Apostila

MAT 27 - ÁLGEBRA LINEAR

Fernanda Pereira, Tiara Martini

INSTITUTO DE TECNOLÓGICO DE AERONÁUTICA

Departamento de Matemática

Sumário

Introdução

Esse texto foi elaborado a partir das notas de aula das autoras para o curso de MAT-27 - Álgebra Linear, disciplina obrigatória aos cursos de engenharia do ITA. Os pré-requisitos para sua leitura são os conteúdos do ensino médio, além de certa maturidade matemática. O texto é autossuficiente e contém todas as demonstrações dos resultados apresentados, com exceção do Capítulo 1, que é considerado uma revisão, e por isso decidimos omitir suas demonstrações, para que ficasse mais sucinto.

Esse material aborda todo o conteúdo que normalmente é dado nas disciplinas de Álgebra Linear em cursos de graduação nas áreas de exatas. Nele são considerados espaços vetoriais reais e complexos, com foco nos espaços de dimensão finita. Apesar disso, quando possível e sem grandes complicações, apresentamos os resultados de maneira geral e fazemos alguns comentários sobre as diferenças, quando há, com o caso de dimensão infinita.

O Capítulo 1 apresenta os conceitos e principais propriedades de matrizes, determinantes e sistemas lineares. Apesar de se tratar de um conteúdo de ensino médio, normalmente ele não é visto de maneira formal e geral, como descrito no texto. Dada a importância de seu domínio para se trabalhar com os assuntos de álgebra linear, recomendamos fortemente a leitura dele a todos os alunos, seja para revisar ou mesmo aprofundar seus conhecimentos.

O Capítulo 2 apresenta o conceito de espaço vetorial e seus fundamentos essenciais: subespaços, conjunto gerador, conjuntos linearmente dependentes e independentes, bases, dimensão e coordenadas.

O Capítulo 3 apresenta o conceito de produto interno num espaço vetorial e as propriedades e conceitos relacionados: ortogonalidade, Gram-Schmidt e projeção ortogonal.

O Capítulo 4 se refere ao estudo das transformações lineares, que são as "funções interessantes" entre os espaços vetoriais. Serão apresentados conceitos como: núcleo, imagem, isomorfismos e matriz de transformação linear. Por fim, estuda-se diagonalização de operadores lineares.

Vale mencionar que os Capítulos 3 e 4 são independentes, tendo como base apenas o Capítulo 2.

O Capítulo 5 se refere ao estudo das transformações lineares entre espaços com produto interno, e esse sim envolve todo o conteúdo dos capítulos anteriores. O objetivo principal desse capítulo é apresentar o Teorema Espectral, nas suas versões para espaços reais e complexos.

O Capítulo 6 apresenta uma aplicação da teoria de Álgebra Linear: o reconhecimento de quádricas a partir das suas equações algébricas na forma geral.

Ao longo do texto, o conjunto dos números naturais será denotado por $\mathbb{N} = \{1, 2, \dots\}$ e os corpos dos números reais e complexos serão denotados por $\mathbb{R}$ e $\mathbb{C}$, respectivamente. Quando dissermos um *número*, *constante* ou *escalar*, estaremos nos referindo a elementos de $\mathbb{R}$ ou $\mathbb{C}$, salvo menção contrária. De um modo mais geral, denotaremos o corpo de escalares por $\mathbb{K}$, significando $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. A cardinalidade de um conjunto X será denotada por $\#X$.

Capítulo 1

Revisão: matrizes, determinantes e sistemas lineares

Neste capítulo, serão apresentados os conceitos e principais propriedades de matrizes, determinantes e sistemas lineares. Esses assuntos serão muito utilizados ao logo deste curso, por isso é preciso dominar muito bem o conteúdo deste capítulo. Apesar de se tratar de um conteúdo de ensino médio, nem sempre ele é visto da maneira mais geral como se encontra aqui. Assim, recomendamos a todos os alunos a leitura deste capítulo, a fim de revisar, ou mesmo se aprofundar no assunto.

As demonstrações dos resultados deste capítulo serão omitidas, e podem ser encontradas [Reg], cuja versão digital encontra-se disponível na página do autor: Reginaldo Santos - livros.

1.1 Matrizes

Uma *matriz* A, $m \times n$, $m,n \in \mathbb{N}$, é uma tabela de mn números, denotados por a_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$, dispostos em m linhas e n colunas:

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

A *i-ésima linha* de A é

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}, \quad i = 1, \ldots, m.$$

A *j-ésima coluna* de A é

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad j = 1, \ldots, n.$$

Também usaremos a notação $A = (a_{ij})_{m\times n}$. Dizemos que a_{ij} é o *elemento* ou *entrada* de posição i, j de A, $m \times n$ é a *ordem* de A e que A é *quadrada de ordem n* se $m = n$.

A *diagonal principal* é formada pelos elementos a_{ii}.

A matriz A é dita ser *triangular superior* (respectivamente *triangular inferior*) se todos as entradas abaixo (resp. acima) da diagonal principal são nulas, isto é, $a_{ij} = 0$ para todo $i > j$ (resp. $i < j$).

A matriz A é dita ser *diagonal* se todos as entradas fora da diagonal principal são nulas, isto é, $a_{ij} = 0$ para todo $i \neq j$. Neste caso, também denota-se $A = \operatorname{diag}(a_{11}, \ldots, a_{nn})$.

A *matriz identidade de ordem n* é $I_n = (a_{ij})_{n \times n}$ onde $a_{ii} = 1$ para todo $i = 1, \ldots, n$ e $a_{ij} = 0$ se $i \neq j$.

Definição 1.1.1 Sejam $A = (a_{ij})_{m \times n}, B = (b_{ij})_{m \times n}, C = (c_{jk})_{n \times p}$ matrizes e α um escalar. A *soma* $A + B$ é a matriz $D = (d_{ij})_{m \times n}$ tal que

$$d_{ij} = a_{ij} + b_{ij} \quad \forall\, i = 1, \ldots, m,\ j = 1, \ldots, n.$$

A *multiplicação por escalar* αA é a matriz $E = (e_{ij})_{m \times n}$ tal que

$$e_{ij} = \alpha a_{ij} \quad \forall\, i = 1, \ldots, m,\ j = 1, \ldots, n.$$

A *multiplicação* AC é a matriz $F = (f_{ik})_{m \times p}$ tal que

$$f_{ik} = \sum_{j=1}^{n} a_{ij} c_{jk} \quad \forall\, i = 1, \ldots, m,\ k = 1, \ldots, p.$$

A *matriz transposta de A* é a matriz $A^\top = (g_{ij})_{n \times m}$ tal que

$$g_{ij} = a_{ji} \quad \forall\, i = 1, \ldots, m,\ j = 1, \ldots, n.$$

A proposição a seguir sintetiza as principais propriedades envolvendo as operações com matrizes definidas anteriormente.

Proposição 1.1.2 Sejam A, B, C matrizes $m \times n$, D matriz $n \times p$, E matriz $p \times q$, F matriz $r \times m$ e α, β escalares. São válidas as seguintes propriedades.

(i) Comutatividade da soma:
$$A + B = B + A.$$

(ii) Associatividade da soma:
$$A + (B + C) = (A + B) + C.$$

(iii) Existência de elemento neutro para a soma:
$$A + 0 = A,$$
onde 0 aqui denota a matriz nula de ordem $m \times n$. Usaremos a mesma notação para o escalar nulo, mas pelo contexto ficará óbvio a qual deles estamos nos referindo.

(iv) Existência de inverso aditivo: para cada matriz $A = (a_{ij})_{m \times n}$ existe única $-A = (-a_{ij})_{m \times n}$ tal que
$$A + (-A) = 0.$$

(v) Associatividade da multiplicação por escalar:
$$\alpha(\beta A) = (\alpha\beta)A \qquad \text{e} \qquad \alpha(AD) = (\alpha A)D = A(\alpha D).$$

(vi) Distributividade da soma com multiplicação por escalar:
$$(\alpha + \beta)A = \alpha A + \beta A \qquad \text{e} \qquad \alpha(A + B) = \alpha A + \alpha B.$$

(vii) Associatividade do produto:
$$A(DE) = (AD)E.$$

(viii) Existência de elemento neutro para o produto:
$$AI_n = I_m A = A,$$
onde I_s é a matriz identidade de ordem s.

(ix) Distributividade da soma com produto:
$$F(A+B) = FA + FB \qquad \text{e} \qquad (A+B)D = AD + BD.$$

(x) $(A^\top)^\top = A$.

(xi) $(A+B)^\top = A^\top + B^\top$.

(xii) $(\alpha A)^\top = \alpha A^\top$.

(xiii) $(AD)^\top = D^\top A^\top$.

(xiv) Se $X = (x_i)_{n\times 1}$ é uma matriz $n \times 1$, então
$$AX = \sum_{j=1}^{n} x_j A_j,$$
onde A_j é a j-ésima coluna de A. Em termos matriciais,
$$[A_1\ A_2\ \cdots\ A_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 A_1 + \cdots + x_n A_n.$$

Uma matriz quadrada $A = (a_{ij})_{n\times n}$ é *inversível* ou *não singular* se existe uma matriz $B = (b_{ij})_{n\times n}$ tal que
$$AB = BA = I_n.$$
Neste caso a matriz B é chamada de *inversa* de A. Se A não tem inversa dizemos que A é *não inversível* ou *singular*.

Proposição 1.1.3 Sejam $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ matrizes.

(i) Se A é inversível, então a inversa é única. Neste caso, denotaremos A^{-1} a matriz inversa de A.

(ii) Se $AB = I_n$, então $BA = I_n$.

(iii) Se A é inversível, então A^{-1} também o é, e
$$(A^{-1})^{-1} = A.$$

(iv) Se A e B são inversíveis, então AB também o é, e
$$(AB)^{-1} = B^{-1}A^{-1}.$$

(v) Se A é inversível, então $A^\top$ também o é, e

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

(vi) Se AB é inversível, então A e B também o são.

