 = -3$, respectivamente. Podemos observar facilmente que os autovetores X_1, X_2 e X_3 são linearmente independentes. Logo, a matriz A é diagonalizável.

Portanto, a matriz P que realiza a diagonalização da matriz A, sua respectiva inversa e a matriz diagonal Λ são dadas por:

$$P = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad , \quad P^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Finalmente, sabemos que a solução do sistema dinâmico é dada por:

$$X(t) = P \exp(\Lambda t) P^{-1} X(0) \qquad \text{para todo} \qquad t \geq 0 .$$

Portanto, obtemos

$$X(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} 3 + e^{-3t} - 3e^{-3t} \\ 3 - 3e^{-3t} \\ 3 + 3e^{-3t} \end{bmatrix} \qquad \text{para todo} \qquad t \geq 0 .$$

Na Figura 6.1 temos os gráficos das soluções do sistema dinâmico do Exemplo 6.6.12. A curva azul representa o gráfico da solução $x(t)$, a curva verde representa o gráfico da solução $y(t)$ e a curva vermelha representa o gráfico da solução $z(t)$, para $0 \leq t \leq 2$.

Exemplo 6.6.13 *Determine a solução do seguinte sistema dinâmico*

$$\begin{cases} x'(t) = x(t) - 2y(t) \\ y'(t) = 3x(t) - 4y(t) \end{cases}$$

com a condição inicial

$$\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Neste caso, a matriz do sistema dinâmico é dada por:

$$A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix} .$$

Vamos determinar os autovalores e os autovetores da matriz A. O polinômio característico da matriz A é dada por:

$$p(\lambda) = \det(A - \lambda I) = \lambda^2 + 3\lambda + 2$$

Assim, os autovalores da matriz A são $\lambda_1 = -1$ e $\lambda_2 = -2$. Portando, a matriz A é diagonalizável.

Podemos verificar facilmente que os autovetores da matriz A são

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{e} \quad X_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = -1$ e $\lambda_2 = -2$, respectivamente.

Portanto, a matriz P que realiza a diagonalização da matriz A, sua respectiva inversa e a matriz diagonal Λ são dadas por:

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \quad , \quad P^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} .$$

Finalmente, sabemos que a solução do sistema dinâmico é dada por:

$$X(t) = P \exp(\Lambda t) P^{-1} X(0) \quad \text{para todo} \quad t \geq 0 .$$

Portanto, obtemos

$$X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} -2e^{-t} + 4e^{-2t} \\ -2e^{-t} + 6e^{-2t} \end{bmatrix} \quad \text{para todo} \quad t \geq 0 .$$

Na Figura 6.2 temos os gráficos das soluções do sistema dinâmico do Exemplo 6.6.13. A curva azul representa o gráfico da solução $x(t)$ e a curva verde representa o gráfico da solução $y(t)$, para $0 \leq t \leq 6$.

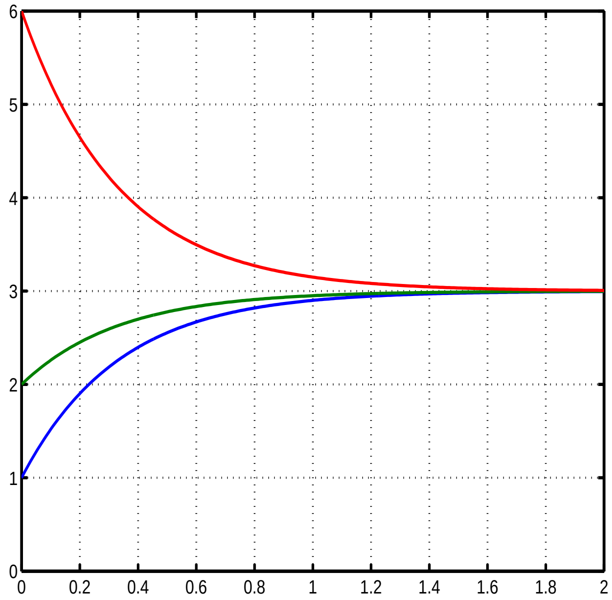

Figura 6.1: Gráficos dos soluções do sistema dinâmico do Exemplo 6.6.12. A curva azul representa o gráfico da solução $x(t)$, a curva verde representa o gráfico da solução $y(t)$ e a curva vermelha representa o gráfico da solução $z(t)$, para $0 \leq t \leq 2$.

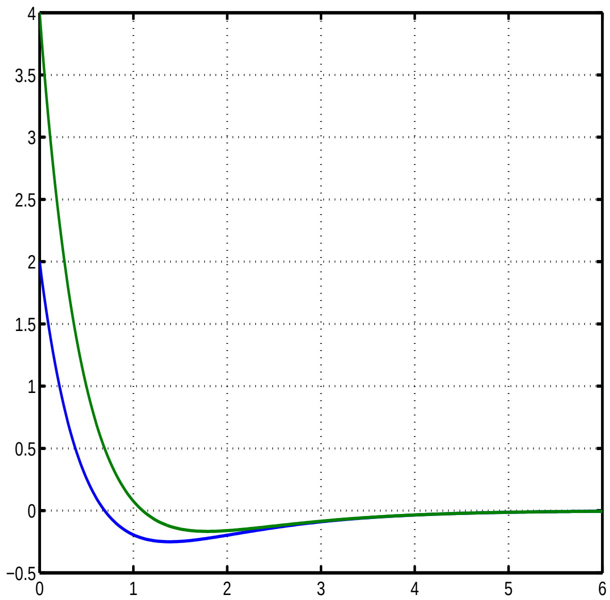

Figura 6.2: Gráficos das soluções do sistema dinâmico do Exemplo 6.6.13. A curva azul representa o gráfico da solução $x(t)$, a curva verde representa o gráfico da solução $y(t)$, para $0 \leq t \leq 6$.

Exercícios

Exercício 6.65 *Seja T o operador linear sobre $\mathbb{R}^3$ definido por:*

$$T(x, y, z) = (-3x - 4y, 2x + 3y, -z).$$

Encontre os autovalores e os autovetores do operador linear T. O operador linear T é diagonalizável ? Justifique sua resposta.

Exercício 6.66 *Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ e o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por: $T(p(x)) = p(x) + xp'(x)$. Determine uma base ordenada γ para $\mathcal{P}_2(\mathbb{R})$ de modo que $[T]_\gamma^\gamma$ seja uma matriz diagonal.*

Exercício 6.67 *Considere o espaço vetorial real $\mathcal{P}_3(\mathbb{R})$ e o operador linear T sobre $\mathcal{P}_3(\mathbb{R})$ definido por: $T(p(x)) = p'(x) + p''(x)$. Verifique se T é um operador linear diagonalizável. Em caso afirmativo, determine uma base ordenada γ para $\mathcal{P}_3(\mathbb{R})$ de modo que $[T]_\gamma^\gamma$ seja uma matriz diagonal.*

Exercício 6.68 *Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ e o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por: $T(p(x)) = p(x) + (x+1)p'(x)$. Determine uma base ordenada γ para $\mathcal{P}_2(\mathbb{R})$ de modo que $[T]_\gamma^\gamma$ seja uma matriz diagonal.*

Exercício 6.69 *Seja $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_2(\mathbb{R})$ o operador linear dado por:*

$$T(a + bx + cx^2) = (2b + c) + (2b - c)x + 2cx^2.$$

Verifique se T é um operador diagonalizável. Justifique sua resposta.

Exercício 6.70 *Considere o espaço vetorial real $\mathbb{M}_3(\mathbb{R})$ e o operador linear T sobre $\mathbb{M}_3(\mathbb{R})$ definido por: $T(A) = A^t$. Determine uma base ordenada para $\mathbb{M}_3(\mathbb{R})$ tal que $[T]_\beta^\beta$ seja uma matriz diagonal.*

Exercício 6.71 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, digamos que $dim(V) = n$, T um operador linear sobre V que possui somente dois autovalores distintos λ_1 e λ_2 com $dim(V_{\lambda_1}) = n - 1$. Prove que T é um operador diagonalizável.*

Exercício 6.72 *Dê um exemplo de um operador linear diagonalizável T sobre $\mathbb{R}^3$ cujo núcleo é gerado pelo elemento $u = (1, 0, 1)$.*

Exercício 6.73 *Dê um exemplo de um operador linear diagonalizável* T *sobre* $I\!R^3$ *cujo imagem é gerada pelos elementos* $u_1 = (1,1,0)$ *e* $u_2 = (1,0,1)$.

Exercício 6.74 *Considere o espaço vetorial complexo* $\mathbb{C}^2$ *e o operador linear* T *sobre* $\mathbb{C}^2$ *definido por:* $T(x,y) = (x + iy\,,\, ix + y)$. *Verifique se* T *é um operador linear diagonalizável. Em caso afirmativo, determine uma base ordenada* γ *para* $\mathbb{C}^2$ *de modo que* $[T]_\gamma^\gamma$ *seja uma matriz diagonal.*

Exercício 6.75 *Dê um exemplo de um operador linear diagonalizável* T *sobre* $\mathcal{P}_2(I\!R)$ *satisfazendo simultaneamente as seguintes propriedades:*

1. $\lambda_1 = -3$ *é um autovalor de* T.
2. $Ker(T) = [1 - x]$.
3. $T(p(x)) \neq p(x)$ *para todo* $p(x)$ *não–nulo.*

Exercício 6.76 *Considere a matriz* A *dada por:*

$$A = \begin{bmatrix} 0 & 1 & 5 & 9 \\ 2 & 1 & 6 & 8 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}.$$

Determine se possível uma matriz $P \in I\!M_4(I\!R)$ *invertível de modo que* $P^{-1}AP$ *seja uma matriz diagonal.*

Exercício 6.77 *Considere a matriz* A *dada por:*

$$A = \begin{bmatrix} 3 & -5 \\ 1 & -3 \end{bmatrix}.$$

Calcule de maneira eficiente A^n *para* $n \in I\!N$.

Exercício 6.78 *Considere a matriz* A *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ a & 2 & 0 \\ 0 & b & 2 \end{bmatrix}.$$

Determine todos os valores dos parâmetros a *e* b *de modo que a matriz* A *seja diagonalizável. Para estes valores de* a *e* b *, determine uma matriz invertível* P *e a matriz diagonal* D *de modo que* $P^{-1}AP = D$.

Exercício 6.79 *Sejam* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$*, digamos que* $dim(V) = n$*, e* $c \in \mathbb{F}$*. Pede–se:*

(a) Para qualquer base ordenada β *de* V *mostre que* $[cI_V]_\beta^\beta = cI_n$*, onde* I_V *é o operador identidade sobre* V *e* I_n *é a matriz identidade de ordem* n*.*

(b) Determine o polinômio característico do operador cI_V*.*

(c) Mostre que o operador cI_V *possui um único autovalor.*

(d) Mostre que o operador cI_V *é um operador diagonalizável.*

Exercício 6.80 *Considere o espaço vetorial complexo* $\mathbb{C}^2$ *e o operador linear* T *sobre* $\mathbb{C}^2$ *definido por:* $T(x,y) = (ix + y\,,\,2x - iy)$*. Verifique se* T *é um operador linear diagonalizável. Em caso afirmativo, determine uma base ordenada* γ *para* $\mathbb{C}^2$ *de modo que* $[T]_\gamma^\gamma$ *seja uma matriz diagonal.*

Exercício 6.81 *Determine o operador linear* T *sobre o* $\mathbb{R}^4$*, diagonalizável, que satisfaz simultaneamente as seguintes condições:*

(a) $Ker(T) = \{\, (x,y,z,t) \in \mathbb{R}^4 \;/\; x + y - z + t = 0 \quad e \quad z - t = 0 \,\}$.

(b) $T(0,0,1,0) = (0,0,2,0)$.

(c) $(0,1,0,0) \in Im(T)$.

(d) $\lambda = -3$ *é um autovalor do operador* T*.*

Exercício 6.82 *Considerando o operador linear diagonalizável* T *do Exemplo 6.6.6, determine explicitamente a expressão do operador linear* T^{16} *sobre o* $\mathbb{R}^2$*.*

Exercício 6.83 *Considere o operador linear diagonalizável* T *do Exemplo 6.6.7. Dado o polinômio* $p(x) = 2 - x + 3x^2$*, determine o polinômio* $q(x) = T^{16}(p(x))$*.*

Exercício 6.84 *Determine o operador linear* T *sobre o* $\mathbb{R}^4$*, diagonalizável, que satisfaz simultaneamente as seguintes condições:*

(a) $Ker(T) = \{\, (x,y,z,t) \in \mathbb{R}^4 \;/\; x + y - z + t = 0 \quad e \quad z - t = 0 \,\}$.

(b) $Im(T) = [(1,0,0,0)\,,\,(0,1,1,0)]$.

(c) $\lambda = 2$ *é um autovalor de* T *com multiplicidade algébrica igual a* 2*.*

Exercício 6.85 *Considere o operador linear* $T : \mathcal{P}_1(\mathbb{R}) \longrightarrow \mathcal{P}_1(\mathbb{R})$ *dado por:*

$$T(p(x)) \,=\, p'(x) \,+\, (x\,+\,1)p(1)\,.$$

Sejam $\beta = \{\,1,\,7-4x\,\}$ *e* $\gamma = \{\,q(x),\,2x-1\,\}$ *bases para* $\mathcal{P}_1(\mathbb{R})$ *tais que*

$$[T]^{\beta}_{\gamma} \,=\, \begin{bmatrix} 3 & s \\ -1 & 1 \end{bmatrix}.$$

(a) Determine o polinômio $q(x)$ *e o parâmetro* $s \in \mathbb{R}$.

(b) T *é um automorfismo? Em caso afirmativo, determine o automorfismo inverso.*

(c) O operador linear T *é diagonalizável? Justifique sua resposta.*

Exercício 6.86 *Considere o espaço vetorial real* $\mathbb{R}^n$ *munido do produto interno usual, que denotamos por* $\langle\,\cdot\,,\,\cdot\,\rangle$, *e o elemento* $u \in \mathbb{R}^n$ *não–nulo. Definimos as aplicações* P *e* Q *de* $\mathbb{R}^n$ *em* $\mathbb{R}^n$ *da seguinte forma:*

$$P(v) \,=\, \frac{\langle\, u\,,\,v\,\rangle}{\langle\, u\,,\,u\,\rangle}u \qquad e \qquad Q(v) \,=\, v \,-\, 2P(v)$$

para todo $v \in \mathbb{R}^n$.

(a) Mostre que P *e* Q *são operadores lineares sobre* $\mathbb{R}^n$.

(b) Mostre que $P(w) = w$, *com* $w = \alpha u$ *para todo* $\alpha \in \mathbb{R}$.

(c) Mostre que $P(w) = 0_{\mathbb{R}^n}$ *para* $\langle\, u\,,\,w\,\rangle = 0$.

(d) Mostre que $Q(w) = -w$, *com* $w = \alpha u$ *para todo* $\alpha \in \mathbb{R}$.

(e) Mostre que $Q(w) = w$ *para* $\langle\, u\,,\,w\,\rangle = 0$.

(f) Dê uma interpretação geométrica para os operadores lineares P *e* Q.

(g) O operador linear P *é diagonalizável? Justifique sua resposta.*

(h) O operador linear Q *é diagonalizável? Justifique sua resposta.*

Exercício 6.87 *Considere o operador linear* $T : \mathbb{M}_2(\mathbb{R}) \longrightarrow \mathbb{M}_2(\mathbb{R})$ *definido por:*

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \,=\, \begin{bmatrix} 2a+b & 2b \\ 2c & 3d \end{bmatrix}.$$

o operador linear T *é diagonalizável? Justifique sua resposta.*

6.7 Diagonalização de Operadores Hermitianos

Considere V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$. Seja T um operador Hermitiano sobre V, isto é,

$$\langle T(u) , v \rangle = \langle u , T(v) \rangle \quad ; \quad \forall \; u, v \in V .$$

Pelo Teorema 5.13.1, sabemos que a matriz $A = [T]_{\beta}^{\beta}$ é uma matriz Hermitiana, onde β é uma base ortonormal de V. Assim, O problema de diagonalização de uma matriz Hermitiana é equivalente ao problema de diagonalização de um operador Hermitiano.

Teorema 6.7.1 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$ e T um operador Hermitiano sobre V. Então, todo autovalor de T é real. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – Seja λ um autovalor de T com v o autovetor associado. Usando a hipótese que T é Hermitiano, obtemos

$$\lambda \langle v , v \rangle = \langle \lambda v , v \rangle = \langle T(v) , v \rangle = \langle v , T(v) \rangle = \langle v , \lambda v \rangle = \overline{\lambda} \langle v , v \rangle .$$

Logo, $(\lambda - \overline{\lambda}) \langle v , v \rangle = 0$. Como v é não–nulo, temos que $\lambda = \overline{\lambda}$. Portanto, os autovalores do operador T são reais.

Sejam λ_1 e λ_2 autovalores distintos do operador T, com v_1 e v_2 os autovetores associados, respectivamente. Desse modo, temos que

$$\lambda_1 \langle v_1 , v_2 \rangle = \langle T(v_1) , v_2 \rangle = \langle v_1 , T(v_2) \rangle = \langle v_1 , \lambda_2 v_2 \rangle = \lambda_2 \langle v_1 , v_2 \rangle .$$

Portanto, tem–se que $(\lambda_1 - \lambda_2) \langle v_1 , v_2 \rangle = 0$. Logo, $\langle v_1 , v_2 \rangle = 0$, pois λ_1 e λ_2 são distintos. Assim, completamos a demonstração. ■

Teorema 6.7.2 *Sejam V um espaço vetorial complexo munido do produto interno $\langle \cdot , \cdot \rangle$, T um operador Hermitiano sobre V e S um subespaço de V invariante sob T, isto é, $T(v) \in S$ para todo $v \in S$. Então, o subespaço $S^{\perp}$ é também invariante sob T.*

Demonstração – Seja $u \in S^{\perp}$, para todo $v \in S$ temos que

$$\langle T(u) , v \rangle = \langle u , T(v) \rangle = 0 .$$

Assim, provamos que $T(u) \in S^{\perp}$ para todo $u \in S^{\perp}$. ■

Teorema 6.7.3 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$, com $dim(V) = n$, e S o subespaço de V gerado pelo elemento unitário $u \in V$. Então, o subespaço $S^{\perp}$ tem dimensão $(n - 1)$.*

Demonstração – Seja $\beta = \{ u, v_2, \cdots , v_n \}$ uma base ortonormal para V. Dado um elemento $v \in S^{\perp} \subset V$ temos que

$$v = c_1 u + c_2 v_2 + \cdots + c_n v_n \quad ; \quad c_j \in \mathbb{F}$$

Note que $c_1 = \langle u , v \rangle = 0$. Portanto, todo elemento do subespaço $S^{\perp}$ pode ser escrito como uma combinação linear dos elemento $v_2, \cdots , v_n$ da base ortonormal de V. Assim, provamos que $dim(S^{\perp}) = (n - 1)$. ■

Teorema 6.7.4 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$, com $dim(V) = n$, e T um operador Hermitiano sobre V. Então, existe uma base ortonormal para V formada de autovetores de T.*

Demonstração – A prova é feita por indução sobre a dimensão do espaço V. Se $n = 1$, T possui exatamente um autovalor λ_1 com v_1 o autovetor associado. Assim, podemos considerar $\| v_1 \|_2 = 1$. Agora consideramos que o resultado seja válido para um espaço de dimensão $n - 1$, com $n > 2$, e vamos mostrar que o resultado é válido para um espaço de dimensão n.

Seja (λ_1, v_1) um autopar de T, isto é, $T(v_1) = \lambda_1 v_1$. Vamos considerar o subespaço $S = [v_1]$. Temos que S é invariante sob T. Pelo Teorema 6.7.2, temos que $S^{\perp}$ é invariante sob T. Desse modo, T é um operador Hermitiano sobre $S^{\perp}$. Como $dim(S^{\perp}) = n - 1$ e pela hipótese de indução, existem autovetores $v_2, \cdots , v_n$ de T os quais formam uma base ortonormal para o subespaço $S^{\perp}$.

Como $V = S \oplus S^{\perp}$, os autovetores $v_1, v_2, \cdots , v_n$ formam uma base ortonormal para o espaço V, o que completa a demonstração. ■

Tomando $\beta = \{ v_1, \cdots , v_n \}$ a base ortonormal de autovetores para o espaço V. Sabemos que a matriz do operador T com relação à base ortonormal β de autovetores é a matriz diagonal $[T]^{\beta}_{\beta} = diag(\lambda_1, \cdots , \lambda_n)$, onde v_k é o autovetor associado ao autovalor λ_k para $k = 1, \cdots , n$. Portanto, o operador T é diagonalizável.

Teorema 6.7.5 *Sejam* V *um espaço vetorial complexo de dimensão finita munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador Hermitiano sobre* V. *Então,* T *é um operador diagonalizável.*

Demonstração – A prova segue do Teorema 6.6.6 e do Teorema 6.7.4. □

Proposição 6.7.1 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana. Então,* A *é uma matriz diagonalizável, isto é,* A *é unitáriamente similar a uma matriz diagonal.*

Demonstração – A prova segue do Corolário 6.6.1 e do Teorema 6.7.5. □

Proposição 6.7.2 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *um matriz simétrica. Então,* A *é uma matriz diagonalizável, isto é,* A *é ortogonalmente similar a uma matriz diagonal.*

Demonstração – A prova segue da Proposição 6.7.1 e do fato que uma matriz simétrica real é um caso particular de uma matriz Hermitiana. □

Proposição 6.7.3 *Considere o espaço vetorial complexo* $\mathbb{C}^n$ *munido do produto interno usual. Sejam* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana e* T_A *o operador linear sobre* $\mathbb{C}^n$ *associado a matriz* A. *Então, o espaço vetorial complexo* $\mathbb{C}^n$ *possui uma base ortonormal de autovetores do operador linear* T_A.

Demonstração – A prova segue aplicando o Teorema 6.7.4 no operador linear T_A que é um operador Hermitiano sobre $\mathbb{C}^n$. □

Proposição 6.7.4 *Considere o espaço vetorial real* $\mathbb{R}^n$ *com o produto interno usual. Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz simétrica e* T_A *o operador linear sobre* $\mathbb{R}^n$ *associado a matriz* A. *o espaço vetorial real* $\mathbb{R}^n$ *possui uma base ortonormal de autovetores do operador linear* T_A.

Demonstração – A prova segue da Proposição 6.7.3 e do fato que uma matriz simétrica real é um caso particular de uma matriz Hermitiana. □

Exemplo 6.7.1 *Considere a matriz simétrica* $A \in \mathbb{M}_3(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

para fazer uma ilustração da Proposição 6.7.4.

Seja T_A o operador linear sobre $\mathbb{R}^3$ associado a matriz A, isto é,

$$T_A = (x + 2z\,,\, y\,,\, 2x + z)\,.$$

Assim, $A = [T]_{\beta}^{\beta}$, onde β é a base canônica do $\mathbb{R}^3$. Desse modo, os autovalores da matriz a são os autovalores do operador T_A, e os autovetores são os autovetores do operador T_A, representados como vetor coluna.

Temos que o polinômio característico da matriz $A = [T]_{\beta}^{\beta}$ é dado por:

$$p(\lambda) = \det(A - \lambda I) = (1 - \lambda)(\lambda^2 - 2\lambda - 3)\,.$$

Assim, os autovalores da matriz A são $\lambda_1 = -1$, $\lambda_2 = 1$ e $\lambda_3 = 3$.

Os autovetores de T_A associados ao autovalor $\lambda_1 = -1$ são do tipo $v = (x, 0, -x)$ para $x \in \mathbb{R}$ não–nulo.

Os autovetores de T_A associados ao autovalor $\lambda_2 = 1$ são do tipo $v = (0, y, 0)$ para $y \in \mathbb{R}$ não–nulo.

Os autovetores de T_A associados ao autovalor $\lambda_3 = 3$ são do tipo $v = (x, 0, x)$ para $x \in \mathbb{R}$ não–nulo.

Assim, temos a seguinte base ortonormal $\gamma = \{ q_1, q_2, q_3 \}$ para $\mathbb{R}^3$, onde

$$q_1 = \frac{\sqrt{2}}{2}(1, 0, -1) \quad , \quad q_2 = (0, 1, 0) \qquad \text{e} \qquad q_3 = \frac{\sqrt{2}}{2}(1, 0, 1)\,,$$

formada por autovetores do operador linear T_A. Podemos observar facilmente que

$$X_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad X_3 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\,,$$

são os autovetores da matriz A associados aos autovalores $\lambda_1 = -1$, $\lambda_2 = 1$ e $\lambda_3 = 3$, respectivamente.

Exemplo 6.7.2 *Considere o operador linear T sobre $\mathbb{C}^2$ definido por:*

$$T(x,y) = (4x + 2iy\,,\, -2ix + 4y) \qquad \textit{para todo} \qquad (x,y) \in \mathbb{C}^2\,.$$

Mostre que T é um operador Hermitiano sobre $\mathbb{C}^2$ e determine seus autovalores e autovetores.

Seja $\beta = \{\, e_1\,,\, e_2 \,\}$ a base canônica de $\mathbb{C}^2$, a matriz $A = [T]^{\beta}_{\beta}$ é dada por:

$$A = \begin{bmatrix} 4 & 2i \\ -2i & 4 \end{bmatrix}.$$

Podemos observar facilmente que A é uma matriz Hermitiana, isto é, $A^* = A$. Logo, pelo Teorema 5.13.1, temos que T é um operador Hermitiano.

O polinômio característico do operador linear T é dado por:

$$p(\lambda) = \det(\, A - \lambda\, I\,) = \lambda^2 - 8\lambda + 12\,.$$

Portanto, $\lambda_1 = 2$ e $\lambda_2 = 6$ são os autovalores do operador linear T.

Para determinar os autovalores de T associados ao autovalor $\lambda_1 = 2$, temos que encontrar os elementos não–nulos de $Ker(T - 2I)$. Assim, temos que obter as soluções não nulas do sistema linear homogêneo

$$\begin{cases} 2x + 2iy = 0 \\ -2ix + 2y = 0 \end{cases} \iff -2ix + 2y = 0 \implies y = ix$$

Portanto, todo elemento $v_1 = (a,\, i\,a) \in \mathbb{C}^2$, para $a \in \mathbb{C}$ não–nulo, é um autovetor do operador T associado ao autovalor $\lambda_1 = 2$. Assim, podemos escolher $v_1 = (1, i)$ o autovetor associado ao autovalor $\lambda_1 = 2$.

Para determinar os autovalores de T associados ao autovalor $\lambda_2 = 6$, temos que encontrar os elementos não–nulos de $Ker(T - 6I)$. Assim, temos que obter as soluções não nulas do sistema linear homogêneo

$$\begin{cases} -2x + 2iy = 0 \\ -2ix - 2y = 0 \end{cases} \iff -2ix - 2y = 0 \implies y = -ix$$

Portanto, todo elemento $v_1 = (a,\, -i\,a) \in \mathbb{C}^2$, para $a \in \mathbb{C}$ não–nulo, é um autovetor do operador T associado ao autovalor $\lambda_2 = 6$. Assim, podemos escolher $v_1 = (1, -i)$ o autovetor associado ao autovalor $\lambda_2 = 6$.

Temos que $\gamma = \{ v_1 , v_2 \}$ é uma base ortogonal de autovetores para o espaço vetorial complexo $\mathbb{C}^2$. Além disso, sabemos que a matriz $\Lambda = [T]^{\gamma}_{\gamma}$ é dada por:

$$\Lambda = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} .$$

Desse modo, temos que $\gamma^* = \{ q_1 , q_2 \}$, onde

$$q_1 = \frac{\sqrt{2}}{2}(1 , i) \qquad \text{e} \qquad q_2 = \frac{\sqrt{2}}{2}(1 , -i) ,$$

é uma base ortonormal de autovetores para $\mathbb{C}^2$.

Assim, temos que

$$X_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad \text{e} \qquad X_2 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

são os autovetores da matriz A associados aos autovetores $\lambda_1 = 2$ e $\lambda_2 = 6$, respectivamente. Podemos observar que $A\,U = U\Lambda$, onde

$$U = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} .$$

Desse modo, a matriz $A = [T]^{\beta}_{\beta}$ pode ser representada da seguinte forma:

$$A = U\,\Lambda\,U^* \qquad \text{ou} \qquad \Lambda = U^*\,A\,U .$$

Note que a matriz U é uma matriz unitária, isto é, $UU^* = U^*U = I$, e realiza a diagonalização da matriz $A = [T]^{\beta}_{\beta}$.

Definição 6.7.1 *Sejam V um espaço vetorial complexo munido do produto interno $\langle \cdot , \cdot \rangle$ e T um operador Hermitiano sobre V. Dizemos que T é um* ***operador positivo*** *sobre V se*

$$\langle T(u) \, , \, u \rangle \; > \; 0 \qquad \textit{para todo} \quad u \in V \quad \textit{não–nulo.}$$

Desse modo, temos que a aplicação

$$\begin{aligned} p : \; V \times V \; &\longrightarrow \; \mathbb{C} \\ (u, v) \; &\longrightarrow \; p(u, v) \; = \; \langle T(u) \, , \, v \rangle \end{aligned}$$

define um produto interno no espaço vetorial complexo V.

Teorema 6.7.6 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$, $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal ordenada para V, T um operador Hermitiano sobre V e $A = [T]_\beta^\beta$ a matriz do operador T com relação à base β. Então, T é um operador positivo se, e somente se, A é uma matriz positiva–definida.*

Demonstração – Seja $A = [a_{ij}]$ a representação matricial do operador T com relação à base ortonormal β, isto é, $a_{ij} = \langle T(q_j) \, , \, q_i \rangle$. Para todo $u \in V$ temos que

$$u \; = \; \sum_{j=1}^{n} \langle \, u \, , \, q_j \, \rangle \, q_j \; = \; \sum_{j=1}^{n} b_j \, q_j \; .$$

Podemos escrever $\langle T(u) \, , \, u \rangle$ da seguinte forma:

$$\begin{aligned} \langle T(u) \, , \, u \rangle \; &= \; \sum_{j=1}^{n} \sum_{i=1}^{n} b_j \, \overline{b}_i \, \langle T(q_j) \, , \, q_i \rangle \\ &= \; \sum_{j=1}^{n} \sum_{i=1}^{n} b_j \, \overline{b}_i \, a_{ij} \\ &= \; ([u]_\beta)^* \, A \, [u]_\beta \end{aligned}$$

Portanto, para todo $u \in V$, não–nulo, obtemos

$$\langle T(u) \, , \, u \rangle \; > \; 0 \qquad \Longleftrightarrow \qquad ([u]_\beta)^* \, A \, [u]_\beta \; > \; 0 \; ,$$

o que completa a demonstração. ■

Teorema 6.7.7 *Seja* $A \in I\!\!M_n(\mathbb{C})$ *uma matriz Hermitiana. Então, a matriz* A *é positiva–definida se, e somente se, seus autovalores são todos positivos.*

Demonstração

$(\Longrightarrow)$ Tomando A positiva–definida, do Teorema 6.4.4, temos que seus autovalores são todos positivos.

$(\Longleftarrow)$ Considerando A uma matriz Hermitiana e seus autovalores todos positivos.

Seja T_A o operador linear sobre $\mathbb{C}^n$ associado a matriz A. Da Proposição 6.7.3, temos que existe uma base ortonormal para o espaço vetorial complexo $\mathbb{C}^n$ de autovetores do operador T_A. Sejam $\lambda_1, \cdots, \lambda_n$ os autovalores do operador T_A, que são também os autovalores da matriz A, com $v_1, \cdots, v_n$ os autovetores associados.

Tomando um elemento não–nulo $u \in \mathbb{C}^n$, que não seja um autovetor de T_A, sabemos que pode ser escrito de modo único da seguinte forma:

$$u = \sum_{i=i}^{n} c_i v_i \,.$$

Desse modo, temos que

$$\langle\, T_A(u)\,,\, u\,\rangle = \langle \sum_{i=i}^{n} c_i T_A(v_i)\,,\, \sum_{j=i}^{n} c_j v_j \,\rangle = \langle \sum_{i=i}^{n} c_i \lambda_i v_i\,,\, \sum_{j=i}^{n} c_j v_j \,\rangle \,.$$

Logo, temos que

$$\langle\, T_A(u)\,,\, u\,\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} c_i\, \overline{c}_j\, \lambda_i \,\langle\, v_i\,,\, v_j\,\rangle \,.$$

Como $v_1, \cdots, v_n$ são mutuamente ortonormais, obtemos

$$\langle\, T_A(u)\,,\, u\,\rangle = \sum_{i=1}^{n} |\, c_i\,|^2\, \lambda_i > 0 \,.$$

Logo, T_A é um operador positivo. Assim, mostramos que a matriz A é positiva–definida, pois $A = [T]_{\beta}^{\beta}$, onde β é a base canônica de $\mathbb{C}^n$. ■

Corolário 6.7.1 *Seja $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz simétrica. Então, a matriz A é positiva–definida se, e somente se, seus autovalores são todos positivos.*

Demonstração – A prova segue do Teorema 6.7.7 e do fato que uma matriz simétrica real é um caso particular de uma matriz Hermitiana.

Exemplo 6.7.3 *Considere a matriz Hermitiana A dada por:*

$$A = \begin{bmatrix} 3 & -i \\ i & 3 \end{bmatrix}$$

para fazer uma ilustração do Teorema 6.7.7.

Podemos verificar facilmente que o polinômio característico da matriz A é dado por:

$$p(\lambda) = \det(A - \lambda I) = \lambda^2 - 6\lambda + 8 .$$

Assim, temos que os autovalores da matriz A são $\lambda_1 = 2$ e $\lambda_2 = 4$. Logo, pelo Teorema 6.7.7, temos que A é uma matriz positiva–definida.

Exemplo 6.7.4 *Considere a matriz simétrica A dada por:*

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

para fazer uma ilustração do Corolário 6.7.1.

Podemos verificar facilmente que o polinômio característico da matriz A é dado por:

$$p(\lambda) = \det(A - \lambda I) = (2 - \lambda)\{\,(2 - \lambda)(3 - \lambda) - 2\,\} = (2 - \lambda)(\lambda^2 - 5\lambda + 4) .$$

Assim, temos que os autovalores da matriz A são $\lambda_1 = 1$, $\lambda_2 = 2$ e $\lambda_3 = 4$. Logo, pelo Corolário 6.7.1, temos que A é uma matriz positiva–definida.

A seguir, enunciamos um resultado geométrico para uma matriz positiva–definida, que será muito importante na análise de convergência do Método dos Gradientes Conjugados.

Teorema 6.7.8 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz positiva–definida. Então, a equação*

$$x^t A x = 1 \tag{6.1}$$

representa um hiper–elipsóide em $I\!R^n$ *com centro na origem e cujos semi–eixos tem comprimentos*

$$\frac{1}{\sqrt{\lambda_1}}, \cdots, \frac{1}{\sqrt{\lambda_n}}$$

nas direções dos autovetores $q_1, \cdots, q_n$ *associados aos autovalores*

$$0 < \lambda_1 \leq \ldots \leq \lambda_n .$$

Demonstração – Como A é positiva–definida, vamos utilizar a sua diagonalização

$$A = Q \Lambda Q^t ,$$

onde

$$\Lambda = diag(\lambda_1, \ldots, \lambda_j, \ldots \lambda_n)$$

é uma matriz diagonal e

$$Q = [q_1 \cdots q_j \cdots q_n]$$

é uma matriz ortogonal. Note que (λ_j, q_j) é um autopar da matriz A.

Desse modo, podemos escrever a equação (6.1) da seguinte forma:

$$x^t A x = (Q^t x)^t \Lambda (Q^t x) = 1 .$$

Fazendo a mudança de variável $y = Q^t x$, obtemos a seguinte equação

$$y^t \Lambda y = \sum_{j=1}^{n} \lambda_j y_j^2 = \sum_{j=1}^{n} \frac{y_j^2}{a_j^2} = 1 ,$$

onde

$$a_j = \frac{1}{\sqrt{\lambda_j}}$$

é o comprimento do semi–eixo na direção do autovetor q_j.

Portanto, o maior eixo está na direção do autovetor associado ao menor autovalor e o menor eixo está na direção do autovetor associado ao maior autovalor, o que completa a demonstração. ■

Lei de Inércia de Sylvester

Definição 6.7.2 *Seja $A \in \mathbb{M}_n(\mathbb{C})$ uma matriz Hermitiana. A **inércia** da matriz A, que indicamos por $i(A)$, é o terno ordenado*

$$i(A) \;=\; (\, i_+(A)\,,\, i_-(A)\,,\, i_0(A)\,)\;,$$

onde $i_+(A)$ é o número de autovalores positivos de A, $i_-(A)$ é o número de autovalores negativos de A, e $i_0(A)$ é o número de autovalores iguais a zero de A, considerando a multiplicidade de cada um dos autovalores.

Definição 6.7.3 *Sejam $A, B \in \mathbb{M}_n(\mathbb{C})$. Dizemos que a matriz B é **congruente** com a matriz A se existe uma matriz invertível $P \in \mathbb{M}_n(\mathbb{C})$ tal que $B = PAP^*$.*

Exemplo 6.7.5 *Seja $A \in \mathbb{M}_m(\mathbb{C})$ uma matriz Hermitiana. Então, $A = U\Lambda U^*$, onde*

$$\Lambda \;=\; diag(\lambda_1\,, \ldots\,, \lambda_j\,, \ldots\,, \lambda_m)$$

é uma matriz diagonal real e

$$U \;=\; [u_1 \cdots u_j \cdots u_m]$$

é uma matriz unitária, com $(\lambda_j\,, u_j)$ um autopar da matriz A. Vamos mostrar que toda matriz Hermitiana A é congruente com a matriz diagonal $\widehat{\Lambda} \in \mathbb{M}_m(\mathbb{R})$ dada por:

$$\widehat{\Lambda} \;=\; diag(1\,, \ldots\,, 1\,, -1\,, \ldots\,, -1\,, 0\,, \ldots\,, 0)\;, \qquad (6.2)$$

*denominada **matriz de inércia** da matriz A, onde o número de 1 é igual a $i_+(A)$, o número de -1 é igual a $i_-(A)$, e o número de 0 é igual a $i_0(A)$.*

De fato, para simplificar a prova, vamos organizar os autovalores de A em três grupos. Os autovalores positivos que vamos denotar por $\lambda_1^+, \cdots, \lambda_p^+$, os autovalores negativos que vamos denotar por $\lambda_1^-, \cdots, \lambda_n^-$, e os autovalores nulos que vamos denotar por $\lambda_1^0, \cdots, \lambda_r^0$, com $p + n + r = m$. Assim, representamos a matriz diagonal Λ da seguinte forma:

$$\Lambda \;=\; diag(\lambda_1^+\,, \ldots\,, \lambda_p^+\,, \lambda_1^-\,, \ldots\,, \lambda_n^-\,, \lambda_1^0, \ldots\,, \lambda_r^0)\;.$$

Vamos definir uma matriz diagonal $D \in \mathbb{M}_m(\mathbb{R})$ da seguinte forma:

$$D \;=\; diag\left(\sqrt{\lambda_1^+}\,, \ldots\,, \sqrt{\lambda_p^+}\,, \sqrt{-\lambda_1^-}\,, \ldots\,, \sqrt{-\lambda_n^-}\,, 1\,, \ldots\,, 1\right)\;,$$

com a qual podemos escrever a matriz diagonal $\Lambda \in \mathbb{M}_m(\mathbb{R})$ da seguinte forma:

$$\Lambda \;=\; D\,\widehat{\Lambda}\,D\;.$$

Portanto, a matriz Hermitiana A pode ser escrita da forma:

$$A = U\Lambda U^* = U\,D\,\widehat{\Lambda}\,D\,U^* = (U\,D)\,\widehat{\Lambda}\,(UD)^* = P\,\widehat{\Lambda}\,P^* \,, \tag{6.3}$$

onde a matriz invertível $P = UD$ realiza a relação de congruência. Assim, mostramos que toda matriz Hermitiana A é congruente com a matriz diagonal $\widehat{\Lambda}$, em uma forma mais simples. Logo, conhecida a matriz diagonal $\widehat{\Lambda}$ sabemos a inércia da matriz A. Reciprocamente, sabendo a inércia da matriz A conhecemos a matriz diagonal $\widehat{\Lambda}$.

Proposição 6.7.5 *Sejam* $A, B \in I\!M_n(\mathbb{C})$ *matrizes congruentes. Então,*

$$posto(A) = posto(B) \,.$$

Demonstração – A prova segue imediata do Exercício 4.50. □

Teorema 6.7.9 (Lei de Inércia de Sylvester) *Sejam* $A, B \in I\!M_n(\mathbb{C})$ *matrizes Hermitianas. Então, existe uma matriz invertível* $P \in I\!M_n(\mathbb{C})$ *tal que* $A = PBP^*$ *se, e somente se,* A *e* B *tem a mesma inércia.*

Demonstração

($\Longrightarrow$) Tomando a hipótese que as matrizes A e B são congruentes, isto é, $A = PBP^*$, para alguma matriz invertível $P \in I\!M_n(\mathbb{C})$. Como matrizes congruentes tem o mesmo posto, $posto(A) = posto(B)$, temos que $i_0(A) = i_0(B)$. Assim, precisamos mostrar somente que $i_+(A) = i_+(B)$.

Sejam $u_1, \cdots, u_p$ autovetores ortonormais da matriz Hermitiana A associados aos autovalores positivos $\lambda_1^+, \cdots, \lambda_p^+$. E denotamos por $E_+(A)$ o subespaço gerado por esse conjunto de autovetores ortonormais, isto é,

$$E_+(A) = [u_1 \,, \cdots, u_p] \,.$$

Note que $dim(\, E_+(A)\,) = i_+(A)$.

Agora consideramos um elemento $w \in E_+(A)$ não–nulo, isto é,

$$w = \alpha_1 u_1 + \cdots + \alpha_1 u_p \neq 0_{\mathbb{C}^n} \,.$$

Assim, temos que

$$w^* A w = \lambda_1^+ \,|\, \alpha_1 \,|^2 + \cdots + \lambda_p^+ \,|\, \alpha_p \,|^2 > 0 \,.$$

Portanto, obtemos

$$x^*(\, PBP^* \,)x \;=\; (\, P^*x \,)^*B(\, P^*x \,) \;=\; z^*Bz \;>\; 0\,,$$

para todo elemento z não–nulo no subespaço gerado pelo conjunto

$$\{\, P^*u_1\,,\,\cdots\,,\,P^*u_p \,\}$$

linearmente independente, desde que P é uma matriz invertível. Desse modo, podemos concluir que $i_+(B) \geq i_+(A)$.

Trocando as posições das matrizes A e B, e fazendo as mesmas argumentações, temos que $i_+(A) \geq i_+(B)$. Portanto, provamos que $i_+(A) = i_+(A)$.

($\Longleftarrow$) Tomando como hipótese que as matrizes Hermitianas A e B tem a mesma inércia, sabemos que podem ser representadas como em (6.3), possuindo a mesma matriz de inércia $\widehat{\Lambda}$, isto é,

$$A \;=\; P\,\widehat{\Lambda}\,P^* \qquad \text{e} \qquad B \;=\; S\,\widehat{\Lambda}\,S^*\,,$$

onde $P,\, S \in I\!\!M_n(\mathbb{C})$ são matrizes invertíveis.

Pela propriedade transitiva da relação de congruência, veja Exemplo 2.8.3, e pelo fato que as matrizes A e B são congruentes a mesma matriz $\widehat{\Lambda}$, temos que as matrizes A e B são congruentes. De fato, podemos escrever a matriz $\widehat{\Lambda}$ da seguinte forma:

$$\widehat{\Lambda} \;=\; Q\,B\,Q^* \qquad \Longrightarrow \qquad A \;=\; (SQ)B(SQ)^*\,,$$

onde $Q = S^{-1}$, com SQ uma matriz invertível, desde que S e Q são matrizes invertíveis, o que completa a demonstração. ■

Corolário 6.7.2 *Sejam $A,\, B \in I\!\!M_n(I\!\!R)$ matrizes simétricas. Então, existe uma matriz invertível $P \in I\!\!M_n(I\!\!R)$ tal que $A = PBP^t$ se, e somente se, A e B tem a mesma inércia.*

Demonstração – A prova segue do resultado do Teorema 6.7.9, e do fato que uma matriz simétrica real é um caso particular de uma matriz Hermitiana. ■

Corolário 6.7.3 *Seja $A \in I\!\!M_n(I\!\!R)$ uma matriz simétrica. Então, existe uma matriz $P \in I\!\!M_n(I\!\!R)$ invertível tal que $D = PAP^t$ é uma matriz diagonal. Além disso, o número de elementos na diagonal de D que são positivos, negativos e nulos é sempre o mesmo, independente da matriz P que realiza a relação de congruência.*

Definição 6.7.4 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz simétrica. Dizemos que* A *é uma* ***matriz indefinida*** *quando possui autovalores de ambos os sinais.*

Exemplo 6.7.6 *Seja* $B \in I\!M_n(I\!R)$ *uma matriz invertível. Mostre que a matriz simétrica* H*, de ordem* $2n$*, dada por:*

$$H = \begin{bmatrix} I_n & B^t \\ B & 0_n \end{bmatrix}$$

é uma matriz indefinida.

Inicialmente vamos mostrar que H é uma matriz invertível. De fato, considere o sistema linear homogêneo

$$\begin{bmatrix} I_n & B^t \\ B & 0_n \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} = \begin{bmatrix} 0_{I\!R^n} \\ 0_{I\!R^n} \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{cases} u + B^t w = 0_{I\!R^n} \\ Bu = 0_{I\!R^n} \end{cases}$$

Portanto, como B é uma matriz invertível, obtemos

$$u = 0_{I\!R^n} \qquad \text{e} \qquad w = 0_{I\!R^n} \, .$$

Assim, mostramos que o sistema linear homogêneo possui somente a solução trivial.

Finalmente, vamos determinar os autovalores da matriz simétrica H, isto é, determinar os escalares $\lambda \in I\!R$ tais que

$$HX = \lambda X \qquad \Longleftrightarrow \qquad \begin{bmatrix} I_n & B^t \\ B & 0_n \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} = \lambda \begin{bmatrix} u \\ w \end{bmatrix} ,$$

para $u \, , w \in I\!R^n$ não–nulos.

Desse modo, obtemos as seguintes equações

$$\begin{cases} u + B^t w = \lambda u \\ Bu = \lambda w \end{cases}$$

Da segunda equação, obtemos

$$w = \frac{1}{\lambda} B u \, ,$$

com $\lambda \neq 0$, pois H é uma matriz invertível. Substituindo w na primeira equação, obtemos

$$B^t B u = (\lambda^2 - \lambda) u \, .$$

Como B é uma matriz invertível, sabemos que B^tB é uma matriz positiva–definida. Logo, seus autovalores são todos positivos, que vamos denotar por $\alpha_1, \cdots, \alpha_n$.

Desse modo, considerando (α_i, v_i) um autopar da matriz B^tB, obtemos

$$\lambda_i^2 - \lambda_i = \alpha_i$$

para $i = 1, \cdots n$.

Assim, para cada autovalor α_i de B^tB, temos dois autovalores associados para H dados por:

$$\lambda_i^+ = \frac{1 + \sqrt{1 + 4\alpha_i}}{2} > 0 \quad \text{e} \quad \lambda_i^- = \frac{1 - \sqrt{1 + 4\alpha_i}}{2} < 0 .$$

para $i = 1, \cdots n$.

Portanto, mostramos que H é uma matriz indefinida, e determinamos seus autovalores.

Exemplo 6.7.7 *Seja $B \in I\!M_{m\times n}(I\!R)$, com $m \leq n$ e $posto(B) = m$. Mostre que a matriz simétrica H, de ordem $n + m$, dada por:*

$$H = \begin{bmatrix} I_n & B^t \\ B & 0_m \end{bmatrix}$$

é uma matriz indefinida.

Inicialmente vamos mostrar que H é uma matriz invertível. De fato, considere o sistema linear homogêneo

$$\begin{bmatrix} I_n & B^t \\ B & 0_m \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} = \begin{bmatrix} 0_{I\!R^n} \\ 0_{I\!R^m} \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{cases} u + B^tw = 0_{I\!R^n} \\ Bu = 0_{I\!R^m} \end{cases}$$

Multiplicando a primeira equação por B, obtemos a equação

$$BB^tw = 0_{I\!R^m}$$

Como $posto(B) = m$, sabemos que BB^t, de ordem m, é uma matriz positiva–definida. Logo, BB^t é uma matriz invertível. Desse modo, obtemos

$$w = 0_{I\!R^m} \quad \text{e} \quad u = 0_{I\!R^n} .$$

Assim, mostramos que o sistema linear homogêneo possui somente a solução trivial.

Finalmente, vamos determinar os autovalores da matriz simétrica H, isto é, determinar os escalares $\lambda \in I\!R$ tais que

$$HX = \lambda X \qquad \Longleftrightarrow \qquad \begin{bmatrix} I_n & B^t \\ B & 0_m \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} = \lambda \begin{bmatrix} u \\ w \end{bmatrix} ,$$

para $u \in I\!R^n$ e $w \in I\!R^m$ não–nulos.

Desse modo, obtemos as seguintes equações

$$\begin{cases} \quad u \; + \; B^t w \; = \; \lambda u \\ \\ Bu \qquad\qquad = \; \lambda w \end{cases}$$

Da segunda equação, obtemos

$$w = \frac{1}{\lambda} Bu ,$$

com $\lambda \neq 0$, pois H é uma matriz invertível. Substituindo w na primeira equação, obtemos

$$B^t B u = (\lambda^2 - \lambda) u .$$

Como $posto(B) = m$ e $m \leq n$, sabemos que a matriz $B^t B$, de ordem n, é uma matriz semipositiva–definida e tem $posto(B^t B) = m$.

Logo, a matriz $B^t B$ possui m autovalores positivos, que vamos denotar por $\alpha_1, \cdots, \alpha_m$, e um autovalor $\beta = 0$ com multiplicidade algébrica $p = n - m$.

Desse modo, considerando $(\alpha_i \, , u_i)$ um autopar da matriz $B^t B$, obtemos

$$\lambda_i^2 - \lambda_i = \alpha_i$$

para $i = 1, \cdots m$.

Assim, para cada autovalor α_i de $B^t B$, temos dois autovalores associados para H dados por:

$$\lambda_i^+ = \frac{1 + \sqrt{1 + 4\alpha_i}}{2} > 0 \qquad \text{e} \qquad \lambda_i^- = \frac{1 - \sqrt{1 + 4\alpha_i}}{2} < 0 .$$

para $i = 1, \cdots m$.

Para o autovalor $\beta = 0$, obtemos um autovalor $\widehat{\lambda} = 1$ com multiplicidade algébrica $p = n - m$ da matriz H.

Portanto, mostramos que H é uma matriz indefinida, e determinamos seus autovalores.

Exemplo 6.7.8 *Sejam* $A \in I\!M_n(I\!R)$ *uma matriz positiva–definida e* $B \in I\!M_{m\times n}(I\!R)$*, com* $m \leq n$ *e* $posto(B) = m$*. Mostre que a matriz simétrica* H*, de ordem* $n + m$*, dada por:*

$$H = \begin{bmatrix} A & B^t \\ B & 0_m \end{bmatrix}$$

é uma matriz indefinida.

Como $A \in I\!M_n(I\!R)$ é uma matriz positiva–definida, sabemos que $A = Q\Lambda Q^t$, onde

$$\Lambda = diag(\lambda_1 , \ldots , \lambda_j , \ldots , \lambda_n)$$

é uma matriz diagonal real, onde os autovalores $\lambda_1 , \ldots , \lambda_j , \ldots , \lambda_n$ são todos positivos, e $Q \in I\!M_n(I\!R)$ é uma matriz ortogonal. Desse modo, pelo Exemplo 6.7.5, sabemos que a matriz positiva–definida A é congruente com a matriz identidade, isto é,

$$A = (QD)\,\widehat{\Lambda}(QD)^t = P\widehat{\Lambda}P^t = P\,I_n\,P^t \;,$$

onde a matriz invertível $P = QD$ realiza a relação de congruência, e as matrizes D e $\widehat{\Lambda}$ são dadas por:

$$D = diag\left(\sqrt{\lambda_1} \,, \ldots , \sqrt{\lambda_n}\right) \qquad \text{e} \qquad \widehat{\Lambda} = diag(1 , \cdots , 1) = I_n \,.$$

Desse modo, temos que

$$I_n = P^{-1}\,A\,P^{-t} \;,$$

onde $P^{-1} = D^{-1}Q^t$ e $P^{-t} = Q\,D^{-1}$.

Considere a matriz invertível S, de ordem $n + m$, definida da seguinte forma:

$$S = \begin{bmatrix} P^{-1} & 0^t_{m\times n} \\ & \\ 0_{m\times n} & I_m \end{bmatrix} .$$

Podemos verificar facilmente que a matriz simétrica $\widehat{H} = SHS^t$, onde

$$\widehat{H} = SHS^t = \begin{bmatrix} P^{-1}\,A\,P^{-t} & P^{-1}\,B^t \\ & \\ BP^{-t} & 0_m \end{bmatrix} = \begin{bmatrix} I_n & P^{-1}\,B^t \\ & \\ BP^{-t} & 0_m \end{bmatrix} ,$$

é uma matriz indefinida, pelo resultado do Exemplo 6.7.7.

Desse modo, pelo Corolário 6.7.2, as matrizes simétricas H e $\widehat{H}$ tem a mesma inércia. Portanto, mostramos que a matriz simétrica H é uma matriz indefinida.

Diagonalização de Operadores anti–Hermitianos

Considere V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle\cdot\,,\cdot\rangle$. Seja T um operador anti–Hermitiano sobre V, isto é,

$$\langle T(u)\,, v\rangle = -\langle u\,, T(v)\rangle \quad ; \quad \forall\ u, v \in V\,.$$

Pelo Teorema 5.13.3, sabemos que a matriz $A = [T]^\beta_\beta$ é uma matriz anti–Hermitiana, onde β é uma base ortonormal de V. Assim, O problema de diagonalização de uma matriz anti–Hermitiana é equivalente ao problema de diagonalização de um operador anti–Hermitiano.

Teorema 6.7.10 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle\cdot\,,\cdot\rangle$, T um operador anti–Hermitiano sobre V e λ um autovalor de T. Então, λ é imaginário puro, isto é, $\lambda = -\overline{\lambda}$. Além disso, autovetores associados a autovalores distintos são ortogonais.*

Demonstração – Seja λ um autovalor de T com v o autovetor associado. Usando a hipótese que T é anti–Hermitiano, obtemos

$$\lambda\langle v\,, v\rangle = \langle \lambda v\,, v\rangle = \langle T(v)\,, v\rangle = -\langle v\,, T(v)\rangle = -\langle v\,, \lambda v\rangle = -\overline{\lambda}\langle v\,, v\rangle\,.$$

Logo, $(\lambda + \overline{\lambda})\langle v\,, v\rangle = 0$. Como v é não–nulo, temos que $\lambda + \overline{\lambda} = 0$, isto é, $\lambda = -\overline{\lambda}$. Portanto, mostramos que o autovalor λ é imaginário puro.

Sejam λ_1 e λ_2 autovalores distintos do operador T, com v_1 e v_2 os autovetores associados, respectivamente. Desse modo, temos que

$$\lambda_1\langle v_1\,, v_2\rangle = \langle T(v_1)\,, v_2\rangle = -\langle v_1\,, T(v_2)\rangle = -\langle v_1\,, \lambda_2 v_2\rangle = -\overline{\lambda}_2\langle v_1\,, v_2\rangle\,.$$

Portanto, tem–se que $(\lambda_1 - \lambda_2)\langle v_1\,, v_2\rangle = 0$. Logo, $\langle v_1\,, v_2\rangle = 0$, pois λ_1 e λ_2 são distintos. Assim, completamos a demonstração. ■

Teorema 6.7.11 *Sejam V um espaço vetorial complexo munido do produto interno $\langle\cdot\,,\cdot\rangle$, T um operador anti–Hermitiano sobre V e S um subespaço de V invariante sob T, isto é, $T(v) \in S$ para todo $v \in S$. Então, o subespaço $S^\perp$ é também invariante sob T.*

Demonstração – Seja $u \in S^\perp$, para todo $v \in S$ temos que

$$\langle T(u)\,, v\rangle = -\langle u\,, T(v)\rangle = 0\,.$$

Assim, provamos que $T(u) \in S^\perp$ para todo $u \in S^\perp$. ■

Teorema 6.7.12 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$, com $dim(V) = n$, e T um operador anti–Hermitiano sobre V. Então, existe uma base ortonormal para V formada de autovetores de T.*

Demonstração – A prova é feita de modo análogo ao Teorema 6.7.4. □

Tomando $\beta = \{ v_1, \cdots , v_n \}$ a base ortonormal de autovetores para o espaço V. Sabemos que a matriz do operador T com relação à base ortonormal β de autovetores é a matriz diagonal $[T]^{\beta}_{\beta} = diag(\lambda_1, \cdots , \lambda_n)$, onde v_k é o autovetor associado ao autovalor λ_k para $k = 1, \cdots , n$. Portanto, o operador T é diagonalizável.

Teorema 6.7.13 *Sejam V um espaço vetorial complexo de dimensão finita munido do produto interno $\langle \cdot , \cdot \rangle$ e T um operador anti–Hermitiano sobre V. Então, T é um operador diagonalizável.*

Demonstração – A prova segue do Teorema 6.6.6 e do Teorema 6.7.12. □

Proposição 6.7.6 *Seja $A \in I\!M_n(\mathbb{C})$ uma matriz anti–Hermitiana. Então, A é uma matriz diagonalizável, isto é, A é unitáriamente similar a uma matriz diagonal.*

Demonstração – A prova segue do Corolário 6.6.1 e do Teorema 6.7.13. □

Proposição 6.7.7 *Considere o espaço vetorial complexo $\mathbb{C}^n$ munido do produto interno usual. Sejam $A \in I\!M_n(\mathbb{C})$ uma matriz anti–Hermitiana e T_A o operador linear sobre $\mathbb{C}^n$ associado a matriz A. Então, o espaço vetorial complexo $\mathbb{C}^n$ possui uma base ortonormal de autovetores do operador linear T_A.*

Demonstração – A prova segue aplicando o Teorema 6.7.12 no operador linear T_A que é um operador anti–Hermitiano sobre $\mathbb{C}^n$. □

Exemplo 6.7.9 *Considere a matriz anti–Hermitiana* A *dada por:*

$$A = \begin{bmatrix} i & 1-i \\ -1-i & 0 \end{bmatrix}$$

para fazer uma ilustração da Proposição 6.7.6.

Podemos verificar facilmente que o polinômio característico da matriz A é dado por:

$$p(\lambda) \;=\; \det(A \,-\, \lambda\, I) \;=\; \lambda^2 \,-\, i\lambda \,+\, 2\,.$$

Podemos verificar que os autovalores da matriz A são $\lambda_1 \,=\, 2i$ e $\lambda_2 \,=\, -i$.

Os autovetores da matriz A são do tipo

$$X_1 \;=\; \begin{bmatrix} -(1+i)y \\ y \end{bmatrix} \qquad \text{e} \qquad X_2 \;=\; \begin{bmatrix} x \\ (1-i)x \end{bmatrix}$$

associados aos autovalores $\lambda_1 \,=\, 2i$ e $\lambda_2 \,=\, -i$, respectivamente.

Portanto, podemos escolher os seguintes autovetores para a matriz A

$$X_1 \;=\; \begin{bmatrix} 1+i \\ -1 \end{bmatrix} \qquad \text{e} \qquad X_2 \;=\; \begin{bmatrix} 1 \\ 1-i \end{bmatrix}$$

associados aos autovalores $\lambda_1 \,=\, 2i$ e $\lambda_2 \,=\, -i$, respectivamente.

Desse modo, podemos representar a matriz A da seguinte forma:

$$A \;=\; U\,\Lambda\,U^* \qquad \text{ou} \qquad \Lambda \;=\; U^*\,A\,U\,,$$

onde

$$U \;=\; \frac{\sqrt{3}}{3}\begin{bmatrix} 1+i & 1 \\ -1 & 1-i \end{bmatrix} \qquad \text{e} \qquad \Lambda \;=\; \begin{bmatrix} 2i & 0 \\ 0 & -i \end{bmatrix}.$$

Note que a matriz U é uma matriz unitária, isto é, $UU^* \,=\, U^*U \,=\, I$, e realiza a diagonalização da matriz A.

Exercícios

Exercício 6.88 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana. Mostre que* $x^* A x \in \mathbb{R}$ *para todo elemento* $x \in \mathbb{C}^n$.

Exercício 6.89 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana. Mostre que* SAS^* *é uma matriz Hermitiana para toda matriz* $S \in \mathbb{M}_n(\mathbb{C})$.

Exercício 6.90 *Seja* $A \in \mathbb{M}_n(\mathbb{C})$ *uma matriz Hermitiana. Mostre que existe um escalar* $\lambda \in \mathbb{R}$ *de modo que a matriz* $\lambda I + A$ *seja positiva–definida.*

Exercício 6.91 *Sejam* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador anti–Hermitiano sobre* V. *Mostre que* $\langle T(u) , u \rangle$ *é imaginário puro para todo* $u \in V$.

Exercício 6.92 *Seja* $A \in \mathbb{M}_n(\mathbb{F})$ *uma matriz diagonalizável. Pede–se:*

(a) Mostre que o $\det(A)$ *é igual ao produto de seus autovalores.*

(b) Mostre que o $tr(A)$ *é igual a soma de seus autovalores.*

Exercício 6.93 *Sejam* V *um espaço vetorial complexo munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* T *um operador positivo sobre* V. *Mostre que a aplicação*

$$\begin{aligned} p : \ V \times V &\longrightarrow \mathbb{C} \\ (u, v) &\longrightarrow p(u, v) = \langle T(u) , v \rangle \end{aligned}$$

define um produto interno no espaço vetorial complexo V.

Exercício 6.94 *Considere a matriz simétrica* $A \in \mathbb{M}_3(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Determine uma matriz ortogonal $Q \in \mathbb{M}_3(\mathbb{R})$ *que realiza a diagonalização da matriz* A*, isto é,* $\Lambda = Q^t A Q$ *é uma matriz diagonal.*

Exercício 6.95 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *positiva–definida. Mostre que* $\det(A)$ *é positivo.*

Exercício 6.96 *Considere a matriz simétrica B dada por:*

$$B = \begin{bmatrix} 2 & 4 & 4 \\ 4 & 2 & 4 \\ 4 & 4 & 2 \end{bmatrix}.$$

Determine os autovalores e os autovetores da matriz B utilizando os resultados do Exercício 6.5 e do Exercício 6.94.

Exercício 6.97 *Determine a solução do seguinte sistema dinâmico*

$$\begin{cases} x'(t) &= -3x(t) + y(t) \\ y'(t) &= x(t) - 3y(t) \end{cases}$$

com a condição inicial

$$\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Exercício 6.98 *Determine a solução do seguinte sistema dinâmico*

$$\begin{cases} x'(t) &= -5x(t) \\ y'(t) &= \quad - 4y(t) + 3z(t) \\ z'(t) &= \quad + 3y(t) - 4z(t) \end{cases}$$

com a condição inicial

$$\begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Exercício 6.99 *Considere a seguinte matriz simétrica*

$$A = \begin{bmatrix} -4 & -1 & 2 \\ -1 & -1 & -1 \\ 2 & -1 & -4 \end{bmatrix}.$$

Determine os autovalores e os autovetores da matriz A.

Exercício 6.100 *Seja $A \in \mathbb{M}_n(\mathbb{R})$ anti–simétrica. Mostre que as matriz $I - A$ e $I + A$ são invertíveis e que a matriz $(I - A)(I + A)^{-1}$ é uma matriz ortogonal.*

Exercício 6.101 *Considere a seguinte matriz simétrica*

$$A = \begin{bmatrix} 5 & -1 & 2 \\ -1 & 8 & -1 \\ 2 & -1 & 5 \end{bmatrix}.$$

Verifique se A *é uma matriz positiva–definida. Determine o* $\det(A)$ *e o* $tr(A)$*, utilizando os resultados do Exercício 6.92.*

Exercício 6.102 *Considere a matriz diagonal em blocos* $T \in I\!M_4(I\!R)$ *dada por:*

$$T = \begin{bmatrix} A & 0_2 \\ 0_2 & U \end{bmatrix},$$

onde $A \in I\!M_2(I\!R)$ *é uma matriz simétrica e* $U \in I\!M_2(I\!R)$ *é uma matriz triangular superior, representadas por:*

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \qquad e \qquad U = \begin{bmatrix} d & e \\ 0 & f \end{bmatrix},$$

com $a, b, c, d, e \in R$ *não–nulos. Determine as condições para que a matriz* T *seja diagonalizável. Justifique sua resposta.*

Exercício 6.103 *Considere a matriz diagonal em blocos* T *dada por:*

$$\begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Encontre os autovalores e os autovetores da matriz T*. A matriz* T *é diagonalizável ?*

Exercício 6.104 *Considere a matriz simétrica* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}.$$

Determine uma matriz ortogonal $Q \in I\!M_3(I\!R)$ *e uma matriz diagonal* $\Lambda \in I\!M_3(I\!R)$ *tais que* $A = Q \Lambda Q^t$*. A matriz* A *é positiva–definida?*

Exercício 6.105 *Considere a matriz simétrica* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

Determine uma matriz ortogonal $Q \in I\!M_3(I\!R)$ *e uma matriz diagonal* $\Lambda \in I\!M_3(I\!R)$ *tais que* $A = Q \Lambda Q^t$. *A matriz* A *é positiva–definida?*

Exercício 6.106 *Sejam* $A \in I\!M_n(I\!R)$ *uma matriz positiva–definida e* $C \in I\!M_n(I\!R)$ *uma matriz invertível. Mostre que a matriz* $B = CAC^t$ *é positiva–definida.*

Exercício 6.107 *Sejam* $A \in I\!M_n(I\!R)$ *uma matriz positiva–definida e* $Q \in I\!M_n(I\!R)$ *uma matriz ortogonal. Mostre que a matriz* $C = QAQ^t$ *é uma matriz positiva–definida.*

Exercício 6.108 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz simétrica. Mostre que* A *é congruente com a matriz identidade se, e somente se, todos os autovalores de* A *são positivos.*

Exercício 6.109 *Seja* $B \in I\!M_n(I\!R)$ *uma matriz invertível. Mostre que a matriz simétrica* H, *de ordem* $2n$, *definida da seguinte forma:*

$$H = \begin{bmatrix} 0_n & B^t \\ B & 0_n \end{bmatrix}$$

é uma matriz indefinida.

Exercício 6.110 *Seja* $B \in I\!M_{m \times n}(I\!R)$, *com* $m \leq n$ *e* $posto(B) = m$. *Mostre que a matriz simétrica* H, *de ordem* $n + m$, *definida da seguinte forma:*

$$H = \begin{bmatrix} 0_n & B^t \\ B & 0_m \end{bmatrix}$$

é uma matriz indefinida.

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] LUENBERGER, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} = Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/∼pulino/ALESA/

Janeiro de 2012

Conteúdo

7
Funcionais Lineares e Espaço Dual

Conteúdo

7.1 Introdução

7.2 Funcionais Lineares

Definição 7.2.1 *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$*. Um* ***Funcional Linear*** *sobre* V *é uma aplicação* $J : V \longrightarrow \mathbb{F}$ *com as seguintes propriedades:*

(a) $J(u + v) = J(u) + J(v) \quad ; \quad \forall\, u,\, v \in V$

(b) $J(\lambda u) = \lambda J(u) \quad ; \quad \forall\, u \in V$ *e* $\lambda \in \mathbb{F}$

Podemos observar facilmente que um funcional linear é uma transformação linear de V em $\mathbb{F}$, onde estamos considerando $\mathbb{F}$ como um espaço vetorial sobre o corpo $\mathbb{F}$. Assim, estamos indicando por $\mathbb{F}$ tanto o corpo como o espaço vetorial.

Exemplo 7.2.1 *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$ *e* $\beta = \{ v_1, \cdots, v_n \}$ *uma base ordenada para* V*. A aplicação*

$$
\begin{aligned}
T_i : \ V &\longrightarrow \mathbb{F} \\
u &\longrightarrow T_i(u) = \alpha_i
\end{aligned}
$$

onde α_i *é a i–ésima coordenada do elemento* u *com relação à base ordenada* β*, é um funcional linear sobre* V*.*

Exemplo 7.2.2 *Considere o espaço vetorial real* $\mathbb{M}_n(\mathbb{R})$*. A aplicação*

$$
\begin{aligned}
Tr : \ \mathbb{M}_n(\mathbb{R}) &\longrightarrow \mathbb{R} \\
A = [a_{ij}] &\longrightarrow Tr(A) = \sum_{i=1}^{n} a_{ii}
\end{aligned}
$$

que é o traço da matriz A*, é um funcional linear sobre* $\mathbb{M}_n(\mathbb{R})$*.*

Exemplo 7.2.3 *Considere o espaço vetorial real* $\mathcal{C}([a,b])$*. A aplicação*

$$
\begin{aligned}
T : \ \mathcal{C}([a,b]) &\longrightarrow \mathbb{R} \\
f &\longrightarrow T(f) = \int_a^b f(x)\, dx
\end{aligned}
$$

é um funcional linear sobre $\mathcal{C}([a,b])$*.*

Definição 7.2.2 *Seja V um espaço vetorial sobre o corpo $I\!F$ munido da norma $\|\cdot\|$. Dizemos que o funcional $J : V \longrightarrow I\!F$ é limitado, se existe uma constante $c \in I\!R$ positiva tal que*

$$|J(u)| \leq c\|u\| \quad \text{para todo} \quad u \in V.$$

Definição 7.2.3 *Seja V um espaço vetorial sobre o corpo $I\!F$ munido da norma $\|\cdot\|$. A **norma** do funcional $J : V \longrightarrow I\!F$, induzida pela norma $\|\cdot\|$, é definida por:*

$$|\!|\!| J |\!|\!| = \max\left\{ \frac{|J(u)|}{\|u\|} \quad ; \quad \|u\| \neq 0 \right\}$$

Exemplo 7.2.4 *Na Definição 7.2.3, se J é um funcional linear sobre V, podemos verificar facilmente que uma forma alternativa para a definição da norma do funcional linear J é dada por:*

$$|\!|\!| J |\!|\!| = \max\{ |J(u)| \quad ; \quad \|u\| = 1 \}$$

Podemos verificar facilmente que da Definição 7.2.3, segue que

$$|J(u)| \leq |\!|\!| J |\!|\!| \|u\| \quad \text{para todo} \quad u \in V.$$

Essa desigualdade será muito utilizada nas nossas análises.

Exemplo 7.2.5 *Seja V um espaço vetorial sobre o corpo $I\!F$ munido do produto interno $\langle \cdot, \cdot \rangle$. Considerando $\overline{v} \in V$ fixo, porém arbitrário. A aplicação definida por:*

$$\begin{aligned} T : \ V &\longrightarrow I\!F \\ u &\longrightarrow T(u) = \langle u, \overline{v} \rangle \end{aligned}$$

é um funcional linear limitado sobre V. Além disso, $|\!|\!| T |\!|\!|_2 = \|\overline{v}\|_2$.

Exemplo 7.2.6 *Considere o espaço vetorial real $\mathcal{C}([a,b])$ munido do produto interno*

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

A aplicação $J : \mathcal{C}([a,b]) \longrightarrow I\!R$ definida por:

$$J(f) = \int_a^b f(x)dx$$

é um funcional linear limitado sobre $\mathcal{C}([a,b])$. Além disso, $|\!|\!| J |\!|\!|_2 = \sqrt{b-a}$.

Exemplo 7.2.7 *Considere o espaço vetorial real* $\mathcal{C}([a,b])$ *munido da norma*

$$\| f \|_\infty = \max\{ \, | f(x) | \quad ; \quad x \in [a,b] \, \} .$$

A aplicação $J : \mathcal{C}([a,b]) \longrightarrow I\!R$ *definida por:*

$$J(f) = \int_a^b f(x) dx$$

é um funcional linear limitado sobre $\mathcal{C}([a,b])$*. Além disso,* $\| J \|_\infty = b - a$*.*

Exemplo 7.2.8 *Considere o espaço vetorial real* $\mathcal{C}([-1,1])$ *munido da norma*

$$\| f \|_\infty = \max\{ \, | f(x) | \quad ; \quad x \in [-1,1] \, \} .$$

A aplicação $J : \mathcal{C}([-1,1]) \longrightarrow I\!R$ *definida por:*

$$J(f) = \int_{-1}^0 f(x) dx - \int_0^1 f(x) dx$$

é um funcional linear limitado sobre $\mathcal{C}([-1,1])$*. Além disso,* $\| J \|_\infty = 2$*.*

Exemplo 7.2.9 *Considere o espaço vetorial real* $\mathcal{C}([a,b])$ *munido da norma*

$$\| f \|_\infty = \max\{ \, | f(x) | \quad ; \quad x \in [a,b] \, \} .$$

A aplicação $J : \mathcal{C}([a,b]) \longrightarrow I\!R$ *definida por:*

$$J(f) = f(x_0) \qquad \textit{para} \qquad x_0 \in [a,b] \ \textit{fixo}$$

é um funcional linear limitado sobre $\mathcal{C}([a,b])$*. Além disso,* $\| J \|_\infty = 1$*.*

Exemplo 7.2.10 *Sejam* V *um espaço vetorial sobre o corpo* $I\!F$ *munido da norma* $\| \cdot \|$*. A aplicação* ***norma*** $\| \cdot \| : V \longrightarrow I\!R$ *é um funcional sobre* V*, entretanto, não é linear.*

Exercícios

Exercício 7.1 *Considere o espaço vetorial real* $\mathcal{P}(I\!R)$. *Mostre que a aplicação*

$$\begin{array}{rccl} J: & \mathcal{P}(I\!R) & \longrightarrow & I\!R \\ & p(x) & \longrightarrow & J(p(x)) \;=\; 2p'(0) \;+\; p''(1) \end{array}$$

é um funcional linear sobre $\mathcal{P}(I\!R)$.

Exercício 7.2 *Considere o espaço vetorial real* $\mathcal{P}_2(I\!R)$. *Mostre que a aplicação*

$$\begin{array}{rccl} J: & \mathcal{P}_2(I\!R) & \longrightarrow & I\!R \\ & p(x) & \longrightarrow & J(p(x)) \;=\; \displaystyle\int_0^1 p(x)dx \end{array}$$

é um funcional linear sobre $\mathcal{P}_2(I\!R)$.

Exercício 7.3 *Sejam o espaço vetorial real* $I\!R^3$ *com a base ordenada*

$$\gamma \;=\; \{\, (1,0,1),\, (1,1,0),\, (0,1,1) \,\}$$

e o espaço vetorial real $I\!R$ *com a base* $\alpha \;=\; \{\, -2 \,\}$. *Considere o funcional linear* J *sobre o* $I\!R^3$ *definido por* $J(x,y,z) \;=\; x - 2y + 3z$. *Determine a representação matricial do funcional* J, *isto é, a matriz* $[J]^{\gamma}_{\alpha}$.

Exercício 7.4 *Sejam o espaço vetorial real* $\mathcal{P}_2(I\!R)$ *com a base canônica*

$$\beta \;=\; \{\, 1,\, x,\, x^2 \,\}$$

e o espaço vetorial real $I\!R$ *com a base* $\gamma \;=\; \{\, 1 \,\}$. *Considere o funcional linear*

$$\begin{array}{rccl} J: & \mathcal{P}_2(I\!R) & \longrightarrow & I\!R \\ & p(x) & \longrightarrow & J(p(x)) \;=\; \displaystyle\int_{-1}^1 p(x)dx \end{array}$$

Determine a representação matricial do funcional J, *isto é, a matriz* $[J]^{\beta}_{\gamma}$.

Exercício 7.5 *Sejam o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *com a base ordenada*

$$\gamma \;=\; \{\, 1,\, 1 \,-\, x,\, 1 \,+\, x^2 \,\}$$

e o espaço vetorial real $\mathbb{R}$ *com a base* $\alpha \,=\, \{\, 2 \,\}$*. Considere o funcional linear*

$$\begin{aligned} J :\;\; \mathcal{P}_2(\mathbb{R}) \;\; &\longrightarrow \;\; \mathbb{R} \\ p(x) \quad &\longrightarrow \;\; J(p(x)) \;=\; \int_{-1}^{1} p(x)dx \end{aligned}$$

Determine a representação matricial do funcional J*, isto é, a matriz* $[J]^{\gamma}_{\alpha}$ *.*

Exercício 7.6 *Sejam o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *com a base canônica*

$$\beta \;=\; \{\, 1,\, x,\, x^2 \,\}$$

e o espaço vetorial real $\mathbb{R}$ *com a base* $\gamma \,=\, \{\, -1 \,\}$*. Considere o funcional linear*

$$\begin{aligned} J :\;\; \mathcal{P}_2(\mathbb{R}) \;\; &\longrightarrow \;\; \mathbb{R} \\ p(x) \quad &\longrightarrow \;\; J(p(x)) \;=\; 2p'(0) \;+\; p''(1) \end{aligned}$$

Determine a representação matricial do funcional J*, isto é, a matriz* $[J]^{\beta}_{\gamma}$ *.*

Exercício 7.7 *Considere o espaço vetorial real* $\mathcal{C}([0,1])$ *munido da norma* $\|\cdot\|_{\infty}$*. Mostre que a aplicação* $J : \mathcal{C}([0,1]) \longrightarrow \mathbb{R}$ *definida por:*

$$J(f) \;\;=\;\; \int_a^b f(x)g(x)dx$$

para $g \in \mathcal{C}([0,1])$ *fixa, porém arbitrária, é um funcional linear limitado sobre* $\mathcal{C}([a,b])$ *e determine* $\|\!|\, J \,|\!\|_{\infty}$ *.*

Exercício 7.8 *Considere o espaço vetorial real* $\mathcal{C}([0,1])$ *munido da norma* $\|\cdot\|_{\infty}$*. Mostre que a aplicação* $J : \mathcal{C}([0,1]) \longrightarrow \mathbb{R}$ *definida por:*

$$J(f) \;\;=\;\; \alpha\, f(0) \;\;+\;\; \beta\, f(1) \qquad \textit{para} \qquad \alpha,\, \beta \,\in\, \mathbb{R}$$

é um funcional linear limitado sobre $\mathcal{C}([0,1])$ *e determine* $\|\!|\, J \,|\!\|_{\infty}$ *.*

Exercício 7.9 *Sejam o espaço vetorial* $\mathbb{R}^3$ *e o funcional linear* J *definido por:*

$$J(u) = 2x + y - z \qquad \text{para todo} \qquad u = (x, y, z) \in \mathbb{R}^3 .$$

Determine uma base para o subespaço $Ker(J)$.

Exercício 7.10 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$, *com* $dim(V) = n$, *e* $J : V \longrightarrow \mathbb{F}$ *um funcional linear. Quais são as possíveis dimensões do subespaço vetorial* $Ker(J)$ *?*

Exercício 7.11 *Sejam o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *e o funcional* J *definido por:*

$$J(p(x)) = \int_{-1}^{1} p(x)dx + p'(0) \qquad \text{para todo} \qquad p(x) \in \mathcal{P}_2(\mathbb{R}) .$$

Mostre que J *é um funcional linear sobre* $\mathcal{P}_2(\mathbb{R})$ *e determine uma base para* $Ker(J)$.

7.3 Espaço Dual

Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. Denotamos por $L(V, \mathbb{F})$ o conjunto de todos os funcionais lineares sobre V, isto é,

$$L(V, \mathbb{F}) \;=\; \{\; J : V \longrightarrow \mathbb{F} \;\;/\;\; J \text{ é um funcional linear } \} \,.$$

Pelo Teorema 4.6.1, sabemos que $L(V, \mathbb{F})$ é um espaço vetorial sobre o corpo $\mathbb{F}$.

Exemplo 7.3.1 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e $\beta = \{ v_1, \cdots, v_n \}$ uma base ordenada para V. O funcional linear definido por:*

$$\begin{aligned} J_i : \;\; V &\longrightarrow \mathbb{F} \\ v &\longrightarrow J_i(v) = c_i \end{aligned}$$

onde

$$[v]_\beta \;=\; \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

*é o vetor de coordenadas do elemento $v \in V$, tem um importante papel na teoria de espaço dual, como veremos a seguir. O funcional J_i é denominado i–**ésima função coordenada** com respeito à base ordenada β.*

Note que $J_i(v_j) = \delta_{ij}$, onde δ_{ij} é o delta de Kronecker, isto é,

$$\delta_{ij} \;=\; \begin{cases} 1 & \text{para} \quad i = j \\[2ex] 0 & \text{para} \quad i \neq j \end{cases}$$

Definição 7.3.1 *Seja V um espaço vetorial sobre o corpo $\mathbb{F}$. O espaço vetorial $L(V, \mathbb{F})$ sobre o corpo $\mathbb{F}$ é denominado o **espaço dual** do espaço vetorial V, que denotamos por V^*.*

Observamos no Exemplo 7.3.1, que os funcionais J_i, para $i = 1, \cdots, n$, pertencem ao espaço dual V^*. Além disso, pelo Teorema 4.2.1, existe um único funcional linear J_i, para cada i, tal que $J_i(v_j) = \delta_{ij}$. Desse modo, a partir de uma base ordenada β, obtemos um único conjunto de n funcionais distintos $J_1, \cdots, J_n$ sobre V.

Teorema 7.3.1 *Considere V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ com $\beta = \{ v_1, \cdots, v_n \}$ uma base ordenada para V. Sejam J_i a i–ésima função coordenada com respeito à base ordenada β, para $i = 1, \cdots n$, e $\beta^* = \{ J_1, \cdots, J_n \}$. Então, β^* é uma base ordenada para o espaço dual V^*, denominada* ***base dual*** *da base β. Além disso, todo funcional linear $T \in V^*$ é representado da seguinte forma:*

$$T = \sum_{i=1}^{n} T(v_i) J_i$$

e cada elemento $v \in V$ é escrito como:

$$v = \sum_{i=1}^{n} J_i(v) v_i \,.$$

Assim, temos que $dim(V^*) = n$.

Demonstração – Devemos mostrar que o conjunto $\beta^* = \{ J_1, \cdots, J_n \}$ é linearmente independente em V^* e gera o espaço dual V^*.

Primeiramente, vamos mostrar que o conjunto β^* é linearmente independente no espaço dual V^*. Para isso, consideramos a seguinte combinação linear

$$T = \sum_{i=1}^{n} c_i J_i \qquad \text{para} \qquad c_i \in \mathbb{F} \,.$$

Desse modo, tomando

$$\begin{aligned} T(v_j) &= \sum_{i=1}^{n} c_i J_i(v_j) \\ &= \sum_{i=1}^{n} c_i \delta_{ij} \\ &= c_j \end{aligned}$$

para $j = 1, \cdots, n$. Em particular, se T é o funcional linear nulo, isto é, T é o elemento neutro do espaço dual V^*, temos que $T(v_j) = 0$ para cada j. Logo, os escalares c_j são todos nulos. Portanto, mostramos que o conjunto β^* é linearmente independente no espaço dual V^*.

Finalmente, vamos mostrar que β^* é um sistema de geradores para o espaço dual V^*.

Como $\beta = \{ v_1, \cdots, v_n \}$ é uma base para V, temos que todo elemento $v \in V$ é representado de modo único por:

$$v = \sum_{i=1}^{n} c_i v_i = \sum_{i=1}^{n} J_i(v) v_i .$$

Assim, dado um funcional linear $T \in V^*$, com $T(v_i) = \alpha_i$, temos que

$$T(v) = \sum_{i=1}^{n} J_i(v) T(v_i) = \sum_{i=1}^{n} \alpha_i J_i(v) \qquad \text{para todo} \qquad v \in V .$$

Desse modo, mostramos que todo funcional $T \in V^*$ é escrito de modo único como:

$$T = \sum_{i=1}^{n} \alpha_i J_i .$$

Portanto, provamos que $V^* = [J_1, \cdots, J_n]$. Logo, $\beta^* = \{ J_1, \cdots, J_n \}$ é uma base para o espaço dual V^* e $dim(V^*) = n$, o que completa a demonstração. ■

Pela observação feita logo abaixo da Definição 7.3.1, podemos concluir que existe uma única base dual β^* relativa à base ordenada β do espaço vetorial V.

Podemos observar facilmente que

$$T(v_j) = \sum_{i=1}^{n} \alpha_i J_i(v_j) = \alpha_j \qquad \text{para} \qquad j = 1, \cdots, n .$$

Assim, temos que a representação matricial do funcional linear $T \in V^*$ em relação às bases ordenadas β de V e $\gamma = \{ 1 \}$ de $\mathbb{F}$ é dada por:

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix},$$

que é uma matriz de ordem $1 \times n$. Além disso, podemos observar que

$$[T]_{\beta^*} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix},$$

é o vetor de coordenadas do funcional T com relação à base dual β^*.

Exemplo 7.3.2 *Podemos verificar facilmente que todo funcional linear sobre o espaço vetorial* $\mathbb{F}^n$*, isto é,* $T : \mathbb{F}^n \longrightarrow \mathbb{F}$*, é representado da forma:*

$$T(x_1, \cdots, x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n \qquad \textit{para todo} \qquad (x_1, \cdots, x_n) \in \mathbb{F}^n ,$$

para certos escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ *definidos por uma base ordenada de* $\mathbb{F}^n$*.*

De fato, considerando o espaço vetorial $\mathbb{F}^n$ com a base canônica $\beta = \{ e_1, \cdots, e_n \}$, temos que todo elemento $u = (x_1, \cdots, x_n) \in \mathbb{F}^n$ é escrito de modo único como:

$$u = x_1 e_1 + \cdots + x_n e_n .$$

Desse modo, temos que

$$\begin{aligned} T(x_1, \cdots, x_n) &= x_1 T(e_1) + \cdots + x_n T(e_n) \\ &= \alpha_1 x_1 + \cdots + \alpha_n x_n \end{aligned}$$

onde $\alpha_i = T(e_i) \in \mathbb{F}$ para $i = 1, \cdots, n$.

Por outro lado, dados os escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$, podemos verificar facilmente que a aplicação $T : \mathbb{F}^n \longrightarrow \mathbb{F}$ definida por:

$$T(x_1, \cdots, x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

é um funcional linear sobre $\mathbb{F}^n$, o que completa a resolução da questão.

Exemplo 7.3.3 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com a base canônica*

$$\beta = \{ (1,0,0), (0,1,0), (0,0,1) \}$$

e o funcional linear T *sobre o* $\mathbb{R}^3$ *definido por* $T(x,y,z) = 2x - 3y + z$*.*

Temos que todo elemento $u = (x,y,z) \in \mathbb{R}^3$ é escrito como:

$$\begin{aligned} (x,y,z) &= x e_1 + y e_2 + z e_3 \\ (x,y,z) &= x(1,0,0) + y(0,1,0) + z(0,0,1) . \end{aligned}$$

Desse modo, para todo $(x,y,z) \in \mathbb{R}^3$, temos que

$$T(x,y,z) = x\,T(e_1) + y\,T(e_2) + z\,T(e_3) = 2x - 3y + z .$$

Podemos verificar facilmente que a representação matricial do funcional linear T em relação às bases ordenadas β de $\mathbb{R}^3$ e $\gamma = \{ 1 \}$ de $\mathbb{R}$ é dada por:

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 2 & -3 & 1 \end{bmatrix} .$$

Exemplo 7.3.4 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, $\beta = \{ v_1, \cdots, v_n \}$ uma base ordenada para V e T um funcional linear sobre V. Podemos verificar facilmente que o funcional T é representado da forma:*

$$T(v) = \alpha_1 c_1 + \cdots + \alpha_n c_n \quad \text{para todo} \quad v \in V ,$$

para certos escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ definidos pela base β e o elemento v escrito de modo único como:

$$v = \sum_{j=1}^{n} c_j v_j .$$

De fato, aplicando o funcional T no elemento v, obtemos

$$\begin{aligned} T(v) &= c_1 T(v_1) + \cdots + c_n T(v_n) \\ &= \alpha_1 c_1 + \cdots + \alpha_n c_n \end{aligned}$$

onde $\alpha_i = T(v_i) \in \mathbb{F}$ para $i = 1, \cdots, n$.

Por outro lado, dados os escalares $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$, podemos verificar facilmente que a aplicação $T : V \longrightarrow \mathbb{F}$ definida por:

$$T(v) = \alpha_1 c_1 + \cdots + \alpha_n c_n$$

é um funcional linear sobre V.

Exemplo 7.3.5 *Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ com a base canônica*

$$\beta = \{ p_1(x) = 1 \,,\; p_2(x) = x \,,\; p_3(x) = x^2 \}$$

e o funcional linear T sobre o $\mathcal{P}_2(\mathbb{R})$ definido por $T(p(x)) = p(1) + p'(1)$.

Temos que todo polinômio $p(x) \in \mathcal{P}_2(\mathbb{R})$ é escrito como:

$$p(x) = a + bx + cx^2 \quad \text{para} \quad a, b, c \in \mathbb{R} .$$

Desse modo, para todo $p(x) \in \mathcal{P}_2(\mathbb{R})$, temos que

$$\begin{aligned} T(p(x)) &= a\,T(p_1(x)) + b\,T(p_2(x)) + c\,T(p_3(x)) \\ &= a + 2b + 3c \end{aligned}$$

onde $T(p_1(x)) = 1$, $T(p_2(x)) = 2$ e $T(p_3(x)) = 3$.

Exemplo 7.3.6 *Considere o espaço vetorial real* $I\!R^3$ *e os seguintes elementos*

$$v_1 = (1,0,1) \quad , \quad v_2 = (0,1,-2) \quad e \quad v_3 = (-1,-1,0) \,.$$

Pede–se:

(a) Considere o funcional linear T *sobre o* $I\!R^3$ *tal que*

$$T(v_1) = 1 \quad , \quad T(v_2) = -1 \quad e \quad T(v_3) = 3 \,.$$

Determine explicitamente a expressão do funcional T*.*

(b) Seja T *um funcional linear sobre o* $I\!R^3$ *tal que*

$$T(v_1) = T(v_2) = 0 \quad e \quad T(v_3) \neq 0 \,.$$

Determine explicitamente a expressão do funcional T*.*

(c) Seja T *um funcional linear sobre o* $I\!R^3$ *tal que*

$$T(v_1) = T(v_2) = 0 \quad e \quad T(v_3) \neq 0 \,.$$

Mostre que $T(2,3,-1) \neq 0$*.*

(a) Podemos verificar facilmente que $\gamma = \{ v_1, v_2, v_3 \}$ é uma base ordenada para $I\!R^3$. Assim, todo elemento $(x,y,z) \in I\!R^3$ é escrito de modo único como:

$$\begin{aligned} (x,y,z) &= a\,v_1 + b\,v_2 + c\,v_3 \\ &= a(1,0,1) + b(0,1,-2) + c(-1,-1,0) \end{aligned}$$

com $a = 2x - 2y - z$, $b = x - y - z$ e $c = x - 2y - z$.

Desse modo, temos que

$$T(x,y,z) = a\,T(v_1) + b\,T(v_2) + c\,T(v_3) = 4x - 7y - z \,.$$

para todo $(x,y,z) \in I\!R^3$.

(b) Considerando que $T(v_3) = \alpha \neq 0$ e do resultado do item **(a)**, temos que

$$T(x,y,z) = \alpha\,(x - 2y - z) \qquad \text{para todo} \qquad (x,y,z) \in I\!R^3 \,.$$

(c) Considerando o resultado do item **(b)**, temos que

$$T(2,3,-1) = -3\alpha \neq 0$$

pois $\alpha \neq 0$.

Exemplo 7.3.7 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com a base ordenada*

$$\gamma = \{ (1,0,1), (1,1,0), (0,1,1) \}$$

e o funcional linear T *sobre o* $\mathbb{R}^3$ *definido por* $T(x,y,z) = 2x - 3y + z$*.*

Vamos determinar a representação matricial do funcional linear T *em relação às bases ordenadas* γ *de* $\mathbb{R}^3$ *e* $\alpha = \{2\}$ *de* $\mathbb{R}$*.*

Inicialmente vamos obter a representação de $T(v_j)$, onde v_j são os elementos da base ordenada γ, em relação à base $\alpha = \{2\}$ de $\mathbb{R}$

$$\begin{aligned} T(1,0,1) &= 3 = 2\,\frac{3}{2} \\ T(1,1,0) &= -1 = 2\left(-\frac{1}{2}\right) \\ T(0,1,1) &= -2 = 2\,(-1) \end{aligned}$$

Desse modo, temos que a representação matricial do funcional linear T em relação às bases ordenadas γ de $\mathbb{R}^3$ e $\alpha = \{2\}$ de $\mathbb{R}$ é dada por:

$$[T]^{\gamma}_{\alpha} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} & -1 \end{bmatrix}.$$

Note que todo elemento $(x,y,z) \in \mathbb{R}^3$ é escrito de modo único como:

$$\begin{aligned} (x,y,z) &= a\,v_1 + b\,v_2 + c\,v_3 \\ &= a(1,0,1) + b(1,1,0) + c(0,1,1) \end{aligned}$$

com

$$a = \frac{x - y + z}{2} \quad , \quad b = \frac{x + y - z}{2} \quad \text{e} \quad c = \frac{-x + y + z}{2}.$$

Considerando o elemento $u = (1,2,1) \in \mathbb{R}^3$ temos que

$$[u]_{\gamma} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Logo, temos que $[T(u)]_{\alpha} = [T]^{\gamma}_{\alpha}\,[u]_{\gamma}$, isto é,

$$[T(u)]_{\alpha} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = -\frac{3}{2}.$$

De fato, $T(1,2,1) = -3 = 2\left(-\dfrac{3}{2}\right)$.

Exemplo 7.3.8 *Considere o espaço vetorial real* $I\!R^2$ *com a base ordenada*

$$\gamma = \{ (2,1), (3,1) \}$$

e vamos denotar por $\gamma^* = \{ J_1, J_2 \}$ *a base dual. Determinar explicitamente as expressões dos funcionais* J_1 *e* J_2.

Para isso, procedemos da seguinte forma:

$$\begin{aligned} J_1(v_1) &= J_1(2e_1 + e_2) = 2J_1(e_1) + J_1(e_2) = 1 \\ J_1(v_2) &= J_1(3e_1 + e_2) = 3J_1(e_1) + J_1(e_2) = 0 \end{aligned}$$

onde $v_1 = (2,1)$, $v_2 = (3,1)$, são os elementos da base ordenada γ, e $\beta = \{ e_1, e_2 \}$ é a base canônica do $I\!R^2$.

Desse modo, obtemos $J_1(e_1) = -1$ e $J_1(e_2) = 3$. Logo, temos que

$$J_1(x,y) = -x + 3y \qquad \text{para todo} \qquad (x,y) \in I\!R^2 .$$

De modo análogo, determinamos a expressão do funcional J_2.

$$\begin{aligned} J_2(v_1) &= J_2(2e_1 + e_2) = 2J_2(e_1) + J_2(e_2) = 0 \\ J_2(v_2) &= J_2(3e_1 + e_2) = 3J_2(e_1) + J_2(e_2) = 1 \end{aligned}$$

Assim, obtemos $J_2(e_1) = 1$ e $J_2(e_2) = -2$. Logo, temos que

$$J_2(x,y) = x - 2y \qquad \text{para todo} \qquad (x,y) \in I\!R^2 .$$

Exemplo 7.3.9 *Considere o espaço vetorial real* $I\!R^2$ *com a base ordenada*

$$\gamma = \{ (1,1), (1,-1) \}$$

e vamos denotar por $\gamma^* = \{ J_1, J_2 \}$ *a base dual. De modo análogo, determinamos explicitamente as expressões dos funcionais* J_1 *e* J_2. *Assim, temos que*

$$J_1(x,y) = \frac{1}{2}(x + y) \qquad e \qquad J_2(x,y) = \frac{1}{2}(x - y)$$

para todo $(x,y) \in I\!R^2$.

Exemplo 7.3.10 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com a base canônica*

$$\beta = \{ (1,0,0), (0,1,0), (0,0,1) \}$$

e vamos denotar por $\beta^* = \{ J_1, J_2, J_3 \}$ *a base dual. De modo análogo, determinamos explicitamente as expressões dos funcionais da base dual.*

$$J_1(x,y,z) = x \quad , \quad J_2(x,y,z) = y \quad \text{e} \quad J_3(x,y,z) = z$$

para todo $(x,y,z) \in \mathbb{R}^3$*. Note que,* $J_i(e_j) = \delta_{ij}$ *para* $i, j = 1, 2, 3$.

Exemplo 7.3.11 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *com a base canônica*

$$\beta = \{ 1, x, x^2 \}$$

e vamos denotar por $\beta^* = \{ J_1, J_2, J_3 \}$ *a base dual. De modo análogo, determinamos explicitamente as expressões dos funcionais da base dual.*

Para isso, procedemos da seguinte forma:

$$\begin{aligned} J_1(p_1(x)) &= 1 \quad , \quad J_1(p_2(x)) = 0 \quad \text{e} \quad J_1(p_3(x)) = 0 \\ J_2(p_1(x)) &= 0 \quad , \quad J_2(p_2(x)) = 1 \quad \text{e} \quad J_2(p_3(x)) = 0 \\ J_3(p_1(x)) &= 0 \quad , \quad J_3(p_2(x)) = 0 \quad \text{e} \quad J_3(p_3(x)) = 1 \end{aligned}$$

onde $p_1(x) = 1$, $p_2(x) = x$ e $p_3(x) = x^2$, são os elementos da base canônica β.

Desse modo, para todo $p(x) = a + bx + cx^2 \in \mathcal{P}_2(\mathbb{R})$, temos que

$$J_1(p(x)) = a \quad , \quad J_2(p(x)) = b \quad \text{e} \quad J_3(p(x)) = c$$

Exemplo 7.3.12 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com a base ordenada*

$$\gamma = \{ (1,0,1), (1,1,0), (0,1,1) \}$$

e vamos denotar por $\gamma^* = \{ J_1, J_2, J_3 \}$ *a base dual. De modo análogo, determinamos explicitamente as expressões dos funcionais da base dual. Assim, temos que*

$$\begin{aligned} J_1(x,y,z) &= \frac{1}{2}(x - y + z) \\ J_2(x,y,z) &= \frac{1}{2}(x + y - z) \\ J_3(x,y,z) &= \frac{1}{2}(-x + y + z) \end{aligned}$$

para todo $(x,y,z) \in \mathbb{R}^3$.

Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$. Neste momento uma questão importante que surge é se toda base de V^* é a base dual de alguma base de V. Para isso, vamos fazer um rápido estudo do espaço dual de V^*, isto é, $V^{**} = (V^*)^*$.

Considere um elemento $v \in V$, arbitrário. A aplicação L_v sobre V^* dada por:

$$L_v(T) = T(v) \qquad \text{para} \qquad T \in V^*$$

é um funcional linear sobre V^*. De fato,

$$L_v(\lambda T_1 + T_2) = (\lambda T_1 + T_2)(v) = \lambda T_1(v) + T_2(v) = \lambda L_v(T_1) + L_v(T_2) ,$$

para $T_1, T_2 \in V^*$ e $\lambda \in \mathbb{F}$. Assim, temos que $L_v \in V^{**}$.