O *traço* de uma matriz quadrada $A = (a_{ij})_{n\times n}$ é definido por

$$\mathrm{tr}\,(A) = a_{11} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

A proposição a seguir sintetiza as principais propriedades relacionadas ao traço de matrizes.

Proposição 1.1.4 Sejam $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ matrizes e α um escalar.

(i) $\mathrm{tr}\,(A+B) = \mathrm{tr}\,(A) + \mathrm{tr}\,(B)$.

(ii) $\mathrm{tr}\,(\alpha A) = \alpha\,\mathrm{tr}\,(A)$.

(iii) $\mathrm{tr}\,(A^\top) = \mathrm{tr}\,(A)$.

(iv) $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$.

Observação 1.1.5 Não é verdade, em geral, que $\mathrm{tr}\,(AB) = \mathrm{tr}\,(A)\,\mathrm{tr}\,(B)$ (encontre um contra-exemplo). Não confunda isso com o item (iv) da Proposição 1.1.4!

1.2 Determinantes

Determinante, nada mais é que uma função, que associa a cada matriz quadrada $A \in M_n(\mathbb{K})$, um escalar $\det A \in \mathbb{K}$ (ou $\det(A)$). Essa função possui algumas propriedades interessantes, que muitas vezes permitem obter informações importantes sobre as matrizes economizando muitos cálculos. Existem algumas maneiras distintas e equivalentes de definir o determinante de uma matriz. Aqui apresentaremos a definição por cofatores, que é dada via indução sobre a ordem da matriz.

Se $A = (a_{11})$ é uma matriz 1×1, define-se o *determinante* de A por

$$\det A = a_{11}.$$

Para matrizes quadradas de ordem maior que 1, definiremos o determinante por indução. Seja $A = (a_{ij})_{n\times n}$ uma matriz com $n \geq 2$ e suponha que o determinante de matrizes de ordem $(n-1)\times(n-1)$ foi definido. O *menor* do elemento a_{ij}, denotado por $\tilde{A}_{ij}$, é a submatriz de ordem $(n-1)\times(n-1)$ obtida de A retirando-se a i-ésima linha e a j-ésima coluna de A. O *cofator* do elemento a_{ij}, denotado por $\tilde{a}_{ij}$, é definido por

$$\tilde{a}_{ij} = (-1)^{i+j}\det(\tilde{A}_{ij}).$$

O *determinante* de A é definido por

$$\det A = \sum_{k=1}^{n} a_{1k}\tilde{a}_{1k}. \tag{1.1}$$

A expressão (1.1) é chamada de *desenvolvimento do determinante de A em cofatores na primeira linha*.

Teorema 1.2.1 Seja $A = (a_{ij})_{n\times n}$ uma matriz. O determinante de A pode ser calculado fazendo-se o desenvolvimento em cofatores em qualquer linha ou qualquer coluna, isto é,

$$\det A = \sum_{k=1}^{n} a_{ik}\tilde{a}_{ik} \quad \forall\ i = 1,\dots,n, \tag{1.2}$$

$$= \sum_{k=1}^{n} a_{kj}\tilde{a}_{kj} \quad \forall\ j = 1,\dots,n. \tag{1.3}$$

A expressão (1.2) é chamada de *desenvolvimento do determinante de A em cofatores na i-ésima linha*, e a expressão (1.3) é chamada de *desenvolvimento do determinante de A em cofatores na j-ésima coluna*.

A seguir, apresentamos as principais propriedades sobre determinantes.

Proposição 1.2.2 Sejam $A = (a_{ij})_{n\times n}, B = (b_{ij})_{n\times n}$ matrizes e α um escalar. São válidas as seguintes propriedades.

(i) Se A é uma matriz triangular superior ou inferior, então

$$\det A = \prod_{i=1}^{n} a_{ii}.$$

(ii) Se B é obtida de A multiplicando-se uma linha de A por α, então

$$\det B = \alpha \det A.$$

(iii) Se B é obtida de A trocando-se duas linhas distintas de posição, então

$$\det B = -\det A.$$

(iv) Se B é obtida de A substituindo-se a l-ésima linha por ela somada a um múltiplo escalar da k-ésima linha, com $k \neq l$, então

$$\det B = \det A.$$

(v) Escrevendo A em termos das suas linhas:

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_n \end{bmatrix},$$

onde $A_k = [\, a_{k1} \;\; a_{k2} \;\; \cdots \;\; a_{kn} \,]$ é a k-ésima linha de A, se $A_k = X + Y$, onde X, Y são matrizes

$1 \times n$, então

$$\det A = \det \begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ X \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix} + \det \begin{bmatrix} A_1 \\ \vdots \\ A_{k-1} \\ Y \\ A_{k+1} \\ \vdots \\ A_n \end{bmatrix}.$$

(vi) $\det(AB) = \det A \det B$.

(vii) $\det(A^\top) = \det A$.

(viii) A é inversível se, e somente se, $\det A \neq 0$.

Observação 1.2.3
- Pode-se, usando as propriedades (ii), (iii) e (iv), transformar o cálculo do determinante de uma matriz no cálculo do determinante de uma matriz triangular superior, o que é bem mais simples pela propriedade (i).
- Pelo item (vii), todas as propriedades de determinantes referentes às linhas da matriz também são válidas em relação as colunas.

■ **Exemplo 1.2.4** Calcule $\det A$, onde $A = \begin{bmatrix} 0 & 1 & 4 \\ 5 & -10 & 15 \\ 2 & -1 & 7 \end{bmatrix}$.

Solução. Primeiro efetuamos algumas operações com a matriz para simplificar o cálculo do determinante. Para facilitar o entendimento, denotamos por L_i a i-ésima linha.

$$\det A \overset{(L_1 \leftrightarrow L_2)}{=} -\det \begin{bmatrix} 5 & -10 & 15 \\ 0 & 1 & 4 \\ 2 & -1 & 7 \end{bmatrix} = -5 \cdot \det \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 4 \\ 2 & -1 & 7 \end{bmatrix} \overset{(L_3 \leftarrow L_3 - 2L_1)}{=} -5 \cdot \det \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 4 \\ 0 & 3 & 1 \end{bmatrix}$$

$$\overset{(L_3 \leftarrow L_3 - 3L_2)}{=} -5 \cdot \det \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & -11 \end{bmatrix}$$

Como essa última matriz é triangular superior, o seu determinante é o produto das entradas da diagonal principal, logo

$$\det A = -5 \cdot 1 \cdot 1 \cdot (-11) = 55.$$

■

1.3 Sistemas lineares

Uma *equação linear* em n variáveis (ou incógnitas) $x_1, x_2, \ldots, x_n$ é uma equação da forma

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b,$$

onde $a_1, a_2, \ldots, a_n$ e b são constantes. Um *sistema de equações lineares* ou simplesmente *sistema linear* é um conjunto de equacões lineares, ou seja, é um conjunto de equações da forma:

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right. \tag{1.4}$$

onde a_{ij} e b_k são constantes, para $i,k = 1,\ldots,m$ e $j = 1,\ldots,n$. Usando o produto de matrizes, o sistema linear (1.4) pode ser escrito como uma equação matricial $AX = B$, onde

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Uma *solução* do sistema linear é uma matriz $S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$ tal que as equações do sistema são satisfeitas quando substituímos $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$, ou equivalentemente, $AS = B$. Também, costuma-se denotar a solução como uma n-upla $S = (s_1, s_2, \ldots, s_n) \in \mathbb{K}^n$. O conjunto de todas as soluções do sistema é chamado *conjunto solução* ou *solução geral* do sistema. A matriz A á chamada *matriz do sistema linear*. A *matriz aumentada do sistema linear* é a seguinte:

$$[A|B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Dois sistemas lineares são ditos *equivalentes* se possuem exatamente as mesmas soluções.

Proposição 1.3.1 **(i)** Se um sistema linear $AX = B$ possui duas soluções distintas, então ele possui infinitas soluções. Daí segue que uma, e apenas uma, das três possibilidades a seguir pode ocorrer: o sistema tem uma única solução, tem infinitas soluções, ou não tem nenhuma solução.

(ii) Se A é uma matriz quadrada, então o sistema $AX = B$ tem uma única solução se, e somente se, A é inversível. Neste caso a solução é $X = A^{-1}B$.

Se $b_1 = b_2 = \cdots = b_m = 0$, o sistema (1.4) é chamado de *homogêneo*. Tal sistema pode ser escrito na forma matricial $AX = 0$ (onde 0 aqui denota a matriz $m \times 1$ com todas entradas nulas). Todo sistema linear homogêneo admite pelo menos a solução $S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, chamada de *solução trivial*. Assim, um sistema linear homogêneo ou tem somente uma solução (a trivial) ou tem infinitas soluções.

Teorema 1.3.2 Seja $A = (a_{ij})_{m\times n}$ uma matriz com $m < n$. Então, o sistema linear homogêneo $AX = 0$ (que tem m equações e n incógnitas) tem infinitas soluções.

1.4 Operações elementares

Definição 1.4.1 Uma *operação elementar* sobre as linhas de uma matriz é uma das seguintes operações:

(i) trocar a posição de duas linhas da matriz;

(ii) multiplicar uma linha da matriz por um escalar diferente de zero;

(iii) somar a uma linha da matriz um múltiplo escalar de outra linha.

Definição 1.4.2 Duas matrizes são *linha-equivalentes* se uma é obtida da outra após uma quantidade finita de operações elementares sobre suas as linhas.

Teorema 1.4.3 Dois sistemas lineares que possuem matrizes aumentadas linha-equivalentes são equivalentes.

Definição 1.4.4 Uma matriz $A = (a_{ij})_{m\times n}$ está na forma *escalonada reduzida* quando satisfaz as seguintes condições:

(i) Todas as linhas nulas (formadas inteiramente por zeros) ocorrem abaixo das linhas não nulas.

(ii) O *pivô* (primeiro elemento não nulo de uma linha) de cada linha não nula é igual a 1.

(iii) O pivô de cada linha não nula ocorre à direita do pivô da linha anterior.

(iv) Se uma coluna contém um pivô, então todos os seus outros elementos são iguais a zero.

Se A satisfaz as propriedades (i) e (iii), mas não necessariamente (ii) e (iv), dizemos que ela está na forma *escalonada*.

■ **Exemplo 1.4.5** As matrizes A e B a seguir estão na forma escalonada:

$$A = \begin{bmatrix} 3 & 9 & 1 \\ 0 & -1 & -5 \\ 0 & 0 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 6 & -2 & 5 \\ 0 & 0 & -4 & 11 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

As matrizes C e D a seguir estão na forma escalonada reduzida:

$$C = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

■

Devido o Teorema 1.4.3, para se resolver um sistema linear pode-se aplicar o método do *escalonamento*, que consiste em aplicar operações elementares sobre as linhas de sua matriz aumentada para se obter uma matriz na forma escalonada (também conhecido como *método de Gauss*) ou escalonada reduzida (também conhecido como *método de Gauss-Jordan*). O sistema linear que tem essa nova matriz como matriz aumentada tem as mesmas soluções e é bem mais fácil de resolver.

1.5 Método para inversão de matrizes

Seja A uma matriz $n \times n$. Para verificarmos se A é inversível basta verificarmos se existe uma matriz B tal que $AB = I_n$. Vamos denotar as colunas de B por $X_1, X_2, \ldots, X_n$, ou seja, $B = [\, X_1 \ \cdots \ X_n \,]$, onde

$$X_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}, \quad \ldots \quad , \quad X_n = \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{bmatrix},$$

e as colunas da matriz identidade I_n por $E_1, \ldots, E_n$, ou seja, $I_n = [\, E_1 \ \cdots \ E_n \,]$, onde

$$E_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots \quad , \quad E_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Assim,

$$AB = I_n \Leftrightarrow A[\, X_1 \ \cdots \ X_n \,] = [\, AX_1 \ \cdots \ AX_n \,] = [\, E_1 \ \cdots \ E_n \,].$$

Analisando coluna a coluna a equação anterior, vemos que encontrar B é equivalente a resolver n sistemas lineares

$$AX_j = E_j, \quad j = 1, \ldots, n.$$

Cada um dos sistemas pode ser resolvido usando o método de Gauss-Jordan. Para isso, formamos as matrizes aumentadas $[A|E_1], [A|E_2], \ldots, [A|E_n]$. Como as matrizes destes sistemas são todas iguais a A, podemos resolver todos os sistemas simultaneamente formando a matriz $n \times 2n$ $[A \mid E_1 \ E_2 \ \cdots \ E_n] = [A \,|\, I_n]$. Transformando $[A \,|\, I_n]$ na sua forma escalonada reduzida, que vamos denotar por $[R \,|\, S]$, vamos chegar a duas situações possíveis: ou a matriz R é a matriz identidade, ou não é.

- Se $R = I_n$, então a forma escalonada reduzida da matriz $[A \,|\, I_n]$ é da forma $[I_n \,|\, S]$. Se escrevemos a matriz S em termos das suas colunas $S = [S_1 \ S_2 \ \cdots \ S_n]$, então as soluções dos sistemas $I_n X_j = S_j$ são $X_j = S_j$, que são as mesmas soluções dos sistemas $AX_j = E_j$, e assim $B = S$ é tal que $AB = I_n$, logo A é inversível.

- Se $R \neq I_n$, então pode-se mostrar que a matriz R tem uma linha nula. Isso implica que cada um dos sistemas $AX_j = E_j$ ou tem infinitas soluções ou não tem solução. Daí, segue que a matriz A não tem inversa, pois as colunas da (única) inversa seriam X_j, para $j = 1, \ldots, n$.

Com isso, obtemos não somente uma forma de descobrir se uma matriz A tem inversa, mas também, como encontrar a inversa, no caso em que ela exista. Ou seja, escalonamos a matriz $[A \,|\, I_n]$ e encontramos a sua forma escalonada reduzida $[R \,|\, S]$. Se $R = I_n$, então a matriz A é inversível e a inversa é S. Caso contrário, a matriz A não é inversível. Além disso, obtemos o seguinte corolário.

Corolário 1.5.1 Uma matriz A $n \times n$ é inversível se, e somente se, A é linha-equivalente à matriz identidade I_n.

■ **Exemplo 1.5.2** Considere a matriz do Exemplo 1.2.4: $A = \begin{bmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{bmatrix}$. Escalonando a matriz $[A \,|\, I_3]$, temos

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 5 & 1 & 0 & 0 \\ 3 & -6 & 9 & 0 & 1 & 0 \\ 2 & 6 & 1 & 0 & 0 & 1 \end{array}\right] \overset{(L_1 \leftrightarrow L_2)}{\sim} \left[\begin{array}{ccc|ccc} 3 & -6 & 9 & 0 & 1 & 0 \\ 0 & 1 & 5 & 1 & 0 & 0 \\ 2 & 6 & 1 & 0 & 0 & 1 \end{array}\right] \overset{(\frac{1}{3}L_1 \to L_1)}{\sim} \left[\begin{array}{ccc|ccc} 1 & -2 & 3 & 0 & 1/3 & 0 \\ 0 & 1 & 5 & 1 & 0 & 0 \\ 2 & 6 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{(L_3-2L_1\to L_3)}{\sim}\left[\begin{array}{ccc|ccc} 1 & -2 & 3 & 0 & 1/3 & 0\\ 0 & 1 & 5 & 1 & 0 & 0\\ 0 & 10 & -5 & 0 & -2/3 & 1\end{array}\right] \overset{(L_3-10L_2\to L_3)}{\sim}\left[\begin{array}{ccc|ccc} 1 & -2 & 3 & 0 & 1/3 & 0\\ 0 & 1 & 5 & 1 & 0 & 0\\ 0 & 0 & -55 & -10 & -2/3 & 1\end{array}\right]$$

$$\overset{(-\frac{1}{55}L_3\to L_3)}{\sim}\left[\begin{array}{ccc|ccc} 1 & -2 & 3 & 0 & 1/3 & 0\\ 0 & 1 & 5 & 1 & 0 & 0\\ 0 & 0 & 1 & 2/11 & 2/165 & -1/55\end{array}\right]$$

$$\overset{(L_2-5L_3\to L_2)}{\sim}\left[\begin{array}{ccc|ccc} 1 & -2 & 3 & 0 & 1/3 & 0\\ 0 & 1 & 0 & 1/11 & -2/33 & 1/11\\ 0 & 0 & 1 & 2/11 & 2/165 & -1/55\end{array}\right]$$

$$\overset{(L_1-3L_3\to L_1)}{\sim}\left[\begin{array}{ccc|ccc} 1 & -2 & 0 & -30/55 & 49/165 & 3/55\\ 0 & 1 & 0 & 1/11 & -2/33 & 1/11\\ 0 & 0 & 1 & 2/11 & 2/165 & -1/55\end{array}\right]$$

$$\overset{(L_1+2L_2\to L_1)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4/11 & 29/165 & 13/55\\ 0 & 1 & 0 & 1/11 & -2/33 & 1/11\\ 0 & 0 & 1 & 2/11 & 2/165 & -1/55\end{array}\right].$$

Essa última matriz é a forma escalonada reduzida de $[A\,|\,I_3]$. Assim, A é linha-equivalente a matriz identidade I_3 e

$$S=\begin{bmatrix} -4/11 & 29/165 & 13/55\\ 1/11 & -2/33 & 1/11\\ 2/11 & 2/165 & -1/55\end{bmatrix}=A^{-1}.$$

■

■ **Exemplo 1.5.3** Considere a matriz $A=\begin{bmatrix} 1 & 10 & 4\\ -1 & 2 & 0\\ 2 & 11 & 5\end{bmatrix}$. Escalonando a matriz $[A\,|\,I_3]$, temos

$$\left[\begin{array}{ccc|ccc} 1 & 10 & 4 & 1 & 0 & 0\\ -1 & 2 & 0 & 0 & 1 & 0\\ 2 & 11 & 5 & 0 & 0 & 1\end{array}\right] \overset{(L_3-2L_1\to L_3)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 10 & 4 & 1 & 0 & 0\\ -1 & 2 & 0 & 0 & 1 & 0\\ 0 & -9 & -3 & -2 & 0 & 1\end{array}\right]$$

$$\overset{(L_2+L_1\to L_2)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 10 & 4 & 1 & 0 & 0\\ 0 & 12 & 4 & 1 & 1 & 0\\ 0 & -9 & -3 & -2 & 0 & 1\end{array}\right] \overset{(\frac{1}{12}L_2\to L_2)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 10 & 4 & 1 & 0 & 0\\ 0 & 1 & 1/3 & 1/12 & 1/12 & 0\\ 0 & -9 & -3 & -2 & 0 & 1\end{array}\right]$$

$$\overset{(L_3+9L_2\to L_3)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 10 & 4 & 1 & 0 & 0\\ 0 & 1 & 1/3 & 1/12 & 1/12 & 0\\ 0 & 0 & 0 & -5/4 & 3/4 & 1\end{array}\right] \overset{(L_1-10L_2\to L_1)}{\sim}\left[\begin{array}{ccc|ccc} 1 & 0 & 2/3 & 1/6 & -5/6 & 0\\ 0 & 1 & 1/3 & 1/12 & 1/12 & 0\\ 0 & 0 & 0 & -5/4 & 3/4 & 1\end{array}\right]$$

Essa última matriz é a forma escalonada reduzida de $[A\,|\,I_3]$. Assim, a forma escalonada reduzida de A é a matriz

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

que tem uma linha nula. Portanto, a matriz A não é inversível. ■

No caso de matrizes 2×2, às vezes é útil lembrar da fórmula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (\text{se } ad-bc \neq 0). \tag{1.5}$$

Capítulo 2

Espaços vetoriais

Neste capítulo, apresentaremos o conceito de espaço vetorial, assim como seus fundamentos essenciais: bases e dimensão. Basicamente, os espaços vetoriais, de dimensão finita - que são o foco desse curso, são um conjunto que se comportam como $\mathbb{R}^n$ (ou $\mathbb{C}^n$): conjunto de n-uplas de números reais (ou complexos), onde é possível fazer somas de seus elementos e multiplicação por escalar (no caso de n-uplas essas operações são feitas coordenada a coordenada).