Lema 7.3.1 *Considere V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$. Seja $v \in V$ um elemento não–nulo. Então, existe um funcional linear T sobre V tal que $T(v) \neq 0$. Equivalentemente, se $L_v(T) = 0$ para todo $T \in V^*$, então $v = 0$.*

Demonstração – Como v é um elemento não–nulo de V, existe uma base ordenada $\beta = \{ v_1, \cdots, v_n \}$ de V tal que $v_1 = v$. Considerando que $\beta^* = \{ J_1, \cdots, J_n \}$ é a base dual da base β, temos que $J_1(v_1) = J_1(v) = 1 \neq 0$. ■

Teorema 7.3.2 *Considere V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ e para cada elemento $v \in V$ definimos o funcional linear L_v sobre V^* por:*

$$L_v(T) = T(v) \qquad \textit{para} \qquad T \in V^* .$$

Então, a aplicação $\varphi : V \longrightarrow V^{**}$ definida por $\varphi(v) = L_v$ é um isomorfismo.

Demonstração – Primeiramente vamos mostrar que a aplicação φ é linear. De fato, sejam $u, v \in V$ e $\lambda \in \mathbb{F}$. Chamando $w = \lambda u + v$, para todo $T \in V^*$, temos que

$$\varphi(\lambda u + v)(T) = L_w(T) = \lambda T(u) + T(v) = \lambda L_u(T) + L_v(T) .$$

Assim, temos que $\varphi(\lambda u + v) = \lambda L_u + L_v = \lambda \varphi(u) + \varphi(v)$. Logo, mostramos que a aplicação φ é uma transformação linear de V em V^{**}.

Pelo Lema 7.3.1, temos que a aplicação φ é injetora, pois $\varphi(v) = L_v$ é o funcional nulo se, e somente se, $v = 0_V$. Pelo Teorema 7.3.1, sabemos que

$$dim(V^{**}) = dim(V^*) = dim(V) .$$

Portanto, φ é sobrejetora. Logo, φ é um isomorfismo de V em V^{**}. ■

Corolário 7.3.1 *Sejam* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$ *e* V^* *o espaço dual de* V*. Então, toda base ordenada de* V^* *é a base dual de alguma base ordenada de* V*.*

Demonstração – Seja $\beta^* = \{ J_1, \cdots, J_n \}$ uma base ordenada de V^*. Sabemos, pelo Teorema 7.3.1, que existe uma base ordenada $\beta^{**} = \{ L_1, \cdots, L_n \}$ de V^{**}, que é a base dual da base β^*, isto é,

$$L_i(J_j) = \delta_{ij} .$$

Pelo Teorema 7.3.2, sabemos que os espaços vetoriais V e V^{**} são isomorfos, isto é, para cada i existe um elemento $v_i \in V$ tal que $\varphi(v_i) = L_{v_i}$, onde φ é um isomorfismo de V em V^{**}. Desse modo, temos que

$$L_i(T) = L_{v_i}(T) = T(v_i) \qquad \text{para} \qquad T \in V^* .$$

Assim, temos que $\beta = \{ v_1, \cdots, v_n \}$ é uma base ordenada para V de modo que β^* é a sua base dual, o que completa a demonstração. ■

Exemplo 7.3.13 *Considere o espaço vetorial* $\mathbb{R}^2$. *Para todo* $(x, y) \in \mathbb{R}^2$ *definimos os seguintes funcionais lineares*

$$J_1(x, y) = \frac{x - y}{2} \quad e \quad J_2(x, y) = \frac{x + y}{2}.$$

Podemos verificar facilmente que $\gamma^* = \{ J_1, J_2 \}$ *é uma base para o espaço dual* $(\mathbb{R}^2)^*$, *bastando mostra que* γ^* *é linearmente independente em* $(\mathbb{R}^2)^*$, *e que* $\gamma = \{ v_1, v_2 \}$, *onde* $v_1 = (1, -1)$ *e* $v_2 = (1, 1)$, *é a base ordenada de* $\mathbb{R}^2$ *tal que* γ^* *é a sua base dual.*

Considerando o funcional linear $J(x, y) = x - 2y$ sobre o $\mathbb{R}^2$, temos que $J(v_1) = 3$ e $J(v_2) = -1$. Logo, sabemos que

$$[J]^\gamma_\alpha = \begin{bmatrix} 3 & -1 \end{bmatrix} \quad \text{e} \quad [J]_{\gamma^*} = \begin{bmatrix} 3 \\ -1 \end{bmatrix},$$

onde estamos tomando $\alpha = \{ 1 \}$ a base de $\mathbb{R}$.

Desse modo, dado o elemento $u = (1, 3) \in \mathbb{R}^2$, temos que

$$J(u) = 3J_1(u) - J_2(u) = -3 - 2 = -5.$$

De fato, avaliando J no ponto $(1, 3) \in \mathbb{R}^2$, obtemos $J(1, 3) = -5$.

Note que, podemos obter o mesmo resultado utilizando o fato que

$$[J]_\alpha = [J]^\gamma_\alpha \, [u]_\gamma .$$

Podemos verificar facilmente que

$$[u]_\gamma = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Assim, temos que

$$[J]_\alpha = \begin{bmatrix} 3 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -5.$$

Portanto, obtemos $J(1, 3) = -5$, lembrando que $\alpha = \{ 1 \}$ é a base de $\mathbb{R}$, o que completa uma primeira exemplificação da teoria exposta anteriormente.

Exemplo 7.3.14 *Considere o espaço vetorial real* $\mathcal{P}_1(I\!R)$*. Para todo* $p(x) \in \mathcal{P}_1(I\!R)$*, definimos os funcionais lineares*

$$J_1(p(x)) = p(1) \qquad e \qquad J_2(p(x)) = p'(1) \,.$$

Verificamos facilmente que $\gamma^* = \{ J_1, J_2 \}$ *é uma base para o espaço dual* $(\mathcal{P}_1(I\!R))^*$*, bastando mostra que* γ^* *é linearmente independente em* $(\mathcal{P}_1(I\!R))^*$*. Denotando a base ordenada* $\gamma = \{ q_1(x), q_2(x) \}$ *para* $\mathcal{P}_1(I\!R)$ *de modo que* γ^* *seja sua base dual, temos que* $J_i(q_j(x)) = \delta_{ij}$*. Desse modo, obtemos*

$$q_1(x) = 1 \qquad e \qquad q_2(x) = x - 1 \,.$$

Considerando o funcional linear $J \in (\mathcal{P}_1(I\!R))^*$ definido por:

$$J(p(x)) = \int_0^1 p(x)dx \,,$$

sabemos que $J = \alpha_1 J_1 + \alpha_2 J_2$, onde

$$\alpha_1 = J(q_1) = \int_0^1 q_1(x)dx = 1 \qquad \text{e} \qquad \alpha_2 = J(q_2) = \int_0^1 q_2(x)dx = -\frac{1}{2} \,.$$

Desse modo, temos que

$$J(p(x)) = J_1(p(x)) - \frac{1}{2}J_2(p(x)) \qquad \text{para todo} \qquad p(x) \in \mathcal{P}_1(I\!R) \,.$$

Assim, o vetor de coordenadas do funcional J em relação à base dual γ^* é dado por:

$$[J]_{\gamma^*} = \begin{bmatrix} 1 \\ -\dfrac{1}{2} \end{bmatrix} .$$

Para exemplificar, consideramos o polinômio $q(x) = 2 + 3x$ e vamos calcular o valor $J(q(x))$ que é dado por:

$$J(q(x)) = J_1(q(x)) - \frac{1}{2}J_2(q(x)) = 5 - \frac{3}{2} = \frac{7}{2} \,.$$

Por outro lado, fazendo o calculo diretamente da definição do funcional J obtemos

$$J(q(x)) = \int_0^1 (2 + 3x)dx = \frac{7}{2} \,.$$

Exemplo 7.3.15 *Considere o espaço vetorial real* $\mathbb{R}^3$*. Definimos os funcionais lineares*

$$J_1(x,y,z) = x \quad , \quad J_2(x,y,z) = -x + y \quad e \quad J_3(x,y,z) = -x + y - z$$

para todo elemento $(x,y,z) \in \mathbb{R}^3$*.*

Mostre que $\gamma^* = \{ J_1, J_2, J_3 \}$ *é uma base para o espaço dual* $(\mathbb{R}^3)^*$*. Determine uma base ordenada* γ *para* $\mathbb{R}^3$ *de modo que* γ^* *seja sua base dual.*

Sabemos que a dimensão do espaço dual $(\mathbb{R}^3)^*$ é igual a três. Assim, basta mostrar que γ^* é linearmente independente. Tomando a combinação linear nula

$$aJ_1(x,y,z) + bJ_2(x,y,z) + cJ_3(x,y,z) = 0$$

e avaliando nos elementos da base canônica $\beta = \{ e_1, e_2, e_3 \}$, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a - b - c = 0 \\ \quad\; b + c = 0 \\ \quad\;\; - c = 0 \end{cases}$$

que possui somente solução trivial $a = b = c = 0$. Desse modo, provamos que γ^* é linearmente independente em $L(\mathbb{R}^3, \mathbb{R})$.

Por simplicidade, denotamos a base ordenada $\gamma = \{ v_1, v_2, v_3 \}$. Inicialmente, vamos determinar o primeiro elemento $v_1 = (a,b,c)$ que pode ser representado por:

$$v_1 = ae_1 + be_2 + ce_3 \qquad \text{para} \qquad a,\, b,\, c,\, \in \mathbb{R}^3 .$$

Sabemos que

$$\begin{aligned} J_1(v_1) &= aJ_1(e_1) + bJ_1(e_2) + cJ_1(e_3) = 1 \\ J_2(v_1) &= aJ_2(e_1) + bJ_2(e_2) + cJ_2(e_3) = 0 \\ J_3(v_1) &= aJ_3(e_1) + bJ_3(e_2) + cJ_3(e_3) = 0 \end{aligned}$$

Assim, obtemos o seguinte sistema linear

$$\begin{cases} a = 1 \\ -a + b = 0 \\ -a + b - c = 0 \end{cases}$$

que possui como única solução $a = 1$, $b = 1$ e $c = 0$. Logo, $v_1 = (1,1,0)$. De modo análogo, determinamos os elementos v_2 e v_3. Desse modo, o espaço vetorial real $\mathbb{R}^3$ tem como base ordenada $\gamma = \{ (1,1,0), (0,1,1), (0,0,-1) \}$, cuja base dual é γ^*.

Exemplo 7.3.16 *Considere o espaço vetorial* $\mathcal{P}_1(\mathbb{R})$*. Definimos os funcionais lineares*

$$J_1(p(x)) = p(-1) \qquad e \qquad J_2(p(x)) = p(1)$$

para todo polinômio $p(x) \in \mathcal{P}_1(\mathbb{R})$ *.*

Mostre que $\gamma^* = \{ J_1, J_2 \}$ *é uma base para o espaço dual* $(\mathcal{P}_1(\mathbb{R}))^*$*. Determine uma base ordenada* γ *para* $\mathcal{P}_1(\mathbb{R})$ *de modo que* γ^* *seja sua base dual.*

Sabemos que a dimensão do espaço dual $(\mathcal{P}_1(\mathbb{R}))^*$ é igual a dois. Assim, basta mostrar que γ^* é linearmente independente. Tomando a combinação linear nula

$$aJ_1(p(x)) + bJ_2(p(x)) = 0$$

e avaliando nos elementos da base canônica $\beta = \{ p_1(x) = 1 , p_2(x) = x \}$, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a + b = 0 \\ -a + b = 0 \end{cases}$$

que possui somente solução trivial $a = b = 0$. Assim, mostramos que γ^* é linearmente independente em $L(\mathcal{P}_1(\mathbb{R}), \mathbb{R})$.

Por simplicidade, denotamos a base ordenada $\gamma = \{ q_1(x), q_2(x) \}$. Primeiramente, vamos determinar o primeiro elemento q_1 que pode ser escrito da seguinte forma:

$$q_1(x) = ap_1(x) + bp_2(x) \qquad \text{para} \qquad a, b \in \mathbb{R} .$$

Sabemos que

$$J_1(q_1(x)) = aJ_1(p_1(x)) + bJ_1(p_2(x)) = 1$$

$$J_2(q_1(x)) = aJ_2(p_1(x)) + bJ_2(p_2(x)) = 0$$

Assim, obtemos o seguinte sistema linear

$$\begin{cases} a - b = 1 \\ a + b = 0 \end{cases}$$

que possui como única solução

$$a = \frac{1}{2} \qquad \text{e} \qquad b = -\frac{1}{2} .$$

Logo, temos o elemento $q_1(x) = \dfrac{1}{2}(1 - x)$.

De modo análogo, determinamos o elemento $q_2(x) = \dfrac{1}{2}(1 + x)$.

Exercícios

Exercício 7.12 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com a base ordenada*

$$\gamma = \{ (1,0,1), (1,2,1), (0,0,1) \} .$$

Determine a base dual $\gamma^* = \{ J_1, J_2, J_3 \}$ *da base ordenada* γ*.*

Exercício 7.13 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com a base ordenada*

$$\gamma = \{ (1,-1,3), (0,1,-1), (0,3,-2) \} .$$

Determine a base dual $\gamma^* = \{ J_1, J_2, J_3 \}$ *da base ordenada* γ*.*

Exercício 7.14 *Considere o espaço vetorial* $\mathcal{P}_1(\mathbb{R})$*. Definimos os funcionais lineares*

$$J_1(p(x)) = \int_0^1 p(x)dx \qquad e \qquad J_2(p(x)) = \int_0^2 p(x)dx$$

para todo polinômio $p(x) \in \mathcal{P}_1(\mathbb{R})$*. Mostre que* $\gamma^* = \{ J_1, J_2 \}$ *é uma base para o espaço dual* $(\mathcal{P}_1(\mathbb{R}))^*$*. Determine uma base ordenada* γ *para* $\mathcal{P}_1(\mathbb{R})$ *de modo que* γ^* *seja sua base dual.*

Exercício 7.15 *Considere o espaço vetorial* $\mathcal{P}_1(\mathbb{R})$*. Definimos os funcionais lineares*

$$J_1(p(x)) = \int_0^1 p(x)dx \qquad e \qquad J_2(p(x)) = \int_{-1}^1 p(x)dx$$

para todo polinômio $p(x) \in \mathcal{P}_1(\mathbb{R})$*. Mostre que* $\gamma^* = \{ J_1, J_2 \}$ *é uma base para o espaço dual* $(\mathcal{P}_1(\mathbb{R}))^*$*. Determine uma base ordenada* γ *para* $\mathcal{P}_1(\mathbb{R})$ *de modo que* γ^* *seja sua base dual.*

Exercício 7.16 *Considere o espaço vetorial real* $\mathbb{R}^3$*. Definimos os funcionais lineares*

$$J_1(x,y,z) = x - 2y \quad , \quad J_2(x,y,z) = x + y + z \qquad e \qquad J_3(x,y,z) = y - 3z$$

para todo elemento $(x,y,z) \in \mathbb{R}^3$*. Mostre que* $\gamma^* = \{ J_1, J_2, J_3 \}$ *é uma base para o espaço dual* $(\mathbb{R}^3)^*$*. Determine uma base ordenada* γ *para* $\mathbb{R}^3$ *de modo que* γ^* *seja sua base dual.*

Exercício 7.17 *Considere o espaço vetorial* $\mathcal{P}_2(\mathbb{R})$. *Definimos os funcionais lineares*

$$J_1(p(x)) \;=\; \int_0^1 p(x)dx\,, \quad J_2(p(x)) \;=\; \int_0^2 p(x)dx \quad e \quad J_3(p(x)) \;=\; \int_0^{-1} p(x)dx$$

para todo polinômio $p(x) \in \mathcal{P}_2(\mathbb{R})$. *Mostre que* $\gamma^* = \{ J_1, J_2, J_3 \}$ *é uma base para o espaço dual* $(\mathcal{P}_2(\mathbb{R}))^*$. *Determine uma base ordenada* γ *para* $\mathcal{P}_2(\mathbb{R})$ *de modo que* γ^* *seja sua base dual.*

Exercício 7.18 *Considere o espaço vetorial* $\mathcal{P}_2(\mathbb{R})$. *Definimos os funcionais lineares*

$$J_1(p(x)) \;=\; p(-1)\,, \quad J_3(p(x)) \;=\; p(0) \quad e \quad J_3(p(x)) \;=\; p(1)$$

para todo polinômio $p(x) \in \mathcal{P}_2(\mathbb{R})$. *Mostre que* $\gamma^* = \{ J_1, J_2, J_3 \}$ *é uma base para o espaço dual* $(\mathcal{P}_2(\mathbb{R}))^*$. *Determine uma base ordenada* γ *para* $\mathcal{P}_2(\mathbb{R})$ *de modo que* γ^* *seja sua base dual.*

Exercício 7.19 *Considere o espaço vetorial* $\mathcal{P}_3(\mathbb{R})$. *Definimos os funcionais lineares*

$$J_1(p(x)) \;=\; p(-1)\,, \quad J_2(p(x)) \;=\; p'(-1)\,, \quad J_3(p(x)) \;=\; p(1) \quad e \quad J_4(p(x)) \;=\; p'(1)$$

para todo polinômio $p(x) \in \mathcal{P}_3(\mathbb{R})$. *Mostre que* $\gamma^* = \{ J_1, J_2, J_3, J_4 \}$ *é uma base para o espaço dual* $(\mathcal{P}_3(\mathbb{R}))^*$. *Encontre uma base ordenada* γ *para* $\mathcal{P}_3(\mathbb{R})$ *de modo que* γ^* *seja sua base dual.*

Exercício 7.20 *Considerando os Exercícios 7.17 e 7.18 e o funcional linear*

$$J(p(x)) \;=\; \int_{-1}^1 p(x)dx \;+\; p'(0)\,,$$

determine $[J]_{\gamma^*}$ *e o valor de* $J(q(x))$ *para* $q(x) = 1 - 2x + x^2$.

Exercício 7.21 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$, $\beta = \{ v_1, v_2, v_3 \}$ *a base ordenada de* V *e* $\beta^* = \{ J_1, J_2, J_3 \}$ *a base dual da base* β. *Sabendo que*

$$[v]_\beta \;=\; \begin{bmatrix} 2a \\ a \\ -5a \end{bmatrix} \qquad e \qquad [J]_{\beta^*} \;=\; \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}\,,$$

encontre o valor de a *para que* $J(v) = -46$, *onde* $v \in V$ *e* $J \in V^*$. *Considere que* $\alpha = \{ 1 \}$ *é a base de* $\mathbb{F}$.

7.4 Teorema de Representação de Riesz

Teorema 7.4.1 (Teorema de Riesz) *Sejam* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$ *munido do produto interno* $\langle \cdot , \cdot \rangle$ *e* $J : V \longrightarrow \mathbb{F}$ *um funcional linear. Então, existe um único elemento* $\overline{v} \in V$ *de modo que o funcional linear* J *é representado da seguinte forma:*

$$J(u) = \langle u , \overline{v} \rangle \qquad \textit{para todo} \qquad u \in V .$$

Além disso, $\| J \|_2 = \| \overline{v} \|_2$.

Demonstração – Seja $\beta = \{ q_1, \cdots , q_n \}$ uma base ortonormal para V. Vamos definir o elemento $\overline{v}$ da seguinte forma:

$$\overline{v} = \sum_{j=1}^{n} \overline{J(q_j)}\, q_j .$$

Vamos considerar F um funcional linear sobre V definido por:

$$F(u) = \langle u , \overline{v} \rangle .$$

Assim, temos que

$$F(q_i) = \langle q_i , \sum_{j=1}^{n} \overline{J(q_j)}\, q_j \rangle = \sum_{j=1}^{n} J(q_j) \langle q_i , q_j \rangle = J(q_i) .$$

para todo $i = 1, \cdots , n$. Como $F(q_i) = J(q_i)$ para todo elemento da base β, temos que os funcionais F e J são os mesmos.

Agora vamos mostrar a unicidade do elemento $\overline{v} \in V$. Supomos que os elementos $\overline{v}$, $\overline{w} \in V$ satisfazem

$$J(u) = \langle u , \overline{v} \rangle \qquad \text{e} \qquad J(u) = \langle u , \overline{w} \rangle \qquad ; \qquad \forall\, u \in V .$$

Assim, temos que

$$\langle u , \overline{v} \rangle - \langle u , \overline{w} \rangle = 0 \qquad \Longrightarrow \qquad \langle u , \overline{v} - \overline{w} \rangle = 0 \qquad ; \qquad \forall\, u \in V .$$

Fazendo $u = \overline{v} - \overline{w}$, obtemos

$$\| \overline{v} - \overline{w} \|_2^2 = 0 .$$

Assim, temos que $\overline{v} - \overline{w} = 0_V$. Logo, $\overline{v} = \overline{w}$, provando a unicidade do elemento $\overline{v}$.

Finalmente, vamos mostrar que $\|\| J \|\|_2 = \| \overline{v} \|_2$. Da Definição 7.2.3, temos que

$$| J(u) | \leq \|\| J \|\|_2 \| u \|_2 \qquad \text{para todo} \qquad u \in V .$$

Fazendo $u = \overline{v}$, obtemos

$$\| \overline{v} \|_2^2 = \langle \overline{v} , \overline{v} \rangle = | J(\overline{v}) | \leq \|\| J \|\|_2 \| \overline{v} \| .$$

Considerando que J é diferente do funcional nulo, isto é, $\overline{v} \neq 0_V$, temos que

$$\| \overline{v} \|_2 \leq \|\| J \|\|_2 . \tag{7.1}$$

Da Representação de Riesz e da desigualdade de Cauchy–Schwarz, vemos que

$$| J(u) | = | \langle u , \overline{v} \rangle | \leq \| u \|_2 \| \overline{v} \|_2 . \tag{7.2}$$

Da definição da norma do funcional J e da desigualdade (7.2), segue que

$$\|\| J \|\|_2 = \max\{ | J(u)) | \quad ; \quad \| u \|_2 = 1 \} \leq \| \overline{v} \|_2 . \tag{7.3}$$

Portanto, das desigualdades (7.1) e (7.3), obtemos $\|\| J \|\|_2 = \| \overline{v} \|_2$. ■

Exemplo 7.4.1 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *. Considere o funcional linear* $J : I\!R^3 \longrightarrow I\!R$ *definido por:*

$$J(u) = 2x + y - z \qquad \textit{para todo} \qquad u = (x, y, z) \in I\!R^3 ,$$

Assim, pelo Teorema de Representação de Riesz, temos que

$$J(u) = \langle u , \overline{v} \rangle \qquad \text{para todo} \qquad u = (x, y, z) \in I\!R^3 .$$

Podemos observar que o elemento $\overline{v} = (2, 1, -1)$. Assim, temos que

$$\|\| J \|\|_2 = \| \overline{v} \|_2 = \sqrt{6} .$$

Exemplo 7.4.2 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$ com o produto interno $\langle \cdot , \cdot \rangle$ e $\beta = \{ q_1 , \cdots , q_n \}$ uma base ortonormal para V. Sabemos que todo elemento $v \in V$ é escrito de modo único como:*

$$v = \sum_{j=1}^{n} c_j \, q_j \, .$$

No Exemplo 7.3.4, mostramos que todo funcional linear T sobre V é representado da forma $T(v) = \alpha_1 c_1 + \cdots + \alpha_n c_n$ para os escalares $\alpha_i = T(q_i) \in \mathbb{F}$. Utilizando essa representação do funcional linear T em relação à base ortonormal β de V, podemos apresentar uma nova demonstração para o Teorema de Riesz.

De fato, Desejamos encontrar um elemento $\overline{v} \in V$, que é escrito de modo único como:

$$\overline{v} = \sum_{j=1}^{n} b_j \, q_j \, ,$$

tal que $T(v) = \langle v , \overline{v} \rangle$ para todo $v \in V$.

Desse modo, para todo $v \in V$, temos que

$$\begin{aligned} T(v) &= T(q_1) \, c_1 + \cdots + T(q_n) \, c_n \\ \langle v , \overline{v} \rangle &= c_1 \, \overline{b}_1 + \cdots + c_n \, \overline{b}_n \end{aligned}$$

Comparando as expressões acima, obtemos $b_j = \overline{T(q_j)}$ para $j = 1, \cdots , n$.

Assim, o elemento $\overline{v} \in V$, que estamos procurando, é escrito como:

$$\overline{v} = \sum_{j=1}^{n} \overline{T(q_j)} \, q_j$$

que é o elemento de V que realiza a representação do funcional linear T com relação ao produto interno, isto é,

$$T(v) = \langle v , \overline{v} \rangle$$

para todo $v \in V$. A prova da unicidade do elemento $\overline{v}$ e que $\| T \|_2 = \| \overline{v} \|_2$ é a mesma prova feita no Teorema 7.4.1.

Exemplo 7.4.3 *Considere o espaço vetorial real* $I\!R^n$*, com a base canônica* β*, munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *. No Exemplo 7.3.2, mostramos que todo funcional linear* T *sobre o espaço vetorial* $I\!R^n$ *é representado da forma:*

$$T(x_1, \cdots , x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n \qquad \textit{para todo} \qquad u = (x_1, \cdots , x_n) \in I\!R^n ,$$

para os escalares $\alpha_i = T(e_i) \in I\!R$*, onde* e_i *é o i–ésimo elemento da base canônica.*

Desse modo, pelo Teorema de Riesz, temos que T é representado da forma:

$$T(u) = \langle u , \overline{v} \rangle \qquad \text{para todo} \qquad u = (x_1, \cdots , x_n) \in I\!R^n ,$$

onde o elemento $\overline{v} = (\alpha_1, \cdots , \alpha_n) \in I\!R^n$. Logo, pelo Teorema de Riesz, sabemos que o elemento $\overline{v}$ é único e que $|\!|\!| T |\!|\!|_2 = \| \overline{v} \|_2$.

Exemplo 7.4.4 *Considere o espaço vetorial real* $I\!R^n$*, com a base canônica* β*, munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *e o funcional linear* $T(x, y, z) = x - 2y + 4z$ *definido sobre* $I\!R^n$*. Sabemos que* T *é representado com relação à base* β *da seguinte forma:*

$$T(u) = x\,T(e_1) + y\,T(e_2) + z\,T(e_3)$$

para todo $u = (x, y, z) \in I\!R^n$*. Temos também que o elemento*

$$\overline{v} = (T(e_1)\,, T(e_2)\,, T(e_3)) = (1, -2, 4)$$

realiza a representação do funcional linear T *com relação ao produto interno, isto é,*

$$T(u) = \langle u , \overline{v} \rangle = x - 2y + 4z$$

para todo $u = (x, y, z) \in I\!R^n$*. Assim,* $|\!|\!| T |\!|\!|_2 = \| \overline{v} \|_2 = \sqrt{21}$.

Exercícios

Exercício 7.22 *Considere o espaço vetorial real* $\mathcal{P}_1(\mathbb{R})$ *com o produto interno*

$$\langle\, p\,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx \qquad ; \qquad \forall\;\, p,\, q \;\in\; \mathcal{P}_1(\mathbb{R})\,.$$

Seja $T : \mathcal{P}_1(\mathbb{R}) \longrightarrow \mathbb{R}$ *o funcional linear definido por:* $T(p(x)) \;=\; p(1)$*. Determine o elemento* $q(x) \in \mathcal{P}_1(\mathbb{R})$ *tal que* $T(p(x)) \;=\; \langle\, p\,,\, q\,\rangle$ *para todo* $p(x) \in \mathcal{P}_1(\mathbb{R})$*.*

Exercício 7.23 *Considere o espaço vetorial real* $\mathcal{P}_2(\mathbb{R})$ *munido do produto interno*

$$\langle\, p\,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx$$

com a base ortogonal $\beta \;=\; \{\, q_1(x)\,,\, q_2(x)\,,\, q_3(x)\,\}$*, onde*

$$q_1(x) \;=\; 1 \quad , \quad q_2(x) \;=\; x \qquad e \qquad q_3(x) \;=\; x^2 \;-\; \frac{1}{3}\,.$$

Dado o funcional linear T *sobre o* $\mathcal{P}_2(\mathbb{R})$ *definido por:*

$$T(p(x)) \;=\; p(1) \;+\; p'(1)\,.$$

Determine o elemento $\overline{p}(x) \in \mathcal{P}_2(\mathbb{R})$ *de modo que* $T(p(x)) = \langle\, p\,,\, \overline{p}\,\rangle$ *para todo* $p(x)$ *em* $\mathcal{P}_2(\mathbb{R})$*.*

Exercício 7.24 *Considere o espaço vetorial complexo* $\mathbb{C}^2$ *munido do produto interno usual* $\langle\, \cdot\,,\, \cdot\,\rangle$ *e com a base canônica* $\beta \;=\; \{\, e_1,\, e_2\,\}$*. Dado o funcional linear* T *sobre* $\mathbb{C}^2$ *definido por:* $T(z_1, z_2) \;=\; 2z_1 \;-\; z_2$*. Determine o elemento* $\overline{v} \in \mathbb{C}^2$ *tal que* $T(u) \;=\; \langle\, u\,,\, \overline{v}\,\rangle$ *para todo* $u \in \mathbb{C}^2$*.*

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] LUENBERGER, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \quad = \quad Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t\,Q \quad = \quad \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/~pulino/ALESA/

Janeiro de 2012

Conteúdo

8

Álgebra Linear Computacional

Conteúdo

8.1 Introdução

8.2 Decomposição de Schur. Teorema Espectral

Teorema 8.2.1 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz que possui um conjunto linearmente independente de autovetores* $v_1, \cdots, v_n$ *associados aos autovalores* $\lambda_1, \cdots, \lambda_n$. *Vamos definir uma matriz diagonal* $D = diag(\lambda_1, \cdots, \lambda_n)$ *e uma matriz* $V = [v_1 \cdots v_n]$, *invertível. Então,* $V^{-1}AV = D$. *Reciprocamente, Se* $V^{-1}AV = D$, *onde* D *é uma matriz diagonal e* V *é uma matriz invertível. Então, as colunas da matriz* V *formam um conjunto linearmente independente de autovetores de* A *e os elementos da diagonal principal de* D *são os autovalores de* A.

Demonstração – Temos que $Av_i = \lambda_i v_i$ para $i = 1, \ldots, n$. Escrevendo na forma matricial, tem–se $AV = VD$. Como V é uma matriz invertível, obtemos $V^{-1}AV = D$. A prova da recíproca é feita com os argumentos de forma reversa. ■

Teorema 8.2.2 (Decomposição de Schur) *Seja* $A \in I\!M_n(\mathbb{C})$. *Então, existe uma matriz unitária* $U \in I\!M_n(\mathbb{C})$ *e uma matriz triangular superior* $T \in I\!M_n(\mathbb{C})$ *tais que* $U^*AU = T$.

Demonstração – A prova é feita por indução sobre a ordem da matriz da A. Para $n = 1$ o resultado é obtido trivialmente. Supomos que o resultado seja válido para $n = k - 1$, e vamos mostrar que é válido para $n = k$.

Seja A uma matriz de ordem k, e (λ, v) um autopar da matriz A, com $\langle v, v \rangle = 1$. Considere U_1 uma matriz unitária que possui v como sua primeira coluna, e as outras colunas consideramos como sendo o completamento para uma base ortonormal do espaço vetorial complexo $\mathbb{C}^k$. Seja W uma submatriz de ordem $k \times (k-1)$ de U_1 considerando da segunda coluna até a k–ésima coluna, isto é, $U_1 = [v \;\; W]$. Note que, $W^*v = 0$.

Definimos agora a matriz $A_1 = U_1^* A U_1$, que é representada da seguinte forma:

$$A_1 = \begin{bmatrix} v^* \\ W^* \end{bmatrix} A \begin{bmatrix} v & W \end{bmatrix} = \begin{bmatrix} v^* A v & v^* A W \\ W^* A v & W^* A W \end{bmatrix}$$

Como $Av = \lambda v$, obtemos $v^*Av = \lambda$ e $W^* A v = \lambda W^* v = 0$.

Assim, tomando a matriz $\widehat{A} = W^* A W$, de ordem $(k-1)$, e o elemento $w = W^* A^* v$, temos que a matriz A_1 tem a seguinte forma:

$$A_1 = \begin{bmatrix} \lambda & w^* \\ 0 & \widehat{A} \end{bmatrix}$$

Pela hipótese de indução, temos que existe uma matriz unitária $\widehat{U}_2$ e uma matriz triangular superior $\widehat{T}$, de ordem $(k-1)$, tais que $\widehat{T} = \widehat{U}_2^* \widehat{A} \widehat{U}_2$. Definimos uma matriz U_2, de ordem k, da seguinte forma:

$$U_2 = \begin{bmatrix} 1 & 0^t \\ 0 & \widehat{U}_2 \end{bmatrix}.$$

Desse modo, temos que U_2 é uma matriz unitária e $U_2^* A_1 U_2 = T$ é uma matriz triangular superior, que tem a seguinte forma:

$$T = \begin{bmatrix} \lambda & w^* \widehat{U}_2 \\ 0 & \widehat{T} \end{bmatrix}.$$

Fazendo $U = U_1 U_2$, e utilizando a expressão da matriz A_1, obtemos

$$T = U_2^* A_1 U_2 = (U_1 U_2)^* A (U_1 U_2) = U^* A U,$$

o que completa a demonstração. ■

Note que os elementos da diagonal principal da matriz T são os autovalores da matriz A, de acordo com o Teorema 6.2.1. O Teorema de Schur mostra que podemos construir uma matriz triangular superior similar a matriz A através de transformações unitárias, obtendo assim os seus autovalores.

Teorema 8.2.3 (Teorema Espectral) *Seja $A \in \mathbb{M}_n(\mathbb{C})$ uma matriz Hermitiana. Então, existe uma matriz unitária $U \in \mathbb{M}_n(\mathbb{R})$ e uma matriz diagonal $D \in \mathbb{M}_n(\mathbb{R})$ tais que $U^* A U = D$. Além disso, as colunas da matriz U são os autovetores de A e os elementos da diagonal de D são os autovalores de A.*

Demonstração – A prova segue do Teorema de Schur. De fato, sabemos que existe uma matriz unitária U e uma matriz triangular superior T tais que $T = U^* A U$. Como A é uma matriz Hermitiana temos que $T^* = U^* A U$. Logo, tem–se $T^* = T$, implicando que T é uma matriz diagonal real. A segunda parte do teorema segue imediatamente do Teorema 8.2.1, o que completa a demonstração. ■

Teorema 8.2.4 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz simétrica. Então, existe uma matriz ortogonal* $Q \in \mathbb{M}_n(\mathbb{R})$ *e uma matriz diagonal* $D \in \mathbb{M}_n(\mathbb{R})$ *tais que* $Q^t A Q = D$. *Além disso, as colunas da matriz* Q *são os autovetores de* A *e os elementos da diagonal de* D *são os autovalores de* A.

Demonstração – A prova é feita seguindo a demonstração do Teorema de Schur, que é o caso complexo, observando que para matriz simétrica devemos fazer as construções considerando o espaço vetorial real $\mathbb{R}^n$. A prova é feita por indução sobre a ordem da matriz A.

É fácil ver que o resultado é valido para matrizes de ordem 1. Supomos que o resultado seja válido para matrizes de ordem $(n-1)$, para $n \geq 2$, e vamos mostrar que o resultado é valido para matrizes de ordem n.

Como A é uma matriz simétrica real, sabemos que qualquer autovalor λ é real, $v \in \mathbb{R}^n$ o autovetor de A associado, escolhido de modo que $\langle v , v \rangle = 1$. Considere Q_1 uma matriz ortogonal com v sua primeira coluna, e as outras colunas consideramos como sendo o completamento para uma base ortonormal do espaço vetorial real $\mathbb{R}^n$, que vamos representar por $Q_1 = [v \;\; W]$.

Definimos agora uma matriz $A_1 = Q_1^t \, A \, Q_1$, que é uma matriz simétrica real, e podemos representa–la da seguinte forma:

$$A_1 = \begin{bmatrix} v^t \\ W^t \end{bmatrix} A \begin{bmatrix} v & W \end{bmatrix} = \begin{bmatrix} v^t \, A \, v & (\, W^t \, A \, v \,)^t \\ W^t \, A \, v & W^t \, A \, W \end{bmatrix}$$

Como $Av = \lambda v$, obtemos $v^t A v = \lambda$ e $W^t A v = \lambda W^t v = 0$.

Tomando a matriz $\widehat{A} = W^t \, A \, W$, de ordem $(n-1)$, temos que a matriz A_1 possui a seguinte forma:

$$A_1 = \begin{bmatrix} \lambda & 0^t \\ 0 & \widehat{A} \end{bmatrix}$$

Como $\widehat{A}$ é uma matriz simétrica real, de ordem $(n-1)$, pela hipótese de indução, temos que existe uma matriz ortogonal $\widehat{Q}_2$ e uma matriz diagonal $\widehat{D}$, de ordem $(n-1)$, tais que $\widehat{D} = \widehat{Q}_2^t \, \widehat{A} \, \widehat{Q}_2$.

Assim, podemos definir uma matriz ortogonal Q_2, de ordem n, representada na forma:

$$Q_2 = \begin{bmatrix} 1 & 0^t \\ 0 & \widehat{Q}_2 \end{bmatrix} .$$

Desse modo, temos que $Q_2^t \, A_1 \, Q_2 = D$ é uma matriz diagonal, que tem a forma:

$$D = \begin{bmatrix} \lambda & 0^t \\ 0 & \widehat{D} \end{bmatrix} .$$

Fazendo $Q = Q_1 \, Q_2$, e utilizando a expressão da matriz A_1, obtemos

$$D = Q_2^t \, A_1 \, Q_2 = (\, Q_1 \, Q_2 \,)^t \, A \, (\, Q_1 \, Q_2 \,) = Q^t \, A \, Q \, .$$

A segunda parte do teorema segue imediatamente do Teorema 8.2.1, o que completa a demonstração. ■

Este resultado mostra que podemos obter os autovalores e autovetores de uma matriz simétrica real através de uma seqüência de transformações ortogonais.

Caracterização de Matrizes Positiva–Definidas

Teorema 8.2.5 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz simétrica. As seguintes afirmações são equivalentes*

(1) A *é positiva–definida, isto é,* $\langle Ax , x \rangle > 0$ *para todo* $x \in I\!R^n$ *não–nulo.*

(2) Os autovalores de A *são positivos.*

(3) Os autovalores das submatrizes principais A_k *são positivos.*

(4) Existe uma matriz $R \in I\!M_n(I\!R)$ *invertível tal que* $A = R^t R$*.*

Demonstração – Inicialmente vamos mostrar que a condição (*1*) implica na condição (*2*). Para isso, seja λ um autovalor de A com v o autovetor associado, considerando $\langle v , v \rangle = 1$. Assim, temos que

$$0 < \langle A v , v \rangle = \langle \lambda v , v \rangle = \lambda ,$$

mostrando que os autovalores da matriz A são positivos.

Agora, consideramos que os autovalores de uma matriz simétrica real A são positivos, e vamos mostrar que a matriz A é positiva–definida.

Para isso, tomamos um elemento $x \in I\!R^n$ não–nulo, que não seja um autovetor de A. Sabemos que como A é simétrica, possui um conjunto completo de autovetores mutuamente ortonormais. Assim, podemos escrever o elemento x de modo único como uma combinação linear desses autovetores, isto é,

$$x = c_1 v_1 + \cdots + c_n v_n .$$

Desse modo, temos que

$$\langle A x , x \rangle = \langle (c_1 \lambda_1 v_1 + \cdots + c_n \lambda_n v_n) , (c_1 v_1 + \cdots + c_n v_n) \rangle$$

Como os autovetores $v_1 , \cdots , v_n$ são mutuamente ortonormais, obtemos

$$\langle A x , x \rangle = c_1^2 \lambda_1 + \cdots + c_n^2 \lambda_n > 0$$

o que mostra que a matriz A é positiva–definida.

Assim, mostramos que a condição (*2*) implica na condição (*1*), completando a prova que as condições (*1*) e (*2*) são equivalentes.

Agora vamos mostrar que as condições (*1*) e (*3*) são equivalentes. Inicialmente, vamos considerar A uma matriz positiva–definida e mostrar que a submatriz principal A_k é também positiva–definida, para todo $k = 1, 2, \cdots, n$. Desse modo, mostramos que a condição (*1*) implica na condição (*3*). Para isso, vamos tomar um elemento $x \in I\!\!R^n$ não–nulo, com as $(n - k)$ componentes, a partir da $(k+1)$–ésima, todas nulas. Por simplicidade, vamos representar esse elemento da seguinte forma:

$$x = \begin{bmatrix} y \\ \\ 0_{r \times 1} \end{bmatrix},$$

para $y \in I\!\!R^k$ não–nulo, onde $r = n - k$. Assim, temos que

$$0 < \langle A x , x \rangle = \begin{bmatrix} y^t & 0^t_{r \times 1} \end{bmatrix} \begin{bmatrix} A_k & \star \\ \\ \star & \star \end{bmatrix} \begin{bmatrix} y \\ \\ 0_{r \times 1} \end{bmatrix} = \langle A_k y , y \rangle$$

provando que a submatriz principal A_k é positiva–definida, e que os seus autovalores são todos positivos. Mostrar que a condição (*3*) implica na condição (*1*) é imediato, pois podemos trabalhar com a própria matriz A. Assim, acabamos de mostrar que as condições (*1*) e (*3*) são equivalentes.

Finalmente, vamos mostrar que as condições (*1*) e (*4*) são equivalentes. Tomando por hipótese que A é uma matriz positiva–definida, sabemos pelo Teorema 8.2.4 que existe uma matriz ortogonal $Q \in I\!\!M_n(I\!\!R)$ e uma matriz diagonal $\Lambda = diag(\lambda_1, \cdots, \lambda_n)$ tais que $A = Q \Lambda Q^t$, onde os elementos da diagonal da matriz Λ são os autovalores da matriz A, que são todos positivos. Assim, podemos escolher a matriz $R = Q \sqrt{\Lambda} Q^t$, que é denominada **raiz quadrada positiva–definida** da matriz A.

Fica claro que podemos fazer mais de uma escolha da matriz R. Assim, uma outra forma de escolher é a seguinte $R = \sqrt{\Lambda} Q^t$, que é denominada **raiz quadrada** da matriz A. Isto mostra que a condição (*1*) implica na condição (*4*).

Considerando agora que existe uma matriz R invertível tal que $A = R^t R$, vamos mostrar que A é uma matriz positiva definida. A propriedade de simetria é imediata. Considerando um elemento $x \in I\!\!R^n$ não–nulo, temos que

$$0 < \langle R x , R x \rangle = \langle R^t R x , x \rangle = \langle A x , x \rangle$$

provando que a A é uma matriz positiva–definida. Assim mostramos que a condição (*4*) implica na condição (*1*). Desse modo, acabamos de mostrar que as condições (*1*) e (*4*) são equivalentes, o que completa a demonstração. ■

8.3 Normas Consistentes em Espaços de Matrizes

Para uma análise de convergência dos métodos numéricos para sistemas lineares, bem como para a análise de estabilidade dos esquemas de diferenças finitas necessitaremos do conceito de norma de matriz, e de suas propriedades para matrizes especiais, como por exemplo, simétrica, positiva–definida, e sua relação com os autovalores dessas matrizes. É importante observar, que no espaço vetorial real $\mathbb{R}^n$ todas as normas são equivalentes, veja referência [1]. Assim a prova de convergência de uma seqüência pode ser feita com uma norma genérica.

Definição 8.3.1 *Considere o espaço vetorial* $\mathbb{M}_n(\mathbb{R})$*. Uma* ***norma consistente*** *em* $\mathbb{M}_n(\mathbb{R})$*, é uma aplicação* $\|\cdot\|$ *que para cada matriz* A *associa o escalar* $\|A\| \in \mathbb{R}$*, com as seguintes propriedades:*

1. $\|A\| > 0 \quad com \quad \|A\| = 0 \iff A = 0$ *(Positividade)*
2. $\|\lambda A\| = |\lambda| \, \|A\| \quad para\ todo \quad \lambda \in \mathbb{F}$ *(Homogeneidade Absoluta)*
3. $\|A + B\| \leq \|A\| + \|B\|$ *(Desigualdade Triangular)*
4. $\|AB\| \leq \|A\| \, \|B\|$ *(Consistência)*

Teorema 8.3.1 *Considere o espaço vetorial* $\mathbb{M}_n(\mathbb{R})$*. A aplicação* $\|\cdot\|_F$ *definida sobre* $\mathbb{M}_n(\mathbb{R})$ *da seguinte forma:*

$$\|A\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

define uma norma consistente em $\mathbb{M}_n(\mathbb{R})$*, denominada de* ***norma de Frobenius****.*

Demonstração – Podemos observar facilmente que as propriedades *1–3* são satisfeitas, pois a norma de Frobenius pode ser vista como a norma vetorial Euclidiana.

Vamos mostrar a propriedade de consistência. Definimos a matriz $C = AB$, assim tem–se que

$$\|AB\|_F^2 = \|C\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |c_{ij}|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2$$

Aplicando a desigualdade de Cauchy–Schwarz na expressão $|\cdot|^2$, obtemos

$$\begin{aligned} \| A\,B \|_F^2 &\leq \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{k=1}^{n} | a_{ik} |^2 \sum_{k=1}^{n} | b_{kj} |^2 \right) \\ &= \left(\sum_{i=1}^{n} \sum_{k=1}^{n} | a_{ik} |^2 \right) \left(\sum_{j=1}^{n} \sum_{k=1}^{n} | b_{kj} |^2 \right) \end{aligned}$$

Portanto, mostramos que

$$\| A\,B \|_F^2 \leq \| A \|_F^2 \; \| B \|_F^2 \,,$$

o que completa a demonstração. ■

Definição 8.3.2 *Seja* $A \in I\!M_n(I\!R)$. *Definimos o* ***traço*** *da matriz* A, *que vamos denotar por* $tr(A)$, *da seguinte forma:*

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

Exemplo 8.3.1 *Mostre que o traço é uma aplicação linear de* $I\!M_n(I\!R)$ *em* $I\!R$.

Exemplo 8.3.2 *Mostre que a norma de Frobenius pode ser escrita na forma:*

$$\| A \|_F = \sqrt{tr(A^t\,A)} = \sqrt{tr(A\,A^t)} \,. \tag{8.1}$$

Teorema 8.3.2 *Sejam* $A \in I\!M_n(I\!R)$ *e* $Q_1,\ Q_2 \in I\!M_n(I\!R)$ *matrizes ortogonais. Então,*

$$\| A \|_F = \| Q_1\,A\,Q_2 \|_F \,,$$

isto é, a norma de Frobenius é invariante por transformações ortogonais.

Demonstração – O resultado é imediato utilizando a norma de Frobenius escrita em função do traço da matriz A. ■

Teorema 8.3.3 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz simétrica. Então,*

$$\| A \|_F = \sqrt{\sum_{i=1}^{n} \lambda_i^2} \ ,$$

onde $\lambda_1, \cdots, \lambda_n$ *são os autovalores da matriz* A.

Demonstração – Como A é uma matriz simétrica, sabemos que existe uma matriz diagonal $\Lambda = diag(\lambda_1, \dots, \lambda_n)$ e uma matriz ortogonal Q tais que $A = Q \Lambda Q^t$.

Assim, o resultado é obtido pelo Teorema 8.3.2, o que completa a demonstração. ■

Definição 8.3.3 *Considere o espaço vetorial* $I\!M_n(I\!R)$ *e* $\| \cdot \|_\alpha$ *uma norma em* $I\!R^n$. *Definimos uma aplicação* $||| \cdot |||_\alpha$ *sobre* $I\!M_n(I\!R)$ *da seguinte forma:*

$$||| A |||_\alpha = \max \left\{ \frac{\| Ax \|_\alpha}{\| x \|_\alpha} \quad ; \quad x \neq 0_{I\!R^n} \right\}$$

ou de modo análogo

$$||| A |||_\alpha = \max \{ \ \| Ax \|_\alpha \quad ; \quad \| x \|_\alpha = 1 \ \}$$

Teorema 8.3.4 *Sejam a norma* $\| \cdot \|_\alpha$ *em* $I\!R^n$ *e a aplicação* $||| \cdot |||_\alpha$ *em* $I\!M_n(I\!R)$. *Então, vale a seguinte desigualdade*

$$\| A x \|_\alpha \leq ||| A |||_\alpha \| x \|_\alpha \, .$$

Demonstração – Para $x = 0_{I\!R^n}$ a igualdade é trivialmente satisfeita. Para $x \neq 0_{I\!R^n}$, tem–se que

$$\frac{\| A x \|_\alpha}{\| x \|_\alpha} \leq \max \left\{ \frac{\| A z \|_\alpha}{\| z \|_\alpha} \quad ; \quad z \neq 0_{I\!R^n} \right\} = ||| A |||_\alpha$$

Portanto, obtemos $\| Ax \|_\alpha \leq ||| A |||_\alpha \| x \|_\alpha$ para $x \in I\!R^n$ não nulo. ■

Teorema 8.3.5 *A aplicação* $\||\cdot\||_\alpha$ *satisfaz as propriedades de norma consistente*

1. $\||\, A \,\||_\alpha \;>\; 0$
2. $\||\, \lambda A \,\||_\alpha \;=\; |\,\lambda\,| \;\||\, A \,\||_\alpha$
3. $\||\, A \,+\, B \,\||_\alpha \;\leq\; \||\, A \,\||_\alpha + \||\, B \,\||_\alpha$
4. $\||\, A\, B \,\||_\alpha \;\leq\; \||\, A \,\||_\alpha \;\||\, B \,\||_\alpha$

Demonstração – Podemos observar facilmente que as propriedades *1–3* são obtidas pelas propriedades *1–3* de norma vetorial, respectivamente. A propriedade *4* é obtida pelo resultado do Teorema 8.3.4. De fato,

$$\|\, A\,B\,x \,\|_\alpha \;\leq\; \||\, A \,\||_\alpha \|\, B\,x \,\|_\alpha \;\leq\; \||\, A \,\||_\alpha \;\||\, B \,\||_\alpha \|\, x \,\|_\alpha \,.$$

para todo $x \in I\!R^n$ não nulo. Desse modo, temos que

$$\||\, A\,B \,\||_\alpha \;=\; \max\left\{ \frac{\|\, A\,B\,x \,\|_\alpha}{\|\, x \,\|_\alpha} \quad ; \quad \|\, x \,\|_\alpha \;\neq\; 0_{I\!R^n} \right\} \;\leq\; \||\, A \,\||_\alpha \;\||\, B \,\||_\alpha \,,$$

o que completa a demonstração. ■

Exemplo 8.3.3 *Considere a matriz*

$$A \;=\; \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$$

Faça a representação gráfica das imagens das bolas unitárias em $I\!R^2$*, com relação às normas vetoriais* $\|\,\cdot\,\|_\infty$ *e* $\|\,\cdot\,\|_1$ *, pela transformação linear* $T_A : I\!R^2 \longrightarrow I\!R^2$ *definida pela matriz* A*. Determine* $\||\, A \,\||_\infty$ *e* $\||\, A \,\||_1$ *utilizando a representação gráfica obtida anteriormente.*

Teorema 8.3.6 *A norma em* $\mathbb{M}_n(\mathbb{R})$ *definida por:*

$$\| A \|_1 \;=\; \max\left\{ \sum_{i=1}^{n} | a_{ij} | \;\;;\;\; j \;=\; 1, \cdots n \right\}$$

é induzida (subordinada) pela norma $\| \cdot \|_1$ *em* $\mathbb{R}^n$*.*

Demonstração – Para todo $x \in \mathbb{R}^n$ não–nulo, temos que

$$\| A\, x \|_1 \;=\; \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij}\, x_j \right| \;\leq\; \sum_{i=1}^{n} \sum_{j=1}^{n} | a_{ij} |\, | x_j | \;=\; \sum_{j=1}^{n} \sum_{i=1}^{n} | a_{ij} |\, | x_j |$$

$$\leq\; \sum_{j=1}^{n} \left(\max_{1 \leq k \leq n} \left\{ \sum_{i=1}^{n} | a_{ik} | \right\} \right) | x_j | \;=\; \left(\max_{1 \leq k \leq n} \left\{ \sum_{i=1}^{n} | a_{ik} | \right\} \right) \| x \|_1$$

Assim, mostramos que

$$\frac{\| A\, x \|_1}{\| x \|_1} \;\leq\; \max_{1 \leq j \leq n} \left\{ \sum_{i=1}^{n} | a_{ij} | \right\}$$

Para obtermos o resultado desejado, temos que exibir um elemento $\overline{x} \in \mathbb{R}^n$ não–nulo que realiza a igualdade, isto é,

$$\frac{\| A\, \overline{x} \|_1}{\| \overline{x} \|_1} \;=\; \max_{1 \leq j \leq n} \left\{ \sum_{i=1}^{n} | a_{ij} | \right\}$$

Vamos supor que o máximo seja atingido na k–ésima coluna de A. Assim, basta tomar $\overline{x} = e_k$, onde e_k é o k–ésimo elemento da base canônica do $\mathbb{R}^n$. ■

Exemplo 8.3.4 *Considere a matriz* $A \in \mathbb{M}_3(\mathbb{R})$ *dada por:*

$$A \;=\; \begin{bmatrix} -2 & 3 & 2 \\ -4 & 1 & 3 \\ 4 & 5 & -2 \end{bmatrix}.$$

Temos que

$$\| A \|_1 \;=\; \max\left\{ \sum_{i=1}^{3} | a_{ij} | \;\;;\;\; j \;=\; 1, \cdots 3 \right\} \;=\; \max\{ 10 \,,\, 9 \,,\, 7 \} \;=\; 10\,.$$

Teorema 8.3.7 *A norma em* $\mathbb{M}_n(\mathbb{R})$ *definida por:*

$$\| A \|_\infty = \max \left\{ \sum_{j=1}^{n} | a_{ij} | \ ; \ i = 1, \cdots n \right\}$$

é induzida (subordinada) pela norma $\| \cdot \|_\infty$ *em* $\mathbb{R}^n$.

Demonstração – Para todo $x \in \mathbb{R}^n$ não–nulo, temos que

$$\| A x \|_\infty = \max_{1 \leq i \leq n} \left\{ \left| \sum_{j=1}^{n} a_{ij} \, x_j \right| \right\} \leq \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^{n} | a_{ij} | \, | x_j | \right\}$$

$$\leq \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^{n} | a_{ij} | \right\} \| x \|_\infty$$

Assim, mostramos que

$$\frac{\| A x \|_\infty}{\| x \|_\infty} \leq \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^{n} | a_{ij} | \right\}$$

Para obtermos o resultado desejado, temos que exibir um elemento $\overline{x} \in \mathbb{R}^n$ não–nulo que realiza a igualdade, isto é,

$$\frac{\| A \overline{x} \|_\infty}{\| \overline{x} \|_\infty} = \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^{n} | a_{ij} | \right\}$$

Vamos supor que o máximo seja atingido na k–ésima linha da matriz A.

Tomando $\overline{x}$ da seguinte forma:

$$\overline{x}_j = \begin{cases} 1 & \text{se} \quad a_{kj} \geq 0 \\ -1 & \text{se} \quad a_{kj} < 0 \end{cases} \qquad \text{para} \qquad j = 1, \ldots, n$$

temos que

$$\sum_{j=1}^{n} a_{kj} \, \overline{x}_j = \sum_{j=1}^{n} | a_{kj} | = \| A \overline{x} \|_\infty$$

pois $\| \overline{x} \|_\infty = 1$, o que completa a demonstração. ■

Exemplo 8.3.5 *Considere a matriz* $A \in \mathbb{M}_3(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} -1 & 2 & 0 \\ -3 & 1 & 2 \\ 4 & 5 & -6 \end{bmatrix}.$$

Temos que

$$\|A\|_\infty = \max\left\{ \sum_{j=1}^{3} |a_{ij}| \ ; \ i = 1, \cdots 3 \right\} = \max\{ 3, 6, 15 \} = 15.$$

Teorema 8.3.8 *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *e* $Q_1, Q_2 \in \mathbb{M}_n(\mathbb{R})$ *matrizes ortogonais. Então,*

$$\|A\|_2 = \|Q_1 A Q_2\|_2,$$

isto é, a norma $\|\cdot\|_2$ *é invariante por transformações ortogonais.*

Demonstração – A prova segue da definição de norma matricial induzida e do fato que a norma Euclidiana em $\mathbb{R}^n$ é invariante por transformação ortogonal. $\square$

Teorema 8.3.9 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz simétrica. Então,*

$$\| A \|_2 \;=\; \lambda_{max} \;=\; \max\{\, |\lambda_j| \quad ; \quad j = 1, \cdots, n \,\},$$

onde $\lambda_1, \cdots, \lambda_n$ *são os autovalores da matriz* A.

Demonstração – Como A é uma matriz simétrica, sabemos que existe uma matriz $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$ diagonal e uma matriz ortogonal Q tais que $A = Q \Lambda Q^t$. Fazendo uso da definição de norma induzida, temos que

$$\begin{aligned} \| A \|_2 &= \max\{\, \| A x \|_2 \quad ; \quad \| x \|_2 = 1 \,\} \\ &= \max\{\, \| (Q \Lambda Q^t) x \|_2 \quad ; \quad \| x \|_2 = 1 \,\} \end{aligned}$$

Chamando $Q^t x = y$ e usando o fato que $\| y \|_2 = \| x \|_2 = 1$, pois a norma Euclidiana é invariante por transformação ortogonal, temos que

$$\begin{aligned} \| A \|_2 &= \max\{\, \| Q (\Lambda y) \|_2 \quad ; \quad \| y \|_2 = 1 \,\} \\ &= \max\{\, \| \Lambda y \|_2 \quad ; \quad \| y \|_2 = 1 \,\} \end{aligned}$$

Podemos verificar facilmente que

$$\| \Lambda y \|_2^2 \;=\; \sum_{j=1}^{n} \lambda_j^2 y_j^2 \;\leq\; \lambda_{max}^2 .$$

Para obtermos o resultado desejado temos que exibir um elemento $\overline{y} \in \mathbb{R}^n$ que realiza a igualdade. Para isso, vamos supor de $\lambda_{max} = |\lambda_k|$ para algum k, com $1 \leq k \leq n$.

Desse modo, basta tomar $\overline{y} = e_k$, onde e_k é o k–ésimo elemento da base canônica do $\mathbb{R}^n$, para obtermos o resultado desejado.

Finalmente, podemos observar que

$$\begin{aligned} \| A \|_2 &= \max\{\, \| A x \|_2 \quad ; \quad \| x \|_2 = 1 \,\} \\ &= \| A q_k \|_2 \;=\; |\lambda_k| \;=\; \lambda_{max} , \end{aligned}$$

onde q_k é o autovetor associado ao autovalor λ_k, de maior valor em módulo, o que completa a demonstração. ■

É importante observar que a norma de Frobenius não é uma norma induzida por nenhuma norma em $\mathbb{R}^n$.

Teorema 8.3.10 *Seja* $A \in I\!M_n(I\!R)$. *Então,* $\| A \|_2 = \sqrt{\| A^t A \|_2}$.

Demonstração – Vamos usar a definição da norma $\| \cdot \|_2$ como ponto de partida

$$\| A \|_2 = \max\{ \| A x \|_2 \quad ; \quad \| x \|_2 = 1 \}$$

Para simplificar a análise definimos uma função auxiliar $g(x) = \| A x \|_2^2$, que pode ser escrita também como $g(x) = \langle A x , A x \rangle = \langle A^t A x , x \rangle$. Desse modo, temos que encontrar o ponto de máximo da função g na bola unitária e seu respectivo valor.

Temos que a matriz $C = A^t A$ é simétrica. Assim, sabemos que $C = Q \Lambda Q^t$, onde $\Lambda = diag(\lambda_1 , \dots , \lambda_n)$ uma matriz diagonal e $Q = [q_1 \cdots q_n]$ uma matriz ortogonal, com (λ_j , q_j) um autopar da matriz C para $j = 1, \cdots , n$.

Como a matriz C pode ser semipositiva–definida, veja Exercício 8.7, temos que seus autovalores são todos não–negativos. Assim, podemos ordena–los da seguinte forma:

$$0 \leq \lambda_{min} = \lambda_1 < \dots < \lambda_n = \lambda_{max} .$$

Desse modo, podemos escrever a função g da seguinte forma:

$$g(x) = \langle A^t A x , x \rangle = x^t C x = (Q^t x)^t \Lambda (Q^t x)$$

Fazendo a mudança de variável $y = Q^t x$, obtemos

$$g(x(y)) = y^t \Lambda y = \sum_{j=1}^{n} \lambda_j y_j^2 \leq \lambda_{max} \sum_{j=1}^{n} y_j^2 = \lambda_{max} \| y \|_2^2 = \lambda_{max}$$

Note que $\| y \|_2^2 = \| Q^t x \|_2^2 = \| x \|_2^2 = 1$.

Portanto, a função g tem seu ponto de máximo em $\overline{x} = q_n$, que corresponde a $\overline{y} = e_n$, que é o n–ésimo vetor da base canônica do $I\!R^n$, e assume o valor máximo $g(v_n) = \lambda_{max}$.

Agora voltando ao problema inicial, temos

$$\begin{aligned} \| A \|_2 &= \max\{ \| A x \|_2 \quad ; \quad \| x \|_2 = 1 \} \\ &= \max\{ \sqrt{g(x)} \quad ; \quad \| x \|_2 = 1 \} \\ &= \sqrt{\lambda_{max}} = \sqrt{\| C \|_2} \end{aligned}$$

pelo Teorema 8.3.9, o que completa a demonstração. ∎

Definição 8.3.4 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível. Definimos o* ***número de condição*** *da matriz com relação à norma* $\|\cdot\|$ *em* $\mathbb{M}_n(\mathbb{R})$ *na forma:*

$$\mathcal{K}(A) = \|A\| \; \|A^{-1}\| .$$

Podemos mostrar que para toda norma induzida $\|\cdot\|_\alpha$ em $\mathbb{M}_n(\mathbb{R})$, tem–se que

1. $\|I\|_\alpha = 1$.

2. $\mathcal{K}_\alpha(A) = \|A\|_\alpha \; \|A^{-1}\|_\alpha \geq 1$.

Teorema 8.3.11 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz simétrica invertível. Então,*

$$\mathcal{K}_2(A) = \frac{\lambda_{max}}{\lambda_{min}}$$

onde

$$\lambda_{min} = \min\{ \, |\lambda_j| \quad ; \quad j = 1, \cdots, n \,\}$$

$$\lambda_{max} = \max\{ \, |\lambda_j| \quad ; \quad j = 1, \cdots, n \,\}$$

com $\lambda_1, \cdots, \lambda_n$ *os autovalores da matriz* A.

Demonstração – Como A é uma matriz simétrica invertível, temos que A^{-1} é uma matriz simétrica.

Assim, a prova segue do fato de que se λ é um autovalor de A, então $\dfrac{1}{\lambda}$ é um autovalor de A^{-1}, pois A é invertível, e do Teorema 8.3.9. □

Exemplo 8.3.6 *Calcular o número de condição* $\mathcal{K}_2(A)$ *da seguinte matriz simétrica*

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}.$$

A matriz A *é positiva–definida ?*

Exemplo 8.3.7 *Seja* ϵ *uma constante positiva muito pequena, e considere a matriz*

$$A = \begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon \end{bmatrix}.$$

Calcule $\mathcal{K}_2(A)$ *e* $\det(A)$. *O que podemos concluir ?*

Exercícios

Exercício 8.1 *Seja* $A \in I\!M_n(I\!R)$. *Mostre que*

$$\| A \|_F^2 = \| a_1 \|_2^2 + \cdots + \| a_j \|_2^2 + \cdots + \| a_n \|_2^2 ,$$

onde $a_j \in I\!R^n$ *é a* j*–ésima coluna de* A *e* $\| \cdot \|_2$ *é a norma Euclidiana em* $I\!R^n$.

Exercício 8.2 *Seja* $A \in I\!M_n(I\!R)$. *Mostre que*

$$\| A \|_F = \sqrt{tr(A^t A)} .$$

Exercício 8.3 *Sejam* $A \in I\!M_n(I\!R)$ *e* $Q_1 , Q_2 \in I\!M_n(I\!R)$ *matrizes ortogonais. Mostre que*

$$\| A \|_F = \| Q_1 A Q_2 \|_F ,$$

isto é, a norma de Frobenius é invariante por transformação ortogonal.

Exercício 8.4 *Sejam* $A \in I\!M_n(I\!R)$ *e* $Q_1 , Q_2 \in I\!M_n(I\!R)$ *matrizes ortogonais. Mostre que*

$$\| A \|_2 = \| Q_1 A Q_2 \|_2 ,$$

isto é, a norma $\| \cdot \|_2$ *em* $I\!M_n(I\!R)$ *é invariante por transformações ortogonais.*

Exercício 8.5 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz simétrica. Mostre que*

$$\| A \|_F = \sqrt{\sum_{i=1}^{n} \lambda_i^2} ,$$

onde $\lambda_1 , \cdots , \lambda_n$ *são os autovalores da matriz* A.

Exercício 8.6 *Considere a matriz simétrica* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix} .$$

Calcule $\| A \|_2$ *e* $\| A \|_F$.

Exercício 8.7 *Considere* $A \in I\!M_n(I\!R)$. *Mostre que* $C = A^t A$ *é uma matriz semipositiva–definida se, e somente se,* A *é uma matriz singular.*

Exercício 8.8 *Considere a matriz simétrica* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} -3 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -3 \end{bmatrix}.$$

Calcule $\| A \|_2$ *e* $\| A \|_F$.

Exercício 8.9 *Considere a matriz* $A \in I\!M_3(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix},$$

Calcule $\| A \|_2$ *e* $\| A \|_F$.

Exercício 8.10 *Seja* $P \in I\!M_n(I\!R)$ *uma matriz* ***idempotente****, isto é,* $P^2 = P$*. Mostre que para qualquer norma consistente* $\| \cdot \|$ *em* $I\!M_n(I\!R)$ *tem–se que* $\| P \| \geq 1$.

Exercício 8.11 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz* ***auto–reflexiva****, isto é,* $A^2 = I$*. Mostre que* $\| A \|_F^2 \geq \sqrt{n}$.

Exercício 8.12 *Sejam* $A, B \in I\!M_n(I\!R)$ ***matrizes similares****, isto é, existe uma matriz invertível* $P \in I\!M_n(I\!R)$ *tal que* $B = P^{-1}AP$*. Mostre que*

$$tr(A) = tr(B).$$

Exercício 8.13 *Sejam* $A, B \in I\!M_n(I\!R)$ ***matrizes ortogonalmente similares****, isto é, existe uma matriz ortogonal* $Q \in I\!M_n(I\!R)$ *tal que* $B = Q^t AQ$*. Mostre que*

(a) $\| A \|_F = \| B \|_F$.

(b) $\| A \|_2 = \| B \|_2$.

Exercício 8.14 *Sejam* $D = diag(d_1, \ldots, d_n) \in I\!M_n(I\!R)$ *uma matriz diagonal e* $\| \cdot \|_p$ *a norma induzida em* $I\!M_n(I\!R)$ *pela norma* $\| \cdot \|_p$ *em* $I\!R^n$*. Mostre que*

$$\| D \|_p = \max\{ \, | d_i | \quad ; \quad i = 1, \ldots, n \, \}.$$

Exercício 8.15 *Sejam* $A \in I\!M_n(I\!R)$ *uma matriz invertível e* $\|\cdot\|$ *uma norma consistente em* $I\!M_n(I\!R)$*. Mostre que*

(a) $\mathcal{K}(A) = \mathcal{K}(A^{-1})$.

(b) $\mathcal{K}(cA) = \mathcal{K}(A)$ *para* $c \in I\!R$ *não–nulo.*

Exercício 8.16 *Seja* $\|\cdot\|_\alpha$ *uma norma induzida em* $I\!M_n(I\!R)$ *pela norma* $\|\cdot\|_\alpha$ *em* $I\!R^n$*. Mostre que*

(a) $\mathcal{K}_\alpha(I) = 1$.

(b) $\mathcal{K}_\alpha(A) \geq 1$ *para* $A \in I\!M_n(I\!R)$ *invertível.*

Exercício 8.17 *Seja* $I_n \in I\!M_n(I\!R)$ *a matriz identidade. Mostre que*

$$\mathcal{K}_F(I) = n ,$$

onde $\mathcal{K}_F$ *é o número de condição com relação à norma de Frobenius.*

Exercício 8.18 *Considere a matriz* $A \in I\!M_2(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Calcule $\|A\|_2$, $\|A^{-1}\|_2$ *e* $\mathcal{K}_2(A)$*, através dos autovalores da matriz* A*.*

Exercício 8.19 *Seja* $Q \in I\!M_n(I\!R)$ *uma matriz ortogonal. Mostre que* $\mathcal{K}_2(Q) = 1$.

Exercício 8.20 *Considere* $A \in I\!M_n(I\!R)$ *uma matriz positiva–definida e* G *o seu fator de Cholesky. Mostre que*

(a) $\|G\|_2 = \sqrt{\|A\|_2}$.

(b) $\mathcal{K}_2(G) = \sqrt{\mathcal{K}_2(A)}$.

Exercício 8.21 *Sejam* $A \in I\!M_n(I\!R)$ *uma matriz invertível e a fatoração* $A = QR$*, onde* $Q \in I\!M_n(I\!R)$ *é uma matriz ortogonal e* $R \in I\!M_n(I\!R)$ *é uma matriz triangular superior. Mostre que* $\mathcal{K}_2(A) = \mathcal{K}_2(R)$.

8.4 Análise de Sensibilidade de Sistemas Lineares

Nesta seção apresentamos alguns resultados de como o número de condição da matriz do sistema linear afeta sua solução com relação a uma perturbação tanto na matriz quanto no elemento do lado direito. Estas perturbações podem ter várias origens. Mostraremos que a pior situação depende, por exemplo, da direção do elemento do lado direito e da direção do elemento que fornece a perturbação do lado direito. Faremos uma análise para o caso particular de uma matriz simétrica, pois fica bem simples encontrar as direções que irão causar a maior perturbação na solução do sistema linear.

Sejam $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz invertível e o elemento $b \in \mathbb{R}^n$ não–nulo. Considere o **Sistema Linear**: encontrar $x^* \in \mathbb{R}^n$ solução da equação

$$Ax = b\,. \tag{8.2}$$

Como A é uma matriz invertível, tem–se

$$x^* = A^{-1}b$$

a única solução do sistema linear (8.2).

Vamos considerar que a perturbação do sistema (8.2) seja dada no elemento b. Assim, temos o seguinte **Sistema Linear Perturbado**: encontrar $\widehat{x} \in \mathbb{R}^n$ solução da equação

$$A(x + \delta x) = b + \delta b \tag{8.3}$$

onde $\delta b \in \mathbb{R}^n$ é a perturbação em b. Indicamos por $\widehat{x} = x^* + \delta x^*$ a solução do sistema perturbado (8.3), onde δx^* é a solução do sistema linear

$$A(\delta x) = \delta b \tag{8.4}$$

isto é, $\delta x^* = A^{-1}(\delta b)$ é que vai fornecer a perturbação da solução do sistema linear (8.2) em função da perturbação δb do elemento b.

Sejam $\|\cdot\|$ uma norma vetorial em $\mathbb{R}^n$ e $|\!|\!|\cdot|\!|\!|$ a norma matricial em $\mathbb{M}_n(\mathbb{R})$, induzida pela norma vetorial. Desse modo, temos que

$$\|\,\delta x^*\,\| = \|\,A^{-1}(\delta b)\,\| \leq |\!|\!|\,A^{-1}\,|\!|\!|\;\|\,\delta b\,\|\,. \tag{8.5}$$

Com uma simples manipulação em (8.5), obtemos

$$\frac{\|\,\delta x^*\,\|}{\|\,x^*\,\|} \leq \frac{|\!|\!|\,A^{-1}\,|\!|\!|\;|\!|\!|\,A\,|\!|\!|\;\|\,\delta b\,\|}{|\!|\!|\,A\,|\!|\!|\;\|\,x^*\,\|}\,. \tag{8.6}$$

Como x^* é a solução do sistema linear (8.2), temos que

$$\| b \| = \| Ax^* \| \leq \| A \| \, \| x^* \| \, . \tag{8.7}$$

Substituindo (8.7) em (8.6), e utilizando a definição do **número de condição** da matriz A, que é dado por:

$$\mathcal{K}(A) = \| A \| \; \| A^{-1} \| \, ,$$

obtemos

$$\frac{\| \delta x^* \|}{\| x^* \|} \leq \mathcal{K}(A) \frac{\| \delta b \|}{\| b \|} \, . \tag{8.8}$$

Desse modo, temos um majorante para o **erro relativo** da solução do sistema linear (8.2), quando consideramos uma perturbação no elemento b.

Direções de Máximo Erro Relativo

Vamos analisar a possibilidade de encontrar uma direção para o elemento $b \in \mathbb{R}^n$ e uma direção para o elemento de perturbação $\delta b \in \mathbb{R}^n$ de modo que tenhamos a igualdade em (8.8), isto é,

$$\frac{\| \delta x^* \|}{\| x^* \|} = \mathcal{K}(A) \frac{\| \delta b \|}{\| b \|} \, . \tag{8.9}$$

Nesta situação, concluímos que, se $\mathcal{K}(A)$ for muito grande, uma pequena perturbação no elemento b vai significar um grande erro relativo na solução do sistema linear (8.2).

Para facilitar a análise e possibilitar uma visualização geométrica do número de condição, vamos introduzir as seguintes definições:

Definição 8.4.1 (Máxima Ampliação) *Para $A \in \mathbb{M}_n(\mathbb{R})$, definimos*

$$maxmag(A) = max \left\{ \frac{\| Ax \|}{\| x \|} \quad ; \quad x \in \mathbb{R}^n \ \text{não–nulo} \right\} = \| A \|$$

é a máxima ampliação dada pela matriz A e o elemento $\overline{x} \in \mathbb{R}^n$ que satisfaz

$$maxmag(A) = \frac{\| A\overline{x} \|}{\| \overline{x} \|} = \| A \|$$

a direção de máxima ampliação dada pela matriz A.

Definição 8.4.2 (Mínima Ampliação) *Para* $A \in \mathbb{M}_n(\mathbb{R})$, *definimos*

$$minmag(A) \; = \; min \left\{ \frac{\| Ax \|}{\| x \|} \quad ; \quad x \in \mathbb{R}^n \;\; \text{não–nulo} \right\}$$

é a mínima ampliação dada pela matriz A *e o elemento* $\overline{x} \in \mathbb{R}^n$ *que satisfaz*

$$minmag(A) \; = \; \frac{\| A\overline{x} \|}{\| \overline{x} \|}$$

a direção de mínima ampliação dada pela matriz A.

Proposição 8.4.1 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível. Então,*

1. $maxmag(A) \; = \; \dfrac{1}{minmag(A^{-1})}$

2. $maxmag(A^{-1}) \; = \; \dfrac{1}{minmag(A)}$

3. $\mathcal{K}(A) \; = \; \dfrac{maxmag(A)}{minmag(A)}$

Demonstração – A prova pode ficar a cargo do leitor. □

Voltando ao problema da análise de sensibilidade do sistema linear, vamos procurar uma direção para o elemento b e uma direção para o elemento δb de modo que tenhamos a igualdade (8.9). Da equação (8.4), tem–se que

$$\| \delta x^* \| \; = \; \| A^{-1}(\delta b) \| \, .$$

Vamos analisar a máxima ampliação dada pela matriz A^{-1}, isto é,

$$|\!|\!| A^{-1} |\!|\!| \; = \; \max \left\{ \frac{\| A^{-1} z \|}{\| z \|} \quad ; \quad z \in \mathbb{R}^n \;\; \text{não–nulo} \right\} \; = \; \frac{\| A^{-1}\overline{z} \|}{\| \overline{z} \|} \, . \tag{8.10}$$

Portanto, $\overline{z}$ é a direção que realiza a máxima ampliação dada pela matriz A^{-1}.

Desse modo, escolhendo o elemento de perturbação $\delta b \; = \; \epsilon \overline{z}$, obtemos

$$\| \delta x^* \| \;\; = \;\; \| A^{-1}(\delta b) \| \;\; = \; |\!|\!| A^{-1} |\!|\!| \, \| \delta b \| \, . \tag{8.11}$$

Analisando agora a máxima ampliação dada pela matriz A, isto é,

$$\| A \| = \max \left\{ \frac{\| A\, w \|}{\| w \|} \quad ; \quad w \in I\!\!R^n \text{ não–nulo} \right\} = \frac{\| A\overline{w} \|}{\| \overline{w} \|} . \tag{8.12}$$

Portanto, $\overline{w}$ é a direção que realiza a máxima ampliação dada pela matriz A.

Escolhendo o elemento $b \in I\!\!R^n$ não–nulo, de modo que a solução x^* do sistema linear (8.2) esteja na direção de máxima ampliação dada pela matriz A, obtemos

$$\| b \| = \| Ax^* \| = \| A \| \, \| x^* \| \iff \| x^* \| = \frac{\| b \|}{\| A \|} . \tag{8.13}$$

Portanto, das equações (8.11) e (8.13), temos a igualdade procurada

$$\frac{\| \delta x^* \|}{\| x^* \|} = \mathcal{K}(A) \, \frac{\| \delta b \|}{\| b \|} . \tag{8.14}$$

A seguir apresentamos uma proposição considerando a que matriz do sistema linear é uma matriz positiva–definida, que mostra as direções que levam a igualdade (8.14). Esse resultado será muito interessante para a análise de sensibilidade dos sistemas lineares que são provenientes de discretizações de problemas de valores de contorno.

Proposição 8.4.2 *Considere o sistema linear* (8.2) *e o sistema linear perturbado* (8.3), *no caso em que* $A \in I\!\!M_n(I\!\!R)$ *é uma matriz positiva–definida. Então, a igualdade* (8.9) *é obtida para a norma–2 quando o elemento* $b = \alpha\, v_n$ *não–nulo, com* v_n *o autovetor associado ao autovalor* λ_n *de maior valor, e o elemento de perturbação* $\delta b = \epsilon\, v_1$*, com* v_1 *o autovetor associado ao autovalor* λ_1 *de menor valor.*

Demonstração – Como A é uma matriz positiva–definida, seus autovalores são todos positivos. Vamos representa–los da seguinte forma: $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, com $v_1 , v_2 , \cdots , v_n$ os autovetores associados, respectivamente.

Da equação (8.4), temos que

$$\| \delta x^* \|_2 = \| A^{-1} \, (\delta b) \|_2 . \tag{8.15}$$

Sabemos que

$$\begin{aligned} maxmag(A^{-1}) &= \max \left\{ \frac{\| A^{-1}\, z \|_2}{\| z \|_2} \quad ; \quad z \in I\!\!R^n \text{ não–nulo} \right\} \\ &= \| A^{-1} \|_2 = \frac{1}{\lambda_1} . \end{aligned} \tag{8.16}$$

Portanto, o elemento $\overline{z} = v_1$ é a direção que realiza a máxima ampliação dada pela matriz A^{-1}. Desse modo, considerando a perturbação $\delta b = \epsilon\, v_1$, temos que

$$\| \delta x^* \|_2 = \| A^{-1} \|_2 \| \delta b \|_2 = \frac{\| \delta b \|_2}{\lambda_1} \,. \tag{8.17}$$

Podemos observar facilmente que a equação (8.17) está dizendo que o erro da solução do sistema linear é ampliado pelo maior autovalor da matriz A^{-1}.

Da equação (8.2), temos que

$$\| Ax^* \|_2 = \| b \|_2 \,. \tag{8.18}$$

Sabemos que

$$maxmag(A) = \max \left\{ \frac{\| A\, z \|_2}{\| z \|_2} \quad ; \quad z \in I\!R^n \;\text{ não nulo} \right\} = \| A \|_2 = \lambda_n \,. \tag{8.19}$$

Portanto, o elemento $\overline{z} = v_n$ é a direção que realiza a máxima ampliação dada pela matriz A. Desse modo, escolhendo o elemento $b = \alpha\, v_n$ não–nulo, temos que x^* está na direção de máxima ampliação da matriz A. Assim, obtemos

$$\| x^* \|_2 = \frac{\| A\, x^* \|_2}{\| A \|_2} = \frac{\| b \|_2}{\lambda_n} \,. \tag{8.20}$$

Das equações (8.17) e (8.20), obtemos a igualdade desejada

$$\frac{\| \delta x^* \|_2}{\| x^* \|_2} = \mathcal{K}_2(A)\, \frac{\| \delta b \|_2}{\| b \|_2} \tag{8.21}$$

o que completa a demonstração. ∎

Como a matriz A é positiva–definida, das equações (8.16) e (8.19), temos que o número de condição da matriz A com relação à norma $\| \cdot \|_2$ é dado por:

$$\mathcal{K}_2(A) = \frac{\lambda_n}{\lambda_1} \,. \tag{8.22}$$

Desse modo, a igualdade (8.21) é um caso extremo do seguinte resultado:

Proposição 8.4.3 *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida e* $b \in \mathbb{R}^n$*. Então, a solução do sistema linear* $Ax = b$*,* $x^* = A^{-1}b$*, e a solução do sistema linear perturbado* $\delta x^* = A^{-1}\delta b$ *satisfazem as seguintes relações*

$$\| x^* \|_2 \;\geq\; \frac{\| b \|_2}{\lambda_n} \qquad e \qquad \| \delta x^* \|_2 \;\leq\; \frac{\| \delta b \|_2}{\lambda_1}\,, \tag{8.23}$$

e o erro relativo é limitado por

$$\frac{\| \delta x^* \|_2}{\| x^* \|_2} \;\leq\; \frac{\lambda_n}{\lambda_1}\frac{\| \delta b \|_2}{\| b \|_2} \;=\; \mathcal{K}_2(A)\frac{\| \delta b \|_2}{\| b \|_2}\,. \tag{8.24}$$

Demonstração – A prova pode ficar a cargo do leitor. □

Proposição 8.4.4 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz simétrica invertível. Considere o sistema linear* $Ax = b$ *e o sistema linear perturbado* $A\,(x + \delta x) = b + \delta b$*. Então, obtemos a igualdade*

$$\frac{\| \delta x^* \|_2}{\| x^* \|_2} \;=\; \mathcal{K}_2(A)\;\frac{\| \delta b \|_2}{\| b \|_2} \tag{8.25}$$

quando o elemento $b = \alpha\, v_n$ *não–nulo, com* v_n *o autovetor associado ao autovalor* λ_n *de maior valor em módulo, e o elemento de perturbação* $\delta b = \epsilon\, v_1$*, com* v_1 *o autovetor associado ao autovalor* λ_1 *de menor valor em módulo.*

Demonstração – A prova pode ficar a cargo do leitor. □

Exemplo 8.4.1 *Considere* $A \in \mathbb{M}_3(\mathbb{R})$ *uma matriz positiva–definida dada por:*

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Pede–se:

(a) calcule $\| A \|_2$ *e* $\| A^{-1} \|_2$ *através de seus autovalores.*

(b) Determine as direções para os elementos b *e* δb *em* $\mathbb{R}^3$ *de modo que o erro relativo da solução do sistema linear* $Ax = b$ *seja o maior possível, isto é,*

$$\frac{\| x^* - \widehat{x} \|_2}{\| x^* \|_2} = \frac{\lambda_{max}}{\lambda_{min}} \frac{\| \delta b \|_2}{\| b \|_2} \tag{8.26}$$

onde $\widehat{x} = x^* + \delta x^*$ *é a solução do sistema linear perturbado*

$$A\,(x + \delta x) = b + \delta b\,.$$

Exemplo 8.4.2 *Considere o sistema linear* $Ax = b$*, onde*

$$A = \begin{bmatrix} 1.00 & 0.99 \\ 0.99 & 0.98 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1.99 \\ 1.97 \end{bmatrix} \tag{8.27}$$

que tem como solução exata

$$x^* = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix} .$$

Pede–se:

(a) *Calcule* $\mathcal{K}_2(A)$

(b) *Encontre a solução* $\widehat{x} = x^* + \delta x^*$ *do sistema linear perturbado*

$$A\,(x + \delta x) = b + \delta b \tag{8.28}$$

considerando uma perturbação do elemento b *dada por*

$$\delta b = \begin{bmatrix} -9.7 \times 10^{-8} \\ 1.06 \times 10^{-7} \end{bmatrix} \tag{8.29}$$

(c) *Compare as grandezas*

$$\frac{\|\,\delta x^*\,\|_2}{\|\,x^*\,\|_2} \qquad e \qquad \mathcal{K}_2(A)\,\frac{\|\,\delta b\,\|_2}{\|\,b\,\|_2} \tag{8.30}$$

Calcule o ângulo entre o elemento b *e o autovetor* v_2 *associado ao autovalor* λ_2 *de maior valor em módulo.*

Calcule o ângulo entre o elemento δb *e o autovetor* v_1 *associado ao autovalor* λ_1 *de menor valor em módulo.*

O que podemos concluir ?

Os autovalores da matriz A são

$$\lambda_1 = -5.050376 \times 10^{-5} \qquad \text{e} \qquad \lambda_2 = 1.980051 \, .$$

Podemos observar que a matriz A não é positiva–definida. Entretanto, A é uma matriz simétrica invertível. Assim, temos que

$$\mathcal{K}_2(A) = \frac{\lambda_{max}}{\lambda_{min}} = 3.920601 \times 10^4 \, , \tag{8.31}$$

indicando que A é uma matriz *mal–condicionada*.

Note que a solução exata do sistema linear é

$$x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

Assim, basta resolver o sistema linear $A(\delta x) = \delta b$ para obtermos δx^*.

Encontramos

$$\delta x^* = \begin{bmatrix} 2.0 \\ -2.02030 \end{bmatrix} \times 10^{-3} \qquad \text{e} \qquad \widehat{x} = \begin{bmatrix} 1.0020 \\ 0.99797970 \end{bmatrix} .$$

Desse modo, temos que

$$\frac{\| \delta x^* \|_2}{\| x^* \|_2} = 2.010176 \times 10^{-3} \qquad \text{e} \qquad \mathcal{K}_2(A) \frac{\| \delta b \|_2}{\| b \|_2} = 2.011751 \times 10^{-3} \tag{8.32}$$

Calculando o ângulo entre os elementos b e v_2, obtemos

$$\cos(\theta_1) = \frac{\langle \, b \, , \, v_2 \, \rangle}{\| \, b \, \|_2 \, \| \, v_2 \, \|_2} = 3.1415925 \qquad \Longrightarrow \qquad \theta_1 = 180.000^{\,o} \, . \tag{8.33}$$

Calculando o ângulo entre os elementos δb e v_1, obtemos

$$\cos(\theta_2) = \frac{\langle \, \delta b \, , \, v_1 \, \rangle}{\| \, \delta b \, \|_2 \, \| \, v_1 \, \|_2} = 3.1023370 \qquad \Longrightarrow \qquad \theta_2 = 177.751^{\,o} \, . \tag{8.34}$$

Analisando os resultados apresentados em (8.32)–(8.34), podemos concluir que estamos muito próximos da pior situação.

Finalmente, considerando o sistema linear $A^{-1}b = x$ e o sistema linear perturbado $A^{-1}(b + \delta b) = (x + \delta x)$, obtemos a seguinte desigualdade

$$\frac{1}{\mathcal{K}(A)} \frac{\| \delta b \|}{\| b \|} \leq \frac{\| \delta x^* \|}{\| x^* \|} \,. \tag{8.35}$$

Desse modo, temos as seguintes desigualdades

$$\frac{1}{\mathcal{K}(A)} \frac{\| \delta b \|}{\| b \|} \leq \frac{\| \delta x^* \|}{\| x^* \|} \leq \mathcal{K}(A) \frac{\| \delta b \|}{\| b \|} \,. \tag{8.36}$$

Fixamos uma matriz A e um elemento b, ambos aleatórios. Consideramos vários erros relativos do elemento b, também aleatórios. Na Figura (8.1) temos uma ilustração para as desigualdades apresentadas em (8.36). Construímos o sistema linear (8.2) escolhendo o elemento b de modo que a solução exata seja

$$x^* = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in I\!R^n \,.$$

Nesta simulação temos $\mathcal{K}_2(A) = 80.9018$.

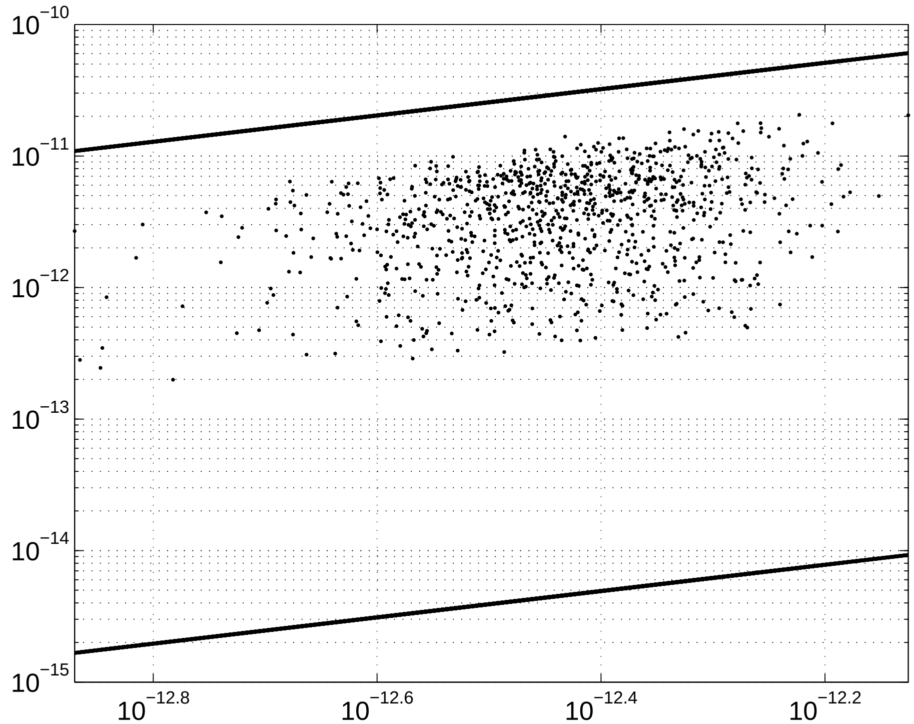

Figura 8.1: Ilustração das desigualdades definidas em (8.36). Consideramos uma matriz aleatória A, de dimensão $n = 8$, e vários erros relativos δb do elemento b, também aleatórios. Neste caso, temos que $\mathcal{K}_2(A) = 80.9018$.

Exemplo 8.4.3 *Analisar a possibilidade de determinar uma direção para o elemento* $b \in \mathbb{R}^n$ *e uma direção para a perturbação* $\delta b \in \mathbb{R}^n$ *de modo que tenhamos a igualdade em* (8.35)*, isto é,*

$$\frac{\| \delta x^* \|}{\| x^* \|} = \frac{1}{\mathcal{K}(A)} \frac{\| \delta b \|}{\| b \|} . \tag{8.37}$$

Nesta situação teremos o mínimo erro relativo na solução do sistema linear (8.2), decorrente da perturbação δb*.*

Exemplo 8.4.4 *Considere o sistema linear* $Ax = b$, *onde*

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad e \quad b = \begin{bmatrix} 0.5858 \\ 0.8284 \\ 0.5858 \end{bmatrix} \tag{8.38}$$

que tem como solução exata

$$x^* = \begin{bmatrix} 1.0 \\ 1.4142 \\ 1.0 \end{bmatrix}.$$

Pede–se:

(a) *Calcule* $\mathcal{K}_2(A)$.

(b) *Encontre a solução* $\widehat{x} = x^* + \delta x^*$ *do sistema linear perturbado*

$$A(x + \delta x) = b + \delta b \tag{8.39}$$

considerando uma perturbação do elemento b *dada por*

$$\delta b = \begin{bmatrix} 0.1250 \times 10^{-6} \\ -0.1768 \times 10^{-6} \\ 0.1250 \times 10^{-6} \end{bmatrix} \tag{8.40}$$

(c) *Compare as grandezas*

$$\frac{\| \delta x^* \|_2}{\| x^* \|_2} \quad , \quad \frac{1}{\mathcal{K}_2(A)} \frac{\| \delta b \|_2}{\| b \|_2} \quad e \quad \mathcal{K}_2(A) \frac{\| \delta b \|_2}{\| b \|_2} \tag{8.41}$$

Determine o ângulo entre o elemento b *e o autovetor* v_1 *associado ao menor autovalor* λ_1.

Determine o ângulo entre o elemento δb *e o autovetor* v_3 *associado ao maior autovalor* λ_3.

O que podemos concluir ?

Considerar agora que a perturbação do sistema (8.2) seja dada na matriz A. Assim, temos o seguinte *Sistema Linear Perturbado*: encontrar $\widehat{x} \in \mathbb{R}^n$ solução da equação

$$(A + \Delta A)(x + \delta x) = b\,, \tag{8.42}$$

onde $\Delta A \in \mathbb{M}_n(\mathbb{R})$ é uma perturbação da matriz A, de modo que $A + \Delta A$ seja uma matriz invertível. Vamos denotar por $\widehat{x} = x^* + \delta x^*$ a solução do sistema linear perturbado (8.22). Desse modo, temos que

$$x^* + \delta x^* = (A + \Delta A)^{-1}\, b\,. \tag{8.43}$$

Logo, a perturbação δx^* é obtido da seguinte forma:

$$\delta x^* = (\,(A + \Delta A)^{-1} - A^{-1}\,)\, b\,. \tag{8.44}$$

Chamando $B = A + \Delta A$, temos a seguinte identidade

$$B^{-1} - A^{-1} = A^{-1}(A - B)\, B^{-1}\,. \tag{8.45}$$

Substituindo (8.45) em (8.44), e utilizando (8.43), obtemos

$$\delta x^* = -A^{-1}\, \Delta A\, (A + \Delta A)^{-1}\, b = A^{-1}\, \Delta A\, (x^* + \delta x^*)\,. \tag{8.46}$$

Tomando a norma de ambos os membros de (8.46), tem–se que

$$\|\, \delta x^* \,\| \leq \|\!|\, A^{-1} \,|\!\| \;\; \|\!|\, \Delta A \,|\!\| \, \|\, x^* + \delta x^* \,\|\,. \tag{8.47}$$

Finalmente, obtemos

$$\frac{\|\, \delta x^* \,\|}{\|\, x^* + \delta x^* \,\|} \leq \mathcal{K}(A)\, \frac{|\!\|\, \Delta A \,|\!\|}{|\!\|\, A \,|\!\|}\,. \tag{8.48}$$

De forma análoga, podemos mostrar que

$$\frac{1}{\mathcal{K}(A)}\, \frac{|\!\|\, \Delta A \,|\!\|}{|\!\|\, A \,|\!\|} \leq \frac{\|\, \delta x^* \,\|}{\|\, x^* + \delta x^* \,\|}\,. \tag{8.49}$$

Das desigualdades (8.48) e (8.49), obtemos

$$\frac{1}{\mathcal{K}(A)}\, \frac{|\!\|\, \Delta A \,|\!\|}{|\!\|\, A \,|\!\|} \leq \frac{\|\, \delta x^* \,\|}{\|\, x^* + \delta x^* \,\|} \leq \mathcal{K}(A)\, \frac{|\!\|\, \Delta A \,|\!\|}{|\!\|\, A \,|\!\|}\,. \tag{8.50}$$

Fixamos uma matriz A e um elemento b, ambos aleatórios. Consideramos vários erros relativos da matriz A, também aleatórios. Na Figura (8.2) temos uma ilustração para as desigualdades (8.50). Construímos o sistema linear (8.2) escolhendo o elemento b de modo que a solução exata seja

$$x^* = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n .$$

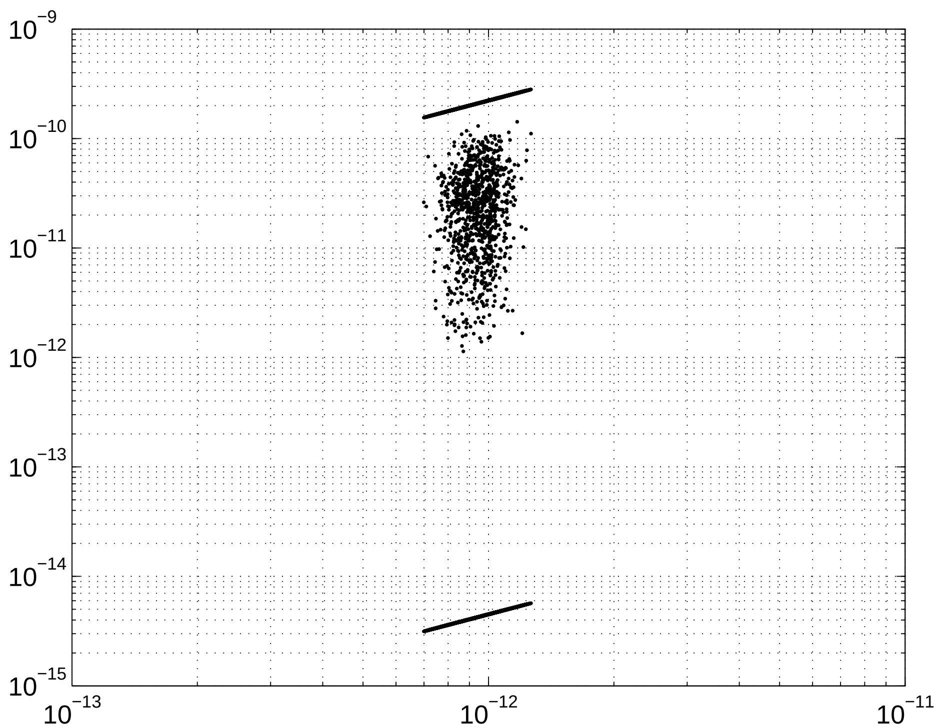

Figura 8.2: Ilustração das desigualdades definidas em (8.50). Consideramos uma matriz aleatória A, de dimensão $n = 10$, e vários erros relativos ΔA da matriz A, também aleatórios. Neste caso, temos que $\mathcal{K}_2(A) = 80.5347$.

Perturbações na Forma Paramétrica

Nesta seção apresentamos uma nova maneira para a *Análise de Sensibilidade*, onde a perturbação é dada tanto na matriz A quanto no elemento b de maneira paramétrica.

Consideramos que a perturbação no sistema linear (8.2) seja dada na matriz A e no elemento b de forma paramétrica. Assim, temos o seguinte *Sistema Linear Perturbado*: encontrar $\widehat{x}(\epsilon) \in I\!R^n$ solução da equação

$$(A + \epsilon F)x(\epsilon) = b + \epsilon f \quad ; \quad \epsilon \in I\!R\,, \tag{8.51}$$

onde $F \in I\!M_n(I\!R)$, $f \in I\!R^n$ são fixas, porém arbitrárias, e $\widehat{x}(0) = x^*$.

Como A é uma matriz invertível, temos que $\widehat{x}(\epsilon)$ é uma função diferenciável em uma vizinhança do parâmetro $\epsilon = 0$.