Primeiramente apresentaremos a definição de espaço vetorial, e a partir disso toda a teoria será desenvolvida. É interessante acompanhar como todas as proposições e teoremas vão sendo construídos passo a passo a partir das definições, sem precisar apelar para resultados muito técnicos e com provas complicadas. Por esse motivo, recomendamos a leitura atenta das demonstrações, pois elas auxiliam no entendimento dos conceitos e também fornecem ferramentas para resolver os exercícios. Lembramos que denotaremos o corpo de escalares por $\mathbb{K}$, significando $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

A discussão apresentada nesse capítulo foi motivada pelos livros [Axl], [Coe] e [Nic].

2.1 Definição e propriedades básicas

Definição 2.1.1 Um *espaço vetorial sobre* $\mathbb{K}$, ou $\mathbb{K}$-*espaço vetorial* é um conjunto não vazio V (cujos elementos são chamados de *vetores*) munido de duas operações: uma chamada de *soma* ou *adição*, $+ : V \times V \to V$, que a cada par de vetores $u, v \in V$ associa um (único) novo vetor $u + v \in V$, e outra chamada de *multiplicação por escalar*, $\cdot : \mathbb{K} \times V \to V$, que a cada elemento $\alpha \in \mathbb{K}$ (chamado de *escalar*) e a cada vetor $v \in V$ associa um (único) vetor $\alpha \cdot v = \alpha v \in V$, que devem satisfazer as propriedades a seguir.

(i) $u + v = v + u$, para todos $u, v \in V$ (comutatividade da soma);

(ii) $(u + v) + w = u + (v + w)$, para todos $u, v, w \in V$ (associatividade da soma);

(iii) existe $0 \in V$ tal que $0 + u = u$, para todo $u \in V$ (vetor nulo);

(iv) para cada $v \in V$ existe $-v \in V$ tal que $v + (-v) = 0$ (vetor simétrico);

(v) $1 \cdot v = v$, para todo $v \in V$;

(vi) $\alpha(\beta u) = (\alpha\beta)u$, para todos $\alpha, \beta \in \mathbb{K}$ e $u \in V$ (associatividade da multiplicação por escalar);

(vii) $(\alpha + \beta)u = \alpha u + \beta u$, para todos $\alpha, \beta \in \mathbb{K}$ e $u \in V$ (distributividade 1);

(viii) $\alpha(u + v) = \alpha u + \alpha v$, para todos $\alpha \in \mathbb{K}$ e $u, v \in V$ (distributividade 2).

Observação 2.1.2 O símbolo "0" também é usado para denotar o número zero de $\mathbb{K}$. Isso não causará confusão, pois ficará claro pelo contexto a que a notação se refere.

Listaremos a seguir, as propriedades mais imediatas das operações de um espaço vetorial. As demonstrações seguem dos itens da Definição 2.1.1, sendo um bom exercício tentar prová-las sem olhar a demonstração apresentada.

Proposição 2.1.3 — Propriedades básicas. Seja V um $\mathbb{K}$-espaço vetorial.

(i) O vetor nulo é único.

(ii) O vetor simétrico de um vetor dado é único.

(iii) Para todo $v \in V$, tem-se $0 \cdot v = 0$.

(iv) Para todo $\alpha \in \mathbb{K}$, tem-se $\alpha \cdot 0 = 0$.

(v) Para todo $v \in V$, tem-se $(-1)v = -v$.

(vi) Se $\alpha \in \mathbb{K}$ e $v \in V$ são tais que $\alpha v = 0$, então $\alpha = 0$ ou $v = 0$.

Demonstração. **(i)** Suponha que 0_1 e 0_2 sejam dois vetores nulos, então

$$0_1 \overset{0_2 \text{ é v.n.}}{=} 0_1 + 0_2 \overset{0_1 \text{ é v.n.}}{=} 0_2.$$

(ii) Seja $v \in V$ suponha que v_1 e v_2 são vetores simétricos de v, então

$$v_1 = v_1 + 0 = v_1 + (v + v_2) = (v_1 + v) + v_2 = 0 + v_2 = v_2.$$

(iii) $0 \cdot v = (0+0)v = 0 \cdot v + 0 \cdot v \Rightarrow 0 = 0 \cdot v + (-0 \cdot v) = 0 \cdot v + 0 \cdot v + (-0 \cdot v) = 0 \cdot v + 0 = 0 \cdot v.$

(iv) $\alpha \cdot 0 = \alpha \cdot (0+0) = \alpha \cdot 0 + \alpha \cdot 0 \Rightarrow 0 = \alpha \cdot 0 + (-\alpha \cdot 0) = \alpha \cdot 0 + \alpha \cdot 0 + (-\alpha \cdot 0) = \alpha \cdot 0 + 0 = \alpha \cdot 0.$

(v) $v + (-1)v = 1 \cdot v + (-1)v = (1-1)v = 0 \cdot v = 0 \Rightarrow (-1)v = -v.$

(vi) Se $\alpha \neq 0$, então

$$\alpha v = 0 \Rightarrow \frac{1}{\alpha}(\alpha v) = \frac{1}{\alpha} \cdot 0 = 0 \Rightarrow v = 1 \cdot v = \left(\frac{1}{\alpha}\alpha\right) v = 0.$$

■

A seguir, serão listados alguns exemplos de espaços vetoriais. Fica como exercício para o leitor verificar que de fato eles satisfazem as oito condições da Definição 2.1.1.

■ **Exemplo 2.1.4 — Espaço nulo.** $V = \{0\}$, que consiste do espaço com um só elemento (obrigatoriamente o vetor nulo). ■

■ **Exemplo 2.1.5 — Espaço $\mathbb{K}^n$.** Dado $n \in \mathbb{N}$, o conjunto das n-uplas em $\mathbb{K}$,

$$\mathbb{K}^n = \{(a_1, \ldots, a_n) : a_i \in \mathbb{K}, i = 1, \ldots, n\}$$

é um $\mathbb{K}$-espaço vetorial com as seguintes operações:

- dados $u = (a_1, \dots, a_n)$ e $v = (b_1, \dots, b_n)$ em $\mathbb{K}^n$, $u + v = (a_1 + b_1, \dots, a_n + b_n)$;
- dados $u = (a_1, \dots, a_n) \in \mathbb{K}^n$ e $\alpha \in \mathbb{K}$, $\alpha u = (\alpha a_1, \dots, \alpha a_n)$.

Se $n = 1$, teremos o próprio $\mathbb{K}$ como $\mathbb{K}$-espaço vetorial.

Se $\mathbb{K} = \mathbb{R}$ e $n = 2$ ou 3, as operações com vetores em $\mathbb{R}^2$ e $\mathbb{R}^3$ podem ser visualizadas geometricamente.

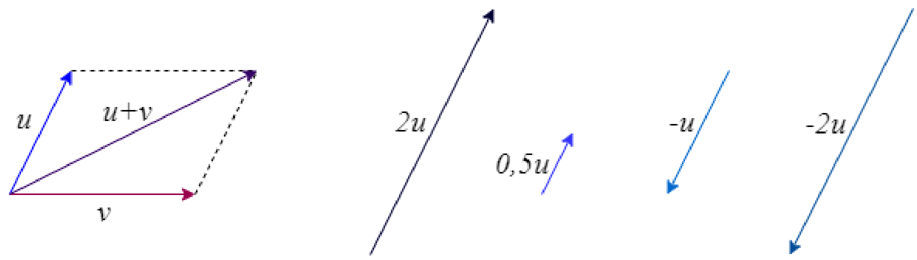


Figura 2.1: Vetores

$\mathbb{C}^n$ também é um $\mathbb{R}$-espaço vetorial com as operações definidas da mesma maneira. ■

■ **Exemplo 2.1.6 — Espaço de matrizes.** O conjunto $M_{m\times n}(\mathbb{K})$ formado pelas matrizes $m \times n$ com entradas em $\mathbb{K}$ é um $\mathbb{K}$-espaço vetorial com as operações usuais de soma de matrizes e multiplicação por escalar.

Quando $m = n$ denotaremos $M_{n\times n}(\mathbb{K})$ simplesmente por $M_n(\mathbb{K})$. ■

■ **Exemplo 2.1.7 — Espaço de polinômios.** O conjunto

$$P(\mathbb{K}) = \{p(x) = a_0 + a_1x + \cdots + a_nx^n : a_i \in \mathbb{K} \text{ e } n \in \mathbb{N}\}$$

é um $\mathbb{K}$-espaço vetorial com as operações definidas a seguir.

- Sejam $p(x) = a_0 + a_1x + \cdots + a_nx^n$ e $q(x) = b_0 + b_1x + \cdots + b_mx^m$ em $P(\mathbb{K})$. Sem perda de generalidade, suponha $m \leq n$. Então
$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + \cdots + a_nx^n.$$
- Dados $p(x) = a_0 + a_1x + \cdots + a_nx^n \in P(\mathbb{K})$ e $\alpha \in \mathbb{K}$,
$$\alpha p(x) = (\alpha a_0) + (\alpha a_1)x + \cdots + (\alpha a_n)x^n.$$

■

■ **Exemplo 2.1.8 — Espaço das soluções de um sistema linear homogêneo.** Dado um sistema linear homogêneo

$$L : \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

onde $a_{ij} \in \mathbb{K}$ para todos $i = 1, \ldots, m,\ j = 1, \ldots, n$, o conjunto formado por todas as soluções $S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \in M_{n\times 1}(\mathbb{K})$ de L é um $\mathbb{K}$-espaço vetorial, com a operações usuais de $M_{n\times 1}(\mathbb{K})$. ■

■ **Exemplo 2.1.9 — Espaço de funções.** Seja X um conjunto não vazio. O conjunto $\mathscr{F}(X,\mathbb{K})$ das funções $f : X \to \mathbb{K}$ é um $\mathbb{K}$-espaço vetorial com as operações definidas a seguir.

- Dados $f, g \in \mathscr{F}(X,\mathbb{K})$, a função $f + g : X \to \mathbb{K}$ é definida por $(f+g)(x) = f(x) + g(x)$ para todo $x \in X$.
- Dados $f \in \mathscr{F}(X,\mathbb{K})$ e $\alpha \in \mathbb{K}$, a função $\alpha f : X \to \mathbb{K}$ é definida por $(\alpha f)(x) = \alpha f(x)$ para todo $x \in X$.

Observe que a estrutura de $\mathbb{K}$-espaço vetorial de $\mathscr{F}(X,\mathbb{K})$ depende somente das operações em $\mathbb{K}$ (X é um conjunto qualquer).

Um caso particular importante, é quando $X = \mathbb{N}$. Neste caso $\mathscr{F}(\mathbb{N},\mathbb{K})$ é o *espaço de sequências*, e é mais comum denotar $f : \mathbb{N} \to \mathbb{K}$ por $(x_n)_{n\in\mathbb{N}}$, onde $x_n = f(n)$. Assim,

$$(x_n)_{n\in N} + (y_n)_{n\in N} = (x_n + y_n)_{n\in\mathbb{N}} \quad \text{e} \quad \alpha(x_n)_{n\in N} = (\alpha x_n)_{n\in N}.$$

■

■ **Exemplo 2.1.10 — Produto cartesiano.** Sejam U e V $\mathbb{K}$-espaços vetoriais. Então o produto cartesiano $U \times V = \{(u,v) : u \in U \text{ e } v \in V\}$ é um $\mathbb{K}$-espaço vetorial com as operações definidas coordenada a coordenada, isto é, dados $(u_1,v_1),(u_2,v_2) \in U \times V$ e $\alpha \in \mathbb{K}$,

$$(u_1,v_1) + (u_2,v_2) = (u_1+u_2, v_1+v_2) \quad \text{e} \quad \alpha(u_1,v_1) = (\alpha u_1, \alpha v_1).$$

■

2.2 Subespaços vetoriais, soma de subespaços, soma direta

A seguir apresentamos o conceito de subespaço vetorial, que basicamente é um espaço vetorial dentro de outro.

Definição 2.2.1 Seja V um $\mathbb{K}$-espaço vetorial. Um subconjunto não vazio W de V é um *subespaço vetorial* de V se a restrição das operações de V a W torna esse conjunto um $\mathbb{K}$-espaço vetorial.
Se $W \subseteq V$ é um subespaço e $W \neq V$, dizemos que W é um subespaço *próprio* de V.

Observe que, para W ser um espaço vetorial com as operações de V, precisa-se verificar as oito condições da Definição 2.1.1 para W. Mas todas, exceto (iii) e (iv), já serão automaticamente satisfeitas por V ser espaço vetorial e $W \subseteq V$. Sendo assim, resta verificar essas duas condições, e se de fato as restrições das operações a W ficam bem definidas em W, isto é, se a soma de dois vetores de W resulta em um vetor de W (e não de $V \setminus W$) e o mesmo para a multiplicação por escalar. O próximo resultado diz que na verdade precisamos mostrar um pouco menos que isso.

Proposição 2.2.2 Sejam V é um $\mathbb{K}$-espaço vetorial e $W \subseteq V$. Então W é um subespaço vetorial se, e somente se, as seguintes condições são satisfeitas:

(i) $0 \in W$;

(ii) se $v_1, v_2 \in W$, então $v_1 + v_2 \in W$;

(iii) se $v \in W$ e $\alpha \in \mathbb{K}$, então $\alpha v \in W$.

Demonstração. Suponha que W é subespaço vetorial. Então as condições (ii) e (iii) são satisfeitas pois as restrições devem se tornar operações bem definidas em W. Além disso deve existir um vetor nulo $0_W \in W$, e este por sua vez deve ser igual ao vetor nulo de V, pois caso contrário V teria dois vetores nulos, contrariando a unicidade. Portanto $0 \in W$.

Reciprocamente, suponha que as condições (i), (ii) e (iii) são satisfeitas. As condições (ii) e (iii) garantem que as as restrições das operações a W ficam bem definidas em W. A condição (i) garante que $W \neq \emptyset$ e que a condição (iii) da Definição 2.1.1 é satisfeita para W. Resta verificar que dado $v \in W$, tem-se $-v \in W$. Mas $-v = (-1) \cdot v \in W$ por (iii). ■

■ **Exemplo 2.2.3** Se V é um $\mathbb{K}$-espaço vetorial, então $\{0\}$ e V são subespaços vetoriais de V, chamados de *subespaços triviais*. ■

■ **Exemplo 2.2.4** Mostremos que $U = \{(x,y) \in \mathbb{R}^2 : x+y=0\}$ é subespaço vetorial de $\mathbb{R}^2$.
É suficiente verificar as três condições da Proposição 2.2.2.

(i) $0 = (0,0) \in U$ pois $0+0=0$.

(ii) Se $u = (x_1,y_1)$ e $v = (x_2,y_2) \in U$, então $x_1+y_1 = 0 = x_2+y_2$. Daí, $u+v = (x_1+x_2, y_1+y_2)$ satisfaz
$$(x_1+x_2)+(y_1+y_2) = (x_1+y_1)+(x_2+y_2) = 0+0 = 0.$$
Logo $u+v \in U$.

(iii) Se $u = (x,y) \in U$ e $\alpha \in \mathbb{K}$, então $\alpha u = (\alpha x, \alpha y)$ e $\alpha x + \alpha y = \alpha(x+y) = \alpha \cdot 0 = 0$. Logo $\alpha u \in U$.

Mais geralmente, dados $a, b \in \mathbb{R}$, o subconjunto $W = \{(x,y) \in \mathbb{R}^2 : ax+by=0\}$ é um subespaço vetorial de $\mathbb{R}^2$ (verifique). Geometricamente, W representa uma reta no plano, que passa pela origem. ■

Exercício 2.1 Mostre que, dados $a, b, c \in \mathbb{R}$, o conjunto $U = \{(x,y,z) \in \mathbb{R}^3 : ax+by+cz=0\}$ é subespaço vetorial de $\mathbb{R}^3$. Geometricamente, U representa um plano que passa pela origem.
Um plano em $\mathbb{R}^3$ que não passa pela origem é um subespaço vetorial?

■ **Exemplo 2.2.5** O subconjunto $U = \{(x,y) \in \mathbb{R}^2 : x \geq 0 \text{ e } y \geq 0\}$ não é um subespaço vetorial de $\mathbb{R}^2$, pois, por exemplo, $(1,1) \in U$ e $-1 \in \mathbb{R}$, mas $-1 \cdot (1,1) = (-1,-1) \notin U$. ■

Nos próximos exemplos, deixaremos a cargo do leitor a verificação de que os subconjuntos apresentados são, de fato, subespaços vetoriais.

■ **Exemplo 2.2.6** Dado $n \in \mathbb{N}$, considere $P_n(\mathbb{K}) = \{p(x) = a_0 + a_1 x + \cdots + a_n x^n : a_i \in \mathbb{K}\}$, isto é, o conjunto dos polinômios com coeficientes em $\mathbb{K}$ e grau no máximo n. Então, $P_n(\mathbb{K})$ é subespaço vetorial do espaço de polinômios $P(\mathbb{K})$ (Exemplo 2.1.7). ■

■ **Exemplo 2.2.7** Considere $U = \mathscr{C}([0,1],\mathbb{R})$ o conjunto das funções $f : [0,1] \to \mathbb{R}$ contínuas. Então, U é subespaço vetorial do espaço de funções $\mathscr{F}([0,1],\mathbb{R})$ (Exemplo 2.1.9). ■

■ **Exemplo 2.2.8** Seja $V = \mathscr{C}([0,1],\mathbb{R})$. O subconjunto $U = \left\{ f \in V : \int_0^1 f(x)\,dx = 0 \right\}$ é um subespaço vetorial de V. ■

■ **Exemplo 2.2.9** Seja $V = M_n(\mathbb{K})$. Os subconjuntos

$$W_1 = \{A \in V : A = A^\top\} \quad \text{e} \quad W_2 = \{A \in V : A = -A^\top\}$$

são subespaços vetoriais de V. As matrizes em W_1 são chamadas de *simétricas* e as matrizes em W_2 são chamadas de *antissimétricas*.

Se $\mathbb{K} = \mathbb{C}$, então o conjunto $U = \{A \in V : A = \overline{A}^\top\}$ é um $\mathbb{R}$-subespaço vetorial de $M_n(\mathbb{C})$ (visto como espaço sobre $\mathbb{R}$), onde $\overline{A}$ denota a matriz cujas entradas são os conjugados complexos das respectivas entradas em A. As matrizes em U são chamadas de *hermitianas*. Se $\mathbb{K} = \mathbb{R}$ o espaço U também poderia ser definido, no entanto ele seria exatamente W_1. ■

■ **Exemplo 2.2.10** Dados V um espaço vetorial e U,W subespaços vetoriais de V, então $U \cap W$ é um subespaço vetorial de V (verifique). Mais geralmente, pode-se provar que a interseção de qualquer família de subespaços, mesmo que uma quantidade infinita, também é sempre subespaço vetorial.

Em geral, $U \cup W$ não é um subespaço vetorial de V (encontre um contra-exemplo). ■

Definição 2.2.11 Sejam V um $\mathbb{K}$-espaço vetorial e $U,W \subseteq V$ subespaços.

(i) Dizemos a *soma* de U e W é o conjunto de todas as possíveis somas entre um elemento de U e um de W, isto é $U + W = \{u + w : u \in U,\ w \in W\}$.

(ii) Dizemos que a soma é *direta* quando $U \cap W = 0$. Neste caso, denotamos $U \oplus W$.

Observação 2.2.12 A soma de dois subespaços é sempre um espaço vetorial (verifique).

■ **Exemplo 2.2.13** Considere $V = M_n(\mathbb{K})$ e os subespaços $U = \{A \in V : A = A^\top\}$ e $W = \{A \in V : A = -A^\top\}$. Verifique que $V = U \oplus W$.