Derivando a equação (8.51) com relação ao parâmetro ϵ, obtemos

$$Fx(\epsilon) + (A + \epsilon F)\frac{dx}{d\epsilon}(\epsilon) = f \quad ; \quad \epsilon \in I\!R\,. \tag{8.52}$$

Derivando a equação (8.52) com relação ao parâmetro ϵ, obtemos

$$2\,F\frac{dx}{d\epsilon}(\epsilon) + (A + \epsilon F)\frac{d^2x}{d\epsilon^2}(\epsilon) = 0 \quad ; \quad \epsilon \in I\!R\,. \tag{8.53}$$

Fazendo $\epsilon = 0$ na equação (8.52) e na equação (8.53), obtemos

$$\begin{aligned} \frac{d\widehat{x}}{d\epsilon}(0) &= A^{-1}(f - Fx^*) \\ \frac{d^2\widehat{x}}{d\epsilon^2}(0) &= -2\,(A^{-1}F)\frac{d\widehat{x}}{d\epsilon}(0) \end{aligned} \tag{8.54}$$

Consideramos a Fórmula de Taylor da função $\widehat{x}(\epsilon)$ numa vizinhança de $\epsilon = 0$

$$\begin{aligned} \widehat{x}(\epsilon) &= x^* + \frac{d\widehat{x}}{d\epsilon}(0)\epsilon + \mathcal{O}(\epsilon^2) \\ \widehat{x}(\epsilon) &= x^* + A^{-1}(f - Fx^*)\epsilon + \mathcal{O}(\epsilon^2) \end{aligned} \tag{8.55}$$

Tomando a norma de ambos os membros de (8.55), tem–se que

$$\begin{aligned} \| \widehat{x}(\epsilon) - x^* \| &\leq |\epsilon| \, \| A^{-1} \| \, \{ \| F \| \| x^* \| + \| f \| \} + \mathcal{O}(\epsilon^2) \\ \frac{\| \widehat{x}(\epsilon) - x^* \|}{\| x^* \|} &\leq |\epsilon| \, \| A^{-1} \| \left\{ \| F \| + \frac{\| f \|}{\| x^* \|} \right\} + \mathcal{O}(\epsilon^2) \end{aligned} \tag{8.56}$$

Como x^* é a solução exata do sistema linear (8.2), temos que

$$\| x^* \| \geq \frac{\| b \|}{\| A \|} . \tag{8.57}$$

Utilizando a desigualdade (8.57) na desigualdade (8.56), obtemos

$$\frac{\| \widehat{x}(\epsilon) - x^* \|}{\| x^* \|} \leq \| A \| \, \| A^{-1} \| \left\{ |\epsilon| \frac{\| F \|}{\| A \|} + |\epsilon| \frac{\| f \|}{\| b \|} \right\} + \mathcal{O}(\epsilon^2) . \tag{8.58}$$

Utilizando a definição do *número de condição* da matriz A, em relação à norma $\| \cdot \|$, que é dado por:

$$\mathcal{K}(A) = \| A \| \, \| A^{-1} \| , \tag{8.59}$$

obtemos

$$\frac{\| \widehat{x}(\epsilon) - x^* \|}{\| x^* \|} \leq \mathcal{K}(A) \left\{ |\epsilon| \frac{\| F \|}{\| A \|} + |\epsilon| \frac{\| f \|}{\| b \|} \right\} + \mathcal{O}(\epsilon^2) . \tag{8.60}$$

Na desigualdade (8.60), temos que

$$\begin{aligned} \rho_A &= |\epsilon| \frac{\| F \|}{\| A \|} \quad ; \quad \epsilon \in \mathbb{R} \\ \rho_b &= |\epsilon| \frac{\| f \|}{\| b \|} \quad ; \quad \epsilon \in \mathbb{R} \end{aligned} \tag{8.61}$$

representam o erro relativo na matriz A e o erro relativo no elemento b, respectivamente.

Portanto, o erro relativo na solução do sistema sistema linear (8.2) é proporcional ao erro relativo $(\rho_A + \rho_b)$, onde $\mathcal{K}(A)$ é o fator de proporcionalidade. Neste sentido, o número de condição $\mathcal{K}(A)$ quantifica a sensibilidade do problema $Ax = b$.

Vamos apresentar através de exemplos uma comparação entre o número de condição e o valor do determinante, de uma matriz invertível.

Considere a matriz $B \in \mathbb{M}_n(\mathbb{R})$ triangular inferior definida da seguinte forma:

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ -1 & 1 & 0 & \cdots & \cdots & 0 \\ -1 & -1 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & -1 & 1 \end{bmatrix} \tag{8.62}$$

Temos que $\det(B) = 1$ para todo n. Entretanto, $\mathcal{K}_1(B) = n\,2^{n-1}$ o que torna a matriz *mal–condicionada* quando n cresce.

Exemplo 8.4.5 *Faça a construção da matriz* B^{-1} *mostrando que é dada por*

$$B^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & 1 & 0 & \cdots & \cdots & 0 \\ 2 & 1 & 1 & \cdots & \cdots & 0 \\ 4 & 2 & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 2^{n-2} & 2^{n-3} & \cdots & \cdots & 1 & 1 \end{bmatrix} \tag{8.63}$$

e que $\|\!| B^{-1} |\!\|_1 = 2^{n-1}$.

Por outro lado, podemos dar exemplo de uma matriz que é *bem–condicionada*, entretanto o determinante é muito pequeno.

De fato, considere a matriz diagonal $D \in \mathbb{M}_n(\mathbb{R})$ definida da seguinte forma:

$$D = diag(10^{-1}, \cdots, 10^{-1}, \cdots, 10^{-1})\,. \tag{8.64}$$

Temos que $\det(D) = 10^{-n}$, que é muito pequeno quando n cresce. Entretanto, temos que $\mathcal{K}_\alpha(D) = 1$, que é o número de condição em relação à uma norma $\|\!| \cdot |\!\|_\alpha$ que é induzida por uma norma $\| \cdot \|_\alpha$.

Portanto, a sensibilidade do problema $Ax = b$ não está relacionada com o $\det(A)$, mas especificamente com o número de condição da matriz do sistema linear.

Exemplo 8.4.6 *Considere a matriz positiva–definida* $A \in I\!M_2(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Determine as direções para os elementos b *e* δb *em* $I\!R^2$ *de modo que o erro relativo da solução do sistema linear* $Ax = b$ *seja o maior possível, isto é,*

$$\frac{\| \delta x^* \|_2}{\| x^* \|_2} = \mathcal{K}_2(A) \frac{\| \delta b \|_2}{\| b \|_2} ,$$

quando consideramos uma perturbação δb *no elemento* b.

8.5 Sistema Linear Positivo–Definido

Em muitos problemas em Análise Numérica são formulados como sistemas lineares e vários outros problemas, após uma linearização ou uma discretização, são transformados também num problema de sistema linear. Como exemplo de linearização, mencionamos o Método de Newton para sistemas de equações algébricas não lineares, no qual em cada iteração temos que obter a solução de um sistema linear. Entre os problemas de valores de contorno, temos vários métodos de discretização para obtermos uma solução numérica, e desse modo, após a discretização, temos que resolver um sistema linear. Em cada caso, a matriz do sistema linear tem propriedades específicas, dependendo da formulação do problema e do método de discretização escolhido. Por exemplo: pode ser simétrica, positiva–definida, diagonalmente dominante, etc.

De um modo geral, devemos analisar se a matriz é densa, se possui uma estrutura de esparsidade, ou se é simplesmente esparsa, sem uma estrutura bem definida. Uma matriz pode ser considerada esparsa no caso em que possui poucos elementos não–nulos, e esses elementos não estão em posições que seguem uma certa ordem. Matrizes esparsas surgem nos Métodos de Elementos Finitos para problemas de valores de contorno bidimensional, por exemplo. Consideramos que uma matriz é densa, se possui poucos elementos nulos. Matrizes densas estão associadas, por exemplo, ao Método dos Quadrados Mínimos para ajuste de curvas. Os casos mais interessantes são as matrizes que possuem uma certa estrutura de esparsidade, por exemplo: tridiagonal, banda, tridiagonal por blocos. Essas estruturas de esparsidade aparecem com freqüência nos sistemas lineares provenientes da discretização de problemas de valores de contorno através dos Esquemas de Diferenças Finitas.

Os métodos numéricos para sistemas lineares podem ser classificados em dois grupos, a saber: Métodos Diretos e Métodos Iterativos. Os métodos diretos são aqueles que, com um número finito de operações elementares encontramos a solução exata. Claro que não estamos considerando erros de arredondamento da aritmética de ponto flutuante. Como exemplo de métodos diretos, temos a Fatoração LU (Eliminação Gaussiana) e a Fatoração de Cholesky. Apresentamos também o Método dos Gradientes Conjugados, que do ponto de vista computacional é tratado como um método iterativo, mas mostraremos que na sua construção é um método direto, de acordo com a definição acima.

Os métodos iterativos são aqueles que, a partir de uma aproximação inicial para a solução do sistema linear, geram uma seqüência de novas aproximações que converge para a solução exata. Assim, a solução exata é obtida como o limite da seqüência gerada pelo método iterativo. Fica evidente que nos métodos iterativos temos que encontrar as condições para as quais a seqüência seja convergente. Note que os métodos iterativos necessitam de um número infinito de operações elementares para obter a solução exata, mesmo não considerando os erros de arredondamento da aritmética de ponto flutuante. Vamos analisar os Métodos Iterativos de Jacobi, Gauss-Seidel, Relaxação Sucessiva e o Método da Máxima Descida, que também é denominado Método do Gradiente Otimizado. Inicialmente vamos estudar os métodos numéricos para sistema linear positivo–definido.

Sejam $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz positiva–definida e $b \in \mathbb{R}^n$. Considere o Sistema Linear Positivo–Definido: encontrar $x^* \in \mathbb{R}^n$ solução da equação

$$Ax = b . \tag{8.65}$$

Como A é positiva-definida, portanto invertível, o sistema linear (8.65) possui uma única solução $x^* = A^{-1}b$.

Em geral, nos problemas de interesse prático, a matriz A é de grande porte e esparsa. Os métodos de decomposição tendem a modificar a estrutura de esparsidade, tornando elementos nulos em elementos não–nulos. Devemos observar que a Fatoração de Cholesky preserva as estruturas de esparsidade de banda e de envelope. Esta última requer uma reserva de posições de elementos nulos que irão se tornar não–nulos no fator de Cholesky, que estão no envelope. Nos métodos iterativos, é necessário somente o cálculo do produto da matriz A por elementos do $\mathbb{R}^n$, com isso não temos necessidade de modificar a estrutura de esparsidade da matriz do sistema linear. Assim, podemos dizer que esta é uma das grandes vantagens dos métodos iterativos.

A seguir apresentamos com todo detalhe o Problema de Minimização associado ao Sistema Linear Positivo–Definido, para que possamos construir o Método da Máxima Descida e o Método dos Gradientes Conjugados.

Problema de Minimização

Vamos associar ao **Sistema Linear Positivo–Definido**

$$Ax = b$$

o **Problema de Minimização:** encontrar $x^* \in \mathbb{R}^n$ tal que

$$J(x^*) = \min\{ J(x) \quad , \quad x \in \mathbb{R}^n \}, \tag{8.66}$$

onde o funcional $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ é definido da seguinte forma:

$$J(x) = \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle \tag{8.67}$$

com $\langle \cdot, \cdot \rangle$ o produto interno usual do $\mathbb{R}^n$.

Podemos verificar facilmente que

$$J(x^*) = -\frac{1}{2} b^t A^{-1} b = -\frac{1}{2}\langle A^{-1} b, b \rangle,$$

onde $x^* = A^{-1} b$ é a solução exata do sistema linear.

Vamos mostrar que o sistema linear positivo–definido e o problema de minimização são equivalentes, isto é, eles possuem a mesma solução. Essa equivalência vem essencialmente do fato da matriz A ser positiva–definida.

Teorema 8.5.1 *Sejam $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz positiva–definida e $b \in \mathbb{R}^m$. Então, o **Problema de Minimização**: encontrar $x^* \in \mathbb{R}^n$ tal que*

$$J(x^*) = \min\{ J(x) \quad ; \quad x \in \mathbb{R}^n \} \tag{8.68}$$

*é equivalente ao **Sistema Linear Positivo–Definido***

$$Ax = b. \tag{8.69}$$

Demonstração – Inicialmente vamos calcular a *Derivada Direcional* do funcional J no ponto $\overline{x}$ na direção do vetor $v \in \mathbb{R}^n$, que é definida da seguinte forma:

$$J'(\overline{x})(v) = \left\{ \frac{d}{dt} J(\overline{x} + tv) \right\}_{t=0}. \tag{8.70}$$

Utilizando a hipótese da matriz A ser simétrica, temos que

$$J(\overline{x} + tv) = J(\overline{x}) + t\langle A\overline{x}, v \rangle + \frac{t^2}{2}\langle Av, v \rangle - t\langle b, v \rangle \quad ; \quad t \in I\!R \tag{8.71}$$

Portanto, derivando (8.71) com relação a t e fazendo $t = 0$ obtemos

$$J'(\overline{x})(v) = \langle A\overline{x} - b, v \rangle \tag{8.72}$$

que é a *Derivada Direcional* do funcional J no ponto $\overline{x}$ na direção do vetor $v \in I\!R^n$.

De (8.72) temos a definição de gradiente do funcional J em um ponto $x \in I\!R^n$.

Definição 8.5.1 *O Gradiente do funcional J no ponto $x \in I\!R^n$ é definido por:*

$$\nabla J(x) = A\,x - b \tag{8.73}$$

Desse modo, definimos o Ponto Crítico do funcional J, como segue.

Definição 8.5.2 *Dizemos que $x^* \in I\!R^n$ é um Ponto Crítico do funcional J se, e somente se,*

$$J'(x^*)(v) = 0 \quad \textit{para todo} \quad v \in I\!R^n . \tag{8.74}$$

Desse modo, temos que um ponto crítico do funcional J é a solução do sistema linear positivo–definido $Ax = b$. Portanto, temos um único ponto crítico para J, que é dado por $x^* = A^{-1}b$.

Para classificar o ponto crítico devemos calcular a *Segunda Variação* do funcional J no ponto $\overline{x}$ na direção do vetor $w \in I\!R^n$, que é definida da seguinte forma:

$$J''(\overline{x}; v)(w) = \left\{ \frac{d}{dt} J'(\overline{x} + tw)(v) \right\}_{t=0} \tag{8.75}$$

Temos que

$$J'(\overline{x} + tw)(v) = \langle A\overline{x}, v \rangle + t\langle Aw, v \rangle - \langle b, v \rangle \quad ; \quad t \in I\!R \tag{8.76}$$

Portanto, derivando (8.76) com relação a t e fazendo $t = 0$, obtemos

$$J''(\overline{x}; v)(w) = \langle Aw, v \rangle \tag{8.77}$$

que é a *Segunda Variação* do funcional J no ponto $\overline{x}$ na direção do vetor $w \in I\!R^n$.

De (8.77) temos a definição da matriz Hessiana do funcional J em um ponto $x \in I\!R^n$.

Definição 8.5.3 *A matriz Hessiana do funcional J no ponto $x \in I\!R^n$ é definida por:*

$$H(x) = A. \tag{8.78}$$

Como A é uma matriz positiva-definida, temos que

$$J''(x^*; v)(v) > 0$$

para todo $v \in I\!R^n$ não–nulo.

Desse modo, $x^* = A^{-1}b$ é um Ponto de Mínimo Global para o funcional J, que é a única solução do sistema linear positivo–definido. Como a matriz Hessiana não depende da variável x, temos que J é um funcional quadrático, como ilustra a Figura 8.3.

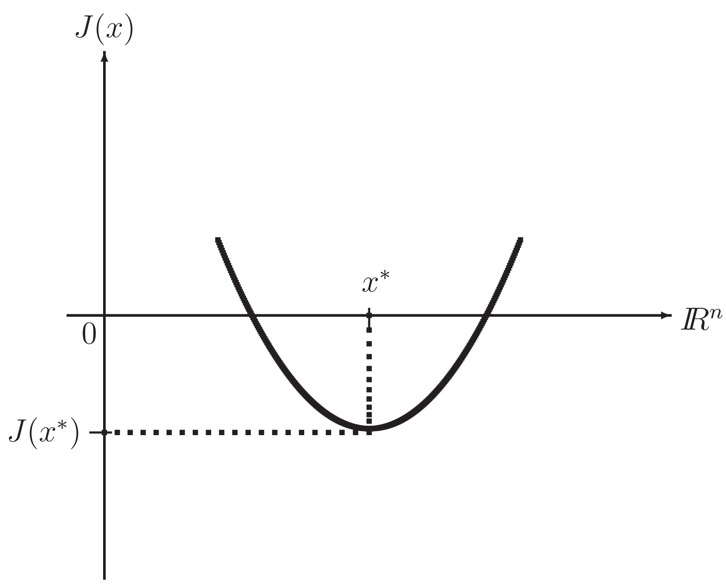

Figura 8.3: Ilustração gráfica do funcional quadrático J.

Assim, mostramos a equivalência entre o **Problema de Minimização**, descrito em (8.68), e o **Sistema Linear Positivo–Definido**, dado em (8.69).

8.6 Métodos dos Gradientes Conjugados

Vamos obter uma solução numérica para o sistema linear positivo–definido (8.65) através de um procedimento iterativo para encontrar o ponto de mínimo do funcional J, isto é, através de uma solução numérica do problema de minimização (8.66).

Método da Máxima Descida

Sabemos que o ponto de mínimo do funcional J pode ser encontrado andando na direção de $-\nabla J(x^{(k)}) = b - Ax^{(k)}$ a partir do ponto $x^{(k)}$, que é a direção que o funcional decresce mais rapidamente. Desse modo, se o resíduo $r^{(k)} = b - Ax^{(k)}$ é não–nulo, então existe um parâmetro $\lambda \in I\!R$ tal que

$$J(x^{(k)} + \lambda r^{(k)}) < J(x^{(k)}) \, .$$

No Método da Máxima Descida o parâmetro λ é escolhido de modo a obtermos o ponto de mínimo de J na variedade linear S_k definida da seguinte forma:

$$S_k = \{ \, z \in I\!R^n \; / \; z = x^{(k)} + \lambda r^{(k)} \quad ; \quad \lambda \in I\!R \, \} \, .$$

Desse modo, vamos encontrar um ponto $x^{(k+1)} = x^{(k)} + \lambda_k r^{(k)}$ tal que

$$J(x^{(k+1)}) = \min\{ \, J(z) \quad ; \quad z \in S_k \, \} \, ,$$

o que é equivalente a escrever

$$J(x^{(k+1)}) = \min\{ \, J(x^{(k)} + \lambda r^{(k)}) \quad ; \quad \lambda \in I\!R \, \}$$

onde λ_k é o parâmetro que realiza o mínimo.

Definindo a função auxiliar $\varphi : I\!R \longrightarrow I\!R$ da seguinte forma:

$$\varphi(\lambda) = J(x^{(k)} + \lambda r^{(k)}) \qquad ; \qquad \lambda \in I\!R$$

e fazendo $\varphi'(\lambda) = 0$, obtemos o parâmetro λ_k, que realiza o mínimo de J na variedade linear S_k. Assim, tem–se que

$$\lambda_k = \frac{\langle \, r^{(k)} \, , \, r^{(k)} \, \rangle}{\langle \, Ar^{(k)} \, , \, r^{(k)} \, \rangle}$$

para $r^k \in I\!R^n$ não–nulo.

Como A é uma matriz positiva–definida, temos que $\lambda_k > 0$. Este fato garante que estamos andando sempre na direção correta, como ilustra a Figura 8.4.

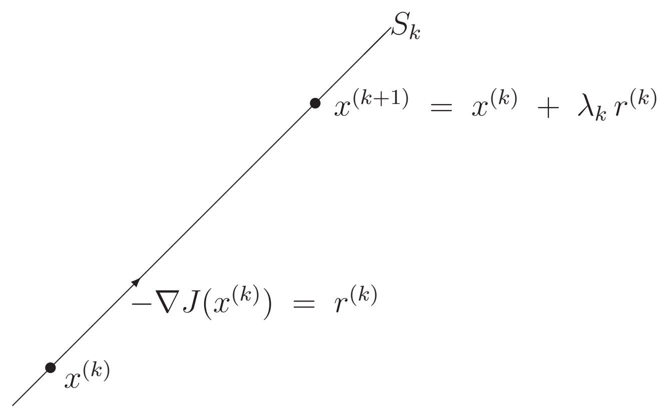

Figura 8.4: Escolha do passo λ_k.

No Método da Máxima Descida, temos que

$$x^{(k+1)} \in span\{ r^{(0)}, r^{(1)}, \cdots, r^{(k)} \} + x^{(0)}$$

e $\{ x^{(k)} \}$ é uma seqüência minimizante para o funcional J, isto é,

$$J(x^{(0)}) > \cdots > J(x^{(k)}) > J(x^{(k+1)}) \cdots > J(x^*)$$

para todo $k = 0, 1, 2, \cdots$.

A convergência global do Método da Máxima Descida segue da seguinte desigualdade

$$J(x^{(k+1)}) - J(x^*) \leq \left(1 - \frac{1}{\mathcal{K}_2(A)} \right) \left(J(x^{(k)}) - J(x^*) \right) \tag{8.79}$$

Note que, a escolha da aproximação inicial $x^{(0)}$ não é relevante para a convergência da seqüência $\{ x^{(k)} \}$. Desse modo, podemos considerar $x^{(0)} = 0$.

Exemplo 8.6.1 *Calcular as equações das curvas de níveis do funcional* J.
Sugestão: *utilizar a diagonalização da matriz* A, *isto é,* $A = Q \Lambda Q^t$, *e acompanhar a demonstração do Teorema 6.7.8.*

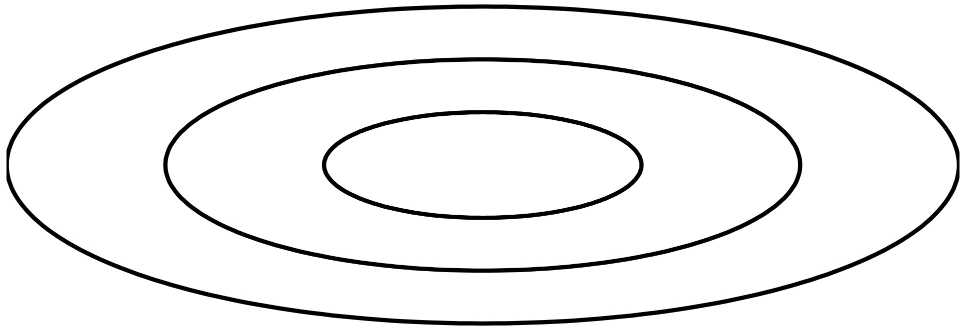

Figura 8.5: Ilustração das curvas de níveis do funcional J para $\mathcal{K}_2(A) \gg 1$, cujo centro é a solução única x^* do sistema linear positivo–definido (8.65).

Desse modo, podemos descrever o algoritmo do Método da Máxima Descida, que também é conhecido como **Método do Gradiente Otimizado**.

Algoritmo 8.6.1 (Método da Máxima Descida)

$x^{(0)} \in I\!R^n$ *uma aproximação inicial para* x^*

for $k = 0, 1, 2, 3, \cdots$

$$r^{(k)} = b - A\,x^{(k)}$$

$$\lambda_k = \frac{\langle\, r^{(k)} \,,\, r^{(k)} \,\rangle}{\langle\, A\,r^{(k)} \,,\, r^{(k)} \,\rangle}$$

$$x^{(k+1)} = x^{(k)} + \lambda_k\, r^{(k)}$$

end

Devemos observar que a convergência fica muito lenta no caso em que

$$\mathcal{K}_2(A) = \frac{\lambda_{max}}{\lambda_{min}}$$

é muito grande, pois as curvas de níveis do funcional J são elipsóides alongados com centro em x^*, como ilustra a Figura 8.5, conforme o resultado do Teorema 6.7.8.

Algebricamente, esta dificuldade vem do fato que as direções dos resíduos são muito próximas. Para superar esta situação descrevemos o Método dos Gradientes Conjugados.

Método dos Gradientes Conjugados

De um modo geral, podemos considerar sucessivas minimizações do funcional J ao longo de direções $\{ p^0, \ p^1, \ \cdots , \ p^k \}$ que não são necessariamente as direções dos resíduos. Isto é,

$$J(x^{(k+1)}) \ = \ \min\{ \ J(x^{(k)} \ + \ \lambda p^{(k)}) \quad ; \quad \lambda \in I\!R \ \} \tag{8.80}$$

com

$$x^{(k+1)} \ = \ x^{(k)} \ + \ \lambda_k \, p^{(k)}$$

onde o parâmetro λ_k é obtido do mesmo modo que no Método da Máxima Descida, como ilustra a Figura 8.6,

$$\lambda_k \ = \ \frac{\langle \, p^{(k)} \, , \, r^{(k)} \, \rangle}{\langle \, Ap^{(k)} \, , \, p^{(k)} \, \rangle} \tag{8.81}$$

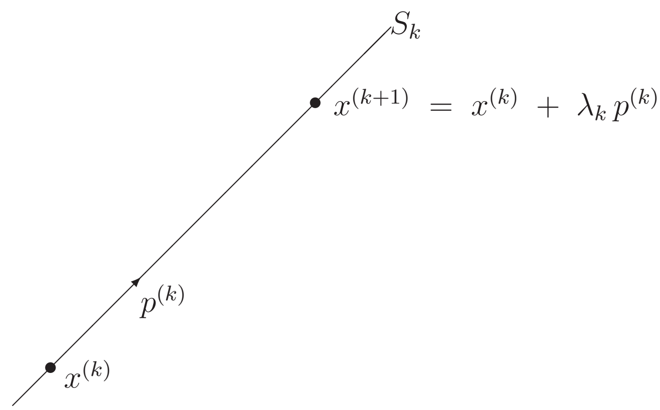

Figura 8.6: Escolha do passo λ_k.

Note que neste caso, tem–se

$$x^{(k+1)} \ \in \ span\left\{ \ p^{(0)}, \ p^{(1)}, \ \cdots \ , \ p^{(k)} \ \right\} \ + \ x^{(0)} \qquad \text{para} \qquad k = 0, 1, 2, \cdots$$

Em seguida, escolhemos as direções $p^{(k)}$, $k = 0, 1, 2, \cdots$, de modo a obter a convergência do processo iterativo em n passos, onde n é a dimensão do sistema linear.

Inicialmente, construímos direções consecutivas $p^{(k)}$ e $p^{(k+1)}$ A–conjugadas,

$$\langle \, A\, p^{(k)} \, , \, p^{(k+1)} \, \rangle \; = \; 0 \, ,$$

com $p^{(0)} = -\nabla J(x^{(0)}) = r^{(0)}$. Assim, vamos determinar as direções da forma:

$$p^{(k+1)} \; = \; r^{(k+1)} \; + \; \beta_k \, p^{(k)} \tag{8.82}$$

onde o parâmetro β_k é obtido impondo a condição que as direções $p^{(k)}$ e $p^{(k+1)}$ sejam A–conjugadas. Assim, tem–se que

$$\beta_k \; = \; - \frac{\langle \, A\, r^{(k+1)} \, , \, p^{(k)} \, \rangle}{\langle \, A\, p^{(k)} \, , \, p^{(k)} \, \rangle} \; = \; - \frac{\langle \, r^{(k+1)} \, , \, A\, p^{(k)} \, \rangle}{\langle \, A\, p^{(k)} \, , \, p^{(k)} \, \rangle} \tag{8.83}$$

Desse modo, temos que a direção $p^{(k+1)}$ é a projeção ortogonal do resíduo $r^{(k+1)}$ sobre o complemento ortogonal do subespaço gerado pela direção $p^{(k)}$, com relação ao produto interno energia $\langle \cdot \, , \, \cdot \rangle_A$.

Podemos calcular o resíduo $r^{(k+1)}$ de uma maneira mais econômica. De fato,

$$r^{(k+1)} \; = \; b \; - \; A\, x^{(k+1)} \; = \; b \; - \; A\, x^{(k)} \; - \; \lambda_k \, A\, p^{(k)} \tag{8.84}$$

Assim, temos que

$$r^{(k+1)} \; = \; r^{(k)} \; - \; \lambda_k \, A\, p^{(k)} \tag{8.85}$$

Portanto, obtemos $r^{(k+1)}$ sem nenhum custo, pois já temos o cálculo de $A\, p^{(k)}$.

No Algoritmo 8.6.2, apresentamos uma primeira versão do **Método dos Gradientes Conjugados**.

Agora, vamos reescrever os parâmetros λ_k e β_k de uma maneira mais econômica. Inicialmente, para encontrar uma maneira mais barata para o parâmetro λ_k , utilizando as propriedades do Método dos Gradientes Conjugados. Note que

$$\langle \, p^{(k)} \, , \, r^{(k)} \, \rangle \; = \; \langle \, r^{(k)} \, , \, r^{(k)} \, \rangle \; + \; \beta_{k-1} \, \langle \, p^{(k-1)} \, , \, r^{(k)} \, \rangle \; = \; \langle \, r^{(k)} \, , \, r^{(k)} \, \rangle$$

pois $\langle \, p^{(k-1)} \, , \, r^{(k)} \, \rangle = 0$, tendo em vista que $r^{(k)} = -\nabla J(x^{(k)})$, onde $x^{(k)}$ é o ponto de mínimo do funcional J na direção de $p^{(k-1)}$. Assim, temos que

$$\lambda_k \; = \; \frac{\langle \, r^{(k)} \, , \, r^{(k)} \, \rangle}{\langle \, A p^{(k)} \, , \, p^{(k)} \, \rangle} \tag{8.86}$$

Podemos mostrar agora, que dois resíduos consecutivos também são ortogonais, isto é, $r^{(k)}$ e $r^{(k+1)}$ são ortogonais. Note que,

$$\langle p^{(k)} , A p^{(k)} \rangle = \langle r^{(k)} , A p^{(k)} \rangle + \beta_{k-1} \langle p^{(k-1)} , A p^{(k)} \rangle = \langle r^{(k)} , A p^{(k)} \rangle \qquad (8.87)$$

Assim, podemos escrever o parâmetro λ_k da seguinte forma alternativa

$$\lambda_k = \frac{\langle r^{(k)} , r^{(k)} \rangle}{\langle Ap^{(k)} , r^{(k)} \rangle} \qquad (8.88)$$

Finalmente, da relação (8.85) e utilizando a relação (8.88) para o parâmetro λ_k, obtemos

$$\langle r^{(k+1)} , r^{(k)} \rangle = \langle r^{(k)} , r^{(k)} \rangle - \lambda_k \langle A p^{(k)} , r^{(k)} \rangle = 0 \qquad (8.89)$$

Portanto, obtemos o resultado desejado.

Vamos obter uma maneira mais econômica de calcular o parâmetro β_k. Utilizando a relação (8.85), temos uma expressão para $A p^{(k)}$ que, substituída na relação (8.83) do parâmetro β_k , e com a nova expressão (8.86) para o parâmetro λ_k , implica

$$\beta_k = \frac{\langle r^{(k+1)} , r^{(k+1)} \rangle}{\langle r^{(k)} , r^{(k)} \rangle} \qquad (8.90)$$

Finalmente, devemos mostrar que $\{ r^{(0)}, r^{(1)}, \cdots , r^{(k)}, \cdots \}$ é um conjunto ortogonal e que $\{ p^{(0)}, p^{(1)}, \cdots , p^{(k)}, \cdots \}$ é um conjunto A–conjugado. Assim, mostramos que o Método dos Gradientes Conjugados converge para a solução exata do sistema linear positivo–definido (8.65) em n passos. Desse modo, podemos dizer que o Método dos Gradientes Conjugados é um método exato.

Vamos obter os resultados acima por indução matemática. Sabemos que esses resultados são válidos para $k = 1$, pois as direções $p^{(0)}$ e $p^{(1)}$ foram construídas de modo a serem A–conjugadas. Temos também que $p^{(0)} = r^{(0)}$ e $r^{(1)}$ são ortogonais, isto é, $\langle p^{(0)} , r^{(1)} \rangle = 0$, tendo em vista que $r^{(1)} = -\nabla J(x^{(1)})$, onde $x^{(1)}$ é o ponto de mínimo do funcional J na direção de $p^{(0)}$.

Supomos, pela hipótese de indução, que os resíduos $r^{(0)}, r^{(1)}, \cdots , r^{(k-1)}$ são mutuamente ortogonais e que as direções

$$p^{(0)} , p^{(1)} , \cdots , p^{(k-1)}$$

são mutuamente A–conjugadas, para $k \geq 2$.

Da relação (8.85) , para $0 \leq i \leq (k-2)$, e da hipótese de indução, temos que

$$\begin{aligned} \langle\, r^{(k)} \,,\, r^{(i)} \,\rangle \quad &= \quad \langle\, r^{(k-1)} \,,\, r^{(i)} \,\rangle \;-\; \lambda_{k-1} \,\langle\, A\,p^{(k-1)} \,,\, r^{(i)} \,\rangle \\ &= \quad -\lambda_{k-1} \,\langle\, A\,p^{(k-1)} \,,\, r^{(i)} \,\rangle \end{aligned} \tag{8.91}$$

Da relação (8.82), podemos escrever $r^{(i)}$ da seguinte forma:

$$r^{(i)} \;=\; p^{(i)} \;-\; \beta_{i-1}\, p^{(i-1)} \tag{8.92}$$

que, substituindo na relação (8.91), e da hipótese de indução para as direções, implica

$$\langle\, r^{(k)} \,,\, r^{(i)} \,\rangle \;=\; -\lambda_{k-1} \,\langle\, A\,p^{(k-1)} \,,\, p^{(i)} \;-\; \beta_{i-1}\, p^{(i-1)} \,\rangle \;=\; 0 \tag{8.93}$$

para $i = 0, 1, \cdots, (k-2)$.

Assim, obtemos que os resíduos $r^{(0)}$, $r^{(1)}$, $\cdots$, $r^{(k)}$ são mutuamente ortogonais, pois de (8.89), sabemos que $r^{(k)}$ e $r^{(k-1)}$ são ortogonais.

Sabemos que $p^{(k)}$ e $p^{(k-1)}$ são A–conjugados pela própria construção do método. Da relação (8.82), para $0 \leq i \leq (k-2)$, e utilizando a hipótese de indução para as direções, temos que

$$\langle\, p^{(k)} \,,\, A\,p^{(i)} \,\rangle \;=\; \langle\, r^{(k)} \,,\, A\,p^{(i)} \,\rangle \;+\; \beta_{k-1} \,\langle\, p^{(k-1)} \,,\, A\,p^{(i)} \,\rangle \;=\; \langle\, r^{(k)} \,,\, A\,p^{(i)} \,\rangle \tag{8.94}$$

Da relação (8.85) , temos que

$$A\,p^{(i)} \;=\; \frac{r^{(i)} \;-\; r^{(i+1)}}{\lambda_i} \tag{8.95}$$

Substituindo na relação (8.94) e utilizando o fato que os resíduos são mutuamente ortogonais, segue que

$$\langle\, p^{(k)} \,,\, A\,p^{(i)} \,\rangle \;=\; 0 \tag{8.96}$$

para $i = 0, 1, \cdots, (k-2)$.

Assim, obtemos o resultado que as direções

$$p^{(0)} \,,\, p^{(1)} \,,\, \cdots \,,\, p^{(k)}$$

são mutuamente A–conjugadas, o que completa a demonstração. ■

No caso em que A é uma matriz esparsa sem uma estrutura bem definida, podemos armazena-la com o Esquema de Coordenadas, isto é, `(ilin(k),jcol(k))` que é a posição do k-ésimo elemento não–nulo e `Avet(k)` é o valor deste elemento. Neste esquema de armazenamento não é necessário uma ordem para a entrada dos elementos não–nulos. Os esquemas para armazenamento de uma matriz esparsa estão descritos na seção 8.6. Neste caso, é facilmente calculado o produto da matriz $A \in \mathbb{M}_n(\mathbb{R})$ por um elemento do $x \in \mathbb{R}^n$. Os procedimentos para o produto de uma matriz esparsa por um vetor estão descritos na seção 8.6.

Uma estrutura de esparsidade que iremos trabalhar mais a frente, e que está ilustrada na Figura 8.7, está relacionada com a discretização da equação de Poisson com condição de Dirichlet num retângulo, pelo Esquema de Diferenças Finitas Centrada numa malha com 5 nós internos em cada uma das direções x e y.

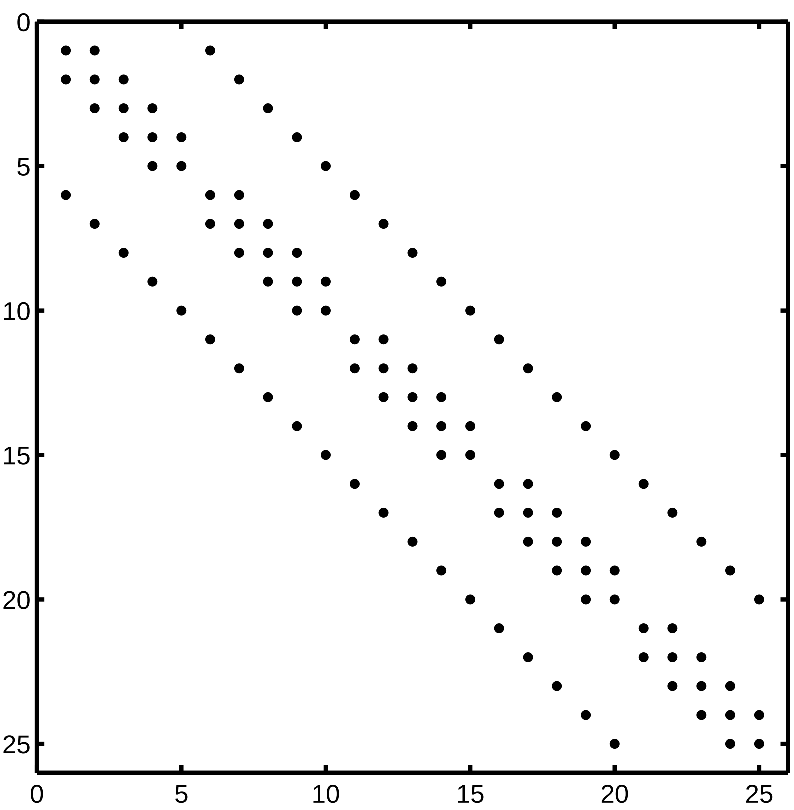

Figura 8.7: Matriz associada a discretização do Problema de Poisson no retângulo.

Algoritmo 8.6.2 (Método dos Gradientes Conjugados)

$x^{(0)} \in I\!R^n$ *uma aproximação inicial para* x^*

$r^{(0)} = b - A\,x^{(0)} \quad ; \quad p^{(0)} = r^{(0)}$

for $k = 0, 1, 2, 3, \cdots$

$$\lambda_k = \frac{\langle\, p^{(k)} \,,\, r^{(k)} \,\rangle}{\langle\, A\,p^{(k)} \,,\, p^{(k)} \,\rangle}$$

$$x^{(k+1)} = x^{(k)} + \lambda_k\, p^{(k)}$$

$$r^{(k+1)} = r^{(k)} - \lambda_k\, A\, p^{(k)}$$

$$\beta_k = -\frac{\langle\, A\,p^{(k)} \,,\, r^{(k+1)} \,\rangle}{\langle\, A\,p^{(k)} \,,\, p^{(k)} \,\rangle}$$

$$p^{(k+1)} = r^{(k+1)} + \beta_k\, p^{(k)}$$

end

Complexidade Computacional

Podemos verificar facilmente que a complexidade computacional, número de operações elementares, do Método dos Gradientes Conjugados, em cada iteração, é dada por:

- **um** produto de matriz por vetor para o cálculo de $A\,p^{(k)}$.
- **três** produtos internos.

e **um** produto de matriz por vetor para o cálculo do resíduo $r^{(0)}$, para iniciar as iterações. Desse modo, a complexidade computacional é do tipo polinomial e da ordem de n^2, em cada iteração.

Finalmente, temos uma versão econômica para o Método dos Gradientes Conjugados.

Algoritmo 8.6.3 (Método dos Gradientes Conjugados)

$x^{(0)} \in \mathbb{R}^n$ *uma aproximação inicial para* x^*

$r^{(0)} = b - A\,x^{(0)} \quad ; \quad p^{(0)} = r^{(0)}$

for $k = 0, 1, 2, 3, \cdots$

$$\lambda_k = \frac{\langle\, r^{(k)} \,,\, r^{(k)} \,\rangle}{\langle\, A\,p^{(k)} \,,\, p^{(k)} \,\rangle}$$

$$x^{(k+1)} = x^{(k)} + \lambda_k\, p^{(k)}$$

$$r^{(k+1)} = r^{(k)} - \lambda_k\, A\, p^{(k)}$$

$$\beta_k = \frac{\langle\, r^{(k+1)} \,,\, r^{(k+1)} \,\rangle}{\langle\, r^{(k)} \,,\, r^{(k)} \,\rangle}$$

$$p^{(k+1)} = r^{(k+1)} + \beta_k\, p^{(k)}$$

end

Exemplo 8.6.2 *Faça uma análise da complexidade computacional do Algoritmo 8.6.3, versão econômica para o Método dos Gradientes Conjugados, e faça uma comparação com a complexidade computacional do Algoritmo 8.6.2.*

Gradientes Conjugados Precondicionado

Vamos apresentar alguns resultados sobre a velocidade de convergência do Método dos Gradientes Conjugados, tendo em vista que computacionalmente devemos trata–lo como um método iterativo, isto é, estabelecer um critério de parada. Assim, é muito interessante conhecer sua taxa de convergência. A seguir, temos um resultado para a velocidade de convergência do Método do Gradiente Otimizado.

Teorema 8.6.1 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *simétrica e positiva–definida e* $b \in \mathbb{R}^n$. *Se* $\{x^{(k)}\}$ *é a seqüência gerada pelo Método do Gradiente Otimizado. Então,*

$$\| x^* - x^{(k)} \|_a \leq 2 \| x^* - x^{(0)} \|_a \left[\frac{\sqrt{\mathcal{K}_2(A)} - 1}{\sqrt{\mathcal{K}_2(A)} + 1} \right]^k \tag{8.97}$$

Demonstração Veja Luenberger (1973). □

Note que, a taxa de convergência na norma energia do Método do Gradiente Otimizado é bastante sensível ao número de condição espectral da matriz A, o que torna o método não muito atraente pois o sucesso do método iterativo depende exclusivamente da sua velocidade de convergência. A seguir, temos dois resultados que nos fornecem um panorama de como deve ser a velocidade de convergência do Método dos Gradientes Conjugados.

Teorema 8.6.2 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *simétrica e positiva–definida e* $b \in \mathbb{R}^n$. *Se* $\{x^{(k)}\}$ *é a seqüência gerada pelo Método dos Gradientes Conjugado. Então,*

$$\| x^* - x^{(k)} \|_a \leq \| x^* - x^{(0)} \|_a \left[\frac{\sqrt{\mathcal{K}_2(A)} - 1}{\sqrt{\mathcal{K}_2(A)} + 1} \right]^{2k} \tag{8.98}$$

Demonstração Veja Luenberger (1973). □

Teorema 8.6.3 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *simétrica e positiva–definida com* k *autovalores distintos e* $b \in \mathbb{R}^n$. *Então, a seqüência* $\{x^{(k)}\}$ *gerada pelo Método dos Gradientes Conjugados converge para a solução exata* $x^* = A^{-1} b$ *com no máximo* $(k+1)$ *iterações.*

Demonstração Veja Luenberger (1973). □

Podemos observar que o Método dos Gradientes Conjugados converge rapidamente na norma energia no caso em $\mathcal{K}_2(A) \approx 1$ ou no caso em que a matriz A tiver poucos autovalores distintos, e terá uma convergência muito lenta se $\mathcal{K}_2(A) \gg 1$. Assim, vamos necessitar de muitas iterações para atingir uma precisão desejada.

A seguir, vamos mostrar como podemos substituir o sistema linear simétrico e positivo–definido $A\,x = b$ por um sistema linear simétrico e positivo–definido equivalente $\widetilde{A}\,\widetilde{x} = \widetilde{b}$, onde a matriz $\widetilde{A}$ tem a propriedade de $\mathcal{K}_2(\widetilde{A})$ bem menor que $\mathcal{K}_2(A)$. Assim, podemos atingir a precisão requerida com poucas iterações.

A estratégia apresentada acima é denominada precondicionamento do sistema linear, onde devemos encontrar uma matriz C simétrica e positiva–definida para ser o precondicionador da matriz A, isto é, $\widetilde{A} = C^{-1}\,A\,C^{-1}$. Em seguida aplicamos o Método dos Gradientes Conjugados no sistema linear transformado

$$\widetilde{A}\,\widetilde{x} = \widetilde{b} \tag{8.99}$$

onde $\widetilde{x} = C\,x$ e $\widetilde{b} = C^{-1}\,b$. Neste processo a dificuldade está em escolher uma matriz C de modo que $\mathcal{K}_2(\widetilde{A}) \ll \mathcal{K}_2(A)$. A escolha depende muito do problema que deu origem ao sistema linear $A\,x = b$. A escolha de um bom precondicionador pode fazer uma grande diferença na velocidade de convergência. Fazendo a seguinte mudança de variáveis

$$\begin{aligned}
M &= C^2 \\
p^{(k)} &= C^{-1}\,\widetilde{p}^{(k)} \\
x^{(k)} &= C^{-1}\,\widetilde{x}^{(k)} \\
z^{(k)} &= C^{-1}\,\widetilde{r}^{(k)} \\
r^{(k)} &= C\,\widetilde{r}^{(k)} = b - A\,x^{(k)}
\end{aligned}$$

no Método dos Gradientes Conjugados para o sistema linear transformado (8.99), obtemos o Método dos Gradientes Conjugados Precondicionado para o sistema linear $A\,x = b$, onde a matriz simétrica e positiva–definida M é o precondicionador. A seguir, apresentamos o algoritmo.

Algoritmo 8.6.4 (Gradientes Conjugados Precondicionado)

$x^{(0)} \in I\!R^n$ *uma aproximação inicial para* x^*

$r^{(0)} = b - A\,x^{(0)}$

$Mz^{(0)} = r^{(0)}$ $\quad$ *(resolver em* $z^{(0)}$*)*

$p^{(0)} = z^{(0)}$

for $k = 0, 1, 2, 3, \cdots$

$$\lambda_k = \frac{\langle\, z^{(k)} \,,\, r^{(k)} \,\rangle}{\langle\, A\,p^{(k)} \,,\, p^{(k)} \,\rangle}$$

$$x^{(k+1)} = x^{(k)} + \lambda_k\, p^{(k)}$$

$$r^{(k+1)} = r^{(k)} - \lambda_k\, A\, p^{(k)}$$

$Mz^{(k+1)} = r^{(k+1)}$ $\quad$ *(resolver em* $z^{(k+1)}$*)*

$$\beta_k = \frac{\langle\, z^{(k+1)} \,,\, r^{(k+1)} \,\rangle}{\langle\, z^{(k)} \,,\, r^{(k)} \,\rangle}$$

$$p^{(k+1)} = z^{(k+1)} + \beta_k\, p^{(k)}$$

end

Na escolha da matriz M simétrica e positiva–definida para ser o precondicionador, observamos que o sistema linear $Mz = r$ deve ser de fácil resolução, para que tenhamos um algoritmo eficiente. Assim, uma primeira escolha para M é a matriz $D = diag(a_{11}, \cdots, a_{nn})$ que é a diagonal principal de A. Desse modo, temos que o precondicionador do sistema linear $A\,x = b$ é a matriz $C = diag(\sqrt{a_{11}}, \cdots, \sqrt{a_{nn}})$.

Armazenamento de Matrizes Esparsas

Nesta seção apresentamos dois esquemas para armazenamento de matrizes esparsas que geralmente são utilizados para a implementação computacional dos métodos da família dos gradientes conjugados, assim como nos principais pacotes computacionais de *Álgebra Linear Computacional*. Por simplicidade, consideramos como exemplo a matriz esparsa

$$A = \begin{bmatrix} 1.0 & 0.0 & 0.0 & -1.0 & 0.0 \\ 2.0 & 0.0 & -2.0 & 0.0 & 3.0 \\ 0.0 & -3.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 4.0 & 0.0 & -4.0 & 0.0 \\ 5.0 & 0.0 & -5.0 & 0.0 & 6.0 \end{bmatrix}$$

que será utilizada para descrever os dois esquemas de armazenamento.

Esquema de Coordenadas

Vamos denotar por `(ilin(k),jcol(k))` a posição do k-ésimo elemento não–nulo e por `Avet(k)` o valor do k-ésimo elemento não-nulo da matriz A.

k	1	2	3	4	5	6	7	8	9	10	11
ilin(k)	1	5	2	4	2	1	3	5	4	5	2
jcol(k)	4	5	5	2	1	1	2	3	4	1	3
Avet(k)	-1.0	6.0	3.0	4.0	2.0	1.0	-3.0	-5.0	-4.0	5.0	-2.0

Esquema de Coleção de Vetores Esparsos

Vamos denotar por `irowst(i)` o início da i-ésima linha nos vetores `jcol` e `A`, denotamos `lenrow(i)` o número de elementos não nulos na i-ésima linha, por `jcol` as posições das colunas dos elementos não nulos e por `Avet` o vetor que contém os elementos não nulos da matriz.

k	1	2	3	4	5	6	7	8	9	10	11
jcol(k)	4	1	5	1	3	4	2	5	1	3	2
Avet(k)	-1.0	1.0	3.0	2.0	-2.0	-4.0	4.0	6.0	5.0	-5.0	-3.0

linha i	1	2	3	4	5
lenrow(i)	2	3	1	2	3
irowst(i)	1	3	11	6	8

Produto de Matriz Esparsa por Vetor

Vamos descrever o algoritmo para calcular o produto de uma matriz $A \in I\!\!M_n(I\!\!R)$ por um vetor $x \in I\!\!R^n$, isto é, vamos calcular o vetor $y = Ax$, para o caso em que a matriz esparsa está armazenada pelo Esquema de Coordenadas. Indicamos por $nelem$ o número de elementos não–nulos da matriz A.

Algoritmo 8.6.5 (Produto de Matriz Esparsa por um Vetor)

```
for i = 1,2, ... , n
y(i) = 0.0
end

for k = 1,2, ... , nelem
y(ilin(k))  =  y(ilin(k))  +  Avet(k)*x(jcol(k))
end
```

Vamos descrever o algoritmo para calcular o produto de uma matriz $A \in I\!\!M_n(I\!\!R)$ por um vetor $x \in I\!\!R^n$, isto é, vamos calcular o vetor $y = Ax$, para o caso em que a matriz esparsa está armazenada pelo Esquema de Coleção de Vetores Esparsos.

Algoritmo 8.6.6 (Produto de Matriz Esparsa por um Vetor)

```
for i = 1,2, ... , n
    y(i) = 0.0
    for j = 1, ... , lenrow(i)
    k = irowst(i) + j - 1
    y(i)  =  y(i)  +  Avet(k)*x(jcol(k))
    end
end
```

Os algoritmos apresentados acima, para o cálculo do produto de uma matriz esparsa por um vetor, são os mais utilizados na implementação computacional dos métodos da família dos gradientes conjugados, por sua simplicidade e eficiência.

Exercícios

Exercício 8.22 *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível e* $b \in \mathbb{R}^n$. *Calcular o gradiente e a Hessiana do funcional*

$$J(x) = \frac{1}{2} \langle A\,x\,,\,x \rangle - \langle b\,,\,x \rangle \quad ; \quad x \in \mathbb{R}^n \,. \tag{8.100}$$

Exercício 8.23 *Considere o Problema de Valor de Contorno com condição periódica*

$$\begin{aligned} -u''(x) + \sigma\, u(x) &= f(x) \quad ; \quad x \in (0, L) \\ u(0) &= u(L) \end{aligned} \tag{8.101}$$

com $\sigma > 0$ *e* f *uma função contínua.*

O sistema linear proveniente da discretização do problema de valor de contorno (8.101) pelo Esquema de Diferenças Finitas Centrada, para uma partição regular

$$\Pi : 0 = x_1 < \cdots < x_{n-1} < x_n = L \,,$$

com espaçamento h, *é dado por:*

$$\begin{bmatrix} d & -1 & & & & -1 \\ -1 & d & -1 & & & \\ & -1 & d & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & d & -1 \\ -1 & & & & -1 & d \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} h^2 f(x_1) \\ \vdots \\ h^2 f(x_i) \\ \vdots \\ h^2 f(x_{n-1}) \end{bmatrix} \tag{8.102}$$

onde

$$d = (\, 2 + \sigma\, h^2 \,) \quad e \quad h = \frac{L}{(n-1)} \,.$$

Estamos denotando por u_i *uma aproximação para o valor* $u(x_i)$ *fornecida pelo Esquema de Diferenças Finitas. Note que* $u_n = u_1$ *devido a condição de contorno periódica.*

Determine uma solução numérica do sistema linear (8.102) pelo Método dos Gradientes Conjugados, com um resíduo relativo inferior a 10^{-5}.

Como exemplo, considere $L = 1$, $f(x) = \sin(\pi x)$ *e* $\sigma = 0.1$. *Para observar o desempenho dos métodos numéricos utilizar vários valores de* n.

Exercício 8.24 *Considere a matriz A positiva–definida dada por:*

$$A = \begin{bmatrix} 16 & 4 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 25 & -3 & 6 & 0 & 0 & 0 & 10 & 0 & 0 \\ 4 & -3 & 30 & 8 & 0 & 2 & 0 & -9 & 0 & 0 \\ 0 & 6 & 8 & 47 & 8 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 17 & 3 & 2 & -1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 3 & 15 & 1 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 21 & -4 & -7 & 0 \\ 0 & 10 & -9 & 0 & -1 & -3 & -4 & 52 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 3 & -7 & -1 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 24 \end{bmatrix}$$

Faça a representação da matriz A pelo Esquema de Coordenadas e pelo Esquema de Coleção de Vetores Esparsos.

Exercício 8.25 *Faça uma implementação computacional para o Método dos Gradientes, em uma linguagem de sua preferência, considerando que a matriz A está armazenada pelo Esquema de Coordenadas.*

Exercício 8.26 *Faça uma implementação computacional para o Método dos Gradientes Conjugados, em uma linguagem de sua preferência, considerando que a matriz A está armazenada pelo Esquema de Coordenadas.*

Exercício 8.27 *Considerando a matriz do Exercício 8.24 e o vetor b dado por:*

$$b = \begin{bmatrix} 14.00 \\ -24.50 \\ 87.70 \\ 142.50 \\ 7.90 \\ 47.35 \\ -46.45 \\ 8.45 \\ -3.50 \\ 80.00 \end{bmatrix}$$

obter uma solução numérica para o sistema linear $Ax = b$ utilizando as implementações computacionais dos Exercícios 8.25 e 8.26.

Exercício 8.28 *Sejam* $A \in I\!M_{m \times n}(I\!R)$, *com* $m \geq n$ *e* *posto*$(A) = n$, *isto é, as colunas de* A *formam um conjunto linearmente independente em* $I\!R^m$, *e um elemento* $b \in I\!R^m$. *Considere o sistema linear sobredeterminado*

$$A\,x \;=\; b\,, \tag{8.103}$$

cuja ***Solução de Quadrados Mínimos*** *é o elemento* $x^* \in I\!R^n$ *tal que*

$$J(x^*) \;=\; \min\{\; J(x) \quad ; \quad x \in I\!R^n \;\}\,, \tag{8.104}$$

onde $J : I\!R^n \longrightarrow I\!R^+$ *é um funcional quadrático definido da seguinte forma:*

$$J(x) \;=\; \langle\, Ax \,-\, b\,,\, Ax \,-\, b\,\rangle \qquad ; \qquad x \in I\!R^n\,. \tag{8.105}$$

Pede-se

1. *Determine o gradiente,* $\nabla J(x)$, *do funcional* J.

2. *Determine a matriz Hessiana,* $H(x)$, *do funcional* J.

3. *Mostre que o funcional* J *possui um único ponto de mínimo global.*

4. *Determine o valor do funcional* J *no ponto de mínimo.*

5. *Faça as alterações necessárias no algoritmo do Método dos Gradientes Conjugados para obtermos uma solução numérica do Sistema Linear Normal* $A^tAx = A^tb$, *de modo que não seja feito o cálculo de* A^tA *explicitamente.*

6. *Quais são as vantagens em não fazer o cálculo explicitamente de* A^tA *?*

8.7 Fatoração de Cholesky

Sejam $A \in I\!M_n(I\!R)$ uma matriz positiva–definida e um elemento $b \in I\!R^n$. Vamos considerar o problema de encontrar $x^* \in I\!R^n$ solução do sistema linear positivo–definido

$$Ax = b.$$

Podemos obter uma solução numérica através da Fatoração de Cholesky da matriz A, garantida pelo teorema abaixo.

Teorema 8.7.1 *Seja $A \in I\!M_n(I\!R)$ uma matriz positiva–definida. Então, existe uma única matriz triangular superior G, com os elementos da diagonal principal positivos, tal que $A = G^t G$.*

Demonstração – Vamos fazer uma construção do fator de Cholesky por indução sobre a ordem da matriz A. Para $A = [a_{11}]$, com $a_{11} > 0$, temos que $G = [\sqrt{a_{11}}]$. Pela hipótese de indução, supomos a existência do fator de Cholesky para uma matriz de ordem $n-1$, e vamos construir o fator de Cholesky para uma matriz $A \in I\!M_n(I\!R)$.

Inicialmente, escrevemos a matriz A na forma particionada

$$A = \begin{bmatrix} A_{n-1} & \omega \\ \omega^t & a_{nn} \end{bmatrix} \tag{8.106}$$

onde a matriz A_{n-1}, de ordem $(n-1)$, é positiva–definida e $a_{nn} > 0$, pois A é uma matriz positiva–definida, e o elemento $\omega \in I\!R^{n-1}$.

Usando a hipótese de indução, sabemos que existe o fator de Cholesky da matriz A_{n-1}, isto é, $A = G_{n-1}^t G_{n-1}$. Desse modo, vamos procurar o fator de Cholesky da matriz A da seguinte forma:

$$G = \begin{bmatrix} G_{n-1} & c \\ 0^t & \alpha \end{bmatrix} \tag{8.107}$$

com o elemento $c \in I\!R^{n-1}$ e $\alpha > 0$, que são determinados pela equação matricial

$$\begin{bmatrix} A_{n-1} & \omega \\ \omega^t & a_{nn} \end{bmatrix} = \begin{bmatrix} G_{n-1}^t & 0 \\ c^t & \alpha \end{bmatrix} \begin{bmatrix} G_{n-1} & c \\ 0^t & \alpha \end{bmatrix} \tag{8.108}$$

resultando nas seguintes equações

$$c^t c \; + \; \alpha^2 \; = \; a_{nn} \qquad \text{e} \qquad G_{n-1}^t c \; = \; \omega \; .$$

Como G_{n-1} é invertível, existe um único elemento $c \in \mathbb{R}^{n-1}$ solução do sistema triangular inferior

$$G_{n-1}^t c \; = \; \omega \; .$$

Sabemos que

$$\det(A) \; = \; \alpha^2 \, (\, det(G_{n-1}) \,)^2 \; > \; 0$$

implicando na existência de um único escalar $\alpha > 0$ solução da equação

$$c^t c \; + \; \alpha^2 \; = \; a_{nn} \qquad \Longrightarrow \qquad \alpha \; = \; \sqrt{a_{nn} \; - \; c^t c}$$

Portanto, temos a existência e unicidade do Fator de Cholesky. Além disso, mostramos uma forma de construí–lo.

Note que, se α for um complexo puro, então a matriz A não é positiva–definida. Desse modo, durante o processo da Fatoração de Cholesky podemos fazer a verificação se a matriz A é positiva–definida, o que completa a demonstração. ∎

A seguir apresentamos o algoritmo da fatoração de Cholesky descrito no Teorema 8.7.1. Note que neste procedimento o fator de Cholesky é construído por coluna, que é obtida resolvendo um sistema triangular inferior, para depois calcular o elemento da diagonal.

Algoritmo 8.7.1 (Fatoração de Cholesky)

```
for j = 1,2, ... , n
    G(j,j) = A(j,j)
    for i = 1,2, ... , (j-1)
        G(i,j) = A(i,j)
        for k = 1,2, ... , (i-1)
            G(i,j) = G(i,j) - G(k,i)*G(k,j)
        end
        G(i,j) = G(i,j)/G(i,i)
        G(j,j) = G(j,j) - G(i,j)*G(i,j)
    end
    G(j,j) = sqrt(G(j,j))
end
```

A seguir, apresentamos uma outra forma de construção do fator de Cholesky, através da comparação entre os elementos da parte triangular superior da equação matricial

$$A = G^t G .$$

Por simplicidade, vamos considerar a seguinte situação

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22} & a_{23} & a_{24} \\ & & a_{33} & a_{34} \\ & & & a_{44} \end{bmatrix} = \begin{bmatrix} g_{11} & & & \\ g_{12} & g_{22} & & \\ g_{13} & g_{23} & g_{33} & \\ g_{14} & g_{24} & g_{34} & g_{44} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ & g_{22} & g_{23} & g_{24} \\ & & g_{33} & g_{34} \\ & & & g_{44} \end{bmatrix}$$

Primeiramente calculamos os elementos da primeira linha do fator de Cholesky G, a partir da diagonal principal, da seguinte forma:

$$\begin{aligned} (g_{11})^2 &= a_{11} \quad &\Longrightarrow \quad g_{11} &= \sqrt{a_{11}} \\ g_{11}g_{12} &= a_{12} \quad &\Longrightarrow \quad g_{12} &= \frac{a_{12}}{g_{11}} \\ g_{11}g_{13} &= a_{13} \quad &\Longrightarrow \quad g_{13} &= \frac{a_{13}}{g_{11}} \\ g_{11}g_{14} &= a_{14} \quad &\Longrightarrow \quad g_{14} &= \frac{a_{14}}{g_{11}} \end{aligned}$$

Em seguida calculamos os elementos da segunda linha do fator de Cholesky G, a partir da diagonal principal, da seguinte forma:

$$\begin{aligned} (g_{12})^2 + (g_{22})^2 &= a_{22} \quad &\Longrightarrow \quad g_{22} &= \sqrt{a_{22} - (g_{12})^2} \\ g_{12}g_{13} + g_{22}g_{23} &= a_{23} \quad &\Longrightarrow \quad g_{23} &= \frac{a_{23} - g_{12}g_{13}}{g_{22}} \\ g_{12}g_{14} + g_{22}g_{24} &= a_{24} \quad &\Longrightarrow \quad g_{24} &= \frac{a_{24} - g_{12}g_{14}}{g_{22}} \end{aligned}$$

De modo análogo, calculamos os elementos da terceira linha do fator de Cholesky G, a partir da diagonal principal, da seguinte forma:

$$(g_{13})^2 + (g_{23})^2 + (g_{33})^2 = a_{33}$$

Assim, escolhemos o elemento g_{33} na forma:

$$g_{33} = \sqrt{a_{33} - (g_{13})^2 - (g_{23})^2}$$

Em seguida encontramos o elemento g_{34} na forma:

$$g_{13}g_{14} + g_{23}g_{24} + g_{33}g_{34} = a_{34} \implies g_{34} = \frac{a_{34} - g_{13}g_{14} - g_{23}g_{24}}{g_{22}}$$

Finalmente, calculamos o elemento g_{44} da seguinte forma:

$$(g_{14})^2 + (g_{24})^2 + (g_{34})^2 + (g_{44})^2 = a_{44}$$

Assim, escolhemos o elemento g_{44} na forma:

$$g_{44} = \sqrt{a_{44} - (g_{14})^2 - (g_{24})^2 - (g_{34})^2}$$

De um modo geral, temos as seguintes equações

$$a_{ij} = \sum_{k=1}^{i} g_{ki}\, g_{kj}$$

para $i = 1, 2, \ldots, n$ e $j = i, (i+1), \cdots, n$.

Organizando de forma adequada as equações acima, determinamos inicialmente o i–ésimo elemento da diagonal principal da matriz G através da equação

$$a_{ii} = (g_{ii})^2 + \sum_{k=1}^{i-1} (g_{ki})^2 .$$

Assim, escolhemos o elemento g_{ii} da seguinte forma:

$$g_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} (g_{ki})^2}$$

Em seguida, encontramos os elementos da i–ésima linha a partir da diagonal principal, através das seguintes equações

$$a_{ij} = g_{ii}g_{ij} + \sum_{k=1}^{i-1} g_{ki}\, g_{kj}$$

Desse modo, temos que os elementos g_{ij} são dados por:

$$g_{ij} = \frac{a_{ij} - \displaystyle\sum_{k=1}^{i-1} g_{ki}\, g_{kj}}{g_{ii}} \qquad \text{para} \qquad j = (i + 1), \ldots, n$$

e para todo $i = 1, 2, \ldots, n$.

Podemos observar que para determinar cada elemento g_{ii} necessitamos do escalar

$$\Delta \;=\; a_{ii} \;-\; \sum_{k=1}^{i-1} (\, g_{ki} \,)^2 \,.$$

Como $g_{ii} = \sqrt{\Delta}$, temos que se $\Delta < 0$ a matriz A não é positiva–definida. Assim, não existe seu fator de Cholesky.

Utilizando a Fatoração de Cholesky da matriz do sistema linear positivo–definido, obtemos sua solução resolvendo dois sistemas triangulares

$$\begin{cases} G^t\, y \;=\; b \\ \\ G\, x \;=\; y \end{cases} \tag{8.109}$$

No procedimento da fatoração de Cholesky o número de operações elementares realizadas é da ordem de $n^3/6$ e são necessárias n raízes quadradas. Em várias referências a decomposição de Cholesky é denominada método da raíz quadrada. No procedimento para obtenção da solução de um sistema triangular o número de operações elementares realizadas é da ordem de $n^2/2$. Mais a frente, apresentamos o algoritmo da fatoração de Cholesky e os algoritmos para a resolução dos sistemas triangulares. No algoritmo da fatoração de Cholesky o fator de Cholesky pode ser armazenado na parte triangular superior da própria matriz A, para economia de memória.

A caracterização de uma matriz positiva–definida fornecida pelo Teorema 8.2.5 é um resultando muito importante que utilizamos para obter vários outros resultados teóricos, entretanto, tem uma alta complexidade computacional. Assim, a fatoração de Cholesky é um procedimento simples, eficiente e de baixa complexidade computacional, pelo qual podemos verificar se uma matriz simétrica e positiva–definida.

Exemplo 8.7.1 *Verifique se a matriz simétrica A dada por:*

$$A \;=\; \begin{bmatrix} 16 & 4 & 8 \\ 4 & 10 & 2 \\ 8 & 2 & 8 \end{bmatrix}$$

é positiva–definida através da fatoração de Cholesky.

Exemplo 8.7.2 *Determine a solução do sistema linear positivo–definido*

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ 4 \end{bmatrix}$$

através da Decomposição de Cholesky.

O fator de Cholesky da matriz A é dado por:

$$G = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Assim, primeiramente temos que resolver o Sistema Triangular Inferior

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ 4 \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}.$$

Finalmente, resolvemos o Sistema Triangular Superior

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix},$$

obtendo a solução do sistema linear positivo–definido.

A seguir apresentamos o algoritmo da fatoração de Cholesky, o algoritmo para a resolução do sistema triangular inferior por substituição avançada descrito no Teorema 2.9.11, e o o algoritmo para a resolução do sistema triangular superior por substituição atrasada descrito no Teorema 2.9.12, obtendo assim a solução do sistema linear positivo–definido.

Algoritmo 8.7.2 (Fatoração de Cholesky)

```
for i = 1,2, ... , n
    soma = 0.0
    for j = 1,2, ... , (i - 1)
        soma = soma + G(j,i)*G(j,i)
    end
    delta = A(i,i) - soma
    G(i,i) = sqrt(delta)
         for j = (i+1), ... , n
             soma = 0.0
             for k = 1,2, ... , (i - 1)
                 soma = soma + G(k,j)*G(k,i)
             end
         G(i,j) = ( A(i,j) - soma ) / G(i,i)
         end
end
```

Algoritmo 8.7.3 (Sistemas Triangulares)

```
% Sistema Triangular Inferior G'y = b
for i = 1,2, ... , n
    soma = 0.0
    for j = 1,2, ... ,(i - 1)
        soma = soma + G(j,i)*y(j)
    end
    y(i) = ( b(i) - soma ) / G(i,i)
end

% Sistema Triangular Superior Gx = y
for i = n, ... , 1
    soma = 0.0
    for j = (i + 1), ... , n
        soma = soma + G(i,j)*x(j)
    end
    x(i) = ( y(i) - soma ) / G(i,i)
end
```

Exercícios

Exercício 8.29 *Sejam* $L \in \mathbb{M}_4(\mathbb{R})$ *uma matriz triangular inferior invertível e o elemento* $b \in \mathbb{R}^4$ *dados por:*

$$L = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 1 & 5 & 0 & 0 \\ 2 & 1 & 4 & 0 \\ 1 & 2 & 3 & 6 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 4 \\ 11 \\ 8 \\ 26 \end{bmatrix}.$$

Obter a solução do sistema linear triangular inferior $Lx = b$.

Exercício 8.30 *Sejam* $U \in \mathbb{M}_4(\mathbb{R})$ *uma matriz triangular superior invertível e o elemento* $b \in \mathbb{R}^4$ *dados por:*

$$U = \begin{bmatrix} 6 & 2 & 3 & 1 \\ 0 & 4 & 1 & 2 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 16 \\ 15 \\ 8 \\ 12 \end{bmatrix}.$$

Obter a solução do sistema linear triangular superior $Ux = b$.

Exercício 8.31 *Sejam* $L \in \mathbb{M}_n(\mathbb{R})$ *uma matriz triangular inferior invertível e o elemento* $b \in \mathbb{R}^n$. *Escrever um algoritmo para resolução do sistema linear triangular inferior*

$$Lx = b$$

por ***substituição avançada****.*

Exercício 8.32 *Sejam* $U \in \mathbb{M}_n(\mathbb{R})$ *uma matriz triangular superior invertível e o elemento* $b \in \mathbb{R}^n$. *Escrever um algoritmo para resolução do sistema linear triangular superior*

$$Ux = b$$

por ***substituição atrasada****.*

Exercício 8.33 *Seja* $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida. Mostre que os elementos da diagonal principal são todos positivos, isto é,*

$$a_{ii} > 0 \qquad para\ todo \qquad i = 1, \cdots, n\,.$$

Exercício 8.34 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida. Mostre que*

$$\det(A) > 0 .$$

Exercício 8.35 *Considere* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida. Mostre que a submatriz principal* A_k *é positiva–definida, para todo* $k = 1, \cdots, n$.

Exercício 8.36 *Considere* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida e uma matriz* $B \in \mathbb{M}_{n\times p}(\mathbb{R})$, *com* $n \geq p$ *e posto*$(B) = p$. *Mostre que a matriz*

$$C = B^t A B \in \mathbb{M}_p(\mathbb{R})$$

é positiva–definida.

Exercício 8.37 *Considere* $C \in \mathbb{M}_4(\mathbb{R})$ *uma matriz positiva–definida e o elemento* $b \in \mathbb{R}^4$ *dados por:*

$$C = \begin{bmatrix} 4 & -2 & 4 & 2 \\ -2 & 10 & -2 & -7 \\ 4 & -2 & 8 & 4 \\ 2 & -7 & 4 & 7 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 8 \\ 2 \\ 16 \\ 6 \end{bmatrix}.$$

Obter a solução do sistema linear $Cx = b$ *utilizando a Fatoração de Cholesky.*

Exercício 8.38 *Considere as seguintes matrizes simétricas*

$$A = \begin{bmatrix} 9 & 3 & 3 \\ 3 & 10 & 5 \\ 3 & 5 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 4 & 8 \\ 4 & -4 & 1 \\ 8 & 1 & 6 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Verifique quais matrizes são positiva–definidas.

Exercício 8.39 *Considere* $A \in \mathbb{M}_n(\mathbb{R})$ *simétrica invertível. Escreva um algoritmo para calcular sua decomposição na forma* $A = LDL^t$, *onde* D *é uma matriz diagonal e* L *é uma matriz triangular inferior com os elementos da diagonal principal iguais a 1.*

Exercício 8.40 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida. Qual a relação entre a fatoração* $A = LDL^t$ *e a fatoração de Cholesky* $A = G^tG$?

Exercício 8.41 *Considere $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz positiva–definida. Escreva um procedimento eficiente, e faça a implementação computacional, para calcular a matriz inversa A^{-1}, utilizando a fatoração de Cholesky da matriz A.*

Exercício 8.42 *Considere a matriz A dada por:*

$$A = \begin{bmatrix} \alpha & 1 & 0 \\ \beta & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Determine todos os valores dos parâmetros α e β para os quais

(a) A matriz A é singular.

(b) A matriz A é positiva–definida.

Exercício 8.43 *Sejam $A, B \in \mathbb{M}_n(\mathbb{R})$ matrizes positiva–definidas.*

(a) A matriz $A + B$ é positiva–definida?

(b) A matriz A^2 é positiva–definida?

(c) A matriz $A - B$ é positiva–definida?

Exercício 8.44 *Mostre que a função quadrática*

$$f(x,y,z) = 2x^2 - 2xy + 2y^2 - 2yz + 2z^2$$

possui um ponto de mínimo.

Exercício 8.45 *Sejam $A \in \mathbb{M}_3(\mathbb{R})$ uma matriz positiva–definida, $G \in \mathbb{M}_3(\mathbb{R})$ o fator de Cholesky da matriz A e o elemento $b \in \mathbb{R}^3$ dados por:*

$$A = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 2 & 5 \\ 6 & 5 & \beta \end{bmatrix} \quad , \quad G = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 1 & \alpha \\ 0 & 0 & 5 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} -2 \\ 3 \\ 30 \end{bmatrix}.$$

(a) Determine os parâmetros α e β.

(b) Calcule $\det(A)$.

(c) Determine a solução do sistema linear positivo–definido $Ax = b$.

Exercício 8.46 *Sejam* $A \in I\!M_n(I\!R)$ *uma matriz positiva–definida e* $G \in I\!M_n(I\!R)$ *o fator de Cholesky da matriz* A. *Mostre que* $\mathcal{K}_2(G) = \sqrt{\mathcal{K}_2(A)}$.

Exercício 8.47 *Considere a matriz positiva–definida* A *dada por:*

$$A = \begin{bmatrix} 4 & 2 & 12 & 2 \\ 2 & 17 & 6 & 1 \\ 12 & 6 & 72 & 6 \\ 2 & 1 & 6 & 10 \end{bmatrix}.$$

(a) Determine a fatoração de Cholesky $A = GG^t$.

(b) Determine uma matriz invertível P *de modo que*

$$D = PAP^t$$

seja uma matriz diagonal, isto é, P *é uma matriz que realiza a diagonalização da matriz* A *através da transformação de congruência, veja a seção 2.8.*

(c) Determine a matriz $L = P^{-1}$ *tal que* $A = LDL^t$.

(d) Determine a relação entre a fatoração $A = LDL^t$*, obtida no item (c), e a fatoração de Cholesky* $A = GG^t$.

Exercício 8.48 *Seja* $A \in I\!M_n(I\!R)$ *positiva–definida. Escreva um procedimento para determinar uma matriz triangular superior* $R \in I\!M_n(I\!R)$*, com os elementos da diagonal principal todos positivos, tal que* $A = RR^t$.

Exercício 8.49 *Considere que* $H = A + iB \in I\!M_n(\mathbb{C})$ *é uma matriz positiva–definida, onde* $A, B \in I\!M_n(I\!R)$*, isto é, para* $X \in \mathbb{C}^n$,

$$X^*HX > 0 \qquad \textit{sempre que} \qquad X \neq 0_{\mathbb{C}^n} .$$

(a) Mostre que a matriz real C *definida por:*

$$C = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$$

é uma matriz positiva–definida.

(b) Escreva um procedimento para determinar a solução do sistema linear

$$(A + iB)(x + iy) = (b + ic),$$

onde $x, y, b, c \in I\!R^n$.

8.8 Métodos Iterativos para Sistemas Lineares

Iteração de Ponto Fixo. Matriz Convergente

Seja $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz invertível e $b \in \mathbb{R}^n$. Considere o Sistema Linear: encontrar $x^* \in \mathbb{R}^n$ solução da equação

$$Ax = b \tag{8.110}$$

Como A é uma matriz invertível, o sistema linear (8.110) possui uma única solução, que vamos denotar por $x^* = A^{-1}b$. Podemos escrever o sistema linear (8.110) em uma forma equivalente

$$x = P\,x + d \tag{8.111}$$

Desse modo, um método iterativo consiste em considerar uma aproximação inicial, que vamos denotar por $x^{(0)} \in \mathbb{R}^n$, para a solução x^* e construir uma seqüência

$$x^{(k+1)} = P\,x^{(k)} + d \qquad \text{para} \qquad k = 0, 1, 2, \cdots \tag{8.112}$$

Neste ponto, podemos fazer as seguintes perguntas:

1. Qual a condição de convergência do processo iterativo ?
2. A seqüência $\left\{x^{(k)}\right\}_{k \in \mathbb{N}}$ converge para a solução do sistema linear (8.110) ?
3. A aproximação inicial $x^{(0)} \in \mathbb{R}^n$ pode ser arbitrária ?

Para responder as questões acima, vamos necessitar das seguintes definições e resultados.

Definição 8.8.1 (Raio Espectral) *Seja $A \in \mathbb{M}_n(\mathbb{F})$. Definimos o raio espectral da matriz A, que denotamos por $\rho(A)$, da seguinte forma:*

$$\rho(A) = \max_{1 \leq j \leq n} \{\, |\lambda_j| \quad ; \quad \lambda_j \ \text{autovalor de} \ A \,\}\,.$$

Definição 8.8.2 *Dizemos que a matriz $A \in \mathbb{M}_n(\mathbb{R})$ é convergente se*

$$\lim_{k \to \infty} A^k = 0\,,$$

onde $A^k = A\,A^{k-1}$ com $A^0 = I$.

Teorema 8.8.1 *Sejam* $A \in \mathbb{M}_n(\mathbb{F})$ *e* $\| \cdot \|$ *uma norma matricial consistente. Então,* $\rho(A) \leq \| A \|$.

Demonstração – Considere o autovalor λ_{max} da matriz A tal que $| \lambda_{max} | = \rho(A)$, e o elemento v o autovetor associado.

Seja $V \in \mathbb{M}_n(\mathbb{F})$ uma matriz cujas colunas são todas iguais ao autovetor v, isto é, $V = [v \cdots v \cdots v]$. Assim, temos que

$$\| \lambda_{max} V \| = | \lambda_{max} | \, \| V \| = \| A V \| \leq \| A \| \, \| V \|$$

Portanto, obtemos $| \lambda_{max} | \leq \| A \|$, o que completa a demonstração. ■

Teorema 8.8.2 *Sejam* $A \in \mathbb{M}_n(\mathbb{F})$ *e* $\epsilon > 0$. *Então, existe uma norma matricial consistente* $\| \cdot \|$ *tal que* $\| A \| \leq \rho(A) + \epsilon$.

Demonstração – Pelo Teorema 8.2.2, Teorema da Decomposição de Schur, existe uma matriz unitária $U \in \mathbb{M}_n(\mathbb{F})$ e uma matriz triangular superior $T \in \mathbb{M}_n(\mathbb{F})$ tais que $A = U T U^*$.

Considere uma matriz diagonal $D = diag(\delta, \ldots, \delta^n)$ para $\delta \in \mathbb{R}_+$. Construímos a partir da matriz D e da decomposição de Schur da matriz A, a seguinte matriz

$$D T D^{-1} = \begin{cases} \lambda_j & \text{na diagonal principal} \\[2ex] \dfrac{t_{ij}}{\delta^{j-i}} & \text{fora da diagonal principal} \end{cases}$$

Note que, λ_j para $j = 1, \ldots, n$, são os autovalores da matriz A que aparecem na diagonal principal da matriz T. Para um determinado valor $\epsilon > 0$ escolhemos um valor adequado para o parâmetro δ de modo que

$$\sum_{j=i+1}^{n} \left| \frac{t_{ij}}{\delta^{j-i}} \right| \leq \epsilon \qquad \text{para todo} \qquad i = 1, \ldots, (n-1)$$

Assim, vamos definir uma norma matricial $\| \cdot \|$ em $\mathbb{M}_n(\mathbb{F})$ da seguinte forma:

$$\| B \| = \| (D U^*) B (U D^{-1}) \|_{\infty} \qquad \text{para toda} \qquad B \in \mathbb{M}_n(\mathbb{F})$$

que claramente depende da matriz A e do valor de ϵ.

Desse modo, construímos uma norma matricial $|||\cdot|||$ tal que

$$|||A||| \;=\; |||D\,T\,D^{-1}|||_{\infty} \;\leq\; \rho(A) \,+\, \epsilon$$

Note que, a norma matricial $|||\cdot|||$ é induzida pela seguinte norma vetorial

$$\|\,(\,D\,U^{*}\,)\,x\,\|_{\infty} \qquad \text{para todo} \qquad x \in \mathbb{F}^{n} \,,$$

o que completa a demonstração. ■

Teorema 8.8.3 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$. *As seguintes afirmações são equivalentes:*

(1) A *é uma matriz convergente.*

(2) $\lim\limits_{k \to \infty} A^k\,x \;=\; 0 \quad ; \quad \forall x \in \mathbb{R}^n$.

(3) O raio espectral da matriz A *satisfaz* $\rho(A) \,<\, 1$.

(4) Existe uma norma matricial $|||\cdot|||$ *tal que* $|||A||| \,<\, 1$.

Demonstração – Para a prova vamos considerar as normas $\|\cdot\|$ e $|||\cdot|||$ compatíveis, isto é, satisfazendo o Teorema 8.3.4. Em espaço vetorial normado de dimensão finita as normas são equivalentes. Assim a convergência de uma seqüência independe da escolha da norma. Note que estamos usando também o fato que as normas $\|\cdot\|$ e $|||\cdot|||$ são funções contínuas.

Inicialmente, vamos mostrar que a condição (*1*) implica na condição (*2*). Para isso, tomamos

$$\lim_{k \to \infty} \|\,A^k\,x\,\| \;\leq\; \lim_{k \to \infty} |||A^k|||\,\|\,x\,\| \,.$$

Pela hipótese, como A é uma matriz convergente, temos que

$$\lim_{k \to \infty} A^k \,=\, 0 \quad \Longrightarrow \quad \lim_{k \to \infty} |||A^k||| \,=\, 0 \quad \Longrightarrow \quad \lim_{k \to \infty} A^k\,x \,=\, 0 \,,$$

provando que a condição (*1*) implica na condição (*2*).

Vamos mostrar agora que a condição (*2*) implica na condição (*3*). Para isso, supomos que $\rho(A) \,\geq\, 1$. Tomando um autopar $(\lambda\,,\,v)$ da matriz A, temos

$$0 \;=\; \lim_{k \to \infty} \|\,A^k\,v\,\| \;=\; \lim_{k \to \infty} |\,\lambda\,|^k\,\|\,v\,\| \;\neq\; 0 \,,$$

que é uma contradição. Logo, $\rho(A) < 1$, mostrando que a condição (*2*) implica na condição (*3*).

Vamos mostrar que a condição (*3*) implica na condição (*4*). Se $\rho(A) < 1$, utilizando o resultado do Teorema 8.8.2, podemos escolher de maneira conveniente uma constante $\epsilon > 0$ e uma norma consistente $\|\!| \cdot |\!\|$ tais que $\|\!| A |\!\| \leq \rho(A) + \epsilon < 1$. Assim, mostramos que a condição (*3*) implica na condição (*4*).

Finalmente, vamos mostrar que a condição (*4*) implica na condição (*1*). Considerando que $\|\!| A |\!\| < 1$, obtemos

$$\lim_{k\to\infty} \|\!| A^k |\!\| \leq \lim_{k\to\infty} \|\!| A |\!\|^k = 0 \implies \lim_{k\to\infty} A^k = 0 \,,$$

provando que a condição (*4*) implica na condição (*1*), o que completa a demonstração. ■

Teorema 8.8.4 *Considere* $A, \in, I\!M_n(I\!R)$.

(1) A série geométrica

$$I + A + A^2 + \cdots + A^k + \cdots$$

converge se, e somente se, a matriz A *for convergente.*

(2) Se $A \in I\!M_n(I\!R)$ *é uma matriz convergente, então* $(I - A)$ *é invertível, e*

$$(I - A)^{-1} = I + A + A^2 + \cdots + A^k + \cdots$$

Demonstração – Basta provar o item (*2*), que o item (*1*) fica automaticamente provado. Como A é uma matriz convergente, sabemos que $\rho(A) < 1$. Assim, os autovalores da matriz $(I - A)$ são tais que $(1 - \lambda) \neq 0$, onde λ é um autovalor da matriz A. Desse modo, temos que a matriz $(I - A)$ é invertível. Considerando a identidade

$$(I - A)(I + A + A^2 + \cdots + A^k) = I - A^{k+1} \,,$$

obtemos

$$(I - A) \lim_{k\to\infty} (I + A + A^2 + \cdots + A^k) = \lim_{k\to\infty} (I - A^{k+1}) = I \,,$$

o que completa a demonstração. ■

Teorema 8.8.5 *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *e* $\|\cdot\|$ *uma norma induzida, com* $\|A\| < 1$. *Então,*

$$\frac{1}{1 + \|A\|} \leq \|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

Demonstração – Como $\|A\| < 1$, sabemos que $(I - A)$ é uma matriz invertível, e

$$1 = \|I\| = \|(I - A)(I - A)^{-1}\| \leq (1 + \|A\|)\,\|(I - A)^{-1}\|,$$

obtemos

$$\|(I - A)^{-1}\| \geq \frac{1}{1 + \|A\|}.$$

Consideremos agora a identidade

$$(I - A)^{-1} = I + A(I - A)^{-1},$$

tem–se que

$$\|(I - A)^{-1}\| \leq 1 + \|A\|\,\|(I - A)^{-1}\|.$$

Agrupando os termos em comum, temos

$$(1 - \|A\|)\,\|(I - A)^{-1}\| \leq 1,$$

resultando na desigualdade

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|},$$

o que completa a demonstração. ■

Método de Jacobi

Inicialmente, vamos descrever o *Método iterativo de Jacobi*. Para isso, representamos a matriz do sistema linear da seguinte forma $A = L + D + U$, onde $L = [l_{ij}]$ é a matriz triangular inferior dada por:

$$l_{ij} = \begin{cases} a_{ij} & ; \quad i > j \\ \\ 0 & ; \quad i \leq j \end{cases} \tag{8.113}$$

$D = [d_{ij}]$ é a matriz diagonal dada por:

$$d_{ij} = \begin{cases} a_{ij} & ; \quad i = j \\ \\ 0 & ; \quad i \neq j \end{cases} \tag{8.114}$$

$U = [u_{ij}]$ é a matriz triangular superior dada por:

$$u_{ij} = \begin{cases} a_{ij} & ; \quad i < j \\ \\ 0 & ; \quad i \geq j \end{cases} \tag{8.115}$$

Assim, o sistema linear (8.110) pode ser escrito da seguinte forma:

$$(L + D + U)x = b \tag{8.116}$$

$$Dx = -(L + U)x + b \tag{8.117}$$

$$x = -D^{-1}(L + U)x + D^{-1}b \tag{8.118}$$

Portanto, temos que no processo iterativo (8.112), a matriz $P = -D^{-1}(L + U)$ e o vetor $d = D^{-1}b$, desde que D seja uma matriz invertível. Com o objetivo de escrever o algoritmo do Método de Jacobi, vamos representa–lo da seguinte forma:

$$x_i^{(k+1)} = \frac{\left(b_i - \sum_{j=1}^{i-1} a_{ij}\, x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}\, x_j^{(k)} \right)}{a_{ii}} \quad ; \quad i = 1, 2, \ldots, n \tag{8.119}$$

para $k = 0, 1, 2, \cdots$

Desse modo, escrevemos o algoritmo do Método Iterativo de Jacobi, como segue abaixo.

Algoritmo 8.8.1 (Método Iterativo de Jacobi)

Dada uma aproximação inicial $x^{(0)} \in I\!R^n$ *para* $x^* = A^{-1} b$

for $k = 0, 1, 2, 3, \cdots$

for $i = 1, 2, 3, \cdots, n$

$x_i^{(k+1)} = b_i$

for $j = 1, \cdots, (i-1)$

$x_i^{(k+1)} = x_i^{(k+1)} - a_{ij}\, x_j^{(k)}$

end

for $j = (i+1), \cdots, n$

$x_i^{(k+1)} = x_i^{(k+1)} - a_{ij}\, x_j^{(k)}$

end

$$x_i^{(k+1)} = \frac{x_i^{(k+1)}}{a_{ii}}$$

end

end

Note que podemos escrever o processo iterativo da seguinte forma:

dada uma aproximação inicial $x^{(0)} = d = D^{-1} b$, tem–se que

$$x^{(k+1)} = (I + P + P^2 + \cdots + P^{k+1}) d \qquad \text{para} \qquad k = 0, 1, 2, \cdots \tag{8.120}$$

Logo, a seqüência gerada pelo Método de Jacobi converge se, e somente se, P for uma matriz convergente. Vamos denotar o ponto de convergência por $\widehat{x}$, que é dado por:

$$\lim_{k \to \infty} x^{(k+1)} = \widehat{x} = (I - P)^{-1} d \,. \tag{8.121}$$

Agora, basta mostrar que o elemento $\widehat{x}$ é a única solução do sistema linear (8.110).

Substituindo as expressões da matriz P e do vetor d na equação (8.121), obtemos

$$\widehat{x} = (I + D^{-1} (L + U))^{-1} (D^{-1} b) \tag{8.122}$$

$$= (D (I + D^{-1} (L + U)))^{-1} b \tag{8.123}$$

$$= (D + L + U)^{-1} b = A^{-1} b = x^* \tag{8.124}$$

Portanto, mostramos que o processo iterativo de Jacobi converge para a única solução x^* do sistema linear (8.110). Além disso, podemos observar que a convergência do Método de Jacobi não depende da escolha da aproximação inicial. Podemos enunciar o seguinte resultado de convergência do Método de Jacobi.

Teorema 8.8.6 *O Método Iterativo de Jacobi converge para a solução exata do sistema linear* (8.110) *para qualquer aproximação inicial* $x^{(0)}$ *se, e somente se, a matriz de iteração* $P = -D^{-1} A + I$ *for convergente.*

Vamos mostrar que a seqüência $\{x^{(k)}\}$ gerada pelo processo iterativo de Jacobi possui uma convergência linear, isto é,

$$\| x^{(k+1)} - x^* \| \leq \beta \| x^{(k)} - x^* \| ,$$

quando da convergência do método, onde a constante β é a taxa de convergência.

De fato, vamos considerar o processo iterativo de Jacobi

$$x^{(k+1)} = P x^{(k)} + d ,$$

onde $P = -D^{-1} A + I$ e $d = D^{-1} b$. Note que, a solução exata x^* satisfaz $x^* = P x^* + d$. Desse modo, temos que

$$x^{(k+1)} - x^* = P (x^{(k)} - x^*) .$$

Escolhendo de forma conveniente as normas $\| \cdot \|$ e $|\!|\!| \cdot |\!|\!|$ compatíveis, isto é, satisfazendo o Teorema 8.3.4, obtemos

$$\| x^{(k+1)} - x^* \| = \| P (x^{(k)} - x^*) \| \leq |\!|\!| P |\!|\!| \, \| x^{(k)} - x^* \| .$$

Portanto, temos a convergência linear do Método de Jacobi, com a taxa de convergência $\beta = |\!|\!| P |\!|\!|$, desde que $|\!|\!| P |\!|\!| < 1$. Este resultado mostra a velocidade de convergência do processo iterativo de Jacobi. Quanto menor for $|\!|\!| P |\!|\!|$ mais rápida será a convergência.

Devemos observar que o fato da matriz P ser convergente é equivalente à existência de uma norma de modo que $\| P \| < 1$, ou ainda, é equivalente a dizer que o raio espectral da matriz P satisfaz $\rho(P) < 1$. Todos essas propriedades são muito úteis do ponto de vista teórico, mas na prática tem um alto custo computacional para determina–las. Vamos mostrar condições mais simples de ser verificadas para a convergência do Método de Jacobi. Para isso, necessitamos dos seguintes conceitos.

Definição 8.8.3 *Dizemos que* $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz* ***Estritamente Diagonalmente Dominante por Linhas*** *se*

$$| a_{ii} | > \sum_{\substack{j=1 \\ j \neq i}}^{n} | a_{ij} | \qquad ; \qquad i = 1, \ldots, n \,.$$

Teorema 8.8.7 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz Estritamente Diagonalmente Dominante por Linhas. Então,* A *é invertível.*

Demonstração – Vamos considerar a matriz $B = I - D^{-1} A$, onde D é a matriz diagonal descrita em (8.114). Pela propriedade da matriz A, temos que D é invertível. Podemos observar que $\| B \|_\infty < 1$. Assim, pelo Teorema 8.8.4 da seção 8.8, temos que a matriz $I - B = D^{-1} A$ é não singular. Desse modo, podemos concluir que A é uma matriz invertível, o que completa a demonstração. ■

Desse modo, temos o seguinte resultado de convergência para o Método de Jacobi, que é facilmente verificado.

Teorema 8.8.8 *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *Estritamente Diagonalmente Dominante por Linhas e* $P = -D^{-1}A + I$ *a matriz de iteração de Jacobi. Então,* $\| P \|_\infty < 1$.

Utilizando o conceito abaixo, vamos propor uma modificação no método de Jacobi, e mostrar a sua convergência para a solução do sistema linear.

Definição 8.8.4 *Dizemos que* $A = [a_{ij}] \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz* ***Estritamente Diagonalmente Dominante por Colunas*** *se*

$$| a_{jj} | > \sum_{\substack{i=1 \\ i \neq j}}^{n} | a_{ij} | \qquad ; \qquad j = 1, \ldots, n \, .$$

Teorema 8.8.9 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz Estritamente Diagonalmente Dominante por Colunas. Então,* A *é invertível.*

Demonstração – A prova é feita de forma análoga a do Teorema 8.8.7, considerando a matriz $B = I - A D^{-1}$. □

Note que no caso da matriz A for estritamente diagonalmente dominante por linhas, todos os elementos da diagonal principal são não–nulos. Logo, não ocorre uma divisão por zero no procedimento do Método de Jacobi.

Considerando que a matriz $A \in \mathbb{M}_n(\mathbb{R})$ do sistema linear (8.110) seja estritamente diagonalmente dominante por colunas, vamos propor um *Método de Jacobi Modificado*, que é descrito da seguinte forma:

dada uma aproximação inicial $y^{(0)} \in \mathbb{R}^n$, construímos a seguinte seqüência

$$y^{(k+1)} = P y^{(k)} + b \qquad \text{para} \qquad k = 0, 1, 2, \ldots \, , \tag{8.125}$$

onde $P = -(L + U) D^{-1}$ é a matriz de iteração.

É fácil mostrar que se a matriz P for convergente, então a seqüência descrita em (8.125) converge para $y^* = D x^*$, onde $x^* = A^{-1} b$ é a única solução do sistema linear.

A seguir temos um resultado de convergência do Método de Jacobi Modificado.

Teorema 8.8.10 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz Estritamente Diagonalmente Dominante por Colunas e* $P = -A D^{-1} + I$ *a matriz de iteração para o Método de Jacobi Modificado. Então,* $\| P \|_1 < 1$.

Exemplo 8.8.1 *Para fazer uma apresentação do desempenho do Método Iterativo de Jacobi, vamos considerar o sistema linear* $A\,x = b$

$$A = \begin{bmatrix} 100 & 30 & 20 & 30 \\ 10 & 100 & 20 & 40 \\ 10 & 10 & 60 & 20 \\ 5 & 10 & 5 & 30 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 205 \\ 190 \\ 140 \\ 54 \end{bmatrix}$$

Como A *é estritamente diagonalmente dominante por linhas o que implica na convergência do Método Iterativo de Jacobi. Considerando a aproximação inicial* $x^{(0)} = b$*, foram realizadas* 54 *iterações no Método de Jacobi, para obter uma solução numérica do sistema linear com um erro relativo de* 8.8596×10^{-7}.

Método de Gauss–Seidel

Observamos que no Método de Jacobi quando vamos calcular a i–ésima componente da solução na $(k+1)$–ésima iteração, todas as componentes anteriores já estão com uma nova aproximação, que no caso de convergência do processo iterativo, estas componentes estarão mais próximas da solução. Assim, com o objetivo de acelerar a convergência do processo iterativo de Jacobi, utilizamos as componentes anteriores a i–ésima componente na $(k+1)$–ésima iteração. Desse modo, temos o *Método de Gauss–Seidel*

$$x_i^{(k+1)} = \frac{\left(b_i - \sum_{j=1}^{i-1} a_{ij}\, x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}\, x_j^{(k)} \right)}{a_{ii}} \quad ; \quad i = 1, 2, \ldots, n \tag{8.126}$$

para $k = 0, 1, 2, \cdots$. A seguir, apresentamos o algoritmo do Método de Gauss–Seidel.

Algoritmo 8.8.2 (Método Iterativo de Gauss–Seidel)

Dada uma aproximação inicial $x^{(0)} \in I\!R^n$ *para* $x^* = A^{-1}\, b$

for $k = 0, 1, 2, 3, \cdots$

for $i = 1, 2, 3, \cdots, n$

$x_i^{(k+1)} = b_i$

for $j = 1, \cdots, (i-1)$

$x_i^{(k+1)} = x_i^{(k+1)} - a_{ij}\, x_j^{(k+1)}$

end

for $j = (i+1), \cdots, n$

$x_i^{(k+1)} = x_i^{(k+1)} - a_{ij}\, x_j^{(k)}$

end

$x_i^{(k+1)} = \dfrac{x_i^{(k+1)}}{a_{ii}}$

end

end

Note que as mesmas condições de convergência para o Método de Jacobi também servem para o Método de Gauss–Seidel, isto é, a convergência do Método de Jacobi implica na convergência do Método de Gauss–Seidel. Podemos representar o Método de Gauss–Seidel na forma matricial, para uma análise mais teórica, da seguinte forma:

$$x^{(k+1)} = G\,x^{(k)} + d \qquad \text{para} \qquad k = 0, 1, 2, \cdots , \tag{8.127}$$

onde a matriz de iteração é dada por:

$$G = -(\,L + D\,)^{-1}\,U$$

e o vetor $d = (\,L + D\,)^{-1}\,b$, desde que a matriz $(\,L + D\,)$ seja invertível. Assim, podemos enunciar o seguinte resultado de convergência do Método Iterativo de Gauss–Seidel.

Teorema 8.8.11 *O Método de Gauss–Seidel converge para a única solução do sistema linear* (8.110) *para qualquer aproximação inicial* $x^{(0)}$ *se, e somente se, a matriz de iteração* G *for convergente.*

Como no Método de Jacobi, o Método de Gauss–Seidel possui uma convergência linear, onde a taxa de convergência é dada pela constante $\beta = \| G \|$, desde que $0 < \| G \| < 1$.

Nos sistemas lineares provenientes da discretização de problemas de valores contorno elípticos, tanto pelo Método de Diferenças Finitas quanto pelo Método dos Elementos Finitos, a matriz do sistema linear é positiva–definida. Assim, apresentamos o seguinte resultado de convergência para o Método de Gauss–Seidel.