Solução. Para demonstrar a igualdade de conjuntos $V = U + W$, a rigor, precisamos verificar que $V \subseteq U + W$ e $U + W \subseteq V$. A última continência é direta, pois $U,W \subseteq V$ são subespaços. Precisamos verificar que $V \subseteq U + W$. Seja $A \in V$. Considere $B,C \in V$ definidos por

$$B = \frac{1}{2}(A + A^\top) \quad \text{e} \quad C = \frac{1}{2}(A - A^\top).$$

É fácil ver que $A = B + C$. Mostremos que $B \in U$ e $C \in W$.

$$B^\top = \left(\frac{1}{2}(A + A^\top)\right)^\top = \frac{1}{2}(A + A^\top)^\top = \frac{1}{2}(A^\top + (A^\top)^\top) = \frac{1}{2}(A^\top + A) = B \Rightarrow B \in U,$$

e,

$$C^\top = \left(\frac{1}{2}(A - A^\top)\right)^\top = \frac{1}{2}(A - A^\top)^\top = \frac{1}{2}(A^\top - (A^\top)^\top) = \frac{1}{2}(A^\top - A) = -C \Rightarrow C \in W.$$

Assim, todo elemento de V também é um elemento de $U+W$, donde segue que $V \subseteq U+W$. Portanto,

$$V = U+W.$$

Para concluir que a soma é direta, resta verificar que $U \cap W = \{0\}$. A inclusão $\{0\} \subseteq U \cap W$ é imediata, pois U e W são subespaços, logo $0 \in U, W$. Suponha $A \in U \cap W$. Então

$$A^\top = A = -A^\top \Rightarrow 2A^\top = 0 \Rightarrow A^\top = 0 \Rightarrow A = 0.$$

Logo, $U \cap W = \{0\}$. Portanto, $V = U \oplus W$. ■

Proposição 2.2.14 Sejam V um $\mathbb{K}$-espaço vetorial e $U, W \subseteq V$ subespaços. A soma $U+W$ é direta se, e somente se, cada elemento $v \in U+W$ pode ser escrito de modo único como uma soma $v = u+w$ com $u \in U$ e $w \in W$.

Demonstração. Suponha que a soma $U+W$ é direta. Então, $U \cap W = 0$. Dado $v \in U+W$, suponha que existam $u_1, u_2 \in U$ e $w_1, w_2 \in W$ tais que

$$v = u_1 + w_1 = u_2 + w_2.$$

Então,

$$u_1 - u_2 = w_1 - w_2.$$

Como U e W são subespaços, $u_1 - u_2 \in U$ e $w_2 - w_1 \in W$. Logo, $u_1 - u_2 = w_1 - w_2 \in U \cap W = \{0\}$. Portanto,

$$u_1 - u_2 = w_1 - w_2 = 0 \Rightarrow u_1 = u_2 \text{ e } w_1 = w_2.$$

Reciprocamente, suponha que cada elemento $v \in U+W$ pode ser escrito de modo único como uma soma $v = u+w$ com $u \in U$ e $w \in W$. Temos que mostrar que $U \cap W = \{0\}$. A inclusão $\{0\} \subseteq U \cap W$ é imediata, já que U e V são subespaços. Seja $v \in u \cap W$.

Por um lado, $v \in U$ e $0 \in W$ e assim $v = v+0$ é uma decomposição como soma de um elemento de U e um de W.

Por outro lado, $0 \in U$ e $v \in W$ e assim $v = 0+v$ é uma decomposição como soma de um elemento de U e um de W.

Pela unicidade da decomposição, deve ocorrer $v = 0$. Assim, provamos que $U \cap W \subseteq \{0\}$, donde segue que $U \cap W = \{0\}$. ■

Definição 2.2.15 Sejam $U_1, \dots, U_m$ subespaços de um espaço vetorial V. A *soma* de $U_1, \dots, U_m$, é o seguinte subespaço de V:

$$U_1 + \cdots + U_m = \{u_1 + \cdots + u_m : u_i \in U_i,\ i = 1, \dots, m\}.$$

Dizemos que a soma $U_1 + \cdots + U_m$ é *direta*, e denotamos $U_1 \oplus U_2 \oplus \cdots \oplus U_m$, se

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m) = \{0\} \text{ para todo } i = 1, 2, \dots, m.$$

Também é válido resultado análogo ao da Proposição 2.2.14, isto é, a soma $U_1 + \cdots + U_m$ é direta se, e somente se, todo elemento se escreve de maneira única como uma soma $u_1 + \cdots + u_m$, com $u_i \in U_i$.

2.3 Subespaço gerado

Seja V um $\mathbb{K}$-espaço vetorial. Uma *combinação linear* dos vetores $v_1, \ldots, v_n \in V$, é um vetor da forma

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n, \quad \text{com } \alpha_1, \ldots, \alpha_n \in \mathbb{K}.$$

Para ilustrar geometricamente o conceito de combinação linear, considere 2 vetores no plano $\mathbb{R}^2$. Por exemplo, sejam $u = (1,2)$ e $v = (-1,1)$. Um vetor $w \in \mathbb{R}^2$ é uma combinação linear de u e v se pode ser obtido "andando" um tanto na direção de u mais um tanto na direção de v.

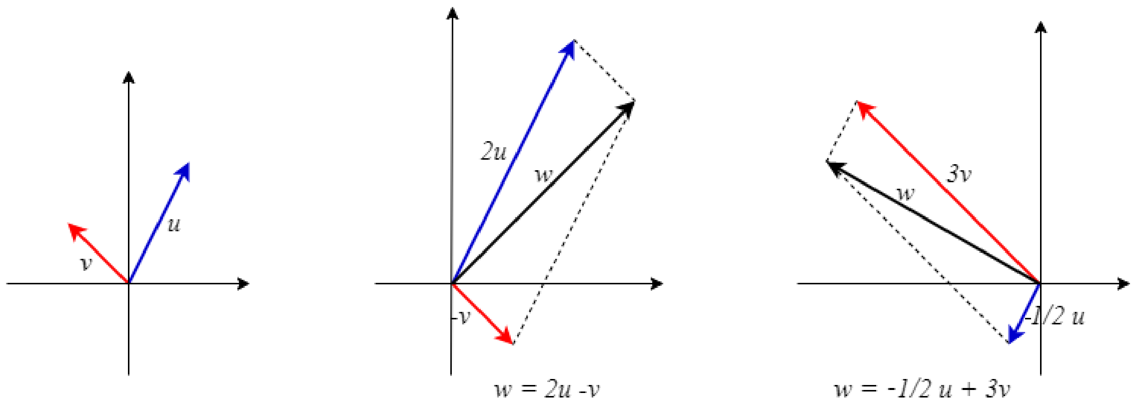


Figura 2.2: Combinação linear

É possível provar que qualquer vetor em $\mathbb{R}^2$ pode ser obtido como uma combinação linear de u e v (veja Exemplo 2.3.3). Esse é um exemplo onde $\{u,v\}$ *gera* $\mathbb{R}^2$ (o conceito de conjunto gerador será introduzido a seguir).

Agora, se $u = (1,1)$ e $v = (2,2)$ (ambos tem a mesma direção), todas as combinações lineares de u e v vão gerar apenas um reta, que passa pela origem e tem a direção de u (e também de v).

Definição 2.3.1 Sejam V um $\mathbb{K}$-espaço vetorial e S um subconjunto de V. Define-se o *subespaço de V gerado por S* por

$$[S] = \{\alpha_1 v_1 + \cdots + \alpha_n v_n : n \in \mathbb{N},\ v_i \in S,\ \alpha_i \in \mathbb{K}\},$$

ou seja, o subconjunto formado por todas as combinações lineares envolvendo vetores de S. Também, dizemos que S é um *conjunto gerador* de $[S]$ ou que S *gera* $[S]$.

Observação 2.3.2
- Mesmo que S seja um conjunto infinito, as combinações lineares em $[S]$ são sempre somas finitas!
- Por convenção, $\{0\}$ é gerado pelo conjunto vazio.
- $[S]$ é sempre um subespaço vetorial de V, mesmo que S não o seja (verifique).

■ **Exemplo 2.3.3** Seja $S = \{(1,2), (-1,1)\} \subseteq \mathbb{R}^2$. Verifique que $[S] = \mathbb{R}^2$.

Solução. A inclusão $[S] \subseteq \mathbb{R}^2$ é imediata. Mostremos que $\mathbb{R}^2 \subseteq [S]$. Seja $v = (a,b) \in \mathbb{R}^2$. Precisamos encontrar $\alpha, \beta \in \mathbb{R}$ tais que

$$v = \alpha(1,2) + \beta(-1,1) \Leftrightarrow (a,b) = (\alpha, 2\alpha) + (-\beta, \beta) = (\alpha - \beta, 2\alpha + \beta).$$

Ou seja, precisamos encontrar uma solução do sistema linear

$$\begin{cases} \alpha - \beta = a \\ 2\alpha + \beta = b \end{cases}.$$

Resolvendo o sistema, temos a solução $\alpha = \dfrac{a+b}{3}$ e $\beta = \dfrac{b-2a}{3}$. Portanto, $v \in [S]$, donde concluímos que $[S] = \mathbb{R}^2$. ■

Exercício 2.2 Seja $S = \{(1,1),(2,2)\} \subseteq \mathbb{R}^2$. Verifique que $[S] = \{(a,a) : a \in \mathbb{R}\}$.

■ **Exemplo 2.3.4** Seja $S = \{(1,0,0),(0,1,0),(0,0,1)\} \subseteq \mathbb{R}^3$. Verifique que $[S] = \mathbb{R}^3$.

Solução. A inclusão $[S] \subseteq \mathbb{R}^3$ é imediata. Para provar $\mathbb{R}^3 \subseteq [S]$, considere $v = (a,b,c) \in \mathbb{R}^3$. Temos,

$$v = a(1,0,0) + b(0,1,0) + c(0,0,1) \in [S].$$

Portanto, $\mathbb{R}^3 \subseteq [S]$, donde segue que $[S] = \mathbb{R}^3$. ■

■ **Exemplo 2.3.5** Considere o $\mathbb{R}$-espaço vetorial $V = P_2(\mathbb{R})$. Verifique quais dos conjuntos

$$S_1 = \{1, x, x^2\}, \ \ S_2 = \{2, 1+x, 1+x^2\}, \ \ S_3 = \{1+x+x^2, 2-x+3x^2, 4+x+5x^2\}$$

geram V.