Teorema 8.8.12 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *é simétrica e positiva–definida. Então, o Método de Gauss–Seidel converge para a única solução do sistema linear* $A\,x = b$, $b \in \mathbb{R}^n$, *para toda aproximação inicial* $x^{(0)} \in \mathbb{R}^n$.

Demonstração – Como A é uma matriz simétrica, temos a decomposição

$$A = L + D + L^t\,.$$

Assim, a matriz de iteração do Método de Gauss-Seidel é dada por:

$$G = -(\,L + D\,)^{-1}\,L^t\,.$$

Vamos mostrar que o raio espectral da matriz G satisfaz $\rho(G) < 1$.

Inicialmente definimos uma matriz auxiliar G_1 da seguinte forma:

$$
\begin{aligned}
G_1 &= D^{\frac{1}{2}}\, G\, D^{-\frac{1}{2}} \;=\; -D^{\frac{1}{2}}\,(\, L \,+\, D\,)^{-1}\, L^t\, D^{-\frac{1}{2}} \\
&= -(\,(\, L \,+\, D\,)\, D^{-\frac{1}{2}}\,)^{-1}\, L^t\, D^{-\frac{1}{2}} \\
&= -(\, D^{\frac{1}{2}}\, D^{-\frac{1}{2}}\,(\, L \,+\, D\,)\, D^{-\frac{1}{2}}\,)^{-1}\, L^t\, D^{-\frac{1}{2}} \\
&= -\,(\, I \,+\, L_1\,)^{-1}\, L_1^t
\end{aligned}
\tag{8.128}
$$

onde a matriz auxiliar L_1 é dada por:

$$
L_1 \;=\; D^{-\frac{1}{2}}\, L\, D^{-\frac{1}{2}}\,. \tag{8.129}
$$

Como a matriz auxiliar G_1 foi obtida da matriz G através de uma transformação de similaridade, temos que elas possuem os mesmos autovalores, isto é, o mesmo raio espectral. Assim, temos que mostrar que a matriz G_1 satisfaz $\rho(G_1) \;<\; 1$. Para isso, consideramos (λ, v) um autopar da matriz G_1, com v unitário. Utilizando a equação (8.128), temos que

$$
-L_1^t\, v \;=\; \lambda\,(\, I \,+\, L_1\,)\, v\,. \tag{8.130}
$$

Fazendo o produto interno de ambos os membros de (8.130) pelo autovetor v, obtemos

$$
-\langle\, L_1^t v\,,\, v\,\rangle \;=\; \lambda\,(\, 1 \,+\, \langle\, L_1 v\,,\, v\,\rangle\,)\,. \tag{8.131}
$$

Chamando $\langle\, L_1 v\,,\, v\,\rangle \,=\, a \,+\, i\,b$, da equação (8.131), tem–se que

$$
|\,\lambda\,|^2 \;=\; \frac{a^2 \,+\, b^2}{1 \,+\, 2\,a \,+\, a^2 \,+\, b^2}\,. \tag{8.132}
$$

Vamos observar que

$$
D^{-\frac{1}{2}}\, A\, D^{-\frac{1}{2}} \;=\; I \,+\, L_1 \,+\, L_1^t\,. \tag{8.133}
$$

Como as matrizes A e D são positiva–definidas, a matriz definida em (8.133) também é positiva–definida. Desse modo, temos que

$$
1 \,+\, \langle\, L_1 v\,,\, v\,\rangle \,+\, \langle\, L_1^t v\,,\, v\,\rangle \;=\; 1 \,+\, 2\,a \;>\; 0\,. \tag{8.134}
$$

Portanto, do resultado (8.134) e da relação (8.132), mostramos que o raio espectral da matriz G_1 satisfaz $\rho(G_1) \;<\; 1$. O que é equivalente a $\rho(G) \;<\; 1$. Desse modo, a matriz de iteração de Gauss–Seidel é convergente, o que completa a demonstração. ■

Método da Relaxação Sucessiva

O Método de Gauss–Seidel é muito atrativo pela sua simplicidade. Entretanto, em geral, o raio espectral da sua matriz de iteração é muito próximo da unidade, o que provoca uma convergência muito lenta. Com o objetivo de acelerar sua convergência propomos uma modificação. Seja $\omega \in I\!R$ um parâmetro arbitrário, consideramos o procedimento

$$x_i^{(k+1)} = \omega \frac{\left(b_i - \sum_{j=1}^{i-1} a_{ij}\, x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}\, x_j^{(k)} \right)}{a_{ii}} + (1 - \omega)\, x_i^{(k)} \tag{8.135}$$

para $i = 1, 2, \cdots, n$ e $k = 0, 1, 2, \cdots$. Este procedimento é denominado *Método da Relaxação Sucessiva*, cujo algoritmo apresentamos a seguir.

Algoritmo 8.8.3 (Método da Relaxação Sucessiva)

Dada uma aproximação inicial $x^{(0)} \in I\!R^n$ *para* $x^* = A^{-1}\, b$

for $k = 0, 1, 2, 3, \cdots$

  for $i = 1, 2, 3, \cdots, n$

  $x_i^{(k+1)} = b_i$

    for $j = 1, \cdots, (i-1)$

    $x_i^{(k+1)} = x_i^{(k+1)} - a_{ij}\, x_j^{(k+1)}$

    end

    for $j = (i+1), \cdots, n$

    $x_i^{(k+1)} = x_i^{(k+1)} - a_{ij}\, x_j^{(k)}$

    end

  $$x_i^{(k+1)} = \omega \frac{x_i^{(k+1)}}{a_{ii}} + (1 - \omega)\, x_i^{(k)}$$

  end

end

Vamos representar o Método da Relaxação Sucessiva na forma matricial, para uma análise mais teórica, da seguinte forma:

$$M_\omega \, x^{(k+1)} \;=\; N_\omega \, x^{(k)} \;+\; \omega \, b \qquad \text{para} \qquad k = 0, 1, 2, \cdots \,, \tag{8.136}$$

onde $M_\omega = D + \omega L$ e $N_\omega = (1 - \omega) D - \omega U$.

Assim a matriz de iteração é dada por:

$$G_\omega \;=\; M_\omega^{-1} N_\omega \,,$$

desde que a matriz M_ω seja invertível. No caso em que a matriz G_ω for convergente, isto é, $\rho(G_\omega) < 1$, o processo iterativo (8.136) converge para a única solução do sistema linear (8.110), para qualquer aproximação inicial $x^{(0)}$. Desse modo, a função do parâmetro ω é de minimizar $\rho(G_\omega)$.

Exemplo 8.8.2 *Para exemplificar a função do parâmetro ω, vamos considerar o sistema linear $Ax = b$, onde a matriz A é dada por:*

$$A \;=\; \begin{bmatrix} 100 & 30 & 20 & 30 \\ 10 & 100 & 20 & 40 \\ 10 & 10 & 60 & 20 \\ 5 & 10 & 5 & 30 \end{bmatrix}$$

Podemos verificar que a matriz A é estritamente diagonalmente dominante por linhas. Desse modo, temos a convergência do Método de Gauss–Seidel. Na Figura 8.8, temos o comportamento do raio espectral da matriz G_ω em função do parâmetro de relaxação ω. Para $\omega = 1.0$ temos $\rho(G_\omega) = 0.2056$, que é a matriz de iteração de Gauss–Seidel. Para o Método da Relaxação Sucessiva, temos que $\rho(G_\omega) \geq \rho(G)$. Este resultado é devido ao fato da matriz A ser estritamente diagonalmente dominante por linhas.

Para melhor observar o resultado obtido, vamos considerar um vetor b dado por:

$$b \;=\; \begin{bmatrix} 205 \\ 190 \\ 140 \\ 54 \end{bmatrix}.$$

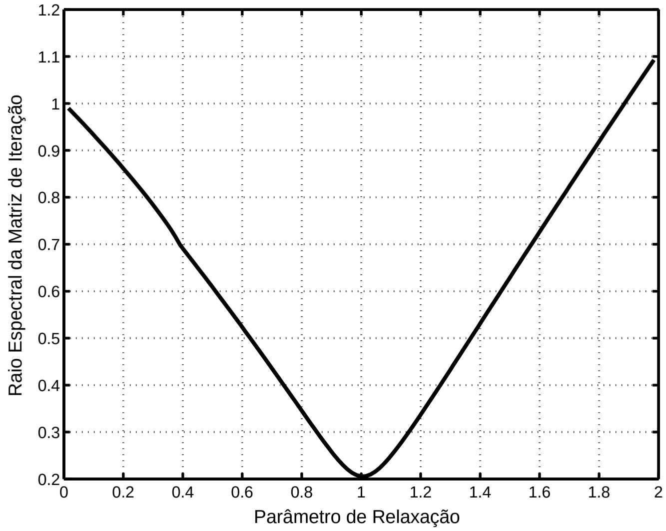

Figura 8.8: Raio espectral da matriz G_ω em função do parâmetro ω.

Considerando a aproximação inicial $x^{(0)} = b$*, foram realizadas* 13 *iterações no Método de Gauss–Seidel, para obter uma solução numérica do sistema linear com um erro relativo de* 8.913×10^{-7}*. No Método Iterativo de Jacobi foram realizadas* 54 *iterações para obter a mesma precisão, com uma taxa de convergência* $\rho(P) = 0.6989$*. Este exemplo mostra que o Método de Gauss–Seidel é uma aceleração para o Método de Jacobi.*

Exemplo 8.8.3 *Consideramos um sistema linear onde a matriz é dada por:*

$$A = \begin{bmatrix} 2.1 & -1.0 & 0 & 0 & 0 & 0 & 0 & -1.0 \\ -1.0 & 2.1 & -1.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1.0 & 2.1 & -1.0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.0 & 2.1 & -1.0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1.0 & 2.1 & -1.0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1.0 & 2.1 & -1.0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1.0 & 2.1 & -1.0 \\ -1.0 & 0 & 0 & 0 & 0 & 0 & -1.0 & 2.1 \end{bmatrix}.$$

Neste caso a matriz A *é positiva–definida. Pelo Teorema 8.8.12, temos que o processo iterativo de Gauss–Seidel é convergente. Na Figura 8.9, temos o comportamento do raio espectral da matriz* G_ω *em função do parâmetro de relaxação* ω*. Para* $\omega = 1.0$ *temos* $\rho(G_\omega) = 0.9082$*, que é a matriz de iteração de Gauss–Seidel. Para* $\omega = 1.5313$ *temos* $\rho(G_\omega) = 0.6991$*, que é o melhor parâmetro para o Método da Relaxação Sucessiva.*

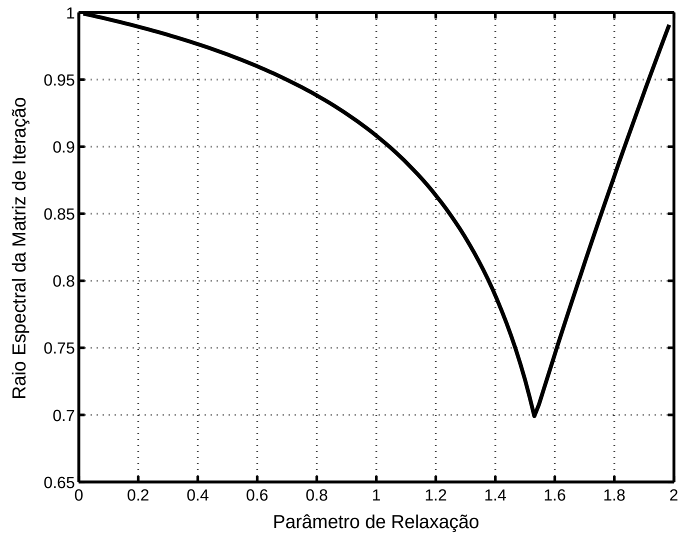

Figura 8.9: Raio espectral da matriz G_ω em função do parâmetro ω.

Para melhor observar o resultado obtido, vamos considerar um vetor b *dado por:*

$$b = \begin{bmatrix} 4.1 \\ -4.1 \\ 4.1 \\ -4.1 \\ 4.1 \\ -4.1 \\ 4.1 \\ -4.1 \end{bmatrix}.$$

Considerando a aproximação inicial $x^{(0)} = b$*, foram realizadas* 118 *iterações no Método de Gauss–Seidel, para obter uma solução numérica do sistema linear com um erro relativo de* 9.4871×10^{-7}*. Para o Método da Relaxação Sucessiva com* $\omega = 1.5313$*, foram realizadas* 45 *iterações para obter uma solução numérica com um erro relativo de* 9.7103×10^{-7}*. Note que o Método da Relaxação Sucessiva teve um desempenho muito superior ao Método de Gauss–Seidel, mostrando a real função do parâmetro* ω*.*

Finalmente, concluímos que a pesquisa do parâmetro ótimo para o Método da Relaxação Sucessiva, para uma determinada matriz associada a um problema prático, deve ser feita em duas situações. A primeira situação é quando temos vários sistemas lineares com a mesma matriz mudando somente o vetor do lado direito. A segunda situação é no caso em que não temos a convergência do Método de Gauss–Seidel.

Exercícios

Exercício 8.50 *Mostre que o Método Iterativo de Jacobi pode ser escrito da forma:*

$$x^{(k+1)} = x^{(k)} + M\,r^{(k)} ,$$

onde $M \in \mathbb{M}_n(\mathbb{R})$ *e* $r^{(k)} = b - A\,x^{(k)}$*. Faça uma comparação com o Método do Gradiente Otimizado, no caso em que a matriz* A *seja simétrica e positiva–definida.*

Exercício 8.51 *Necessita-se adubar um terreno acrescentando a cada* $10m^2$*,* $140g$ *de nitrato,* $190g$ *de fosfato e* $205g$ *de potássio. Dispõe-se de quatro qualidades de adubos com as seguintes características*

(i) Cada quilograma do adubo I custa 5 upc e contém $10g$ *de nitrato,* $10g$ *de fosfato e* $100g$ *de potássio.*

(ii) Cada quilograma do adubo II custa 10 upc e contém $10g$ *de nitrato,* $100g$ *de fosfato e* $30g$ *de potássio.*

(iii) Cada quilograma do adubo III custa 5 upc e contém $60g$ *de nitrato,* $20g$ *de fosfato e* $20g$ *de potássio.*

(iv) Cada quilograma do adubo IV custa 30 upc e contém $20g$ *de nitrato,* $40g$ *de fosfato e* $30g$ *de potássio.*

Quanto de cada adubo devemos misturar para conseguir uma boa aplicação de adubo se desejamos gastar somente $54\,upc$ *a cada* $10m^2$ *?*

Escrever o modelo matemático. Podemos obter uma solução numérica através do Método Iterativo de Gauss–Seidel ? Justifique a sua resposta. Em caso afirmativo obter uma solução numérica com um erro relativo inferior a 10^{-3}.

Exercício 8.52 *Considere* $A \in \mathbb{M}_n(\mathbb{R})$ *e a decomposição* $A = M - N$*, com* M *uma matriz invertível. Se* A *é uma matriz singular, então* $\rho(M^{-1}N) \geq 1$.

Exercício 8.53 *Considere* $A \in \mathbb{M}_n(\mathbb{R})$ *e* $b \in \mathbb{R}^n$*, e a decomposição* $A = M - N$*, com* M *uma matriz invertível. Se* $M\,x^{(k+1)} = N\,x^{(k)} + b$ *converge para* $x^* = A^{-1}b$*, então* $\rho(M^{-1}N) < 1$.

Exercício 8.54 *Faça a demonstração do Teorema 8.8.9, apresentando todos os detalhes.*

Exercício 8.55 *Sejam* $L \in I\!M_n(I\!R)$ *triangular inferior invertível e* $b \in I\!R^n$. *Podemos aplicar o Método de Jacobi para obter uma solução de* $L\,x = b$ *? Em caso afirmativo, quantas iterações serão necessárias ? Idem para o Método de Gauss–Seidel.*

Exercício 8.56 *Considere o Problema de Valor de Contorno com condição periódica*

$$\begin{aligned} -u''(x) \;+\; \sigma\, u(x) &= f(x) \qquad ; \qquad x \in (0, L) \\ u(0) &= u(L) \end{aligned} \tag{8.137}$$

com $\sigma > 0$ *e* f *uma função contínua.*

O sistema linear proveniente da discretização do problema de valor de contorno (8.137) pelo Esquema de Diferenças Finitas Centrada, para uma partição regular

$$\Pi : 0 = x_1 < \cdots < x_{n-1} < x_n = L \,,$$

com espaçamento h, *é dado por:*

$$\begin{bmatrix} d & -1 & & & & -1 \\ -1 & d & -1 & & & \\ & -1 & d & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & d & -1 \\ -1 & & & & -1 & d \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} h^2\, f(x_1) \\ \vdots \\ h^2\, f(x_i) \\ \vdots \\ h^2\, f(x_{n-1}) \end{bmatrix} \tag{8.138}$$

onde

$$d = (\, 2 + \sigma\, h^2 \,) \qquad e \qquad h = \frac{L}{(n-1)} \,.$$

Estamos denotando por u_i *uma aproximação para o valor* $u(x_i)$ *fornecida pelo Esquema de Diferenças Finitas. Note que* $u_n = u_1$ *devido a condição de contorno periódica.*

Determine uma solução numérica do sistema linear (8.138) pelo Método da Relaxação Sucessiva, com um resíduo relativo inferior a 10^{-5}, *para vários valores de* $\omega \in [1, 2)$.

Como exemplo, considere $L = 1$, $f(x) = \sin(\pi x)$ *e* $\sigma = 0.1$. *Para observar o desempenho dos métodos numéricos utilizar vários valores de* n.

Exercício 8.57 *Considere o sistema linear* $Ax = b$*, com a matriz* A *dada por:*

$$A = \begin{bmatrix} 1 & \rho \\ -\rho & 1 \end{bmatrix} \quad ; \quad \rho \in I\!R \,. \tag{8.139}$$

Qual a condição de convergência do Método de Gauss–Seidel ? Qual a escolha ótima para o parâmetro ω *do Método de Relaxação Sucessiva ?*

Exercício 8.58 *Sejam* $\alpha, \beta \in I\!R$ *e* A *uma matriz de ordem* n *. Prove que se* λ *é um autovalor de* A *, então* $\alpha\lambda + \beta$ *é um autovalor da matriz* $\alpha A + \beta I$*.*

Exercício 8.59 *Seja* $A \in I\!M_n(I\!R)$ *tridiagonal com os elementos da diagonal principal todos nulos (* $a_{ii} = 0$ *para* $i = 1, \cdots, n$ *) e os elementos das subdiagonais todos iguais a* 1 *, isto é,* $a_{i,(i+1)} = a_{(i+1),i} = 1$ *para* $i = 1, \cdots, (n-1)$ *. Mostre que, os autovalores da matriz* A *são dados por:*

$$\lambda_j = 2\cos\left(\frac{j\,\pi}{(n+1)}\right) \qquad para \qquad j = 1, \cdots, n \,,$$

e os autovetores associados são dados por:

$$v_j = \begin{bmatrix} \sin\left(\dfrac{j\,\pi}{(n+1)}\right) \\ \vdots \\ \sin\left(\dfrac{k\,j\,\pi}{(n+1)}\right) \\ \vdots \\ \sin\left(\dfrac{n\,j\,\pi}{(n+1)}\right) \end{bmatrix} .$$

Exercício 8.60 *Seja* $A = [a_{ij}]$ *uma matriz tridiagonal de ordem* n *com os elementos da diagonal principal todos iguais a* d *(* $a_{ii} = d$ *para* $i = 1, \cdots, n$ *) e os elementos das subdiagonais todos iguais a* e *, isto é,*

$$a_{i,(i+1)} = a_{(i+1),i} = e \qquad para \qquad i = 1, \cdots, (n-1) \,.$$

Prove que

$$\lambda_j = d + 2e\cos\left(\frac{j\,\pi}{(n+1)}\right) \qquad para \qquad j = 1, \cdots, n$$

são os autovalores da matriz A *.*

Exercício 8.61 *Considere a matriz*

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

com autovalores $\lambda_1 = 2 - \sqrt{2}$, $\lambda_2 = 2$ *e* $\lambda_3 = 2 + \sqrt{2}$. *Calcular a matriz de Iteração do Método de Jacobi e seus autovalores.*

Exercício 8.62 *Seja* $A = [a_{ij}]$ *uma matriz tridiagonal de ordem* n *com os elementos da diagonal principal todos iguais a* d ($a_{ii} = d$ *para* $i = 1, \cdots, n$), *os elementos da diagonal superior todos iguais a* e ($a_{i,(i+1)} = e$ *para* $i = 1, \cdots, (n-1)$) *e os elementos da diagonal inferior todos iguais a* f ($a_{(i+1),i} = f$ *para* $i = 1, \cdots, (n-1)$). *Mostre que, os autovalores da matriz* A *são dados por:*

$$\lambda_j = d + 2e\sqrt{\frac{f}{e}} \cos\left(\frac{j\,\pi}{(n+1)}\right) \qquad \textit{para} \qquad j = 1, \cdots, n$$

e os autovetores associados são dados por:

$$v_j = \begin{bmatrix} 2\left(\sqrt{\dfrac{f}{e}}\right) \sin\left(\dfrac{j\,\pi}{(n+1)}\right) \\ \vdots \\ 2\left(\sqrt{\dfrac{f}{e}}\right)^k \sin\left(\dfrac{k\,j\,\pi}{(n+1)}\right) \\ \vdots \\ 2\left(\sqrt{\dfrac{f}{e}}\right)^n \sin\left(\dfrac{n\,j\,\pi}{(n+1)}\right) \end{bmatrix}$$

Exercício 8.63 *Seja* $A \in I\!M_n(I\!R)$ *uma matriz tridiagonal e positiva-definida*

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \cdot & \\ & & \cdot & \cdot & -1 \\ & & & -1 & 2 \end{bmatrix}$$

Determinar $\mathcal{K}_2(A)$. *O que podemos concluir quando* n *cresce ?*

Exercício 8.64 *Considere a Equação de Difusão–Advecção*

$$-\varepsilon\, u''(x) \;+\; \beta\, u'(x) \;=\; f(x) \qquad ; \qquad x \in (0,1) \tag{8.140}$$

sujeita a condição de contorno

$$u(0) \;=\; 10 \qquad e \qquad u(1) \;=\; 10\,\exp(\beta) \tag{8.141}$$

Considere uma partição regular $\Pi \,:\, 0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = 1$, *com espaçamento* h. *O sistema linear proveniente da discretização do problema de valor de contorno (8.140)–(8.141) utilizando uma Fórmula de Diferenças Finitas Centrada para discretizar tanto o termo de difusão, quanto o termo de advecção, é dado por:*

$$\begin{bmatrix} d & b & & & & \\ c & d & b & & & \\ & c & d & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & d & b \\ & & & & c & d \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_n \end{bmatrix} \;=\; \begin{bmatrix} h^2\, f(x_1) + u_0 \\ \vdots \\ h^2\, f(x_i) \\ \vdots \\ h^2\, f(x_n) + u_{n+1} \end{bmatrix} \tag{8.142}$$

onde $d \;=\; 2\,\varepsilon\,,\;\; b \;=\; (\,P_e \,-\, \varepsilon\,)\,,\;\; c \;=\; -(\,P_e \,+\, \varepsilon\,)$ *e* $h \;=\; \dfrac{1}{(n+1)}$.

Definimos o Número de Péclet Local da seguinte forma:

$$P_e \;=\; \frac{|\,\beta\,|\,h}{2\varepsilon} \tag{8.143}$$

Estamos denotando por u_i *uma aproximação para o valor* $u(x_i)$ *fornecida pelo Esquema de Diferenças Finitas. Note que* $u_0 = u(0)$ *e* $u_{n+1} = u(1)$ *devido a imposição da condição de contorno. Sabemos que o Esquema de Diferenças Finitas Centrada é estável para* $P_e < 1$.

Sejam $A \in \mathbb{M}_n(\mathbb{R})$ *e* $b \in \mathbb{R}^n$ *a matriz e o vetor do lado direito do sistema linear (8.142), respectivamente. Definimos o seguinte processo iterativo*

$$M\,u^{(k+1)} \;=\; -N\,u^{(k)} \;+\; b\,, \tag{8.144}$$

onde M *e* N *são as partes simétrica e anti–simétrica de* A, *respectivamente, isto é,*

$$M \;=\; \frac{A \,+\, A^t}{2} \qquad e \qquad N \;=\; \frac{A \,-\, A^t}{2}\,.$$

Qual a taxa de convergência do processo iterativo (8.144) ?

Teorema 8.8.13 (Círculos de Gershgorin) *Seja* $A \in \mathbb{M}_n(\mathbb{C})$. *Então, todo autovalor de* A *está em pelo menos um dos círculos* $C_1, \cdots, C_n$ *, onde o* i*–ésimo círculo* C_i *tem centro em* $(a_{ii}, 0)$ *e raio* r_i *dado por:*

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad ; \quad i = 1, \ldots, n \tag{8.145}$$

Demonstração – Seja λ um autovalor de A e v o autovetor associado, isto é, $Av = \lambda v$. Seja i o índice da componente de v tal que

$$|v_i| = \max_{1 \leq j \leq n} \{ |v_j| \}$$

Examinando a i–ésima componente da equação vetorial $Av = \lambda v$, temos que

$$(\lambda - a_{ii}) v_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} v_j$$

tomando o módulo em ambos os lados da igualdade acima, temos

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \frac{|v_j|}{|v_i|} \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| = r_i$$

Desse modo, temos que o autovalor λ pertence ao i–ésimo círculo de Gershgorin de A, o que completa a demonstração. ■

Exercício 8.65 *Considere a matriz* A *estritamente diagonalmente dominante por linhas*

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 4 & 1 \\ 2 & 2 & 5 \end{bmatrix}$$

Mostre que o zero não está em nenhum círculo de Gershgorin de A, *isto é, a matriz* A *tem autovalores não–nulos. Portanto, concluímos que* A *é invertível.*

Exercício 8.66 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *estritamente diagonalmente dominante por linhas. Escreva a matriz de iteração de Jacobi* P *associada a matriz* A. *Mostre que os raios dos círculos de Gershgorin de* P *satisfazem* $r_i < 1$, *e concluindo que a matriz* P *é convergente.*

Exercício 8.67 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *estritamente diagonalmente dominante por linhas. Mostre que* A *é uma matriz invertível, utilizando os círculos de Gershgorin.*

Exercício 8.68 *Considere a matriz simétrica* A *dada por:*

$$A = \begin{bmatrix} 16 & 4 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 25 & -3 & 6 & 0 & 0 & 0 & 10 & 0 & 0 \\ 4 & -3 & 30 & 8 & 0 & 2 & 0 & -9 & 0 & 0 \\ 0 & 6 & 8 & 47 & 8 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 17 & 3 & 2 & -1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 3 & 15 & 1 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 21 & -4 & -7 & 0 \\ 0 & 10 & -9 & 0 & -1 & -3 & -4 & 52 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 3 & -7 & -1 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 24 \end{bmatrix}$$

Mostre que A *é uma matriz positiva–definida, utilizando os círculos de Gershgorin.*

8.9 Sistema Linear Sobredeterminado

Sejam $A \in I\!M_n(I\!R)$ uma matriz invertível e $b \in I\!R^n$ um elemento não–nulo. Vamos considerar o *Sistema Linear*: encontrar $x^* \in I\!R^n$ solução da equação

$$Ax = b. \tag{8.146}$$

Como a matriz A é invertível, temos que

$$x^* = A^{-1}b \tag{8.147}$$

é a única solução do sistema linear (8.146). Neste caso, temos que o elemento $x^* \in I\!R^n$ satisfaz exatamente todas as equações do sistema linear (8.146).

Utilizamos o Método de Decomposição LU (Eliminação Gaussiana) para encontrar uma solução numérica $\widehat{x}$. Note que durante o processo de Eliminação Gaussiana podemos identificar se a matriz A é singular.

Consideramos agora uma matriz $A \in I\!M_{m \times n}(I\!R)$, com $m > n$ e $posto(A) = n$. Desse modo, temos que o conjunto

$$\{ v_1 , \cdots , v_j , \cdots , v_n \} ,$$

onde $v_j \in I\!R^m$ é a j–ésima coluna da matriz A, é *linearmente independente* em $I\!R^m$.

Dado um elemento $b \in I\!R^m$, definimos o *Sistema Linear Sobredeterminado*

$$Ax = b. \tag{8.148}$$

Neste caso, temos duas situações possíveis. Na primeira situação, consideramos que o elemento b pertence ao subespaço gerado pelas colunas da matriz A. Assim, podemos encontrar um único elemento $x^* \in I\!R^n$ que satisfaz todas as equações do sistema linear sobredeterminado (8.148), isto é, o elemento $b \in I\!R^m$ pode ser escrito de modo único como uma combinação linear dos vetores colunas da matriz A. Dizemos que $x^* \in I\!R^n$ é uma solução clássica para o sistema linear (8.148).

Na segunda situação, consideramos que o elemento b não pertence ao subespaço gerado pelas colunas da matriz A. Neste caso, não temos um solução clássica para o sistema linear (8.148). Portanto, temos que definir o que seja uma solução *matematicamente satisfatória* para o sistema linear sobredeterminado (8.148).

Exemplo 8.9.1 *Mostre que o sistema linear sobredeterminado*

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 1 \end{bmatrix}$$

possui uma única solução, isto é, o elemento $b \in I\!R^3$ *é uma combinação linear das colunas da matriz do sistema linear.*

Exemplo 8.9.2 *Mostre que o sistema linear sobredeterminado*

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$$

não possui uma solução clássica, isto é, o elemento $b \in I\!R^3$ *não pode ser escrito como uma combinação linear das colunas da matriz do sistema linear.*

Antes de passar ao estudo de uma solução matematicamente satisfatória para o sistema linear sobredeterminado, vamos necessitar de alguns resultados que apresentamos a seguir.

Teorema 8.9.1 *Seja* $A \in I\!M_{m \times n}(I\!R)$, *com* $m > n$ *e* $posto(A) = n$. *Então,* $A^t A$ *é uma matriz positiva–definida.*

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 8.9.2 *Seja* $A \in I\!M_{m \times n}(I\!R)$, *com* $m > n$ *e* $posto(A) = r < n$. *Então,* $A^t A$ *é uma matriz semipositiva–definida.*

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 8.9.3 *Seja* $A \in I\!M_{m \times n}(I\!R)$, *com* $m < n$ *e* $posto(A) = m$. *Então,* $A A^t$ *é uma matriz positiva–definida.*

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 8.9.4 *Seja* $A \in I\!M_{m \times n}(I\!R)$, *com* $m < n$ *e* $posto(A) = m$. *Então,* $A^t A$ *é uma matriz semipositiva–definida.*

Demonstração – A prova pode ficar a cargo do leitor. □

Solução de Quadrados Mínimos

Definição 8.9.1 *Seja* $A \in I\!M_{m\times n}(I\!R)$*, com* $m > n$ *e posto*$(A) = n$*. Dado* $b \in I\!R^m$*, definimos o seguinte problema de minimização: encontrar um elemento* $x^* \in I\!R^n$ *de modo que*

$$\| Ax^* - b \|_2 \;=\; \min\{ \| Ax - b \|_2 \quad ; \quad x \in I\!R^n \} . \tag{8.149}$$

Dizemos que o elemento x^* *é uma* ***solução de quadrados mínimos*** *para o sistema linear* $Ax = b$ *e que o elemento* $z^* = Ax^*$ *é a* ***melhor aproximação*** *do elemento* b *no subespaço gerado pelas colunas da matriz* A *com relação à norma Euclidiana* $\| \cdot \|_2$ *.*

A seguir vamos apresentar um resultado de caracterização da solução de quadrados mínimos para um sistema linear sobredeterminado.

Teorema 8.9.5 *Sejam* $A \in I\!M_{m\times n}(I\!R)$*, com* $m > n$ *e posto*$(A) = n$*, e* $b \in I\!R^m$*. Definimos o funcional* $J : I\!R^n \longrightarrow I\!R$ *da seguinte forma:*

$$J(x) \;=\; \langle\, Ax - b \,,\, Ax - b \,\rangle \qquad ; \qquad x \in I\!R^n . \tag{8.150}$$

Então, o ***Problema de Minimização****: encontrar* $x^* \in I\!R^n$ *tal que*

$$J(x^*) \;=\; \min\{ J(x) \quad ; \quad x \in I\!R^n \} \tag{8.151}$$

é equivalente ao ***Sistema Normal***

$$A^t A x \;=\; A^t b . \tag{8.152}$$

Demonstração – Inicialmente vamos calcular a *Derivada Direcional* do funcional J no ponto $\overline{x}$ na direção do vetor $v \in I\!R^n$, que é definida da seguinte forma:

$$J'(\overline{x})(v) \;=\; \left\{ \frac{d}{dt} J(\overline{x} + tv) \right\}_{t=0} \tag{8.153}$$

Primeiramente, vamos calcular $J(x + tv)$ para $t \in R$

$$\begin{aligned} J(x + tv) \;&=\; \langle\, A(x + tv) - b \,,\, A(x + tv) - b \,\rangle \\ &=\; \langle\, Ax \,,\, Ax \,\rangle \;+\; 2\,t \,\langle\, Ax \,,\, Av \,\rangle \;+\; t^2 \,\langle\, Av \,,\, Av \,\rangle \\ &\quad -2 \,\langle\, Ax \,,\, b \,\rangle \;-\; 2\,t \,\langle\, Av \,,\, b \,\rangle \;+\; \langle\, b \,,\, b \,\rangle \end{aligned} \tag{8.154}$$

Podemos escrever $J(x+tv)$ da seguinte forma:

$$\begin{aligned} J(x+tv) \;&=\; \langle\, A^tAx\,,\, x\,\rangle \;+\; 2\,t\,\langle\, A^tAx\,,\, v\,\rangle \;+\; t^2\,\langle\, A^tAv\,,\, v\,\rangle \\ &\quad -2\,\langle\, Ax\,,\, b\,\rangle \;-\; 2\,t\,\langle\, A^tb\,,\, v\,\rangle \;+\; \langle\, b\,,\, b\,\rangle \end{aligned} \tag{8.155}$$

Derivando (8.155) com relação a t e fazendo $t = 0$, obtemos

$$J'(\overline{x})(v) \;=\; 2\,\langle\, A^tA\overline{x} \;-\; A^tb\,,\, v\,\rangle \tag{8.156}$$

que é a *Derivada Direcional* de J no ponto $\overline{x}$ na direção do vetor $v \in I\!R^n$, também denominada *Primeira Variação* do funcional J.

De (8.156) temos a definição do gradiente de J em um ponto $x \in I\!R^n$, como segue.

Definição 8.9.2 *O Gradiente do funcional J no ponto $x \in I\!R^n$ é definido por:*

$$\nabla J(x) \;=\; 2\,A^tA\,x \;-\; 2\,A^tb\,. \tag{8.157}$$

Desse modo, definimos o Ponto Crítico do funcional J, como segue.

Definição 8.9.3 *Dizemos que x^* é um Ponto Crítico do funcional J se, e somente se,*

$$J'(x^*)(v) \;=\; 0 \qquad \textit{para todo} \qquad v \in I\!R^n\,. \tag{8.158}$$

Desse modo, um ponto crítico do funcional J é uma solução do **sistema normal**

$$A^tAx \;=\; A^tb\,. \tag{8.159}$$

Como A tem posto completo, a matriz $C = A^tA$ é positiva–definida.

Portanto, o único ponto crítico do funcional J é caracterizado da seguinte forma:

$$x^* \;=\; (A^tA)^{-1}A^tb\,,$$

uma vez que A^tA é uma matriz invertível, veja Teorema 8.9.1.

Para classificar o ponto crítico x^* devemos calcular a *Segunda Variação* do funcional J no ponto $\overline{x}$ na direção do vetor $w \in I\!R^n$, que é definida da seguinte forma:

$$J''(\overline{x};v)(w) \;=\; \left\{ \frac{d}{dt} J'(\overline{x} + tw)(v) \right\}_{t=0} . \tag{8.160}$$

De (8.156), temos que

$$J'(\overline{x} + tw)(v) \;=\; 2\,\langle\, A^tA\overline{x}\,,\, v\,\rangle \;+\; 2\,t\langle\, A^tAw\,,\, v\,\rangle \;-\; 2\,\langle\, A^tb\,,\, v\,\rangle \tag{8.161}$$

para todo $t \in I\!R$.

Derivando (8.161) com relação a t e fazendo $t = 0$, obtemos

$$J''(\overline{x};v)(w) \;=\; 2\,\langle\, A^tAw\,,\, v\,\rangle \tag{8.162}$$

que é a *Segunda Variação* do funcional J no ponto $\overline{x}$ na direção do vetor $w \in I\!R^n$. De (8.162) temos a definição da matriz Hessiana do funcional J em um ponto $x \in I\!R^n$, como segue.

Definição 8.9.4 *A matriz Hessiana do funcional J no ponto $x \in I\!R^n$ é definida por:*

$$H(x) \;=\; 2\,A^tA\,. \tag{8.163}$$

Como A^tA é uma matriz positiva–definida, temos que $J''(x^*;v)(v) > 0$ para todo $v \in I\!R^n$ não–nulo. Desse modo, $x^* = (A^tA)^{-1}A^tb$ é um Ponto de Mínimo Global para o funcional J.

Assim, mostramos a equivalência entre o problema de minimização (8.151) e o sistema normal (8.152), que fornece uma caracterização da solução de quadrados mínimos para o sistema linear sobredeterminado $Ax = b$, o que completa a demonstração. ■

Definição 8.9.5 *Seja $J : I\!R^n \longrightarrow R$ um funcional. Dizemos que J é um **funcional quadrático**, se a matriz Hessiana não depende da variável x.*

Portanto, a solução de quadrados mínimos x^* para o sistema linear sobredeterminado $Ax = b$ é a solução do sistema normal

$$A^tAx \;=\; A^tb\,, \tag{8.164}$$

que pode ser obtida através da *Fatoração de Cholesky*, apresentada na seção 8.5.

Definição 8.9.6 *Seja* $A \in I\!M_{m \times n}(I\!R)$*, com* $m > n$ *e* $posto(A) = n$*. A matriz*

$$A^{\dagger} = (A^t A)^{-1} A^t$$

*é a **inversa a esquerda**, ou **pseudo–inversa**, da matriz* A*, que satisfaz as seguintes propriedades:*

1. $(AA^{\dagger})^t = AA^{\dagger}$.
2. $(A^{\dagger}A)^t = A^{\dagger}A$.
3. $AA^{\dagger}A = A$.
4. $A^{\dagger}AA^{\dagger} = A^{\dagger}$.

*A pseudo–inversa é denominada **inversa generalizada de Moore–Penrose**.*

Exemplo 8.9.3 *A partir da primeira variação do funcional* J *dada em (8.156), podemos calcular facilmente as derivadas parciais de primeira ordem do funcional* J*. De fato, sabemos que*

$$\frac{\partial J(x)}{\partial x_j} = J'(x)(e_j) = 2\langle A^t A x - A^t b \,,\, e_j \rangle \quad ; \quad j = 1, \cdots, n \tag{8.165}$$

onde e_j *é o* j*–ésimo elemento da base canônica do* $I\!R^n$*. Desse modo, temos que*

$$\frac{\partial J(x)}{\partial x_j} = 2\, v_j^t A x - 2\, v_j^t b \quad ; \quad j = 1, \cdots, n \tag{8.166}$$

onde v_j *é a* j*–ésima coluna da matriz* A*.*

Exemplo 8.9.4 *Considere o sistema linear sobredeterminado*

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}.$$

Determine uma solução de quadrados mínimos através da fatoração de Cholesky.

8.10 Subespaços Fundamentais de uma Matriz

Definição 8.10.1 *Seja* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$*. Definimos os seguintes subespaços:*

1. Espaço Coluna de A *é o subconjunto do* $\mathbb{R}^m$ *definido por:*

$$\mathcal{R}(A) = \{ z \in \mathbb{R}^m \ / \ z = Ax \quad ; \quad x \in \mathbb{R}^n \} .$$

Podemos observar que $\mathcal{R}(A)$ *é um subespaço de* $\mathbb{R}^m$ *gerado pelas coluna da matriz* A*. De fato, utilizando a notação* $A = [v_1 \cdots v_j \cdots v_n]$*, onde* $v_j \in \mathbb{R}^m$ *é a* j*–ésima coluna da matriz* A*, temos que, todo elemento* $z \in \mathcal{R}(A)$ *pode ser escrito da forma:*

$$z = Ax = \sum_{j=1}^{n} c_j \, v_j \qquad ; \qquad x = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n .$$

2. Espaço Nulo de A *é o subconjunto do* $\mathbb{R}^n$ *definido por:*

$$\mathcal{N}(A) = \{ x \in \mathbb{R}^n \ / \ Ax = 0_{\mathbb{R}^m} \} .$$

Podemos observar facilmente que $\mathcal{N}(A)$ *é um subespaço de* $\mathbb{R}^n$*, que é o conjunto solução do sistema linear homogêneo* $Ax = 0_{\mathbb{R}^m}$*.*

3. Espaço Coluna de A^t *é o subconjunto do* $\mathbb{R}^n$ *definido por:*

$$\mathcal{R}(A^t) = \{ x \in \mathbb{R}^n \ / \ x = A^t y \quad ; \quad y \in \mathbb{R}^m \} .$$

Podemos observar que $\mathcal{R}(A^t)$ *é um subespaço de* $\mathbb{R}^n$ *gerado pelas linhas da matriz* A*, que também pode ser denominado* ***espaço linha de*** A*.*

4. Espaço Nulo de A^t *é o subconjunto do* $\mathbb{R}^m$ *definido por:*

$$\mathcal{N}(A^t) = \{ z \in \mathbb{R}^m \ / \ A^t z = 0_{\mathbb{R}^n} \} .$$

Observamos que $\mathcal{N}(A^t)$ *é um subespaço de* $\mathbb{R}^m$*. Podemos fazer a seguinte observação. Se o elemento* $z \in \mathcal{N}(A^t)$ *temos que*

$$A^t z = 0_{\mathbb{R}^n} \qquad \Longleftrightarrow \qquad z^t A = 0^t_{\mathbb{R}^n} .$$

Assim, podemos denominar o espaço nulo de A^t *como o* ***espaço nulo esquerdo de*** A*.*

Proposição 8.10.1 *Sejam* $A, U \in \mathbb{M}_{m\times n}(\mathbb{R})$, *onde* U *é a matriz na forma escalonada linha equivalente a matriz* A. *Então,*

1. $\mathcal{R}(A^t) = \mathcal{R}(U^t)$.

2. *Uma base para o subespaço* $\mathcal{R}(A^t)$ *é formada pelas linhas não–nulas da matriz* U.

3. $dim(\,\mathcal{R}(A^t)\,) = posto(A)$, *que é igual ao número de linhas não–nulas de* U.

Demonstração – *1.* Sabemos que cada operação elementar de linhas realizada sobre a matriz A, substitui uma linha de A por um combinação linear de duas outras linhas de A, ou permutação de linhas. Desse modo, cada linha da matriz U é uma combinação linear das linhas da matriz A. Além disso, a matriz A pode ser obtida novamente da matriz U através das correspondentes operações inversas de linhas realizadas para obter U. Assim, as linhas da matriz A é também uma combinação linear das linha de U. Portanto, podemos concluir que as linhas de A e as linhas de U geram o mesmo subespaço, isto é, $\mathcal{R}(A^t) = \mathcal{R}(U^t)$.

2. Podemos observar facilmente que as linhas não–nulas da matriz U são linearmente independentes. Caso contrário, poderíamos obter uma outra linha nula através de uma operação elementar de linhas. Entretanto, isso seria um contradição com a hipótese de que a matriz U está na forma escalonada. Portanto, uma base para o subespaço $\mathcal{R}(U^t) = \mathcal{R}(A^t)$ é formada pelas linhas não–nulas de U.

3. Sabemos que a dimensão do subespaço $\mathcal{R}(U^t) = \mathcal{R}(A^t)$ é igual ao número de linhas não–nulas da matriz U. Portanto, temos que $dim(\,\mathcal{R}(A^t)\,) = posto(A)$, o que completa a demonstração. ■

Exemplo 8.10.1 *Considere a matriz* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 3 \\ 4 & -1 & 3 \\ 1 & 1 & 7 \end{bmatrix}.$$

Determine uma base para o subespaço $\mathcal{R}(A^t)$.

Uma matriz reduzida U, linha equivalente a matriz A, é dada por:

$$U = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Desse modo, temos que $dim(\,\mathcal{R}(A^t)\,) = posto(A) = 2$, que é igual ao número de linhas não–nulas da matriz U.

Assim, temos que os elementos

$$u = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \quad \text{e} \quad v = \begin{bmatrix} 0 \\ 1 \\ 5 \end{bmatrix},$$

que são as linhas não–nulas da matriz U, formam uma base para o subespaço $\mathcal{R}(A^t)$.

Proposição 8.10.2 *Sejam* $A, U \in \mathbb{M}_{m \times n}(\mathbb{R})$, *onde* U *é a matriz na forma escalonada linha equivalente a matriz* A. *Então,*

1. $\mathcal{N}(A) = \mathcal{N}(U)$.

2. *Uma base para o subespaço* $\mathcal{N}(A)$ *pode ser obtida através da solução geral do sistema linear reduzido* $Ux = 0_{\mathbb{R}^m}$.

3. $dim(\mathcal{N}(A)) = n - posto(A)$, *que é igual ao número de variáveis livres do sistema linear reduzido* $Ux = 0_{\mathbb{R}^m}$.

Demonstração – *1.* Como A e U são equivalentes por linha, pelo Teorema 2.9.4, sabemos que os sistemas lineares homogêneos $Ax = 0_{\mathbb{R}^m}$ e $Ux = 0_{\mathbb{R}^m}$ possuem o mesmo conjunto solução. Portanto, podemos concluir que $\mathcal{N}(A) = \mathcal{N}(U)$.

2. Determinamos uma base para o subespaço $\mathcal{N}(A)$ da seguinte forma: para cada variável livre o sistema linear homogêneo $Ux = 0_{\mathbb{R}^m}$, temos associada uma solução básica, que é obtida atribuindo o valor 1 a essa variável, e o valor zero para as variáveis livres restantes, e em seguida calculando os valores das varáveis básica. Desse modo, construímos uma base para o subespaço $\mathcal{N}(A)$ formada pelas soluções básicas.

3. Temos que uma base para o subespaço $\mathcal{N}(A)$ é formada pelas soluções básicas, isto é, pelos elementos que aparecem na expressão da solução geral do sistema linear homogêneo $Ax = 0_{\mathbb{R}^m}$. Portanto, temos que $dim(\mathcal{N}(A)) = n - posto(A)$, pois o $posto(A)$ é igual ao número de linhas não–nulas da matriz U, o que completa a demonstração. ■

Exemplo 8.10.2 *Considere a matriz* $A \in \mathbb{M}_{3\times5}(\mathbb{R})$

$$A = \begin{bmatrix} 1 & 1 & 3 & 2 & 0 \\ 2 & 2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -2 & 1 \end{bmatrix}.$$

Determine uma base para o subespaço $\mathcal{N}(A)$.

Uma matriz reduzida U, linha equivalente a matriz A, é dada por:

$$U = \begin{bmatrix} 1 & 1 & 3 & 2 & 0 \\ 0 & 0 & 5 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Assim, temos que $dim(\mathcal{N}(A)) = n - posto(A) = 3$, onde $n = 5$ e $posto(A) = 2$, que é igual ao número de linhas não–nulas da matriz U.

O sistema reduzido equivalente ao sistema linear homogêneo $Ax = 0_{\mathbb{R}^m}$ é dado por:

$$\begin{cases} x_1 + x_2 + 3x_3 + 2x_4 = 0 \\ 5x_3 + 4x_4 - x_5 = 0 \end{cases}$$

Escolhendo as variáveis básica x_1 e x_5 e as variáveis livres x_2 , x_3 e x_4, a solução geral do sistema linear homogêneo é escrito da seguinte forma:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 5 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 4 \end{bmatrix} \quad ; \quad x_2, x_3, x_4 \in \mathbb{R}.$$

Desse modo, encontramos uma base para o subespaço $\mathcal{N}(A)$, formada pelos elementos

$$u_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad , \quad u_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 5 \end{bmatrix} \quad \text{e} \quad u_3 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 4 \end{bmatrix}.$$

Proposição 8.10.3 *Sejam* $A, U \in \mathbb{M}_{m\times n}(\mathbb{R})$, *onde* U *é a matriz na forma escalonada linha equivalente a matriz* A. *Então,*

1. *Sejam* $c_1, \cdots, c_n$ *escalares reais, então*

$$c_1u_1 + c_2u_2 + \cdots + c_nu_n = 0_{\mathbb{R}^m} \iff c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0_{\mathbb{R}^m},$$

   onde $u_j, v_j \in \mathbb{R}^m$ *são a* j*–ésima coluna da matriz* U *e a* j*–ésima coluna da matriz* A, *respectivamente.*

2. *Para cada conjunto de colunas da matriz* U *que seja linearmente independente, as correspondentes colunas da matriz* A *também são linearmente independentes.*

Demonstração – *1.* A validade da equivalência vem do fato que os sistemas lineares homogêneos $Ax = 0_{\mathbb{R}^m}$ e $Ux = 0_{\mathbb{R}^m}$ possuem o mesmo conjunto solução.

2. Sejam $w_1, \cdots, w_k$ um conjunto de colunas de U linearmente independentes, e $a_1, \cdots, a_k$ as correspondentes colunas de A. Se $c_1, \cdots, c_n$ escalares reais tais que

$$c_1a_1 + c_2a_2 + \cdots + c_na_k = 0_{\mathbb{R}^m},$$

então pelo resultado do item *1.* temos que

$$c_1w_1 + c_2w_2 + \cdots + c_nw_k = 0_{\mathbb{R}^m}.$$

Como $w_1, \cdots, w_k$ são linearmente independentes, isso implica em $c_1 = \cdots = c_k = 0$. Portanto, $a_1, \cdots, a_k$ são linearmente independentes, o que completa a demonstração. ■

Proposição 8.10.4 *Sejam* $A, U \in \mathbb{M}_{m\times n}(\mathbb{R})$, *onde* U *é a matriz na forma escalonada linha equivalente a matriz* A. *Então,*

1. *Uma base para o subespaço* $\mathcal{R}(A)$ *é formada pelas colunas da matriz* A *correspondentes à aquelas colunas da matriz* U *que contem pivo.*
2. $dim(\mathcal{R}(A)) = posto(A)$.

Demonstração – *1.* O resultado segue do item *2.* da Proposição 8.10.3, e do fato que as colunas da matriz U que contem pivo são linearmente independentes, considerando que a matriz U está na forma escalonada.

2. Temos que $dim(\mathcal{R}(A))$ é igual ao número de colunas da matriz U que contem pivo, que por sua vez é igual ao $posto(A)$, o que completa a demonstração. ■

Desse modo, dos resultados das Proposições 8.10.4 e 8.10.1, segue que para qualquer matriz $A \in \mathbb{M}_n(\mathbb{R})$ tem–se que

$$dim(\mathcal{R}(A)) = dim(\mathcal{R}(A^t)) .$$

Portanto, podemos concluir que

(a) $posto(A) = posto(A^t)$.

(b) O número máximo de colunas da matriz A linearmente independente, é igual ao número máximo de linhas da matriz A linearmente independente.

Exemplo 8.10.3 *Considere a matriz* $A \in \mathbb{M}_{3\times 4}(\mathbb{R})$

$$A = \begin{bmatrix} 1 & 3 & -1 & 0 \\ 2 & -1 & 0 & 4 \\ 1 & 2 & -1 & 0 \end{bmatrix} .$$

Determine uma base para o subespaço $\mathcal{R}(A)$.

Uma matriz reduzida U, linha equivalente a matriz A, é dada por:

$$U = \begin{bmatrix} 1 & 3 & -1 & 0 \\ 0 & -7 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix}.$$

Desse modo, temos que $dim(\mathcal{R}(A)) = posto(A) = 3$, que é igual ao número de linhas não–nulas da matriz U.

Podemos observar que a primeira, a segunda e a terceira coluna da matriz U são as que contem pivo. Assim, temos que os elementos

$$u_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad , \quad u_2 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} \quad \text{e} \quad u_3 = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix},$$

que correspondem a primeira, a segunda e a terceira coluna da matriz A, formam uma base para o subespaço $\mathcal{R}(A)$.

Teorema 8.10.1 *Seja $A \in I\!M_{m \times n}(I\!R)$. Então,*

1. *O subespaço $\mathcal{R}(A^t)$ e o subespaço $\mathcal{N}(A)$ são subespaços ortogonais em $I\!R^n$.*
2. *O subespaço $\mathcal{R}(A)$ e o subespaço $\mathcal{N}(A^t)$ são subespaços ortogonais em $I\!R^m$.*

Demonstração – *1.* Seja $x \in \mathcal{N}(A)$, isto é, $Ax = 0$. Para todo $y \in I\!R^m$, tem–se

$$0 = \langle Ax \, , \, y \rangle = \langle x \, , \, A^t y \rangle$$

Como o elemento $z = A^t y \in \mathcal{R}(A^t)$, temos que todo elemento do subespaço $\mathcal{N}(A)$ é ortogonal aos elementos do subespaço $\mathcal{R}(A^t)$.

2. Seja $y \in \mathcal{N}(A^t)$, isto é, $A^t y = 0$. Para todo $x \in I\!R^n$, temos que

$$0 = \langle A^t y \, , \, x \rangle = \langle y \, , \, Ax \rangle$$

Como o elemento $z = Ax \in \mathcal{R}(A)$, temos que todo elemento do subespaço $\mathcal{N}(A^t)$ é ortogonal aos elementos do subespaço $\mathcal{R}(A)$, o que completa a demonstração. ∎

Teorema 8.10.2 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,*

1. $\mathcal{R}(A^t)^\perp = \mathcal{N}(A)$.

2. $\mathcal{N}(A)^\perp = \mathcal{R}(A^t)$.

3. $\mathcal{R}(A)^\perp = \mathcal{N}(A^t)$.

4. $\mathcal{N}(A^t)^\perp = \mathcal{R}(A)$.

Demonstração

1. Pelo Teorema 8.10.1, temos que o subespaço $\mathcal{N}(A)$ é ortogonal ao subespaço $\mathcal{R}(A^t)$. Logo, obtemos que $\mathcal{N}(A) \subset \mathcal{R}(A^t)^\perp$.

Vamos provar agora que $\mathcal{R}(A^t)^\perp \subset \mathcal{N}(A)$. Seja $x \in \mathcal{R}(A^t)^\perp$, isto é, o elemento x é ortogonal a todo elemento $z = A^t y$ para $y \in \mathbb{R}^m$. Desse modo, temos que

$$0 = \langle x , A^t y \rangle = \langle Ax , y \rangle \quad ; \quad \forall \ y \in \mathbb{R}^m$$

Logo, $Ax = 0$ o que implica em $x \in \mathcal{N}(A)$.

2. Sabemos que $(\mathcal{R}(A^t)^\perp)^\perp = \mathcal{R}(A^t)$. Logo, $\mathcal{N}(A)^\perp = \mathcal{R}(A^t)$.

3. Pelo Teorema 8.10.1, temos que o subespaço $\mathcal{N}(A^t)$ é ortogonal ao subespaço $\mathcal{R}(A)$. Logo, obtemos que $\mathcal{N}(A^t) \subset \mathcal{R}(A)^\perp$.

Vamos provar agora que $\mathcal{R}(A)^\perp \subset \mathcal{N}(A^t)$. Seja $y \in \mathcal{R}(A)^\perp$, isto é, o elemento y é ortogonal a todo elemento $z = Ax$ para $x \in \mathbb{R}^n$. Desse modo, temos que

$$0 = \langle y , Ax \rangle = \langle A^t y , x \rangle \quad ; \quad \forall \ x \in \mathbb{R}^n$$

Logo, $A^t y = 0$ o que implica em $y \in \mathcal{N}(A^t)$.

4. Sabemos que $(\mathcal{R}(A)^\perp)^\perp = \mathcal{R}(A)$. Logo, $\mathcal{N}(A^t)^\perp = \mathcal{R}(A)$, o que completa a demonstração. ■

Observações:

1. Podemos fazer a seguinte interpretação do item *1.* do Teorema 8.10.2: o elemento $z \in \mathbb{R}^n$ é ortogonal ao subespaço $\mathcal{R}(A^t)$ se, e somente se, for solução do sistema homogêneo $Ax = 0$.

2. Podemos fazer a seguinte interpretação do item *1.* do Teorema 8.10.2: para $b \in \mathbb{R}^m$, o sistema linear $Ax = b$ possui solução se, e somente se, o elemento b for ortogonal a toda solução do sistema homogêneo $A^t y = 0$.

3. Se conhecemos um conjunto gerador para um subespaço S do $\mathbb{R}^n$ e queremos encontrar $S^\perp$, basta construir uma matriz A cujas linhas são os elementos do conjunto gerador. Assim, temos que $S = \mathcal{R}(A^t)$ e $S^\perp = \mathcal{N}(A)$.

4. Conhecendo um subespaço S mediante um conjunto de equações lineares, podemos encontrar uma matriz A tal que $S = \mathcal{N}(A)$, e assim $S^\perp = \mathcal{R}(A^t)$.

Exemplo 8.10.4 *Seja S um subespaço do $\mathbb{R}^4$ gerado pelo seguinte conjunto*

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}.$$

Vamos construir uma matriz A da forma:

$$A = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 1 & -1 & 0 \end{bmatrix}.$$

Assim, temos que o subespaço $S = \mathcal{R}(A^t)$ e o subespaço $S^\perp = \mathcal{N}(A)$.

Uma matriz reduzida U, linha equivalente a matriz A, é dada por:

$$U = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Desse modo, encontramos uma base para o subespaço $\mathcal{R}(A^t)$ dada por:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} \right\}.$$

Encontramos uma base para o subespaço $S^\perp = \mathcal{N}(A)$ *resolvendo o sistema reduzido* $Ux = 0$, *equivalente ao sistema homogêneo* $Ax = 0$, *cuja solução geral é dada por:*

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ -3 \\ 0 \\ 1 \end{bmatrix} \quad ; \quad x_3, x_4 \in I\!\!R\,.$$

Dessa forma, encontrando uma base para o subespaço $S^\perp = \mathcal{N}(A)$.

Exemplo 8.10.5 *Considere o subespaço* S *do* $I\!\!R^4$ *definido da seguinte forma:*

$$S = \{ (x, y, z, w) \in I\!\!R^4 \ / \ x + z = 0 \quad e \quad y + w = 0 \}\,.$$

Encontrar uma base para o subespaço S *e uma base para o subespaço* $S^\perp$.

Temos que, $(x, y, z, w) \in S$ *se, e somente se, suas componentes satisfazem o sistema homogêneo*

$$\begin{cases} x + z = 0 \\ y + w = 0 \end{cases}$$

O subespaço $S = \mathcal{N}(A)$ *e o subespaço* $S^\perp = \mathcal{R}(A^t)$, *onde a matriz* A *é dada por:*

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Desse modo, encontramos uma base para o subespaço $S^\perp = \mathcal{R}(A^t)$ *dada por:*

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} , \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Encontramos uma base para o subespaço $S = \mathcal{N}(A)$ *resolvendo o sistema homogêneo* $Ax = 0$, *cuja solução geral é dada por:*

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = z \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \quad ; \quad z, w \in I\!\!R\,.$$

Desse forma, encontramos uma base para o subespaço $S = \mathcal{N}(A)$.

Neste momento é muito importante recordar o **Teorema da Decomposição Ortogonal**, Teorema 5.9.1. Sejam V um espaço vetorial munido do produto interno $\langle \cdot , \cdot \rangle$ e S um subespaço de dimensão finita de V. Então, $V = S \oplus S^{\perp}$, isto é, todo elemento $u \in V$ pode ser escrito de modo único da seguinte forma:

$$u = v + w \qquad \text{com} \qquad v \in S \quad \text{e} \quad w \in S^{\perp} \, .$$

Além disso, a norma do elemento u é dada pela *Fórmula de Pitágoras*

$$\| u \|_2^2 = \| v \|_2^2 + \| w \|_2^2 \, .$$

Desse resultado, temos a definição de projeção ortogonal sobre os subespaços S e $S^{\perp}$.

Definição 8.10.2 *Sejam V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$, S um subespaço de dimensão finita de V e $u \in V$. Se $v \in S$ e $w \in S^{\perp}$ são os únicos elementos de V tais que $u = v + w$, dizemos que*

1. *$v \in S$ é a* ***projeção ortogonal*** *do elemento u sobre o subespaço S.*
2. *$w \in S^{\perp}$ é a* ***projeção ortogonal*** *do elemento u sobre o subespaço $S^{\perp}$.*

Teorema 8.10.3 *Seja $A \in I\!M_{m\times n}(I\!R)$. Então,*

(a) $dim(\mathcal{R}(A^t)) + dim(\mathcal{N}(A)) = n$.

(b) $dim(\mathcal{R}(A)) + dim(\mathcal{N}(A^t)) = m$.

Demonstração – *(a)* Pelo Teorema 8.10.2, tem–se $\mathcal{R}(A^t)^{\perp} = \mathcal{N}(A)$, e do Teorema 5.9.1, que é o Teorema da Decomposição Ortogonal, temos que $I\!R^n = \mathcal{R}(A^t) \oplus \mathcal{R}(A^t)^{\perp}$.

Portanto, obtemos

$$dim(\mathcal{R}(A^t)) + dim(\mathcal{R}(A^t)^{\perp}) = n \qquad \Longrightarrow \qquad dim(\mathcal{R}(A^t)) + dim(\mathcal{N}(A)) = n \, ,$$

o que completa a prova do item *(a)*.

(b) Pelo Teorema 8.10.2, sabemos que $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^t)$, e do Teorema 5.9.1, que é o Teorema da Decomposição Ortogonal, temos que $I\!R^m = \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$. Portanto, obtemos

$$dim(\mathcal{R}(A)) + dim(\mathcal{R}(A)^{\perp}) = m \qquad \Longrightarrow \qquad dim(\mathcal{R}(A)) + dim(\mathcal{N}(A^t)) = m \, ,$$

o que completa a prova do item *(b)*. ■

Utilizando as conexões geométricas entre os subespaços fundamentais e o Teorema do Núcleo e da Imagem, apresentamos uma outra demonstração para o fato que a dimensão do subespaço coluna de A e a dimensão do espaço linha de A são iguais. Sendo assim, provamos novamente que o posto das matrizes A e A^t são iguais.

Teorema 8.10.4 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,* $dim(\,\mathcal{R}(A)\,) \,=\, dim(\,\mathcal{R}(A^t\,)$.

Demonstração – Pelos resultados do Teorema 8.10.3, temos que

$$dim(\,\mathcal{R}(A)\,) \,=\, m \,-\, dim\mathcal{N}(A^t)) \qquad e \qquad dim(\,\mathcal{R}(A^t\,) \,=\, n \,-\, dim(\mathcal{N}(A))\,.$$

Agora, consideramos a transformação linear $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ associada a matriz A definida por:

$$T_A(x) \,=\, Ax\,,$$

lembrando que os espaços vetoriais $\mathbb{R}^n$ e $\mathbb{M}_{n\times 1}(\mathbb{R})$ são isomorfos, assim como os espaços vetoriais $\mathbb{R}^m$ e $\mathbb{M}_{m\times 1}(\mathbb{R})$.

Podemos verificar facilmente que $Im(T_A) = \mathcal{R}(A)$ e $Ker(T_A) = \mathcal{N}(A)$. Desse modo, pelo Teorema 4.4.2, que é o Teorema do Núcleo e da Imagem, sabemos que

$$dim(\,Im(T_A)\,) \,+\, dim(\,Ker(T_A) \,=\, dim(\,\mathbb{R}^n\,) \,=\, n\,.$$

Desse modo, temos que

$$dim(\mathcal{R}(A)) \,+\, dim(\mathcal{N}(A)) \,=\, n \qquad \Longrightarrow \qquad dim(\mathcal{R}(A)) \,=\, n \,-\, dim(\mathcal{N}(A))\,.$$

Como $dim(\,\mathcal{R}(A^t\,) \,=\, n \,-\, dim(\mathcal{N}(A))$, provamos que

$$dim(\mathcal{R}(A)) \,=\, dim(\,\mathcal{R}(A^t\,)\,,$$

o que completa a demonstração. ■

No Teorema 8.10.4 é importante observar que o subespaço $\mathcal{R}(A) \subset \mathbb{R}^m$ e que o subespaço $\mathcal{R}(A^t) \subset \mathbb{R}^n$, isto é, eles estão contidos em espaços vetoriais diferentes, entretanto, possuem a mesma dimensão.

Teorema 8.10.5 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,* $\mathcal{N}(A^tA) = \mathcal{N}(A)$.

Demonstração – É imediato que $\mathcal{N}(A) \subset \mathcal{N}(A^tA)$. De fato, considere um elemento $x \in \mathcal{N}(A)$, isto é, $Ax = 0_{\mathbb{R}^m}$. Logo, temos que

$$A^tAx = A^t0_{\mathbb{R}^m} = 0_{\mathbb{R}^n} \qquad \Longrightarrow \qquad x \in \mathcal{N}(A^tA)\,.$$

Assim, provamos que $\mathcal{N}(A) \subset \mathcal{N}(A^tA)$.

Considere agora um elemento $x \in \mathcal{N}(A^tA)$, isto é, $A^tAx = 0_{\mathbb{R}^n}$. Desse modo, fazendo

$$x^tA^tAx = 0_{\mathbb{R}} \quad \Longleftrightarrow \quad (Ax)^t(Ax) = \|\,Ax\,\|_2^2 = 0_{\mathbb{R}} \quad \Longleftrightarrow \quad Ax = 0_{\mathbb{R}^m}$$

obtemos $x \in \mathcal{N}(A)$. Assim, provamos que $\mathcal{N}(A^tA) \subset \mathcal{N}(A)$. Portanto, mostramos que $\mathcal{N}(A^tA) = \mathcal{N}(A)$, o que completa a demonstração. ■

Teorema 8.10.6 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,* $\mathcal{N}(AA^t) = \mathcal{N}(A^t)$.

Demonstração – A prova pode ficar a cargo do leitor. □

Corolário 8.10.1 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,*

$$posto(A^tA) = posto(A) = posto(A^t) = posto(AA^t)\,.$$

Demonstração – Fazendo uso dos resultados anteriores, obtemos

$$\begin{aligned} posto(A^tA) &= n - dim(\mathcal{N}(A^tA)) = n - dim(\mathcal{N}(A)) \\ &= posto(A) = posto(A^t) \\ &= m - dim(\mathcal{N}(A^t)) = m - dim(\mathcal{N}(AA^t)) \\ &= posto(AA^t) \end{aligned}$$

o que completa a demonstração. ■

Exercícios

Exercício 8.69 *Considere a matriz* $A \in I\!M_{4\times3}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 1 \\ 1 & 0 & 1 \\ -2 & 3 & -1 \end{bmatrix}.$$

Determine uma base para cada um dos subespaços fundamentais da matriz A.

Exercício 8.70 *Considere a matriz* $A \in I\!M_{3\times4}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ -1 & 1 & 2 & 1 \end{bmatrix}.$$

Determine uma base para cada um dos subespaços fundamentais da matriz A.

Exercício 8.71 *Considere a matriz* $A \in I\!M_{3\times5}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 & 1 \\ 1 & 2 & 4 & 0 & -1 \\ -1 & 1 & 5 & 1 & -2 \end{bmatrix}.$$

Determine uma base para cada um dos subespaços fundamentais da matriz A.

Exercício 8.72 *Sejam* $A \in I\!M_{m\times n}(I\!R)$ *e* $b \in I\!R^m$. *Mostre que apenas um dos sistemas lineares abaixo possui solução*

$$Ax = b$$

$$A^t y = 0 \quad \text{com} \quad \langle\, b \,,\, y \,\rangle \neq 0$$

Faça uma interpretação geométrica.

Exercício 8.73 *Sejam* $A \in I\!M_{m\times n}(I\!R)$, $x \in I\!R^n$ *e* $b, y \in I\!R^m$. *Mostre que se* $Ax = b$ *e* $y^t A = 0$, *então* $\langle\, b \,,\, y \,\rangle = 0$. *Faça uma interpretação geométrica.*

Exercício 8.74 *Sejam* $A \in I\!M_{m\times n}(I\!R)$, $y \in I\!R^m$ *e* $x, b \in I\!R^n$. *Mostre que se* $Ax = 0$ *e* $A^t y = b$, *então* $\langle\, b \,,\, x \,\rangle = 0$. *Faça uma interpretação geométrica.*

Exercício 8.75 *Dada a matriz* $A \in I\!M_{3\times4}(I\!R)$

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ -1 & 1 & 2 & 1 \end{bmatrix}.$$

Determine uma base para os subespaços $\mathcal{R}(A^t)$ *e* $\mathcal{N}(A)$.

Exercício 8.76 *Dada a matriz* $A \in I\!M_{4\times3}(I\!R)$

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \\ 3 & 2 & 8 \end{bmatrix}.$$

Determine uma base para os subespaços $\mathcal{R}(A)$ *e* $\mathcal{N}(A^t)$.

Exercício 8.77 *Sejam a matriz* $A \in I\!M_{3\times5}(I\!R)$ *e o elemento* $b \in I\!R^5$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 2 & -1 & 0 & 1 \end{bmatrix} \quad e \quad b = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 1 \\ 9 \end{bmatrix}.$$

Verifique se o elemento b *pertence ao subespaço* $\mathcal{R}(A^t)$.

Exercício 8.78 *Seja* $A \in I\!M_{m\times n}(I\!R)$. *Mostre que* $\mathcal{R}(AA^t) = \mathcal{R}(A)$.

Exercício 8.79 *Seja* $A \in I\!M_{m\times n}(I\!R)$. *Mostre que* $\mathcal{R}(A^tA) = \mathcal{R}(A^t)$.

Exercício 8.80 *Sejam* $X \in I\!R^m$, $Y \in I\!R^n$ *e* $A \in I\!M_{m\times n}(I\!R)$ *dada por* $A = XY^t$. *Mostre que* $posto(A) = posto(A^t) = 1$.

Exercício 8.81 *Considere a matriz* $A \in I\!M_{3\times4}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 3 & 0 & -2 & 1 \\ -6 & 0 & 4 & -2 \\ 6 & 0 & -4 & 2 \end{bmatrix}.$$

Escreva a matriz $A = XY^t$, *com* $X \in I\!R^3$ *e* $Y \in I\!R^4$. *Determine uma base para o subespaço* $\mathcal{N}(A)$ *e uma base para o subespaço* $\mathcal{R}(A^t)$.

Exercício 8.82 *Sejam* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *e* $b \in \mathbb{R}^m$. *Mostre que o sistema linear* $Ax = b$ *possui solução se, e somente se,* $posto(A) = posto(M)$, *onde* $M = [A \,|\, b]$ *é a matriz ampliada do sistema linear.*

Exercício 8.83 *Seja* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$. *Considere que a única solução do sistema linear homogêneo* $Ax = 0_{\mathbb{R}^m}$ *é a solução trivial* $x = 0_{\mathbb{R}^n}$. *Qual é o posto da matriz* A?

Exercício 8.84 *Seja uma matriz* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *tal que o sistema linear homogêneo* $Ax = 0_{\mathbb{R}^m}$ *possui solução não trivial. Mostre que o sistema linear* $A^t y = b$ *não possui solução para alguns elementos* $b \in \mathbb{R}^n$. *Dê um exemplo considerando uma matriz* $A \in \mathbb{M}_{4 \times 3}(\mathbb{R})$ *e um elemento* $b \in \mathbb{R}^3$.

Exercício 8.85 *Seja* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, *com* $m > n$. *Considere que a única solução do sistema linear homogêneo* $Ax = 0_{\mathbb{R}^m}$ *é a solução trivial* $x = 0_{\mathbb{R}^n}$. *Mostre que o sistema linear* $A^t y = b$ *possui solução para todo* $b \in \mathbb{R}^n$.

Exercício 8.86 *Existe uma matriz* A *tal que* $\mathcal{R}(A) = \mathbb{R}^4$ *e* $\mathcal{R}(A^t) = \mathbb{R}^3$?

Exercício 8.87 *Considere o espaço vetorial* $\mathbb{R}^3$ *munido do produto interno usual. Seja* W *o subespaço de* $\mathbb{R}^3$ *gerado pelos elementos*

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad , \quad w_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \qquad e \qquad w_3 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} .$$

Determine uma matriz A *e uma matriz* B *tais que*

$$\mathcal{R}(A^t) = W \qquad e \qquad \mathcal{N}(B) = W .$$

Inicialmente, faça a representação geométrica dos subespaços fundamentais de cada uma das matrizes A *e* B *que estão contidos em* $\mathbb{R}^3$.

Exercício 8.88 *Considere o espaço vetorial* $\mathbb{R}^4$ *munido do produto interno usual. Seja* W *o subespaço de* $\mathbb{R}^4$ *gerado pelos elementos*

$$w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad w_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \qquad e \qquad w_3 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} .$$

Determine uma matriz A *e uma matriz* B *tais que*

$$\mathcal{R}(A^t) = W \qquad e \qquad \mathcal{N}(B) = W .$$

Exercício 8.89 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$*,* $b \in \mathbb{R}^m$ *e* $\overline{x} \in \mathbb{R}^n$ *uma solução do sistema linear* $Ax = b$*, isto é,* $A\overline{x} = b$*. Pede–se:*

(a) Mostre que qualquer outra solução $\widehat{x}$ *do sistema linear* $Ax = b$ *pode ser escrita como* $\widehat{x} = \overline{x} + z$*, com* $z \in \mathcal{N}(A)$*.*

(b) Mostre que qualquer elemento $\widehat{x} = \overline{x} + z$*, com* $z \in \mathcal{N}(A)$*, é também uma solução do sistema linear* $Ax = b$*.*

(c) Mostre que o sistema linear $Ax = b$ *possui uma única solução se, e somente se,* $\mathcal{N}(A) = \{ 0_{\mathbb{R}^n} \}$*.*

Exercício 8.90 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e* $b \in \mathbb{R}^m$*. Tomando* $m = 3$ *e* $n = 2$*, dê exemplo de um sistema linear* $Ax = b$ *para cada uma das seguintes situações:*

(a) O sistema linear possui infinitas soluções.

(b) O sistema linear possui uma única solução.

(c) O sistema linear não possui solução.

Faça uma interpretação geométrica em cada um dos casos.

8.11 Projeções Ortogonais

Teorema 8.11.1 *Sejam* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *e* $b \in \mathbb{R}^m$. *A* ***projeção ortogonal*** *do elemento* b *sobre o subespaço* $\mathcal{R}(A)$, *é o único elemento* $z^* = Ax^* \in \mathcal{R}(A)$ *cuja distância ao elemento* b *é a menor possível com relação à norma Euclidiana* $\| \cdot \|_2$, *isto é,*

$$\| b - z^* \|_2 \leq \| b - z \|_2 \quad ; \quad \forall \; z \in \mathcal{R}(A) \, .$$

Demonstração – Pelo Teorema da Decomposição Ortogonal, sabemos que o elemento $b \in \mathbb{R}^m$ pode ser escrito de modo único na forma:

$$b = z^* + w^* \, ,$$

onde $z^* \in \mathcal{R}(A)$ e $w^* \in \mathcal{R}(A)^{\perp} = \mathcal{N}(A^t)$ são as projeções ortogonais do elemento b nos subespaços $\mathcal{R}(A)$ e $\mathcal{N}(A^t)$, respectivamente.

Desse modo, temos que

$$b - z = (b - z^*) + (z^* - z) \quad \text{para todo} \quad z \in \mathcal{R}(A) \, ,$$

uma vez que

$$(z^* - z) \in \mathcal{R}(A) \quad \text{e} \quad (b - z^*) \in \mathcal{R}(A)^{\perp} = \mathcal{N}(A^t) \, ,$$

que é a decomposição ortogonal do elemento $b - z$.

Pela Fórmula de Pitágoras, obtemos

$$\| b - z \|_2^2 = \| b - z^* \|_2^2 + \| z^* - z \|_2^2 \, .$$

Portanto, temos que

$$\| b - z^* \|_2 \leq \| b - z \|_2 \quad \text{para todo} \quad z \in \mathcal{R}(A) \, ,$$

o que completa da demonstração. ■

Considerando uma matriz $A \in I\!M_{m\times n}(I\!R)$, $m > n$ e *posto*$(A) = n$, e um elemento $b \in I\!R^m$, sabemos que o único elemento $z^* = Ax^* \in \mathcal{R}(A)$, onde $x^* \in I\!R^n$ é a solução de quadrados mínimos para o sistema linear $Ax = b$, é a **melhor aproximação** do elemento b no subespaço $\mathcal{R}(A)$ com relação à norma Euclidiana $\|\cdot\|_2$.

Desse modo, o elemento $z^* = Ax^*$ é a projeção ortogonal do elemento b sobre o subespaço $\mathcal{R}(A)$. Como o elemento $x^* \in I\!R^n$ é a solução do sistema normal

$$A^tAx^* = A^tb \qquad \Longleftrightarrow \qquad A^t(b - Ax^*) = 0_{I\!R^n} ,$$

o elemento $r^* = b - Ax^*$ pertence ao subespaço $\mathcal{N}(A^t) = \mathcal{R}(A)^{\perp}$.

Assim, temos que o elemento $r^* = b - z^*$ é a projeção ortogonal do elemento b sobre o subespaço $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^t)$.

Exemplo 8.11.1 *Sejam a matriz $A \in I\!M_{3\times 5}(I\!R)$ e o elemento $b \in I\!R^5$ dados por:*

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 2 & -1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Encontrar a projeção ortogonal do elemento b no subespaço $\mathcal{N}(A) \subset I\!R^5$.

Sabemos que $\mathcal{N}(A) = \mathcal{R}(A^t)^{\perp}$. Desse modo, basta encontrar a projeção ortogonal do elemento b no subespaço $\mathcal{R}(A^t)$. De fato, vamos denotar por $z^* = A^tx^* \in \mathcal{R}(A^t)$ a projeção ortogonal de b sobre $\mathcal{R}(A^t)$, onde o elemento $x^* \in I\!R^3$ é a solução do **sistema normal** $AA^tx = Ab$. Logo, o elemento $r^* = b - z^*$ é a projeção ortogonal do elemento b sobre o subespaço $\mathcal{N}(A) = \mathcal{R}(A^t)^{\perp}$.

Temos que

$$AA^t = \begin{bmatrix} 7 & 2 & 4 \\ 2 & 7 & 3 \\ 4 & 3 & 7 \end{bmatrix} \qquad \text{e} \qquad Ab = \begin{bmatrix} 4 \\ 3 \\ 6 \end{bmatrix}$$

Portanto, podemos obter $x^* \in I\!R^3$, que é a solução do sistema normal, através da fatoração de Cholesky da matriz normal AA^t.

Exercícios

Exercício 8.91 *Considere a matriz* $A \in \mathbb{M}_{3\times5}(\mathbb{R})$

$$A = \begin{bmatrix} 1 & 1 & 3 & 2 & 0 \\ 2 & 2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -2 & 1 \end{bmatrix}.$$

(a) Determine uma base para cada um dos subespaços $\mathcal{R}(A^t)$ *e* $\mathcal{N}(A)$.

(b) Determine um elemento $b \in \mathbb{R}^5$ *de modo que o sistema linear*

$$A^t x = b$$

tenha solução.

Exercício 8.92 *Sejam a matriz* $A \in \mathbb{M}_{2\times4}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{bmatrix} \quad e \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Determine a projeção ortogonal do elemento b *sobre o subespaço* $\mathcal{N}(A)$.

Exercício 8.93 *Determine a solução de quadrados mínimos para o sistema linear*

$$\begin{bmatrix} 3 & 1 \\ -3 & 1 \\ 3 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix},$$

e verifique que o resíduo $r^* = b - Ax^*$ *é ortogonal ao subespaço* $\mathcal{R}(A)$.

Exercício 8.94 *Determine a solução de quadrados mínimos para o sistema linear*

$$\begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix},$$

e verifique que o resíduo $r^* = b - Ax^*$ *é ortogonal ao subespaço* $\mathcal{R}(A)$.

Exercício 8.95 *Sejam a matriz* $A \in I\!M_{2\times4}(I\!R)$ *e o elemento* $b \in I\!R^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

Faça a representação do elemento b *na forma:*

$$b = b_r + b_n ,$$

onde $b_r \in \mathcal{R}(A^t)$ *e* $b_n \in \mathcal{N}(A)$.

Exercício 8.96 *Considere o espaço vetorial* $I\!R^n$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Sejam um elementos genérico* $Y \in I\!R^n$ *e o elemento* $X \in I\!R^n$ *dados por:*

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} \qquad , \qquad X = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Determine a projeção ortogonal do elemento Y *sobre o subespaço gerado pelo elemento* X. *O que podemos concluir?*

Exercício 8.97 *Sejam* V *um espaço vetorial real munido do produto interno* $\langle \cdot , \cdot \rangle$, *os elementos ortonormais* $u , v \in V$ *e um elemento arbitrário* $w \in V$. *Considere o funcional* $J : I\!R^2 \longrightarrow I\!R$ *definido da seguinte forma:*

$$J(\alpha, \beta) = \| \, w - (\, \alpha u + \beta v \,) \, \|_2^2 \qquad ; \qquad (\, \alpha \, , \, \beta \,) \in I\!R^2 ,$$

onde $\| \cdot \|_2$ *é a norma proveniente do produto interno* $\langle \cdot , \cdot \rangle$. *Pede–se:*

(a) Mostre que o funcional J *possui um único ponto de mínimo* (α^* , β^*).

(b) Dê uma interpretação geométrica para o elemento $w^* = \alpha^* u + \beta^* v$.

(c) Dê uma interpretação geométrica para o elemento $r^* = w - w^*$.

Exercício 8.98 *Considere o espaço vetorial* $\mathbb{R}^n$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida,* $b \in \mathbb{R}^n$ *e o funcional* $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ *definido da seguinte forma:*

$$J(x) = \langle Ax , x \rangle - 2\langle b , x \rangle \quad ; \quad x \in \mathbb{R}^n .$$

Mostre que o ***Problema de Minimização:*** *Encontrar* $x^* \in \mathbb{R}^n$ *tal que*

$$J(x^*) = \min\{ J(x) \quad ; \quad x \in \mathbb{R}^n \}$$

é equivalente ao ***Sistema Linear Positivo–Definido*** $Ax = b$.

O ponto de mínimo x^* *é único? Justifique sua resposta.*

Exercício 8.99 *Considere o espaço vetorial real* $\mathbb{R}^n$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$. *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível,* $b \in \mathbb{R}^n$ *e* $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ *o funcional definido da seguinte forma:*

$$J(x) = \langle Ax , x \rangle - 2\langle b , x \rangle .$$

Pede–se:

(a) Determine o gradiente e a matriz Hessiana do funcional J.

(b) Determine a caracterização dos pontos críticos de J.

(c) Considere a matriz A *e o vetor* b *dados por:*

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 2 \\ 1 \end{bmatrix} .$$

Classifique os pontos críticos de J. *Faça um gráfico das curvas de nível de* J.

Exercício 8.100 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e* T_r *a transformação linear definida por:*

$$\begin{aligned} T_r : \ \mathcal{R}(A^t) &\longrightarrow \mathcal{R}(A) \\ x_r &\longrightarrow T_r(x_r) = Ax_r \end{aligned}$$

Mostre que T_r *é um isomorfismo de* $\mathcal{R}(A^t)$ *em* $\mathcal{R}(A)$.

Exercício 8.101 *Sejam V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$, com $dim(V) \geq 2$, os elementos linearmente independentes $u , v \in V$ e um elemento arbitrário $w \in V$. Considere o funcional $J : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definido da forma:*

$$J(\alpha, \beta) = \| w - (\alpha u + \beta v) \|_2^2 \qquad ; \qquad (\alpha , \beta) \in \mathbb{R}^2 ,$$

onde $\| \cdot \|_2$ é a norma proveniente do produto interno $\langle \cdot , \cdot \rangle$. Pede–se:

(a) Mostre que o funcional J possui um único ponto de mínimo (α^* , β^*).

(b) Dê uma interpretação geométrica para o elemento $w^* = \alpha^* u + \beta^* v$.

(c) Dê uma interpretação geométrica para o elemento $r^* = w - w^*$.

Exercício 8.102 *Considere a matriz $A \in \mathbb{M}_{3\times5}(\mathbb{R})$ dada por:*

$$A = \begin{bmatrix} 1 & -1 & -2 & 0 & 1 \\ -3 & 4 & 5 & 1 & -2 \\ -2 & 2 & 3 & 1 & -1 \end{bmatrix} .$$

Pede–se:

(a) Determine uma base para cada um dos subespaços $\mathcal{R}(A^t)$ e $\mathcal{N}(A)$.

(b) Determine a projeção ortogonal do elemento $b \in \mathbb{R}^5$ dado por:

$$b = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix} ,$$

no subespaço $\mathcal{N}(A)$.

Exercício 8.103 *Considere a matriz $A \in \mathbb{M}_{4\times3}(\mathbb{R})$ e o elemento $b \in \mathbb{R}^4$ dados por:*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} .$$

Encontre a projeção ortogonal do elemento b sobre o subespaço $\mathcal{R}(A)$, através da fatoração de Cholesky para obter a solução do sistema normal.

8.12 Matriz de Projeção Ortogonal

Sejam uma matriz $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m > n$ e $posto(A) = n$, e o elemento $b \in \mathbb{R}^m$. Sabemos que a solução de quadrados mínimos x^* para o sistema linear sobredeterminado $Ax = b$ é a solução do sistema normal $A^tAx = A^tb$, isto é, $x^* = (A^tA)^{-1}A^tb$. Desse modo, o elemento $z^* = Ax^*$ é a projeção ortogonal do elemento b sobre o subespaço $\mathcal{R}(A)$, que é representado na forma $z^* = (\, A(A^tA)^{-1}A^t \,)b$. Portanto, a matriz

$$P = A(A^tA)^{-1}A^t = AA^\dagger$$

é a **matriz de projeção ortogonal** sobre o subespaço $\mathcal{R}(A)$.

Desse modo, o elemento $r^* = b - z^* = b - P\,b$ é a projeção ortogonal do elemento b sobre o subespaço $\mathcal{R}(A)^\perp = \mathcal{N}(A^t)$. Logo, a matriz $(I - P)$ é a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)^\perp = \mathcal{N}(A^t)$.

Proposição 8.12.1 *A matriz $P \in \mathbb{M}_m(\mathbb{R})$ é simétrica e idempotente se, e somente se, P projeta ortogonalmente cada elemento $b \in \mathbb{R}^m$ sobre o subespaço $\mathcal{R}(P)$.*

Demonstração

$(\Longrightarrow)$ Considerando a hipótese de P simétrica $(P^t = P)$ e idempotente $(P^2 = P)$, vamos mostra que para cada elemento $b \in \mathbb{R}^m$ temos que o elemento $P\,b \in \mathbb{R}^m$ é a projeção ortogonal do elemento b no subespaço $\mathcal{R}(P)$. É fácil ver que $P\,b \in \mathcal{R}(P)$. Desse modo, basta mostrar que o elemento $(b - P\,b) \in \mathcal{R}(P)^\perp$. De fato,

$$P^t\,(b - P\,b) = P^t\,(I - P)\,b = (P^t - P^t\,P)\,b = (P - P^2)\,b = 0\,.$$

Portanto, mostramos que o elemento $(b - P\,b) \in \mathcal{N}(P^t) = \mathcal{R}(P)^\perp$.

$(\Longleftarrow)$ Considerando que P projeta ortogonalmente cada elemento $b \in \mathbb{R}^m$ sobre o subespaço $\mathcal{R}(P)$, vamos mostrar que a matriz P é simétrica e idempotente. De fato, temos que $P\,b \in \mathcal{R}(P)$ e que $(b - P\,b) \in \mathcal{N}(P^t) = \mathcal{R}(P)^\perp$, portanto,

$$P^t\,(b - P\,b) = (P^t - P^t\,P)\,b = 0 \qquad ; \qquad \forall\; b \in \mathbb{R}^m$$

o que implica em $(P^t - P^t\,P) = 0$. Desse modo, $P^t = P^t\,P$ e $P = P^t\,P$, logo, a matriz P é simétrica. Agora, utilizando a simetria da matriz P e a igualdade $P = P^t\,P$, obtemos $P^2 = P$. O que completa a demonstração. ■

Definição 8.12.1 *Dizemos que* $P \in \mathbb{M}_m(\mathbb{R})$ *é uma matriz de* ***projeção ortogonal*** *se* P *for uma matriz simétrica e idempotente.*

Proposição 8.12.2 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, *com* $m > n$ *e* $posto(A) = n$, *e a matriz*

$$P = A(A^tA)^{-1}A^t = AA^\dagger \in \mathbb{M}_m(\mathbb{R}) .$$

Então,

1. $\mathcal{R}(P) = \mathcal{R}(A)$.

2. P *é uma matriz simétrica e idempotente.*

Demonstração

1. Seja $z \in \mathcal{R}(P)$, então o elemento z é escrito da seguinte forma $z = Px$ para algum $x \in \mathbb{R}^m$. Assim, temos que

$$z = Px = A(A^tA)^{-1}A^tx = A(\,(A^tA)^{-1}A^tx\,) = Ay$$

onde o elemento $y = (A^tA)^{-1}A^tx$.

Desse modo, mostramos que o elemento $z = Px \in \mathcal{R}(A)$, isto é, $\mathcal{R}(P) \subset \mathcal{R}(A)$.

Tomando agora um elemento $z \in \mathcal{R}(A)$. Podemos escrever $x = Px$, pois P é a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$. Logo, o elemento $x \in \mathcal{R}(P)$. Portanto, $\mathcal{R}(A) \subset \mathcal{R}(P)$, o que completa a prova do item *1.*

2. O fato da matriz P ser simétrica e idempotente pode ser verificado diretamente ou como um conseqüência imediata do item *1.* e da Proposição 8.12.1, pois sabemos que P projeta ortogonalmente cada elemento $b \in \mathbb{R}^m$ sobre o subespaço $\mathcal{R}(A)$. O que completa a demonstração. ■

Exemplo 8.12.1 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *e o subespaço* $S = [v]$ *com*

$$v = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} .$$

Seja P *o operador linear sobre* $I\!R^3$ *de modo que o elemento* $w \in I\!R^3$ *dado por:*

$$w = P(u) \qquad para \qquad u \in I\!R^3 ,$$

é a projeção ortogonal do elemento u *sobre o subespaço* S*. Vamos determinar os autovalores e autovetores de* P*.*

Sabemos que $w = P(u) = \alpha^* v \in S$ com

$$\alpha^* = \frac{\langle u , v \rangle}{\langle v , v \rangle} = \frac{v^t u}{v^t v} .$$

Assim, temos que o elemento w pode ser escrito da seguinte forma:

$$w = P(u) = \frac{v^t u}{v^t v} v = \frac{v v^t}{v^t v} u$$

Considerando o espaço vetorial $I\!R^3$ com a base canônica β, temos que

$$[P]^{\beta}_{\beta} = \frac{v v^t}{v^t v} = \frac{1}{6} \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{bmatrix} .$$

Sabemos que, $P(z) = z$ para todo $z \in S$. Portanto, $\lambda_1 = 1$ é um autovalor do operador linear P com o elemento $v_1 \in I\!R^3$ dada por:

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

o autovetor associado. Logo, S é o subespaço associado ao autovalor $\lambda_1 = 1$.

O complemento ortogonal do subespaço S em $I\!R^3$, $S^\perp$, é o hiperplano dado por:

$$S^\perp = H = \{ u \in I\!R^3 \ / \ \langle u , v \rangle = 0 \}$$

Note que o subespaço $S^\perp$ é um plano em $I\!R^3$ dado pela equação

$$x - y + 2z = 0 .$$

Desse modo, temos que

$$P(u) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

para todo $u \in S^\perp$, isto é, $Ker(P) = S^\perp$.

Desse modo, temos que $P(u) = 0\,u$ para todo $u \in S^\perp$. Assim, podemos concluir que $\lambda_2 = 0$ é um autovalor de P e $S^\perp$ é o subespaço associado ao autovalor λ_2. Assim, quaisquer dois elementos v_2 e v_3 linearmente independentes em $S^\perp$ são autovetores associados ao autovalor $\lambda_2 = 0$. Desse modo, podemos escolher os elementos

$$v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad v_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

como sendo os autovetores do operador linear P associados ao autovalor $\lambda_2 = 0$.

Exemplo 8.12.2 *Considere o espaço vetorial real* $I\!R^3$ *munido do produto interno usual* $\langle \cdot , \cdot \rangle$ *e o subespaço* $S = [v]$ *com*

$$v = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} .$$

Seja R *o operador linear sobre* $I\!R^3$ *de modo que o elemento* $w \in I\!R^3$ *dado por:*

$$w = R(u) \qquad para \qquad u \in I\!R^3 ,$$

é a reflexão do elemento u *em torno do subespaço* $S^\perp$*. Vamos determinar os autovalores e autovetores de* R.

Do Exemplo 8.12.1, sabemos que o operador P de projeção ortogonal sobre o subespaço S é dado por:

$$P(u) = \frac{v^t\,u}{v^t\,v}\,v = \frac{v\,v^t}{v^t\,v}\,u \qquad \text{para todo} \qquad u \in I\!R^3 .$$

Desse modo, o operador T de projeção ortogonal sobre o subespaço $S^{\perp}$ é dado por:

$$T(u) = u - P(u) = u - \frac{v^t\,u}{v^t\,v}\,v = \left(I - \frac{v\,v^t}{v^t\,v} \right) u .$$

Temos que o operador linear R de reflexão em torno do subespaço $S^{\perp}$ é dado por:

$$R(u) = T(u) - P(u) = u - 2\,P(u) = = \left(I - 2\,\frac{v\,v^t}{v^t\,v} \right) u .$$

Assim, temos que $R(u) = u$ para todo $u \in S^{\perp}$. Logo, concluímos que $\lambda_1 = 1$ é um autovalor de R e $S^{\perp}$ é o subespaço associado ao autovalor λ_1. Desse modo, quaisquer dois elementos v_1 e v_2 linearmente independentes em $S^{\perp}$ são autovetores associados ao autovalor $\lambda_1 = 1$. Portanto, podemos escolher os elementos

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad v_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

como sendo os autovetores do operador linear R associados ao autovalor $\lambda_1 = 1$.

Sabemos que, $R(w) = -w$ para todo $w \in S$. Portanto, $\lambda_2 = -1$ é um autovalor do operador linear R com o elemento $v_3 \in I\!R^3$ dada por:

$$v_3 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

o autovetor associado. Logo, S é o subespaço associado ao autovalor $\lambda_2 = -1$.

Podemos fazer as seguintes considerações sobre o operador $P : I\!R^n \longrightarrow I\!R^n$ de projeção ortogonal sobre um subespaço $S \subset I\!R^n$:

(a) $Im(P) = S$ *(b)* $P^2 = P$ *(c)* A matriz $[P]^{\beta}_{\beta}$ é simétrica

Note que para o operador linear $R : I\!R^n \longrightarrow I\!R^n$ de reflexão em torno de um subespaço $S \subset I\!R^n$ temos que $R^2 = I$.

Exercícios

Exercício 8.104 *Sejam a matriz* $A \in I\!M_{2\times 4}(I\!R)$ *e o elemento* $b \in I\!R^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

Calcule a projeção ortogonal do elemento b *sobre o subespaço* $\mathcal{N}(A)$ *e a respectiva matriz de projeção ortogonal.*

Exercício 8.105 *Considere o espaço vetorial real* $I\!R^3$ *munido com o produto interno usual. Encontre a matriz de projeção ortogonal sobre o plano coordenado* xy.

Exercício 8.106 *Considere o espaço vetorial real* $I\!R^3$ *munido com o produto interno usual. Encontre a matriz de projeção ortogonal sobre o plano coordenado* xz.

Exercício 8.107 *Sejam* $b \in I\!R^m$, $A \in I\!M_{m\times n}(I\!R)$ *com* $m > n$ *e posto*$(A) = n$, *e o elemento* $x^* \in I\!R^n$ *solução de quadrados mínimos do sistema linear* $Ax = b$. *Mostre que o elemento* $r^* = b - Ax^*$ *pertence ao subespaço* $\mathcal{N}(A^t)$. *Qual é a pseudo–inversa da matriz* A? *Justifique sua resposta. Mostre que a pseudo–inversa* $A^\dagger$ *satisfaz as propriedades:*

1. $AA^\dagger$ *é uma matriz simétrica.*

2. $A^\dagger A$ *é uma matriz simétrica.*

3. $AA^\dagger A = A$.

4. $A^\dagger AA^\dagger = A^\dagger$.

Exercício 8.108 *Seja* $P \in I\!M_m(I\!R)$ *uma matriz de projeção ortogonal, isto é,*

$$P^t = P \qquad e \qquad P^2 = P.$$

Mostre que $R = 2P - I$ *é uma matriz ortogonal. Faça a interpretação geométrica para os elementos* $z^* = Pb$ *e* $v^* = Rb$, *onde* $b \in I\!R^m$.

Exercício 8.109 *Seja* $P \in \mathbb{M}_m(\mathbb{R})$ *uma matriz de projeção ortogonal. Mostre que a matriz* $Q = I - P$ *é uma matriz de projeção ortogonal. Faça a interpretação geométrica para os elementos* $z^* = Pb$ *e* $v^* = Qb$*, onde* $b \in \mathbb{R}^m$.

Exercício 8.110 *Considere o espaço vetorial* $\mathbb{R}^m$ *com o produto interno usual. Sejam* S_1 *e* S_2 *subespaços do* $\mathbb{R}^m$, $P_1 \in \mathbb{M}_m(\mathbb{R})$ *a matriz de projeção ortogonal sobre o subespaço* S_1 *e* $P_2 \in \mathbb{M}_m(\mathbb{R})$ *a matriz de projeção ortogonal sobre o subespaço* S_2*, com a propriedade* $P_1P_2 = P_2P_1 = 0_m$*. Pede–se:*

(a) Mostre que $P_1 + P_2$ *é uma matriz de projeção ortogonal.*

(b) Mostre que os subespaços S_1 *e* S_2 *são ortogonais.*

(c) Mostre que $P_1 + P_2$ *é a matriz de projeção ortogonal sobre o subespaço*

$$W = S_1 \oplus S_2 .$$

Exercício 8.111 *Considere o espaço vetorial* $\mathbb{R}^n$ *com o produto interno usual. Seja* S *o subespaço gerado pelo elemento* $u \in \mathbb{R}^n$ *não–nulo. Determine a matriz* P *de projeção ortogonal sobre o subespaço* S *e a matriz de projeção ortogonal sobre o subespaço* $S^{\perp}$.

Exercício 8.112 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com o produto interno usual. Seja* S *o subespaço do* $\mathbb{R}^3$ *definido pela equação*

$$x - 2y + 3z = 0 .$$

Determine a matriz P *de projeção ortogonal sobre o subespaço* S.

Exercício 8.113 *Considere o espaço vetorial real* $\mathbb{R}^3$ *com o produto interno usual. Seja* S *o subespaço do* $\mathbb{R}^3$ *definido pela equação*

$$x - 2y + 3z = 0 .$$

Determine a matriz Q *de reflexão sobre o subespaço* S.