Solução. Dado $p(x) = a + bx + cx^2 \in V$, é fácil ver que

$$p(x) = a \cdot 1 + b \cdot x + c \cdot x^2 \in [S_1].$$

Daí, não é difícil concluir que $[S_1] = V$.

Também, pode-se observar que

$$p(x) = \left(\frac{a}{2} - \frac{b}{2} - \frac{c}{2}\right) \cdot 2 + b \cdot (1+x) + c \cdot (1+x^2) \in [S_2],$$

e concluir que $[S_2] = V$.

No caso de S_3, se tentássemos escrever

$$p(x) = \alpha(1+x+x^2) + \beta(2-x+3x^2) + \gamma(4+x+5x^2) = (\alpha+2\beta+4\gamma) + (\alpha-\beta+\gamma)x + (\alpha+3\beta+5\gamma)x^2,$$

cairíamos no sistema linear

$$\begin{cases} \alpha + 2\beta + 4\gamma = a \\ \alpha - \beta + \gamma = b \\ \alpha + 3\beta + 5\gamma = c \end{cases}.$$

que não tem solução para a,b,c quaisquer, pois a matriz do sistema

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & -1 & 1 \\ 1 & 3 & 5 \end{bmatrix}$$

é singular! Assim, já concluímos que S_3 não gera V. Neste caso, $[S_3]$ é um subespaço vetorial próprio de V. Observe que

$$4 + x + 5x^2 = 2 \cdot (1+x+x^2) + 1 \cdot (2-x+3x^2) \in [\{1+x+x^2, 2-x+3x^2\}].$$

Assim,

$$[S_3] = [\{1+x+x^2, 2-x+3x^2\}].$$

■

O exemplo a seguir será importante mais adiante.

■ **Exemplo 2.3.6 — Espaço coluna.** Seja $A \in M_{m \times n}(\mathbb{K})$. Se $A_1, \ldots, A_n \in \mathbb{K}^m$ denotam as colunas de A, o subespaço de $\mathbb{K}^m$ gerado pelas colunas de A, é chamado de *espaço coluna de* A e será denotado por $R(A)$. Assim,

$$R(A) = [A_1, A_2, \ldots, A_n].$$

Segue da Proposição 1.1.2(**xiv**),

$$R(A) = \{v \in \mathbb{R}^m : v = Au \text{ para algum } u \in \mathbb{R}^n\}.$$

■

2.4 Dependência linear

Definição 2.4.1 Sejam $v_1, \ldots, v_m$ vetores distintos de um $\mathbb{K}$-espaço vetorial V. Dizemos que $v_1, \ldots, v_m$ são *linearmente dependente* (LD) se existem $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ não todos nulos tais que

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0.$$

Se $v_1, \ldots, v_m$ não são LD, dizemos que eles são *linearmente independentes* (LI).

Observe que, $v_1, \ldots, v_m$ são LI se, e somente se, dada uma combinação linear linear nula

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0,$$

necessariamente ocorre $\lambda_1 = \cdots = \lambda_m = 0$, isto é, o vetor nulo só pode ser escrito de uma única maneira como combinação linear de $v_1, \ldots, v_m$.

Definição 2.4.2 Seja S um subconjunto (finito ou infinito) de $\mathbb{K}$-espaço vetorial V. Dizemos que S é *linearmente dependente* (LD) se existem $v_1, \ldots, v_m \in S$ distintos e $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ não todos nulos tais que

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0.$$

Lembre que combinações lineares são sempre somas finitas! Caso contrário, S é *linearmente independentes* (LI).

Observação 2.4.3 Por convenção, o conjunto vazio é LI.

■ **Exemplo 2.4.4** O subconjunto $S = \{(1,1), (0,2)\} \subseteq \mathbb{R}^2$ é LI ou LD?

Solução. Suponha $\alpha, \beta \in \mathbb{R}$ tais que

$$\alpha(1,1) + \beta(0,2) = (0,0).$$

Então,

$$(\alpha, \alpha + 2\beta) = (0,0) \Leftrightarrow \alpha = 0 \text{ e } \alpha + \beta = 0 \Leftrightarrow \alpha = \beta = 0.$$

Portanto, a única combinação linear nula de vetores em S é a trivial. Logo S é LI. ■

■ **Exemplo 2.4.5** O subconjunto $S = \{(1,1,0), (0,1,0), (2,1,0)\} \subseteq \mathbb{R}^3$ é LI ou LD?

Solução. Suponha $\alpha, \beta, \gamma \in \mathbb{R}$ tais que

$$\alpha(1,1,0) + \beta(0,1,0) + \gamma(2,1,0) = (0,0,0).$$

Então,

$$(\alpha+2\gamma, \alpha+\beta+\gamma, 0) = (0,0,0) \Leftrightarrow \begin{cases} \alpha+2\beta = 0 \\ \alpha+\beta+\gamma = 0 \end{cases}.$$

Observe que esse sistema linear possui solução não trivial. Por exemplo, se $\alpha = -2$ e $\beta = \gamma = 1$, temos uma combinação linear nula

$$-2\cdot(1,1,0) + 1\cdot(0,1,0) + 1\cdot(2,1,0) = (0,0,0)$$

com coeficientes não todos nulos. Portanto, S é LD. ■

■ **Exemplo 2.4.6** Considere $S = \{(1,0),(i,0),(0,1),(0,i)\} \subseteq \mathbb{C}^2$. Vamos analisar se S é LI ou LD, considerando as duas possibilidades para $\mathbb{C}^2$: espaço vetorial sobre $\mathbb{R}$ e $\mathbb{C}$.

Como $\mathbb{C}$-espaço vetorial: observe que

$$1\cdot(1,0) + i\cdot(i,0) + 0\cdot(0,1) + 0\cdot(0,i) = (0,0)$$

é uma combinação linear nula de vetores em S, com coeficientes não todos nulos. Portanto, S é LD.

Como $\mathbb{R}$-espaço vetorial: suponha

$$\alpha_1(1,0) + \alpha_2(i,0) + \alpha_3(0,1) + \alpha_4(0,i) = (0,0),\ \alpha_i \in \mathbb{R}.$$

Então,

$$(\alpha_1 + i\alpha_2, \alpha_3 + i\alpha_4) = (0,0) \Leftrightarrow \alpha_1 + i\alpha_2 = 0 \text{ e } \alpha_3 + i\alpha_4 = 0 \overset{\alpha_i \in \mathbb{R}}{\Leftrightarrow} \alpha_i = 0\ \forall\, i = 1,2,3,4.$$

Portanto, S é LI. ■

■ **Exemplo 2.4.7** O subconjunto $S = \{\cos x, e^x\} \subseteq \mathscr{F}(\mathbb{R},\mathbb{R})$ é LI ou LD?

Solução. Suponha $\alpha, \beta \in \mathbb{R}$ tais que

$$\alpha \cos x + \beta e^x = 0 \quad \text{(função nula)}. \tag{2.1}$$

Como estamos no espaço das funções de $\mathbb{R}$ em $\mathbb{R}$, essa combinação linear igual a função nula implica que, para todo $x \in \mathbb{R}$, vale a igualdade (2.1). Em particular, para $x = \pi/2$, temos

$$\beta e^{\pi/2} = \alpha\cdot 0 + \beta e^{\pi/2} = 0 \Rightarrow \beta = 0.$$

Substituindo $\beta = 0$ em (2.1), resta $\alpha \cos x = 0$. Agora, considerando $x = 0$, obtemos $\alpha = 0$. Portanto, a única possibilidade para (2.1) é $\alpha = \beta = 0$. Logo, S é LI. ■

■ **Exemplo 2.4.8** O subconjunto $S = \{1, \cos(2x), \cos^2 x\} \subseteq \mathscr{F}(\mathbb{R},\mathbb{R})$ é LI ou LD?

Solução. Como

$$\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos(2x) \quad \forall\, x \in \mathbb{R},$$

temos que

$$\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot\cos(2x) - 1\cdot\cos^2 x = 0$$

é uma combinação linear nula de vetores em S, com coeficientes não todos nulos. Portanto, S é LD. ■

Exemplo 2.4.9 O subconjunto $S = \{1, x, x^2, x^3, \ldots\} \subseteq P(\mathbb{K})$ é LI ou LD?

Solução. Uma combinação linear nula de vetores em S:

$$\alpha_0 \cdot 1 + \alpha_1 x + \cdots + \alpha_n x^n = 0$$

nada mais é que um polinômio identicamente nulo. Isso só ocorre se $\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0$. Portanto, S é LI. ■

Proposição 2.4.10 — Propriedades básicas sobre dependência linear. Sejam V um $\mathbb{K}$-espaço vetorial e $S \subseteq V$.

(i) Suponha $S = \{v\}$. Então S é LI se $v \neq 0$ e LD se $v = 0$.

(ii) Se S contém um subconjunto LD, então S é LD. Em particular, se $0 \in S$, então S é LD.

(iii) Se S é LI, então todo subconjunto S também é LI.

(iv) Suponha que S possua 2 ou mais vetores. Então S é LD se, e somente se, existe $v \in S$ e $v_1, \ldots, v_n \in S \backslash \{v\}$ tais que $v \in [v_1, \ldots, v_n]$.

(v) Se S é LI e $v \in V \setminus S$ é tal que $S \cup \{v\}$ é LD, então $v \in [S]$.

Demonstração. Provaremos apenas o item **(iv)**. As demonstrações dos demais itens são mais simples e serão deixadas como exercício.