Exercício 8.114 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ *uma matriz ortogonal e* $b \in \mathbb{R}^m$*. Determine a projeção ortogonal do elemento* b *sobre o subespaço* $\mathcal{R}(A)$ *e a respectiva matriz de projeção ortogonal. Qual é a dimensão do subespaço* $\mathcal{R}(A)$*?*

Exercício 8.115 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$*, com* $m > n$ *e* $posto(A) = n$*, e* $b \in \mathcal{N}(A^t)$*. Determine a solução de quadrados mínimos para o sistema linear* $Ax = b$.

Exercício 8.116 *Considere o espaço vetorial* $\mathbb{R}^m$ *com o produto interno usual* $\langle \cdot , \cdot \rangle$ *e o elemento* $u \in \mathbb{R}^m$ *tal que* $\langle u , u \rangle = 1$. *Definimos as seguintes matrizes*

$$P = uu^t \qquad e \qquad Q = I - 2P ,$$

onde a matriz Q *é denominada* ***matriz de Householder***. *Pede–se:*

(a) Mostre que $Pw = w$, *com* $w = \alpha u$ *para todo* $\alpha \in \mathbb{R}$. *Dê uma interpretação geométrica.*

(b) Mostre que $Pv = 0_{\mathbb{R}^m}$ *para* $\langle u , v \rangle = 0$. *Dê uma interpretação geométrica.*

(c) Mostre que $Q^t = Q$ *e* $Q^2 = I$. *O que podemos dizer da matriz* Q*?*

(d) Mostre que $Qw = -w$, *com* $w = \alpha u$ *para todo* $\alpha \in \mathbb{R}$. *Dê uma interpretação geométrica.*

(e) Mostre $Qv = v$ *para* $\langle u , v \rangle = 0$. *Dê uma interpretação geométrica.*

8.13 Fatoração QR

Nesta seção apresentamos um procedimento para obter a fatoração QR de uma matriz $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, com $m > n$ e $posto(A) = n$, utilizando o Processo de Ortogonalização de Gram–Schmidt, isto é, vamos determinar uma matriz ortogonal $Q \in \mathbb{M}_{m\times n}(\mathbb{R})$ e uma matriz triangular superior $R \in \mathbb{M}_n(\mathbb{R})$, com os elementos da diagonal principal positivos, tais que $A = QR$. Esse processo é muito importante na determinação da solução de quadrados mínimos para sistemas lineares, de uma maneira geral. Procuramos apresentar interpretações geométricas e fazer as relações com os subespaços fundamentais da matriz A. Todos os resultados desenvolvidos nessa seção estão fortemente baseados na Teorema 5.7.1, que é o Processo de Ortogonalização de Gram-Schmidt.

Nesse momento é importante recordar o Teorema 5.7.2 e sua demonstração. Desse modo, sabemos que todo espaço vetorial de dimensão finita munido de um produto interno tem uma base ortonormal. Sejam V um espaço vetorial munido de um produto interno e $\beta = \{ v_1 , \cdots , v_n \}$ uma base ordenada para V. A partir dessa base, vamos obter uma base ortogonal, através do processo de ortogonalização de Gram–Schmidt.

Do Teorema 5.7.2, temos que, para $j = 2, \cdots n$ e $q_1 = v_1$,

$$q_j = v_j - \sum_{i=1}^{j-1} \alpha_{ij}\, q_i \quad \Longrightarrow \quad \alpha_{ij} = \frac{\langle v_j , q_i \rangle}{\| q_i \|_2^2} \qquad \text{para} \qquad i = 1, \cdots , j-1$$

Como $v_1 , \cdots , v_j$ são linearmente independentes, tem–se que o elemento $q_j \neq 0_V$. Além disso, sabemos que o subespaço $S_j = [v_1, \cdots , v_j] = W_j = [q_1, \cdots , q_j]$.

Assim, obtemos uma base ortogonal $\beta' = \{ q_1 , \cdots , q_n \}$. Finalmente, fazendo

$$q_j^* = \frac{q_j}{\| q_j \|_2} \qquad \text{para} \qquad j = 1, \cdots , n \,,$$

obtemos uma base ortonormal $\beta^* = \{ q_1^* , \cdots , q_n^* \}$.

Finalmente, podemos observar que o elemento w_j escrito na forma:

$$w_j = \sum_{i=1}^{j-1} \frac{\langle v_j , q_i \rangle}{\langle q_i , q_i \rangle}\, q_i$$

é a projeção ortogonal do elemento v_j sobre o subespaço W_{j-1}. Assim, o elemento q_j é a projeção ortogonal do elemento v_j sobre o complemento ortogonal de W_{j-1}.

Matriz de Mudança de Base

Durante o processo de ortogonalização de Gram–Schmidt obtemos a construção da matriz de mudança da base ordenada β para a base ordenada β^*. De fato, os elementos da base $\beta = \{ v_1 , \cdots , v_n \}$ são escritos da seguinte forma:

$$v_j = \sum_{i=1}^{j} r_{ij}\, q_i^* \qquad \text{para} \qquad j = 1, \cdots , n \,,$$

onde os coeficientes da combinação linear r_{ij} são dados por:

$$\begin{aligned} r_{ij} &= \langle v_j , q_i^* \rangle \quad &;& \quad i = 1, \cdots , j-1 \quad \text{e} \quad j = 2, \cdots , n \\ r_{ij} &= 0 \quad &;& \quad i > j \\ r_{jj} &= \| q_j \|_2 \quad &;& \quad j = 1, \cdots , n \end{aligned}$$

Desse modo, $R = [r_{ij}] \in \mathbb{M}_n(\mathbb{R})$ é a matriz de mudança da base ordenada β para a base ordenada β^*. Podemos observar facilmente que R é uma matriz triangular superior. Observamos, também, que a matriz R é construída durante o processo de ortogonalização por colunas, isto é, na construção do elemento q_j^* construímos a j–ésima coluna da matriz R. Assim, a matriz R é construída da seguinte forma:

$$R = \begin{bmatrix} \| q_1 \|_2 & \langle v_2 , q_1^* \rangle & \langle v_3 , q_1^* \rangle & \cdots & \langle v_n , q_1^* \rangle \\ 0 & \| q_2 \|_2 & \langle v_3 , q_2^* \rangle & \cdots & \langle v_n , q_2^* \rangle \\ 0 & 0 & \| q_3 \|_2 & \cdots & \langle v_n , q_3^* \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \| q_n \|_2 \end{bmatrix} = [I]^{\beta}_{\beta^*} \,,$$

onde $[I]^{\beta}_{\beta^*}$ denota a matriz de mudança da base ordenada β para a base ordenada β^*.

Fatoração QR. Método de Gram–Schmidt

Teorema 8.13.1 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$*, com* $m > n$ *e* $posto(A) = n$*. Então, existe uma matriz ortogonal* $Q \in \mathbb{M}_{m\times n}(\mathbb{R})$ *e uma matriz triangular superior* $R \in \mathbb{M}_n(\mathbb{R})$*, com todos os elementos da diagonal principal positivos, tais que* $A = QR$*.*

Demonstração – A prova é feita de modo construtivo, isto é, vamos exibir uma maneira de construir as matrizes Q e R.

Para isso, vamos utilizar o processo de ortogonalização de Gram–Schmidt para obter a fatoração QR de uma matriz $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, com $m > n$ e $posto(A) = n$. Desse modo, temos que o conjunto $\beta = \{ v_1 , \cdots , v_j , \cdots , v_n \}$, onde $v_j \in \mathbb{R}^m$ é a j–ésima coluna da matriz A, é *linearmente independente* em $\mathbb{R}^m$, isto é, β é uma base ordenada para o subespaço $\mathcal{R}(A)$.

Através do processo de ortogonalização de Gram–Schmidt obtemos uma base ortonormal $\beta^* = \{ q_1^* , \cdots , q_j^* , \cdots , q_n^* \}$, e a matriz de mudança de base $R \in \mathbb{M}_n(\mathbb{R})$, que é uma matriz triangular superior com os elementos da diagonal principal todos positivos. Assim, temos a fatoração $A = Q\,R$, onde $Q \in \mathbb{M}_{m\times n}(\mathbb{R})$ é uma matriz ortogonal que tem como j–ésima coluna o elemento q_j^* da base ortonormal β^*. De fato, todo elemento $v \in \mathcal{R}(A)$ pode ser escrito das seguintes forma:

$$v = Ax \qquad \text{e} \qquad v = Qz = Q(Rx) = (QR)x$$

para todo $x \in \mathbb{R}^n$, uma vez que $z = Rx$.

Portanto, temos a fatoração $A = QR$, o que completa a demonstração. ■

Teorema 8.13.2 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$*, com* $m \geq n$ *e* $posto(A) = n$*. Então, existe uma única matriz* $Q \in \mathbb{M}_{m\times n}(\mathbb{R})$ *ortogonal e uma única matriz* $R \in \mathbb{M}_n(\mathbb{R})$*, com os elementos da diagonal principal positivos, tais que* $A = QR$*.*

Demonstração – A prova é baseada na conexão entre a fatoração QR da matriz A e a fatoração de Cholesky da matriz A^tA, considerando a unicidade do fator de Cholesky e do fato que a matriz R é invertível. □

A seguir, apresentamos o algoritmo para obter a fatoração QR de uma matriz A, com $m \geq n$ e $posto(A) = n$, pelo processo de ortogonalização de Gram–Schmidt.

Algoritmo 8.13.1 (Método de Gram–Schmidt)

```
for j = 1,2, ... ,n
    Q(:,j) = A(:,j)
    for i = 1,2, ... ,(j - 1)
        R(i,j) = dot(A(:,j),Q(:,i))
        Q(:,j) = Q(:,j) - R(i,j)*Q(:,i)
    end
    R(j,j) = sqrt(dot(Q(:,j),Q(:,j)))
    Q(:,j) = Q(:,j) / R(j,j)
end
```

No algoritmo acima as funções `dot(·,·)` e `sqrt(·)` denotam os procedimentos para o cálculo do produto interno e da raíz quadrada, respectivamente.

Fatoração QR. Método de Gram–Schmidt Modificado

Nessa seção apresentamos um algoritmo para obter a fatoração QR de uma matriz $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, com $m \geq n$ e $posto(A) = n$, através de uma pequena alteração no processo de ortogonalização de Gram–Schmidt. O objetivo dessa modificação é tornar o método numericamente mais estável. Em alguns casos os elementos que formam as colunas da matriz Q obtida pelo Método de Gram–Schmidt apresentam uma perda de ortogonalidade, devido os erros da aritmética de ponto flutuante.

Para exemplificar vamos considerar o espaço vetorial real $\mathbb{R}^m$ com o produto interno usual e os elementos $v_1 , \cdots , v_4$ linearmente independentes em $\mathbb{R}^m$.

Inicialmente tomamos os elementos

$$\begin{array}{lclcl} q_1 & = & v_1 & \text{e} & q_1^* = \dfrac{q_1}{\| q_1 \|_2} \\[2ex] q_2 & = & v_2 - \langle v_2 , q_1^* \rangle q_1^* & \text{e} & q_2^* = \dfrac{q_2}{\| q_2 \|_2} \end{array} \tag{8.167}$$

O elemento q_3 é construído em partes da seguinte forma:

$$\begin{aligned} q_3^{(1)} &= v_3 - \langle v_3 , q_1^* \rangle q_1^* \\ q_3 &= q_3^{(1)} - \langle q_3^{(1)} , q_2^* \rangle q_2^* \\ q_3^* &= \frac{q_3}{\| q_3 \|_2} \end{aligned} \tag{8.168}$$

O elemento q_4 é construído em partes da seguinte forma:

$$\begin{aligned} q_4^{(1)} &= v_4 - \langle v_4 , q_1^* \rangle q_1^* \\ q_4^{(2)} &= q_4^{(1)} - \langle q_4^{(1)} , q_2^* \rangle q_2^* \\ q_4 &= q_4^{(2)} - \langle q_4^{(2)} , q_3^* \rangle q_3^* \\ q_4^* &= \frac{q_4}{\| q_4 \|_2} \end{aligned} \tag{8.169}$$

O procedimento descrito em (8.167)–(8.169) é denominado **Método de Gram–Schmidt Modificado**.

Para um conjunto $\{ v_1 , \cdots , v_n \}$ linearmente independente em $\mathbb{R}^m$, consideramos que já foram construídos os elementos $q_1^* , \cdots , q_{j-1}^*$, através do procedimento descrito em (8.167)–(8.169), o elemento q_j é construído da seguinte forma:

$$
\begin{aligned}
q_j &= v_j \\
q_j &= q_j - \langle q_j , q_k^* \rangle q_k^* \qquad \text{para} \qquad k = 1, \cdots (j-1) \\
q_j^* &= \frac{q_j}{\| q_j \|_2}
\end{aligned}
\tag{8.170}
$$

para $j = 1, \cdots , n$.

A seguir, apresentamos um algoritmo eficiente para obter a fatoração QR de uma matriz A, com $m \geq n$ e $posto(A) = n$, pelo processo de ortogonalização de Gram–Schmidt modificado.

Algoritmo 8.13.2 (Método de Gram–Schmidt Modificado)

```
    Q = A
for j = 1,2, ... ,n
    R(j,j) = sqrt(dot(Q(:,j),Q(:,j)))
    Q(:,j) = Q(:,j) / R(j,j)
    for i = (j + 1), ... ,n
        R(j,i) = dot(Q(:,j),Q(:,i))
        Q(:,i) = Q(:,i) - R(j,i)*Q(:,j)
    end
end
```

No algoritmo acima as funções `dot(·,·)` e `sqrt(·)` denotam os procedimentos para o cálculo do produto interno e da raíz quadrada, respectivamente.

Sejam $A \in I\!M_{m\times n}(I\!R)$, com $m \geq n$ e $posto(A) = n$, e a fatoração $A = QR$, onde $Q \in I\!M_m(I\!R)$ e $R \in I\!M_{m\times n}(I\!R)$, descritas da seguinte forma:

$$Q = \begin{bmatrix} \widehat{Q} & \widetilde{Q} \end{bmatrix} \qquad \text{e} \qquad R = \begin{bmatrix} \widehat{R} \\ 0_{r\times n} \end{bmatrix},$$

onde $\widehat{Q} \in I\!M_{m\times n}(I\!R)$ e $\widetilde{Q} \in I\!M_{m\times r}(I\!R)$ são matrizes ortogonais, $\widehat{R} \in I\!M_n(I\!R)$ é uma matriz triangular superior com os elementos da diagonal principal positivos e $0_{r\times n}$ é a matriz nula de ordem $r \times n$, com $r = m - n$.

Vamos mostrar como construir as matrizes ortogonais $\widehat{Q}$ e $\widetilde{Q}$ fazendo uma relação geométrica da fatoração $A = QR$ com os subespaços fundamentais $\mathcal{R}(A)$ e $\mathcal{N}(A^t)$.

Por simplicidade, vamos denotar as matrizes $A \in I\!M_{m\times n}(I\!R)$ e $Q \in I\!M_m(I\!R)$ da seguinte forma:

$$A = [v_1 \cdots v_j \cdots v_n] \qquad \text{e} \qquad Q = [q_1^* \cdots q_j^* \cdots q_n^* \;\; w_1^* \cdots w_j^* \cdots w_r^*] \,,$$

onde $v_j \in I\!R^m$ é a j–ésima coluna da matriz A, $q_j^* \in I\!R^m$ é a j–ésima coluna da matriz $\widehat{Q}$ e $w_j^* \in I\!R^m$ é a j–ésima coluna da matriz $\widetilde{Q}$. Vamos denotar a matriz $\widehat{R} \in I\!M_n(I\!R)$ da seguinte forma:

$$\widehat{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{bmatrix}.$$

Tomando um elemento $z \in \mathcal{R}(A)$, temos que $z = Ax$ para algum $x \in I\!R^n$. Desse modo, substituindo a fatoração $A = QR$ na expressão de z, obtemos

$$\begin{aligned} z &= Ax = \sum_{j=1}^{n} x_j v_j \\ &= QRx \\ &= \widehat{Q}y = \sum_{j=1}^{n} y_j q_j^* \end{aligned}$$

onde $y = \widehat{R}x \in I\!R^n$. Note que, como $posto(A) = n$, temos que $\beta = \{ v_1 \,, \cdots \,, v_n \}$, cujos elementos são as colunas da matriz A, é uma base ordenada de $\mathcal{R}(A)$.

Portanto, observamos que as colunas da matriz $\widehat{Q}$ formam uma base ortonormal para o subespaço $\mathcal{R}(A)$. Assim, a base ortonormal $\beta^* = \{ q_1^* , \cdots , q_n^* \}$ pode ser obtida da base ordenada $\beta = \{ v_1 , \cdots , v_n \}$ de $\mathcal{R}(A)$ através do processo de ortogonalização de Gram–Schmidt. Além disso, $\widehat{R}$ é a matriz de mudança da base ordenada β para a base ortonormal β^*.

Finalmente, como $Q \in \mathbb{M}_m(\mathbb{R})$ é uma matriz ortogonal, observamos facilmente que $\Gamma^* = \{ w_1^* , \cdots , w_j^* , \cdots , w_r^* \}$, cujos elementos são as colunas da matriz $\widetilde{Q}$, é uma base ortonormal para o subespaço $\mathcal{N}(A^t) = \mathcal{R}(A)^\perp$. Assim, podemos encontrar uma base ordenada $\Gamma = \{ w_1 , \cdots , w_j , \cdots , w_r \}$ para o subespaço $\mathcal{N}(A^t)$, que é definido por:

$$\mathcal{N}(A^t) = \{ z \in \mathbb{R}^m \ / \ A^t z = 0_{\mathbb{R}^n} \} ,$$

obtendo a solução geral do sistema linear homogêneo

$$A^t z = 0_{\mathbb{R}^n} \qquad \Longleftrightarrow \qquad \langle v_j , z \rangle = 0 \quad \text{para} \quad j = 1, \cdots , n .$$

Em seguida, determinamos a base ortonormal Γ^* a partir da base ordenada Γ através do processo de ortogonalização de Gram–Schmidt. Note que os elementos

$$w_1 , \cdots , w_j , \cdots , w_r \in \mathbb{R}^m$$

são as soluções básicas do sistema linear homogêneo $A^t z = 0_{\mathbb{R}^n}$, que possui n equações lineares com $r = m - n$ variáveis livres. Todo esse procedimento está justificado pelos resultados apresentados na seção 5.8.

Fatoração QR. *Solução de Quadrados Mínimos*

Teorema 8.13.3 *Considere uma matriz* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, *com* $m > n$ *e* $posto(A) = n$, *e um elemento* $b \in \mathbb{R}^m$. *Se conhecemos a fatoração* $A = QR$, *com* $Q \in \mathbb{M}_{m \times n}(\mathbb{R})$, *uma matriz ortogonal, e* $R \in \mathbb{M}_n(\mathbb{R})$, *uma matriz triangular superior, então*

(a) A solução de quadrados mínimos $x^* \in \mathbb{R}^n$, *para o sistema linear* $Ax = b$, *é a única solução do sistema triangular superior* $Rx = Q^t b$.

(b) A matriz $A^\dagger = R^{-1} Q^t$ *é a pseudo–inversa da matriz* A.

(c) A matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$ *é dada por* $P = QQ^t$.

Demonstração – *(a)* A prova segue do fato que a solução de quadrados mínimos é a única solução do sistema normal e utilizar a fatoração $A = QR$, considerando que R é uma matriz invertível.

Para mostrar o ítem *(b)*, basta utilizar a decomposição $A = QR$ na expressão da pseudo–inversa da matriz A. A prova do item *(c)* é feita de modo análogo. $\square$

Finalmente, apresentamos o procedimento de como obter a solução de quadrados mínimos para o sistema linear sobredeterminado $Ax = b$ utilizando a fatoração $A = QR$, a partir da sua definição.

Sejam $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m \geq n$ e $posto(A) = n$, e a fatoração $A = QR$, onde $Q \in \mathbb{M}_m(\mathbb{R})$ e $R \in \mathbb{M}_{m \times n}(\mathbb{R})$, descritas da seguinte forma:

$$Q = \begin{bmatrix} \widehat{Q} & \widetilde{Q} \end{bmatrix} \quad \text{e} \quad R = \begin{bmatrix} \widehat{R} \\ 0_{r \times n} \end{bmatrix},$$

onde $\widehat{Q} \in \mathbb{M}_{m \times n}(\mathbb{R})$ e $\widetilde{Q} \in \mathbb{M}_{m \times r}(\mathbb{R})$ são matrizes ortogonais, $\widehat{R} \in \mathbb{M}_n(\mathbb{R})$ é uma matriz triangular superior invertível e $0_{r \times n}$ é a matriz nula de ordem $r \times n$, com $r = m - n$.

A solução de quadrados mínimos $x^* \in \mathbb{R}^n$ é definida como:

$$\| Ax^* - b \|_2 = \min\{ \| Ax - b \|_2 \quad ; \quad x \in \mathbb{R}^n \} . \tag{8.171}$$

Fazendo uso do fato que a norma Euclidiana em $I\!R^n$ é invariante por transformação ortogonal, e substituindo a fatoração $A = QR$ em (8.171), obtemos

$$\begin{aligned} \| Ax^* - b \|_2 &= \min\{ \| Ax - b \|_2 \quad ; \quad x \in I\!R^n \} \\ &= \min\{ \| Q^t(Ax - b) \|_2 \quad ; \quad x \in I\!R^n \} \\ &= \min\{ \| Q^t(QRx - b) \|_2 \quad ; \quad x \in I\!R^n \} \\ &= \min\{ \| Rx - Q^t b \|_2 \quad ; \quad x \in I\!R^n \} \end{aligned} \tag{8.172}$$

Inicialmente vamos analisar com mais detalhe $\| Rx - Q^t b \|_2$. Para isso, utilizamos a seguinte representação

$$Rx - Q^t b = \begin{bmatrix} \widehat{R} \\ 0_{r \times n} \end{bmatrix} \begin{bmatrix} x \end{bmatrix} - \begin{bmatrix} \widehat{Q}^t \\ \widetilde{Q}^t \end{bmatrix} \begin{bmatrix} b \end{bmatrix} = \begin{bmatrix} \widehat{R}x \\ 0_{r \times 1} \end{bmatrix} - \begin{bmatrix} \widehat{Q}^t b \\ \widetilde{Q}^t b \end{bmatrix}$$

Note que podemos escrever a igualdade acima da seguinte forma:

$$Rx - Q^t b = \begin{bmatrix} \widehat{R}x \\ 0_{r \times 1} \end{bmatrix} - \left(\begin{bmatrix} \widehat{Q}^t b \\ 0_{r \times 1} \end{bmatrix} + \begin{bmatrix} 0_{n \times 1} \\ \widetilde{Q}^t b \end{bmatrix} \right) = \begin{bmatrix} \widehat{R}x - \widehat{Q}^t b \\ 0_{r \times 1} \end{bmatrix} + \begin{bmatrix} 0_{n \times 1} \\ \widetilde{Q}^t b \end{bmatrix},$$

onde os elementos do membro direito são ortogonais em $I\!R^m$.

Desse modo, pela Fórmula de Pitágoras, podemos escrever $\| Rx - Q^t b \|_2$ da forma:

$$\| Rx - Q^t b \|_2^2 = \left\| \begin{matrix} \widehat{R}x - \widehat{Q}^t b \\ 0_{r \times 1} \end{matrix} \right\|_2^2 + \left\| \begin{matrix} 0_{n \times 1} \\ \widetilde{Q}^t b \end{matrix} \right\|_2^2,$$

que de forma simplificado, obtemos

$$\| Rx - Q^t b \|_2^2 = \| \widehat{R}x - \widehat{Q}^t b \|_2^2 + \| \widetilde{Q}^t b \|_2^2 .$$

Finalmente, voltando ao problema original, temos que

$$\begin{aligned} \| Ax^* - b \|_2 &= \min\{ \| Ax - b \|_2 \quad ; \quad x \in I\!R^n \} \\ &= \min\{ \| \widehat{R}x - \widehat{Q}^t b \|_2 + \| \widetilde{Q}^t b \|_2 \quad ; \quad x \in I\!R^n \} \\ &= \min\{ \| \widehat{R}x - \widehat{Q}^t b \|_2 \quad ; \quad x \in I\!R^n \} + \| \widetilde{Q}^t b \|_2 \end{aligned} \tag{8.173}$$

Podemos observar facilmente que o ponto de mínimo $x^* \in I\!R^n$ é tal que

$$\| \widehat{R}x^* - \widehat{Q}^t b \|_2 = 0 \iff \widehat{R}x^* - \widehat{Q}^t b = 0_{I\!R^n} \iff \widehat{R}x^* = \widehat{Q}^t b .$$

Portanto, a solução de quadrados mínimos $x^* \in I\!R^n$ é a única solução do sistema linear triangular superior

$$\widehat{R}x = \widehat{Q}^t b \,,$$

sabendo que $\widehat{R}$ é uma matriz invertível, pois $posto(A) = n$.

É importante observar que o elemento $r^* = b - Ax^*$ é a projeção ortogonal do elemento b sobre o subespaço $\mathcal{R}(A)^\perp = \mathcal{N}(A^t)$, e o elemento $z^* = Ax^*$ é a projeção ortogonal do elemento b no subespaço $\mathcal{R}(A)$, que é a melhor aproximação do elemento b no subespaço $\mathcal{R}(A)$. Desse modo, dizemos que o elemento r^* é o **vetor de resíduo** da melhor aproximação, e que $\| r^* \|_2$ é o **resíduo** dessa aproximação. Sendo assim, da equação (8.173), podemos concluir que

$$\| b - Ax^* \|_2 = \| r^* \|_2 = \| \widetilde{Q}^t b \|_2 \,,$$

lembrando a segunda igualdade é válida somente em norma, não significando a igualdade dos elementos envolvidos.

Exercícios

Exercício 8.117 *Sejam a matriz* $A \in \mathbb{M}_{4\times 2}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}.$$

(a) Determine uma base ortogonal para o subespaço $\mathcal{R}(A)$.

(b) Determine a projeção ortogonal do elemento b *sobre o subespaço* $\mathcal{N}(A^t)$.

Exercício 8.118 *Sejam* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz invertível e* $b \in \mathbb{R}^n$. *Descreva como podemos obter a solução do sistema linear* $Ax = b$ *através da fatoração* $A = QR$. *Mostre que* $\mathcal{K}_2(R) = \mathcal{K}_2(A)$. *O que podemos concluir ?*

Exercício 8.119 *Encontre a fatoração* QR *da matriz*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

através do Processo de Ortogonalização de Gram–Schmidt.

Exercício 8.120 *Considere a matriz* $A \in \mathbb{M}_{3\times 5}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^5$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 2 & -1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Encontre a projeção ortogonal do elemento b *no subespaço* $\mathcal{N}(A)$, *utilizando o Método de Gram–Schmidt Modificado. Faça a implementação computacional em uma linguagem de sua preferência. Apresente uma pequena introdução teórica justificando a resolução do problema.*

Exercício 8.121 *Considere a matriz* $A \in I\!M_{4\times 2}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

(a) Determine uma base ortonormal para o subespaço $\mathcal{R}(A)$.

(b) Determine uma base ortonormal para o subespaço $\mathcal{N}(A^t)$.

(c) Determine a fatoração $A = QR$, *onde* $Q \in I\!M_4(I\!R)$ *e* $R \in I\!M_{4\times 2}(I\!R)$, *descritas da seguinte forma:*

$$Q = \begin{bmatrix} \widehat{Q} & \widetilde{Q} \end{bmatrix} \qquad e \qquad R = \begin{bmatrix} \widehat{R} \\ 0_2 \end{bmatrix},$$

onde $\widehat{Q}\,,\, \widetilde{Q} \in I\!M_{4\times 2}(I\!R)$ *são matrizes ortogonais,* $\widehat{R} \in I\!M_2(I\!R)$ *é uma matriz triangular superior com os elementos da diagonal principal positivos e* $0_2 \in I\!M_2(I\!R)$ *é a matriz nula.*

(d) Dê uma interpretação para a matriz $\widehat{R}$.

Exercício 8.122 *Considere a matriz* $A \in I\!M_{4\times 3}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Determine a fatoração $A = QR$, *através do Processo de Gram–Schmidt, onde a matriz* $Q \in I\!M_{4\times 3}(I\!R)$ *é uma matriz ortogonal e* $R \in I\!M_3(I\!R)$ *é uma matriz triangular superior com os elementos da diagonal principal positivos.*

Exercício 8.123 *Considere a matriz* $A \in I\!M_{2\times 4}(I\!R)$ *e o elemento* $b \in I\!R^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix}$$

Determine a projeção ortogonal do elemento b *sobre o subespaço* $\mathcal{N}(A)$ *e a respectiva matriz de projeção ortogonal, utilizando a fatoração* $A^t = \widehat{Q}\widehat{R}$.

Exercício 8.124 *Considere a matriz* $A \in \mathbb{M}_{2\times4}(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Determine a matriz de Reflexão sobre $\mathcal{R}(A^t)$, *utilizando a fatoração* $A^t = \widehat{Q}\widehat{R}$.

Exercício 8.125 *Considere a matriz* $A \in \mathbb{M}_{4\times3}(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 2 & -1 \\ 2 & 0 \end{bmatrix}.$$

(a) Determine a decomposição $A = QR$, *através do Processo de Gram–Schmidt, onde* $Q \in \mathbb{M}_{4\times3}(\mathbb{R})$ *é uma matriz ortogonal e* $R \in \mathbb{M}_3(\mathbb{R})$ *é uma matriz triangular superior com os elementos da diagonal principal positivos.*

(b) Determine a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$.

(c) Determine a pseudo–inversa da matriz A.

(d) Determine a projeção ortogonal do elemento $b \in \mathbb{R}^4$ *dado por:*

$$b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

no subespaço $\mathcal{R}(A)$.

Exercício 8.126 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, *com* $m > n$ *e* $posto(A) = n$, *e a fatoração* $A = QR$, *com* $Q \in \mathbb{M}_{m\times n}(\mathbb{R})$ *uma matriz ortogonal e* $R \in \mathbb{M}_n(\mathbb{R})$ *uma matriz triangular superior com os elementos da diagonal principal positivos. Pede–se:*

(a) Mostre que $\mathcal{R}(A) = \mathcal{R}(Q)$.

(b) Mostre que $\mathcal{N}(A^t) = \mathcal{N}(Q^t)$.

(c) Mostre que $\mathcal{R}(A^t) = \mathcal{R}(Q^t) = \mathbb{R}^n$.

(d) Mostre que $\mathcal{N}(A) = \mathcal{N}(Q) = \{ 0_{\mathbb{R}^n} \}$.

Exercício 8.127 *Considere a matriz* $A \in I\!M_{4\times 3}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

(a) Determine a decomposição $A = QR$, *através do Processo de Gram–Schmidt, onde* $Q \in I\!M_{4\times 3}(I\!R)$ *é uma matriz ortogonal e* $R \in I\!M_{3}(I\!R)$ *é uma matriz triangular superior com os elementos da diagonal principal positivos.*

(b) Determine a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$.

(c) Determine a pseudo–inversa da matriz A.

(d) Determine a projeção ortogonal do elemento $b \in I\!R^4$ *dado por:*

$$b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

no subespaço $\mathcal{R}(A)$.

Exercício 8.128 *Sejam* $A \in I\!M_{m\times n}(I\!R)$, *com* $m > n$ *e posto*$(A) = n$, $b \in I\!R^m$ *e a fatoração* $A = QR$, *com* $Q \in I\!M_{m\times n}(I\!R)$ *uma matriz ortogonal e* $R \in I\!M_{n}(I\!R)$ *uma matriz triangular superior com os elementos da diagonal principal positivos. Pede–se:*

(a) Mostre que a solução de quadrados mínimos $x^* \in I\!R^n$, *para o sistema linear* $Ax = b$, *é a única solução do sistema triangular superior* $Rx = Q^t b$.

(b) Mostre que $A^\dagger = R^{-1}Q^t$ *é a pseudo–inversa da matriz* A.

(c) Mostre que $P = QQ^t$ *é a matriz de projeção ortogonal sobre* $\mathcal{R}(A)$.

Exercício 8.129 *Considere a matriz* $A \in I\!M_{4\times 3}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 2 & -1 \\ 2 & 0 \end{bmatrix}.$$

(a) Determine a decomposição $A = QR$*, através do Processo de Gram–Schmidt, onde* $Q \in I\!M_{4\times 3}(I\!R)$ *é uma matriz ortogonal e* $R \in I\!M_3(I\!R)$ *é uma matriz triangular superior com os elementos da diagonal principal positivos.*

(b) Determine a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$*.*

(c) Determine a pseudo–inversa da matriz A*.*

(d) Determine a projeção ortogonal do elemento $b \in I\!R^4$ *dado por:*

$$b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

no subespaço $\mathcal{R}(A)$*.*

Exercício 8.130 *Considere a matriz* $A \in I\!M_{4\times 3}(I\!R)$ *dada por:*

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

(a) Determine a decomposição $A = QR$*, através do Processo de Gram–Schmidt, onde* $Q \in I\!M_{4\times 3}(I\!R)$ *é uma matriz ortogonal e* $R \in I\!M_3(I\!R)$ *é uma matriz triangular superior com os elementos da diagonal principal positivos.*

(b) Determine a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$*.*

(c) Determine a pseudo–inversa da matriz A*.*

(d) Determine a projeção ortogonal do elemento $b \in I\!R^4$ *dado por:*

$$b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

no subespaço $\mathcal{R}(A)$.

Exercício 8.131 *Sejam* $A \in I\!M_n(I\!R)$ *não singular e* $b \in I\!R^n$. *Descreva como obter a solução do sistema linear* $Ax = b$ *através da fatoração QR da matriz* A. *Escreva um procedimento em Matlab para obter uma solução numérica do sistema linear* $Ax = b$ *através da fatoração QR da matriz* A, *utilizando o Método de Gram–Schmidt. Tome como exemplo a matriz de Hilbert* $A = [a_{ij}]$, *onde*

$$a_{ij} = \frac{1}{i + j - 1},$$

para $i, j = 1, \cdots, n$, *escolhendo o vetor* b *de modo que a solução exata* $x^* \in I\!R^n$ *seja dada por:*

$$x^* = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix},$$

para vários valores de n, *calculando em cada caso o erro relativo*

$$\frac{\| x^* - \widehat{x} \|_2}{\| x^* \|_2},$$

onde $\widehat{x}$ *é a solução numérica.*

Exercício 8.132 *Considerando o procedimento em Matlab desenvolvido no Exercício 8.131, faça a implementação de um procedimento eficiente para calcular a matriz inversa* A^{-1} *de uma matriz* $A \in I\!M_n(I\!R)$ *invertível.*

Exercício 8.133 *Considere a matriz* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

(a) Determine uma base ortonormal para o subespaço $\mathcal{R}(A)$.

(b) Determine uma base ortonormal para o subespaço $\mathcal{N}(A^t)$.

(c) Determine a fatoração $A = QR$, *onde* $Q \in \mathbb{M}_4(\mathbb{R})$ *é uma matriz ortogonal e* $R \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *é uma matriz triangular superior.*

Exercício 8.134 *Considere a matriz* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *dada por:*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

(a) Determine uma base ortonormal para o subespaço $\mathcal{R}(A)$.

(b) Determine uma base ortonormal para o subespaço $\mathcal{N}(A^t)$.

(c) Determine a fatoração $A = QR$, *onde* $Q \in \mathbb{M}_4(\mathbb{R})$ *é uma matriz ortogonal e* $R \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *é uma matriz triangular superior.*

Exercício 8.135 *Seja* $Q \in \mathbb{M}_4(\mathbb{R})$ *a matriz de reflexão sobre o subespaço* $\mathcal{R}(A)$, *onde* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *é dada por:*

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

(a) Mostre que os autovalores da matriz Q *são* $\lambda_1 = 1$ *e* $\lambda_2 = -1$.

(b) Mostre que $E_{\lambda_1} = \mathcal{R}(A)$ *e que* $E_{\lambda_2} = \mathcal{N}(A^t)$, *que são os subespaços associados aos autovalores* λ_1 *e* λ_2, *respectivamente.*

(c) Determine uma base ortonormal para $\mathbb{R}^4$ *formada por autovetores de* Q.

8.14 Modelos de Regressão Linear

Nesta seção vamos apresentar vários modelos de regressão linear, como aplicação da solução de quadrados mínimos para sistema linear sobredeterminado. O objetivo principal é estudar como podemos relacionar uma variável de observação y, denominada variável resposta, com outras variáveis $x_1, \cdots, x_n$, denominadas variáveis regressoras. Além disso, desejamos encontrar o melhor modelo, denominado Modelo de Regressão Linear, e que tenha uma expressão funcional com menos complexidade possível. Para um estudo mais detalhado sobre o tema podemos consultar a referência [8].

Uma vez escolhido um determinado modelo, apresentamos como obter o ajuste desse modelo a um conjunto de observações da variável y, através do Método dos Quadrados Mínimos. Finalmente, vamos construir uma ferramenta para analisar a qualidade desse ajuste, procurando fazer uma conexão com as propriedades dos subespaços fundamentais e a solução de quadrados mínimos para um sistema linear sobredeterminado.

Modelo de Regressão Linear Simples

Inicialmente estamos interessados em analisar o comportamento de uma variável resposta y com relação a uma única variável regressora x. Assim, queremos determinar qual é a relação entre as variáveis y e x. Caso seja possível encontrar essa relação, desejamos formular o melhor modelo de regressão linear simples.

Um modelo de regressão linear, muito utilizado em várias situações de interesse prático, relacionando uma variável resposta y a uma única variável regressora x é dado por:

$$y(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_n x^n , \tag{8.174}$$

isto é, um Modelo de Regressão Linear Polinomial Determinístico, pois não consideramos a possibilidade de erros aleatórios nas observações, ou os erros nas medidas são muito pequenos.

De uma maneira geral, escolhemos um conjunto de funções linearmente independente, definidas no intervalo da variável regressora e que sejam pelo menos contínuas,

$$S = \{ \varphi_i , \cdots , \varphi_n \}$$

e propomos o seguinte modelo de regressão linear simples

$$y(x) = \beta_1 \varphi_1(x) + \beta_2 \varphi_2(x) + \cdots + \beta_n \varphi_n(x) . \tag{8.175}$$

É importante observar que o termo linear está relacionado aos parâmetros do modelo, e não à relação funcional entre a variável resposta e a variável regressora.

Finalmente, para ajustar o modelo de regressão linear, isto é, determinar os parâmetros do modelo, necessitamos de um conjunto de m observações, ou medidas, nas duas variáveis x e y, que vamos denotar por:

$$(x_1\,,\,y_1) \quad , \quad (x_2\,,\,y_2) \quad , \quad \cdots \quad , \quad (x_m\,,\,y_m)\,.$$

Desse modo, o gráfico do modelo de regressão linear, que ajusta os dados observados, é denominado **curva de regressão**.

Assim, para cada um dos pares de medidas, obtemos uma equação

$$y_i \;=\; \beta_1\varphi_1(x_i) \;+\; \beta_2\varphi_2(x_i) \;+\; \cdots \;+\; \beta_n\varphi_n(x_i)\,, \tag{8.176}$$

para $i \,=\, 1,\, 2,\, \cdots,\, m$. É importante ressaltar que para uma melhor estimativa dos parâmetros do modelo devemos ter $m \gg n$, isto é, o número de medidas dever ser muito maior que o número de parâmetros a serem estimados.

Por simplicidade, representamos esse conjunto de equações na forma matricial

$$AX \;=\; Y\,,$$

onde

$$A \;=\; \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_1(x_i) & \varphi_2(x_i) & \cdots & \varphi_n(x_i) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{bmatrix} \quad , \quad X \;=\; \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_n \end{bmatrix} \quad \text{e} \quad Y \;=\; \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{bmatrix},$$

que é um **sistema linear sobredeterminado**.

Vamos obter a estimativa dos parâmetros do modelo de regressão através da **solução de quadrados mínimos** para o sistema linear sobredeterminado $AX \,=\, Y$, que indicamos pelo elemento $\widehat{X} \,\in\, I\!\!R^n$, que é a solução do **problema de minimização**

$$\|\, A\widehat{X} \,-\, Y \,\|_2 \;=\; \min\{\; \|\, AX \,-\, Y \,\|_2 \quad ; \quad X \in I\!\!R^n \;\}\,, \tag{8.177}$$

que é equivalente ao **sistema normal**

$$A^t A X \;=\; A^t Y \,. \tag{8.178}$$

Encontrada a solução de quadrados mínimos, que indicamos por:

$$\widehat{X} \;=\; \begin{bmatrix} \widehat{\beta}_1 \\ \vdots \\ \widehat{\beta}_i \\ \vdots \\ \widehat{\beta}_n \end{bmatrix} \,,$$

obtemos o modelo que melhor se ajusta aos dados observados, que é dado por:

$$\widehat{y}(x) \;=\; \widehat{\beta}_1 \varphi_1(x) \;+\; \widehat{\beta}_2 \varphi_2(x) \;+\; \cdots \;+\; \widehat{\beta}_n \varphi_n(x) \,. \tag{8.179}$$

Podemos determinar a solução de quadrados mínimos através da Fatoração de Cholesky, para obter uma solução numérica do sistema normal, ou através da Fatoração QR, para obter uma solução numérica do problema de minimização.

Considerando um conjunto de m observações,

$$(x_1\,,\,y_1) \quad , \quad (x_2\,,\,y_2) \quad , \quad \cdots \quad , \quad (x_m\,,\,y_m)\,,$$

com medidas distintas na variável regressora, isto é, $x_1, \cdots, x_i, \cdots, x_m$ um conjunto de pontos distintos, podemos mostrar facilmente que $dim(\,\mathcal{R}(A)\,) \;=\; posto(A) \;=\; n$, onde A é a matriz do sistema linear sobredeterminado associado ao problema de regressão linear, que é dada por:

$$A \;=\; \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_1(x_i) & \varphi_2(x_i) & \cdots & \varphi_n(x_i) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{bmatrix} \,,$$

desde que o conjunto $S \;=\; \{\,\varphi_i\,,\,\cdots\,,\,\varphi_n\,\}$ seja linearmente independente, de acordo com Definição 3.5.4.

Modelo de Regressão Linear Múltipla

Nesta seção nosso principal objetivo é estudar um **Modelo de Regressão Linear Múltipla** através da análise de uma possível relação do **peso** (kg) dos indivíduos de uma determinada população, denotamos essa variável por P, em função da **altura** (m), indicamos essa variável por x, do fato do indivíduo ser ou não **fumante**, indicamos essa variável por y, que assume os valores $y = 0$ para não fumante e $y = 1$ para fumante, e do **sexo**, indicamos essa variável por z, que assume os valores $z = 0$ para o sexo masculino e $z = 1$ para o sexo feminino.

A variável P é denominada **variável dependente** ou **variável resposta** e as variáveis x, y e z são denominadas **variáveis regressoras**. Desse modo, a variável resposta é aquela que estamos tentando explicar em função das variáveis regressoras. Portanto, queremos relacionar a variável resposta P em termos das variáveis regressoras, através do modelo

$$P(x,y,z) \;=\; \beta_1 \,+\, \beta_2 x \,+\, \beta_3 y \,+\, \beta_4 z \,+\, \beta_5 xy \,+\, \beta_6 xz \,, \qquad (8.180)$$

que é chamado um *Modelo de Regressão Linear Múltipla*. É importante observar que o termo **linear** faz referência à relação entre a variável resposta e os parâmetros do modelo, que neste caso são $\beta_1\,, \cdots, \beta_6$, e não à relação entre a variável resposta e as variáveis regressoras.

Finalmente, para determinar os parâmetros do modelo de regressão linear múltipla, vamos coletar as informações necessárias entre os alunos matriculados na disciplina MS 512, no segundo semestre de 2007. Desse modo, obtemos um conjunto de m equações lineares, onde m é o número de alunos que responderam nosso questionário, com n incógnitas, neste caso temos $n = 6$, que é o número de parâmetros do modelo, dadas por:

$$P_i \;=\; \beta_1 \,+\, \beta_2 x_i \,+\, \beta_3 y_i \,+\, \beta_4 z_i \,+\, \beta_5 x_i y_i \,+\, \beta_6 x_i z_i \,, \qquad (8.181)$$

para $i = 1, 2, \cdots, m$, onde indicamos por P_i os valores observados em nossa população para a variável resposta.

Vamos escrever o sistema de equações lineares (8.181) na forma matricial

$$AX \;=\; Y \;,$$

onde

$$A = \begin{bmatrix} 1 & x_1 & y_1 & z_1 & x_1y_1 & x_1z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_i & y_i & z_i & x_iy_i & x_iz_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & y_m & z_m & x_my_m & x_mz_m \end{bmatrix} \quad , \quad X = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_6 \end{bmatrix} \quad \text{e} \quad Y = \begin{bmatrix} P_1 \\ \vdots \\ P_i \\ \vdots \\ P_m \end{bmatrix},$$

que é um **sistema linear sobredeterminado**, tendo em vista que possui mais equações do que incógnitas, isto é, $m \gg n$, que é uma situação ideal para análise de regressão.

Obtemos a estimativa dos parâmetros do modelo de regressão através da **solução de quadrados mínimos** para o sistema linear sobredeterminado $AX = Y$, que indicamos pelo elemento $\widehat{X} \in I\!R^6$, que é a solução do **problema de minimização**

$$\| A\widehat{X} - Y \|_2 = \min\{ \| AX - Y \|_2 \quad ; \quad X \in I\!R^6 \}, \tag{8.182}$$

que é equivalente ao **sistema normal**

$$A^tAX = A^tY. \tag{8.183}$$

Encontrada a solução de quadrados mínimos, que vamos indicar por:

$$\widehat{X} = \begin{bmatrix} \widehat{\beta}_1 \\ \vdots \\ \widehat{\beta}_6 \end{bmatrix},$$

encontramos o modelo que melhor se ajusta aos dados que coletamos na nossa população, que é dado por:

$$\widehat{P}(x,y,z) = \widehat{\beta}_1 + \widehat{\beta}_2x + \widehat{\beta}_3y + \widehat{\beta}_4z + \widehat{\beta}_5xy + \widehat{\beta}_6xz. \tag{8.184}$$

A seguir apresentamos uma pequena introdução às grandezas estatísticas utilizadas na análise do modelo proposto, verificando a qualidade do ajuste dos dados observados, com a finalidade de verificar a qualidade de previsões que podemos realizar com o modelo. Além disso, mostramos algumas propriedades envolvidas no processo de regressão linear.

Análise do Ajuste do Modelo de Regressão

Essa avaliação é feita através do **Coeficiente de Determinação do Modelo**

$$R^2 = \frac{SQR}{SQT}, \tag{8.185}$$

onde SQR é a **Soma de Quadrados da Regressão**

$$SQR = \sum_{j=1}^{m} (\widehat{P}_j - \overline{p})^2 , \tag{8.186}$$

e SQT é a **Soma de Quadrados Total**

$$SQT = \sum_{j=1}^{m} (P_j - \overline{p})^2 . \tag{8.187}$$

Por simplicidade, estamos indicando $\widehat{P}_j = \widehat{P}(x_j, y_j, z_j)$ os valores estimados pelo modelo, e $\overline{p}$ é a **média** das observações da variável resposta P, isto é,

$$\overline{p} = \frac{\displaystyle\sum_{j=1}^{m} P_j}{m} .$$

Temos também uma outra grandeza importante para a verificação da qualidade do ajuste, que é a **Soma de Quadrados dos Resíduos**

$$SQE = \sum_{j=1}^{m} (P_j - \widehat{P}_j)^2 . \tag{8.188}$$

A grandeza SQR fornece a variabilidade dos valores estimados em torno da média $\overline{p}$, e a grandeza SQT fornece a variabilidade das observações da variável resposta em torno da média $\overline{p}$. A grandeza SQE fornece a variabilidade do resíduo proveniente do ajuste do modelo.

Através das interpretações geométricas relacionadas com a solução de quadrados mínimos para um sistema linear sobredeterminado, vamos mostrar que

$$SQT = SQR + SQE . \tag{8.189}$$

Observado as grandezas definidas acima, concluímos que $0 \leq R^2 \leq 1$. Além disso, melhor será o ajuste do modelo quando $R^2 \approx 1$, tendo em vista que teremos $SQE \approx 0$.

Desse modo, o coeficiente de determinação do modelo R^2 representa a **porcentagem** da variabilidade total dos valores observados explicada pelo modelo.

Problema 8.14.1 *Podemos observar facilmente que a nossa população está dividida em quatro grupos distintos. Desse modo, podemos determinar um modelo que melhor se ajusta aos dados da Tabela 8.1 para cada um dos grupos e comparar os resultados com o modelo descrito em (8.184). Faça a identificação de cada um dos modelos e apresente uma conclusão.*

Uma primeira análise estatística de um conjunto de observações é o seu **vetor de média**. Inicialmente, apresentamos a interpretação geométrica para o vetor de média fazendo uso de projeções ortogonais e das conexões entre os subespaços fundamentais. Para isso, considere o espaço vetorial real $I\!R^m$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Sejam $Y \in I\!R^m$ o **vetor de observações** da variável resposta P e o **vetor unidade** $U \in I\!R^m$ dados por:

$$Y = \begin{bmatrix} P_1 \\ \vdots \\ P_m \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Sabemos que a projeção ortogonal do elemento Y sobre o subespaço gerado pelo elemento U, que vamos denotar por $\overline{Y}$, é dada por:

$$\overline{Y} = \frac{\langle U , Y \rangle}{\langle U , U \rangle} U = \frac{\displaystyle\sum_{j=1}^{m} y_j}{m} U = \begin{bmatrix} \overline{p} \\ \vdots \\ \overline{p} \end{bmatrix},$$

onde $\overline{p}$ é a média do vetor de observações, isto é,

$$\overline{p} = \frac{\displaystyle\sum_{j=1}^{m} P_j}{m}.$$

Assim, concluímos que $\overline{Y}$ é o **vetor de média** do vetor de observações, isto é,

$$\overline{Y} = \begin{bmatrix} \overline{p} \\ \vdots \\ \overline{p} \end{bmatrix}.$$

Podemos verificar facilmente que o vetor $\overline{Y}$ pertence ao subespaço $\mathcal{R}(A)$, tendo em vista que o subespaço $S = [U]$ está contido em $\mathcal{R}(A)$.

Além disso, o complemento ortogonal $S^{\perp}$ do subespaço S é um hiperplano em $\mathbb{R}^m$, que tem dimensão $(m-1)$, isto é,

$$S^{\perp} = \{ Z \in \mathbb{R}^m \ / \ \langle U , Z \rangle = 0 \} .$$

Assim, o elemento $(Y - \overline{Y})$ é a projeção ortogonal do elemento Y no subespaço $S^{\perp}$.

Considerando o fato que $\overline{Y} \in \mathcal{R}(A)$ e as interpretações geométricas relacionadas com a solução de quadrados mínimos para um sistema linear sobredeterminado, vamos mostrar a relação (8.189), e assim provamos também que o coeficiente de determinação do modelo satisfaz

$$0 \leq R^2 \leq 1 . \tag{8.190}$$

Chamando $\widehat{X}$ o **vetor de parâmetros**, que é a solução de quadrados mínimos, isto é,

$$\widehat{X} = \begin{bmatrix} \widehat{\beta}_1 \\ \vdots \\ \widehat{\beta}_n \end{bmatrix} . \tag{8.191}$$

Desse modo, o vetor $\widehat{Y} = A\widehat{X}$ que é a projeção ortogonal do vetor de observações Y no subespaço $\mathcal{R}(A)$, isto é, $\widehat{Y}$ é a **melhor aproximação** do vetor Y no subespaço $\mathcal{R}(A)$, é dado por:

$$\widehat{Y} = \begin{bmatrix} \widehat{P}_1 \\ \vdots \\ \widehat{P}_m \end{bmatrix} .$$

O elemento $(Y - \widehat{Y})$ é a projeção ortogonal do elemento Y no subespaço $\mathcal{N}(A^t)$.

Sendo assim, podemos reescrever as grandezas SQR, SQT e SQE na forma:

$$\begin{aligned} SQR &= \| \widehat{Y} - \overline{Y} \|_2^2 \\ SQT &= \| Y - \overline{Y} \|_2^2 \\ SQE &= \| Y - \widehat{Y} \|_2^2 \end{aligned} \tag{8.192}$$

respectivamente.

Finalmente, pelo Teorema de Pitágoras, obtemos

$$\begin{aligned} \| Y - \overline{Y} \|_2^2 &= \| (Y - \widehat{Y}) + (\widehat{Y} - \overline{Y}) \|_2^2 \\ &= \| Y - \widehat{Y} \|_2^2 + \| \widehat{Y} - \overline{Y} \|_2^2 , \end{aligned} \tag{8.193}$$

desde que o elemento $(Y - \widehat{Y}) \in \mathcal{N}(A^t)$ e o elemento $(\widehat{Y} - \overline{Y}) \in \mathcal{R}(A)$. Assim, provamos a relação (8.189).

Portanto, da equação (8.193), temos a seguinte relação

$$SQT = \| Y - \overline{Y} \|_2^2 \geq SQR = \| \widehat{Y} - \overline{Y} \|_2^2 . \tag{8.194}$$

Logo, mostramos que $0 \leq R^2 \leq 1$.

Como a matriz $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, com $m > n$, tem posto completo, isto é, $posto(A) = n$, sabemos que a matriz A^tA é positiva–definida. Desse modo, temos que A^tA é uma matriz invertível.

Portanto, o vetor de parâmetros $\widehat{X}$ tem a seguinte **caracterização**:

$$\widehat{X} = (A^tA)^{-1}A^tY , \tag{8.195}$$

que é a solução do sistema normal (8.183).

Desse modo, o elemento $\widehat{Y} = A\widehat{X}$, que é a projeção ortogonal do vetor de observações Y no subespaço $\mathcal{R}(A)$, é representado na seguinte forma:

$$\widehat{Y} = A(A^tA)^{-1}A^tY = PY , \tag{8.196}$$

onde $P = A(A^tA)^{-1}A^t$ é a **matriz de projeção ortogonal** sobre o subespaço $\mathcal{R}(A)$. Sendo assim, sabemos que

(a) $\mathcal{R}(P) = \mathcal{R}(A)$.

(b) $P^t = P$ (P é uma matriz simétrica).

(c) $P^2 = P$ (P é uma matriz idempotente).

Finalmente, vamos mostrar que

$$\sum_{j=1}^{m} P_j = \sum_{j=1}^{m} \widehat{P}_j \iff \frac{\displaystyle\sum_{j=1}^{m} P_j}{m} = \frac{\displaystyle\sum_{j=1}^{m} \widehat{P}_j}{m} . \tag{8.197}$$

De fato, podemos observar que

$$Y^t U = \langle U , Y \rangle = \sum_{j=1}^{m} P_j \quad \text{e} \quad \widehat{Y}^t U = \langle U , \widehat{Y} \rangle = \sum_{j=1}^{m} \widehat{P}_j .$$

Como P é a matriz de projeção ortogonal sobre o subespaço $\mathcal{R}(A)$, obtemos

$$\widehat{Y}^t U = (PY)^t U = Y^t P^t U = Y^t (PU) = Y^t U , \tag{8.198}$$

desde que $PU = U$, pois o vetor unidade $U \in \mathcal{R}(A)$, o que prova do igualdade (8.197).

Da relação (8.197), podemos fazer a interpretação geométrica: *a projeção ortogonal do vetor $\widehat{Y}$ sobre o subespaço $S = [U] \subset \mathcal{R}(A)$ é igual ao vetor $\overline{Y}$, que é a projeção ortogonal do vetor de observações Y sobre o subespaço S.*

É importante ressaltar que utilizamos fortemente o fato do subespaço $S = [U]$ estar contido no subespaço $\mathcal{R}(A)$, isto é, o vetor unidade U é uma das colunas da matriz A. Esse fato significa que o modelo de regressão linear possui o **intercepto**, que é o parâmetro β_1, veja o modelo de regressão linear descrito em (8.180).

Problema 8.14.2 *Determinar a estimativa dos parâmetros do modelo descrito em (8.180) utilizando os dados da nossa população, apresentados na Tabela 8.1, através da solução de quadrados mínimos para o sistema linear sobredeterminado*

$$AX = Y , \tag{8.199}$$

obtida das seguintes maneiras:

- *Através da Fatoração QR da matriz A, obtemos a solução para o problema de minimização (8.182). Num primeiro momento, vamos determinar a Fatoração QR da matriz A através do* ***Método de Gram–Schmidt Modificado****, que é proveniente do* ***Processo de Ortogonalização de Gram–Schmidt****.*

- *Através da* ***Fatoração de Cholesky*** *da matriz $C = A^t A$, obtemos a solução para o sistema normal (8.183).*

Comparar com os resultados obtidos no Problema 8.14.1.

Tabela 8.1: Tabela dos dados observados em nossa população

Peso (kg)	Altura (m)	Fumante (0) Não (1) Sim	Sexo (0) M (1) F
68.6	1.85	0	0
64.1	1.75	0	1
72.0	1.88	0	0
70.2	1.79	1	0
50.3	1.53	0	1
48.5	1.66	0	1
62.0	1.72	0	1
69.0	1.77	0	0
50.0	1.63	0	1
78.5	1.78	0	0
55.0	1.70	0	0
66.0	1.71	0	0
60.0	1.71	0	1
60.0	1.71	0	1
67.0	1.70	0	1
73.0	1.74	0	0
74.0	1.80	0	0
56.0	1.59	0	1
62.0	1.64	0	1
54.0	1.55	0	1
71.0	1.60	1	1
53.0	1.61	0	1
70.0	1.61	0	1
60.0	1.70	0	1
67.0	1.68	1	1
60.0	1.68	0	0
63.0	1.52	0	1
49.0	1.60	0	1
63.0	1.65	0	1
72.0	1.70	0	0
64.0	1.80	0	0
85.0	1.86	0	0

		Fumante	Sexo
Peso (kg)	Altura (m)	(0) Não (1) Sim	(0) M (1) F
76.0	1.68	0	0
97.0	1.77	0	0
82.0	1.82	0	0
89.0	1.75	0	0
66.0	1.77	0	1
63.0	1.72	0	1
78.0	1.80	0	0
63.0	1.79	0	0
83.0	1.89	0	0
53.0	1.60	0	1
84.0	1.83	0	0
68.0	1.85	0	0
62.0	1.65	0	1
61.0	1.68	0	1
51.0	1.65	0	1
52.0	1.67	1	1
53.0	1.55	0	1
70.0	1.86	0	0
65.0	1.70	0	0
70.0	1.71	0	0
72.0	1.73	0	0
65.0	1.70	0	0
85.0	1.80	1	0
80.0	1.80	0	0
97.0	1.73	0	0
96.0	1.65	0	0
57.0	1.62	0	1
125.0	1.89	0	0
76.0	1.79	0	0
65.0	1.82	0	0
51.0	1.57	0	1
60.0	1.50	1	1

Exemplo 8.14.1 *Considere um experimento conduzido com a finalidade de analisar a variação do* ***calor específico*** *da glicerina*[1] *($C_3H_5(OH)_3$), indicamos essa variável por Y, em função da temperatura T. Os resultados do experimento estão na Tabela 8.2.*

Tabela 8.2: Calor específico da glicerina

Temperatura $[K]$ T	Calor Específico $\left[\dfrac{kJ}{kg}K\right]$ Y
273.0	2.261
280.0	2.298
290.0	2.367
300.0	2.427
310.0	2.490
320.0	2.564

Nosso objetivo é estudar a relação da variável resposta Y, que representa o calor específico da glicerina, em função da variável regressora T, que é a temperatura, através do seguinte ***Modelo de Regressão Linear Simples***

$$Y(T) \;=\; \beta_1 \;+\; \beta_2 T \,.$$

Apresentar uma análise do modelo proposto, verificando a qualidade do ajuste dos dados observados, através do ***Coeficiente de Determinação do Modelo****, da* ***Soma de Quadrados dos Resíduos*** *e do* ***gráfico de dispersão****. O que podemos concluir?*

Inicialmente vamos encontrar solução de quadrados mínimos para o sistema linear

$$AX \;=\; Y \,,$$

onde

$$A \;=\; \begin{bmatrix} 1 & 273.0 \\ 1 & 280.0 \\ 1 & 290.0 \\ 1 & 300.0 \\ 1 & 310.0 \\ 1 & 320.0 \end{bmatrix} \quad , \quad X \;=\; \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \qquad \text{e} \qquad Y \;=\; \begin{bmatrix} 2.261 \\ 2.298 \\ 2.367 \\ 2.427 \\ 2.490 \\ 2.564 \end{bmatrix} \,.$$

[1] F. P. Incropera e D. P. DeWitt, *Fundamentos de Transferência de Calor e de Massa*, LTC, 1998.

A solução de quadrados mínimos $\widehat{X}$ é a solução do sistema normal $A^tAX = A^tY$, onde

$$A^tA = \begin{bmatrix} 6 & 1773 \\ 1773 & 525529 \end{bmatrix} \qquad \text{e} \qquad A^tY = \begin{bmatrix} 14.4 \\ 4267.6 \end{bmatrix} .$$

Como $posto(A) = 2$, sabemos que a matriz do sistema normal é positiva–definida. Assim, através da Fatoração de Cholesky obtemos o vetor de parâmetros $\widehat{X}$ que é dado por:

$$\widehat{X} = \begin{bmatrix} \widehat{\beta}_1 \\ \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 5.0141876620 \times 10^{-1} \\ \\ 6.4289269051 \times 10^{-3} \end{bmatrix} .$$

Assim, o modelo apresenta um coeficiente de determinação do modelo $R^2 = 0.9986308869$, e uma soma de quadrados dos resíduos $SQE = 9.108823224468478 \times 10^{-5}$. Desse modo, o melhor modelo que ajusta os dados experimentais é dado por:

$$\widehat{Y} = \widehat{\beta}_1 + \widehat{\beta}_2 T .$$

Com os resultados obtidos, concluímos que os dados experimentais são bem representados pelo modelo de regressão linear simples, isto é, a variação do calor específico da glicerina tem um relação linear com a temperatura. Na Figura 8.10, temos o gráfico de dispersão indicando o resultado obtido.

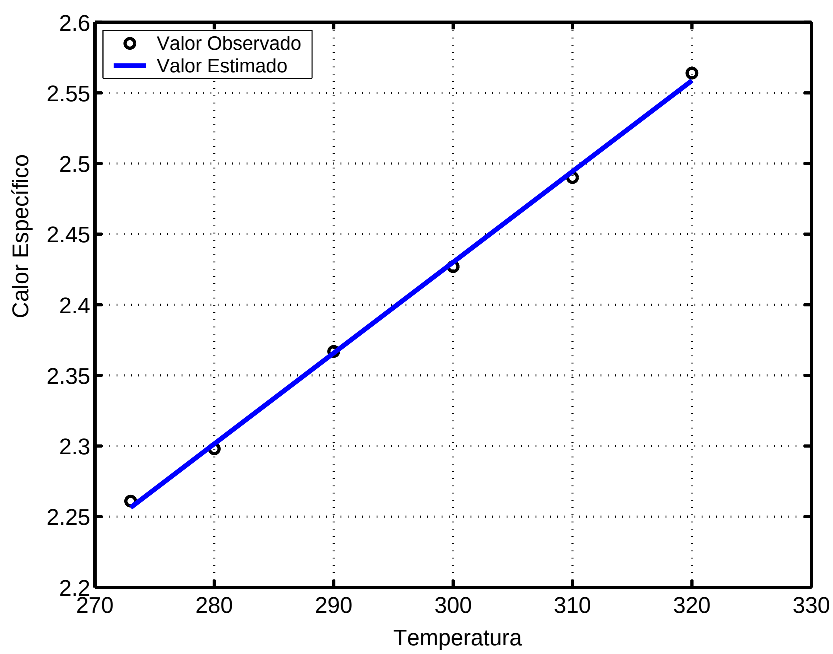

Figura 8.10: Gráfico de dispersão para o Exemplo 8.14.1.

Variância, Covariância e Coeficiente de Correlação

Com os resultados de Álgebra Linear desenvolvidos, vamos apresentar de forma muito básica e como uma aplicação dos conceitos estudados até o momento, algumas medidas estatísticas com as quais estudamos diversas propriedades de variáveis observadas sobre cada indivíduo de uma amostra de uma determinada população. Problemas multivariados surgem em diversas áreas de investigações científicas, tais como, biologia, física, sociologia, ciências médicas, etc, e de forma muito mais natural do que os problemas univariados, quando problemas práticos resultam em coleção de dados na qual mais de uma variável é observada sobre cada um dos indivíduos. Com o objetivo de fazer uma conexão entre os conceitos apresentados no texto, vamos definir e apresentar uma interpretação algébrica das medidas estatísticas como variância amostral, desvio padrão amostral, vetor de média amostral, covariância amostral e coeficiente de correlação amostral. Por simplicidade, e que claramente fica subentendido, omitiremos a palavra amostral nos conceitos que serão apresentados nesta seção. Sempre que possível, e de modo que venha contribuir para uma melhor compreensão dos conceitos, apresentamos uma interpretação geométrica para cada uma dessas medidas.

A forma de como trabalhar com um problema multivariado depende de como vamos estruturar mais de uma variável observada sobre cada indivíduo de uma amostra de uma determinada população. Todo conjunto multivariado de dados $Y_1 , \cdots , Y_j , \cdots , Y_n$, onde cada variável Y_j possui m medidas, pode ser representado através da **matriz de dados**, que denotamos por $Y = [y_{ij}] \in \mathbb{M}_{m\times n}(\mathbb{R})$, definida da seguinte forma:

$$Y = [Y_1 \cdots Y_j \cdots Y_n] = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1j} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2j} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{i1} & y_{i2} & \cdots & y_{ij} & \cdots & y_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mj} & \cdots & y_{mn} \end{bmatrix},$$

onde m é o número de indivíduos observados e n é o número de variáveis sobre cada indivíduo. Desse modo, y_{ij} é a observação da variável aleatória Y_j sobre o i–ésimo indivíduo da amostra.

Assim, uma amostra de m indivíduos sobre os quais foram observadas n variáveis, pode ser identificada com n pontos em $\mathbb{R}^m$, ou espaço dos indivíduos. De modo análogo, podemos identifica–la com m pontos em $\mathbb{R}^n$, ou espaço das variáveis.

Como estamos tratando com uma amostra aleatória de m indivíduos de uma mesma população, as linhas da matriz de dados Y correspondem à observações das variáveis aleatórias $Y_1, \cdots, Y_n$ sobre indivíduos de maneira que as linhas fornecem observações independentes e identicamente distribuídas das variáveis aleatórias.

Assim sendo, quer estejamos fazendo uma análise no espaço dos indivíduos o $\mathbb{R}^m$, ou quer estejamos fazendo uma análise no espaço das variáveis o $\mathbb{R}^n$, estamos representando indivíduos ou variáveis com relação à base canônica de cada um desses espaços vetoriais, respectivamente. De uma maneira mais simples, podemos ver cada variável observada Y_j como um elemento do espaço vetorial real $\mathbb{R}^m$, e a sua respectiva representação com relação à base canônica.

O trabalho com conjuntos de observações com muitas informações apresenta uma certa dificuldade para a extração das propriedades sobre o fenômeno que desejamos estudar. Assim, muitas dessas propriedades que estão contidas neste conjunto de dados, podem ser obtidas pelo cálculo de certas medidas estatísticas conhecidas como *estatística descritiva*, cujos conceitos serão apresentados a seguir, com as devidas conexões com os resultados que foram estudados no texto.

Podemos associar à matriz de dados Y o **vetor de médias amostrais**, ou simplesmente **vetor de médias**, denotado por $\overline{Y}$, que é definido da seguinte forma:

$$\overline{Y} = \begin{bmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_j \\ \vdots \\ \overline{y}_n \end{bmatrix},$$

como extensão da média amostral univariada, onde $\overline{y}_j$ é a média da variável observada $Y_j \in \mathbb{R}^m$ que é dada por:

$$\overline{y}_j = \frac{\displaystyle\sum_{i=1}^{m} y_{ij}}{m} \qquad \text{para} \qquad j = 1, \cdots, n.$$

Quando o vetor de médias amostrais é identificado como um elemento do espaço vetorial $\mathbb{R}^n$, representa o *centro de gravidade* dos pontos amostrais Y_j, $j = 1, \cdots, n$. Assim, dizemos que o vetor de médias $\overline{Y}$ é o *indivíduo médio*, isto é, o representante sumário das observações.

Definição 8.14.1 *Considere* $Y \in I\!R^m$ *uma variável observada. Sejam* $\overline{y}$ *a média e* $\overline{Y}$ *o vetor de média da variável* Y*, respectivamente. A* ***variância*** *da variável* Y*, que denotamos por* $\mathbf{var}(Y)$*, é definida da seguinte forma:*

$$\mathbf{var}(Y) \;=\; \frac{\langle\, Z^* \,,\, Z^* \,\rangle}{m-1}\,,$$

onde $Z^* = Y - \overline{Y}$*, que é denominado* ***vetor de resíduo****.*

Podemos mostrar facilmente que o elemento $Z^* = Y - \overline{Y}$ é a projeção ortogonal do elemento Y no complemento ortogonal do subespaço $S = [U]$, tendo em vista que o vetor de média $\overline{Y}$ é a projeção ortogonal do elemento Y no subespaço S. Além disso, sabemos que $dim(\, S^{\perp}\,) = m - 1$. Desse modo, como $Z^* \in S^{\perp}$, temos que qualquer uma de suas componentes depende das outras $(m-1)$ componentes restantes. Portanto, desse fato surge a motivação da divisão por $(m - 1)$ na definição da variância de uma variável de observações. É importante observar que o conceito de variância está vinculado ao **desvio padrão** da variável Y, que denotamos por $\mathbf{S}_Y$, definido da seguinte forma:

$$\mathbf{S}_Y \;=\; \sqrt{\mathbf{var}(Y)}\;.$$

Uma outra medida muito utilizada em estatística é o conceito de covariância, que é uma medida de associação linear entre duas variáveis observadas.

Definição 8.14.2 *Sejam* $Y_1\,,\, Y_2 \in I\!R^m$ *variável observadas. A* ***covariância*** *entre as variáveis* Y_1 *e* Y_2*, que denotamos por* $\mathbf{cov}(Y_1, Y_2)$*, é definida da seguinte forma:*

$$\mathbf{cov}(Y_1, Y_2) \;=\; \frac{\langle\, Z_1^* \,,\, Z_2^* \,\rangle}{m-1}\,,$$

onde $Z_1^* = Y_1 - \overline{Y}_1$ *e* $Z_2^* = Y_2 - \overline{Y}_2$*.*

Podemos observar que a covariância pode ser vista como um pseudo–produto interno, denominado **produto interno covariante**, que denotamos por $\langle\, \cdot\,,\, \cdot\,\rangle_{\mathbf{cov}}$.

De fato, podemos observar facilmente que os vetores de resíduos

$$Z_1^* = Y_1 - \overline{Y}_1 \qquad \text{e} \qquad Z_2^* = Y_2 - \overline{Y}_2$$

podem ser escritos da forma:

$$Z_j^* = Y_j - \overline{Y}_j = (I - P)Y_j = HY_j \qquad \text{para} \qquad j = 1, 2\,,$$

onde $P \in I\!M_m(I\!R)$ é a matriz de projeção sobre o subespaço $S = [U]$, que é dada por:

$$P \;=\; \frac{UU^t}{U^tU}\,,$$

e $H = I - P$ é a matriz de projeção ortogonal sobre o subespaço $S^{\perp}$.

Assim, a covariância entre as variáveis observadas Y_1 e Y_2 pode ser escrita da forma:

$$\begin{aligned}
\textbf{cov}(Y_1, Y_2) &= \frac{\langle\, Z_1^* \,,\, Z_2^* \,\rangle}{m-1} \\
&= \frac{(\, Z_2^* \,)^t Z_1^*}{m-1} \\
&= \frac{(\, Y_2^t H^t \,)(\, HY_1 \,)}{m-1} \\
&= \frac{Y_2^t(\, H^t H \,)Y_1}{m-1} \\
&= \frac{Y_2^t(\, HY_1 \,)}{m-1} \\
&= \frac{\langle\, HY_1 \,,\, Y_2 \,\rangle}{m-1} \\
&= \langle\, Y_1 \,,\, Y_2 \,\rangle_{\textbf{cov}} \,,
\end{aligned} \tag{8.200}$$

uma vez que H, é uma matriz de projeção ortogonal, satisfaz as propriedades:

$$H^t \,=\, H \qquad \text{e} \qquad H^2 \,=\, H \,,$$

veja a Proposição 8.12.1 e a Definição 8.12.1.

Desse modo, para todo elemento $Y \in S = [U]$, temos que

$$\textbf{cov}(Y, Y) \,=\, \frac{Y^t(\, HY \,)}{m-1} \,=\, \frac{\langle\, HY \,,\, Y \,\rangle}{m-1} \,=\, \langle\, Y \,,\, Y \,\rangle_{\textbf{cov}} \,=\, 0 \,,$$

uma vez que $HY \,=\, 0_{I\!R^m}$.

Portanto, provamos que o produto interno covariante é um pseudo–produto interno em $I\!R^m$ associado à matriz semipositiva–definida dada por:

$$\frac{H}{m-1} \,,$$

onde H é a matriz de projeção ortogonal sobre o subespaço $S^\perp$. É importante observar que estamos considerando o espaço vetorial $I\!R^m$ munido do produto interno usual $\langle \cdot \,,\, \cdot \rangle$, e com a base canônica $\beta = \{\, e_1, \cdots, e_n \,\}$.

Formas Bilineares e Formas Quadráticas

Definição 8.14.3 *Seja* V *um espaço vetorial sobre o corpo* $\mathbb{F}$*. Uma* ***forma bilinear*** *sobre* V *é uma aplicação* $\mathbf{a}(\cdot,\cdot) : V \times V \longrightarrow \mathbb{F}$ *que satisfaz as seguintes propriedades:*

(a) $\mathbf{a}(\alpha u \,+\, v, w) \;=\; \alpha\, \mathbf{a}(u,w) \,+\, \mathbf{a}(v,w)$

(b) $\mathbf{a}(u, \alpha v \,+\, w) \;=\; \alpha\, \mathbf{a}(u,v) \,+\, \mathbf{a}(u,w)$

para todos $u,\, v,\, w \in V$ *e para todo escalar* $\alpha \in \mathbb{F}$*.*

Exemplo 8.14.2 *Sejam* V *um espaço vetorial sobre o corpo* $\mathbb{F}$*,* J_1 *e* J_2 *funcionais lineares sobre* V*, veja Definição 7.2.1. A aplicação* $\mathbf{a}(\cdot,\cdot) : V \times V \longrightarrow \mathbb{F}$ *definida por:*

$$\mathbf{a}(u,v) \;=\; J_1(u)\, J_2(v)$$

para todos $u,\, v \in V$*, é uma forma bilinear sobre* V*.*

Exemplo 8.14.3 *Considere o espaço vetorial real* $\mathbb{M}_{m\times n}(\mathbb{R})$ *e* $A \in \mathbb{M}_m(\mathbb{R})$ *uma matriz fixa, porém arbitrária. A aplicação* $f_A(\cdot,\cdot) : \mathbb{M}_{m\times n}(\mathbb{R}) \times \mathbb{M}_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ *definida da forma:*

$$f_A(X,Y) \;=\; tr(X^t A Y) \qquad \textit{para todos} \qquad X,\, Y \in \mathbb{M}_{m\times n}(\mathbb{R})$$

é uma forma bilinear sobre $\mathbb{M}_{m\times n}(\mathbb{R})$*.*

Teorema 8.14.1 *Considere* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$ *e* $\beta \;=\; \{\, u_1\,,\, \cdots\,,\, u_n \,\}$ *uma base ordenada para* V*. Seja* $\mathbf{a}(\cdot,\cdot)$ *é uma forma bilinear sobre* V*. Então, a matriz de* $\mathbf{a}(\cdot,\cdot)$ *com relação à base ordenada* β *é a matriz* $A \;=\; [a_{ij}]$ *cujos elementos são da forma* $a_{ij} \;=\; \mathbf{a}(u_i, u_j)$*.*

Demonstração – Para todos $u,\, v \in V$, temos que

$$u \;=\; \sum_{i=1}^{n} b_i\, u_i \qquad \text{e} \qquad v \;=\; \sum_{i=1}^{n} c_i\, u_i\,.$$

Desse modo, temos que

$$\mathbf{a}(u,v) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} b_i\, c_j\, \mathbf{a}(u_i,u_j) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} b_i\, c_j\, a_{ij} \;=\; [u]_\beta^t\, A\, [v]_\beta\;,$$

onde $a_{ij} \;=\; \mathbf{a}(u_i, u_j)$, o que completa a demonstração. ■

Portanto, toda forma bilinear $\mathbf{a}(\cdot,\cdot)$ sobre V pode ser representada na forma:

$$\mathbf{a}(u,v) \;=\; (\,[u]_\beta\,)^t\,A\,[v]_\beta\;,$$

onde $[u]_\beta$ e $[v]_\beta$ são os vetores coordenadas dos elementos u e v em relação à base ordenada β. Vamos denotar por $[\mathbf{a}]_\beta$ a matriz da forma bilinear $\mathbf{a}(\cdot,\cdot)$ em relação à base ordenada β. Podemos verificar facilmente que dada uma matriz $A \in \mathbb{M}_n(\mathbb{F})$, podemos definir uma forma bilinear sobre V associada a essa matriz.

Definição 8.14.4 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$,*

$$\beta \;=\; \{\,u_1\,,\,\cdots\,,\,u_j\,,\,\cdots\,,\,u_n\,\}$$

uma base ordenada para V, e $A \in \mathbb{M}_n(\mathbb{F})$ uma matriz fixa, porém arbitrária. A aplicação $\mathbf{a}(\cdot,\cdot) : V \times V \longrightarrow \mathbb{F}$ definida da forma:

$$\mathbf{a}(u,v) \;=\; (\,[u]_\beta\,)^t\,A\,[v]_\beta \qquad \text{para todos} \qquad u\,,\,v \in V\;,$$

*é denominada **forma bilinear** associada à matriz A.*

Definição 8.14.5 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $\mathbf{a}(\cdot,\cdot)$ uma forma bilinear sobre V. Dizemos que $\mathbf{a}(\cdot,\cdot)$ é uma **forma bilinear simétrica** se*

$$\mathbf{a}(u,v) \;=\; \mathbf{a}(v,u) \qquad \text{para todos} \qquad u,\,v \in V\,.$$

Exemplo 8.14.4 *Seja V um espaço vetorial real munido do produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$. Podemos verificar que o produto interno é uma forma bilinear simétrica sobre V.*

Teorema 8.14.2 *Sejam V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{F}$, $\beta = \{\,u_1\,,\,\cdots\,,\,u_n\,\}$ uma base ordenada para V, $\mathbf{a}(\cdot,\cdot)$ uma forma bilinear sobre V, e $A = [\mathbf{a}]_\beta$ a matriz da forma bilinear $\mathbf{a}(\cdot,\cdot)$ com relação à base ordenada β. Então, $\mathbf{a}(\cdot,\cdot)$ é uma forma bilinear simétrica se, e somente se, A é uma matriz simétrica.*

Demonstração – A prova pode ficar a cargo do leitor. □

Definição 8.14.6 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $\mathbf{a}(\cdot,\cdot)$ uma forma bilinear simétrica sobre V. A aplicação $\mathbf{q}(\cdot) : V \longrightarrow \mathbb{F}$ definida da forma:*

$$\mathbf{q}(u) \;=\; \mathbf{a}(u,u) \qquad \text{para todo} \qquad u \in V\;,$$

*é denominada **forma quadrática** associada à forma bilinear simétrica $\mathbf{a}(\cdot,\cdot)$.*

Definição 8.14.7 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, $\mathbf{a}(\cdot,\cdot)$ uma forma bilinear simétrica sobre V, e $\mathbf{q}(\cdot)$ a forma quadrática associada à forma bilinear simétrica $\mathbf{a}(\cdot,\cdot)$. Dizemos que $\mathbf{q}(\cdot)$ é uma **forma quadrática positiva** se*

$$\mathbf{q}(u) = \mathbf{a}(u,u) > 0$$

para todo $u \in V$ diferente do elemento neutro $0_V \in V$.

Definição 8.14.8 *Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$, $\mathbf{a}(\cdot,\cdot)$ uma forma bilinear simétrica sobre V, e $\mathbf{q}(\cdot)$ a forma quadrática associada à forma bilinear simétrica $\mathbf{a}(\cdot,\cdot)$. Dizemos que $\mathbf{q}(\cdot)$ é uma **forma quadrática semipositiva** se*

$$\mathbf{q}(u) = \mathbf{a}(u,u) \geq 0$$

para $u \in V$.

Exemplo 8.14.5 *Seja V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$. Podemos associar ao produto interno $\langle \cdot , \cdot \rangle$, que é uma forma bilinear simétrica sobre V, a forma quadrática positiva dada por:*

$$\mathbf{q}(u) = \langle u , u \rangle \qquad \text{para todo} \qquad u \in V ,$$

onde a norma Euclidiana $\| \cdot \|_2$ em V é definida por:

$$\| u \|_2 = \sqrt{\mathbf{q}(u)} \qquad \text{para todo} \qquad u \in V .$$

Exemplo 8.14.6 *Considere o espaço vetorial $\mathbb{R}^n$ com o produto interno usual $\langle \cdot , \cdot \rangle$, $\beta = \{ e_1 , \cdots , e_n \}$ a base canônica, e $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz positiva–definida. Podemos definir uma forma bilinear simétrica associada à matriz A da seguinte forma:*

$$\mathbf{a}(x,y) = ([x]_\beta)^t A [y]_\beta = x^t A y$$

para todos $x, y \in \mathbb{R}^n$. Desse modo, podemos associar a forma quadrática positiva

$$\mathbf{q}(x) = x^t A x \qquad \text{para todo} \qquad x \in \mathbb{R}^n .$$

É importante lembrar que estamos considerando os elementos de $\mathbb{R}^n$ representados na forma de matriz coluna, devido ao fato dos espaços vetoriais $\mathbb{R}^n$ e $\mathbb{M}_{n\times 1}(\mathbb{R})$ serem isomorfos.

Teorema 8.14.3 *Sejam* V *um espaço vetorial de dimensão finita sobre o corpo* $\mathbb{F}$ *e* $\mathbf{a}(\cdot,\cdot)$ *uma forma bilinear sobre* V*. Então, as matrizes* $A, B \in \mathbb{M}_n(\mathbb{F})$ *representam a forma bilinear* $\mathbf{a}(\cdot,\cdot)$ *em relação a bases ordenadas diferentes se, e somente se, a matriz* B *é congruente com a matriz* A*.*

Demonstração

($\Longrightarrow$) Tomando como hipótese que as matrizes A e B representam a mesma forma quadrática, entretanto, com relação a diferentes bases ordenadas, isto é,

$$\mathbf{a}(u,v) = ([u]_\beta)^t A [v]_\beta = ([u]_\gamma)^t B [v]_\gamma ,$$

onde $[u]_\beta = P[u]_\gamma$ para todo $u \in V$, representa a mudança de coordenadas.

Desse modo, obtemos

$$\begin{aligned} ([u]_\beta)^t A [v]_\beta &= (P[u]_\gamma)^t A (P[v]_\gamma) \\ &= ([u]_\gamma)^t (P^t A P) [v]_\gamma \\ &= ([u]_\gamma)^t B [v]_\gamma \end{aligned}$$

Portanto, $B = P^t A P$, onde P é a matriz de mudança da base ordenada γ para a base ordenada β, que é uma matriz invertível. Assim, mostramos que a matriz B é congruente com a matriz A, veja Definição 2.8.1.

($\Longleftarrow$) Tomando por hipótese que a matriz B é congruente com a matriz A, isto é, existe uma matriz invertível P tal que $B = P^t A P$.

Assim, temos que $A = (P^t)^{-1} B P^{-1}$, e obtemos

$$\begin{aligned} ([u]_\beta)^t A [v]_\beta &= ([u]_\beta)^t \{ (P^t)^{-1} B P^{-1} \} [v]_\beta \\ &= (P^{-1}[u]_\beta)^t B (P^{-1}[v]_\beta) \\ &= ([u]_\gamma)^t B [v]_\gamma \end{aligned}$$

onde $[u]_\beta = P[u]_\gamma$ para todo $u \in V$. Logo, P é a matriz de mudança da base ordenada γ para a base ordenada β, o que completa a demonstração. $\blacksquare$

Podemos verificar facilmente que a covariância é uma **forma bilinear simétrica** sobre o espaço vetorial real $I\!R^m$. De fato, a covariância satisfaz:

(a) $\mathbf{cov}(kY_1 + Y_2, Y) = k\mathbf{cov}(Y_1, Y) + \mathbf{cov}(Y_2, Y)$

(b) $\mathbf{cov}(Y, kY_1 + Y_2) = k\mathbf{cov}(Y, Y_1) + \mathbf{cov}(Y, Y_2)$

(c) $\mathbf{cov}(Y_1, Y_2) = \mathbf{cov}(Y_2, Y_1)$

para todo $k \in I\!R$ e para todos Y, Y_1, $Y_2 \in I\!R^m$.

Desse modo, temos que a variância é uma **forma quadrática** associada à forma bilinear simétrica definida pela covariância, isto é,

$$\mathbf{var}(Y) = \mathbf{cov}(Y, Y) \qquad \text{para todo} \qquad Y \in I\!R^m .$$

Logo, a variância é uma forma quadrática semipositiva, uma vez que

$$\mathbf{cov}(Y, Y) = \mathbf{var}(Y) = 0$$

para todo elemento $Y \in S = [U]$.

Assim sendo, a partir da variância podemos definir pseudo–norma, proveniente do produto interno covariante, denominada **norma covariante**, que vamos denotar por:

$$\| Y \|_{\mathbf{cov}} = \sqrt{\mathbf{var}(Y)} = \sqrt{\mathbf{cov}(Y, Y)} = \sqrt{\langle Y , Y \rangle_{\mathbf{cov}}}$$

para todo $Y \in I\!R^m$.

Definição 8.14.9 *Sejam* $Y_1, \cdots Y_j, \cdots, Y_n \in I\!R^m$ *variáveis observadas. A matriz* $C = [c_{ij}] \in I\!M_n(I\!R)$ *definida por:*

$$c_{ij} = \mathbf{cov}(Y_i, Y_j) \qquad \textit{para} \qquad i, j = 1, \cdots, n ,$$

é denominada ***matriz de covariância****.*

Note que a diagonal principal da matriz de covariância C, isto é,

$$c_{ii} = \mathbf{cov}(Y_i, Y_i) = \mathbf{var}(Y_i) \qquad \text{para} \qquad i = 1, \cdots, n ,$$

é o **vetor de variância** das variáveis observadas.

Considere o espaço vetorial real $I\!R^m$ munido do produto interno usual $\langle \cdot , \cdot \rangle$, e com a base canônica $\beta = \{ e_1 , \cdots , e_j , \cdots , e_m \}$. Como a covariância é uma forma bilinear simétrica sobre o $I\!R^m$, sabemos que sua matriz em relação à base canônica β, que vamos denotar por $A = [\mathbf{cov}]_\beta$, é uma matriz simétrica dada por:

$$a_{ij} = \mathbf{cov}(e_i, e_j) = \begin{cases} \dfrac{1}{m} & \text{para} \quad i = j \\[2ex] \dfrac{-1}{m(m-1)} & \text{para} \quad i \neq j \end{cases} \tag{8.201}$$

De fato, temos que

$$e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad , \quad \overline{e_j} = \frac{1}{m} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{e} \quad Z_j^* = e_j - \overline{e_j} = -\frac{1}{m} \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1-m \\ 1 \\ \vdots \\ 1 \end{bmatrix} .$$

Desse modo, obtemos

$$a_{ii} = \mathbf{cov}(e_i, e_i) = \mathbf{var}(e_i) = \frac{\langle Z_i^* , Z_i^* \rangle}{m - 1} = \frac{1}{m}$$

para $i = 1, \cdots , m$, e

$$a_{ij} = \mathbf{cov}(e_i, e_j) = \frac{\langle Z_i^* , Z_j^* \rangle}{m - 1} = \frac{-1}{m(m-1)}$$

para $i \neq j$.

Portanto, podemos calcular a variância entre as variáveis observadas $Y_1, Y_2 \in I\!R^m$ da seguinte forma:

$$\mathbf{cov}(Y_1, Y_2) = ([Y_1]_\beta)^t \, [\mathbf{cov}]_\beta \, [Y_2]_\beta = (Y_1)^t \, [\mathbf{cov}]_\beta \, Y_2 ,$$

facilitando consideravelmente os cálculos.

Podemos verificar facilmente que a matriz $[\mathbf{cov}]_\beta$, descrita em (8.201), é a mesma matriz obtida pelos cálculos realizados em (8.200) para representar a covariância, isto é,

$$[\mathbf{cov}]_\beta = \frac{H}{m - 1} ,$$

onde H é a matriz de projeção ortogonal sobre o subespaço $S^\perp$.

Sejam $Y_1, \cdots Y_j, \cdots, Y_n \in \mathbb{R}^m$ variáveis observadas. Vamos definir a **matriz de dados**, que denotamos por $Y \in \mathbb{M}_{m\times n}(\mathbb{R})$, da seguinte forma:

$$Y = [Y_1 \cdots Y_j \cdots Y_n],$$

e a **matriz das médias**, que vamos denotar por $\overline{Y} \in \mathbb{M}_{m\times n}(\mathbb{R})$, da seguinte forma:

$$\overline{Y} = [\overline{Y_1} \cdots \overline{Y_j} \cdots \overline{Y_n}],$$

onde $\overline{Y_j}$ é o vetor de média da variável Y_j.

Temos também associada à matriz de dados Y a **matriz de Gramm** definida por:

$$G = Y^t Y,$$

freqüentemente utilizada em análise de sinal e controle, por exemplo.

Finalmente, representamos a matriz de covariância $C \in \mathbb{M}_n(\mathbb{R})$ da seguinte forma:

$$C = \frac{(Y - \overline{Y})^t(Y - \overline{Y})}{m-1} = \frac{Z^t Z}{m-1},$$

onde a matriz $Z = Y - \overline{Y} \in \mathbb{M}_{m\times n}(\mathbb{R})$, denominada **matriz de resíduos**.

Sabemos que a matriz $Z^t Z$ é positiva–definida para $m \geq n$ e $posto(Z) = n$. Assim, a matriz de covariância é positiva–definida.

De modo análogo, sabemos que a matriz $Z^t Z$ é semipositiva–definida para $m \geq n$ e $posto(Z) = r < n$. Assim, a matriz de covariância é semipositiva–definida.

Para o caso que $m < n$, a matriz $Z^t Z$ é semipositiva–definida. Desse modo, a matriz de covariância é semipositiva–definida.

Podemos observar facilmente que a matriz $Z = Y - \overline{Y}$ pode ser escrita da forma:

$$Z = (I - P)Y = HY,$$

onde $P \in \mathbb{M}_m(\mathbb{R})$ é a matriz de projeção sobre o subespaço $S = [U]$, e a matriz $H = I - P$ é a matriz de projeção ortogonal sobre o subespaço $S^\perp$.

Portanto, a matriz de covariância pode ser escrita da seguinte forma:

$$C = \frac{Z^t Z}{m-1} = \frac{Y^t H^t H Y}{m-1} = \frac{Y^t (H^t H) Y}{m-1} = \frac{Y^t H Y}{m-1},$$

uma vez que $H^t = H$ e $H^2 = H$.

Exemplo 8.14.7 *Considere as variáveis observadas* $Y_1, Y_2, Y_3 \in I\!R^4$ *dadas por:*

$$Y_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix} \quad , \quad Y_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix} \quad e \quad Y_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ -4 \end{bmatrix} .$$

Verifique que a matriz de covariância $C = [c_{ij}] \in I\!M_3(I\!R)$ *é dada por:*

$$C = \frac{Z^t Z}{m-1} = \frac{1}{3} \begin{bmatrix} 10 & -3 & -7 \\ -3 & 2 & 0 \\ -7 & 0 & 18 \end{bmatrix} ,$$

onde $Z = Y - \overline{Y}$ *é a matriz de resíduos,* Y *a matriz de dados e* $\overline{Y}$ *a matriz das médias. Determine o posto da matriz de covariância. Neste exemplo, a matriz de covariância é uma matriz positiva–definida?*

Exemplo 8.14.8 *Considere* $D = diag(d_1, \cdots, d_j, \cdots, d_m)$ *uma matriz diagonal, com* $d_i > 0$ *para* $i = 1, \cdots, m$*, e o produto interno energia* $\langle \cdot , \cdot \rangle_D$ *em* $I\!R^m$ *associado à matriz positiva–definida* D*, isto é, para* $Y_1, Y_2 \in I\!R^m$ *temos que*

$$\langle Y_1 , Y_2 \rangle_D = Y_2^t D Y_1 .$$

(a) Determine o vetor de média $\overline{Y}$ *da variável observada* $Y \in I\!R^m$ *com relação ao produto interno* $\langle \cdot , \cdot \rangle_D$*, denominado* ***vetor de média ponderada****.*

(b) Mostre que a matriz de projeção ortogonal sobre o subespaço $S = [U]$ *com relação ao produto interno* $\langle \cdot , \cdot \rangle_D$*, onde* $U \in I\!R^m$ *é o vetor unidade, é dada por:*

$$P = \frac{U U^t D}{U^t D U} = \frac{U (D U)^t}{(D U)^t U} .$$

(c) Considere o espaço vetorial real $I\!R^3$ *munido do produto interno energia associado à matriz diagonal* $D = diag(2, 3, 5)$*. Sejam as variáveis observadas* $Y_1, Y_2 \in I\!R^3$ *dadas por:*

$$Y_1 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \quad e \quad Y_2 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} .$$

Determine a covariância entre as variáveis Y_1 *e* Y_2 *definida da seguinte forma:*

$$\mathbf{cov}(Y_1, Y_2) = \frac{\langle Z_1^* , Z_2^* \rangle_D}{2} ,$$

onde $Z_1^* = Y_1 - \overline{Y_1}$ *e* $Z_2^* = Y_2 - \overline{Y_2}$*.*

Exemplo 8.14.9 *Considere* $D = diag(d_1, \cdots, d_j, \cdots, d_m)$ *uma matriz diagonal, com* $d_i > 0$ *para* $i = 1, \cdots, m$, *e o produto interno energia* $\langle \cdot , \cdot \rangle_D$ *em* $I\!R^m$ *associado à matriz positiva–definida* D, *isto é, para* $Y_1, Y_2 \in I\!R^m$ *temos que*

$$\langle Y_1 , Y_2 \rangle_D = Y_2^t D Y_1 .$$

Sejam $Y_1, Y_2 \in I\!R^m$ *duas variáveis observadas. Temos que a covariância entre as variáveis observadas* Y_1 *e* Y_2 *pode ser escrita da forma:*

$$\begin{aligned}
\mathbf{cov}(Y_1, Y_2) &= \frac{\langle Z_1^* , Z_2^* \rangle_D}{m-1} \\
&= \frac{(Z_2^*)^t (D Z_1^*)}{m-1} \\
&= \frac{(Y_2^t H^t)(D H Y_1)}{m-1} \\
&= \frac{Y_2^t (H^t D H) Y_1}{m-1} \\
&= Y_2^t (A Y_1) \\
&= \langle Y_1 , Y_2 \rangle_A = \langle Y_1 , Y_2 \rangle_{\mathbf{cov}} ,
\end{aligned}$$

onde $Z_1^* = Y_1 - \overline{Y_1}$, $Z_2^* = Y_2 - \overline{Y_2}$ *e a matriz* A *é dada por:*

$$A = \frac{H^t D H}{m-1} ,$$

onde $H = I - P$ *a matriz de projeção sobre o subespaço* $S^\perp$ *e* P *a matriz de projeção sobre o subespaço* $S = [U]$, *com relação ao produto interno energia* $\langle \cdot , \cdot \rangle_D$, *veja Exercício 8.14.8.*

Portanto, o produto interno covariante é um pseudo–produto interno em $I\!R^m$, munido do produto interno energia $\langle \cdot , \cdot \rangle_D$, associado à matriz semipositiva–definida A, uma vez que H é uma matriz semipositiva–definida e D é uma matriz positiva–definida.

Podemos notar que a covariância depende da magnitude e da unidade física das variáveis observadas, quando essas possuem uma unidade. Desse modo, a covariância é utilizada preferencialmente nas aplicações nas quais as variáveis possuem a mesma unidade física. Desse modo, seria muito interessante definir uma medida que não tenha uma dependência da magnitude das variáveis observadas, e portanto, não depende da unidade física de cada uma delas.

Definição 8.14.10 *Sejam* $Y_1\,,\,Y_2\,\in\,\mathbb{R}^m$ *duas variáveis observadas. O* ***coeficiente de correlação*** *entre as variáveis* Y_1 *e* Y_2*, que denotamos por* $\mathbf{corr}(Y_1, Y_2)$*, é definido da seguinte forma:*

$$\mathbf{corr}(Y_1, Y_2) \;=\; \frac{\langle\, Z_1^*\,,\,Z_2^*\,\rangle}{\|\,Z_1^*\,\|_2\,\|\,Z_2^*\,\|_2} \;=\; \frac{\mathbf{cov}(Y_1, Y_2)}{\|\,Y_1\,\|_{\mathbf{cov}}\,\|\,Y_2\,\|_{\mathbf{cov}}}\,.$$

Considerando as interpretações algébricas das medidas de variância e covariância, ou seja, das definições de produto interno covariante e de norma covariante, podemos verificar que o coeficiente de correlação entre as variáveis Y_1 e Y_2 pode ser visto da forma:

$$\mathbf{corr}(Y_1, Y_2) \;=\; \cos(\theta_{12}^*)\,,$$

onde $\theta_{12}^* \in [0, \pi]$ é o ângulo entre os vetores de resíduos Z_1^* e Z_2^* com relação ao produto interno usual do $\mathbb{R}^m$, ou também pode ser interpretado como o ângulo entre as variáveis observadas Y_1 e Y_2 com relação ao produto interno covariante, veja a Definição 5.5.1, uma vez que o produto interno covariante satisfaz a desigualdade de Cauchy–Schwarz, veja o Teorema 5.3.1, isto é,

$$\mathbf{cov}(Y_1, Y_2)^2 \;\leq\; \mathbf{cov}(Y_1, Y_1)\,\mathbf{cov}(Y_2, Y_2)$$

para todos $Y_1\,,\,Y_2\,\in\,\mathbb{R}^m$.

Definição 8.14.11 *Sejam* $Y_1, \cdots Y_j, \cdots, Y_n \,\in\, \mathbb{R}^m$ *variáveis observadas. A matriz* $R = [r_{ij}] \in \mathbb{M}_n(\mathbb{R})$ *definida por:*

$$r_{ij} \;=\; \mathbf{corr}(Y_i, Y_j) \;=\; \cos(\theta_{ij}^*) \qquad \textit{para} \qquad i,\, j \;=\; 1, \cdots, n\,,$$

é denominada ***matriz de correlação****.*

Considere as variáveis observadas $Y_1, \cdots Y_j, \cdots, Y_n \in \mathbb{R}^m$. Sabemos que a matriz de covariância $C \in \mathbb{M}_n(\mathbb{R})$ pode ser representada da seguinte forma:

$$C = \frac{(Y - \overline{Y})^t (Y - \overline{Y})}{m - 1} = \frac{Z^t Z}{m - 1},$$

onde $Z = Y - \overline{Y}$ é a matriz de resíduos, Y é a matriz de dados, e $\overline{Y}$ é a matriz das médias.

Vamos considerar que $Y_j \notin S = [U]$, para $j = 1, \cdots, n$. Assim, temos que

$$Z_j^* = Y_j - \overline{Y_j} \neq 0_{\mathbb{R}^m} \qquad \text{para} \qquad j = 1, \cdots, n.$$

Inicialmente, definimos uma matriz diagonal $D = diag(d_1, \cdots, d_i, \cdots, d_n) \in \mathbb{R}^n$ da seguinte forma:

$$d_i = \frac{1}{(Z_i^*)^t Z_i^*} = \frac{1}{(Y_i - \overline{Y_i})^t (Y_i - \overline{Y_i})} \qquad \text{para} \qquad i = 1, \cdots, n.$$

Desse modo, podemos representar a matriz de correlação da seguinte forma:

$$R = \sqrt{D}\, Z^t Z \sqrt{D} = \left(Z\sqrt{D} \right)^t \left(Z\sqrt{D} \right).$$

Observamos que os elementos da diagonal principal na matriz de correlação são todos iguais a 1, pois

$$\mathbf{corr}(Y_i, Y_i) = \cos(\theta_{ii}^*) = 1,$$

de acordo com a Definição 8.14.10, e conseqüentemente da interpretação geométrica do coeficiente de correlação.

Como $\sqrt{D}$ é uma matriz positiva–definida, temos que a matriz de correlação R é uma matriz positiva–definida se, e somente se, $posto(Z) = n$, para $n < m$. No caso em que $posto(Z) = r < n$, com $n < m$, R é uma matriz semipositiva–definida.

Exemplo 8.14.10 *Considerando as informações do Exemplo 8.14.7, determine a matriz de correlação* $R \in \mathbb{M}_3(\mathbb{R})$*, como descrita acima. Verifique se* R *é positiva–definida.*

Exemplo 8.14.11 *Considere a matriz* $R = [r_{ij}] \in \mathbb{M}_n(\mathbb{R})$ *definida da forma:*

$$r_{ij} = \begin{cases} 1 & \textit{para} \quad i = j \\ \\ r & \textit{para} \quad i \neq j \end{cases}$$

Determine os valores do escalar r *de modo que* R *seja uma matriz positiva–definida.*

Inicialmente vamos considerar a matriz P de projeção ortogonal sobre o subespaço $S = [U]$, onde $U \in \mathbb{R}^n$ é o vetor unidade, que é dada por:

$$P = \frac{1}{n}\begin{bmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 \\ 1 & 1 & 1 & \cdots & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 & \cdots & 1 \end{bmatrix}$$

Sabemos que $\lambda_1 = 1$, com multiplicidades algébrica e geométrica iguais à 1, e $\lambda_2 = 0$, com multiplicidades algébrica e geométrica iguais à $n - 1$, são os autovalores da matriz P, veja Exemplo 6.1.13.

Podemos representar a matriz R da seguinte forma:

$$R = nrP + (1 - r)I_n \,,$$

onde $I_n \in \mathbb{M}_n(\mathbb{R})$ é a matriz identidade.

Podemos mostrar facilmente que se λ é um autovalor de uma matriz A com v o autovetor associado, então $\alpha\lambda + \beta$ é um autovalor da matriz $\alpha A + \beta I_n$ com v o autovetor associado, veja Exercício 6.30. Desse modo, podemos concluir que os autovalores da matriz R são dados por:

$$\mu_1 = nr + (1 - r) \qquad \text{e} \qquad \mu_2 = 1 - r \,.$$

De modo análogo, o autovalor μ_1 tem multiplicidades algébrica e geométrica iguais à 1, e o autovalor μ_2 tem multiplicidades algébrica e geométrica iguais à $n - 1$.

Portanto, impondo a condição que os autovalores de R devem ser positivos, obtemos

$$\frac{1}{1-n} < r < 1 .$$

Assim, como R é uma matriz simétrica, temos que R é uma matriz positiva–definida, veja Corolário 6.7.1. Podemos mostrar facilmente que

$$\det(R) = [\,nr + (1-r)\,](\,1 - r\,)^{n-1} .$$

Observe que a matriz R pode ser considerada uma matriz de correlação onde as variáveis observadas possuem correlações iguais.

Exemplo 8.14.12 *Considere a matriz* $A = [a_{ij}] \in I\!M_n(I\!R)$ *definida da forma:*

$$a_{ij} = \begin{cases} d & para \quad i = j \\ \\ r & para \quad i \neq j \end{cases}$$

com $r \neq d$ *e* $d > 0$*. Determine os valores dos escalares* r *e* d *de modo que* A *seja uma matriz positiva–definida e faça uma representação geométrica no plano numérico* R^2*, e mostre que*

$$\det(A) = [\,d + (n-1)r\,](\,d - r\,)^{n-1} .$$

Exemplo 8.14.13 *Com os resultados do Exemplo 8.14.12, determine os autovalores da matriz* $[\mathbf{cov}]_\beta \in I\!M_m(I\!R)$ *descrita em (8.201), onde* β *é a base canônica do* $I\!R^m$*, e as respectivas multiplicidades algébrica e geométrica de cada um dos autovalores. Descreva o subespaço associado a cada um dos autovalores da matriz* $[\mathbf{cov}]_\beta$*.*

Exemplo 8.14.14 *Considere as variáveis observadas* $Y_1, Y_2 \in I\!R^4$ *dadas por:*

$$Y_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \qquad e \qquad Y_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -5 \end{bmatrix} .$$

Determine a covariância entre as variáveis Y_1 *e* Y_2 *utilizando a matriz* $[\mathbf{cov}]_\beta$ *descrita em (8.201), onde* β *é a base canônica do* $I\!R^4$*, e calcule também pela Definição 8.14.2.*

Exemplo 8.14.15 *Com os resultados do Exemplo 8.14.12, tomando* $d = 1$*,* $r = \dfrac{1}{2}$ *e* $n = 5$*, calcule* $\det(A)$*.*

Sejam $Y_1, Y_2 \in I\!R^m$ duas variáveis observadas que possuem uma relação linear, isto é,

$$Y_2 = \alpha Y_1 + \beta U$$

para $\alpha, \beta \in I\!R$, onde $U \in I\!R^m$ é o vetor unidade. Neste caso, podemos observar que os vetores de resíduos $Z_1^* = Y_1 - \overline{Y_1}$ e $Z_2^* = Y_2 - \overline{Y_2}$ também possuem a mesma relação linear.

Portanto, temos que

$$\textbf{corr}(Y_1, Y_2) = \cos(\theta_{12}^*) = \pm 1 \,,$$

uma vez que o coeficiente de correlação é invariante quando existe uma relação linear entre as variáveis observadas, veja Exercício 8.141.

Reciprocamente, se o coeficiente de correlação entre as variáveis Y_1 e Y_2 é igual a 1, podemos concluir que existe uma relação linear entre as variáveis Y_1 e Y_2 ?

Exemplo 8.14.16 *Considere as variáveis observadas $Y_1, Y_2 \in I\!R^4$ dadas por:*

$$Y_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 4 \end{bmatrix} \qquad e \qquad Y_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 5 \end{bmatrix} .$$

Verifique que o coeficiente de correlação entre as variáveis Y_1 e Y_2 é igual a 1. Podemos afirmar que existe uma relação linear entre as variáveis Y_1 e Y_2 ? Em caso afirmativo, determine essa relação linear, isto é, determine constantes $\alpha, \beta \in I\!R$ tais que

$$Y_2 = \alpha Y_1 + \beta U \,,$$

através de um problema de regressão linear simples.

Portanto, apresentamos uma interpretação algébrica e uma interpretação geométrica para cada uma das medidas estatísticas de variância, covariância e coeficiente de correlação. Dessa forma, podemos analisar de vários outros aspectos as suas aplicações em problemas de interesse pratico.

Exercícios

Exercício 8.136 *Considere o espaço vetorial* $I\!R^m$ *munido do produto interno usual. Sejam um vetor* $Y \in I\!R^m$, $\overline{y}$ *a* ***média*** *do vetor* Y *e o* ***vetor unidade*** $U \in I\!R^m$ *dados por:*

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} \quad , \quad \overline{y} = \frac{\displaystyle\sum_{j=1}^{m} y_j}{m} \qquad e \qquad U = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} .$$

Mostre que a média do vetor Y *pode ser escrita da seguinte forma:*

$$\overline{y} = \frac{1}{\sqrt{m}} \| Y \|_2 \cos(\theta) ,$$

onde θ *é o ângulo entre os vetores* Y *e* U.

Seja $Z^* = Y - \overline{Y}$, *onde* $\overline{Y}$ *é o* ***vetor de média*** *do vetor* Y, *mostre que*

$$\| Z^* \|_2^2 = \| Y \|_2^2 - \| \overline{Y} \|_2^2 = \| Y \|_2^2 - m\,(\,\overline{y}\,)^2 .$$

Além disso, podemos verificar facilmente que a média do vetor Z^*, *que denotamos por* $\overline{z^*}$, *é nula, isto é,* $\overline{z^*} = 0$. *Logo, o vetor de média* $\overline{Z^*} = 0_{I\!R^m}$. *Explique esse resultado.*

Exercício 8.137 *Considere o espaço vetorial* $I\!R^4$ *com o produto interno usual. Sejam* S *o subespaço gerado pelo conjunto* $\beta = \{\, q_1 \,,\, q_2 \,\}$ *e o vetor* $Y \in I\!R^4$ *dados por:*

$$q_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad q_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \qquad e \qquad Y = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix} .$$

(a) Encontre o elemento $\widehat{Y}$ *que é a melhor aproximação do elemento* Y *em* S.

(b) Faça uma análise da qualidade dessa aproximação utilizando o ***Coeficiente de Determinação do Modelo****, isto é,*

$$R^2 = \frac{SQR}{SQT} = \frac{\| \widehat{Y} - \overline{Y} \|_2^2}{\| Y - \overline{Y} \|_2^2} ,$$

onde $\overline{Y}$ *é o vetor de média do vetor* Y.

(c) Determine uma base ortogonal para o subespaço $S^{\perp}$, *que é o complemento ortogonal do subespaço* S *em* $I\!R^4$ *com relação ao produto interno usual.*

Exercício 8.138 *Sejam* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$*, com* $m > n$ *e* $posto(A) = n$*, onde a primeira coluna da matriz* A *é o vetor unidade* U *que é dado por:*

$$U = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

$Y \in \mathbb{R}^m$*, vetor de observações, e* $\widehat{Y} \in \mathbb{R}^m$ *a projeção ortogonal do vetor de observações* Y *sobre o subespaço* $\mathcal{R}(A)$*, vetor de estimativas. Mostre que a média do vetor de resíduos* $Z = Y - \widehat{Y}$ *é igual a zero.*

Exercício 8.139 *Considere o espaço vetorial real* $\mathbb{R}^m$ *munido do produto interno usual, e com a norma Euclidiana* $\|\cdot\|_2$ *proveniente do produto interno. Mostre que a distância Euclidiana é invariante por translação, isto é, para todos* $X, Y \in \mathbb{R}^m$ *tem–se que*

$$d_2(X,Y) = \| X - Y \|_2 = d_2(X',Y') = \| X' - Y' \|_2 ,$$

com

$$X' = X - T \qquad e \qquad Y' = Y - T ,$$

onde $T \in \mathbb{R}^m$ *é o elemento que realiza a translação.*

Exercício 8.140 *Seja* $Y \in \mathbb{R}^m$ *um variável observadas. Mostre que*

(a) *A variância não é invariante com relação à multiplicação por um escalar, isto é,*

$$\mathbf{var}(Y) \neq \mathbf{var}(kY)$$

(b) *A variância é invariante com relação à adição por um elemento constante, isto é,*

$$\mathbf{var}(Y) = \mathbf{var}(Y + kU)$$

para todo $k \in \mathbb{R}$*, onde* $U \in \mathbb{R}^m$ *é o vetor unidade.*

Exercício 8.141 *Sejam* $Y_1, Y_2 \in \mathbb{R}^m$ *variáveis observadas. Mostre que o coeficiente de correlação entre as variáveis* Y_1 *e* Y_2 *é invariante quando existe uma relação linear em cada uma das varáveis, isto é,*

$$\mathbf{corr}(Y_1, Y_2) = \mathbf{corr}((\alpha_1 Y_1 + \beta_1 U), (\alpha_2 Y_2 + \beta_2 U))$$

para todos $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$*, onde* $U \in \mathbb{R}^m$ *é o vetor unidade.*

Exercício 8.142 *Considere as variáveis observadas* $Y_1, Y_2, Y_3 \in I\!R^3$ *dadas por:*

$$Y_1 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} \quad , \quad Y_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad e \quad Y_3 = \begin{bmatrix} 2 \\ 1 \\ -6 \end{bmatrix} .$$

Determine a matriz de covariância $C = [c_{ij}] \in I\!M_3(I\!R)$ *definida por:*

$$c_{ij} = \mathbf{cov}(Y_i, Y_j)$$

para $i, j = 1, 2, 3$.

Exercício 8.143 *Sejam* $Y_1, \cdots Y_j, \cdots , Y_n \in I\!R^m$ *variáveis observadas, e a matriz de covariância* $C = [c_{ij}] \in I\!M_n(I\!R)$ *definida por:*

$$c_{ij} = \mathbf{cov}(Y_i, Y_j) \qquad para \qquad i, j = 1, \cdots , n .$$

Faça uma análise de forma detalhada para cada situação em que a matriz de covariância C *é semipositiva–definida.*

Exercício 8.144 *Sejam* $Y_1, \cdots Y_j, \cdots , Y_n \in I\!R^m$ *variáveis observadas, e tome as seguintes combinações lineares*

$$Y = \sum_{i=1}^{n} a_i \, Y_i \qquad e \qquad Z = \sum_{j=1}^{n} b_j \, Y_j .$$

Mostre que a covariância entre Y *e* Z *é representada da seguinte forma:*

$$\mathbf{cov}(Y, Z) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \, b_j \, \mathbf{cov}(Y_i, Y_j) = (X_1)^t \, C \, X_2 ,$$

onde $C \in I\!M_n(I\!R)$ *é a matriz de covariância, veja Definição 8.14.9, e os elementos* $X_1, X_2 \in I\!M_{n \times 1}(I\!R)$ *são dados por:*

$$X_1 = \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{bmatrix} \qquad e \qquad X_2 = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix} .$$

Exercício 8.145 *Sejam* $Y_1, \cdots Y_j, \cdots, Y_n \in I\!R^m$. *Mostre que a equação*

$$c_1 Y_1 + \cdots + c_j Y_j + \cdots + c_n Y_n = 0_{I\!R^m}$$

é invariante por translação se, e semente se,

$$c_1 + \cdots + c_j + \cdots + c_n = 0 \,,$$

isto é,

$$c_1 Y_1' + \cdots + c_j Y_j' + \cdots + c_n Y_n' = 0_{I\!R^m} \,,$$

onde $Y_j' = Y_j - X$ *para* $j = 1, \cdots, n$, *para qualquer elemento* $X \in I\!R^m$ *que realiza a translação.*

Exercício 8.146 *O* ***centróide*** *do conjunto de variáveis observadas* $Y_1, \cdots, Y_n \in I\!R^m$, *que indicamos por* $\overline{Y}$, *é definido da seguinte forma:*

$$\overline{Y} = \frac{Y_1 + \cdots + Y_j + \cdots + Y_n}{n} \,.$$

Desse modo, temos que

$$n\overline{Y} - Y_1 - \cdots - Y_j - \cdots - Y_n = 0_{I\!R^m} \,.$$

Como uma aplicação do Exercício 8.145, mostre que o centróide do conjunto de variáveis transladadas $Y_j' = Y_j - X$ *para* $j = 1, \cdots, n$, *que indicamos por* $\overline{Y}'$, *é dado por:*

$$\overline{Y}' = \overline{Y} - X \,,$$

para qualquer elemento $X \in I\!R^m$ *que realiza a translação.*

Exercício 8.147 *Faça uma verificação do Exercício 8.146, para o seguinte conjunto de variáveis observadas* $Y_1, \cdots, Y_4 \in I\!R^3$ *dado por:*

$$Y_1 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \,,\; Y_2 = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix} \,,\; Y_3 = \begin{bmatrix} 5 \\ 7 \\ 1 \end{bmatrix} \,,\; Y_4 = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} \,,$$

e $X = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ *o elemento que realiza a translação.*

Exercício 8.148 *Considere um experimento conduzido com a finalidade de analisar a variação do* ***calor específico*** *do etileno glicol*[2] *($C_2H_4(OH)_2$), indicamos essa variável por Y, em função da temperatura, que indicamos por T, e da densidade, que indicamos por D. Os resultados do experimento estão na Tabela 8.2.*

Tabela 8.3: Propriedades Termofísicas do etileno glicol

Temperatura $[K]$ T	Densidade $\left[\dfrac{kg}{m^3}\right]$ D	Calor Específico $\left[\dfrac{kJ}{kg}K\right]$ Y
280.0	1125.8	2.323
290.0	1118.8	2.368
300.0	1114.4	2.415
310.0	1103.7	2.460
320.0	1096.2	2.505
330.0	1089.5	2.549
340.0	1083.8	2.592
350.0	1079.0	2.637
360.0	1074.0	2.682
370.0	1066.7	2.728

Estudar a relação da variável resposta Y, que representa o calor específico do etileno glicol, em função da variável regressora T, que é a temperatura, e da variável regressora D, que é a densidade, através do seguinte ***Modelo de Regressão Linear Múltipla***

$$Y(T,D) \;=\; \beta_1 \;+\; \beta_2 T \;+\; \beta_3 D \;+\; \beta_4 T D\,.$$

Apresentar uma análise do modelo proposto, verificando a qualidade do ajuste dos dados observados, através do ***Coeficiente de Determinação do Modelo****, da* ***Soma de Quadrados dos Resíduos*** *e do* ***gráfico de dispersão****. O que podemos concluir?*

Considere as variáveis observadas T, D, Y, dadas na Tabela 8.3, e determine a matriz de correlação. O que podemos concluir?

[2]F. P. Incropera e D. P. DeWitt, *Fundamentos de Transferência de Calor e de Massa*, LTC, 1998.

8.15 Solução de norma–2 Mínima

Sejam $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m < n$ e $posto(A) = m$, e $b \in \mathbb{R}^m$. Considere o conjunto solução do **Sistema Linear Subdeterminado** $Ax = b$ definido por:

$$S = \{ \, x \in \mathbb{R}^n \; / \; Ax = b \, \} \, .$$

Definimos a solução de norma–2 mínima para o sistema linear $Ax = b$, como sendo o elemento $x^* \in S$ tal que

$$\| \, x^* \, \|_2 \;\leq\; \| \, x \, \|_2 \qquad \text{para todo} \qquad x \in S \, .$$

Vamos determinar a solução de norma–2 mínima x^* através da fatoração $A^t = QR$, onde $Q \in \mathbb{M}_n(\mathbb{R})$ e $R \in \mathbb{M}_{n \times m}(\mathbb{R})$, descritas da seguinte forma:

$$Q = \begin{bmatrix} \widehat{Q} & \widetilde{Q} \end{bmatrix} \qquad \text{e} \qquad R = \begin{bmatrix} \widehat{R} \\ 0_{p \times m} \end{bmatrix} ,$$

onde $\widehat{Q} \in \mathbb{M}_{n \times m}(\mathbb{R})$ e $\widetilde{Q} \in \mathbb{M}_{n \times p}(\mathbb{R})$ são matrizes ortogonais, $\widehat{R} \in \mathbb{M}_m(\mathbb{R})$ é uma matriz triangular superior e $0_{p \times m} \in \mathbb{M}_{p \times m}(\mathbb{R})$ é a matriz nula, com $p = n - m$.

Substituindo a fatoração $A = R^t \, Q^t$ no sistema linear $Ax = b$, obtemos

$$R^t \, Q^t x = b \, .$$

Chamando $z = Q^t x$, temos o seguinte sistema linear triangular inferior

$$R^t \, z = b \qquad \Longleftrightarrow \qquad \widehat{R}^t y + 0^t w = b \, ,$$

onde $y \in \mathbb{R}^m$ e $w \in \mathbb{R}^p$, que são as variáveis livres do sistema linear $R^t z = b$.

Note que o elemento $z \in \mathbb{R}^n$ foi particionado da seguinte forma:

$$z = \begin{bmatrix} y \\ w \end{bmatrix} .$$

Como $\| \, x \, \|_2 = \| \, z \, \|_2$, podemos observar facilmente que para encontrar uma solução de norma–2 mínima para o sistema linear $Ax = b$, basta encontrar uma solução de norma–2 mínima para o sistema linear $R^t z = b$.

Como estamos interessado em determinar uma solução de norma–2 mínima para o sistema linear $R^t z = b$, vamos tomar $w^* = 0_{I\!R^p}$.

Desse modo, ficamos com o seguinte sistema triangular inferior

$$\widehat{R}^t y = b \,,$$

que possui uma única solução y^*, desde que $\widehat{R}$ é invertível, pois $posto(A^t) = m$.

Assim, a solução de norma–2 mínima z^* para o sistema linear $R^t z = b$ é dada por:

$$z^* = \begin{bmatrix} y^* \\ w^* \end{bmatrix} .$$

Portanto, a solução de norma–2 mínima x^* para o sistema linear $Ax = b$ é obtida da seguinte forma:

$$x^* = Qz^* = \begin{bmatrix} \widehat{Q} & \widetilde{Q} \end{bmatrix} \begin{bmatrix} y^* \\ w^* \end{bmatrix} \qquad \Longleftrightarrow \qquad x^* = \widehat{Q} y^* \,.$$

Finalmente, temos que a caracterização da solução de norma–2 mínima para o sistema linear $Ax = b$ é representada da seguinte forma:

$$x^* = \widehat{Q} \widehat{R}^{-t} b \,,$$

pois $y^* = \widehat{R}^{-t} b$.

Desse modo, tomando a fatoração $A^t = \widehat{Q}\widehat{R}$, temos que a matriz $A^\dagger = \widehat{Q}\widehat{R}^{-t}$ é a **inversa a direita**, ou **pseudo–inversa**, da matriz A.

Podemos observar facilmente que a solução de norma–2 mínima $x^* \in S$ é o elemento que está mais próximo do elemento neutro $0_{I\!R^n}$. Assim, temos uma interpretação geométrica para a solução de norma–2 mínima.

Note que o conjunto solução S pode ser representado da seguinte forma:

$$S = \mathcal{N}(A) + x_p \,,$$

onde x_p é uma solução particular para o sistema linear $Ax = b$. Desse modo, dizemos que S é uma **variedade linear**, isto é, um subespaço vetorial transladado.

Problema de Minimização com Restrição

Uma outra maneira para encontrar a solução de norma–2 mínima para o sistema linear subdeterminado

$$Ax = b$$

é através do seguinte **Problema de Programação Quadrática**

$$\begin{cases} & \min\{ \langle x , x \rangle \} \\ \\ \text{sujeito a} & Ax - b = 0_{\mathbb{R}^m} \end{cases}$$

Utilizando o **Método dos Multiplicadores de Lagrange**, associamos ao *Problema de Programação Quadrática com Restrição* um *Problema de Otimização Global*, considerando a função de Lagrange

$$F(x, \lambda) = \frac{1}{2}\langle x , x \rangle + \langle Ax - b , \lambda \rangle ,$$

onde $\lambda \in \mathbb{R}^m$ é o vetor dos multiplicadores de Lagrange.

Sendo assim, temos que determinar os pontos críticos da função de Lagrange, isto é, encontrar $(x^* , \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ tal que

$$\nabla F(x^*, \lambda^*) = \begin{cases} \nabla_x F(x^*, \lambda^*) = 0_{\mathbb{R}^n} \\ \\ \nabla_\lambda F(x^*, \lambda^*) = 0_{\mathbb{R}^m} \end{cases}$$

Inicialmente, precisamos encontrar $\nabla_x F(x, \lambda)$. Para isso, calculamos a *primeira variação* da função de Lagrange F com relação à variável x, que é definida da forma:

$$F'_x(x, \lambda)(v) = \left\{ \frac{d}{dt} F(x + tv, \lambda) \right\}_{t=0}$$

para $v \in \mathbb{R}^n$. Assim, temos que

$$F'_x(x, \lambda)(v) = \langle x + A^t\lambda , v \rangle \qquad \text{para todo} \qquad v \in \mathbb{R}^n ,$$

mostrando que

$$\nabla_x F(x, \lambda) = x + A^t\lambda .$$

De modo análogo, vamos encontrar $\nabla_\lambda F(x,\lambda)$. Para isso, precisamos calcular a *primeira variação* da função de Lagrange F com relação à variável λ, que é definida da forma:

$$F'_\lambda(x,\lambda)(w) \;=\; \left\{ \frac{d}{dt} F(x, \lambda + tw) \right\}_{t=0}$$

para $w \in I\!R^m$. Assim, temos que

$$F'_\lambda(x,\lambda)(w) \;=\; \langle\, Ax - b \,,\, w \,\rangle \qquad \text{para todo} \qquad w \in I\!R^m \,,$$

mostrando que

$$\nabla_\lambda F(x,\lambda) \;=\; Ax \;-\; b \,.$$

Portanto, temos que $\nabla F(x,\lambda)$ é um vetor coluna, de ordem $(n+m) \times 1$, dado por:

$$\nabla F(x,\lambda) \;=\; \begin{bmatrix} \nabla_x F(x,\lambda) \\ \\ \nabla_\lambda F(x,\lambda) \end{bmatrix} \;=\; \begin{bmatrix} x \;+\; A^t\lambda \\ \\ Ax \;-\; b \end{bmatrix},$$

que utilizamos para calcular os pontos críticos da função de Lagrange.

Desse modo, obtemos o seguinte

Problema de Ponto Sela: encontrar $x^* \in I\!R^n$ e $\lambda^* \in I\!R^m$ solução do sistema linear indefinido

$$\begin{cases} \;\; x \;\;+\;\; A^t\lambda \;\;=\;\; 0_{I\!R^n} \\ \\ Ax \qquad\qquad\;\; = \;\; b \end{cases}$$

que na forma matricial é representado por:

$$\begin{bmatrix} I_n & A^t \\ \\ A & 0_m \end{bmatrix} \begin{bmatrix} x \\ \\ \lambda \end{bmatrix} \;=\; \begin{bmatrix} 0_{I\!R^n} \\ \\ b \end{bmatrix},$$

onde $I_n \in I\!M_n(I\!R)$ é a matriz identidade e $0_m \in I\!M_m(I\!R)$ é a matriz nula.

É importante observar que estamos considerando um *sistema linear indefinido* como sendo aquele que possui uma **matriz indefinida**, isto é, a matriz do sistema linear é simétrica e possui autovalores de ambos os sinais.

Neste caso, a matriz do Problema de Ponto Sela que é dada por:

$$H = \begin{bmatrix} I_n & A^t \\ A & 0_m \end{bmatrix}$$

é uma matriz simétrica invertível, de ordem $(n+m)$, que possui n autovalores positivos e m autovalores negativos, de acordo com os resultados do Exemplo 6.7.7.

Sendo assim, podemos mostrar que a solução $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ do Problema de Ponto Sela é um **ponto sela** para a função de Lagrange, isto é, tem–se

$$F(x^*, \lambda) \leq F(x^*, \lambda^*) \leq F(x, \lambda^*),$$

para quaisquer $x \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}^m$. De modo equivalente, temos que

$$F(x^*, \lambda^*) = \min_{x} \max_{\lambda} \{ F(x, \lambda) \}.$$

Finalmente, do sistema linear indefinido, obtemos

$$\lambda^* = -(AA^t)^{-1} b \qquad \text{e} \qquad x^* = -A^t \lambda^* = A^t (AA^t)^{-1} b,$$

onde x^* é a solução de norma–2 mínima para o sistema linear $Ax = b$.

Como $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m < n$ e $posto(A) = m$, sabemos que a matriz AA^t é positiva–definida. Logo, o vetor dos multiplicadores de Lagrange λ^* é a única solução do sistema linear positivo–definido $AA^t \lambda = -b$, que pode ser obtida pela Fatoração de Cholesky, descrita na seção 8.5.

Definição 8.15.1 *Seja $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m < n$ e $posto(A) = m$. A matriz*

$$A^\dagger = A^t (AA^t)^{-1}$$

*é a **inversa a direita**, ou **pseudo–inversa**, da matriz A, que satisfaz as seguintes propriedades:*

1. $(AA^\dagger)^t = AA^\dagger$.
2. $(A^\dagger A)^t = A^\dagger A$.
3. $AA^\dagger A = A$.
4. $A^\dagger A A^\dagger = A^\dagger$.

*A pseudo–inversa é denominada **inversa generalizada de Moore–Penrose**.*

Finalmente, para fazer a classificação do ponto crítico (x^*, λ^*), temos que calcular a segunda variação da função de Lagrange com relação às variáveis x e λ. Assim, obtemos a **matriz Hessiana**

$$H(x,\lambda) = \begin{bmatrix} I_n & A^t \\ A & 0_m \end{bmatrix},$$

onde

$$F''_{xx}(x,\lambda;v)(u) = \left\{ \frac{d}{dt} F'_x(x+tu,\lambda)(v) \right\}_{t=0} = \langle I_n u \,,\, v \rangle$$

$$F''_{\lambda x}(x,\lambda;v)(u) = \left\{ \frac{d}{dt} F'_x(x,\lambda+tu)(v) \right\}_{t=0} = \langle A^t u \,,\, v \rangle$$

para $v\,,\,u \in I\!\!R^n$, e

$$F''_{x\lambda}(x,\lambda;w)(z) = \left\{ \frac{d}{dt} F'_\lambda(x+tz,\lambda)(w) \right\}_{t=0} = \langle Az \,,\, w \rangle$$

$$F''_{\lambda\lambda}(x,\lambda;w)(z) = \left\{ \frac{d}{dt} F'_\lambda(x,\lambda+tz)(w) \right\}_{t=0} = \langle 0_m z \,,\, w \rangle$$

para $w\,,\,z \in I\!\!R^m$.

Como a matriz Hessiana da função de Lagrange F é uma matriz indefinida, temos que o ponto crítico (x^*, λ^*) é um ponto de sela.

Podemos também determinar a matriz Hessiana da função de Lagrange F de uma outra maneira. Para isso, vamos representar $\nabla F(X)$, que é um vetor coluna de ordem $(n+m)\times 1$, da seguinte forma:

$$\nabla F(X) = \begin{bmatrix} x + A^t\lambda \\ Ax - b \end{bmatrix} = \begin{bmatrix} I_n & A^t \\ A & 0_m \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} - \begin{bmatrix} 0_{I\!\!R^n} \\ b \end{bmatrix} = HX - B\,,$$

onde os elementos $X,\,B \in I\!\!R^{n+m}$ são dados por:

$$X = \begin{bmatrix} x \\ \lambda \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0_{I\!\!R^n} \\ b \end{bmatrix}.$$

Desse modo, a primeira variação da função de Lagrange F no ponto X na direção do vetor $Y \in I\!R^{n+m}$, pode ser escrita da seguinte forma:

$$F'(X)(Y) = \langle HX - B, Y \rangle .$$

Portanto, a segunda variação da função de Lagrange F no ponto X na direção do vetor $Z \in I\!R^{n+m}$, é definida da seguinte forma:

$$F''(X;Y)(Z) = \left\{ \frac{d}{dt} F'(X + tZ, \lambda)(Y) \right\}_{t=0} = \langle HZ, Y \rangle .$$

Assim, mostramos que a matriz H é a matriz Hessiana da função de Lagrange.

Vamos indicar por $\alpha_1, \cdots, \alpha_n$ os autovalores positivos da matriz H, com $v_1, \cdots, v_n$ os autovetores associados, e indicamos por $\beta_1, \cdots, \beta_m$ os autovalores negativos da matriz H, com $w_1, \cdots, w_m$ os autovetores associados.

Como H é uma matriz simétrica, sabemos que

$$\Gamma = \{ v_1, \cdots, v_n, w_1, \cdots, w_m \}$$

é uma base ortonormal para o espaço vetorial $I\!R^{n+m}$. Desse modo, todo elemento $Y \in I\!R^{n+m}$ pode ser escrito de modo único como:

$$Y = \sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{m} b_i w_i .$$

Portanto, temos que

$$Y^t H Y = \langle HY, Y \rangle = \sum_{i=1}^{n} a_i^2 \alpha_i + \sum_{i=1}^{m} b_i^2 \beta_i \begin{cases} > & 0 \\ < & 0 \end{cases}$$

dependendo do elemento Y escolhido, pois o primeiro somatório é positivo e o segundo somatório é negativo. Assim, mostramos que H é uma matriz indefinida.

Exemplo 8.15.1 *Considere a matriz* $A \in I\!M_{2\times 4}(I\!R)$ *e o vetor* $b \in I\!R^2$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Encontre a solução de norma–2 mínima do sistema linear $Ax = b$.

Considerando o seguinte

Problema de Ponto Sela: encontrar $x^* \in I\!R^4$ e $\lambda^* \in I\!R^2$ solução do sistema linear indefinido

$$\begin{cases} x + A^t\lambda = 0_{I\!R^4} \\ Ax = b \end{cases}$$

temos que x^* é a solução de norma–2 mínima.

Obtendo explicitamente a solução do sistema linear indefinido, através de equações não acopladas, sabemos que λ^* é a única solução do sistema linear positivo–definido

$$AA^t\lambda = -b\,,$$

desde que $posto(A) = 2$ e AA^t é uma matriz positiva–definida. Para isso, tomamos a fatoração de Cholesky da matriz AA^t, cujo fator de Cholesky é dado por:

$$G = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix},$$

e resolvendo os sistemas triangulares

$$\begin{cases} G^t z = -b \\ G\lambda = z^* \end{cases}$$

temos que o vetor dos multiplicadores de Lagrange λ^* é dado por:

$$\lambda^* = \frac{1}{2}\begin{bmatrix} -2 \\ 3 \end{bmatrix}.$$

Portanto, a solução de norma–2 mínima é determinada da forma:

$$x^* = -A^t\lambda^* = \frac{1}{2}\begin{bmatrix} -1 \\ 2 \\ -1 \\ 2 \end{bmatrix}.$$

Exercícios

Exercício 8.149 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, *com* $m < n$ *e* $posto(A) = m$, *e* $b \in \mathbb{R}^m$. *Considere o conjunto solução do sistema linear* $Ax = b$ *definido por:*

$$S = \{ x \in \mathbb{R}^n \ / \ Ax = b \}.$$

Definimos a solução de norma–2 mínima do sistema $Ax = b$, *como sendo o elemento* $x^* \in S$ *tal que*

$$\| x^* \|_2 \leq \| x \|_2 \qquad \textit{para todo} \qquad x \in S.$$

(a) Mostre que $x^* \in \mathbb{R}^n$ *é dado por* $x^* = A^t(AA^t)^{-1}b$, *utilizando o método dos multiplicadores de Lagrange, considerando a função de Lagrange*

$$F(x,\lambda) = \frac{1}{2}\langle x , x \rangle + \langle Ax - b , \lambda \rangle ,$$

onde $\lambda \in \mathbb{R}^m$ *é o vetor dos multiplicadores de Lagrange.*

(b) Mostre que a matriz $A^\dagger = A^t(AA^t)^{-1}$ *é a pseudo–inversa da matriz* A, *isto é, mostre que* $A^\dagger$ *satisfaz as propriedades de Moore–Penrose.*

Exercício 8.150 *Encontre a solução de norma–2 mínima do sistema linear*

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix},$$

utilizando o procedimento **`qrndmin`** *disponível na página da disciplina no link Matlab. Compare o resultado obtido com a solução de uma calculadora científica e também com o resultado obtido pelo Matlab utilizando o procedimento* **`mldivide`**, *que tem a seguinte sintaxe*

```
>> Xstar = mldivide(A,b)
```

O que podemos dizer a respeito do método utilizado em cada um dos casos?

Exercício 8.151 *Determine a solução de norma–2 mínima para* $Ax = b$, *onde*

$$A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 2 \end{bmatrix}.$$

Dê uma interpretação geométrica.

Exercício 8.152 *Considere o sistema linear apresentado no Exercício 8.150. Determine a solução de norma–2 mínima utilizando a solução geral na sua forma paramétrica.*

Exercício 8.153 *Considere a matriz* $A \in \mathbb{M}_{2\times 4}(\mathbb{R})$ *e o vetor* $b \in \mathbb{R}^2$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Considerando a fatoração $A^t = QR$, *com* $Q \in \mathbb{M}_{4\times 2}(\mathbb{R})$ *e* $R \in \mathbb{M}_2(\mathbb{R})$, *encontre a solução de norma–2 mínima do sistema linear* $Ax = b$. *Determine a pseudo–inversa da matriz* A, *fazendo uso da fatoração* $A^t = QR$.

Exercício 8.154 *Encontre a solução de norma–2 mínima do sistema linear*

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

através do Problema de Ponto Sela.

Exercício 8.155 *Seja a matriz* $A \in \mathbb{M}_{3\times 4}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^3$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 6 \\ 5 \\ 6 \end{bmatrix}.$$

Encontre a solução de norma–2 mínima do sistema linear $Ax = b$ *através do Problema de Ponto Sela.*

Exercício 8.156 *Seja a matriz* $A \in \mathbb{M}_{2\times 4}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^2$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Determine a solução de norma–2 mínima do sistema linear $Ax = b$.

Exercício 8.157 *Seja a matriz* $A \in \mathbb{M}_{3\times 5}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^3$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 2 & -1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Encontre a solução de norma–2 mínima do sistema linear $Ax = b$ *fazendo uso do procedimento* **qr** *do Matlab. Determine a pseudo–inversa* $A^\dagger$ *da matriz* A.

Exercício 8.158 *Seja a matriz* $A \in \mathbb{M}_{3\times 5}(\mathbb{R})$ *e o elemento* $b \in \mathbb{R}^3$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 2 & -1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 3 \\ 11 \\ 13 \end{bmatrix}.$$

(a) *Faça um procedimento em Matlab para encontrar a solução de norma–2 mínima do sistema linear subdeterminado* $Ax = b$, *considerando a fatoração*

$$A^t = QR,$$

com $Q \in \mathbb{M}_{5\times 3}(\mathbb{R})$ *e* $R \in \mathbb{M}_3(\mathbb{R})$.

(b) *Faça um procedimento em Matlab para encontrar a solução de norma–2 mínima do sistema linear subdeterminado* $Ax = b$, *através do Problema de Ponto Sela.*

(c) *Verifique que a matriz do Problema de Ponto Sela, que é dada por:*

$$H = \begin{bmatrix} I_n & A^t \\ A & 0_m \end{bmatrix},$$

é uma matriz simétrica, de ordem $(n + m)$, *que possui* n *autovalores positivos e* m *autovalores negativos, com* $m = 3$ *e* $n = 5$. *Utilize o procedimento* **eig** *do Matlab para calcular os autovalores da matriz* H. *Construa a matriz* H *da seguinte forma:*

```
>> H = [ eye(5)  A' ; A  zeros(3) ];
```

Exercício 8.159 *Encontre a solução de norma–2 mínima do sistema linear*

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

através do Problema de Ponto Sela.

8.16 Problemas de Ponto Sela

Considere o espaço vetorial real $\mathbb{R}^n$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Sejam $A \in \mathbb{M}_n(\mathbb{R})$ positiva–definida, $b \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $B \in \mathbb{M}_{m\times n}(\mathbb{R})$, com $m < n$ e $posto(B) = m$, e o funcional $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ definido por:

$$J(x) = \frac{1}{2}\langle Ax , x \rangle - \langle b , x \rangle .$$

Considere o seguinte **Problema de Programação Quadrática**

$$\begin{cases} & \min\{ J(x) \} \\ & \\ \text{sujeito a} & Bx = d \end{cases}$$

Utilizando o **Método dos Multiplicadores de Lagrange**, associamos ao *Problema de Programação Quadrática com Restrição* um *Problema de Otimização Global*, considerando a função de Lagrange

$$F(x,\lambda) = J(x) + \langle Bx - d , \lambda \rangle ,$$

onde $\lambda \in \mathbb{R}^m$ é o vetor dos multiplicadores de Lagrange.

Sendo assim, temos que determinar os pontos críticos da função de Lagrange, isto é, encontrar $(x^* , \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ tal que

$$\nabla F(x^*,\lambda^*) = \begin{cases} \nabla_x F(x^*,\lambda^*) = 0_{\mathbb{R}^n} \\ \\ \nabla_\lambda F(x^*,\lambda^*) = 0_{\mathbb{R}^m} \end{cases}$$

Calculando a *primeira variação* da função de Lagrange F com relação à variável x, obtemos

$$\nabla_x F(x,\lambda) = Ax + B^t\lambda - b ,$$

e calculando a *primeira variação* da função de Lagrange F com relação à variável λ, obtemos

$$\nabla_\lambda F(x,\lambda) = Bx - d .$$

Desse modo, temos o seguinte

Problema de Ponto Sela: Encontrar $x^* \in \mathbb{R}^n$ e $\lambda^* \in \mathbb{R}^m$ solução do sistema linear indefinido

$$\begin{cases} Ax + B^t\lambda = b \\ \\ Bx = d \end{cases} \tag{8.202}$$

que na forma matricial é representado por:

$$\begin{bmatrix} A & B^t \\ B & 0_m \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix},$$

onde 0_m é a matriz nula de ordem $m \times m$.

Por simplicidade, vamos indicar o sistema linear indefinido por $HX = Y$, onde

$$X = \begin{bmatrix} x \\ \lambda \end{bmatrix} \quad , \quad Y = \begin{bmatrix} b \\ d \end{bmatrix} \qquad \text{e} \qquad H = \begin{bmatrix} A & B^t \\ B & 0_m \end{bmatrix}.$$

Nesta caso, a matriz do Problema de Ponto Sela H é uma matriz simétrica invertível, de ordem $n + m$, com n autovalores positivos e m autovalores negativos, isto é, H é uma matriz indefinida, de acordo com os resultados do Exemplo 6.7.8.

Finalmente, calculando a segunda variação da função de Lagrange F com relação às variáveis x e λ, obtemos a matriz Hessiana de F, que é a matriz H. Assim, o ponto crítico (x^*, λ^*) é um ponto de sela para a função de Lagrange.

Solução de Quadrados Mínimos. Problema Primal–dual

Sejam $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m > n$ e $posto(A) = n$, e $b \in \mathbb{R}^m$. Queremos determinar a solução de quadrados mínimos para o sistema linear sobredeterminado $Ax = b$. Para isso, definimos o funcional $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ da seguinte forma:

$$J(x) = \langle Ax - b , Ax - b \rangle \quad ; \quad x \in \mathbb{R}^n . \tag{8.203}$$

Desse modo, temos o seguinte

Problema Primal: encontrar $x^* \in \mathbb{R}^n$ tal que

$$J(x^*) = \min\{ J(x) \quad ; \quad x \in \mathbb{R}^n \} , \tag{8.204}$$

que é equivalente ao **Sistema Normal**

$$A^t A x = A^t b . \tag{8.205}$$

Assim, a solução de quadrados mínimos $x^* \in \mathbb{R}^n$ é a única solução do sistema normal, isto é,

$$A^t A x^* = A^t b \quad \Longleftrightarrow \quad A^t(b - Ax^*) = 0_{\mathbb{R}^n} ,$$

que podemos concluir que o elemento $w^* = b - Ax^* \in \mathcal{N}(A^t) = \mathcal{R}(A)^\perp$. Logo, w^* é a melhor aproximação do elemento $b \in \mathbb{R}^m$ no subespaço $\mathcal{N}(A^t) = \mathcal{R}(A)^\perp$.

Portanto, queremos determinar o elemento $w^* \in \mathcal{N}(A^t)$ que está mais próximo do elemento $b \in \mathbb{R}^m$ com relação à norma Euclidiana. Sendo assim, temos o seguinte

Problema Dual

$$\begin{cases} & \min\{ \| w - b \|_2 \} \\ \\ \text{sujeito a} & A^t w = 0_{\mathbb{R}^n} \end{cases}$$

Utilizando o **Método dos Multiplicadores de Lagrange**, associamos ao *Problema de Programação Quadrática com Restrição* um *Problema de Otimização Global*, considerando a função de Lagrange

$$F(x, \lambda) = \frac{1}{2} \langle w - b , w - b \rangle + \langle A^t w , \lambda \rangle ,$$

onde $\lambda \in \mathbb{R}^n$ é o vetor dos multiplicadores de Lagrange.

Assim, temos que determinar os pontos críticos da função de Lagrange F, isto é, encontrar $(w^*, \lambda^*) \in I\!R^m \times I\!R^n$ tal que

$$\nabla F(w^*, \lambda^*) = \begin{cases} \nabla_w F(w^*, \lambda^*) & = & 0_{I\!R^m} \\ \\ \nabla_\lambda F(w^*, \lambda^*) & = & 0_{I\!R^n} \end{cases}$$

Calculando a *primeira variação* da função de Lagrange F com relação à variável w, obtemos

$$\nabla_w F(w, \lambda) = w + A\lambda - b,$$

e calculando a *primeira variação* da função de Lagrange F com relação à variável λ, obtemos

$$\nabla_\lambda F(w, \lambda) = A^t w.$$

Desse modo, temos o seguinte

Problema de Ponto Sela: encontrar $w^* \in I\!R^m$ e $\lambda^* \in I\!R^n$ solução do sistema linear indefinido

$$\begin{cases} w + A\lambda = b \\ \\ A^t w = 0_{I\!R^n} \end{cases} \tag{8.206}$$

que na forma matricial é representado por:

$$\begin{bmatrix} I_m & A \\ \\ A^t & 0_n \end{bmatrix} \begin{bmatrix} w \\ \\ \lambda \end{bmatrix} = \begin{bmatrix} b \\ \\ 0_{I\!R^n} \end{bmatrix},$$

onde $I_m \in I\!M_m(I\!R)$ é a matriz identidade e $0_n \in I\!M_n(I\!R)$ é a matriz nula.

Finalmente, resolvendo o sistema linear indefinido (8.206), obtemos

$$\lambda^* = (A^t A)^{-1} A^t b \qquad \text{e} \qquad w^* = b - A(A^t A)^{-1} A^t b.$$

Podemos observar que $\lambda^* = (A^t A)^{-1} A^t b = x^*$ é a solução de quadrados mínimos do sistema linear sobredeterminado $Ax = b$, isto é, $\lambda^* = x^*$ é a solução do *Problema Primal*.

Exemplo 8.16.1 *Considere a matriz* $A \in I\!M_{4\times 3}(I\!R)$ *e o vetor* $b \in I\!R^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Encontre a projeção ortogonal do elemento $b \in I\!R^4$ *no subespaço* $\mathcal{N}(A^t)$ *e a solução de quadrados mínimos do sistema* $Ax = b$*, através do Problema de Ponto Sela.*

Neste exemplo, temos o seguinte

Problema de Ponto Sela: encontrar $w^* \in I\!R^4$ e $\lambda^* \in I\!R^3$ solução do sistema linear indefinido

$$\begin{cases} \quad w \;+\; A\lambda \;=\; b \\ A^t w \qquad\quad\; = \; 0_{I\!R^3} \end{cases} \tag{8.207}$$

que na forma matricial é representado por:

$$\begin{bmatrix} I_4 & A \\ A^t & 0_3 \end{bmatrix} \begin{bmatrix} w \\ \lambda \end{bmatrix} = \begin{bmatrix} b \\ 0_{I\!R^3} \end{bmatrix},$$

onde $I_4 \in I\!M_4(I\!R)$ é a matriz identidade e $0_3 \in I\!M_3(I\!R)$ é a matriz nula.

O elemento $w^* = b - A\lambda^*$ é a projeção ortogonal do elemento $b \in I\!R^4$ no subespaço $\mathcal{N}(A^t) = \mathcal{R}(A)^\perp$, e $\lambda^* = x^*$ é a solução de quadrados mínimos para o sistema $Ax = b$.

Por simplicidade, vamos indicar o sistema linear indefinido por $HX = Y$, onde os elementos $X, Y \in I\!R^7$ são dados por:

$$X = \begin{bmatrix} w \\ \lambda \end{bmatrix} \qquad \text{e} \qquad Y = \begin{bmatrix} b \\ 0_{I\!R^3} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

e a matriz $H \in I\!M_7(I\!R)$ é dada por:

$$H = \begin{bmatrix} I_4 & A \\ A^t & 0_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Finalmente, podemos obter uma solução numérica do sistema linear indefinido através da fatoração $H = QR$, onde $Q \in I\!M_7(I\!R)$ é uma matriz ortogonal e $R \in I\!M_7(I\!R)$ é uma matriz triangular superior, da seguinte forma:

$$HX = Y \iff QRX = Y \iff RX = Q^tY\,.$$

Desse modo, temos que resolver o sistema linear triangular superior $RX = Q^tY$, pelo processo de substituição atrasada, descrito no Algoritmo 2.9.2. A fatoração $H = QR$ pode ser obtida pelo Método de Gram–Schmidt Modificado, descrito no Algoritmo 8.13.2.

Fazendo uso dos procedimentos acima mencionados, obtemos a solução do sistema linear indefinido que é dada por:

$$X^* = \begin{bmatrix} w^* \\ \lambda^* \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 5 \\ -2 \\ -2 \end{bmatrix}$$

Desse modo, temos que a projeção ortogonal do elemento b sobre o subespaço $\mathcal{N}(A^t)$, e a solução de quadrados mínimos para o sistema linear $Ax = b$, são dados por:

$$w^* = \frac{1}{4} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \qquad \text{e} \qquad \lambda^* = x^* = \frac{1}{4} \begin{bmatrix} 5 \\ -2 \\ -2 \end{bmatrix} ,$$

respectivamente.

Quadrados Mínimos com Restrição

Considere o espaço vetorial real $I\!R^n$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Sejam $A \in I\!M_{m \times n}(I\!R)$, com $m > n$ e $posto(A) = n$, $b \in I\!R^m$, $B \in I\!M_{p \times n}(I\!R)$, com $p < n$ e $posto(B) = p$, e $d \in I\!R^p$. Considere o problema:

Encontrar uma solução de Quadrados Mínimos para o Sistema Linear Sobredeterminado $Ax = b$ **com a restrição** $Bx = d$.

Considere o funcional $J : I\!R^n \longrightarrow I\!R$ definido da seguinte forma:

$$J(x) = \frac{1}{2} \langle Ax - b , Ax - b \rangle = \frac{1}{2} \| Ax - b \|_2^2 .$$

Representamos nosso problema com o **Problema de Programação Quadrática**

$$\begin{cases} & \min\{ J(x) \} \\ & \\ \text{sujeito a} & Bx = d \end{cases}$$

Utilizando o **Método dos Multiplicadores de Lagrange**, associamos ao *Problema de Programação Quadrática com Restrição* um *Problema de Otimização Global*, considerando a função de Lagrange

$$F(x, \lambda) = \frac{1}{2} \langle Ax - b , Ax - b \rangle + \langle Bx - d , \lambda \rangle ,$$

onde $\lambda \in I\!R^p$ é o vetor dos multiplicadores de Lagrange.

Sendo assim, temos que determinar os pontos críticos da função de Lagrange, isto é, encontrar $(x^* , \lambda^*) \in I\!R^n \times I\!R^p$ tal que

$$\nabla F(x^*, \lambda^*) = \begin{cases} \nabla_x F(x^*, \lambda^*) = 0_{I\!R^n} \\ \\ \nabla_\lambda F(x^*, \lambda^*) = 0_{I\!R^p} \end{cases}$$

Calculando a *primeira variação* da função de Lagrange F com relação à variável x, obtemos

$$\nabla_x F(x, \lambda) = A^t A x + B^t \lambda - A^t b ,$$

e calculando a *primeira variação* da função de Lagrange F com relação à variável λ, obtemos

$$\nabla_\lambda F(x, \lambda) = Bx - d .$$

Desse modo, temos o seguinte

Problema de Ponto Sela: Encontrar $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^p$ solução do sistema linear indefinido

$$\begin{cases} A^tAx \; + \; B^t\lambda \; = \; A^tb \\ \\ Bx \; = \; d \end{cases} \tag{8.208}$$

que na forma matricial é representado por:

$$\begin{bmatrix} A^tA & B^t \\ & \\ B & 0_p \end{bmatrix} \begin{bmatrix} x \\ \\ \lambda \end{bmatrix} = \begin{bmatrix} A^tb \\ \\ d \end{bmatrix},$$

onde 0_p é a matriz nula de ordem $p \times p$. Como $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m > n$ e $posto(A) = n$, sabemos que A^tA é uma matriz positiva–definida.

Por simplicidade, vamos indicar o sistema linear indefinido por $HX = Y$, onde

$$X = \begin{bmatrix} x \\ \\ \lambda \end{bmatrix} \quad , \quad Y = \begin{bmatrix} A^tb \\ \\ d \end{bmatrix} \quad \text{e} \quad H = \begin{bmatrix} A^tA & B^t \\ & \\ B & 0_p \end{bmatrix}.$$

Nesta caso, a matriz do Problema de Ponto Sela H é uma matriz simétrica invertível, de ordem $n + m$, com n autovalores positivos e m autovalores negativos, isto é, H é uma matriz indefinida, de acordo com os resultados do Exemplo 6.7.8.

Finalmente, calculando a segunda variação da função de Lagrange F com relação às variáveis x e λ, obtemos a matriz Hessiana de F, que é a matriz H. Assim, o ponto crítico (x^*, λ^*) é um ponto de sela para a função de Lagrange.

Exemplo 8.16.2 *Considere a matriz* $A \in \mathbb{M}_{4\times 3}(\mathbb{R})$ *e o vetor* $b \in \mathbb{R}^4$ *dados por:*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Determinar a solução de quadrados mínimos para o sistema linear sobredeterminado $Ax = b$ *com a restrição* $x_1 + x_2 + x_3 = 0$*, isto é,* $Bx = d$*, onde*

$$B = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \quad , \quad d = \begin{bmatrix} 0 \end{bmatrix} \qquad e \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Por simplicidade, vamos indicar o sistema linear indefinido por $HX = Y$, onde os elementos $X, Y \in \mathbb{R}^4$ são dados por:

$$X = \begin{bmatrix} x \\ \lambda \end{bmatrix} \qquad \text{e} \qquad Y = \begin{bmatrix} A^t b \\ d \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 0 \end{bmatrix},$$

e a matriz $H \in \mathbb{M}_4(\mathbb{R})$ é dada por:

$$H = \begin{bmatrix} A^t A & B^t \\ B & 0_1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 & 1 \\ 2 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

Como $posto(A) = 3$, temos que $A^t A$ é uma matriz positiva–definida. Assim, sabemos que H é uma matriz indefinida.

Utilizando a fatoração $H = QR$, que pode ser obtida pelo Método de Gram–Schmidt Modificado, obtemos a solução do sistema linear indefinido, e por conseqüência a solução de quadrados mínimos com restrição, dadas por:

$$X^* = \frac{1}{3}\begin{bmatrix} 4 \\ -2 \\ -2 \\ 1 \end{bmatrix} \qquad \text{e} \qquad x^* = \frac{1}{3}\begin{bmatrix} 4 \\ -2 \\ -2 \end{bmatrix},$$

respectivamente. Note que o multiplicador de Lagrange $\lambda^* = \dfrac{1}{3}$.

Método de Uzawa

Considere o **Problema de Ponto Sela:** encontrar $x^* \in I\!R^n$ e $\lambda^* \in I\!R^m$ solução do sistema linear indefinido

$$\begin{cases} Ax + B^t\lambda = b \\ Bx = d \end{cases}$$

onde $A \in I\!M_n(I\!R)$ positiva–definida, $b \in I\!R^n$, $d \in I\!R^m$ e $B \in I\!M_{m\times n}(I\!R)$, com $m < n$ e $posto(B) = m$.

O **Método de Uzawa com Direções de Gradientes**[3] é largamente utilizado para obter uma solução numérica para o Problema de Ponto Sela, que descrevemos a seguir.

Sejam $\lambda_0 \in I\!R^m$ e $Ax_1 = b - B^t\lambda_0$

$$\begin{aligned} q_k &= d - Bx_k \\ p_k &= B^t q_k \\ Ah_k &= p_k \\ \alpha_k &= \frac{\langle q_k , q_k \rangle}{\langle p_k , h_k \rangle} \\ \lambda_k &= \lambda_{k-1} - \alpha_k q_k \\ x_{k+1} &= x_k + \alpha_k h_k \end{aligned}$$

para $k = 1, 2, 3, \cdots$.

Devemos obter uma solução numérica para o sistema linear positivo–definido

$$Ah_k = p_k ,$$

de maneira eficiente, o que pode ser feito através do Método dos Gradientes Conjugados, nos casos em que a matriz A é esparsa e de grande porte, ou pelo Método de Cholesky, caso contrário. Note que na aproximação inicial temos, também, de obter uma solução numérica para o sistema linear positivo–definido $Ax_1 = b - B^t\lambda_0$.

[3]Dietrich Braess, *Finite Elements*, Cambridge University Press, 1997.

Exemplo 8.16.3 *Sabe–se que uma alimentação diária equilibrada em vitaminas deve conter de* 950 *unidades de vitamina A,* 725 *unidades de vitamina B,* 625 *unidades de vitamina C,* 700 *unidades de vitamina D e* 850 *unidades de vitamina E.*

Com o objetivo de analisar como deve ser uma refeição equilibrada, foram estudados três alimentos. Fixada a mesma quantidade de 1.0 *grama de cada alimento, determinou–se que :*

(i) O alimento I contém 1 unidade de vitamina A, 1 *unidades de vitamina B,* 1 *unidade de vitamina C,* 2 *unidades de vitamina D e* 2 *unidades de vitamina E, e custa* $R\$\,15,00$ *por quilograma.*

(ii) O alimento II contém 2 *unidades de vitamina A,* 1 *unidade de vitamina B,* 1 *unidade de vitamina C,* 1 *unidade de vitamina D e* 1 *unidades de vitamina E, e custa* $R\$\,20,00$ *por quilograma.*

(iii) O alimento III contém 2 *unidades de vitamina A,* 2 *unidades de vitamina B,* 3 *unidades de vitamina C,* 1 *unidade de vitamina D e* 2 *unidades de vitamina E, e custa* $R\$\,10,00$ *por quilograma.*

Quantos gramas de cada um dos alimentos I, II e III devemos ingerir diariamente para que nossa alimentação seja a mais equilibrada possível e desejamos que cada refeição tenha 550 *gramas a um custo de* $R\$\,9,50$ *?*

Temos um **Problema de Quadrados Mínimos com Restrição**, onde

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \\ 2 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix} \quad , \quad b = \begin{bmatrix} 950 \\ 725 \\ 625 \\ 700 \\ 850 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 1.000 & 1.00 & 1.00 \\ 0.015 & 0.02 & 0.01 \end{bmatrix} \quad \text{e} \quad d = \begin{bmatrix} 550 \\ 9.50 \end{bmatrix},$$

cuja solução obtida pelo Método de Uzawa é dada por:

$$x^* = \begin{bmatrix} 167.86 \\ 316.07 \\ 66.07 \end{bmatrix},$$

que representa a quantidade em gramas de cada um dos alimentos I, II e III que devem compor uma refeição.

Considerando somente a solução de quadrados mínimos para o sistema linear $Ax = b$, obtemos a solução

$$x^* = \begin{bmatrix} 179.55 \\ 328.41 \\ 61.36 \end{bmatrix},$$

que representa a quantidade em gramas de cada um dos alimentos I, II e III que devem compor uma refeição mais equilibrada possível.

Neste caso, teremos uma refeição contendo 569.32 gramas a um custo de $R\$\,9,88$, que possui uma coerência com a solução do Problema de Quadrados Mínimos com Restrição, observando que a quantidade de cada um dos alimentos I, II e III tiveram uma modificação no sentido inverso ao custo.

Exemplo 8.16.4 *Considere o* ***Problema de Ponto Sela:*** *encontrar* $x^* \in \mathbb{R}^3$ *e* $\lambda^* \in \mathbb{R}^2$ *solução do sistema linear indefinido*

$$\begin{cases} Ax + B^t\lambda = b \\ Bx = d \end{cases}$$

onde $A \in \mathbb{M}_3(\mathbb{R})$ *é uma matriz positiva–definida,* $B \in \mathbb{M}_{2\times 3}(\mathbb{R})$*, com* $posto(B) = 2$*,* $b \in \mathbb{R}^3$ *e* $d \in \mathbb{R}^2$ *dadas por:*

$$A = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad , \quad b = \begin{bmatrix} -8 \\ -3 \\ -3 \end{bmatrix} \qquad e \qquad d = \begin{bmatrix} 3 \\ 0 \end{bmatrix} ,$$

utilizando o Método de Uzawa.

Sabemos que o Problema de Ponto Sela está associado ao seguinte

Problema de Programação Quadrática com Restrição

$$\begin{cases} \min\{\, J(x) \,\} \\ \text{sujeito a} \quad Bx = d \end{cases}$$

onde o funcional $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ é definido por:

$$J(x) = \frac{1}{2}\langle\, Ax \,,\, x \,\rangle - \langle\, b \,,\, x \,\rangle \,.$$

Utilizando o Método de Uzawa obtemos uma solução numérica com um erro relativo inferior à 10^{-12}, em 20 iterações, dada por:

$$x^* = \begin{bmatrix} 0.769230769230729 \\ -2.230769230769186 \\ 2.230769230769171 \end{bmatrix} \qquad \text{e} \qquad \lambda^* = \begin{bmatrix} -10.38461538461517 \\ 2.15384615384613 \end{bmatrix} .$$

O Método de Uzawa fica interessante quando A é uma matriz esparsa e de grande porte.

Exemplo 8.16.5 *Encontre a solução de norma–2 mínima do sistema linear*

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},$$

através do ***Problema de Ponto Sela:*** *encontrar* $x^* \in I\!R^4$ *e* $\lambda^* \in I\!R^3$ *solução do sistema linear indefinido*

$$\begin{cases} x + B^t\lambda = 0_{I\!R^4} \\ Bx = d \end{cases}$$

utilizando o Método de Uzawa, onde

$$B = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad e \quad d = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Para esse Problema de Ponto Sela, podemos escrever o Método de Uzawa da forma:

Sejam $\lambda_0 \in I\!R^m$ e $x_1 = -B^t\lambda_0$

$$\begin{aligned} q_k &= d - Bx_k \\ p_k &= B^t q_k \\ h_k &= p_k \\ \alpha_k &= \frac{\langle q_k \, , \, q_k \rangle}{\langle p_k \, , \, h_k \rangle} \\ \lambda_k &= \lambda_{k-1} - \alpha_k \, q_k \\ x_{k+1} &= x_k + \alpha_k \, h_k \end{aligned}$$

para $k = 1, 2, 3, \cdots$.

Utilizando o algoritmo de Uzawa, descrito acima, obtemos uma solução numérica, com um erro relativo inferior à 10^{-12} em 60 iterações, para a solução de norma–2 mínima x^* e para o vetor dos multiplicadores do Lagrange λ^*, dadas por:

$$x^* = \frac{1}{4} \begin{bmatrix} 3 \\ 5 \\ -1 \\ -3 \end{bmatrix} \quad \text{e} \quad \lambda^* = \frac{1}{4} \begin{bmatrix} 3 \\ -6 \\ -2 \end{bmatrix}.$$

Exercícios

Exercício 8.160 *Mostre que uma solução* $(x^*, \lambda^*) \in I\!R^n \times I\!R^m$ *do Problema de Ponto Sela definido em (8.202), é um* ***ponto sela*** *para a* ***Função de Lagrange***

$$F(x,\lambda) \;=\; \frac{1}{2}\langle\, Ax\,,\,x\,\rangle \;-\; \langle\, b\,,\,x\,\rangle \;+\; \langle\, Bx \,-\, d\,,\,\lambda\,\rangle\,,$$

isto é, para quaisquer $x \in I\!R^n$ *e* $\lambda \in I\!R^m$*, tem–se*

$$F(x^*,\lambda) \;\leq\; F(x^*,\lambda^*) \;\leq\; F(x,\lambda^*)\,,$$

equivalentemente

$$F(x^*,\lambda^*) \;=\; \min_{x}\,\max_{\lambda}\{\,F(x,\lambda)\,\}\,.$$

Exercício 8.161 *Seja a matriz* $A \in I\!M_{3\times 5}(I\!R)$ *e o elemento* $b \in I\!R^5$ *dados por:*

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 2 & -1 & 0 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Encontre uma aproximação para a projeção ortogonal do elemento b *no subespaço* $\mathcal{N}(A)$*, utilizando o método de Uzawa. Faça a implementação computacional em uma linguagem de sua preferência.*

Exercício 8.162 *Um agricultor deseja adubar a sua plantação e disponha de dois tipos diferentes de adubos, tipo A e Tipo B. O primeiro adubo do tipo A contém* $3.0\,g$ *de fósforo,* $1.0\,g$ *de nitrogênio e* $8.0\,g$ *de potássio, e custa* $R\$\,10,00$ *por quilograma. O segundo adubo do tipo B contém* $2.0\,g$ *de fósforo,* $3.0\,g$ *de nitrogênio e* $2.0\,g$ *de potássio, e custa* $R\$\,8,00$ *por quilograma.*

Sabemos que um quilograma de adubo dá para $10.0\,m^2$ *de terra, e que sua plantação necessita de pelo menos* $300.0\,g$ *de fósforo,* $150.0\,g$ *de nitrogênio e* $400.0\,g$ *de potássio a cada* $10.0\,m^2$.

Quanto o agricultor deve misturar de cada tipo de adubo para conseguir o efeito desejado se está disposto a gastar $R\$\,0,70$ *a cada* $10.0\,m^2$ *com a adubação ?*

Exercício 8.163 *Sejam a matriz* $A \in \mathbb{M}_{10}(\mathbb{R})$ *positiva–definida, o vetor* $b \in \mathbb{R}^{10}$ *e a matriz* $B \in \mathbb{M}_{2\times 10}(\mathbb{R})$, *de posto completo, dadas por:*

$$A = \begin{bmatrix} 16 & 4 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 25 & -3 & 6 & 0 & 0 & 0 & 10 & 0 & 0 \\ 4 & -3 & 30 & 8 & 0 & 2 & 0 & -9 & 0 & 0 \\ 0 & 6 & 8 & 47 & 8 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 17 & 3 & 2 & -1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 3 & 15 & 1 & -3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 21 & -4 & -7 & 0 \\ 0 & 10 & -9 & 0 & -1 & -3 & -4 & 52 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 3 & -7 & -1 & 14 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 24 \end{bmatrix} \quad , \quad b = \begin{bmatrix} 14.00 \\ -24.50 \\ 87.70 \\ 142.50 \\ 7.90 \\ 47.35 \\ -46.45 \\ 8.45 \\ -3.50 \\ 80.00 \end{bmatrix}$$

$$e \qquad B = \begin{bmatrix} 1 & -1 & 1 & -1 & 0 & 1 & 0 & 1 & 1 & -1 \\ -1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & -1 \end{bmatrix}.$$

Considere o funcional $J : \mathbb{R}^{10} \longrightarrow \mathbb{R}$ *definido da seguinte forma:*

$$J(x) = \frac{1}{2}\langle Ax , x \rangle - \langle b , x \rangle .$$

Determine uma solução numérica do ***Problema de Programação Quadrática***

$$\begin{cases} & \min\{ J(x) \} \\ \textit{sujeito a} & x \in \mathcal{N}(B) \end{cases}$$

através do Método de Uzawa.

8.17 Decomposição em Valores Singulares

Nessa seção vamos apresentar uma metodologia para a construção da Decomposição em Valores Singulares de uma matriz de ordem $m \times n$, fazendo conexões algébricas e geométricas com os resultados estudados até o momento, apresentaremos também suas aplicações em vários problemas de interesse prático. Para isso, vamos necessitar de alguns resultados que apresentamos a seguir.

Teorema 8.17.1 *Sejam* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ *e* v *um autovetor da matriz* A^tA *associado a um autovalor* λ *não–nulo. Então,* Av *é um autovetor da matriz* AA^t *associado ao mesmo autovalor* λ.

Demonstração – Temos que $A^tAv = \lambda v$. Assim, obtemos

$$(AA^t)(Av) = A(A^tAv) = A(\lambda v) = \lambda(Av).$$

Portanto, (λ, Av) é um autopar da matriz AA^t, o que completa a demonstração. ■

Corolário 8.17.1 *Seja* $A \in \mathbb{M}_{m \times n}(\mathbb{R})$. *Então, as matriz* A^tA *e* AA^t *possuem os mesmos autovalores não–nulos, contando com as suas multiplicidades.*

Demonstração – A prova segue imediata do Teorema 8.17.1. □

Teorema 8.17.2 *Sejam* v_1 *e* v_2 *autovetores ortogonais da matriz* A^tA*, associados a autovalores não–nulos. Então,* Av_1 *e* Av_2 *são autovetores ortogonais da matriz* AA^t.

Demonstração – Como v_1 e v_2 são autovetores ortogonais da matriz A^tA, associados a autovalores λ_1 e λ_2 não–nulos, respectivamente, temos que

$$\langle A^tAv_1 , A^tAv_2 \rangle = \lambda_1\lambda_2\langle v_1 , v_2 \rangle = 0.$$

Pelo Teorema 8.17.1, sabemos que Av_1 e Av_2 são autovetores da matriz AA^t associados aos autovalores λ_1 e λ_2 não–nulos, respectivamente. Desse modo, temos que

$$\langle A^tAv_1 , A^tAv_2 \rangle = \langle Av_1 , (AA^t)(Av_2) \rangle = \lambda_2\langle Av_1 , Av_2 \rangle = 0.$$

Por hipótese temos que $\lambda_2 \neq 0$, assim provamos que Av_1 e Av_2 são ortogonais, o que completa a demonstração. ■

Teorema 8.17.3 *Seja* $B \in \mathbb{M}_m(\mathbb{R})$ *uma matriz com um conjunto de* m *autovetores*

$$S = \{ v_1 , \cdots , v_m \}$$

linearmente independente em $\mathbb{R}^m$*, associados aos autovalores*

$$\lambda_1 , \cdots , \lambda_m ,$$

respectivamente, e supomos que $\lambda_1 , \cdots , \lambda_r$ *são não–nulos e que* $\lambda_{r+1} , \cdots , \lambda_m$ *são nulos. Então,* $\{ v_1, \cdots , v_r \}$ *é uma base para o subespaço* $\mathcal{R}(B)$*.*

Demonstração – Sabemos que $\{ v_1 , \cdots , v_m \}$ é uma base para $\mathbb{R}^m$. Logo, todo elemento $x \in \mathbb{R}^m$ é escrito de modo único como

$$x = \sum_{i=1}^{m} \alpha_i v_i .$$

Tomando $y \in \mathcal{R}(B)$, isto é, $y = Bx$ para algum $x \in \mathbb{R}^m$, temos que

$$y = Bx = B\left(\sum_{i=1}^{m} \alpha_i v_i \right) = \sum_{i=1}^{m} \alpha_i B v_i = \sum_{i=1}^{r} (\alpha_i \lambda_i) v_i .$$

Portanto, mostramos que $\{ v_1, \cdots , v_r \}$ é uma base para o subespaço $\mathcal{R}(B)$, o que completa a demonstração. ■

Assim sendo, provamos também que

$$posto(B) = dim(\mathcal{R}(B)) = r ,$$

onde r é o número de autovalores não–nulos da matriz B.

Além disso, podemos observar facilmente que os autovetores

$$v_{r+1} , \cdots , v_m ,$$

associados aos autovalores nulos, formam uma base para o subespaço $\mathcal{N}(B)$.

Sabemos que a matriz simétrica AA^t possui um conjunto de m autovetores linearmente independentes, e que a matriz simétrica A^tA possui um conjunto de n autovetores linearmente independentes. Assim, do Corolário 8.17.1 e do Teorema 8.17.3, apresentamos uma outra maneira para provar que

$$posto(AA^t) = posto(A^tA) ,$$

uma vez que AA^t e A^tA possuem o mesmo número de autovalores não–nulos, contando com suas multiplicidades.

Teorema 8.17.4 (Decomposição em Valores Singulares) *Seja* $A \in I\!M_{m \times n}(I\!R)$. *Então, existe uma matriz ortogonal* $U \in I\!M_m(I\!R)$*, uma matriz diagonal* $\Sigma \in I\!M_{m \times n}(I\!R)$ *da seguinte forma:*

$$\Sigma = diag(\sigma_1 , \cdots , \sigma_i , \cdots , \sigma_r , 0 , \cdots , 0) ,$$

com $\sigma_i > 0$*, e uma matriz ortogonal* $V \in I\!M_n(I\!R)$*, tais que* $A = U \Sigma V^t$.

Antes de passarmos à demonstração do Teorema 8.17.4, vamos fazer algumas observações interessantes relacionadas à Decomposição em Valores Singulares.

(1) No caso em que $A \in I\!M_n(I\!R)$ é uma matriz positiva–definida, temos a fatoração $A = Q \Lambda Q^t$, onde $Q \in I\!M_n(I\!R)$ é uma matriz ortogonal e $\Lambda \in I\!M_n(I\!R)$ é uma matriz diagonal, que podemos representar na forma:

$$Q = [q_1 \cdots q_j \cdots q_n] \qquad \text{e} \qquad \Lambda = diag(\lambda_i , \cdots , \lambda_j , \cdots , \lambda_n) ,$$

com (λ_j , q_j) um autopar da matriz A.

(2) Sabemos que $AA^t \in I\!M_m(I\!R)$ é uma matriz semipositiva–definida. Considerando a fatoração $A = U \Sigma V^t$, obtemos

$$AA^t = (U \Sigma V^t)(V \Sigma^t U^t) = U \Sigma \Sigma^t U^t ,$$

onde $\Sigma \Sigma^t \in I\!M_m(I\!R)$ é uma matriz diagonal da forma:

$$\Sigma \Sigma^t = diag(\sigma_1^2 , \cdots , \sigma_i^2 , \cdots , \sigma_r^2 , 0 , \cdots , 0) ,$$

e $U = [u_1 \cdots u_j \cdots u_m] \in I\!M_m(I\!R)$ é uma matriz ortogonal.

Desse modo, temos que $u_1 , \cdots , u_j , \cdots , u_m$ são autovetores da matriz AA^t associados aos autovalores $\sigma_1^2 , \cdots , \sigma_i^2 , \cdots , \sigma_r^2 , 0 , \cdots , 0$, respectivamente.

Logo,

$$\sigma_j = \sqrt{\lambda_j} \qquad \text{para} \qquad j = 1, \cdots , r ,$$

onde $\lambda_1, \cdots , \lambda_r$ são os autovalores não–nulos da matriz AA^t.

Sabemos que $A^tA \in \mathbb{M}_n(\mathbb{R})$ é uma matriz semipositiva–definida. De modo análogo, temos que $A^tA = V\Sigma^t\Sigma V^t$, onde $\Sigma^t\Sigma \in \mathbb{M}_n(\mathbb{R})$ é uma matriz diagonal da forma:

$$\Sigma^t\Sigma = diag(\sigma_1^2, \cdots, \sigma_i^2, \cdots, \sigma_r^2, 0, \cdots, 0),$$

e $V = [v_1 \cdots v_j \cdots v_n] \in \mathbb{M}_n(\mathbb{R})$ é uma matriz ortogonal.

Desse modo, temos que $v_1, \cdots, v_j, \cdots, v_n$ são autovetores da matriz A^tA associados aos autovalores $\sigma_1^2, \cdots, \sigma_i^2, \cdots, \sigma_r^2, 0, \cdots, 0$, respectivamente.

Logo, $\sigma_j = \sqrt{\lambda_j}$ para $j = 1, \cdots, r$, onde $\lambda_1, \cdots, \lambda_r$ são os autovalores não–nulos da matriz A^tA.

Já mostramos que

(3) Pelo Teorema 8.10.6, temos que $\mathcal{N}(AA^t) = \mathcal{N}(A^t)$.

(4) Pelo Teorema 8.10.3, temos que $\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp = \mathcal{R}(A) \oplus \mathcal{N}(A^t)$.

(5) Pelo Teorema 8.17.3, temos que $\mathcal{R}(AA^t) = [u_1, \cdots, u_r]$.

Assim, temos que $u_1, \cdots, u_r$ formam uma base ortonormal para o subespaço $\mathcal{R}(A)$, e $u_{r+1}, \cdots, u_m$ formam uma base ortonormal para o subespaço $\mathcal{N}(A^t)$.

De modo análogo, mostramos que

(6) Pelo Teorema 8.10.5, temos que $\mathcal{N}(A^tA) = \mathcal{N}(A)$.

(7) Pelo Teorema 8.10.3, temos que $\mathbb{R}^n = \mathcal{R}(A^t) \oplus \mathcal{R}(A^t)^\perp = \mathcal{R}(A^t) \oplus \mathcal{N}(A)$.

(8) Pelo Teorema 8.17.3, temos que $\mathcal{R}(A^tA) = [v_1, \cdots, v_r]$.

Assim, temos que $v_1, \cdots, v_r$ formam uma base ortonormal para o subespaço $\mathcal{R}(A^t)$, e $v_{r+1}, \cdots, v_n$ formam uma base ortonormal para o subespaço $\mathcal{N}(A)$.

Finalmente, vamos passar para a demonstração do Teorema 8.17.4, que é feita de forma construtiva.

Como $A^tA \in I\!M_n(I\!R)$ é uma matriz simétrica, sabemos que possui um sistema completo de autovetores ortonormais

$$v_1 \, , \, \cdots \, , \, v_j \, , \, \cdots \, , \, v_n \, ,$$

associados aos autovalores $\lambda_1 \, , \, \cdots \, , \, \lambda_j \, , \, \cdots \, , \, \lambda_n$, respectivamente, e que formam uma base ortonormal para $I\!R^n$.

Vamos construir uma matriz ortogonal $V = [v_1 \cdots v_j \cdots v_n] \in I\!M_n(I\!R)$. Desse modo, temos que

$$v_j^t(\, A^tAv_j \,) \; = \; v_j^t(\, \lambda_j v_j \,) \; = \; \lambda_j v_j^t v_j \; = \; \lambda_j \qquad \Longrightarrow \qquad \| \, Av_j \, \|_2^2 \; = \; \lambda_j \; \geq \; 0$$

para $j = 1, 2, \cdots, n$.

Supomos que $\lambda_1, \cdots, \lambda_r$ são os autovalores não–nulos, e que $\lambda_{r+1}, \cdots, \lambda_n$ são os autovalores nulos. Logo, temos que

$$Av_j \; = \; 0_{I\!R^m} \qquad \text{para} \qquad j \; = \; (r+1), \cdots n \, .$$

Para os autovalores positivos, $\lambda_1, \cdots, \lambda_r$, definimos

$$\sigma_j \; = \; \sqrt{\lambda_j} \qquad \text{e} \qquad u_j \; = \; \frac{Av_j}{\sigma_j}$$

para $j = 1, 2, \cdots, r$.

Os elementos $u_1, \cdots, u_r$ são ortonormais em $I\!R^m$. De fato,

$$u_i^t u_j \;\; = \;\; \left(\frac{Av_i}{\sigma_i} \right)^t \left(\frac{Av_j}{\sigma_j} \right) \;\; = \;\; \frac{v_i^t(\, A^tAv_j \,)}{\sigma_i \sigma_j} \;\; = \;\; \frac{v_i^t(\, \lambda_j v_j \,)}{\sigma_i \sigma_j} \;\; = \;\; \frac{\lambda_j v_i^t v_j}{\sigma_i \sigma_j} \; = \; \delta_{ij} \, ,$$

o que prova o resultado desejado.

Desse modo, podemos completar o conjunto ortonormal $\{ \, u_1, \cdots, u_r \, \}$ para obter uma base ortonormal $\{ \, u_1, \cdots, u_r, u_{r+1}, \cdots, u_m \, \}$ para o $I\!R^m$.

Assim, podemos construir uma matriz ortogonal $U \in I\!M_m(I\!R)$ da forma:

$$U \; = \; [u_1 \cdots u_r \; u_{r+1} \cdots u_m] \; \in \; I\!M_m(I\!R) \, .$$

Finalmente, obtemos a **Decomposição em Valores Singulares** da matriz $A \in I\!M_{m \times n}(I\!R)$ na forma:

$$A = U \Sigma V^t \qquad \Longleftrightarrow \qquad U^t A V = \Sigma = diag(\sigma_1 , \cdots , \sigma_r , 0 , \cdots , 0) \,.$$

Os escalares positivos $\sigma_1 , \cdots , \sigma_r$ são denominados **Valores Singulares** da matriz A.

De fato, denotando $\Sigma = [u_i^t A v_j]$, sabemos que, para $i = 1, \cdots , m$,

$$u_i^t A v_j = 0_{I\!R} \qquad \text{para} \qquad j = (r+1), \cdots , n \,,$$

desde que $Av_j = 0_{I\!R^m}$ para $j = (r+1), \cdots , n$, e que

$$u_i^t A v_j = u_i^t (\sigma_j u_j) = \sigma_j u_i^t u_j \qquad \text{para} \qquad j = 1, \cdots , r \,,$$

desde que $Av_j = \sigma_j u_j$ para $j = 1, \cdots , r$.

Logo, para $i = 1, \cdots , m$, temos que

$$u_i^t A v_j = \sigma_j u_i^t u_j = \sigma_j \delta_{ij} \qquad \text{para} \qquad j = 1, \cdots , r \,.$$

Portanto, obtemos

$$U^t A V = \Sigma = \begin{bmatrix} \widehat{\Sigma} & 0_{r \times q} \\ 0_{p \times r} & \widehat{D} \end{bmatrix} \,,$$

onde $\widehat{\Sigma}$ é uma matriz diagonal de ordem $r \times r$ dada por:

$$\widehat{\Sigma} = diag(\sigma_1 , \cdots , \sigma_r) \,,$$

e $\widehat{D}$ é uma matriz diagonal nula de ordem $p \times q$, com $p = m - r$ e $q = n - r$.

Podemos observar facilmente que

$$posto(A) = r \leq \min\{ m , n \} \,.$$

Assim, completamos a demonstração do Teorema 8.17.4. ∎

Exemplo 8.17.1 *Encontre a Decomposição em Valores Singulares da matriz*

$$A = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix},$$

detalhando todos os passos.

Inicialmente vamos calcular os autovalores e autovetores da matriz A^tA, que é uma matriz positiva–definida, desde que $posto(A) = 2$. O polinômio característico da matriz A^tA é dado por:

$$p(\lambda) = \det(A^tA - \lambda I) = \begin{vmatrix} 2 - \lambda & -1 \\ -1 & 2 - \lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3 .$$

Assim, seus autovalores são $\lambda_1 = 3$ e $\lambda_2 = 1$, e podemos escolher os seguintes autovetores ortonormais associados

$$v_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad v_2 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

que é uma base ortonormal para o subespaço $\mathcal{R}(A^t)$.

Logo, a matriz ortogonal $V \in I\!M_2(I\!R)$ é dada por:

$$V = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Portanto os valores singulares da matriz A são

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3} \quad \text{e} \quad \sigma_2 = \sqrt{\lambda_2} = 1 .$$

Assim, a matriz diagonal $\Sigma \in I\!M_{3\times 2}(I\!R)$ é dada por:

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Finalmente, vamos construir a matriz ortogonal $U = [u_1\, u_2\, u_3] \in I\!M_3(I\!R)$. Os elementos u_1 e u_2 são escolhidos da forma:

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{\sqrt{6}}{6}\begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} \quad \text{e} \quad u_2 = \frac{Av_2}{\sigma_2} = \frac{\sqrt{2}}{2}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Podemos observar que $\{\, u_1, u_2 \,\}$ é uma base ortonormal para o subespaço $\mathcal{R}(A)$.

Escolhemos o elemento u_3 de modo que $\{ u_3 \}$ seja uma base ortonormal para o subespaço $\mathcal{N}(A^t)$, isto é, $\{ u_1, u_2, u_3 \}$ é uma base ortonormal para $\mathbb{R}^3$.

Para isso, temos que obter a solução geral do sistema linear homogêneo

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

que é dada por:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = c \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \text{para} \qquad c \in \mathbb{R}.$$

Assim, o elemento u_3 é dado por:

$$u_3 = \frac{\sqrt{3}}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Desse modo, a matriz ortogonal $U \in \mathbb{M}_3(\mathbb{R})$ é dada por:

$$U = \begin{bmatrix} -\frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix},$$

Portanto, temos a Decomposição em Valores Singulares $A = U \Sigma V^t$.

Exemplo 8.17.2 *Encontre a Decomposição em Valores Singulares da matriz*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix},$$

detalhando todos os passos.

Inicialmente vamos calcular os autovalores e autovetores da matriz A^tA, que é uma matriz positiva–definida, desde que $posto(A) = 2$. O polinômio característico da matriz A^tA é dado por:

$$p(\lambda) = \det(A^tA - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3\,.$$

Assim, seus autovalores são $\lambda_1 = 3$ e $\lambda_2 = 1$, e podemos escolher os seguintes autovetores ortonormais associados

$$v_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{e} \quad v_2 = \frac{\sqrt{2}}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

que é uma base ortonormal para o subespaço $\mathcal{R}(A^t)$.

Logo, a matriz ortogonal $V \in I\!M_2(I\!R)$ é dada por:

$$V = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Portanto os valores singulares da matriz A são

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3} \quad \text{e} \quad \sigma_2 = \sqrt{\lambda_2} = 1\,.$$

Assim, a matriz diagonal $\Sigma \in I\!M_{4\times 2}(I\!R)$ é dada por:

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Finalmente, construímos a matriz ortogonal $U = [u_1\, u_2\, u_3\, u_4] \in I\!M_4(I\!R)$. Os elementos u_1 e u_2 são escolhidos da forma:

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{\sqrt{6}}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix} \quad \text{e} \quad u_2 = \frac{Av_2}{\sigma_2} = \frac{\sqrt{2}}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Podemos observar que $\{ u_1, u_2 \}$ é uma base ortonormal para o subespaço $\mathcal{R}(A)$.

Escolhemos os elementos u_3 e u_4 de modo que $\{ u_3, u_4 \}$ seja uma base ortonormal para o subespaço $\mathcal{N}(A^t)$, isto é, $\{ u_1, u_2, u_3, u_4 \}$ é uma base ortonormal para $\mathbb{R}^4$.

Para isso, temos que obter a solução geral do sistema linear homogêneo

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} ,$$

que é dada por:

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = c \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad \text{para} \qquad c, d \in \mathbb{R} .$$

Assim, os elementos u_3 e u_4 são dados por:

$$u_3 = \frac{\sqrt{3}}{3} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{e} \qquad u_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} .$$

Desse modo, a matriz ortogonal $U \in \mathbb{M}_4(\mathbb{R})$ é dada por:

$$U = \begin{bmatrix} \dfrac{\sqrt{6}}{6} & -\dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{3}}{3} & 0 \\[2ex] \dfrac{\sqrt{6}}{6} & \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{3}}{3} & 0 \\[2ex] \dfrac{\sqrt{6}}{3} & 0 & \dfrac{\sqrt{3}}{3} & 0 \\[2ex] 0 & 0 & 0 & 1 \end{bmatrix} ,$$

Portanto, temos a Decomposição em Valores Singulares $A = U \Sigma V^t$.

Aplicações da Decomposição em Valores Singulares

Neste momento estamos preparados para apresentar várias aplicações da Decomposição em Valores Singulares fazendo uma conexão entre os diversos temas estudados neste texto. Além disso, apresentamos uma maneira simples e elegante para obter a *Solução Ótima de Quadrados Mínimos* para um sistema linear $Ax = b$, que é uma generalização dos resultados até agora apresentados sobre o tema.

Teorema 8.17.5 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,*

$$\| A \|_2 = \sigma_{max} ,$$

onde σ_{max} *é o maior valor singular da matriz* A.

Demonstração – A prova pode ficar a cargo do leitor. □

Corolário 8.17.2 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,* $\| A \|_2 = \| A^t \|_2$.

Teorema 8.17.6 *Seja* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. *Então,*

$$\| A \|_F = \sqrt{\sum_{j=1}^{r} \sigma_j} ,$$

onde $\sigma_1, \cdots, \sigma_r$ *são os valores singulares da matriz* A.

Demonstração – A prova pode ficar a cargo do leitor. □

Teorema 8.17.7 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$. *Então, a matriz* A *pode ser fatorada na forma:*

$$A = QS ,$$

denominada ***Decomposição Polar***, *onde* $Q \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz ortogonal e* $S \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz semipositiva–definida.*

Demonstração – Considerando a Decomposição em Valores Singulares da Matriz A, obtemos

$$A = U\Sigma V^t = U(V^t V)\Sigma V^t = (UV^t)(V\Sigma V^t) = QS ,$$

onde $Q = UV^t \in \mathbb{M}_n(\mathbb{R})$ é uma matriz ortogonal e $S = V\Sigma V^t \in \mathbb{M}_n(\mathbb{R})$ é uma matriz semipositiva–definida, o que completa a demonstração. ■

Exemplo 8.17.3 *Seja* $A = QS$ *a Decomposição Polar da matriz* $A \in \mathbb{M}_n(\mathbb{R})$*, com* A *uma matriz invertível. Mostre que* S *é uma matriz positiva–definida.*

Proposição 8.17.1 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *positiva–definida. Então, existe uma matriz invertível* $R \in \mathbb{M}_n(\mathbb{R})$ *tal que* $A = R^t R$*, denominada* ***raiz quadrada*** *de* A.

Demonstração – A prova será feita através dos Exemplos a seguir. ■

Exemplo 8.17.4 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *semipositiva–definida, e* $G \in \mathbb{M}_n(\mathbb{R})$ *seu fator de Cholesky, isto é,* $A = G^t G$*. Considere a fatoração em Valores Singulares da Matriz* G*, isto é,* $G = U \Sigma V^t$*, e a matriz semipositiva–definida* $X = V \Sigma V^t$*. Desse modo, temos que*

$$A = G^t G = (V \Sigma U^t)(U \Sigma V^t) = V \Sigma^2 V^t = (V \Sigma V^t)(V \Sigma V^t) = X^2 .$$

Portanto, X *é a* ***raiz quadrada semipositiva–definida*** *da matriz* A*. No caso em que* A *é uma matriz positiva–definida,* X *é a sua* ***raiz quadrada positiva–definida****.*

Exemplo 8.17.5 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz semipositiva–definida, e consideramos sua fatoração na forma* $A = Q \Lambda Q^t$*, onde* $Q \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz ortogonal e* $\Lambda \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz diagonal. Tomando a matriz semipositiva–definida*

$$R = Q \sqrt{\Lambda} Q^t ,$$

obtemos

$$R^2 = (Q \sqrt{\Lambda} Q^t)(Q \sqrt{\Lambda} Q^t) = Q \Lambda Q^t = A .$$

Portanto, R *é a* ***raiz quadrada semipositiva–definida*** *da matriz* A*. No caso em que* A *é uma matriz positiva–definida,* R *é a sua* ***raiz quadrada positiva–definida****.*

Exemplo 8.17.6 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz positiva–definida, e consideramos sua fatoração na forma* $A = Q \Lambda Q^t$*, onde* $Q \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz ortogonal e* $\Lambda \in \mathbb{M}_n(\mathbb{R})$ *é uma matriz diagonal. Tomando a matriz invertível*

$$\widehat{R} = \sqrt{\Lambda} Q^t ,$$

obtemos

$$\widehat{R}^t \widehat{R} = (\sqrt{\Lambda} Q^t)^t (\sqrt{\Lambda} Q^t) = Q \Lambda Q^t = A .$$

Portanto, $\widehat{R}$ *é a* ***raiz quadrada*** *da matriz* A.

Solução Ótima de Quadrados Mínimos

Sejam $A \in I\!M_{m\times n}(I\!R)$, com $posto(A) = r \leq \min\{m, n\}$, e o elemento $b \in I\!R^m$. Considere os seguintes problemas equivalentes

(M) Problema de Quadrados Mínimos: encontrar $x^* \in I\!R^n$ tal que

$$\| Ax^* - b \|_2 = \min\{ \| Ax - b \|_2 \quad , \quad x \in I\!R^n \} .$$

(SN) Sistema Normal: Encontrar $x^* \in I\!R^n$ solução do sistema linear

$$A^t Ax = A^t b .$$

Temos que x^* é a solução de quadrados mínimos para o sistema linear $Ax = b$. Entretanto, considerando a possibilidade de que $posto(A) = r \leq \min\{m, n\}$, o problema **(M)** pode possuir infinitas soluções. Vamos indicar por $\mathbb{X}_{LS}$ o conjunto solução do problema **(M)**, isto é,

$$\mathbb{X}_{LS} = \{ x^* \in I\!R^n \ / \ \| Ax^* - b \|_2 \leq \| Ax - b \|_2 \quad , \quad x \in I\!R^n \} .$$

Assim sendo, vamos definir o elemento $x^+ \in \mathbb{X}_{LS}$ tal que

$$\| x^+ \|_2 \leq \| x^* \|_2 \qquad \text{para todo} \qquad x^* \in \mathbb{X}_{LS} ,$$

como a **Solução Ótima de Quadrados Mínimos** para o sistema linear $Ax = b$.

Finalmente, vamos mostrar como utilizar a *Decomposição em Valores Singulares*

$$A = U \Sigma V^t$$

para obtermos a *Solução Ótima de Quadrados Mínimos* para o sistema linear $Ax = b$.

Desse modo, temos que

$$\| Ax - b \|_2 = \| U\Sigma V^t x - b \|_2 = \| U^t (U\Sigma V^t x - b) \|_2 = \| \Sigma V^t x - U^t b \|_2 .$$

Chamando $y = V^t x$ e $d = U^t b$, obtemos

$$\| Ax - b \|_2 = \| \Sigma y - d \|_2 ,$$

que de acordo com os nossos objetivos, podemos considerar

$$\| Ax - b \|_2^2 = \| \Sigma y - d \|_2^2 = \sum_{i=1}^{r} (\sigma_i \, y_i - d_i)^2 + \sum_{i=r+1}^{m} d_i^2 .$$

Como nosso objetivo é minimizar $\| \Sigma y - d \|_2^2$, e o segundo termo da expressão da direita não depende da variável y, basta impor a seguinte condição

$$\sum_{i=1}^{r} (\sigma_i y_i - d_i)^2 = 0 ,$$

que obtemos

$$y_i = \frac{d_i}{\sigma_i} \qquad \text{para} \qquad i = 1, \cdots , r ,$$

e como as variáveis y_i para $i = (r+1), \cdots , n$ estão livres, podemos escolher

$$y_i = 0 \qquad \text{para} \qquad i = (r+1), \cdots , n .$$

Assim, encontramos um elemento $y^+ \in I\!R^n$ dado por:

$$y^+ = \begin{bmatrix} \dfrac{d_1}{\sigma_1} \\ \vdots \\ \dfrac{d_r}{\sigma_r} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

que minimiza $\| \Sigma y - d \|_2^2$ e que possui a menor norma–2 possível, desde que escolhemos as variáveis livres todas nulas.

Portanto, a **solução ótima de quadrados mínimos** para o sistema linear $Ax = b$ é dada por $x^+ = V y^+$, isto é,

$$\| x^+ \|_2 \leq \| x^* \|_2 \qquad \text{para todo} \qquad x^* \in \mathbb{X}_{SL} ,$$

onde as infinitas **soluções de quadrados mínimos** são $x^* = V y^*$, onde

$$y^* = \begin{bmatrix} \dfrac{d_1}{\sigma_1} \\ \vdots \\ \dfrac{d_r}{\sigma_r} \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix}$$

com as variáveis livres y_i para $i = (r+1), \cdots , n$.

Considerando a matriz $\Sigma \in \mathbb{M}_{m \times n}(\mathbb{R})$ dada por:

$$\Sigma = \begin{bmatrix} \widehat{\Sigma} & 0_{r \times q} \\ 0_{p \times r} & \widehat{D} \end{bmatrix},$$

onde $\widehat{\Sigma}$ é uma matriz diagonal de ordem $r \times r$ dada por:

$$\widehat{\Sigma} = diag(\sigma_1, \cdots, \sigma_r),$$

e $\widehat{D}$ é uma matriz nula de ordem $p \times q$, com $p = m - r$ e $q = n - r$.

Definimos a matriz $\Sigma^\dagger \in \mathbb{M}_{n \times m}(\mathbb{R})$ da seguinte forma:

$$\Sigma^\dagger = \begin{bmatrix} \widehat{\Sigma}^\dagger & 0_{r \times p} \\ 0_{q \times r} & \widehat{D}^t \end{bmatrix},$$

onde $\widehat{\Sigma}^\dagger$ é uma matriz diagonal de ordem $r \times r$ dada por:

$$\widehat{\Sigma}^\dagger = diag\left(\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_r}\right),$$

e $\widehat{D}^t$ é uma matriz diagonal nula de ordem $q \times p$. Podemos verificar facilmente que a matriz $\Sigma^\dagger \in \mathbb{M}_{n \times m}(\mathbb{R})$ é a pseudo–inversa da matriz Σ.

Assim, podem os escrever a solução ótima de quadrados mínimos da forma:

$$x^+ = (V\widehat{\Sigma}^\dagger U^t)b = A^\dagger b,$$

onde a matriz $A^\dagger = V\widehat{\Sigma}^\dagger U^t \in \mathbb{M}_{n \times m}(\mathbb{R})$ é a **pseudo–inversa** da matriz A, isto é, a matriz $A^\dagger$ satisfaz as seguintes propriedades

1. $(AA^\dagger)^t = AA^\dagger$.

2. $(A^\dagger A)^t = A^\dagger A$.

3. $AA^\dagger A = A$.

4. $A^\dagger A A^\dagger = A^\dagger$.

que são as propriedades de Moore–Penrose.

Teorema 8.17.8 *Seja* $A \in I\!M_{m \times n}(I\!R)$. *Então,*

$$\| A^{\dagger} \|_2 = \frac{1}{\sigma_{min}} ,$$

onde σ_{min} *é o menor valor singular da matriz* A.

Demonstração – A prova pode ficar a cargo do leitor. □

Definição 8.17.1 *Seja* $A \in I\!M_{m \times n}(I\!R)$. *Definimos o* ***número de condição*** *da matriz* A *com relação à norma* $\| \cdot \|_2$ *na forma:*

$$\mathcal{K}_2(A) = \| A \|_2 \, \| A^{\dagger} \|_2 = \frac{\sigma_{max}}{\sigma_{min}} .$$

Assim, podemos fazer uma análise de sensibilidade para Problemas de Quadrados Mínimos, que é uma generalização do que foi apresentado até o momento para sistema linear cuja matriz é invertível.

Vamos apresentar uma forma interessante, e importante, da aplicação da Decomposição em Valores Singulares em Problemas de Quadrados Mínimos. Considere uma matriz $A \in \mathbb{M}_{m \times n}(\mathbb{R})$, com $m > n$ e $posto(A) = n$, e um elemento $b \in \mathbb{R}^m$. Vamos utilizar a Decomposição em Valores Singulares $A = U \Sigma V^t$ para obter uma *Solução Ótima de Quadrados Mínimos* para o sistema linear sobredeterminado $Ax = b$, fazendo uma redução do posto da matriz A.

Para isso, consideramos uma precisão $\varepsilon > 0$ e ordenamos os valores singulares da matriz A da seguinte forma:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{r(\varepsilon)} > \varepsilon \geq \cdots \geq \sigma_n \, .$$

Vamos escolher uma matriz A_ε de modo que $posto(A_\varepsilon) = r(\varepsilon) < n$, utilizando a decomposição $A = U \Sigma V^t$, da seguinte forma:

$$A_\varepsilon = U \Sigma_\varepsilon V^t$$

onde a matriz Σ_ε é dada da seguinte forma:

$$\Sigma_\varepsilon = \begin{bmatrix} \widehat{\Sigma}_\varepsilon & 0_{r(\varepsilon) \times q} \\ 0_{p \times r(\varepsilon)} & \widehat{D} \end{bmatrix} ,$$

onde $\widehat{\Sigma}_\varepsilon$ é uma matriz diagonal de ordem $r(\varepsilon) \times r(\varepsilon)$ dada por:

$$\widehat{\Sigma} = diag(\sigma_1 \, , \, \cdots \, , \, \sigma_{r(\varepsilon)}) \, ,$$

e $\widehat{D}$ é uma matriz diagonal nula de ordem $p \times q$, com $p = m - r(\varepsilon)$ e $q = n - r(\varepsilon)$.

Portanto, a **solução ótima de quadrados mínimos** para o sistema linear $A_\varepsilon x = b$ é dada por $x_\varepsilon^+ = V y_\varepsilon^+$, onde

$$y_\varepsilon^+ = \begin{bmatrix} \dfrac{d_1}{\sigma_1} \\ \vdots \\ \dfrac{d_{r(\varepsilon)}}{\sigma_{r(\varepsilon)}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} ,$$

isto é, podemos escrever $x_\varepsilon^+ = V \, \Sigma_\varepsilon^\dagger \, U^t \, b$.

Verificamos facilmente que

$$\| A - A_\varepsilon \|_2 < \varepsilon ,$$

o que justifica a escolha da solução ótima de quadrados mínimos, em vez da solução de quadrados mínimos para o sistema linear sobredeterminado $Ax = b$ que é dada por:

$$x^* = V y^* ,$$

onde y^* é dado por:

$$y^* = \begin{bmatrix} \dfrac{d_1}{\sigma_1} \\ \vdots \\ \dfrac{d_{r(\varepsilon)}}{\sigma_{r(\varepsilon)}} \\ \vdots \\ \dfrac{d_n}{\sigma_n} \end{bmatrix} ,$$

desde que $posto(A) = n$.

Assim, a solução de quadrados mínimos x^* é representada da seguinte forma:

$$x^* = V \Sigma^\dagger U^t b ,$$

que não apresenta vantagem em relação à solução obtida através da fatoração QR.

Exemplo 8.17.7 *Determine a solução ótima de quadrados mínimos para* $Ax = b$, *onde*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

Exercícios

Exercício 8.164 *Determine a Decomposição em Valores Singulares das matrizes*

$$(a)\ A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad , \quad (b)\ A = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad , \quad (c)\ A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \\ 0 & 0 \end{bmatrix} ,$$

detalhando todos os passos.

Exercício 8.165 *Determine a Decomposição em Valores Singulares das matrizes*

$$(a)\ A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \quad , \quad (b)\ A = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix} ,$$

detalhando todos os passos.

Exercício 8.166 *Seja* $A \in I\!M_{m\times n}(I\!R)$. *Mostre que*

$$\| A \|_2 = \sigma_{max} ,$$

onde σ_{max} *é o maior valor singular da matriz* A.

Exercício 8.167 *Seja* $A \in I\!M_{m\times n}(I\!R)$. *Mostre que*

$$\| A^{\dagger} \|_2 = \frac{1}{\sigma_{min}} ,$$

onde σ_{min} *é o menor valor singular da matriz* A.

Exercício 8.168 *Determine a solução ótima de quadrados mínimos para* $Ax = b$, *onde*

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad e \qquad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix} .$$

Exercício 8.169 *Seja* $Q \in I\!M_{m\times n}(I\!R)$ *uma matriz ortogonal. Determine a matriz pseudo–inversa da matriz* Q.

Exercício 8.170 *Dada a matriz* $A \in I\!M_2(I\!R)$ *e sua Decomposição Polar* $A = QS$, *onde*

$$A = \begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix} \quad , \quad Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad e \qquad S = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} .$$

Determine a Decomposição em Valores Singulares $A = U\Sigma V^t$

Exercício 8.171 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$*, com* $posto(A) = r \leq \min\{m, n\}$*, e* $b \in \mathbb{R}^m$*. Considere os seguintes problemas:*

(M) Problema de Quadrados Mínimos: *Encontrar* $x^* \in \mathbb{R}^n$ *tal que*

$$\| Ax^* - b \|_2^2 \;=\; \min\{\; \| Ax - b \|_2^2 \quad ; \quad x \in \mathbb{R}^n \;\}$$

que é equivalente ao problema

(SN) Sistema Normal: *Encontrar* $x^* \in \mathbb{R}^n$ *solução do sistema linear*

$$A^t A x \;=\; A^t b \,.$$

Pede–se:

(a) *Faça uma análise do conjunto solução* $\mathbb{X}_{LS}$ *do problema* ***(M)*** $\Longleftrightarrow$ ***(SN)****.*

(b) *Determine a caracterização do elemento* $x^+ \in \mathbb{X}_{LS}$ *tal que*

$$\| x^+ \|_2 \;\leq\; \| x^* \|_2 \qquad \text{para todo} \qquad x^* \in \mathbb{X}_{LS} \,,$$

denominado ***Solução Ótima de Quadrados Mínimos*** *para* $Ax = b$*.*

(c) *Dê a interpretação geométrica para o elemento* $z^+ \in \mathbb{R}^m$ *definido por:*

$$z^+ \;=\; Ax^+ \,,$$

para cada uma das situações analisadas no item (a).

(d) *Considerando*

$$A \;=\; \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ -1 & 1 & -1 & 1 \end{bmatrix} \qquad e \qquad b \;=\; \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix} ,$$

determine a solução ótima de quadrados mínimos para $Ax = b$*.*

(e) *Considerando o resultado do item (d), tome o elemento* w^+ *definido por:*

$$w^+ \;=\; b \;-\; z^+ \,,$$

onde $z^+ = Ax^+$ *e* x^+ *é a solução ótima de quadrados mínimos para o sistema linear* $Ax = b$*. Verifique que o elemento* $w^+ \in \mathcal{N}(A^t)$ *é não–nulo e dê uma interpretação geométrica para os elementos* z^+ *e* w^+*. Justifique esse resultado.*

Exercício 8.172 *Sejam* $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, *com* $posto(A) = r \leq \min\{m, n\}$, *e* $\sigma_1, \cdots, \sigma_r$ *os valores singulares da matriz* A. *Pede–se:*

(a) *Mostre que a norma de Frobenius da matriz* A *é dada por:*

$$\| A \|_F^2 = \sigma_1^2 + \cdots + \sigma_r^2$$

(b) *Para* $m > n$ *e* $posto(A) = n$, *mostre que o subconjunto*

$$S = \{ Ax \in \mathbb{R}^m \; / \; \| x \|_2 = 1 \quad ; \quad x \in \mathbb{R}^n \}$$

é um hiper–elipsóide em $\mathbb{R}^m$, *determinando as direções dos semi–eixos e seus respectivos comprimentos.*

(c) *Utilizando a Decomposição em Valores Singulares da matriz* A, *mostre que para todo* $\epsilon > 0$, *existe uma matriz* $A_\epsilon \in \mathbb{M}_{m\times n}(\mathbb{R})$ *de posto completo tal que*

$$\| A - A_\epsilon \|_2 < \epsilon .$$

(d) *Considerando a fatoração* $A = U\Sigma V^t$, *mostre que*

$$A = \sum_{j=1}^{r} \sigma_j \, u_j \, v_j^t \qquad e \qquad A^\dagger = \sum_{j=1}^{r} \frac{1}{\sigma_j} \, v_j \, u_j^t ,$$

onde $A^\dagger$ *é a pseudo–inversa da matriz* A.

(e) *Mostre que a matriz* $P = A A^\dagger$ *é a matriz de projeção ortogonal em* $\mathcal{R}(A)$.

(f) *Mostre que a matriz* $P = A^\dagger A$ *é a matriz de projeção ortogonal em* $\mathcal{R}(A^t)$.

(g) *Determine a Decomposição em Valores Singulares e a pseudo–inversa da matriz*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

(h) *Mostre que os autovalores da matriz*

$$B = \begin{bmatrix} 0_n & A^t \\ A & 0_m \end{bmatrix},$$

de ordem $m + n$, *são*

$$-\sigma_1 , \cdots , -\sigma_r , \sigma_1 , \cdots , \sigma_r , 0 , \cdots , 0 .$$

Exercício 8.173 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz semipositiva–definida. Considerando a matriz* X *do Exemplo 8.17.4 e a matriz* R *do Exemplo 8.17.5, mostre que* $X = R$. *No caso em que* A *é uma matriz positiva–definida, mostre que as matrizes* X *e* R *são diferentes da matriz* $\widehat{R}$ *do Exemplo 8.17.6.*

Exercício 8.174 *Seja* $A \in \mathbb{M}_n(\mathbb{R})$ *uma matriz semipositiva–definida. Então, existe uma única matriz semipositiva–definida* $R \in \mathbb{M}_n(\mathbb{R})$ *tal que* $A = R^2$.

Bibliografia

[1] Tom M. Apostol, *Análisis Matemático*, Segunda Edición, Editorial Reverté, 1977.

[2] Tom M. Apostol, *Calculus*, Volume I, Second Edition, John Wiley & Sons, 1976.

[3] Tom M. Apostol, *Calculus*, Volume II, Second Edition, John Wiley & Sons, 1976.

[4] Tom M. Apostol, *Linear Algebra–A First Course with Applications to Differential Equations*, John Wiley & Sons, 1997.

[5] Alexander Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, Dover, 1983.

[6] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler, *Álgebra Linear*, Terceira Edição, Editora Harbra Ltda, 1986.

[7] C. A. Callioli, H. H. Domingues e R. C. F. Costa, *Álgebra Linear e Aplicações*, Sexta Edição, Atual Editora, 2003.

[8] R. Charnet, C. A. L. Freire, E. M. R. Charnet e H. Bonvino, *Análise de Modelos de Regressão Linear com Aplicações*, Editora da Unicamp, Segunda Edição, 2008.

[9] F. U. Coelho e M. L. Lourenço, *Um Curso de Álgebra Linear*, edusp, 2001.

[10] S. H. Friedberg, A. J. Insel and L. E. Spence, *Linear Algebra*, Prentice– Hall, Third Edition, 1997.

[11] Gene H. Golub & Charles F. Van Loan, *Matrix Computations*, Third Edition, John Hopkins, 1996.

[12] K. Hoffman e R. Kunze, *Álgebra Linear*, Editora da USP, 1971.

[13] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[14] Bernard Kolman e David R. Hill, *Introdução à Álgebra Lienar com Aplicações*, LTC, Oitava Edição, 2006.

[15] Serge Lang, *Introduction to Linear Algebra*, Second Edition, Springer, 1986.

[16] Elon L. Lima, *Álgebra Linear*, Coleção Matemática Universitária, IMPA, 1996.

[17] Elon L. Lima, *Curso de Análise*, Projeto Euclides, IMPA, 1996.

[18] Seymour Lipschutz, *Álgebra Linear*, Terceira Edição, Makron Books, 1994.

[19] LUENBERGER, D. D. (1973), *Introduction to Linear and Nonlinear Programming*, Addison–Wesley.

[20] Patricia R. de Peláez, Rosa F. Arbeláez y Luz E. M. Sierra, *Algebra Lineal con Aplicaciones*, Universidad Nacional de Colombia, 1997.

[21] Gilbert Strang, *Linear Algebra and its Applications*, Third Edition, Harcourt Brace Jovanovich Publishers, 1988.

[22] David S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, 1991.

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

$$\begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} = Q \begin{bmatrix} -4 & & \\ & 1 & \\ & & 6 \end{bmatrix} Q^t$$

$$Q^t Q = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

PULINUS

Álgebra Linear e suas Aplicações

Notas de Aula

Petronio Pulino

Departamento de Matemática Aplicada

Instituto de Matemática, Estatística e Computação Científica

Universidade Estadual de Campinas

E–mail: pulino@ime.unicamp.br

www.ime.unicamp.br/∼pulino/ALESA/

Janeiro de 2012

Conteúdo

A

Provas e Avaliações

Conteúdo

A.1 Segundo Semestre de 2008

A.1.1 Primeiro Teste

No Teste temos quatro questões enumeradas, 0, 1, 2 e 3. O aluno deve fazer uma questão cujo número é o resto da divisão por 4 da soma dos algarismos de seu RA. Exemplo: RA 0314468, 0+3+1+4+4+6+8 = 26. Como o resto da divisão de 26 por 4 é 2, o aluno com esse RA deve fazer a Questão 2.

Problema:

(1) Mostre que o subconjunto $W \subset \mathbb{R}^4$ é um subespaço vetorial de $\mathbb{R}^4$.

(2) Verifique se os elementos $v_1,\ v_2,\ v_3,\ v_4$ são linearmente dependentes em $\mathbb{R}^4$.

(3) Encontre um sistema de geradores para cada um dos subespaços $U \cap W$ e $U + W$.

Questão 0. $W = \{\ (x,y,z,t) \in \mathbb{R}^4\ /\ 2x - y + 3z + 4t = 0\ \}$

$$v_1 = (1,2,0,0)\ ,\ v_2 = (-5,-3,1,1)\ ,\ v_3 = (0,7,1,1) \quad \text{e} \quad v_4 = (1,0,1,0)$$

$$U = [(0,3,1,0),(1,0,1,0),(-4,6,2,2)]$$

Questão 1. $W = \{\ (x,y,z,t) \in \mathbb{R}^4\ /\ 3x + y - z + 6t = 0\ \}$

$$v_1 = (1,1,10,1)\ ,\ v_2 = (0,2,-4,-1)\ ,\ v_3 = (2,12,0,-3) \quad \text{e} \quad v_4 = (1,1,0,1)$$

$$U = [(1,1,4,0),(1,1,0,1),(2,3,3,-1)]$$

Questão 2. $W = \{\ (x,y,z,t) \in \mathbb{R}^4\ /\ 2x + 2y + z - t = 0\ \}$

$$v_1 = (1,1,0,4)\ ,\ v_2 = (0,1,1,3)\ ,\ v_3 = (3,-1,-4,0) \quad \text{e} \quad v_4 = (1,0,1,1)$$

$$U = [(2,-2,3,3),(1,0,1,1),(1,2,-2,4)]$$

Questão 3. $W = \{\ (x,y,z,t) \in \mathbb{R}^4\ /\ x - 3y - 2z - t = 0\ \}$

$$v_1 = (1,1,-1,0)\ ,\ v_2 = (6,1,1,1)\ ,\ v_3 = (5,0,2,1) \quad \text{e} \quad v_4 = (0,1,0,1)$$

$$U = [(-4,0,-3,2),(0,1,0,1),(12,3,1,1)]$$

A.1.2 Primeira Prova

Questão 1. **(2.5 Pontos)**

Considere o subconjunto U do espaço vetorial real $\mathcal{P}_3(\mathbb{R})$ definido da forma:

$$U = \{ \, p(x) \in \mathcal{P}_3(\mathbb{R}) \ / \ p(-1) + p'(-1) = 0 \quad \text{e} \quad p(1) = 0 \, \} \, .$$

onde p' indica a derivada de p. Verifique se o subconjunto U é um subespaço vetorial de $\mathcal{P}_3(\mathbb{R})$. Em caso afirmativo, determine uma base para o subespaço U.

Questão 2. **(2.0 Pontos)**

Considere o subespaço W do espaço vetorial real $\mathbb{R}^4$ gerado pelos elementos do conjunto S definido por:

$$S = \{ \, (1,0,1,2), \, (2,1,1,2), \, (1,-1,2,4) \, \} \, .$$

Determine um subespaço U de $\mathbb{R}^4$ de modo que $\mathbb{R}^4 = W \oplus U$.

Questão 3. **(2.5 Pontos)**

Sejam V um espaço vetorial real e $\gamma = \{ \, v_1 \, , \, v_2 \, , \, v_3 \, \}$ uma base ordenada de V.

(a) Mostre que $\beta = \{ \, v_1 + v_3 \, , \, v_2 + v_3 \, , \, v_1 + v_2 + v_3 \, \}$ é uma base de V.

(b) Se o elemento $v \in V$ tem como matriz de coordenadas $[v]_\gamma$ dada por:

$$[v]_\gamma = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} ,$$

determine a matriz de coordenadas do elemento v em relação à base ordenada β.

Questão 4. **(3.0 Pontos)**

Considere a transformação linear $T : \mathbb{R}^3 \longrightarrow \mathcal{P}_3(\mathbb{R})$ definida por:

$$T(1,0,1) = 2 + x^2 + x^3 \quad , \quad T(0,1,0) = 1 + x^2 \qquad \text{e} \qquad T(0,0,1) = x^2 - x^3 \, .$$

(a) Calcule $T(a,b,c)$ para a transformação linear T.

(b) Determine uma base para o subespaço $Im(T)$.

(c) A transformação linear T é injetora ?

A.1.3 Segundo Teste

No Teste temos quatro questões enumeradas, 0, 1, 2 e 3. O aluno deve fazer uma questão cujo número é o resto da divisão por 4 do último algarismo de seu RA. Exemplo: RA 0314468, que tem 8 como último algarismo. Como o resto da divisão de 8 por 4 é 0, o aluno com esse RA deve fazer a Questão 0.

Questão 0. Considere a transformação linear $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ tal que

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

onde $\beta = \{ (0,1) , (1,1) \}$ é a base ordenada para $\mathbb{R}^2$ e $\gamma = \{ 1+t , t-1 \}$ é a base ordenada para $\mathcal{P}_1(\mathbb{R})$.

(a) Determine o elemento $(a,b) \in \mathbb{R}^2$ de modo que $T(a,b) = 1+t$.

(b) Determine explicitamente a expressão de $T(a,b)$.

(c) Verifique se T é um isomorfismo de $\mathbb{R}^2$ em $\mathcal{P}_1(\mathbb{R})$. Em caso afirmativo, determine o isomorfismo inverso $T^{-1} : \mathcal{P}_1(\mathbb{R}) \longrightarrow \mathbb{R}^2$.

Questão 1. Considere a transformação linear $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ tal que

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix},$$

onde $\beta = \{ (1,1) , (0,1) \}$ é a base ordenada para $\mathbb{R}^2$ e $\gamma = \{ 1 , t-1 \}$ é a base ordenada para $\mathcal{P}_1(\mathbb{R})$.

(a) Determine o elemento $(a,b) \in \mathbb{R}^2$ de modo que $T(a,b) = t-1$.

(b) Determine explicitamente a expressão de $T(a,b)$.

(c) Verifique se T é um isomorfismo de $\mathbb{R}^2$ em $\mathcal{P}_1(\mathbb{R})$. Em caso afirmativo, determine o isomorfismo inverso $T^{-1} : \mathcal{P}_1(\mathbb{R}) \longrightarrow \mathbb{R}^2$.

Questão 2. Considere a transformação linear $T : I\!R^2 \longrightarrow \mathcal{P}_1(I\!R)$ tal que

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix},$$

onde $\beta = \{ (1,-1), (1,1) \}$ é a base ordenada para $I\!R^2$ e $\gamma = \{ 2t - 1, 1 + t \}$ é a base ordenada para $\mathcal{P}_1(I\!R)$.

(a) Determine o elemento $(a, b) \in I\!R^2$ de modo que $T(a, b) = 1 + t$.

(b) Determine explicitamente a expressão de $T(a, b)$.

(c) Verifique se T é um isomorfismo de $I\!R^2$ em $\mathcal{P}_1(I\!R)$. Em caso afirmativo, determine o isomorfismo inverso $T^{-1} : \mathcal{P}_1(I\!R) \longrightarrow I\!R^2$.

Questão 3. Considere a transformação linear $T : I\!R^2 \longrightarrow \mathcal{P}_1(I\!R)$ tal que

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix},$$

onde $\beta = \{ (-1,-1), (0,1) \}$ é a base ordenada para $I\!R^2$ e $\gamma = \{ t, 2t - 1 \}$ é a base ordenada para $\mathcal{P}_1(I\!R)$.

(a) Determine o elemento $(a, b) \in I\!R^2$ de modo que $T(a, b) = t$.

(b) Determine explicitamente a expressão de $T(a, b)$.

(c) Verifique se T é um isomorfismo de $I\!R^2$ em $\mathcal{P}_1(I\!R)$. Em caso afirmativo, determine o isomorfismo inverso $T^{-1} : \mathcal{P}_1(I\!R) \longrightarrow I\!R^2$.

A.1.4 Segunda Prova

Questão 1. **(3.0 Pontos)**

Considere a transformação linear $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ definida por:

$$T(x, y, z, t) = (x - 2y + t \,,\, 2x + y - z \,,\, 5y - z - 2t) \,.$$

(a) Determine uma base para o subespaço $Ker(T)$.

(b) Determine uma base para o subespaço $Im(T)$.

(c) Determine uma base γ para o espaço vetorial $\mathbb{R}^4$ contendo uma base de $Ker(T)$.

(d) Determine a matriz $[T]^{\gamma}_{\beta}$, onde β é a base ordenada de $\mathbb{R}^3$ dada por:

$$\beta = \{ (1,0,0) \,,\, (1,1,0) \,,\, (0,1,1) \} \,.$$

Questão 2. **(2.0 Pontos)**

Considere $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ a transformação linear tal que

$$T(1,1) = 1 - x \qquad \text{e} \qquad T(1,-1) = 1 + 3x \,.$$

Mostre que T é um isomorfismo de $\mathbb{R}^2$ em $\mathcal{P}_1(\mathbb{R})$. Determine explicitamente a expressão do isomorfismo inverso $T^{-1}(a_0 + a_1 x)$.

Questão 3. **(2.5 Pontos)**

Considere o espaço vetorial real $\mathbb{R}^3$ munido com o seguinte produto interno

$$\langle\, u \,,\, v \,\rangle = 2x_1x_2 + y_1y_2 + 4z_1z_2 \,,$$

onde $u = (x_1 \,,\, y_1 \,,\, z_1)$ e $v = (x_2 \,,\, y_2 \,,\, z_2)$. Dados os elementos

$$w_1 = (1,0,1) \qquad \text{e} \qquad w_2 = (1,1,1) \,,$$

determine dois elementos $v_1, v_2 \in \mathbb{R}^3$ de modo que

$$w_2 = v_1 + v_2 \,,$$

com $\{ v_1 \,,\, w_1 \}$ um conjunto linearmente dependente e $\{ v_2 \,,\, w_1 \}$ um conjunto ortogonal com relação ao produto interno definido acima.

Questão 4. **(2.5 Pontos)**

Considere o seguinte subespaço vetorial W do espaço vetorial real $I\!R^4$ dado por:

$$W = \{ (x, y, z, t) \in I\!R^4 \ / \ z + t = 0 \quad \text{e} \quad x - y = 0 \} .$$

Utilizando o **Processo de Ortogonalização de Gram-Schmidt**, determine uma base ortogonal para o espaço vetorial real $I\!R^4$ contendo uma base ortogonal do subespaço W, com relação ao produto interno canônico de $I\!R^4$.

A.1.5 Segunda Chamada

Questão 1. **(4.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ e o subconjunto U definido por:

$$U = \left\{ p(x) \in \mathcal{P}_2(\mathbb{R}) \ / \ \int_{-1}^{1} p(x)dx \ + \ 2\,p'(0) = 0_{\mathbb{R}} \right\}.$$

(a) Mostre que o subconjunto U é um subespaço vetorial de $\mathcal{P}_2(\mathbb{R})$.

(b) Determine uma base para o subespaço U.

(c) Determine um subespaço W de $\mathcal{P}_2(\mathbb{R})$ de modo que $\mathcal{P}_2(\mathbb{R}) = U \oplus W$.

(d) Dado o polinômio $p(x) = 2 - x$, determine um polinômio $q(x) \in U$ e um polinômio $r(x) \in W$ de modo que $p(x) = q(x) + r(x)$.

Questão 2. **(3.0 Pontos)**

Seja $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ a transformação linear tal que

$$T(1,2) = (1,0,1) \qquad \text{e} \qquad T(2,1) = (1,1,0)\,.$$

(a) Mostre que T é uma transformação linear injetora.

(b) Determine a matriz $[T]_{\gamma}^{\beta}$, onde $\beta = \{\,(1,2)\,,\,(2,1)\,\}$ é a base ordenada de $\mathbb{R}^2$ e γ é a base canônica de $\mathbb{R}^3$.

(c) Exiba uma transformação linear $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ tal que $Ker(P) = Im(T)$.

Questão 3. **(3.0 Pontos)**

(a) Mostre que a aplicação $\langle\,\cdot\,,\,\cdot\,\rangle : \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ dada por:

$$\langle\, p\,,\, q\,\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

define um produto interno no espaço vetorial real $\mathcal{P}_2(\mathbb{R})$.

(b) Determine uma base para o complemento ortogonal do subespaço $U = [1 + x]$ em $\mathcal{P}_2(\mathbb{R})$ com relação ao produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$ definido acima.

(c) Utilizando o **Processo de Ortogonalização de Gram-Schmidt**, determine uma base ortogonal para o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ com relação ao produto interno $\langle\,\cdot\,,\,\cdot\,\rangle$ definido acima.

A.1.6 Exame

Questão 1. **(2.5 Pontos)**

Considere o subconjunto U do espaço vetorial real $\mathbb{M}_n(\mathbb{R})$ dado por:

$$U = \{ A \in \mathbb{M}_n(\mathbb{R}) \ / \ A^t = A \quad \text{e} \quad tr(A) = 0 \}.$$

(a) Mostre que U é um subespaço vetorial de $\mathbb{M}_n(\mathbb{R})$.

(b) Considerando o espaço vetorial $\mathbb{M}_3(\mathbb{R})$, exiba uma base para o subespaço U.

Questão 2. **(2.5 Pontos)**

Considere os subespaços W_1 e W_2 do espaço vetorial real $\mathbb{R}^3$ dados por:

$$W_1 = \{ (x,y,z) \ / \ 3x - 2y + z = 0 \} \quad \text{e} \quad W_2 = \{ (x,y,z) \ / \ 2x + y - 4z = 0 \}.$$

(a) Determine a dimensão dos subespaços $W_1 + W_2$ e $W_1 \cap W_2$.

(b) Encontre uma base de $\mathbb{R}^3$ que contenha uma base do subespaço W_1 e também uma base do subespaço W_2.

Questão 3. **(2.5 Pontos)**

Considere o operador linear $T : \mathcal{P}_1(\mathbb{R}) \longrightarrow \mathcal{P}_1(\mathbb{R})$ dado por:

$$T(p(x)) = p'(x) + (x+1)p(1).$$

Sejam $\beta = \{ 1, 7 - 4x \}$ e $\gamma = \{ q(x), 2x - 1 \}$ bases para $\mathcal{P}_1(\mathbb{R})$ tais que

$$[T]^\beta_\gamma = \begin{bmatrix} 3 & s \\ -1 & 1 \end{bmatrix}.$$

(a) Determine o polinômio $q(x)$ e o parâmetro $s \in \mathbb{R}$.

(b) T é um automorfismo? Em caso afirmativo, determine o automorfismo inverso.

Questão 4. **(2.5 Pontos)**

Seja W o subespaço de $\mathbb{R}^4$ dado por:

$$W = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x + y = 0 \quad \text{e} \quad 2x - y + z = 0 \}.$$

Determine uma base ortogonal para cada um dos subespaços W e $W^\perp$, com relação ao produto interno usual de $\mathbb{R}^4$.

A.2 Segundo Semestre de 2006

A.2.1 Primeira Prova

Questão 1. **(2.5 Pontos)**

Considere o espaço vetorial real $\mathbb{R}^2$ e os seguinte subespaços

$$U = \{ (x,y) \in \mathbb{R}^2 \ / \ y = 3x \} \quad \text{e} \quad W = \{ (x,y) \in \mathbb{R}^2 \ / \ y = -2x \} .$$

Verifique se o seguinte subconjunto

$$U \cup W = \{ (x,y) \in \mathbb{R}^2 \ / \ (x,y) \in U \quad \text{ou} \quad (x,y) \in W \}$$

é um subespaço vetorial de $\mathbb{R}^2$.

Questão 2. **(2.5 Pontos)**

Sejam V um espaço vetorial sobre o corpo $\mathbb{F}$ e $u,\, v,\, w$ elementos distintos de V. Prove que o conjunto $\{ u,\, v,\, w \}$ é linearmente independente em V se, e somente se, o conjunto $\{ u + v,\, u + w,\, v + w \}$ é linearmente independente em V.

Questão 3. **(2.5 Pontos)**

Considere o espaço vetorial real $\mathbb{M}_2(\mathbb{R})$ e os seguintes subespaços

$$U = \left\{ \begin{bmatrix} a & b \\ c & a \end{bmatrix} \quad ; \quad a\,,\, b\,,\, c\, \in \mathbb{R} \right\} \quad \text{e} \quad W = \left\{ \begin{bmatrix} 0 & a \\ -a & b \end{bmatrix} \quad ; \quad a\,,\, b\, \in \mathbb{R} \right\} .$$

(a) Determine uma base para cada um dos seguintes subespaços:

$$U\,, \quad W\,, \quad U \cap W \quad \text{e} \quad U + W\,.$$

(b) $\mathbb{M}_2(\mathbb{R}) = U \oplus W$? Justifique sua resposta.

Questão 4. **(2.5 Pontos)**

Considere o espaço vetorial $\mathbb{R}^2$. A matriz de mudança da base ordenada $\gamma = \{ u_1,\, u_2 \}$, onde $u_1 = (1,1)$ e $u_2 = (-2,2)$, para a base ordenada $\alpha = \{ v_1,\, v_2 \}$ é dada por:

$$[I]^{\gamma}_{\alpha} = \begin{bmatrix} 1 & 0 \\ 4 & -2 \end{bmatrix}.$$

(a) Determine a base ordenada α.

(b) Determine o elemento $u \in \mathbb{R}^2$ tal que $[u]_\alpha = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

A.2.2 Segunda Prova

Questão 1. **(3.0 Pontos)**

Determine explicitamente a expressão de uma transformação linear $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$ satisfazendo simultaneamente as seguintes condições:

(a) O elemento $p(x) = (1 + x) \in Ker(T)$.

(b) O elemento $q(x) = x \notin Ker(T)$.

(c) $Im(T) = [(1, 1, 1)]$.

Questão 2. **(3.0 Pontos)**

Sejam T um operador linear sobre $\mathbb{R}^4$, $\gamma = \{ v_1, v_2, v_3, v_4 \}$ uma base ordenada para o espaço vetorial real $\mathbb{R}^4$ e o subespaço $S = [v_1, v_2, v_3]$. Pede–se:

(a) Sabendo que $T(v) = v$ para todo $v \in S$ e $T(v_4) = v_1 + v_3$, determine $[T]_\gamma^\gamma$.

(b) Sabendo que

$$[I]_\gamma^\beta = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

onde $\beta = \{ e_1, e_2, e_3, e_4 \}$ é a base canônica de $\mathbb{R}^4$, determine $[T(e_1)]_\gamma$.

Questão 3. **(3.0 Pontos)**

Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$, definido por: $T(p(x)) = p'(x) + p(x)$, e a transformação linear $P : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$ definida por: $P(a + bx + cx^2) = (a + b, c, a - b)$.

(a) Determine a transformação linear $P \circ T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$.

(b) Determine a matriz $[P \circ T]_\gamma^\beta$, onde β é a base canônica de $\mathcal{P}_2(\mathbb{R})$ e γ é a base canônica de $\mathbb{R}^3$.

(c) Verifique se P é um isomorfismo de $\mathcal{P}_2(\mathbb{R})$ em $\mathbb{R}^3$. Em caso afirmativo, determine o isomorfismo inverso $P^{-1} : \mathbb{R}^3 \longrightarrow \mathcal{P}_2(\mathbb{R})$.

Questão 4. **(2.0 Pontos)**

Sejam V um espaço vetorial de dimensão finita, com $dim(V) = n$, e T um operador linear sobre V tal que $Im(T) = Ker(T)$. Pede–se:

(a) Mostre que n é par.

(b) Considerando $V = I\!R^4$, determine um operador linear T sobre V com essas propriedades.

A.2.3 Terceira Prova

Questão 1. **(2.0 Pontos)**

Sejam V um espaço vetorial sobre o corpo $I\!\!F$, T um operador linear sobre V, $\lambda \in I\!\!F$ e E_λ o subconjunto de V definido por:

$$E_\lambda = \{ v \in V \ / \ T(v) = \lambda v \} .$$

Prove que $T(E_\lambda) \subset E_\lambda$.

Questão 2. **(3.0 Pontos)**

Sejam V um espaço vetorial de dimensão n sobre o corpo $I\!\!F$ e T um operador linear sobre V. Pede–se:

(a) Se $v \in V$ é um autovetor de T, quantos autovalores associados a v podem existir, no máximo? Justifique sua resposta.

(b) Se $\lambda = 0$ é um autovalor de T, podemos afirmar que T não é um operador injetor? A recíproca é verdadeira? Justifique suas respostas.

(c) Se o operador linear T possui somente dois autovalores distintos λ_1 e λ_2 com $dim(V_{\lambda_1}) = n - 1$, prove que T é um operador diagonalizável.

Questão 3. **(3.0 Pontos)**

Seja $T : I\!\!R^2 \longrightarrow I\!\!R^2$ um operador linear definido por $T(x,y) = (5x - 6y\,,\, x)$. Pede–se:

(a) Calcule os autovalores e os autovetores do operador T.

(b) Exiba uma base para cada um dos autoespaços do operador T.

(c) Utilizando o resultado do item (a), calcule os valores de a, b, c, $d \in I\!\!R$, tais que

$$T^8(x,y) = (ax + by\,,\, cx + dy) ,$$

onde $T^n : I\!\!R^2 \longrightarrow I\!\!R^2$ é o operador linear definido por:

$$T^0 = I \qquad \text{e} \qquad T^n = T^{n-1} \circ T \qquad \text{para todo natural} \qquad n \geq 1 .$$

Questão 4. **(3.0 Pontos)**

Determine explicitamente a expressão do operador linear T sobre $\mathbb{R}^4$, diagonalizável, satisfazendo simultaneamente as seguintes condições:

(a) $Ker(T) = \{ (x, y, z, t) \in \mathbb{R}^4 \ / \ x + y - z + t = 0 \quad \text{e} \quad z - t = 0 \}$.

(b) $T(0, 0, 1, 0) = (0, 0, 2, 0)$.

(c) $(0, 1, 0, 0) \in Im(T)$.

(d) $\lambda = -3$ é um autovalor do operador T.

A.2.4 Segunda Chamada

Questão 1. **(2.5 Pontos)**

Diga se é Falsa ou Verdadeira cada uma das afirmações abaixo, justificando sua resposta.

(a) Existe uma transformação linear $T : I\!R^4 \longrightarrow I\!R^3$ que é injetora.

(b) Existe uma transformação linear $T : I\!R^4 \longrightarrow \mathcal{P}_2(I\!R)$ que é sobrejetora.

(c) Subconjuntos de um conjunto linearmente dependente são linearmente dependentes.

(d) Os espaços vetoriais $\mathcal{P}_4(I\!R)$ e $I\!M_2(I\!R)$ são isomorfos.

(e) Sejam V um espaço vetorial de dimensão finita com $dim(V) = n$, U e W subespaços de V com $dim(U) > \dfrac{n}{2}$ e $dim(W) > \dfrac{n}{2}$. Então, $U \cap W = \{ 0_V \}$.

Questão 2. **(2.5 Pontos)**

Considere o subconjunto U do espaço vetorial real $I\!M_n(I\!R)$ definido por:

$$U = \{ A \in I\!M_n(I\!R) \ / \ A^t = A \quad \text{e} \quad tr(A) = 0 \}$$

(a) Mostre que U é um subespaço vetorial de $I\!M_n(I\!R)$.

(b) Considerando o espaço vetorial $I\!M_3(I\!R)$, exiba uma base para o subespaço U.

Questão 3. **(2.5 Pontos)**

Considere a transformação linear $T : I\!R^2 \longrightarrow I\!R^3$ tal que

$$[T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ -2 & 3 \end{bmatrix},$$

onde β é a base canônica de $I\!R^2$ e $\gamma = \{ (1,0,1), (-1,0,1), (0,1,0) \}$ é uma base ordenada de $I\!R^3$. Pede–se

(a) Determine $T(1,0)$ e $T(0,1)$.

(b) Determine uma base para $Im(T)$.

(c) A transformação T é injetora? Justifique sua resposta.

Questão 4. **(2.5 Pontos)**

Considere o operador linear T sobre $\mathcal{P}_2(\mathbb{R})$ definido por:

$$T(a + bx + cx^2) = (3a + 2b + c) + (b - c)x + 2cx^2 .$$

Determine os autovalores e os autovetores do operador linear T, exibindo uma base para cada um dos autoespaços de T. O operador T é diagonalizável? Justifique sua resposta.

A.2.5 Exame

Questão 1. **(3.0 Pontos)**

Considere o espaço vetorial real $I\!R^4$ e os seguintes subespaços

$$\begin{aligned} U &= \{ (x,y,z,t) \in I\!R^4 \ / \ x+y-z+t = 0 \quad \text{e} \quad z-t = 0 \} \\ W &= \{ (x,y,z,t) \in I\!R^4 \ / \ x+y+z = 0 \} \end{aligned}$$

(a) Determine uma base para o subespaço $U + W$.

(b) O subespaço $U + W$ é uma soma direta dos subespaços U e W? Justifique.

(c) Determine uma base para o subespaço $U \cap W$.

(d) Determine o operador linear T sobre $I\!R^4$ tal que $Im(T) = U \cap W$ e $Ker(T) = W$.

Questão 2. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_3(I\!R)$ e o subconjunto U definido por:

$$U = \{ p(x) \in \mathcal{P}_3(I\!R) \ / \ p(1) + p(-1) = 0 \} .$$

O subconjunto U é um subespaço vetorial de $\mathcal{P}_3(I\!R)$? Justifique sua resposta. Em caso afirmativo, determine uma base para U.

Questão 3. **(3.0 Pontos)**

Considere o operador linear T sobre $I\!R^2$ tal que

$$[T]^{\alpha}_{\gamma} = \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix},$$

onde $\alpha = \{ (0,1) , (1,0) \}$ e $\gamma = \{ (-1,0) , (0,-1) \}$ são bases ordenadas de $I\!R^2$.

(a) Determine $T(1,0)$ e $T(0,1)$.

(b) Determine a matriz $[I]^{\alpha}_{\gamma}$.

(c) Determine explicitamente a expressão do operador linear T.

(d) O operador linear T^2 é diagonalizável? Justifique.

Questão 4. **(3.0 Pontos)**

Sejam A e B matrizes similares de ordem n. Pede–se:

(a) Mostre que A e B possuem os mesmos autovalores.

(b) Determine a relação entre os autovetores das matrizes A e B.

(c) Mostre que se A é diagonalizável, então B é diagonalizável.

A.3 Primeiro Semestre de 2006

A.3.1 Primeira Prova

Questão 1. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{C}([-1,1])$. Dê exemplo de um subconjunto S de $\mathcal{C}([-1,1])$ que é fechado com relação à operação de adição de elementos, mas que não é fechado com relação à operação de multiplicação por escalar. Justifique sua resposta.

Questão 2. **(2.0 Pontos)**

Considere V um espaço vetorial sobre o corpo $\mathbb{F}$. Sejam $S = \{ v_1, v_2, v_3, v_4 \}$ um conjunto linearmente independente em V e um elemento $u \in V$, não nulo. Mostre que o conjunto $\{ v_1, v_2, v_3, v_4, u \}$ é linearmente dependente se, e somente se, o elemento u pertence ao subespaço gerado pelos elementos do conjunto S, isto é, $u \in [v_1, v_2, v_3, v_4]$.

Questão 3. **(2.0 Pontos)**

Considere os seguintes subespaços vetoriais de $\mathbb{R}^3$

$$\begin{aligned} U &= \{ (x,y,z) \in \mathbb{R}^3 \ / \ 2x - 4y + 6z = 0 \} \\ W &= [(1,0,1), (1,1,3)] \end{aligned}$$

Determine um sistema de geradores para cada um dos subespaços $U + W$ e $U \cap W$. O subespaço $U + W$ é uma soma direta dos subespaços U e W ? Justifique sua resposta.

Questão 4. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_3(\mathbb{R})$. Determine uma base para o subespaço vetorial de $\mathcal{P}_3(\mathbb{R})$ definido por:

$$S = \{ p(x) \in \mathcal{P}_3(\mathbb{R}) \ / \ p(-1) + p'(-1) = 0 \quad \text{e} \quad p(1) = 0 \} .$$

Questão 5. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathbb{R}^4$. Seja W o subespaço de $\mathbb{R}^4$ gerado pelos elementos do conjunto $S = \{ (1,0,1,2), (2,-1,1,3), (-1,1,0,-1) \}$. Determine uma base de $\mathbb{R}^4$ contendo uma base do subespaço W.

A.3.2 Segunda Prova

Questão 1. **(2.0 Pontos)**

A matriz de mudança da base ordenada $\alpha = \{ p_1(x), p_2(x), p_3(x) \}$ de $\mathcal{P}_2(\mathbb{R})$, onde

$$p_1(x) = 1 - x , \quad p_2(x) = 1 + x \quad \text{e} \quad p_3(x) = 1 - x^2 ,$$

para uma base ordenada $\gamma = \{ q_1(x), q_2(x), q_3(x) \}$ de $\mathcal{P}_2(\mathbb{R})$ é dada por:

$$[I]^{\alpha}_{\gamma} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Determine a base ordenada γ. Dado o polinômio $p(x) = 3 - x + 2x^2$ determine seu vetor de coordenadas $[p(x)]_{\alpha}$, com relação à base ordenada α.

Questão 2. **(2.0 Pontos)**

Determine explicitamente a expressão de uma transformação linear T de $\mathcal{P}_2(\mathbb{R})$ em $\mathbb{M}_2(\mathbb{R})$ satisfazendo simultaneamente as seguintes condições:

(a) O elemento $p(x) = (1 + x^2) \in Ker(T)$.

(b) O elemento $q(x) = 1 \notin Ker(T)$.

(c) O elemento $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \in Im(T)$.

Questão 3. **(2.0 Pontos)**

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{F}$ e $T : V \longrightarrow W$ uma transformação linear **injetora**. Mostre que se $\{ v_1, \cdots , v_m \}$ é linearmente independente em V, então $\{ T(v_1), \cdots , T(v_m) \}$ é linearmente independente em W.

Questão 4. **(2.0 Pontos)**

Diga se é Falsa ou Verdadeira cada uma das afirmações abaixo, justificando sua resposta.

(a) Existe uma transformação linear $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ que é injetora.

(b) Existe uma transformação linear $T : \mathbb{R}^4 \longrightarrow \mathcal{P}_2(\mathbb{R})$ que é sobrejetora.

(c) Existe uma transformação linear $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_2(\mathbb{R})$ que é bijetora.

Questão 5. **(2.0 Pontos)**

Considere $T : \mathbb{R}^2 \longrightarrow \mathcal{P}_1(\mathbb{R})$ a transformação linear tal que

$$T(1,-1) \,=\, 2 + x \qquad \text{e} \qquad T(0,1) \,=\, x - 1\,.$$

Mostre que T é um isomorfismo de $\mathbb{R}^2$ em $\mathcal{P}_1(\mathbb{R})$. Determine o isomorfismo inverso T^{-1} de $\mathcal{P}_1(\mathbb{R})$ em $\mathbb{R}^2$.

A.3.3 Terceira Prova

Questão 1. **(2.0 Pontos)**

Seja U um subespaço de $\mathcal{P}_3(\mathbb{R})$ tendo como base $\beta = \{\, x - x^2 + x^3,\ 1 + x + x^2 \,\}$. Considere a transformação linear $T : U \longrightarrow \mathcal{P}_2(\mathbb{R})$ dada por: $T(p(x)) = p'(x) + (x+1)p(0)$.

Considere que $[T]_\gamma^\beta = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$, onde γ é uma base para $\mathcal{P}_2(\mathbb{R})$. Pede–se:

(a) Determine $[p(x)]_\beta$ sabendo que $[T(p(x))]_\gamma = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$.

(b) Se $\gamma = \{\, x - 1,\ p_1(x),\ p_2(x) \,\}$, determine o polinômio $q(x) = p_1(x) - p_2(x)$.

Questão 2. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{C}_0^1([0,1])$, isto é,

$$\mathcal{C}_0^1([0,1]) = \{\, f \in \mathcal{C}^1([0,1]) \ / \ f(1) = 0 \,\}.$$

Verifique se cada uma das aplicações

(a) $\langle f , g \rangle = \displaystyle\int_0^1 f'(x)g(x)dx$ **(b)** $\langle f , g \rangle = \displaystyle\int_0^1 f'(x)g'(x)dx$

define um produto interno no espaço vetorial $\mathcal{C}_0^1([0,1])$. Justifique sua resposta.

Questão 3. **(2.0 Pontos)**

Sejam V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$ e $\| \cdot \|_2$ a norma Euclidiana. Pede–se:

(a) Mostre que se θ é o ângulo entre os elementos $u, v \in V$, não nulos, então

$$\| u + v \|_2^2 = \| u \|_2^2 + \| v \|_2^2 + 2 \| u \|_2 \| v \|_2 \cos(\theta).$$

(b) Mostre que se $\beta = \{ q_1, \cdots , q_n \}$ é um conjunto ortonormal em V, então β é um conjunto linearmente independente em V.

Questão 4. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(I\!R)$ munido do produto interno

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} x^2 p(x)q(x)dx \qquad ; \qquad \forall\;\; p,\, q \;\in\; \mathcal{P}_2(I\!R)\,.$$

Determine uma base para o complemento ortogonal do subespaço $S \,=\, [1 \,+\, x]$ em $\mathcal{P}_2(I\!R)$ com relação ao produto interno $\langle\, \cdot\,,\, \cdot\, \rangle$ definido acima.

Questão 5. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(I\!R)$ com o produto interno

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{0}^{1} p(x)q(x)dx \qquad ; \qquad \forall\;\; p,\, q \;\in\; \mathcal{P}_2(I\!R)\,.$$

Determine a melhor aproximação do polinômio $q(x) \,=\, 1 \,-\, x^2$ no subespaço $\mathcal{P}_1(I\!R)$.

A.3.4 Segunda Chamada

Questão 1. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ e a base $\beta = \{1, x, x^2\}$. Dada a matriz

$$P = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Pede–se:

(a) Determine uma base $\gamma = \{p_1(x), p_2(x), p_3(x)\}$ de modo que $P = [I]^{\gamma}_{\beta}$.

(b) Dado o polinômio $q(x) = -3 - 2x + 2x^2$, determine $[q(x)]_{\gamma}$.

Questão 2. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathbb{R}^3$ com o produto interno usual $\langle \cdot , \cdot \rangle$ e S o subespaço definido por:

$$S = \{ (x, y, z) \in \mathbb{R}^3 \ / \ x + y + z = 0 \}.$$

Determine um operador linear $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ tal que $Im(T) = S$ e $Ker(T) = S^{\perp}$.

Questão 3. **(2.0 Pontos)**

Seja $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ o operador linear definido por:

$$T(x, y) = (3x - 2y, -2x + 3y).$$

Pede–se:

(a) Determine uma base para cada um dos seguintes subespaços:

$$U_1 = \{ (x, y) \in \mathbb{R}^2 \ / \ T(x, y) = 5(x, y) \}$$

$$U_2 = \{ (x, y) \in \mathbb{R}^2 \ / \ T(x, y) = (x, y) \}$$

(b) Mostre que o conjunto $\beta = \beta_1 \cup \beta_2$, onde β_1 é uma base para U_1 e β_2 é uma base para U_2, é uma base para $\mathbb{R}^2$ e determine $[T]^{\beta}_{\beta}$.

Questão 4. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathbb{R}^3$ munido do produto interno usual $\langle \cdot , \cdot \rangle$. Dados os elementos $u = (1,1,1)$ e $v = (3,2,1)$. Pede–se:

(a) Determine os elementos w_1 e w_2 tais que $v = w_1 + w_2$, de modo que w_1 seja ortogonal ao elemento u e que o conjunto $\{ w_2, u \}$ seja linearmente dependente.

(b) Decompor o elemento $w = (1,-1,2)$ como a soma de um elemento no subespaço $S = [u, w_1]$ e outro no subespaço $S^{\perp}$.

Questão 5. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathbb{R}^4$ munido do produto interno usual. Seja U o subespaço gerado pelos elementos $u_1 = (1,1,1,1)$ e $u_2 = (-1,1,-1,1)$. Pede–se:

(a) Determine a melhor aproximação do elemento $v = (2,1,3,1)$ no subespaço U.

(b) Determine um subespaço W de modo que $\mathbb{R}^4 = U \oplus W$. Justifique sua resposta.

A.3.5 Exame

Questão 1. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(I\!R)$ e o subconjunto U definido por:

$$U = \left\{ p \in \mathcal{P}_2(I\!R) \ / \ \int_{-1}^{1} p(x)dx \ + \ p'(0) = 0 \right\}.$$

O subconjunto U é um subespaço vetorial de $\mathcal{P}_2(I\!R)$? Justifique sua resposta. Em caso afirmativo, determine uma base para U.

Questão 2. **(2.0 Pontos)**

Considere V um espaço vetorial real e $\beta = \{ u_1, u_2, u_3 \}$ uma base ordenada de V. Seja $\gamma = \{ w_1, w_2, w_3 \}$ cujos elementos estão relacionados com os elementos da base β da seguinte forma:

$$\begin{cases} w_1 = u_1 - u_2 - u_3 \\ w_2 = 2u_2 + 3u_3 \\ w_3 = 3u_1 + u_3 \end{cases}$$

(a) Mostre que γ é uma base para V.

(b) Determine a matriz de mudança de base $[I]_\beta^\gamma$.

(c) Se um elemento $v \in V$ tem por vetor de coordenadas, em relação à base γ,

$$[v]_\gamma = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix},$$

qual é o seu vetor de coordenadas com relação à base ordenada β ?

Questão 3. **(2.0 Pontos)**

Sejam U e W subespaços vetoriais de $I\!R^3$ definidos por:

$$U = \{ (x, y, z) \in I\!R^3 \ / \ x + y + z = 0 \}$$

$$W = [(1, 0, 1), (0, -1, 1)]$$

Determine um operador linear T sobre $I\!R^3$ tal que $Im(T) = U$ e $Ker(T) = U \cap W$.

Questão 4. **(2.0 Pontos)**

Considere o operador linear $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ definido por:

$$T(p(x)) \;=\; p(x) \;+\; (1 \,+\, x)p'(x)\,.$$

Verifique se T é um automorfismo de $\mathcal{P}_3(\mathbb{R})$ e determine a matriz $[T]^{\beta}_{\beta}$, onde β é a base canônica de $\mathcal{P}_3(\mathbb{R})$.

Questão 5. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$ com o produto interno

$$\langle\, p \,,\, q \,\rangle \;=\; \int_{-1}^{1} p(x)q(x)dx \qquad ; \qquad \forall \;\; p,\, q \;\in\; \mathcal{P}_2(\mathbb{R})\,.$$

Determine uma base para o complemento ortogonal do subespaço $S = [1 \,+\, x,\, 1 \,-\, x^2]$ em $\mathcal{P}_2(\mathbb{R})$ com relação ao produto interno $\langle\, \cdot \,,\, \cdot \,\rangle$ definido acima.

A.4 Segundo Semestre de 2005

A.4.1 Primeira Prova

Questão 1. **(2.0 Pontos)**

Considere o espaço vetorial real $V = \{ (x, y) \ / \ x, y \in I\!R \}$, com as operações:

- **adição de elementos:** $(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2 + 5, \, y_1 + y_2)$
- **multiplicação por escalar:** $\alpha \odot (x, y) = (\alpha\, x + 5(\alpha - 1), \, \alpha\, y)$ para $\alpha \in I\!R$.

(a) Exiba o elemento neutro da adição desse espaço.

(b) Exiba o elemento simétrico aditivo do elemento $(x, y) \in V$.

(c) Verifique se $W = \{ (x, y) \in V \ / \ x = -5 \}$ é um subespaço vetorial de V.

Questão 2. **(2.0 Pontos)**

Diga se é Falsa ou Verdadeira cada uma das afirmações, justificando sua resposta.

(a) Seja V um espaço vetorial real. Se $\{ v_1, v_2, v_3 \}$ é LI em V, então o conjunto $\{ v_1 - v_2 \, , \, v_2 + v_3 \, , \, v_1 + v_3 \}$ é LI em V.

(b) O subconjunto $W = \{ A \in I\!M_2(I\!R) \ / \ A^2 = A \}$ é um subespaço de $I\!M_2(I\!R)$.

(c) O subconjunto $S = \{ f \in \mathcal{C}([-a, a]) \ / \ f(-x) = f(x) \ ; \ x \in [-a, a] \}$ é um subespaço de $\mathcal{C}([-a, a])$.

Questão 3. **(2.0 Pontos)**

Considere o sistema linear homogêneo

$$\begin{cases} 2x + 4y + z = 0 \\ x + y + 2z = 0 \\ x + 3y - z = 0 \end{cases} .$$

Pede–se:

(a) Mostre que o conjunto solução S é um subespaço vetorial de $I\!R^3$ e determine uma base para esse subespaço.

(b) Dado o subespaço vetorial $U = \{ (x, \, y, \, z) \in I\!R^3 \ / \ x - y + z = 0 \}$, determine o subespaço $U \cap S$ e uma base para esse subespaço.

(c) Determine o subespaço vetorial $U + S$ e uma base para esse subespaço.

Questão 4. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_2(\mathbb{R})$. Pede–se:

(a) Mostre que o subconjunto $W = \{ p \in \mathcal{P}_2(\mathbb{R}) \ / \ p(2) = 0 \}$ é um subespaço vetorial de $\mathcal{P}_2(\mathbb{R})$.

(b) Exiba uma base β para o subespaço W.

(c) Encontre as coordenadas do polinômio $p(x) = 6 - 5x + x^2$ com relação à base β.

Questão 5. **(2.0 Pontos)**

Considere o subespaço

$$V = \{ (x, y, z) \in \mathbb{R}^3 \ / \ x + 2y + z = 0 \quad \text{e} \quad -x + 3y + 2z = 0 \}$$

do espaço vetorial real $\mathbb{R}^3$. Determine um subespaço W do $\mathbb{R}^3$ tal que $\mathbb{R}^3 = V \oplus W$.

A.4.2 Primeira Prova (*Substitutiva*)

Questão 1. **(2.0 Pontos)**

Considere o espaço vetorial real $(V, \oplus, \odot)$, onde $V = \{ (x,y) \in I\!R^2 \ / \ x > 0 \quad \text{e} \quad y > 0 \}$ munido com as operações

- **adição de elementos:** $(x_1, y_1) \oplus (x_2, y_2) = (x_1 x_2, \, y_1 y_2), \ \forall \, (x_1, y_1), \, (x_2, y_2) \in V.$
- **multiplicação por escalar:** $\alpha \odot (x,y) = (x^{\alpha}, \, y^{\alpha}), \ \forall \, \alpha \in I\!R \quad \text{e} \quad \forall \, (x,y) \in V.$

Pede–se:

(a) Exiba o elemento neutro da adição desse espaço.

(b) Exiba o elemento simétrico aditivo do elemento $u = (x,y) \in V$.

(c) Mostre que $(\alpha + \beta) \odot u = \alpha \odot u \oplus \beta \odot u$, $u = (x,y) \in V$ e $\alpha, \beta \in I\!R$.

(d) $W = \{ (x,y) \in V \ / \ y = 2x \}$ é um subespaço vetorial de V ?

Questão 2. **(2.0 Pontos)**

Considere os seguintes subconjuntos do espaço vetorial real $I\!M_n(I\!R)$

$$U = \{ A \in I\!M_n(I\!R) \ / \ A^t = A \} \qquad \text{e} \qquad W = \{ A \in I\!M_n(I\!R) \ / \ A^t = -A \}$$

Mostre que U e W são subespaços vetoriais de $I\!M_n(I\!R)$ e que $I\!M_n(I\!R) = U \oplus W$. Considerando o espaço vetorial $I\!M_3(I\!R)$, exiba uma base para os subespaços U e W.

Questão 3. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{P}_3(I\!R)$ e os elementos $p(x) = -3 - 5x - 2x^2 + x^3$ e $q(x) = -9 - 4x - 5x^2 + 3x^3$. Mostre que o elemento $r(x) = -6 + 12x - 2x^2 + 2x^3$ pode ser escrito como uma combinação linear dos elementos p e q, mas que o elemento $s(x) = 8 + 7x - 2x^2 + 3x^3$ não pode.

Questão 4. **(2.0 Pontos)**

Seja $(V, +, \cdot)$ um espaço vetorial real. Se $\{ v_1, \, v_2, \, v_3 \}$ é linearmente independente em V, mostre que $\{ v_1 + v_2 + v_3 \, , \, v_2 + v_3 \, , \, v_1 + v_3 \}$ é linearmente independente em V.

Questão 5. **(2.0 Pontos)**

Considere o seguinte subconjunto do espaço vetorial real $I\!\!M_n(I\!\!R)$

$$U = \{ A \in I\!\!M_n(I\!\!R) \ / \ A^t = A \quad \text{e} \quad tr(A) = 0 \}$$

Mostre que U é um subespaço vetorial de $I\!\!M_n(I\!\!R)$. Considerando o espaço vetorial $I\!\!M_3(I\!\!R)$, exiba uma base para o subespaço U.

A.4.3 Segunda Prova

Questão 1. **(2.0 Pontos)**

Seja V um espaço vetorial real e $\gamma = \{ v_1, v_2, v_3 \}$ uma base para V. Pede–se:

(a) Mostre que $\beta = \{ v_1, v_1 + v_2, v_1 + v_2 + v_3 \}$ é uma base para V.

(b) Determine a matriz de mudança da base β para a base γ.

(c) Se o elemento $v \in V$ tem como vetor de coordenadas em relação à base γ

$$[v]_\gamma = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

determine seu vetor de coordenadas $[v]_\beta$.

Questão 2. **(2.0 Pontos)**

Considere os seguintes subespaços de $\mathbb{R}^4$

$$\begin{aligned} U &= [(1,0,1,1),(0,1,1,1)] \\ W &= \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x + y = 0 \quad \text{e} \quad z + t = 0 \} \end{aligned}$$

Determine um operador linear T sobre $\mathbb{R}^4$ tal que $Ker(T) = W$ e $Im(T) = U$.

Questão 3. **(3.0 Pontos)**

Diga se é Falsa ou Verdadeira cada uma das afirmações, justificando sua resposta.

(a) $D : \mathbb{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ com $D(A) = det(A)$ é uma transformação linear.

(b) Não existe transformação linear $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ que é injetora.

(c) Existe uma transformação linear $T : \mathbb{R}^4 \longrightarrow \mathcal{P}_2(\mathbb{R})$ que é sobrejetora.

(d) Existe transformação linear $T : \mathbb{R}^m \longrightarrow \mathbb{R}^n$, com $m < n$, que é bijetora.

Questão 4. **(2.0 Pontos)**

Considere as transformações lineares $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ e $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dadas por: $T(x,y) = (2x , x - y , y)$ e $P(x,y,z) = (y - z , z - x)$. Determine a matriz $[T \circ P]_\beta^\beta$, onde β é a base canônica do $\mathbb{R}^3$, e determine uma base para o subespaço $Im(T \circ P)$. O operador linear $T \circ P$ é um automorfismo de $\mathbb{R}^3$? Justifique sua resposta.

Questão 5. **(2.0 Pontos)**

Considere a transformação linear $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_1(\mathbb{R})$ dada por:

$$T(p(x)) = ap(0) - p'(x)$$

com

$$[T]^{\alpha}_{\beta} = \begin{bmatrix} 3 & -3 & b \\ 3 & -3 & -2 \end{bmatrix},$$

considerando $\alpha = \{ 1, cx + 1, x^2 \}$ a base para $\mathcal{P}_2(\mathbb{R})$ e $\beta = \{ 1 - x, x \}$ a base para $\mathcal{P}_1(\mathbb{R})$. Pede–se:

(a) Determine os parâmetros $a, b, c \in \mathbb{R}$.

(b) Determine $[T(q(x))]_{\beta}$ e $[T(q(x))]_{\gamma}$, com $\gamma = \{ 1, x \}$ a base canônica de $\mathcal{P}_1(\mathbb{R})$, sabendo que

$$[q(x)]_{\alpha} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}.$$

A.4.4 Terceira Prova

Questão 1. **(2.0 Pontos)**

Seja V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$ e T um automorfismo de V. Mostre que a aplicação

$$\begin{aligned} f(\cdot,\cdot): \ V \times V & \longrightarrow \ I\!R \\ (u,v) & \longrightarrow \ f(u,v) \ = \ \langle\, T(u) \,,\, T(v) \,\rangle \end{aligned}$$

define um produto interno em V. Dê um exemplo considerando $V = I\!R^3$ munido do produto interno usual.

Questão 2. **(2.0 Pontos)**

(a) Sejam a e b reais positivos e os elementos $u = (\sqrt{a}, \sqrt{b})$, $v = (\sqrt{b}, \sqrt{a}) \in I\!R^2$. Utilize a desigualdade de Cauchy–Schwarz para comparar a média aritmética $\dfrac{a + b}{2}$ com a média geométrica $\sqrt{a\, b}$.

(b) Seja V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$ e $\| \cdot \|_2$ a norma Euclidiana. Mostre que os elementos $(u - v)$, $(u + v) \in V$ são ortogonais se, e somente se, $\| u \|_2 = \| v \|_2$.

Questão 3. **(2.0 Pontos)**

Considere o espaço vetorial $\mathcal{P}_2(I\!R)$ munido do produto interno

$$\langle\, p \,,\, q \,\rangle \ = \ \int_{-1}^{1} t^2 p(t) q(t) dt \,.$$

A partir da base canônica $\beta = \{ 1, \ t, \ t^2 \}$ do espaço $\mathcal{P}_2(I\!R)$, construir uma base ortogonal $\gamma = \{ P_1, \ P_2, \ P_3 \}$ para o espaço $\mathcal{P}_2(I\!R)$.

Questão 4. **(2.0 Pontos)**

Considere o espaço vetorial $I\!R^4$ munido do produto interno usual. Seja U o subespaço gerado pelos elementos $u_1 = (1, -1, 1, 1)$ e $u_2 = (1, 2, 0, 1)$. Pede–se:

(a) Determine uma base para o subespaço $U^{\perp}$.

(b) Calcule a projeção ortogonal do elemento $u = (2, 1, 1, -1)$ no subespaço U.

(c) Considere o operador linear $P : I\!R^4 \longrightarrow I\!R^4$ que representa a projeção ortogonal sobre o subespaço U. Mostre que $Ker(I - P) = U$.

Questão 5. **(2.0 Pontos)**

Seja V um espaço vetorial real munido do produto interno $\langle \cdot , \cdot \rangle$ e $\| \cdot \|_2$ a norma Euclidiana. Considerando os elementos $u, v \in V$, com $v \neq 0_V$, determine o elemento w^* do conjunto $S = \{ w \in V \ / \ w = u - tv \ , \ t \in I\!R \}$ que possui a menor norma Euclidiana. Dê uma interpretação geométrica para o elemento w^*.

A.4.5 Segunda Chamada

Questão 1. **(2.0 Pontos)**

Seja V um espaço vetorial real e $\gamma = \{ v_1, v_2, v_3, v_4 \}$ uma base ordenada de V. Pede–se:

(a) Mostre que $\beta = \{ v_1 , v_1 + v_2 , v_1 + v_2 + v_3 , v_1 + v_2 + v_3 + v_4 \}$ é uma base de V.

(b) Se o elemento $v \in V$ tem como vetor de coordenadas

$$[v]_\gamma = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 2 \end{bmatrix}$$

determine seu vetor de coordenadas $[v]_\beta$.

Questão 2. **(2.0 Pontos)**

Considere os seguintes elementos do espaço vetorial $\mathbb{R}^4$

$$v_1 = (1, -1, 2, 3) \quad , \quad v_2 = (2, 1, -1, -2) \qquad \text{e} \qquad v_3 = (3, 3, -4, -7) .$$

Sejam U e V subespaços do $\mathbb{R}^4$ tais que $dim(U) = 3$ e $U \cap V = [v_1, v_2, v_3]$. Determine as possíveis dimensões dos subespaços V e $U + V$.

Questão 3. **(2.0 Pontos)**

Seja V o subespaço de $\mathbb{M}_2(\mathbb{R})$ das matrizes simétricas. Considere a transformação linear $T : V \longrightarrow \mathcal{P}_2(\mathbb{R})$ dada por:

$$T \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \right) = (a + b) - b\,x + (c - a + b)\,x^2 .$$

Mostre que T é um isomorfismo. Considerando a base canônica γ para o $\mathcal{P}_2(\mathbb{R})$ e a base canônica β para o subespaço V, determine a matriz $[T]^\beta_\gamma$.

Questão 4. **(2.0 Pontos)**

Mostre que a aplicação $\langle \cdot , \cdot \rangle : \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ dada por:

$$\langle p , q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

define um produto interno no espaço vetorial $\mathcal{P}_2(\mathbb{R})$. Determine uma base para o complemento ortogonal do subespaço $U = [2 - x]$ em $\mathcal{P}_2(\mathbb{R})$ com relação ao produto interno $\langle \cdot , \cdot \rangle$ definido acima.

Questão 5. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathcal{C}([-1,1])$ munido do produto interno usual. Determine o polinômio $p(x) = a + bx$, $a, b \in I\!R$, mais próximo da função $f(x) = \sin(\pi x)$, $x \in [-1,1]$, com relação à norma Euclidiana. Dê uma interpretação geométrica para o polinômio $p(x)$.

A.4.6 Exame

Questão 1. **(2.0 Pontos)**

Considere os elementos $u_1 = (-1, 2, 1, 1)$ e $u_2 = (2, 1, -1, 1)$ do $\mathbb{R}^4$. Pede–se:

(a) Encontre o complemento ortogonal do subespaço $W = [u_1, u_2]$ em $\mathbb{R}^4$ com relação ao produto interno usual.

(b) Encontre dois elementos $u_3, u_4 \in \mathbb{R}^4$ tais que $\beta = \{ u_1, u_2, u_3, u_4\}$ seja uma base para $\mathbb{R}^4$.

(c) Determine a projeção ortogonal do elemento $v = (1, 2, -1, 0) \in \mathbb{R}^4$ sobre W.

Questão 2. **(2.0 Pontos)**

Considere o seguinte subconjunto do espaço vetorial real $\mathbb{M}_n(\mathbb{R})$

$$U = \{ A \in \mathbb{M}_n(\mathbb{R}) \ / \ A^t = A \quad \text{e} \quad tr(A) = 0 \} .$$

Pede–se:

(a) Mostre que U é um subespaço vetorial de $\mathbb{M}_n(\mathbb{R})$.

(b) Considerando o espaço vetorial $\mathbb{M}_3(\mathbb{R})$, exiba uma base para o subespaço U.

Questão 3. **(2.0 Pontos)**

Seja $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ a transformação linear tal que

$$T(2,1) = (3,0,2) \qquad \text{e} \qquad T(1,2) = (1,1,0) .$$

Pede–se:

(a) Mostre que T é injetora.

(b) Exiba uma transformação linear $P : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ tal que $Ker(P) = Im(T)$.

Questão 4. **(2.0 Pontos)**

Considere o operador linear $T : \mathcal{P}_1(I\!R) \longrightarrow \mathcal{P}_1(I\!R)$ dado por:

$$T(p(x)) \;=\; p'(x) \;+\; (x\,+\,1)p(0)\,.$$

Sejam $\beta = \{\, 1,\, 1 - x \,\}$ e $\gamma = \{\, q(x),\, 1 - x \,\}$ bases para $\mathcal{P}_1(I\!R)$ tais que

$$[T]^{\beta}_{\gamma} \;=\; \begin{bmatrix} 2 & 1 \\ 1 & s \end{bmatrix}.$$

Pede–se:

(a) Determine o polinômio $q(x)$ e a constante $s \in I\!R$.

(b) T é um isomorfismo ? Justifique sua resposta.

Questão 5. **(2.0 Pontos)**

(a) Seja V um espaço vetorial real munido do produto interno $\langle \cdot \,,\, \cdot \rangle$ e T um operador linear sobre V. Se $\langle\, T(u)\,,\, T(v)\,\rangle = \langle\, u\,,\, v\,\rangle$ para todo $u,\, v \in V$, então T é um operador injetor.

(b) Considere o espaço vetorial $I\!R^n$ munido do produto interno usual $\langle\, x\,,\, y\,\rangle = y^t x$. Se $A \in I\!M_n(I\!R)$ é uma matriz simétrica, então $\langle\, Ax\,,\, y\,\rangle = \langle\, x\,,\, Ay\,\rangle$ para todo $x,\, y \in I\!R^n$.

A.5 Segundo Semestre de 2004

A.5.1 Primeira Prova

Questão 1. **(2.0 Pontos)**

Considere o espaço vetorial real $V = \{ (x,y) \in \mathbb{R}^2 \ / \ x > 0 \}$ com as operações:

Adição de Elementos: $(x_1,\ y_1) \oplus (x_2,\ y_2) = (x_1 x_2,\ y_1 + y_2)$

Multiplicação por Escalar: $\alpha \odot (x,\ y) = (x^\alpha,\ \alpha y) \quad , \quad \alpha \in \mathbb{R}$

Pede–se:

(a) Exiba o elemento neutro da operação $\oplus$.

(b) Exiba o inverso aditivo do elemento $v = (x,\ y) \in V$.

(c) Mostre que $\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v \quad , \quad u,\ v \in V$ e $\alpha \in \mathbb{R}$.

(d) Verifique se $W = \{ (x,\ y) \in V \ / \ x = 1 \}$ é um subespaço vetorial de V.

Questão 2. **(2.0 Pontos)**

Considere os seguintes subespaços de $\mathbb{R}^4$

$$W = \{ (x,\ y,\ z,\ t) \in \mathbb{R}^4 \ / \ x - y - z = 0 \quad \text{e} \quad x - 3y + t = 0 \}$$

$$U = [(1,\ 2,\ 1,\ 3),\ (3,\ 1,\ -1,\ 4)]$$

Determine uma base para os subespaços $U + W$ e $U \cap W$. O subespaço $U + W$ é uma soma direta dos subespaços U e W? Justifique sua resposta.

Questão 3. **(2.0 Pontos)**

Considere o seguinte subespaço S de $\mathcal{P}_3(\mathbb{R})$ definido da seguinte forma:

$$S = \{ \ p \in \mathcal{P}_3(\mathbb{R}) \ / \ p(-1) = 0 \quad \text{e} \quad p'(1) = 0 \ \} .$$

Qual é a dimensão de S? Encontre uma base para S.

Questão 4. **(2.0 Pontos)**

Seja $\Gamma = \{ v_1, v_2, v_3 \}$ uma base para o espaço vetorial real V. Pede–se:

(a) Mostre que $\beta = \{ v_1, v_1 + v_2, -v_1 + v_2 + v_3 \}$ é também uma base para V.

(b) Se o elemento $v \in V$ tem como vetor de coordenadas

$$[v]_\Gamma = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

em relação à base Γ, quais são as suas coordenadas em relação à base β ?

Questão 5. **(2.0 Pontos)**

Seja V um espaço vetorial sobre o corpo $\mathbb{F}$, com $dim(V) = 9$. Sejam U e W subespaços vetoriais de V tais que $dim(U) = 6$ e $dim(W) = 5$. Mostre que $2 \leq dim(U \cap W) \leq 5$.

A.5.2 Segunda Prova

Questão 1. **(2.0 Pontos)**

Considere o operador linear

$$\begin{array}{rccl} T: & \mathcal{P}_4(\mathbb{R}) & \longrightarrow & \mathcal{P}_4(\mathbb{R}) \\ & p & \longrightarrow & q = T(p) \end{array}$$

com $q(x) = T(p)(x) = x^2\,p''(x) \quad ; \quad x \in \mathbb{R}$. Pede–se:

(a) Determine a representação matricial de T com relação à base canônica.

(b) Determine o núcleo e a imagem do operador T.

(c) T é um operador linear injetor ? Justifique.

Questão 2. **(2.0 Pontos)**

(a) Exiba uma transformação linear $T : \mathbb{R}^3 \longrightarrow \mathcal{P}_2(\mathbb{R})$ tal que $dim(\,Ker(T)\,) = 1$.

(b) A transformação linear T é sobrejetora ? Justifique.

Questão 3. **(2.0 Pontos)**

Sejam F e G operadores lineares de um espaço V, tais que $G \circ F = F \circ G$. Mostre que $Ker(F) + Ker(G) \subset Ker(F \circ G)$.

Questão 4. **(2.0 Pontos)**

Sejam V e W espaços vetoriais sobre o corpo $\mathbb{R}$ e $\langle \cdot , \cdot \rangle$ um produto interno em W. Se $T : V \longrightarrow W$ é uma transformação linear injetora, então a aplicação $f(\cdot,\cdot)$ dada por:

$$f(u,v) = \langle\, T(u)\,,\, T(v)\,\rangle \qquad \text{para todo} \qquad u,\ v \in V$$

define um produto interno em V.

Questão 5. **(2.0 Pontos)**

Sejam $a_1, \cdots, a_n$ números reais estritamente positivos. Aplicando a desigualdade de Cauchy-Schwarz, em dois elementos $u,\ v \in \mathbb{R}^n$ escolhidos adequadamente, mostre que

$$(\, a_1 + \cdots + a_n\,)\left(\frac{1}{a_1} + \cdots + \frac{1}{a_n} \right) \geq n^2$$

A.5.3 Terceira Prova

Questão 1. **(2.0 Pontos)**

Seja W o subespaço de $\mathbb{R}^5$ gerado pelos vetores $u = (1,\ 2,\ 3,\ -1,\ 2)$ e $v = (2,\ 4,\ 7,\ 2,\ -1)$. Encontre uma base ortogonal para o complemento ortogonal $W^{\perp}$ de W em $\mathbb{R}^5$ com relação ao produto interno usual de $\mathbb{R}^5$.

Questão 2. **(2.0 Pontos)**

Determine todos os valores dos parâmetros a e b de modo que a matriz A dada abaixo seja diagonalizável. Para estes valores de a e b, determine uma matriz inversível P e a matriz diagonal D de modo que $P^{-1}AP = D$.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ a & 2 & 0 \\ 0 & b & 2 \end{pmatrix}.$$

Questão 3. **(2.0 Pontos)**

Determinar e classificar os pontos críticos das funções $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$, dadas abaixo, através da análise dos autovalores da matriz Hessiana, justificando sua resposta.

(a) $F(x,y) = 2x^2 - xy - 3y^2 - 3x + 7y$

(b) $F(x,y) = x^3 + y^3 - 3xy$

Questão 4. **(2.0 Pontos)**

Se $B \in \mathbb{M}_n(\mathbb{R})$ é similar a matriz $A \in \mathbb{M}_n(\mathbb{R})$ **auto–reflexiva**, mostre que B é uma matriz **auto–reflexiva**. Estabeleça a relação entre os autovetores de A e B.

Questão 5. **(2.0 Pontos)**

Considere o operador linear $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ dado por:

$$T(p)(x) = x^2 p''(x) + p'(x) + p(x) \quad ; \quad x \in \mathbb{R}.$$

(a) Determine a matriz $[T]_{\beta}^{\beta}$ onde β é a base canônica de $\mathcal{P}_3(\mathbb{R})$.

(b) Determine os autovalores e os autovetores do operador T.

(c) Para cada um dos autovalores do operador T, diga qual é o subespaço associado.

(d) Diga qual é a multiplicidade algébrica e geométrica de cada um dos autovalores do operador T. Justifique sua resposta.

(e) O operador T é diagonalizável ? Justifique sua resposta.

A.5.4 Segunda Chamada

ATENÇÃO: Faça as 2 (duas) questões relativas à prova que você faltou e resolva mais 3 (três) questões, dentre as outras 4 (quatro). Leia as questões com atenção.

Questão 1. – P1 **(2.0 Pontos)**

Sejam U e V subespaços de $I\!R^4$ tais que $dim(U) = 3$ e $dim(V) = 3$ e o subespaço $U \cap V = [(1,1,1,1),\ (1,-3,4,-2),\ (2,-2,5,-1)]$. Pede–se

(a) Qual é a dimensão do subespaço $U + V$? Justifique.

(b) Exiba uma base do $I\!R^4$ que contenha uma base do subespaço V, sabendo que o elemento $(0,0,0,1) \in V$.

Questão 2. – P1 **(2.0 Pontos)**

Seja $\Gamma = \{ v_1 , v_2 , v_3 , v_4 \}$ uma base para o espaço vetorial real $(E, +, \cdot)$.

(a) Mostre que $\beta = \{ v_1 , v_1 + v_2 , v_1 + v_2 + v_3 , v_1 + v_2 + v_3 + v_4 \}$ é também uma base para E.

(b) Se o elemento $v \in E$ tem como vetor de coordenadas

$$[v]_\Gamma = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

em relação à base Γ , quais são as suas coordenadas em relação à base β ?

Questão 3. – P2 **(2.0 Pontos)**

Seja V o subespaço de $I\!M_2(I\!R)$ das matrizes triangulares superiores. Pede–se:

(a) Exiba uma base para V.

(b) Seja $T : V \longrightarrow \mathcal{P}_2(I\!R)$ a transformação linear dada por:

$$T\left(\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \right) = (a + b) + b\,x + (c - a - b)\,x^2 .$$

Mostre que T é uma transformação inversível.

Questão 4. – P2 **(2.0 Pontos)**

Sejam $(V, +, \cdot)$ e $(W, +, \cdot)$ espaços vetoriais reais e $T : V \longrightarrow W$ uma transformação linear. Mostre que:

(a) $Ker(T) = \{ 0_V \}$ se, e somente se, T é injetora.

(b) Se T é injetora e $\{ v_1, v_2, \ldots, v_m \}$ é linearmente independente em V, então $\{ T(v_1), T(v_2), \ldots, T(v_m) \}$ é linearmente independente em W.

Questão 5. – P3 **(2.0 Pontos)**

Seja W o subespaço de $\mathbb{R}^4$ dado por

$$W = \{ (a+c, b+c, -b, -a) \in \mathbb{R}^4 \ / \ a, b, c \in R \} .$$

Encontre uma base ortonormal para o subespaço W com relação ao produto interno usual de $\mathbb{R}^4$.

Questão 6. – P3 **(2.0 Pontos)**

Seja $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ o operador linear dado por

$$T(x, y, z) = (3x - 2y - 4z, 4x - 3y - 4z, -z) . \quad \text{Pede–se :}$$

(a) Encontre $A = [T]_\beta^\beta$ onde β é a base canônica de $\mathbb{R}^3$

(b) Encontre os autovalores e os autovetores A

(c) Calcule A^9 de uma forma eficiente.

A.5.5 Exame

Questão 1. **(2.0 Pontos)**

Considere um subespaço vetorial W de $\mathcal{P}_2(\mathbb{R})$ com base $\beta = \{\, 1 + t,\ 1 - t^2 \,\}$. Sabemos que a matriz de mudança da base β para a base γ é dada por:

$$[I]_\gamma^\beta = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$

Determine a base γ do subespaço W.

Questão 2. **(2.0 Pontos)**

(a) Sejam U e V espaços vetoriais sobre o corpo $\mathbb{F}$ e T uma transformação linear de U em V. Se $dim(U) > dim(V)$, prove que existe um elemento não nulo $u \in U$ tal que $T(u) = 0_V$.

(b) Considerando $U = \mathbb{R}^3$ e $V = \mathcal{P}_1(\mathbb{R})$, dê um exemplo de uma transformação linear T de U em V que seja sobrejetora.

Questão 3. **(2.0 Pontos)**

Seja W o subespaço de $\mathbb{R}^4$ dado por:

$$W = \{\, (x, y, z, t) \in \mathbb{R}^4 \ / \ x + y = 0 \quad \text{e} \quad 2x - y + z = 0 \,\}\,.$$

Determine uma base ortogonal para cada um dos subespaços W e $W^\perp$.

Questão 4. **(2.0 Pontos)**

Seja $A \in \mathbb{M}_n(\mathbb{R})$ uma matriz simétrica. Pede–se:

(a) Considere o espaço vetorial $\mathbb{R}^n$ com o produto interno usual $\langle \cdot , \cdot \rangle$. Mostre que para todo $x,\ y \in \mathbb{R}^n$ temos que

$$\langle A\,x \,,\, y \rangle = \langle x \,,\, A\,y \rangle\,.$$

(b) Mostre que os autovalores de A são reais.

(c) Mostre que autovetores de A associados a autovalores distintos são ortogonais.

(d) Mostre que se a matriz $B \in \mathbb{M}_n(\mathbb{R})$ é ortogonalmente similar à matriz A, isto é, existe uma matriz ortogonal $Q \in \mathbb{M}_n(\mathbb{R})$ tal que $B = Q^t\,A\,Q$, então B é uma matriz simétrica.

Questão 5. **(2.0 Pontos)**

Considere o operador linear $T : \mathbb{M}_2(\mathbb{R}) \longrightarrow \mathbb{M}_2(\mathbb{R})$ definido da seguinte forma:

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2a+b & 2b \\ 2c & 3d \end{bmatrix} .$$

T é um operador linear diagonalizável ? Justifique.

A.6 Segundo Semestre de 1999

A.6.1 Primeiro Teste

Questão 1. **(5.0 Pontos)**

(a) Seja V um espaço vetorial e sejam U e W subespaços vetoriais de V. Mostre que o subconjunto de V dado por

$$U + W = \{ \; v \in V \; / \; v = u + w \quad ; \quad u \in U \quad e \quad w \in W \; \}$$

é também um subespaço vetorial de V.

(b) Dado o subespaço $W = \{ \, (x, \; y, \; z) \in \mathbb{R}^3 \; / \; x + 2y + z = 0 \, \}$ determine um subespaço U do $\mathbb{R}^3$ tal que $\mathbb{R}^3 = U \oplus W$. Justifique sua resposta.

Questão 2. **(5.0 Pontos)**

(a) Mostre que o subconjunto W de $\mathbb{M}_3(\mathbb{R})$ definido por

$$W = \{ \; A \in \mathbb{M}_3(\mathbb{R}) \; / \; A^t = -A \; \}$$

é um subespaço vetorial de $\mathbb{M}_3(\mathbb{R})$.

(b) Encontre uma base para o subespaço W. Qual a dimensão do subespaço W ? Justifique suas respostas.

A.6.2 Primeira Prova

Questão 1. **(3.0 Pontos)**

Seja $\Gamma = \{ v_1, v_2, v_3 \}$ uma base para o espaço vetorial $(E, +, \cdot)$. Pede–se:

(a) Mostre que $\beta = \{ v_1, v_1 + v_2, -v_1 + v_2 + v_3 \}$ é também uma base para E.

(b) Se o elemento $v \in E$ tem coordenadas $[v]_\Gamma = (2, -1, 1)$ em relação à base Γ, quais são as suas coordenadas em relação à base β ?

Questão 2. **(2.0 Pontos)**

Considere o espaço vetorial real

$$\mathcal{C}^\star([a,b]) = \{ f \in \mathcal{C}^1([a,b]) \ / \ f(a) = f(b) \} .$$

A aplicação $F : \mathcal{C}^\star([a,b]) \times \mathcal{C}^\star([a,b]) \longrightarrow I\!R$ dada por

$$F(f,g) = \int_a^b f'(x)g'(x)dx ,$$

define um produto interno em $\mathcal{C}^\star([a,b])$? Justifique sua resposta.

Questão 3. **(2.0 Pontos)**

Seja $(E, +, \cdot)$ um espaço vetorial e S um subconjunto de E com um número finito de elementos. Nas afirmações abaixo, demonstre se for verdadeira ou dê um contra–exemplo se for falsa:

(a) Se S é linearmente independente, então qualquer subconjunto de S é também linearmente independente;

(b) Se S é linearmente dependente, então qualquer subconjunto de S é também linearmente dependente.

Questão 4. **(3.0 Pontos)**

Sejam U e W subespaços vetoriais de dimensão 3 de $I\!R^4$ e seja

$$U \cap W = [(1, 2, 1, 0), (-1, 1, 0, 1), (1, 5, 2, 1)] .$$

Pede-se:

(a) $I\!R^4 = U + W$? Justifique sua resposta;

(b) Determine uma base ortonormal para o subespaço $U \cap W$;

(c) Determine a projeção do elemento $u = (1, 1, 1, 1)$ no subespaço $U \cap W$.

A.6.3 Segundo Teste

Questão 1. **(5.0 Pontos)**

Considere o espaço vetorial real $\mathbb{M}_3(\mathbb{R})$ munido do produto interno usual. Seja S o subespaço dado por: $S = \{ D \in \mathbb{M}_3(\mathbb{R}) \ / \ D$ é uma matriz diagonal $\}$. Determine o subespaço $S^{\perp}$, isto é, o complemento ortogonal de S. Encontre uma base para o subespaço $S^{\perp}$.

Questão 2. **(5.0 Pontos)**

Considere a seguinte transformação linear $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por:

$$T(x,y) = (2x, \, x - y, \, y) \, .$$

Pede–se:

(a) Determine a representação matricial de T com relação às bases canônicas.

(b) Determine uma base para a imagem da transformação T.

(c) T é uma transformação linear injetora ? Justifique.

(d) T é uma transformação linear sobrejetora ? Justifique.

A.6.4 Segunda Prova

Questão 1. **(2.5 Pontos)**

Considere a transformação linear $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ dada por:

$$T(p)(x) = xp''(x) + p(x) \quad ; \quad x \in \mathbb{R}.$$

(a) Determine a matriz $[T]^{\beta}_{\beta}$ onde β é a base canônica de $\mathcal{P}_3(\mathbb{R})$.

(b) Determine o núcleo e a imagem da transformação T.

(c) T é um isomorfismo ? Justifique.

Questão 2. **(2.5 Pontos)**

Sejam A , $B \in \mathbb{M}_n(\mathbb{R})$ duas matrizes similares, isto é, existe uma matriz inversível $P \in \mathbb{M}_n(\mathbb{R})$ tal que $A = P^{-1}\,B\,P$. Estabeleça a relação entre os autopares de A e B.

Questão 3. **(2.5 Pontos)**

(a) Seja $A \in \mathbb{M}_n(\mathbb{R})$ inversível. Estabeleça a relação entre os autopares de A e A^{-1} .

(b) Seja $A \in \mathbb{M}_n(\mathbb{R})$. Mostre que as matrizes A e A^t possuem os mesmos autovalores. **Sugestão:** utilize o polinômio característico.

Questão 4. **(2.5 Pontos)**

Determine todos os valores dos parâmetros a e b de modo que a matriz A dada abaixo seja diagonalizável. Para estes valores de a e b determine uma matriz inversível P e a matriz diagonal D de modo que $P^{-1}AP = D$.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ a & 2 & 0 \\ 0 & b & 2 \end{pmatrix} .$$

A.6.5 Exame

Questão 1. **(2.0 Pontos)**

Considere o seguinte subespaço de $\mathcal{P}_3([-1,1])$

$$S = \{ \; p \in \mathcal{P}_3([-1,1]) \; / \; p(-1) = 0 \;\text{ e }\; p'(1) = 0 \; \}.$$

Qual é a dimensão de S ? Encontre uma base para S.

Questão 2. **(2.0 Pontos)**

Seja $(E, +, \cdot)$ um espaço vetorial real munido do produto interno $\langle \, . \, , \, . \, \rangle$. Considere a função $J : I\!R \longrightarrow I\!R$ definida da seguinte forma: dados os elementos u , $v \in E$, não nulos, seja

$$J(\alpha) = \| \; v - \alpha u \; \|^2 \quad ; \quad \alpha \in I\!R \, ,$$

onde $\| \cdot \|$ é a norma proveniente do produto interno.

(a) Mostre que a função J possui um único ponto de mínimo α^*.

(b) Dê uma interpretação para o elemento $\alpha^* u$.

Questão 3. **(2.0 Pontos)**

Considere a transformação linear $T : \mathcal{P}_2(I\!R) \longrightarrow I\!R^3$ definida da seguinte forma

$$T(\, a_0 + a_1 t + a_2 t^2 \,) = (\, a_0 + a_1 + a_2 \, , \; a_0 + a_1 \, , \; a_0 + a_2 \,) \, .$$

(a) Determine a matriz $[T]^{\Gamma}_{\beta}$ onde Γ é a base canônica de $\mathcal{P}_2(I\!R)$ e β é a base canônica de $I\!R^3$.

(b) Mostre que T é um isomorfismo e determine a expressão da transformação linear $T^{-1}(\, x \, , \, y \, , \, z \,)$ para todo $(\, x \, , \, y \, , \, z \,) \in I\!R^3$.

Questão 4. **(2.0 Pontos)**

Utilize a forma diagonal da matriz A para calcular eficientemente A^n, $n \in I\!N$.

$$A = \begin{pmatrix} 3 & -5 \\ 1 & -3 \end{pmatrix}.$$

Questão 5. **(2.0 Pontos)**

Considere $(E, +, \cdot)$ um espaço vetorial munido do produto interno $\langle \cdot , \cdot \rangle$ e um operador linear **auto-adjunto** $T : \; E \longrightarrow E$, isto é, $\langle \, T(v) \, , \, u \, \rangle = \langle \, v \, , \, T(u) \, \rangle$ para todo $u, v \in E$. Sejam λ_1 e λ_2 autovalores distintos de T e v_1 e v_2 os autovetores associados a λ_1 e λ_2 , respectivamente. Então, v_1 e v_2 são ortogonais.

A.7 Segundo Semestre de 1998

A.7.1 Primeiro Teste

Questão 1. **(5.0 Pontos)**

Considere o espaço vetorial real $\mathbb{M}_2(\mathbb{R})$. Dadas as seguintes matrizes:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

identificar o subespaço gerado pelos elementos do conjunto $S = \{ A_1, A_2, A_3 \}$.

Questão 2. **(5.0 Pontos)**

Seja E um espaço vetorial sobre um corpo $\mathbb{F}$. Mostre que, se U e W são subespaços vetoriais de E , então $U + W$ também é um subespaço vetorial de E.

Questão 3. **(5.0 Pontos)**

Considere o sistema linear homogêneo

$$\begin{cases} 2x + 4y + z = 0 \\ x + y + 2z = 0 \\ x + 3y - z = 0 \end{cases}$$

Mostre que

$$S = \{ (x, y, z) \in \mathbb{R}^3 \ / \ (x, y, z) \text{ é solução do sistema linear homogêneo} \}$$

é um subespaço do $\mathbb{R}^3$. Encontre os vetores geradores do subespaço S.

A.7.2 Primeira Prova

Questão 1. **(2.0 Pontos)**

Seja V um espaço vetorial sobre o corpo F. Mostre que se $\{ v_1, v_2, v_3 \}$ é linearmente independente, então $\{ v_1+v_2, v_1+v_3, v_2+v_3 \}$ também é linearmente independente em V.

Questão 2. **(2.0 Pontos)**

Considere o subespaço W de $I\!R^3$ gerado pelos elementos

$$w_1 = (1,-1,0,0) \quad , \quad w_2 = (0,0,1,1) \quad , \quad w_3 = (-2,2,1,1) \quad \text{e} \quad w_4 = (1,0,1,0) \, .$$

Pede–se:

(a) Determine uma base para W.

(b) O elemento $u = (2,-3,2,2) \in W$?

Questão 3. **(2.0 Pontos)**

(a) Dado o subespaço $S = \{ (x,y,z) \in I\!R^3 \;/\; x + 2y + z = 0 \}$ do $I\!R^3$. Determine um subespaço W do $I\!R^3$ tal que $I\!R^3 = S \oplus W$.

(b) Dê exemplos de dois subespaços, S e W, de dimensão dois em $I\!R^3$ tais que $I\!R^3 = S + W$. Temos uma soma direta ?

(c) Ilustre com um exemplo a seguinte proposição:

"Se S e W são dois subespaços de um espaço vetorial V de dimensão finita, então $dim(S+W) = dim(S) + dim(W) - dim(S \cap W)$".

Questão 4. **(2.0 Pontos)**

Considere o seguinte subconjunto do $I\!R^3$

$$\gamma = \{ (1,0,2), (0,1,-1), (1,0,1) \} \, .$$

Pede–se:

(a) Mostre que γ é uma base para $I\!R^3$ e determine a matriz de mudança da base γ para a base canônica β.

(b) Dado o elemento $u = (1,1,1)$. Determine o vetor de coordenadas $[u]_\gamma$.

Questão 5. **(2.0 Pontos)**

Considere o espaço vetorial real $\mathbb{M}_2(\mathbb{R})$ munido do produto interno usual, isto é, $\langle A , B \rangle = tr(B^t A)$. Seja S o subespaço de $\mathbb{M}_2(\mathbb{R})$ gerado pelo elemento

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} .$$

Pede–se:

(a) Dada a matriz

$$B = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix} .$$

Determine sua projeção ortogonal sobre o subespaço S.

(b) Determine uma matriz $C \in \mathbb{M}_2(\mathbb{R})$ tal que $\langle A , C \rangle = 0$.

A.7.3 Segundo Teste

Questão 1. **(5.0 Pontos)**

Considere o espaço vetorial $I\!R^3$ munido do produto interno usual. Seja S o subespaço gerado pelo vetor $u = (-1,\ 1,\ -1)$. Pede–se:

(a) Determine o complemento ortogonal de S, isto é, o subespaço $S^{\perp}$.

(b) Determine uma base ortogonal para o subespaço $S^{\perp}$.

(c) Dado o vetor $v = (3,\ 1,\ -1)$ calcule sua projeção no subespaço $S^{\perp}$.

Questão 2. **(5.0 Pontos)**

Considere a seguinte transformação linear $T : I\!R^3 \longrightarrow I\!R^3$ dada por:

$$T(x,y,z) = (\ x - y + z\ ,\ x + y + 2z\ ,\ x + 5y + 4z)\ .$$

Pede–se:

(a) Determine a representação matricial de T com relação à base canônica.

(b) Determine $Ker(T)$, $dim(Ker(T))$, $Im(T)$ e $dim(Im(T))$.

(c) T é uma transformação linear injetora ?

(d) Determine o complemento ortogonal de $Im(T)$.

A.7.4 Segunda Prova

Questão 1. **(2.5 Pontos)**

Considere a seguinte transformação linear $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ dada por:

$$T(x, y, z) = (y + z , x + z , x + y + 2z) ,$$

e as seguintes bases para o espaço $\mathbb{R}^3$

$$\beta = \{ (1, 0, 0), (0, 1, 0), (0, 0, 1) \}$$

$$\Gamma = \{ (1, 0, 1), (0, 1, 1), (0, 0, 1) \}$$

Pede–se:

(a) Determine $[T]_{\beta}^{\beta}$ e $[T]_{\beta}^{\Gamma}$.

(b) Determine $Ker(T)$, $dim(Ker(T))$, $Im(T)$ e $dim(Im(T))$.

(c) Encontre os autovalores e autovetores da transformação T.

(d) Encontre uma base α para o $\mathbb{R}^3$ de modo que $[T]_{\alpha}^{\alpha}$ seja uma matriz diagonal.

(e) Dado o vetor $v = (1, 1, -1)$ calcule sua projeção no subespaço $Im(T)$.

Questão 2. **(2.5 Pontos)**

Para quais valores do parâmetro a as matrizes abaixo são diagonalizáveis ?

$$(a) \quad A = \begin{pmatrix} 1 & 1 \\ 0 & a \end{pmatrix} \qquad (b) \quad B = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

Questão 3. **(2.5 Pontos)**

Mostre que a matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ é semelhante à matriz $D = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$.
Calcule de maneira eficiente A^n para $n \in \mathbb{N}$.

Questão 4. **(2.5 Pontos)**

Provar que se λ é um autovalor de A com o autovetor associado v , então λ^n é um autovalor de A^n com o autovetor associado v.

A.7.5 Exame

Questão 1. **(2.0 Pontos)**

Seja E um espaço vetorial sobre um corpo $\mathbb{F}$. Mostre que, se $\{ v_1 , v_2 , v_3 \}$ é um conjunto linearmente independente em E, então $\{ v_1 + v_2 , v_1 + v_3 , v_2 + v_3 \}$ também é um conjunto linearmente independente em E.

Questão 2. **(2.0 Pontos)**

Considere os seguintes subespaços de $\mathbb{R}^4$

$$U = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x - y - z + t = 0 \}$$

$$W = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x + y = 0 \quad e \quad z - t = 0 \}$$

Pede–se:

(a) Determine uma base para o subespaço $U \cap W$.

(b) Determine uma base para o subespaço $U + W$.

(c) Determine uma base ortogonal para o subespaço U.

(d) $\mathbb{R}^4 = U \oplus W$? Justifique sua resposta.

(e) Calcule a projeção ortogonal de $v = (3, 1, -1, 2)$ no subespaço $U \cap W$.

Questão 3. **(2.0 Pontos)**

Considere o espaço Euclidiano $\mathbb{R}^4$ munido do produto interno usual. Seja S o subespaço gerado pelos elementos do seguinte subconjunto:

$$\beta = \{ (1 , 1 , 1 , 1) , (1 , -1 , 1 , 1) \} \subset \mathbb{R}^4 .$$

Determine uma base ortogonal para $S^{\perp}$, através do processo de Gram-Schmidt.

Questão 4. **(2.0 Pontos)**

Considere a transformação linear $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ dada por:

$$T(x,y,z) = (y + z , x + z , x + y + 2z) ,$$

e as seguintes bases para o espaço $\mathbb{R}^3$

$$\beta = \{ (1, 0, 0), (0, 1, 0), (0, 0, 1) \} \quad \text{e} \quad \Gamma = \{ (1, 0, 1), (0, 1, 1), (0, 0, 1) \} .$$

Pede–se:

(a) Determine $[T]^{\beta}_{\beta}$ e $[T]^{\beta}_{\Gamma}$.

(b) Determine $Ker(T)$, $dim(Ker(T))$, $Im(T)$ e $dim(Im(T))$.

(c) Encontre os autovalores e autovetores da transformação T.

(d) Encontre uma base α para o $\mathbb{R}^3$ de modo que $[T]^{\alpha}_{\alpha}$ seja uma matriz diagonal e exiba a matriz $[T]^{\alpha}_{\alpha}$. Justifique sua resposta.

Questão 5. **(2.0 Pontos)**

Estudar quanto à possibilidade de diagonalização das matrizes dadas abaixo. Determine a matriz inversível P e a matriz diagonal D de modo que $P^{-1}AP = D$.

(a) $A = \begin{pmatrix} 1 & 0 & 0 \\ a & 2 & 0 \\ 0 & b & 2 \end{pmatrix}$ **(b)** $A = \begin{pmatrix} 1 & 1 \\ 0 & a \end{pmatrix}$.

B

Gabarito das Avaliações

Conteúdo

B.1 Segundo Semestre de 2006

B.1.1 Primeira Prova

Questão 1.

Vamos verificar se o elemento neutro da adição $0_{\mathbb{R}^2}$ pertence ao subconjunto $U \cup W$ e se o subconjunto $U \cup W$ é fechado com relação à operação de adição de elementos e com relação à operação de multiplicação por escalar.

Como U e W são subespaços vetoriais de $\mathbb{R}^2$, temos que $0_{\mathbb{R}^2}$ pertence tanto a U quanto a W. Logo, $0_{\mathbb{R}^2} \in U \cup W$. Note que, $U \cap W = \{ 0_{\mathbb{R}^2} \}$.

Considere um elemento $v \in U \cup W$, isto é, $v \in U$ ou $v \in W$. Assim, para qualquer $\lambda \in \mathbb{R}$ temos que $\lambda v \in U \cup W$, pois $\lambda v \in U$ ou $\lambda v \in W$.

Finalmente, tomando os elementos $v_1, v_2 \in U \cup W$, temos três possibilidades.

A primeira, consideramos que $v_1, v_2 \in U$. Como U é um subespaço vetorial de $\mathbb{R}^2$, temos que $v_1 + v_2 \in U$. Logo, $v_1 + v_2 \in U \cup W$.

A segunda, consideramos que $v_1, v_2 \in W$. Como W é um subespaço vetorial de $\mathbb{R}^2$, temos que $v_1 + v_2 \in W$. Logo, $v_1 + v_2 \in U \cup W$.

A terceira, consideramos que $v_1 \in U$ e $v_2 \in W$. Assim, temos que

$$v_1 = (x_1 , 3x_1) \qquad \text{e} \qquad v_2 = (x_2 , -2x_2) .$$

Logo, $v_1 + v_2 = (x_1 + x_2 , 3x_1 - 2x_2)$. Portanto, temos que

$$v_1 + v_2 \notin U \qquad \text{e} \qquad v_1 + v_2 \notin W .$$

Desse modo, $v_1 + v_2 \notin U \cup W$. Assim, mostramos que $U \cup W$ não é um subespaço vetorial de $\mathbb{R}^2$, pois o subconjunto $U \cup W$ não é fechado com relação à operação de adição de elementos.

Questão 2.

Inicialmente vamos provar que

$$\{\, u,\, v,\, w \,\} \ \text{LI} \quad \Longrightarrow \quad \{\, u+v,\, u+w,\, v+w \,\} \ \text{LI} \,.$$

Tomando a combinação linear nula

$$a\,(u + v) \ + \ b\,(u + w) \ + \ c\,(v + w) \ = \ 0_V$$

obtemos

$$(a + b)u \ + \ (a + c)v \ + \ (b + c)w \ = \ 0_V \,.$$

Utilizando a hipótese que o conjunto $\{\, u,\, v,\, w \,\}$ é linearmente independente, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a + b = 0 \\ a + c = 0 \\ b + c = 0 \end{cases}$$

que possui somente a solução trivial $a = b = c = 0$. Portanto, provamos que o conjunto $\{\, u+v,\, u+w,\, v+w \,\}$ é linearmente independente.

Finalmente, vamos provar que

$$\{\, u + v,\, u + w,\, v + w \,\} \;\text{ LI } \quad \Longrightarrow \quad \{\, u,\, v,\, w \,\} \;\text{ LI }.$$

Equivalentemente, podemos provar que

$$\{\, u,\, v,\, w \,\} \;\text{ LD } \quad \Longrightarrow \quad \{\, u + v,\, u + w,\, v + w \,\} \;\text{ LD }.$$

Tomando a combinação linear nula

$$a\,(u \,+\, v) \;\;+\;\; b\,(u \,+\, w) \;\;+\;\; c\,(v \,+\, w) \;\;=\;\; 0_V$$

obtemos

$$(a \,+\, b)u \;\;+\;\; (a \,+\, c)v \;\;+\;\; (b \,+\, c)w \;\;=\;\; 0_V\,.$$

Utilizando a hipótese que o conjunto $\{\, u,\, v,\, w \,\}$ é linearmente dependente, temos que os coeficientes da combinação linear acima não são todos nulos, isto é,

$$(a \,+\, b) \quad , \quad (a \,+\, c) \qquad \text{e} \qquad (b \,+\, c)$$

não são todos nulos. Assim, existem escalares $a,\, b,\, c \,\in\, \mathbb{F}$ não todos nulos tais que

$$a\,(u \,+\, v) \;\;+\;\; b\,(u \,+\, w) \;\;+\;\; c\,(v \,+\, w) \;\;=\;\; 0_V\,.$$

Portanto, mostramos que o conjunto $\{\, u + v,\, u + w,\, v + w \,\}$ é linearmente dependente.

Assim, provamos que

$$\{\, u + v,\, u + w,\, v + w \,\} \;\text{ LI } \quad \Longrightarrow \quad \{\, u,\, v,\, w \,\} \;\text{ LI }.$$

completando a resolução da questão.

Questão 3.

(a)

Vamos determinar uma base para o subespaço U. Note que, toda matriz $A \in U$ é escrita da seguinte forma:

$$A = a\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + b\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + c\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} \quad \text{para} \quad a, b, c \in I\!R.$$

Tomando a combinação linear nula

$$a\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + b\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + c\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

obtemos que $a = b = c = 0$, são os únicos escalares que satisfazem o sistema acima. Assim, mostramos que o conjunto

$$\left\{ \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} \right\}$$

é uma base para o subespaço U, pois gera o subespaço U e é linearmente independente. Logo, temos que $dim(U) = 3$.

Vamos determinar uma base para o subespaço W. Note que, toda matriz $A \in W$ é escrita da seguinte forma:

$$A = a\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix} + b\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} \quad \text{para} \quad a, b \in I\!R.$$

Tomando a combinação linear nula

$$a\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix} + b\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

obtemos que $a = b = 0$, são os únicos escalares que satisfazem o sistema acima. Assim, mostramos que o conjunto

$$\left\{ \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} \right\}$$

é uma base para o subespaço W, pois gera o subespaço W e é linearmente independente. Logo, temos que $dim(U) = 2$.

Agora, vamos determinar uma base para o subespaço $U \cap W$. Considere uma matriz $A \in U \cap W$, isto é, $A \in U$ e $A \in W$. Assim, temos que A é escrita como:

$$A = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = d\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + e\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

para $a, b, c, d, e \in \mathbb{R}$.

Desse modo, obtemos o seguinte sistema linear

$$\begin{cases} a = 0 \\ b = d \\ c = -d \\ a = e \end{cases}$$

cuja solução é $a = 0$, $b = d$, $c = -d$ e $e = 0$. Portanto, toda matriz $A \in U \cap W$ é escrita como

$$A = d\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{para} \quad d \in \mathbb{R}.$$

Assim, temos que

$$\left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$

é uma base para o subespaço $U \cap W$. Logo, $dim(U \cap W) = 1$.

Finalmente, vamos determinar uma base para o subespaço $U + W$. Pelos resultados obtidos acima, sabemos que

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W) = 4,$$

e que o subespaço $U + W$ tem por um sistema de geradores o seguinte conjunto

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Podemos observar que

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Portanto, temos que o conjunto

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \;,\; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \;,\; \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \;,\; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é uma base para o subespaço $U + W$.

(b)

Como $U + W$ é um subespaço de $\mathbb{M}_2(\mathbb{R})$ e $dim(U + W) = dim(\mathbb{M}_2(\mathbb{R}))$, temos que $U + W = \mathbb{M}_2(\mathbb{R})$, entretanto, não como soma direta, pois $U \cap W \neq \{ 0_{\mathbb{M}_2(\mathbb{R})} \}$.

Questão 4.

(a)
Conhecendo a matriz de mudança de base $[I]_{\alpha}^{\gamma}$, temos que

$$\begin{cases} u_1 = v_1 + 4v_2 \\ u_2 = \quad - 2v_2 \end{cases}$$

Resolvendo o sistema acima, obtemos

$$\begin{cases} v_1 = u_1 + 2u_2 \\ v_2 = \quad - \dfrac{1}{2}u_2 \end{cases}$$

Logo, $v_1 = (-3, 5)$ e $v_2 = (1, -1)$, que são os elementos da base ordenada α.

(b)
Conhecendo a matriz de mudança de base $[I]_{\alpha}^{\gamma}$ e o vetor de coordenadas $[u]_{\alpha}$, temos que

$$[u]_{\alpha} = [I]_{\alpha}^{\gamma}\,[u]_{\gamma}.$$

Chamando $[u]_{\gamma} = \begin{bmatrix} a \\ b \end{bmatrix}$, obtemos o seguinte sistema linear

$$\begin{bmatrix} 1 & 0 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

cuja única solução é $a = b = 1$. Portanto, temos que

$$u = a\,u_1 + b\,u_2 = u_1 + u_2 = (-1, 3)\,.$$

Observe que podemos obter o elemento u a partir do vetor de coordenadas $[u]_{\alpha} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ e da base ordenada α obtida no item **(a)**. Desse modo, temos que

$$u = v_1 + 2v_2 = (-1, 3)\,.$$

B.1.2 Segunda Prova

Questão 1.

Da condição **(a)**, temos que $T(1 + x) = (0,0,0)$.

Da condição **(b)**, isto é, $q(x) = x \notin Ker(T)$, implica que o elemento $r(x) = 1$ não pode pertencer ao $Ker(T)$, pois podemos escrever $q(x) = p(x) - r(x)$. Claramente, se o elemento $r(x) \in Ker(T)$, então $q(x) \in Ker(T)$, o que contradiz a hipótese.

Da condição **(c)**, isto é, $Im(T) = [(1,1,1)]$, temos que $dim(\, Im(T)\,) = 1$.

Logo, pelo Teorema do Núcleo e da Imagem, devemos ter $dim(\, Ker(T)\,) = 2$, pois $dim(\, \mathcal{P}_2(I\!R)\,) = 3$.

Assim, podemos considerar $T : \mathcal{P}_2(I\!R) \longrightarrow I\!R^3$, com $\{\, 1 + x \,,\, x^2 \,\}$ uma base ordenada para $Ker(T)$, definida da seguinte forma:

$$T(1 + x) = (0,0,0) \quad , \quad T(x^2) = (0,0,0) \quad , \quad T(x) = (1,1,1) \,,$$

onde estamos escolhendo $\gamma = \{\, 1 + x \,,\, x^2 \,,\, x \,\}$ uma base ordenada para $\mathcal{P}_2(I\!R)$, que foi obtida completando a base do $Ker(T)$.

Finalmente, vamos determinar a expressão da transformação linear T definida acima.

Para isso, tomamos um elemento genérico $p(x) = a + bx + cx^2 \in \mathcal{P}_2(I\!R)$ que é representado com relação à base ordenada γ de $\mathcal{P}_2(I\!R)$ da seguinte forma:

$$\begin{aligned} p(x) &= d_1(\, 1 + x \,) + d_2x^2 + d_3x \\ &= d_1 + (d_1 + d_3)x + d_2x^2 \,, \end{aligned}$$

obtendo o seguinte sistema linear

$$\begin{cases} \phantom{d_1 + {}} d_1 = a \\ d_1 + d_3 = b \\ \phantom{d_1 + {}} d_2 = c \,, \end{cases}$$

que possui somente a solução $d_1 = a$, $d_2 = c$ e $d_3 = b - a$. Desse modo, temos que

$$p(x) = a(\, 1 + x \,) + cx^2 + (b - a)x \,.$$

Agora, fazendo $T(p(x)) = T(a + bx + cx^2)$, obtemos

$$T(a + bx + cx^2) = a\,T(1 + x) + c\,T(x^2) + (b - a)\,T(x) = (b - a)(1, 1, 1)\,.$$

Assim, encontramos uma transformação linear T com as propriedades pedidas.

Questão 2.

(a) Sabendo que $T(v) = v$ para todo $v \in S$ e que $T(v_4) = v_1 + v_3$, obtemos

$$\begin{aligned}
T(v_1) &= v_1 &&= 1v_1 + 0v_2 + 0v_3 + 0v_4 \\
T(v_2) &= v_2 &&= 0v_1 + 1v_2 + 0v_3 + 0v_4 \\
T(v_3) &= v_3 &&= 0v_1 + 0v_2 + 1v_3 + 0v_4 \\
T(v_4) &= v_1 + v_3 &&= 1v_1 + 0v_2 + 1v_3 + 0v_4 \,.
\end{aligned}$$

Portanto, temos que

$$[T]_{\gamma}^{\gamma} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(b) Conhecemos as matrizes $[I]_{\gamma}^{\beta}$ e $[T]_{\gamma}^{\gamma}$. Assim, para obter $[T(e_1)]_{\gamma}$, vamos determinar inicialmente $[e_1]_{\gamma}$ da seguinte forma:

$$[e_1]_{\gamma} = [I]_{\gamma}^{\beta}\,[e_1]_{\beta} \quad \Longrightarrow \quad [e_1]_{\gamma} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Finalmente, vamos calcular

$$[T(e_1)]_{\gamma} = [T]_{\gamma}^{\gamma}\,[e_1]_{\gamma} \quad \Longrightarrow \quad [T(e_1)]_{\gamma} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

o que completa a resolução da questão.

Questão 3.

(a) Dado um polinômio $p(x) = a + bx + cx^2 \in \mathcal{P}_2(I\!R)$, vamos calcular

$$T(p(x)) = p'(x) + p(x) = (b+a) + (b+2c)x + cx^2 ,$$

para em seguida calcular

$$\begin{aligned}(P \circ T)(p(x)) &= P(T(p(x)) = P(\, (b+a) + (b+2c)x + cx^2)\,) \\ &= (\, a + 2b + 2c \, , \, c \, , \, a - 2c \,) \, .\end{aligned}$$

Portanto, a transformação linear $P \circ T : \mathcal{P}_2(I\!R) \longrightarrow I\!R^3$ é dada por:

$$(P \circ T)(p(x)) = (\, a + 2b + 2c \, , \, c \, , \, a - 2c \,)$$

para todo $p(x) = a + bx + cx^2 \in \mathcal{P}_2(I\!R)$.

(b) Considerando $\beta = \{\, 1 \, , \, x \, , \, x^2 \,\}$ a base canônica de $\mathcal{P}_2(I\!R)$ e $\gamma = \{\, e_1, \, e_2, \, e_3 \,\}$ a base canônica de $I\!R^3$, vamos determinar a matriz $[P \circ T]^{\beta}_{\gamma}$.

Para isso, vamos calcular

$$\begin{aligned}(P \circ T)(1) &= (1,0,1) &&= 1e_1 + 0e_2 + 1e_3 \\ (P \circ T)(x) &= (2,0,0) &&= 2e_1 + 0e_2 + 0e_3 \\ (P \circ T)(x^2) &= (2,1,-2) &&= 2e_1 + 1e_2 - 2e_3 \, .\end{aligned}$$

Portanto, obtemos

$$[P \circ T]^{\beta}_{\gamma} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & -2 \end{bmatrix} .$$

(c) Vamos verifique se P é um isomorfismo de $\mathcal{P}_2(\mathbb{R})$ em $\mathbb{R}^3$.

Para isso, basta verificar se $Ker(P) = \{ 0_{\mathcal{P}_2(\mathbb{R})} \}$.

Tomando um elemento $p(x) = a + bx + cx^2 \in Ker(P)$, isto é,

$$P(a + bx + cx^2) = (a + b\,,\, c\,,\, a - b) = (0, 0, 0)\ ,$$

obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a + b &= 0 \\ \quad\;\; c &= 0 \\ a - c &= 0\ , \end{cases}$$

que possui somente a solução trivial $a = b = c = 0$. Logo, $Ker(P) = \{ 0_{\mathcal{P}_2(\mathbb{R})} \}$.

Portanto, temos que P é um isomorfismo de $\mathcal{P}_2(\mathbb{R})$ em $\mathbb{R}^3$.

Vamos determinar o isomorfismo inverso. Dado um elemento $(a, b, c) \in \mathbb{R}^3$ tal que

$$P^{-1}(a, b, c) = d_1 + d_2 x + d_3 x^2\ ,$$

temos que

$$P(d_1 + d_2 x + d_3 x^2) = (a, b, c) \quad\Longrightarrow\quad (\, d_1 + d_2\,,\, d_3\,,\, d_1 - d_2\,) = (a, b, c)\ .$$

Assim, obtemos o seguinte sistema linear

$$\begin{cases} d_1 + d_2 &= a \\ \qquad d_3 &= b \\ d_1 - d_2 &= c\ , \end{cases}$$

que possui somente a solução

$$d_1 = \frac{a + c}{2} \quad , \quad d_2 = \frac{a - c}{2} \qquad \text{e} \qquad d_3 = b\ .$$

Portanto, obtemos

$$P^{-1}(a, b, c) = \left(\frac{a + c}{2} \right) + \left(\frac{a - c}{2} \right) x + b x^2\ ,$$

o que completa a resolução da questão.

Questão 4.

(a) Sabemos que $dim(V) = n$ e que $Im(T) = Ker(T)$.

Assim, podemos afirmar que $dim(\,Im(T)\,) = dim(\,Ker(T)\,) = m$.

Pelo Teorema do Núcleo e da Imagem, temos que

$$dim(\,Im(T)\,) \; + \; dim(\,Ker(T)\,) \; = \; dim(V) \,.$$

Portanto, temos que $n = 2m$. Logo, podemos concluir que n é par e que $m = \dfrac{n}{2}$.

(b) Considerando $V = \mathbb{R}^4$, temos que $dim(\,Im(T)\,) = dim(\,Ker(T)\,) = 2$.

Tomando a base canônica $\beta = \{\,e_1,\, e_2,\, e_3,\, e_4\,\}$ para $\mathbb{R}^4$, vamos definir um operador linear T sobre $\mathbb{R}^4$, com as propriedades acima, da seguinte forma:

$$\begin{aligned}
T(e_1) &= (0,0,0,0) = 0_{\mathbb{R}^4} \\
T(e_2) &= (0,0,0,0) = 0_{\mathbb{R}^4} \\
T(e_3) &= (1,0,0,0) = e_1 \\
T(e_4) &= (0,1,0,0) = e_2 \,.
\end{aligned}$$

Podemos observar facilmente que $\{\,e_1,\, e_2\,\}$ é uma base para o subespaço $Ker(T)$ e também para o subespaço $Im(T)$. Logo, temos que

$$dim(\,Ker(T)\,) = dim(\,Im(T)\,) = 2 \qquad \text{e} \qquad Ker(T) = Im(T) \,.$$

Portanto, o operador linear T, definido acima, possui as propriedades desejadas. Podemos verificar facilmente que

$$T(x,y,z,t) = (z,t,0,0) \qquad \text{para todo} \qquad (x,y,z,t) \in \mathbb{R}^4 \,.$$

B.1.3 Terceira Prova

Questão 1.

Seja $w \in T(E_\lambda)$, isto é, existe um elemento $v \in E_\lambda$ tal que $w = T(v)$. Como $v \in E_\lambda$, temos que $w = T(v) = \lambda v$. Logo, $w = \lambda v$.

Aplicando o operador T no elemento w, obtemos $T(w) = \lambda T(v)$.

Como $w = T(v)$, temos que $T(w) = \lambda w$. Assim, podemos concluir que $w \in E_\lambda$. Portanto, provamos que $T(E_\lambda) \subset E_\lambda$.

Questão 2.

(a) Temos somente um autovalor λ associado ao autovetor v. Podemos observar que o autovalor λ é unicamente determinado pelo operador T e pelo autovetor v. De fato, considere que λ e λ' são autovalores do operador T associados ao autovetor v, isto é,

$$T(v) = \lambda\, v \qquad \text{e} \qquad T(v) = \lambda'\, v\,.$$

Assim, temos que

$$\lambda\, v - \lambda'\, v = 0_V \implies (\lambda - \lambda')\, v = 0_V \implies (\lambda - \lambda') = 0 \implies \lambda = \lambda'\,,$$

pois $v \neq 0_V$.

(b) Sim. De fato, se $\lambda = 0$ é um autovalor de T e v um autovetor associado, temos que $v \in Ker(T)$, pois $T(v) = \lambda\, v = 0_V$. Logo, como $v \neq 0_V$, $Ker(T) \neq \{ 0_V \}$. Portanto, T não é um operador injetor.

Reciprocamente, se T não é um operador injetor, sabemos que $Ker(T) \neq \{ 0_V \}$. Logo, os elementos não nulos $v \in Ker(T)$ são autovetores do operador T associados ao autovalor $\lambda = 0$, pois $T(v) = 0_V = \lambda\, v$.

(c) Seja $\{ v_1\,, \cdots, v_{n-1} \}$ uma base para o subespaço V_{λ_1}, desde que $dim(V_{\lambda_1}) = n - 1$. Sabemos que cada elemento v_j é um autovetor de T associado ao autovalor λ_1, pois

$$T(v_j) = \lambda_1\, v_j \qquad \text{para} \qquad j = 1\,, \cdots, (n-1)\,.$$

Assim, temos $(n-1)$ autovetores T linearmente independentes. Tomando v_n o autovetor de T associado ao autovalor λ_2, temos que o conjunto $\{ v_1\,, \cdots, v_{n-1}\,, v_n \}$ também é linearmente independente, pois o autovetor $v_n \notin V_{\lambda_1}$.

Desse modo, temos uma base ordenada $\gamma = \{ v_1\,, \cdots, v_{n-1}\,, v_n \}$ de autovetores de T para o espaço vetorial V. Assim, sabemos que $[T]^{\gamma}_{\gamma} = diag(\lambda_1, \lambda_1, \cdots \lambda_1, \lambda_2)$. Logo, T é um operador diagonalizável.

Questão 3.

(a) Com relação à base canônica $\beta = \{ (1,0), (0,1) \}$ de $\mathbb{R}^2$, temos que

$$[T]_\beta^\beta = \begin{bmatrix} 5 & -6 \\ 1 & 0 \end{bmatrix}.$$

Sabemos que o polinômio característico do operador T é o polinômio característico da matriz $A = [T]_\beta^\beta$ que é dado por:

$$p(\lambda) = \det(A - \lambda I) = -\lambda(5 - \lambda) + 6 = \lambda^2 - 5\lambda + 6 .$$

Portanto, $\lambda_1 = 2$ e $\lambda_2 = 3$ são os autovalores do operador T.

Vamos determinar os autovetores de T associados ao autovalor $\lambda_1 = 2$. Desse modo, temos que encontrar os elementos não nulos do núcleo do operador $T - \lambda_1 I$.

Assim, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} 3x - 6y = 0 \\ x - 2y = 0 \end{cases} \iff x - 2y = 0$$

que possui como solução $x = 2y$ para $y \in \mathbb{R}$ não nulo.

Desse modo, os autovetores do operator T associados ao autovalor $\lambda_1 = 2$ são do tipo $v = (2y, y)$ para $y \in \mathbb{R}$ não nulo. Assim, podemos escolher $v_1 = (2,1)$ o autovetor de T associado ao autovalor $\lambda_1 = 2$.

De modo análogo, para determinar os autovetores de T associados ao autovalor $\lambda_2 = 3$, temos que encontrar os elementos não nulos do núcleo do operador $T - \lambda_2 I$.

Assim, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} 2x - 6y = 0 \\ x - 3y = 0 \end{cases} \iff x - 3y = 0$$

que possui como solução $x = 3y$ para $y \in \mathbb{R}$ não nulo.

Desse modo, os autovetores do operator T associados ao autovalor $\lambda_2 = 3$ são do tipo $v = (3y, y)$ para $y \in \mathbb{R}$ não nulo. Assim, podemos escolher $v_2 = (3,1)$ o autovetor de T associado ao autovalor $\lambda_2 = 3$.

(b) Do item **(a)**, podemos observar facilmente que

$$V_{\lambda_1} = [(2,1)] \qquad \text{e} \qquad V_{\lambda_2} = [(3,1)]$$

são os autoespaços do operador T associados aos autovalores $\lambda_1 = 2$ e $\lambda_2 = 3$, respectivamente.

(c) Do item **(a)**, podemos concluir que T é um operador diagonalizável. Logo, a matriz $A = [T]_\beta^\beta$ é uma matriz diagonalizável.

Além disso, sabemos que os autovetores da matriz A são

$$X_1 = [v_1]_\beta = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \text{e} \qquad X_2 = [v_2]_\beta = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 2$ e $\lambda_2 = 3$, respectivamente.

Temos que a matriz A é similar a matriz diagonal $\Lambda = diag(2,3)$, onde a matriz invertível P que realiza a transformação de similaridade é dada por:

$$P = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \qquad \text{e} \qquad P^{-1} = \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix}.$$

Desse modo, temos que $A = P\,\Lambda\,P^{-1}$. Logo, sabemos que $A^8 = P\,\Lambda^8\,P^{-1}$ e que a matriz do operador T^8 com relação à base canônica β é dada por $[T^8]_\beta^\beta = A^8$.

Temos que a matriz A^8 é obtida da seguinte forma:

$$A^8 = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 \\ 0 & 3^8 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 256 & 0 \\ 0 & 6561 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 19171 & -37830 \\ 6305 & -12354 \end{bmatrix}$$

Finalmente, temos que

$$[T^8(u)]_\beta = [T^8]_\beta^\beta [u]_\beta \qquad \text{para} \qquad u = (x,y) \in I\!R^2 .$$

Portanto, a expressão explícita do operador linear T^8 é dada por:

$$T^8(x,y) = (\, 19171x - 37830y \,,\, 6305x - 12354y \,) \qquad \text{para} \qquad (x,y) \in I\!R^2 .$$

Questão 4.

Da condição (a), sabemos que

$$Ker(T) = \{ (x,y,z,t) \in \mathbb{R}^4 \ / \ x+y-z+t = 0 \quad \text{e} \quad z-t = 0 \} .$$

Podemos verificar facilmente que $\{ (-1,1,0,0) , (0,0,1,1) \}$ é uma base para $Ker(T)$.

Desse modo, podemos concluir que $\lambda_1 = 0$ é um autovalor de T com multiplicidade algébrica igual a 2 e multiplicidade geométrica também igual a 2, pois $V_\lambda = Ker(T)$ e $dim(Ker(T)) = 2$. Assim, podemos escolher $v_1 = (-1,1,0,0)$ e $v_2 = (0,0,1,1)$ os autovetores de T associados aos autovalores $\lambda_1 = 0$ e $\lambda_2 = 0$.

Da condição (b), sabemos que $v_3 = (0,0,1,0)$ é um autovetor do operador T associado ao autovalor $\lambda_3 = 2$. De fato, $T(v_3) = \lambda_3 v_3$, isto é, $T(0,0,1,0) = 2(0,0,1,0)$.

Podemos observar que o conjunto $\{ v_1 , v_2 , v_3 \}$ é linearmente independente em $\mathbb{R}^4$. Assim, estamos precisando de mais um elemento $v_4 \in \mathbb{R}^4$ para autovetor do operador T de modo que $\gamma = \{ v_1 , v_2 , v_3 , v_4 \}$ seja uma base de autovetores para $\mathbb{R}^4$.

Da condição (c), sabemos que o elemento $(0,1,0,0) \in Im(T)$. Assim, podemos escolher $v_4 = (0,1,0,0)$ como um autovetor do operador T associado ao autovalor $\lambda_4 = -3$.

Portanto, temos que

$$v_1 = (-1,1,0,0) \quad , \quad v_2 = (0,0,1,1) \quad , \quad v_3 = (0,0,1,0) \quad \text{e} \quad v_4 = (0,1,0,0)$$

são os autovetores do operador linear T associados aos autovalores $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 2$ e $\lambda_4 = -3$, respectivamente. Desse modo, $\gamma = \{ v_1 , v_2 , v_3 , v_4 \}$ é uma base de autovetores para $\mathbb{R}^4$ e sabemos que $[T]^\gamma_\gamma = diag(0,0,2,-3)$.

Finalmente, vamos determinar a expressão explícita do operador linear T diagonalizável que satisfaz as condições desejadas. Para isso, vamos representar um elemento genérico $(x,y,z,t) \in \mathbb{R}^4$ em relação à base de autovetores $\gamma = \{ v_1 , v_2 , v_3 , v_4 \}$, isto é,

$$(x,y,z,t) = a(-1,1,0,0) + b(0,0,1,1) + c(0,0,1,0) + d(0,1,0,0) .$$

Podemos verificar facilmente que $a = -x$, $b = t$, $c = z - t$ e $d = x + y$.

Portanto, obtemos

$$\begin{aligned} T(x,y,z,t) &= -xT(-1,1,0,0) \,+\, tT(0,0,1,1) \,+\, (z-t)T(0,0,1,0) \,+\, (x+y)T(0,1,0,0) \\ &= (0\,,\,-3x-3y\,,\,2z-2t\,,\,0) \end{aligned}$$

B.1.4 Segunda Chamada

Questão 1.

(a) A afirmação é **Falsa**.
Considere que exista uma transformação linear injetora T de $I\!R^4$ em $I\!R^3$. Desse modo, pelo **Teorema do núcleo e da imagem**, temos $dim(\,Im(T)\,) = 4$, pois $Ker(T) = \{\, 0_V \,\}$. O que não é possível, pois $Im(T)$ é um subespaço de $I\!R^3$ e $dim(\,I\!R^3\,) = 3$. Logo, não existe uma transformação linear injetora T de $I\!R^4$ em $I\!R^3$.

(b) A afirmação é **Verdadeira**.
Considere uma transformação linear T de $I\!R^4$ em $\mathcal{P}_2(I\!R)$ tal que $dim(\,Ker(T)\,) = 1$. Pelo **Teorema do núcleo e da imagem**, obtemos $dim(\,Im(T)\,) = 3$. Como $Im(T)$ é um subespaço de $\mathcal{P}_2(I\!R)$, tem–se que $Im(T) = \mathcal{P}_2(I\!R)$, pois $dim(\,\mathcal{P}_2(I\!R)\,) = 3$. Logo, existe uma transformação linear T sobrejetora de $I\!R^4$ em $\mathcal{P}_2(I\!R)$.

(c) A afirmação é **Falsa**.
Considere o conjunto linearmente dependente S em $I\!R^3$ dado por:

$$S = \{\, (1,1,0)\,,\, (-1,1,0)\,,\, (1,3,0)\,\}\,.$$

Entretanto, o subconjunto $\{\, (1,1,0)\,,\, (-1,1,0)\,\}$ é linearmente independente em $I\!R^3$.

(d) A afirmação é **Falsa**.
De fato, os espaços vetoriais $\mathcal{P}_4(I\!R)$ e $I\!M_2(I\!R)$ tem dimensões diferente, isto é, $dim(\,\mathcal{P}_4(I\!R)\,) = 5$ e $dim(\,I\!M_2(I\!R)\,) = 4$. Assim, não são espaços isomorfos.

(e) A afirmação é **Falsa**.
Considerando que $U \cap W = \{\, 0_V \,\}$, pelo Teorema da dimensão da soma de subespaços, temos que

$$dim(\,U + W\,) = dim(U) + dim(W) - dim(U \cap W) > \frac{n}{2} + \frac{n}{2} = n\,,$$

que é uma contradição, pois $U + W$ é um subespaço de V. Logo, podemos concluir que

$$U \cap W \neq \{\, 0_V \,\}\,.$$

Questão 2.

(a) Podemos verificar que a matriz nula 0_n pertence ao subconjunto U, pois 0_n é uma matriz simétrica e tem traço nulo.

Considerando as matrizes $A\,,\,B \in U$, temos

$$(A + B)^t \;=\; A^t + B^t \;=\; A + B \qquad \text{e} \qquad tr(A + B) \;=\; tr(A) + tr(B) \;=\; 0\,.$$

Portanto, a matriz $A + B \in U$.

Finalmente, considerando $A \in U$ e $\lambda \in \mathbb{R}$, temos

$$(\lambda A)^t \;=\; \lambda A^t \;=\; \lambda A \qquad \text{e} \qquad tr(\lambda A) \;=\; \lambda tr(A) \;=\; 0\,.$$

Portanto, a matriz $\lambda A \in U$.

Assim, mostramos que o subconjunto U é um subespaço de $\mathbb{M}_n(\mathbb{R})$.

(b) Como $A^t = A$ e $tr(A) = 0$, uma matriz genérica $A \in \mathbb{M}_3(\mathbb{R})$ com essas propriedades, pode ser escrita da seguinte forma:

$$A \;=\; \begin{bmatrix} a & x & y \\ x & b & z \\ y & z & (-a-b) \end{bmatrix}\,,$$

para $a,\, b,\, x,\, y,\, z \in \mathbb{R}$.

Por sua vez, podemos escrever

$$\begin{aligned} A \;&=\; \begin{bmatrix} a & x & y \\ x & b & z \\ y & z & (-a-b) \end{bmatrix} \\ &=\; a\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} + x\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + y\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} + z\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \end{aligned}$$

Desse modo, escrevemos $A = aA_1 + xA_2 + yA_3 + bA_4 + zA_5$, onde o conjunto

$$\{\, A_1\,,\, A_2\,,\, A_3\,,\, A_4\,,\, A_5 \,\}$$

é linearmente independente e gera o subespaço U de $\mathbb{M}_3(\mathbb{R})$. Assim, encontramos uma base para U.

Questão 3.

(a) Interpretando a matriz $[T]^{\beta}_{\gamma}$, obtemos

$$\begin{aligned} T(1,0) &= (1,0,1) - 2(0,1,0)) &&= (1,-2,1) \\ T(0,1) &= -(1,0,1) + (-1,0,1) + 3(0,1,0)) &&= (-2,3,0) \end{aligned}$$

(b) Utilizando o resultado do item **(a)**, obtemos

$$T(x,y) = T(x(1,0) + y(0,1)) = xT(1,0) + yT(0,1) = x(1,-2,1) + y(-2,3,0) \,.$$

Desse modo, podemos verificar facilmente que $\{ (1,-2,1) , (-2,3,0) \}$ é uma base para o subespaço $Im(T)$, pois é um conjunto linearmente independente em $\mathbb{R}^3$.

Assim, temos que

$$T(x,y) = (x - 2y \,,\, -2x + 3y \,,\, x) \qquad \text{para} \qquad (x,y) \in \mathbb{R}^2 \,.$$

(c) Para verificar se T é uma transformação injetora, vamos determinar o núcleo da transformação T. Desse modo, vamos encontrar os elementos $(x,y) \in \mathbb{R}^2$ tais que

$$T(x,y) = (0,0,0) \quad \Longrightarrow \quad (x - 2y \,,\, -2x + 3y \,,\, x) = (0,0,0) \,.$$

Assim, temos que determinar a solução do sistema linear homogêneo

$$\begin{cases} x - 2y = 0 \\ -2x + 3y = 0 \\ x = 0 \end{cases}$$

Obtemos, $x = y = z = 0$.

Portanto, $Ker(T) = \{ 0_V \}$. Logo, a transformação linear T é injetora.

Questão 4.

Seja T o operador linear sobre $\mathcal{P}_2(I\!R)$ definido por:

$$T(a + bx + cx^2) = (3a + 2b + c) + (b - c)x + 2cx^2 .$$

Inicialmente, vamos determinar a matriz $A = [T]^{\beta}_{\beta}$, onde $\beta = \{ 1 , x , x^2 \}$ é a base canônica de $\mathcal{P}_2(I\!R)$. Temos que

$$T(1) = 3 \quad , \quad T(x) = 2 + x \qquad \text{e} \qquad T(x^2) = 1 - x + 2x^2 .$$

Portanto, obtemos

$$A = [T]^{\beta}_{\beta} = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} .$$

Desse modo, o polinômio característico do operador linear T é dado por:

$$p(\lambda) = \det(A - \lambda I) = (3 - \lambda)(1 - \lambda)(2 - \lambda) .$$

Logo, $\lambda_1 = 3$, $\lambda_2 = 2$ e $\lambda_3 = 1$ são os autovalores do operador T. Como os autovalores são distintos, sabemos que T é um operador linear diagonalizável.

Os autovetores da matriz $A = [T]^{\beta}_{\beta}$ são

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad , \quad X_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \qquad \text{e} \qquad X_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

associados aos autovalores $\lambda_1 = 3$, $\lambda_2 = 2$ e $\lambda_3 = 1$, respectivamente.

Sabemos que

$$[p_1(x)]_{\beta} = X_1 \quad , \quad [p_2(x)]_{\beta} = X_2 \qquad \text{e} \qquad [p_3(x)]_{\beta} = X_3 ,$$

onde $p_1(x)$, $p_2(x)$ e $p_3(x)$ são os autovetores do operador T associados aos autovalores $\lambda_1 = 3$, $\lambda_2 = 2$ e $\lambda_3 = 1$, respectivamente. Logo, obtemos

$$p_1(x) = 1 \quad , \quad p_2(x) = 1 - x + x^2 \qquad \text{e} \qquad p_3(x) = -1 + x .$$

Finalmente, temos que

$$V_{\lambda_1} = [1] \quad , \quad V_{\lambda_2} = [1 - x + x^2] \qquad \text{e} \qquad V_{\lambda_3} = [-1 + x]$$

são os autoespaços do operador T.

B.1.5 Exame

Questão 1.

(a) Os elementos $(x, y, z, t) \in U$ podem ser escritos como:

$$(x, y, z, t) = y(-1, 1, 0, 0) + t(0, 0, 1, 1) \quad \text{para} \quad y, t \in I\!R \,.$$

Assim, o conjunto $\{ (-1, 1, 0, 0), (0, 0, 1, 1) \}$ é uma base para o subespaço U.

Os elementos $(x, y, z, t) \in W$ podem ser escritos como:

$$(x, y, z, t) = y(-1, 1, 0, 0) + z(-1, 0, 1, 0) + t(0, 0, 0, 1) \quad \text{para} \quad y, z, t \in I\!R \,.$$

Assim, o conjunto $\{ (-1, 1, 0, 0), (-1, 0, 1, 0), (0, 0, 0, 1) \}$ é uma base para W.

Podemos verificar que o conjunto $\{ (-1, 1, 0, 0), (0, 0, 1, 1), (-1, 0, 1, 0), (0, 0, 0, 1) \}$ é uma base para o subespaço $U + W$.

(b) Pelo Teorema da dimensão da soma de subespaços, temos que

$$dim(U + W) = dim(U) + dim(W) - dim(U \cap W) \,.$$

Como $dim(U) = 2$, $dim(W) = 3$ e $dim(U + W) = 4$, obtemos $dim(U \cap W) = 1$. Logo, o subespaço $U + W$ não é uma soma direta dos subespaços U e W.

(c) Como $dim(U \cap W) = 1$ e o elemento $(-1, 1, 0, 0)$ pertence tanto ao subespaço U quanto ao subespaço W, podemos concluir que $U \cap W = [(-1, 1, 0, 0)]$.

(d) Vamos determinar o operador linear T sobre $I\!R^4$ de modo que $Ker(T) = W$ e $Im(T) = U \cap W$. Pelo item **(a)**, temos que $\{ (-1, 1, 0, 0), (-1, 0, 1, 0), (0, 0, 0, 1) \}$ é uma base para o subespaço $Ker(T)$. Completando a base do núcleo de T, obtemos que $\{ (-1, 1, 0, 0), (-1, 0, 1, 0), (0, 0, 0, 1), (0, 0, 1, 0) \}$ é uma base para $I\!R^4$. Assim, o operador linear T é determinado por:

$$\begin{aligned}
T(-1, 1, 0, 0) &= (0, 0, 0, 0) \\
T(-1, 0, 1, 0) &= (0, 0, 0, 0) \\
T(0, 0, 0, 1) &= (0, 0, 0, 0) \\
T(0, 0, 1, 0) &= (-1, 1, 0, 0)
\end{aligned}$$

o que completa a resolução da questão.

Questão 2.

Vamos verificar se o subconjunto U é um subespaço de $\mathcal{P}_3(\mathbb{R})$. Podemos verificar facilmente que o polinômio identicamente nulo, isto é, $p(x) = 0$ para todo x, que é o elemento neutro da operação de adição, pertence ao subconjunto U.

Tomando os elemento $p(x), q(x) \in \mathcal{P}_3(\mathbb{R})$, isto é,

$$p(1) + p(-1) = 0 \qquad \text{e} \qquad q(1) + q(-1) = 0 \,,$$

temos que

$$(p + q)(1) + (p + q)(-1) = (\, p(1) + p(-1)\,) + (\, q(1) + q(-1)\,) = 0 \,.$$

Portanto, o elemento $(\, p(x) + q(x)\,) \in \mathcal{P}_3(\mathbb{R})$.

Agora, tomando o elemento $p(x) \in \mathcal{P}_3(\mathbb{R})$ e $\lambda \in \mathbb{R}$, temos que

$$(\lambda p)(1) + (\lambda p)(-1) = \lambda p(1) + \lambda p(-1) = \lambda(\, p(1) + p(-1)\,) = 0 \,.$$

Portanto, o elemento $\lambda p(x) \in \mathcal{P}_3(\mathbb{R})$.

Assim, mostramos que o subconjunto U é um subespaço vetorial de $\mathcal{P}_3(\mathbb{R})$.

Finalmente, vamos determinar uma base para o subespaço U. Para isso, consideramos um elemento $p(x) = a + bx + cx^2 + dx^3 \in \mathcal{P}_3(\mathbb{R})$ e, impondo a condição $p(1) + p(-1) = 0$, obtemos

$$(\, a + b + c + d\,) + (\, a - b + c - d\,) = 0 \qquad \Longrightarrow \qquad a + c = 0 \,.$$

Assim, temos que $a = -c$ com $b, c, d \in \mathbb{R}$.

Desse modo, os elementos $p(x) \in U$ podem ser representados da seguinte forma:

$$p(x) = c(-1 + x^2) + bx + dx^3 \qquad \text{para} \qquad b, c, d \in \mathbb{R} \,.$$

Portanto, o conjunto $\{\, -1 + x^2 \,,\, x \,,\, x^3 \,\}$ é uma base para o subespaço U.

Questão 3.

(a) Utilizando a matriz $[T]^{\alpha}_{\gamma}$, obtemos

$$
\begin{aligned}
T(0,1) &= -1(-1,0) + 0(0,-1) = (1,0) \\
T(1,0) &= -1(-1,0) - 1(0,-1) = (1,1)
\end{aligned}
$$

Assim, temos que $T(0,1) = (1,0)$ e $T(1,0) = (1,1)$.

(b) Representando o elemento $(0,1)$ da base α em relação à base γ, temos

$$(0,1) = a(-1,0) + b(0,-1) \,.$$

Logo, $a = 0$ e $b = -1$.

Representando o elemento $(1,0)$ da base α em relação à base γ, temos

$$(1,0) = c(-1,0) + d(0,-1) \,.$$

Logo, $c = -1$ e $d = 0$. Portanto, obtemos

$$[I]^{\alpha}_{\gamma} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix},$$

que é a matriz de mudança da base ordenada α para a base ordenada γ.

(c) Utilizando o resultado do item **(b)**, obtemos

$$
\begin{aligned}
T(x,y) &= T(x(1,0) + y(0,1)) \\
&= xT(1,0) + yT(0,1) \\
&= x(1,1) + y(1,0) \\
&= (x+y\,,\,x)
\end{aligned}
$$

Portanto, $T(x,y) = (x+y\,,\,x)$ para todo $(x,y) \in I\!R^2$.

(d) Utilizando o resultado do item **(c)**, obtemos

$$[T]^{\beta}_{\beta} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

onde $\beta = \{ (1,0), (0,1) \}$ é a base canônica de $I\!R^2$.

Sabemos que a matriz $[T^2]^{\beta}_{\beta}$ é dada por:

$$A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Assim, o polinômio característico do operador linear T^2 é dado por:

$$p(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) - 1 = \lambda^2 - 3\lambda + 1 .$$

Portanto, os autovalores do operador T^2 são

$$\lambda_1 = \frac{3 + \sqrt{5}}{2} \quad \text{e} \quad \lambda_2 = \frac{3 - \sqrt{5}}{2} .$$

Como os autovalores de T são distintos, sabemos que T é um operador diagonalizável.

Questão 4.

(a) Considerando que a matriz B é similar à matriz A, existe uma matriz P invertível tal que $B = P^{-1}AP$. Tomando o polinômio característico da matriz B, obtemos

$$\begin{aligned} p(\lambda) &= det(B - \lambda I_n) \\ &= det(P^{-1}AP - \lambda P^{-1} P) \\ &= det(P^{-1}(A - \lambda I_n)P) \\ &= det(P^{-1})\, det(A - \lambda I_n)\, det(P) \\ &= det(A - \lambda I_n)\,, \end{aligned}$$

onde I_n é a matriz identidade de ordem n. Assim, mostramos que as matrizes A e B possuem o mesmo polinômio característico. Como os autovalores são as raízes do polinômio característico, temos que as matrizes A e B possuem os mesmos autovalores.

(b) Seja X um autovetor da matriz B associado ao autovalor λ, isto é,

$$BX = \lambda X \iff P^{-1}APX = \lambda X \iff A(PX) = \lambda(PX)\,.$$

Portanto, obtemos que PX é um autovetor da matriz A associado ao autovalor λ.

(c) Considerando que A é uma matriz diagonalizável, existe uma matriz Q invertível tal que $A = Q\Lambda Q^{-1}$, onde Λ é uma matriz diagonal. Como B é similar a matriz A, obtemos

$$B = P^{-1}AP = P^{-1}(Q\Lambda Q^{-1})P = (P^{-1}Q)\Lambda(Q^{-1}P) = (Q^{-1}P)^{-1}\,\Lambda\,(Q^{-1}P)\,.$$

Assim, mostramos que B é similar à matriz diagonal Λ. Logo, B é diagonalizável.

B.2 Primeiro Semestre de 2006

B.2.1 Primeira Prova

Questão 1.

Considere o subconjunto S de $\mathcal{C}([-1,1])$ das funções monótonas crescentes, isto é,

$$S = \{ f \in \mathcal{C}([-1,1]) \ / \ f(x) > f(y) \quad \text{para} \quad x > y \}.$$

Considerando $f, g \in S$, isto é, $f(x) > f(y)$ e $g(x) > g(y)$ para $x > y$, temos que $f(x)+g(x) > f(y)+g(y)$. Assim, $f + g \in S$. Logo, S é fechado com relação à operação de adição.

Considerando $f \in S$, isto é, $f(x) > f(y)$ para $x > y$, e $\lambda \in I\!R$ negativo. Desse modo, temos que $\lambda f(x) < \lambda g(x)$ para $x > y$. Logo, temos que $(\lambda f) \notin S$. Portanto, S não é fechado com relação à operação de multiplicação por escalar.

Questão 2.

($\Longrightarrow$) Considerando que o conjunto $\{ v_1, v_2, v_3, v_4, u \}$ é linearmente dependente em V, temos que existem escalares $c_1, c_2, c_3, c_4, \alpha$, não todos nulos, tais que

$$c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 + \alpha u = 0_V .$$

Agora temos duas possibilidades. Primeira, se o escalar $\alpha \neq 0$, obtemos

$$u = -\frac{c_1}{\alpha}v_1 - \frac{c_2}{\alpha}v_2 - \frac{c_3}{\alpha}v_3 - \frac{c_4}{\alpha}v_4 \qquad \Longrightarrow \qquad u \in [v_1, v_2, v_3, v_4] .$$

Segunda, se $\alpha = 0$, temos que

$$c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0_V .$$

Como o conjunto $\{ v_1, v_2, v_3, v_4 \}$ é linearmente independente em V, isto implicaria que os escalares c_1, c_2, c_3, c_4 devem ser todos nulos. Entretanto, isso contraria a hipótese do conjunto $\{ v_1, v_2, v_3, v_4, u \}$ ser linearmente dependente em V. Logo, mostramos que $u \in [v_1, v_2, v_3, v_4]$.

($\Longleftarrow$) Considerando que $u \in [v_1, v_2, v_3, v_4]$, existem escalares c_1, c_2, c_3, c_4, não todos nulos, tais que

$$u = c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 \qquad \Longrightarrow \qquad c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 - u = 0_V .$$

Portanto, mostramos que $\{ v_1, v_2, v_3, v_4, u \}$ é linearmente dependente em V.

Questão 3.

Todo elemento $(x, y, z) \in U$ satisfaz a equação $2x - 4y + 6z = 0$. Desse modo, temos que $x = 2y - 3z$ para $y, z \in \mathbb{R}$. Portanto, todo elemento $(x, y, z) \in U$ é escrito da seguinte forma:

$$(x, y, z) = y(2, 1, 0) + z(-3, 0, 1) \quad ; \quad y, z \in \mathbb{R}.$$

Assim, o conjunto $\{ (2, 1, 0), (-3, 0, 1) \}$ é um sistema de geradores para o subespaço U.

Sabemos que o subespaço $U + W = \{ v \in \mathbb{R}^3 \ / \ v = u + w \quad ; \quad u \in U \text{ e } w \in W \}$. Assim, temos que

$$v = a(2, 1, 0) + b(-3, 0, 1) + c(1, 0, 1) + d(1, 1, 3) \quad ; \quad a, b, c, d \in \mathbb{R}.$$

Assim, o conjunto $\{ (2, 1, 0), (-3, 0, 1), (1, 0, 1), (1, 1, 3) \}$ é um sistema de geradores para o subespaço $U + W$.

Vamos determinar um sistema de geradores para o subespaço $U \cap W$. Sabemos que, se $v \in U \cap W$, então $v \in U$ e $v \in W$. Assim, temos que

$$a(2, 1, 0) + b(-3, 0, 1) = c(1, 0, 1) + d(1, 1, 3) \quad \text{para} \quad a, b, c, d \in \mathbb{R}.$$

Desse modo, obtemos o seguinte sistema linear

$$\begin{cases} 2a - 3b = c + d \\ a = d \\ b = c + 3d \end{cases}$$

cuja solução é dada por $a = d$, $b = d$ e $c = -2d$ para $d \in \mathbb{R}$.

Portanto, se $v \in U \cap W$, então ele pode ser escrito como $v = d(-1, 1, 1)$ para $d \in \mathbb{R}$. Logo, $\{ (-1, 1, 1) \}$ é um sistema de geradores para o subespaço $U \cap W$. Como o subespaço $U \cap W \neq \{ 0_{\mathbb{R}^3} \}$, o subespaço $U + W$ não é uma soma direta dos subespaços U e W.

Questão 4.

Consideramos um elemento genérico $p(x) = a + bx + cx^2 + dx^3 \in \mathcal{P}_3(\mathbb{R})$ e vamos impor as condições para que esse elemento pertença ao subespaço S, isto é,

$$\begin{array}{rcrcrcrcrcl} p(-1) + p'(-1) & = & a & & & - & c & + & 2d & = & 0 \\ p(1) & = & a & + & b & + & c & + & d & = & 0 \end{array} .$$

Escalonando o sistema linear homogêneo, obtemos

$$\begin{cases} a + b + c + d = 0 \\ \quad - b - 2c + d = 0 \end{cases} .$$

Assim, obtemos um sistema linear homogêneo com dois graus de liberdade. Desse modo, podemos concluir que o subespaço S tem dimensão dois. Logo, temos uma relação entre os coeficientes dos elementos $p(x) \in S$. Podemos verificar facilmente que

$$b = -2c + d \qquad \text{e} \qquad a = c - 2d$$

para $c, d \in \mathbb{R}$. Substituindo a e b no polinômio $p(x)$, obtemos que todo elemento do subespaço S é escrito como:

$$p(x) = c(1 - 2x + x^2) + d(-2 + x + x^3) \qquad ; \qquad c, d \in \mathbb{R} .$$

Portanto, mostramos que o subespaço S é gerado pelos elementos do conjunto

$$\Gamma = \{ 1 - 2x + x^2 , -2 + x + x^3 \} ,$$

que é linearmente independente em $\mathcal{P}_3(\mathbb{R})$, pois tomando a combinação linear nula

$$a(1 - 2x + x^2) + b(-2 + x + x^3) = 0$$

obtemos o seguinte sistema linear homogêneo

$$\begin{cases} a - 2b = 0 \\ -2a + b = 0 \\ a = 0 \\ b = 0 \end{cases}$$

cuja solução é $a = b = 0$. Logo, o conjunto Γ é uma base para o subespaço S.

Note que os elementos da base Γ satisfazem as condições para que um elemento do espaço vetorial $\mathcal{P}_3(\mathbb{R})$ pertença ao subespaço S.

Questão 5.

Inicialmente, vamos encontrar uma base para o subespaço W. Para isso, construímos uma matriz cujas linhas são os elementos do sistema de geradores do subespaço W e procedemos com o escalonamento

$$A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & -1 & 1 & 3 \\ -1 & 1 & 0 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, podemos escolher

$$\Gamma = \{ (1,0,1,2),\ (2,-1,1,3) \} \qquad \text{ou} \qquad \Gamma' = \{ (1,0,1,2),\ (0,1,1,1) \}$$

para uma base do subespaço W.

Finalmente, vamos completar uma base de W para obter uma base de $\mathbb{R}^4$.

Desse modo, podemos escolher

$$\beta = \{ (1,0,1,2),\ (2,-1,1,3),\ (0,0,1,0),\ (0,0,0,1) \} \qquad \text{ou}$$

$$\beta' = \{ (1,0,1,2),\ (0,1,1,1),\ (0,0,1,0),\ (0,0,0,1) \}$$

para uma base do espaço vetorial $\mathbb{R}^4$, pois construindo a matriz M de ordem 4 associada ao conjunto β'

$$M = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

vemos que está forma escalonada e $posto(M) = 4$.

B.2.2 Segunda Prova

Questão 1.

Da matriz de mudança de base $[I]^{\alpha}_{\gamma}$, sabemos que

$$\begin{aligned} p_1(x) &= q_1(x) + q_2(x) \\ p_2(x) &= 2q_1(x) + q_2(x) \\ p_3(x) &= q_3(x) \end{aligned}$$

Da última equação, temos que $q_3(x) = 1 - x^2$. Das duas primeiras equações, obtemos o seguinte sistema linear

$$\left\{ \begin{array}{rcl} q_1(x) + q_2(x) &=& p_1(x) \\ \\ 2q_1(x) + q_2(x) &=& p_2(x) \end{array} \right. \iff \left\{ \begin{array}{rcrcl} q_1(x) &+& q_2(x) &=& p_1(x) \\ \\ && q_2(x) &=& 2p_1(x) - p_2(x) \end{array} \right.$$

que possui a solução $q_1(x) = p_2(x) - p_1(x)$ e $q_2(x) = 2p_1(x) - p_2(x)$. Assim, obtemos

$$q_1(x) = 2x \,, \quad q_2(x) = 1 - 3x \quad \text{e} \quad q_3(x) = 1 - x^2$$

que são os elementos da base ordenada γ.

Vamos encontrar o vetor de coordenadas do elemento $p(x) = 3 - x + 2x^2$ com relação à base ordenada α. Para isso, basta fazer

$$\begin{aligned} p(x) &= a\,p_1(x) + b\,p_2(x) + c\,p_3(x) \\ 3 - x + 2x^2 &= a(1 - x) + b(1 + x) + c(1 - x^2) \end{aligned}$$

obtendo o seguinte sistema linear

$$\left\{ \begin{array}{rcrcrcr} a &+& b &+& c &=& 3 \\ -a &+& b && &=& -1 \\ && &-& c &=& 2 \end{array} \right.$$

que possui como solução $a = 3$, $b = 2$ e $c = -2$. Assim, temos

$$[p(x)]_{\alpha} = \begin{bmatrix} 3 \\ 2 \\ -2 \end{bmatrix}.$$

Questão 2.

Da condição **(a)** temos que

$$T(1 + x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} .$$

Da condição **(b)**, isto é, $q(x) = 1 \notin Ker(T)$, implica que o elemento $r(x) = x^2$ não pode pertencer ao $Ker(T)$, pois podemos escrever $q(x) = p(x) - r(x)$. Claramente, se o elemento $r(x) \in Ker(T)$, então $q(x) \in Ker(T)$, o que contradiz a hipótese.

Assim, podemos considerar a seguinte transformação linear $T : \mathcal{P}_2(I\!\!R) \longrightarrow I\!\!M_2(I\!\!R)$, com $\{ 1 + x^2 \}$ a base para $Ker(T)$, dada por:

$$T(1 + x^2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} , \qquad T(1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \text{e} \qquad T(x) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

onde estamos escolhendo $\gamma = \{ 1 + x^2, 1, x \}$ uma base ordenada para $\mathcal{P}_2(I\!\!R)$, que foi obtida completando a base do $Ker(T)$.

Vamos tomar um elemento genérico $p(x) = a + bx + cx^2 \in \mathcal{P}_2(I\!\!R)$ e representa–lo com relação à base ordenada γ de $\mathcal{P}_2(I\!\!R)$

$$\begin{aligned} p(x) &= d_1(1 + x^2) + d_2 + d_3 x \\ &= (d_1 + d_2) + d_3 x + d_1 x^2 \end{aligned}$$

obtendo o seguinte sistema linear

$$\begin{cases} d_1 + d_2 = a \\ \phantom{d_1 + {}} d_1 = c \\ \phantom{d_1 + {}} d_3 = b \end{cases}$$

que possui somente a solução $d_1 = c$, $d_2 = a - c$ e $d_3 = b$. Desse modo, temos que

$$p(x) = c(1 + x^2) + (a - c) + bx .$$

Agora, fazendo $T(p(x)) = T(a + bx + cx^2)$, obtemos

$$T(a + bx + cx^2) = c\,T(1 + x^2) + (a - c)\,T(1) + b\,T(x) = \begin{bmatrix} 2b & a - c \\ a - c & b \end{bmatrix} .$$

Assim, encontramos uma transformação T com as propriedades pedidas.

Questão 3.

Tomando uma combinação linear nula

$$\sum_{i=1}^{m} c_i \, T(v_i) \;=\; 0_W \,,$$

e como T é uma transformação linear, podemos escrever

$$T\left(\sum_{i=1}^{m} c_i \, v_i \right) \;=\; 0_W \,.$$

Considerando a hipótese que T é injetora, isto é, $Ker(T) = \{ 0_V \}$, temos que

$$\sum_{i=1}^{m} c_i \, v_i \;=\; 0_V \,.$$

Como $\{ v_1 \,, \cdots \,, v_m \}$ é linearmente independente em V, implica que

$$c_1 = c_2 = \cdots = c_m = 0 \,.$$

Portanto, $\{ T(v_1) \,, \cdots \,, T(v_m) \}$ é linearmente independente em W.

Questão 4.

(a) A afirmação é **Falsa**.
Considere que exista uma transformação linear injetora T de $I\!R^4$ em $I\!R^3$, isto é, $Ker(T) = \{ 0_V \}$. Pelo **Teorema do núcleo e da imagem**, temos $dim(\, Im(T)\,) = 4$. O que não é possível, pois $Im(T)$ é um subespaço de $I\!R^3$ e $dim(\, I\!R^3\,) = 3$. Logo, não existe uma transformação linear injetora T de $I\!R^4$ em $I\!R^3$.

(b) A afirmação é **Verdadeira**.
Considere uma transformação linear T de $I\!R^4$ em $\mathcal{P}_2(I\!R)$ tal que $dim(\, Ker(T)\,) = 1$. Pelo **Teorema do núcleo e da imagem**, temos que $dim(\, Im(T)\,) = 3$. Como $Im(T)$ é um subespaço de $\mathcal{P}_2(I\!R)$, tem–se que $Im(T) = \mathcal{P}_2(I\!R)$, pois $dim(\, \mathcal{P}_2(I\!R)\,) = 3$. Logo, existe uma transformação linear T sobrejetora de $I\!R^4$ em $\mathcal{P}_2(I\!R)$.

(c) A afirmação é **Falsa**.
Considere uma transformação linear injetora T de $I\!R^2$ em $\mathcal{P}_2(I\!R)$. Pelo **Teorema do núcleo e da imagem**, temos que $dim(\, Im(T)\,) = 2$, pois $Ker(T) = \{ 0_V \}$. Logo, tem–se que $Im(T) \neq \mathcal{P}_2(I\!R)$, pois $dim(\, \mathcal{P}_2(I\!R)\,) = 3$. Portanto, não existe uma transformação bijetora T de $I\!R^2$ em $\mathcal{P}_2(I\!R)$.

Questão 5.

Temos que $\gamma = \{ (1,-1), (0,1) \}$ é uma base para o $\mathbb{R}^2$. Vamos mostrar que γ é linearmente independente. Considere a combinação linear nula

$$a(1,-1) \, + \, b(0,1) \, = \, (0,0) \quad \Longleftrightarrow \quad \left\{ \begin{array}{rcrcl} a & & & = & 0 \\ -a & + & b & = & 0 \end{array} \right. .$$

Assim, obtemos $a = b = 0$. Logo, γ é linearmente independente em $\mathbb{R}^2$.

Vamos tomar um elemento genérico $(a,b) \in \mathbb{R}^2$ e representa–lo com relação à base ordenada γ

$$(a,b) \, = \, c(1,-1) \, + \, d(0,1) \, = \, (c, \, -c + d) \, .$$

Assim, obtemos o seguinte sistema linear

$$\left\{ \begin{array}{rcrcl} c & & & = & a \\ -c & + & d & = & b \end{array} \right.$$

que possui como solução $c = a$ e $d = a + b$. Desse modo, temos que

$$(a,b) \, = \, a(1,-1) \, + \, (a+b)(0,1) \, .$$

Agora, fazendo

$$\begin{aligned} T(a,b) \, &= \, aT(1,-1) \, + \, (a+b)T(0,1) \\ &= \, a(2+x) \, + \, (a+b)(x-1) \\ &= \, (a \, - \, b) \, + \, (2a \, + \, b)x \end{aligned}$$

obtemos a transformação linear T.

Para mostrar que T é um isomorfismo, basta mostrar que $Ker(T) = \{ 0_{\mathbb{R}^2} \}$. Assim, considerando um elemento $(a,b) \in Ker(T)$, temos que

$$T(a,b) \, = \, (a \, - \, b) \, + \, (2a \, + \, b)x \, = \, 0_{\mathcal{P}_1(\mathbb{R})} \qquad \text{para todo} \qquad x \in \mathbb{R}$$

Assim, obtemos o seguinte sistema linear homogêneo

$$\left\{ \begin{array}{rcrcl} a & - & b & = & 0 \\ 2a & + & b & = & 0 \end{array} \right.$$

que possui somente a solução trivial $a = b = 0$. Logo, T é um isomorfismo.

Vamos encontrar o isomorfismo inverso. Dado um elemento $p(x) = a + bx \in \mathcal{P}_1(\mathbb{R})$, supomos que $T^{-1}(a + bx) = (c, d)$. Assim, temos que $T(c, d) = a + bx$, isto é,

$$(c - d) + (2c + d)x = a + bx \,,$$

obtendo o seguinte sistema linear

$$\begin{cases} c - d = a \\ 2c + d = b \end{cases}$$

que possui como solução $c = \dfrac{a+b}{3}$ e $d = \dfrac{b-2a}{3}$.

Portanto, temos que o isomorfismo inverso é dado por:

$$T^{-1}(a + bx) = \left(\frac{a+b}{3}, \frac{b-2a}{3} \right) .$$

B.2.3 Terceira Prova

Questão 1.

Chamando $[p(x)]_\beta = \begin{bmatrix} a \\ b \end{bmatrix}$.

Sabemos que $[T(p(x))]_\gamma = [T]^\beta_\gamma [p(x)]_\beta$. Assim, obtemos o seguinte sistema linear

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \iff \begin{cases} a + b = 1 \\ 2a + b = 3 \\ a + 2b = 0 \end{cases} \iff \begin{cases} a + b = 1 \\ \quad - b = 1 \end{cases}$$

que tem uma única solução $a = 2$ e $b = -1$. Logo, $[p(x)]_\beta = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

Chamando $\beta = \{ q_1(x), q_2 \}$, onde

$$q_1(x) = x - x^2 + x^3 \qquad \text{e} \qquad q_2(x) = 1 + x + x^2 .$$

Conhecemos a matriz $[T]^\beta_\gamma$, onde $\gamma = \{ x - 1, p_1(x), p_2(x) \}$. Assim, temos que

$$\begin{aligned} T(q_1(x)) &= (x - 1) + 2p_1(x) + p_2(x) \\ T(q_2(x)) &= (x - 1) + p_1(x) + 2p_2(x) \end{aligned} . \tag{B.1}$$

Tomando $T(p(x)) = p'(x) + (x + 1)p(0)$, vamos calcular

$$\begin{aligned} T(q_1(x)) &= 1 - 2x + 3x^2 \\ T(q_2(x)) &= 1 + 2x + (x + 1) = 3x + 2 \end{aligned} . \tag{B.2}$$

Substituindo (B.2) em (B.1), obtemos um sistema linear nas incógnitas $p_1(x)$ e $p_2(x)$

$$\begin{cases} 2p_1(x) + p_2(x) = 3x^2 - 3x + 2 \\ p_1(x) + 2p_2(x) = 2x + 3 \end{cases}$$

Fazendo a primeira equação menos a segunda equação, obtemos

$$p_1(x) - p_2(x) = 3x^2 - 5x - 1 ,$$

o que completa a resolução da questão.

Questão 2.

(a) Note que a aplicação $\langle \cdot , \cdot \rangle$ definida por:

$$\langle f , g \rangle = \int_0^1 f'(x)g(x)dx \qquad ; \qquad \forall \ f, g \in \mathcal{C}_0^1([0,1])$$

não satisfaz a propriedade de **simetria**. De fato,

$$\langle f , g \rangle = \int_0^1 f'(x)g(x)dx \neq \int_0^1 g'(x)f(x)dx = \langle g , f \rangle .$$

Por exemplo, tomando as funções $f(x) = 1 - x$ e $g(x) = 1 - x^2$, temos que

$$\langle f , g \rangle = \int_0^1 (x^2 - 1)dx = -\frac{2}{3} \qquad \text{e} \qquad \langle g , f \rangle = \int_0^1 (2x^2 - 2x)dx = -\frac{1}{3} .$$

Portanto, $\langle f , g \rangle \neq \langle g , f \rangle$.

Além disso, podemos verificar facilmente que a aplicação $\langle \cdot , \cdot \rangle$ não satisfaz a propriedade de **positividade**. De fato,

$$\begin{aligned} \langle f , f \rangle &= \int_0^1 f'(x)f(x)dx = \frac{1}{2} \int_0^1 (f^2(x))'dx \\ &= \frac{1}{2} (f^2(1) - f^2(0)) = -\frac{1}{2} f^2(0) \leq 0 , \end{aligned}$$

onde $f(1) = 0$.

Logo, a aplicação $\langle \cdot , \cdot \rangle$ não define um produto interno no espaço vetorial $\mathcal{C}_0^1([0,1])$.

(b) Podemos verificar facilmente que a aplicação $\langle \cdot , \cdot \rangle$ definida por:

$$\langle f , g \rangle = \int_0^1 f'(x)g'(x)dx \quad ; \quad \forall \ f, g \in \mathcal{C}_0^1([0,1])$$

satisfaz as propriedades de **simetria**, **homogeneidade** e **distributividade**. De fato,

$$\langle f , g \rangle = \int_0^1 f'(x)g'(x)dx = \int_0^1 g'(x)f'(x)dx = \langle g , f \rangle \quad ; \quad \forall \ f, g \in \mathcal{C}_0^1([0,1]) .$$

$$\langle \lambda f , g \rangle = \int_0^1 (\lambda f)'(x)g'(x)dx = \int_0^1 \lambda f'(x)g'(x)dx = \lambda \int_0^1 f'(x)g'(x)dx = \lambda \langle f , g \rangle$$

para todas funções $f, g \in \mathcal{C}_0^1([0,1])$ e $\lambda \in I\!R$.

$$\begin{aligned} \langle f + g , h \rangle &= \int_0^1 (f+g)'(x)h'(x)dx = \int_0^1 f'(x)h'(x)dx + \int_0^1 g'(x)h'(x)dx \\ &= \langle f , h \rangle + \langle g , h \rangle \quad ; \quad \forall \ f, g \in \mathcal{C}_0^1([0,1]) . \end{aligned}$$

Vamos mostrar que a aplicação $\langle \cdot , \cdot \rangle$ satisfaz a propriedade de **positividade**. De fato,

$$\langle f , f \rangle = \int_0^1 (\, f'(x) \,)^2 dx \geq 0 ,$$

pois o integrando é uma função contínua positiva.

Agora, supomos que

$$\langle f , f \rangle = \int_0^1 (\, f'(x) \,)^2 dx = 0 .$$

Como f' é uma função contínua, temos que $f'(x) = 0$ para todo $x \in [0,1]$. Logo, f é uma função constante em $[0,1]$, entretanto, $f(1) = 0$. Assim, a única função constante no espaço $\mathcal{C}_0^1([0,1])$ é a função identicamente nula ($f \equiv 0$), isto é, $f(x) = 0$ para todo $x \in [0,1]$.

Portanto, a aplicação $\langle \cdot , \cdot \rangle$ define um produto interno no espaço vetorial $\mathcal{C}_0^1([0,1])$.

Questão 3.

(a) Considerando que V é um espaço vetorial real, temos que

$$\| u + v \|_2^2 = \langle u + v , u + v \rangle = \langle u , u \rangle + 2\langle u , v \rangle + \langle v , v \rangle .$$

Utilizando o fato que θ é o ângulo entre os elementos u e v, não nulos, temos que

$$\cos(\theta) = \frac{\langle u , v \rangle}{\| u \|_2 \| v \|_2} \quad \Longrightarrow \quad \langle u , v \rangle = \| u \|_2 \| v \|_2 \cos(\theta) .$$

Portanto, obtemos a relação

$$\| u + v \|_2^2 = \| u \|_2^2 + \| v \|_2^2 + 2 \| u \|_2 \| v \|_2 \cos(\theta)$$

que é denominada **Lei do Paralelogramo**.

(b) Tomando a combinação linear nula dos elementos do conjunto β

$$c_1 q_1 + \cdots + c_i q_i + \cdots + c_n q_n = 0_V ,$$

e fazendo o produto interno de ambos os membros com um elemento $q_j \in \beta$ temos que

$$c_1 \langle q_1 , q_j \rangle + \cdots + c_i \langle q_i , q_j \rangle + \cdots + c_n \langle q_n , q_j \rangle = 0 .$$

Usando o fato que β é um conjunto ortonormal, isto é,

$$\begin{cases} \langle q_i , q_j \rangle = 0 & \text{para} \quad i \neq j \\ \langle q_i , q_j \rangle = 1 & \text{para} \quad i = j \end{cases}$$

obtemos $c_j = 0$ para $j = 1, \cdots , n$. Portanto, mostramos que β é um conjunto linearmente independente em V.

Questão 4.

Chamando $p(x) = 1 + x$, temos que o subespaço $S = [p(x)] \subset \mathcal{P}_2(I\!R)$.

O subespaço $S^{\perp}$ é definido por:

$$S^{\perp} = \{ \, q \in \mathcal{P}_2(I\!R) \ / \ \langle \, r \, , \, q \, \rangle = 0 \quad ; \quad \forall \, r \in S \, \} \, .$$

Tomando um elemento genérico $q(x) = a + bx + cx^2 \in S^{\perp}$, sabemos que $\langle \, p \, , \, q \, \rangle = 0$. Assim, temos que

$$\begin{aligned} \langle \, p \, , \, q \, \rangle &= \int_{-1}^{1} x^2(1 + x)(a + bx + cx^2)dx \\ &= \int_{-1}^{1} (x^2 + x^3)(a + bx + cx^2)dx \\ &= \int_{-1}^{1} (\, ax^2 + bx^3 + cx^4 + ax^3 + bx^4 + cx^5 \,)dx = 0 \\ &= \int_{-1}^{1} (\, ax^2 + cx^4 + bx^4 \,)dx = 0 \end{aligned}$$

Calculando a integral, resulta a seguinte equação

$$\frac{2}{3}a + \frac{2}{5}c + \frac{2}{5}b = 0$$

Resolvendo a equação acima para a incógnita c, temos que

$$c = -\frac{5}{3}a - b \, .$$

Portanto, todo elemento $q(x) \in S^{\perp}$ é escrito como:

$$\begin{aligned} q(x) &= a + bx + \left(-\frac{5}{3}a - b \right) x^2 \\ &= \left(1 - \frac{5}{3}x^2 \right) a + (\, x - x^2 \,) \, b \qquad \text{para} \qquad a, b \in I\!R \, . \end{aligned}$$

Desse modo, uma base para o subespaço $S^{\perp}$ é formada pelos elementos

$$q_1(x) = 1 - \frac{5}{3}x^2 \qquad \text{e} \qquad q_2(x) = x - x^2 \, ,$$

completando a resolução da questão.

Questão 5.

A melhor aproximação do elemento $q(x) = 1 - x^2$ no subespaço $\mathcal{P}_1(\mathbb{R}) \subset \mathcal{P}_2(\mathbb{R})$ é dada pela projeção ortogonal do elemento $q(x)$ sobre o subespaço $\mathcal{P}_1(\mathbb{R})$.

Inicialmente, vamos obter uma base ortogonal $\beta^\star = \{ q_1(x), q_2(x) \}$ para o subespaço $\mathcal{P}_1(\mathbb{R})$ a partir da base canônica $\beta = \{ p_1(x) = 1, p_2(x) = x \}$, através do **Processo de Ortogonalização de Gram-Schmidt**.

Desse modo, escolhemos $q_1(x) = p_1(x) = 1$. Agora, vamos construir o elemento $q_2(x)$ da seguinte forma:

$$q_2(x) = p_2(x) - \alpha_{12}\, q_1(x)$$

ortogonal ao subespaço gerado pelo elemento $q_1(x)$. Assim, temos que

$$\alpha_{12} = \frac{\langle p_2 \,,\, q_1 \rangle}{\langle p_2 \,,\, q_1 \rangle} = \frac{1}{2}\,.$$

Logo, o elemento $q_2(x) = x - \dfrac{1}{2}$, completando a base ortogonal $\beta^\star = \{ q_1(x), q_2(x) \}$.

Finalmente, vamos determinar a projeção ortogonal, $\widetilde{q}(x)$, do elemento $q(x) = 1 - x^2$ no subespaço $\mathcal{P}_1(\mathbb{R})$ que é dada por:

$$\widetilde{q}(x) = \frac{\langle q \,,\, q_1 \rangle}{\langle q_1 \,,\, q_1 \rangle}\, q_1(x) + \frac{\langle q \,,\, q_2 \rangle}{\langle q_2 \,,\, q_2 \rangle}\, q_2(x)$$

onde

$$\langle q_1 \,,\, q_1 \rangle = \int_0^1 dx = 1 \qquad \text{e} \qquad \langle q_2 \,,\, q_2 \rangle = \int_0^1 \left(x - \frac{1}{2} \right)^2 dx = \frac{1}{12}$$

$$\langle q \,,\, q_1 \rangle = \int_0^1 (1 - x^2)dx = \frac{2}{3}$$

$$\langle q \,,\, q_2 \rangle = \int_0^1 (1 - x^2) \left(x - \frac{1}{2} \right) dx = -\frac{1}{12}$$

Portanto, temos que

$$\widetilde{q}(x) = \frac{2}{3} - \left(x - \frac{1}{2} \right) = \frac{7}{6} - x\,,$$

o que completa a resolução da questão.

B.2.4 Segunda Chamada

Questão 1.

(a) Temos que $P = [p_{ij}]$ é a matriz de mudança da base $\gamma = \{\, p_1(x), p_2(x), p_3(x) \,\}$ para a base $\beta = \{\, 1, x, x^2 \,\}$. Desse modo, obtemos

$$\begin{aligned} p_1(x) &= p_{11} + p_{21}x + p_{31}x^2 = 1 + x \\ p_2(x) &= p_{12} + p_{22}x + p_{32}x^2 = x \\ p_3(x) &= p_{13} + p_{23}x + p_{33}x^2 = 2 + 2x + x^2 \end{aligned}$$

Assim, temos que $\gamma = \{\, 1 + x\,,\, x\,,\, 2 + 2x + x^2 \,\}$.

(b) Sabemos que $[q(x)]_\beta = [I]^\gamma_\beta \, [q(x)]_\gamma$. Temos que

$$[q(x)]_\beta = \begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix} \qquad \text{e vamos denotar} \qquad [q(x)]_\gamma = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Assim, obtemos o seguinte sistema linear

$$\begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{cases} a + 2c = -3 \\ a + b + 2c = -2 \\ c = 2 \end{cases}$$

que possui uma única solução $a = -7\,,\ b = 1$ e $c = 2$. Logo, $[q(x)]_\gamma = \begin{bmatrix} -7 \\ 1 \\ 2 \end{bmatrix}$.

Questão 2.

Vamos determinar uma base para o subespaço S. Temos que todo elemento $(x, y, z) \in S$ satisfaz a equação $x + y + z = 0$. Logo, temos que $z = -x - y$. Desse modo, obtemos que todo elemento $(x, y, z) \in S$ é escrito da seguinte forma:

$$(x, y, z) = x(1, 0, -1) + y(0, 1, -1) \quad \text{para} \quad x, y \in \mathbb{R}.$$

Portanto, o conjunto $\beta = \{ (1, 0, -1), (0, 1, -1)$ é uma base para o subespaço S. De fato, tomando uma combinação linear nula dos elementos do conjunto β

$$\alpha_1(1, 0, -1) + \alpha_2(0, 1, -1) = (0, 0, 0)$$

obtemos $\alpha_1 = \alpha_2 = 0$.

Agora vamos determinar uma base para o subespaço $S^\perp$ definido por:

$$S^\perp = \{ w \in \mathbb{R}^3 \ / \ \langle w, v \rangle = 0 \quad ; \quad \forall \ v \in S \}.$$

Sabemos que todo elemento $w = (a, b, c) \in S^\perp$ deve ser ortogonal aos elemento da base β do subespaço S. Assim, temos que

$$\langle w, v_1 \rangle = a - c = 0$$

$$\langle w, v_2 \rangle = b - c = 0$$

onde $v_1 = (1, 0, -1)$ e $v_2 = (0, 1, -1)$.

Desse modo, obtemos $a = c$ e $b = c$ para $c \in \mathbb{R}$. Assim, todo elemento $w = (a, b, c) \in S^\perp$ é escrito da seguinte forma:

$$(a, b, c) = c(1, 1, 1) \quad \text{para} \quad c \in \mathbb{R}.$$

Logo, $S^\perp = [(1, 1, 1)]$.

Considerando o espaço vetorial $\mathbb{R}^3$ com a base $\gamma = \{ (1, 0, 0), (0, 1, 0), (1, 1, 1) \}$, vamos definir o operador linear T sobre o $\mathbb{R}^3$, satisfazendo $Im(T) = S$ e $Ker(T) = S^\perp$, da seguinte forma:

$$T(1, 0, 0) = (1, 0, -1)$$

$$T(0, 1, 0) = (0, 1, -1)$$

$$T(1, 1, 1) = (0, 0, 0)$$

Vamos escrever um elemento genérico $(x, y, z) \in I\!R^3$ com relação à base γ, isto é,

$$(x, y, z) = a(1, 0, 0) + b(0, 1, 0) + c(1, 1, 1) = (a + c\,,\, b + c\,,\, c)$$

obtendo $c = z$, $b = y - z$ e $a = x - z$. Assim, temos que

$$(x, y, z) = (x - z)(1, 0, 0) + (y - z)(0, 1, 0) + z(1, 1, 1)\,.$$

Finalmente, obtemos a expressão do operador T que é dada por:

$$\begin{aligned} T(x, y, z) &= (x - z)T(1, 0, 0) + (y - z)T(0, 1, 0) + zT(1, 1, 1) \\ &= (x - z)(1, 0, -1) + (y - z)(0, 1, -1) + z(0, 0, 0) \\ &= (\, x - z\,,\, y - z\,,\, -x - y + 2z\,)\,. \end{aligned}$$

Portanto, temos que o operador linear

$$T(x, y, z) = (\, x - z\,,\, y - z\,,\, -x - y + 2z\,) \qquad \text{para} \qquad (x, y, z) \in I\!R^3$$

satisfaz as condições exigidas.

Questão 3.

(a) Vamos determinar uma base β_1 para o subespaço U_1. Temos que todo elemento $(x, y) \in U_1$ deve satisfazer a seguinte condição $T(x, y) = 5(x, y)$, isto é,

$$(3x - 2y\,,\, -2x + 3y) = (5x, 5y) \iff (-2x - 2y\,,\, -2x - 2y) = (0, 0)\,.$$

Logo, temos uma única equação $x + y = 0$, isto é, $y = -x$.

Assim, todo elemento $(x, y) \in U_1$ é escrito como:

$$(x, y) = x(1, -1) \qquad \text{para} \qquad x \in I\!R.$$

Portanto, temos que $\beta_1 = \{\, (1, -1) \,\}$.

Agora vamos determinar uma base β_2 para o subespaço U_2. Temos que todo elemento $(x, y) \in U_2$ deve satisfazer a seguinte condição $T(x, y) = (x, y)$, isto é,

$$(3x - 2y\,,\, -2x + 3y) = (x, y) \iff (2x - 2y\,,\, -2x + 2y) = (0, 0)\,.$$

Logo, temos uma única equação $x - y = 0$, isto é, $y = x$.

Assim, todo elemento $(x, y) \in U_2$ é escrito como:

$$(x, y) = x(1, 1) \qquad \text{para} \qquad x \in I\!R.$$

Portanto, temos que $\beta_2 = \{\, (1, 1) \,\}$.

(b) O conjunto $\beta = \beta_1 \cup \beta_2 = \{\, (1, -1)\,,\, (1, 1) \,\}$ é linearmente independente. De fato, podemos observar que β é um conjunto ortogonal. Logo, β é uma base ortogonal para o $I\!R^2$.

Finalmente, vamos determinar a matriz $[T]_\beta^\beta$. Temos que

$$T(1, -1) = (5, -5) = 5(1, -1) + 0(1, 1)$$

$$T(1, 1) = (1, 1) = 0(1, -1) + 1(1, 1)$$

Portanto, obtemos

$$[T]_\beta^\beta = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix},$$

o que completa a resolução do questão.

Questão 4.

(a) Temos os elementos $u = (1, 1, 1)$ e $v = (3, 2, 1)$, e as seguintes condições:

(1) $v = w_1 + w_2$.

(2) O elemento w_1 é ortogonal ao elemento u, isto é $\langle w_1 , u \rangle = 0$.

(3) O conjunto $\{ w_2, u \}$ é linearmente dependente, isto é, existe um escalar $\alpha \in \mathbb{R}$ tal que $w_2 = \alpha u$.

Substituindo o elemento $w_2 = \alpha u$, dado pela condição **(3)**, na condição **(1)** e calculando o produto interno $\langle v , u \rangle$ utilizando a condição **(2)**, obtemos

$$\langle v , u \rangle = \langle w_1 + \alpha u , u \rangle = \langle w_1 , u \rangle + \alpha \langle u , u \rangle \implies \alpha = \frac{\langle v , u \rangle}{\langle u , u \rangle} = \frac{6}{3} = 2 .$$

Assim, temos que

$$w_2 = (2, 2, 2) \qquad \text{e} \qquad w_1 = v - w_2 = (1, 0, -1) .$$

(b) Vamos representar o elemento $w = (1, -1, 2)$ da seguinte forma:

$$w = \widetilde{w} + \overline{w} \qquad \text{onde} \qquad \widetilde{w} \in S \quad \text{e} \quad \overline{w} \in S^{\perp} ,$$

isto é, $\widetilde{w}$ é a projeção ortogonal de w sobre o subespaço S e $\overline{w}$ é a projeção ortogonal de w sobre o subespaço $S^{\perp}$.

Como o conjunto $\{ w_1, u \}$ é uma base ortogonal para o subespaço $S = [w_1, u]$, temos que o elemento $\widetilde{w}$ é calculado da seguinte forma:

$$\widetilde{w} = \frac{\langle w , u \rangle}{\langle u , u \rangle} u + \frac{\langle w , w_1 \rangle}{\langle w_1 , w_1 \rangle} w_1 .$$

Assim, temos que

$$\widetilde{w} = \frac{2}{3}(1, 1, 1) - \frac{1}{2}(1, 0, -1) = \left(\frac{1}{6} , \frac{4}{6} , \frac{7}{6} \right) .$$

Finalmente, temos que o elemento $\overline{w} = w - \widetilde{w}$. Logo,

$$\overline{w} = (1, -1, 2) - \left(\frac{1}{6} , \frac{4}{6} , \frac{7}{6} \right) = \left(\frac{5}{6} , -\frac{10}{6} , \frac{5}{6} \right) .$$

Questão 5.

(a) Temos que o subespaço $U = [u_1, u_2]$, onde $u_1 = (1,1,1,1)$ e $u_2 = (-1,1,-1,1)$. Note que os elementos u_1 e u_2 são ortogonais.

Sabemos que a melhor aproximação do elemento $v = (2,1,3,1) \in I\!R^4$ no subespaço U é dada pela projeção ortogonal do elemento v sobre o subespaço U.

Como o conjunto $\{ u_1, u_2 \}$ é uma base ortogonal para o subespaço U, temos que a projeção ortogonal, $\widetilde{v}$, do elemento v no subespaço U é calculada da seguinte forma:

$$\widetilde{v} = \frac{\langle v \,,\, u_1 \rangle}{\langle u_1 \,,\, u_1 \rangle} u_1 + \frac{\langle v \,,\, u_2 \rangle}{\langle u_2 \,,\, u_2 \rangle} u_2 \,.$$

Assim, temos que

$$\widetilde{v} = \frac{7}{4}(1,1,1,1) - \frac{3}{4}(-1,1,-1,1) = \left(\frac{5}{2} \,,\, 1 \,,\, \frac{5}{2} \,,\, 1 \right).$$

(b) Basta considerar W como sendo o complemento ortogonal do subespaço U em $I\!R^4$ com relação ao produto interno usual. Pelo **Teorema da Decomposição Ortogonal**, temos que $I\!R^4 = U \oplus W$. Como $dim(U) = 2$, temos que $dim(W) = 2$.

Finalmente vamos determinar uma base para o subespaço W. Sabemos que todo elemento $w = (a,b,c,d) \in W = U^{\perp}$ deve ser ortogonal aos elemento da base de U, isto é,

$$\langle w \,,\, u_1 \rangle = a + b + c + c = 0$$

$$\langle w \,,\, u_2 \rangle = -a + b - c + c = 0$$

Assim, obtemos um sistema linear homogêneo

$$\begin{cases} a + b + c + d = 0 \\ -a + b - c + d = 0 \end{cases} \iff \begin{cases} a + b + c + d = 0 \\ 2b + 2d = 0 \end{cases}$$

que possui como solução $b = -d$ e $a = -c$.

Assim, todo elemento $w = (a,b,c,d) \in W = U^{\perp}$ é escrito da seguinte forma:

$$(a,b,c,d) = c(-1,0,1,0) + d(0,-1,0,1) \quad \text{para} \quad c,\, d \in I\!R \,.$$

O conjunto $\beta = \{ (-1,0,1,0) \,,\, (0,-1,0,1) \}$ é claramente linearmente independente.

Portanto, o conjunto β é uma base para o subespaço $W = U^{\perp}$.

B.2.5 Exame

Questão 1.

(a) Podemos verificar facilmente que U é um subconjunto não vazio, pois o polinômio identicamente nulo satisfaz a condição para que um elemento de $\mathcal{P}_2(\mathbb{R})$ pertença a U. Desse modo, $0_{\mathcal{P}_2(\mathbb{R})} \in U$.

Assim, devemos mostrar que U é fechado com relação à operação de adição e fechado com relação à operação de multiplicação por escalar.

Tomando $p(x),\, q(x) \in U$, isto é, satisfazendo

$$\int_{-1}^{1} p(x)dx \; + \; p'(0) = 0 \qquad \text{e} \qquad \int_{-1}^{1} q(x)dx \; + \; q'(0) = 0\,.$$

Logo, temos que

$$\begin{aligned} \int_{1}^{1} (p+q)(x)dx \; + \; (p+q)'(0) &= \int_{1}^{1} (\,p(x) + q(x)\,)dx \; + \; p'(0) \; + \; q'(0) \\ &= \left\{ \int_{-1}^{1} p(x)dx + p'(0) \right\} + \left\{ \int_{-1}^{1} q(x)dx + q'(0) \right\} \\ &= 0 \end{aligned}$$

Assim, mostramos que $(\,p(x) + q(x)\,) \in U$.

Tomando $p(x) \in U$ e $\lambda \in \mathbb{R}$, temos que

$$\begin{aligned} \int_{1}^{1} (\lambda p)(x)dx \; + \; (\lambda p)'(0) &= \int_{1}^{1} \lambda p(x)dx \; + \; \lambda p'(0) \\ &= \lambda \left\{ \int_{-1}^{1} p(x)dx + p'(0) \right\} = 0\,. \end{aligned}$$

Assim, mostramos que $\lambda p(x) \in U$.

Portanto, mostramos que o subconjunto U é um subespaço vetorial de $\mathcal{P}_2(\mathbb{R})$.

(b) Vamos determinar uma base para o subespaço U. Tomando um elemento genérico $p(x) = a + bx + cx^2 \in \mathcal{P}_2(\mathbb{R})$, e impondo a condição que $p(x) \in U$, isto é,

$$\int_{-1}^{1} (a + bx + cx^2)dx + b = 0,$$

obtemos uma equação algébrica

$$6a + 3b + 2c = 0,$$

que possui dois grau de liberdade, de onde concluímos que $dim(U) = 2$.

Assim, temos que

$$c = -3a - \frac{3}{2}b \quad ; \quad a, b \in \mathbb{R}.$$

Logo, todo elemento $p(x) \in U$ é escrito da seguinte forma:

$$p(x) = (1 - 3x^2)a + \left(x - \frac{3}{2}x^2\right)b \quad ; \quad a, b \in \mathbb{R}.$$

Portanto, temos que o conjunto

$$\gamma = \left\{ 1 - 3x^2, x - \frac{3}{2}x^2 \right\}$$

é uma base para o subespaço U.

Questão 2.

(a) Como $dim(V) = 3$, basta mostrar que o conjunto γ é linearmente independente. Para isso, vamos considerar uma combinação linear nula dos elementos do conjunto γ

$$c_1w_1 + c_2w_2 + c_3w_3 = 0_V .$$

Substituindo as relações entre os elementos de γ e os elementos da base β temos que

$$c_1(u_1 - u_2 - u_3) + c_2(2u_2 + 3u_3) + c_3(3u_1 + u_3) = 0_V ,$$

que reorganizando os termos, obtemos

$$(c_1 + 3c_3)u_1 + (-c_1 + 2c_2)u_2 + (-c_1 + 3c_3 + c_3)u_3 = 0_V .$$

Como o conjunto β é linearmente independente, pois é uma base de V, implica que os coeficientes da combinação linear nula acima devem ser todos iguais a zero. Assim, obtemos o seguinte sistema linear homogêneo

$$\begin{cases} c_1 & & + & 3c_3 & = & 0 \\ -c_1 & + & 2c_2 & & = & 0 \\ -c_1 & + & 3c_2 & + & c_3 & = & 0 \end{cases}$$

que possui somente a solução trivial $c_1 = c_2 = c_3 = 0$. Logo, provamos que o conjunto γ é linearmente independente. Portanto, o conjunto γ é uma base para V.

(b) Podemos encontrar facilmente a matriz $[I]^{\gamma}_{\beta}$ utilizando as relações entre os elementos de γ e os elementos da base β. Assim, temos que

$$[I]^{\gamma}_{\beta} = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 2 & 0 \\ -1 & 3 & 1 \end{bmatrix} .$$

(c) Sabemos que $[v]_\beta = [I]^{\gamma}_{\beta}\,[v]_\gamma$. Desse modo, obtemos

$$[v]_\beta = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 2 & 0 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} ,$$

o que completa a resolução da questão.

Questão 3.

Inicialmente vamos determinar um conjunto de geradores para o subespaço U.

Sabemos que todo elemento $(x, y, z) \in U$ satisfaz a equação algébrica

$$x + y + z = 0 \iff z = -x - y .$$

Logo, todo elemento $(x, y, z) \in U$ é escrito da seguinte forma:

$$(x, y, z) = x(1, 0, -1) + y(0, 1, -1) \quad ; \quad x, y \in \mathbb{R} .$$

Portanto, temos que $U = [(1, 0, -1), (0, 1, -1)] = Im(T)$.

Agora considerando o subespaço $W = [(1, 0, 1), (0, -1, 1)]$, vamos determinar uma base para o subespaço $U \cap W$, que é o núcleo do operador que desejamos encontrar.

Tomando um elemento $v \in U \cap W$, isto é, $v \in U$ e $v \in W$. Logo, temos que existem escalares $a, b, c, d \in \mathbb{R}$ tais que

$$v = a(1, 0, -1) + b(0, 1, -1) = c(1, 0, 1) + d(0, -1, 1) .$$

Desse modo, obtemos o seguinte sistema linear

$$\begin{cases} a = c \\ b = -d \\ -a - b = c + d \end{cases}$$

que possui a seguinte solução $c = 0$, $a = 0$ e $b = -d$ para $d \in \mathbb{R}$.

Portanto, todo elemento $v \in U \cap W$ é escrito da seguinte forma:

$$v = b(0, 1, -1) \quad ; \quad b \in \mathbb{R} .$$

Logo, temos que $U \cap W = [(0, 1, -1)] = Ker(T)$.

Completamos a base do subespaço $Ker(T)$ para obter uma base para o espaço vetorial $\mathbb{R}^3$. Assim, podemos escolher $\gamma = \{ (0, 1, -1), (1, 0, 0), (0, 1, 0) \}$ uma base para $\mathbb{R}^3$.

Finalmente definimos o operador linear T com as características solicitadas

$$\begin{aligned} T(0,1,-1) &= (0,0,0) \\ T(1,0,0) &= (1,0,-1) \\ T(0,1,0) &= (0,1,-1) \end{aligned}$$

Para obtermos a expressão do operador T, vamos inicialmente representar um elemento genérico $(x,y,z) \in I\!R^3$ com relação à base ordenada γ, isto é,

$$(x,y,z) = a(0,1,-1) + b(1,0,0) + c(0,1,0)\,,$$

obtendo o seguinte sistema linear nas incógnitas a, b, c

$$\begin{cases} b = x \\ a + c = y \\ -a = z \end{cases}$$

Assim, temos que $a = -z$, $b = x$ e $c = y + z$.

Logo, todo elemento $(x,y,z) \in I\!R^3$ é escrito da seguinte forma:

$$(x,y,z) = -z(0,1,-1) + x(1,0,0) + (y+z)(0,1,0)\,.$$

Portanto,

$$\begin{aligned} T(x,y,z) &= -zT(0,1,-1) + xT(1,0,0) + (y+z)T(0,1,0) \\ T(x,y,z) &= -z(0,0,0) + x(1,0,-1) + (y+z)(0,1,-1) \\ T(x,y,z) &= (x,\, y+z,\, -x-y-z) \end{aligned}$$

o que completa a resolução da questão.

Questão 4.

Vamos verificar se o operador T é um automorfismo de $\mathcal{P}_3(\mathbb{R})$. Para isso, devemos determinar o subespaço $Ker(T)$, isto é,

$$Ker(T) = \{ p(x) \in \mathcal{P}_3(\mathbb{R}) \ / \ T(p(x)) = 0_{\mathcal{P}_3(\mathbb{R})} \}.$$

Tomando um elemento genérico $p(x) = a + bx + cx^2 + dx^3$, vamos impor a condição que $p(x) \in Ker(T)$, isto é,

$$\begin{aligned} T(p(x)) &= (a + bx + cx^2 + dx^3) + (1 + x)(b + 2cx + 3dx^2) = 0_{\mathcal{P}_3(\mathbb{R})} \\ &= (a + b) + (2b + 2c)x + (3c + 3d)x^2 + 4dx^3 = 0_{\mathcal{P}_3(\mathbb{R})} \end{aligned}$$

Isso nos leva a um sistema linear homogêneo cuja solução é $a = b = c = d = 0$.

Logo, $Ker(T) = \{ 0_{\mathcal{P}_3(\mathbb{R})} \}$, isto é, T é um operador injetor. Pelo Teorema do núcleo e da imagem, sabemos que $Im(T) = \mathcal{P}_3(\mathbb{R})$, isto é, T é um operador sobrejetor. Portanto, mostramos que T é um automorfismo de $\mathcal{P}_3(\mathbb{R})$.

Finalmente vamos determinar a representação matricial do operador T com relação à base canônica $\beta = \{ 1, x, x^2, x^3 \}$ de $\mathcal{P}_3(\mathbb{R})$. Para isso, vamos calcular

$$\begin{aligned} T(1) &= 1 \\ T(x) &= 1 + 2x \\ T(x^2) &= 2x + 3x^2 \\ T(x^3) &= 3x^2 + 4x^3 \end{aligned}$$

Desse modo, obtemos que

$$[T]^{\beta}_{\beta} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

o que completa a resolução da questão.

Questão 5.

Chamando $p_1(x) = 1 + x$ e $p_2 = 1 - x^2$, temos que o subespaço $S = [p_1(x), p_2(x)]$.

O subespaço $S^\perp$ é definido por:

$$S^\perp = \{ \, q(x) \in \mathcal{P}_2(I\!\!R) \ / \ \langle \, r \, , \, q \, \rangle = 0 \quad ; \quad \forall \ r(x) \in S \, \} \, .$$

Tomando um elemento genérico $q(x) = a + bx + cx^2 \in S^\perp$, sabemos que

$$\begin{aligned}
\langle \, p_1 \, , \, q \, \rangle &= \int_{-1}^{1} (\, a + bx + cx^2 \,)(\, 1 + x \,) dx \\
&= \int_{-1}^{1} (\, a + bx + cx^2 + ax + bx^2 + cx^3 \,) dx \\
&= \int_{-1}^{1} (\, a + cx^2 + bx^2 \,) dx = 0 \\
\langle \, p_2 \, , \, q \, \rangle &= \int_{-1}^{1} (\, a + bx + cx^2 \,)(\, 1 - x^2 \,) dx \\
&= \int_{-1}^{1} (\, a + bx + cx^2 - ax^2 - bx^3 - cx^4 \,) dx \\
&= \int_{-1}^{1} (\, a + cx^2 - ax^2 - cx^4 \,) dx = 0
\end{aligned}$$

Calculando as integrais acima obtemos o sistema linear homogêneo

$$\left\{ \begin{array}{rcrcrcl} 2a & + & \dfrac{2}{3}c & + & \dfrac{2}{3}b & = & 0 \\ \\ \dfrac{4}{3}a & & & + & \dfrac{4}{15}c & = & 0 \end{array} \right. \iff \left\{ \begin{array}{rcrcrcl} 6a & + & 2c & + & 2b & = & 0 \\ \\ 5a & & & + & c & = & 0 \end{array} \right.$$

que possui a seguinte solução $b = 2a$ e $c = -5a$ para $a \in I\!\!R$.

Desse modo, todo elemento $q(x) \in S^\perp$ é escrito da seguinte forma:

$$q(x) = a(\, 1 + 2x - 5x^2 \,) \qquad ; \qquad a \in I\!\!R \, .$$

Portanto, uma base para o subespaço $S^\perp$ é dada pelo conjunto

$$\gamma = \{ \, 1 + 2x - 5x^2 \, \} \, ,$$

o que completa a resolução da questão.