Suponha que S é LD. Então existem $v_1, \ldots, v_m \in S$ distintos e $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ não todos nulos tais que

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0. \tag{2.2}$$

Se $m = 1$, então deve ocorrer $v_1 = 0$. Assim, qualquer que seja $v \in S \setminus \{0\}$, tem-se $0 \in [v]$. Suponha agora $m \geq 2$. Como os escalares em (2.2) são não todos nulos, vamos assumir, sem perda de generalidade, que $\alpha_1 \neq 0$. Assim,

$$v_1 = \left(-\frac{\alpha_2}{\alpha_1}\right) v_2 + \cdots + \left(-\frac{\alpha_m}{\alpha_1}\right) v_m \in [v_2, \ldots, v_m].$$

Reciprocamente, suponha que $v \in [v_1, \ldots, v_n]$, com $v \in S$ e $v_1, \ldots, v_n \in S \backslash \{v\}$. Então, existem $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \;\Leftrightarrow\; (-1)v + \alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

A última equação é uma combinação linear nula de vetores de S com ao menos um dos escalares não nulo. Portanto, S é LD. ■

Observação 2.4.11 Como caso particular da proposição anterior, dois vetores u e v são LD, se, e somente se, existe $\alpha \in \mathbb{K}$ tal que $u = \alpha v$ ou $v = \alpha u$ (um deles é um múltiplo escalar do outro).

2.5 Base e dimensão

Definição 2.5.1 Seja V um espaço vetorial. Dizemos que um subconjunto $B \subseteq V$ é uma *base* de V se $[B] = V$ e B é LI.

Observação 2.5.2 Por convenção, assumiremos que $\emptyset$ é base do espaço nulo $\{0\}$.

■ **Exemplo 2.5.3** Seja $B = \{(1,2),(-1,1)\} \subseteq \mathbb{R}^2$.

Foi mostrado no Exemplo 2.3.3 que $[B] = \mathbb{R}^2$. É fácil ver que B é LI, já que nenhum dos dois vetores é múltiplo escalar do outro. Portanto, B é uma base de $\mathbb{R}^2$. ■

■ **Exemplo 2.5.4** Seja $B = \{(1,1),(0,2)\} \subseteq \mathbb{R}^2$.

Foi mostrado no Exemplo 2.4.4 que B é LI. Não é difícil verificar que $[B] = \mathbb{R}^2$, uma vez que, dado $(a,b) \in \mathbb{R}^2$, tem-se

$$(a,b) = a(1,1) + \frac{b-a}{2}(0,2).$$

Logo, B é uma base de $\mathbb{R}^2$. ■

A seguir, apresentaremos as chamadas *bases canônicas* de alguns espaços vetoriais que serão trabalhados neste curso. Essas bases levam esse título, por serem facilmente descritas e por ser mais "simples" escrever um vetor qualquer no respectivo espaço como combinação linear dos vetores da base canônica.

■ **Exemplo 2.5.5 — Base canônica de $\mathbb{K}^n$.** Considere o $\mathbb{K}$-espaço vetorial $\mathbb{K}^n$. Para cada $j = 1,\dots,n$, seja $e_j \in \mathbb{K}^n$ o vetor cuja j-ésima coordenada é igual a 1 e as demais são nulas:

$$e_1 = (1,0,\dots,0),\ e_2 = (0,1,0\dots,0),\ \dots\ ,\ e_n = (0,\dots,0,1).$$

O subconjunto $B = \{e_1,e_2,\dots,e_n\} \subseteq \mathbb{K}^n$ é uma base, chamada de *base canônica*. ■

■ **Exemplo 2.5.6 — Base canônica de $M_{m\times n}(\mathbb{K})$.** Considere o $\mathbb{K}$-espaço vetorial $M_{m\times n}(\mathbb{K})$. Para cada $k = 1,\dots,m$ e $l = 1,\dots,n$, seja $E_{kl} = (a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{K})$ a matriz tal que $a_{kl} = 1$ e $a_{ij} = 0$ se $(i,j) \neq (k,l)$. O subconjunto

$$B = \{E_{kl} : k = 1,\dots,m;\ l = 1,\dots,n\}$$

é uma base, chamada de *base canônica* de $M_{m\times n}(\mathbb{K})$.

Por exemplo, se $m = 2$ e $n = 3$, temos

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix},$$

$$E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},\ E_{31} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix},\ E_{32} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$

e o conjunto $B = \{E_{11},E_{12},E_{21},E_{22},E_{31},E_{32}\}$ é a base canônica de $M_{3\times 2}(\mathbb{K})$. ■

■ **Exemplo 2.5.7 — Base canônica de $P_n(\mathbb{K})$ e $P(\mathbb{K})$.** Considere o $\mathbb{K}$-espaço vetorial $P_n(\mathbb{K})$. O subconjunto

$$B = \{1,x,x^2,\dots,x^n\}$$

é uma base, chamada de *base canônica* de $P_n(\mathbb{K})$.

Considere agora $\mathbb{K}$-espaço vetorial $P(\mathbb{K})$. No Exemplo 2.4.9, foi mostrado que o subconjunto

$$B = \{1,x,x^2,x^3,\dots\}$$

é LI. Não é difícil verificar que B gera $P(\mathbb{K})$, e portanto é uma base, chamada de *base canônica* de $P(\mathbb{K})$. ■

Observação 2.5.8 Não existe uma definição geral de base canônica para um espaço vetorial qualquer.

Definição 2.5.9 Dizemos que um $\mathbb{K}$-espaço vetorial é *finitamente gerado*, se existe um subconjunto $S \subseteq V$ finito que gera V.

Exercício 2.3 Verifique que o espaço $P(\mathbb{K})$ não é finitamente gerado.

Teorema 2.5.10 Todo espaço vetorial tem base.

A demonstração do Teorema 2.5.10 em sua versão mais geral, utiliza o chamado Lema de Zorn e é um pouco mais técnica. Faremos aqui a demonstração da existência de base para espaços finitamente gerados. Precisaremos dos resultados apresentados a seguir.

Lema 2.5.11 Sejam V um $\mathbb{K}$-espaço vetorial e $S = \{v_1, \dots, v_n\} \subseteq V$. Se algum $v_i \in [S \setminus \{v_i\}]$, isto é, v_i é combinação linear dos outros vetores em S, então $[S] = [S \setminus \{v_i\}]$.

Demonstração. A inclusão $[S \setminus \{v_i\}] \subseteq [S]$ é imediata. Mostremos que $[S] \subseteq [S \setminus \{v_i\}]$. Sem perda de generalidade, assuma $i = n$. Dado $u \in [S]$, existem $\alpha_1, \dots, \alpha_{n-1}, \alpha_n \in \mathbb{K}$ tais que

$$u = \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1} + \alpha_n v_n.$$

Como, por hipótese, $v_n \in [S \setminus \{v_n\}]$, existem $\beta_1, \dots, \beta_{n-1} \in \mathbb{K}$ tais que

$$v_n = \beta_1 v_1 + \cdots + \beta_{n-1} v_{n-1}.$$

Portanto,

$$u = \alpha_1 v_1 + \cdots + \alpha_{n-1} v_{n-1} + \alpha_n(\beta_1 v_1 + \cdots + \beta_{n-1} v_{n-1}) = (\alpha_1 + \beta_1) v_1 + \cdots + (\alpha_{n-1} + \beta_{n-1}) v_{n-1}$$

que é uma combinação linear dos vetores de $S \setminus \{v_n\}$. Portanto, $u \in [S \setminus \{v_n\}]$, donde segue que $[S] = [S \setminus \{v_n\}]$. ■

Proposição 2.5.12 Sejam V um $\mathbb{K}$-espaço vetorial e $S = \{v_1, \dots, v_n\} \subseteq V$ um conjunto gerador de V, com $n \geq 1$. Então, existe $B \subseteq S$ tal que B é uma base de V.

Demonstração. Suponha $n = 1$, donde $S = \{v_1\}$. Se $v_1 = 0$, então $V = [v_1] = \{0\}$ e $B = \emptyset \subseteq S$ é uma base de V. Se $v_1 \neq 0$, então S é LI, e portanto $B = S$ é uma base de V.

Assuma $n > 1$. Em particular, existe $v_i \neq 0$ e S contém um subconjunto LI não vazio. A ideia que usaremos a seguir, é tomar um subconjunto LI em S com o maior número de vetores possível, eliminando do conjunto gerador os vetores que são combinações lineares dos demais.

Se S é LI, então $B = S$ é base de V.

Se S não é LI, então é LD, logo existem $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ não todos nulos tais que

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

Sem perda de generalidade, suponha $\alpha_n \neq 0$. Logo,

$$v_n = \left(-\frac{\alpha_1}{\alpha_n}\right) v_1 + \cdots + \left(-\frac{\alpha_{n-1}}{\alpha_n}\right) v_{n-1} \in [S \setminus \{v_n\}].$$

Pelo Lema 2.5.11, $V = [S] = [S \setminus \{v_n\}] = [\{v_1, \ldots, v_{n-1}\}]$. Se $\{v_1, \ldots, v_{n-1}\}$ é LI, então $B = \{v_1, \ldots, v_{n-1}\}$ é base de V.

Caso contrário, $S' = \{v_1, \ldots, v_{n-1}\}$ é LD, e então repetimos o processo anterior, a fim obtermos um conjunto gerador contido em S' com $n-2$ vetores. Como S é finito, esse processo termina quando encontrarmos $B \subseteq S$ LI, de modo que $[B] = [S'] = V$, assim obtendo uma base B de V contida em S. ■

A demonstração do Teorema 2.5.10, para um espaço vetorial finitamente gerado, agora é imediata: tome um conjunto gerador finito e aplique a Proposição 2.5.12.

Proposição 2.5.13 Seja $B = \{v_1, v_2, \ldots, v_m\}$ uma base de um espaço vetorial V. Então, todo subconjunto LI de V tem no máximo m vetores.

Demonstração. Vamos mostrar que todo conjunto com mais de m vetores é LD.

Seja $S = \{u_1, u_2, \ldots, u_n\} \subseteq V$ com $n > m$. Como B é base de V, para cada $j = 1, \ldots, n$, existem $\alpha_{ij} \in \mathbb{K}$ tais que

$$u_j = \alpha_{1j} v_1 + \cdots + \alpha_{mj} v_m = \sum_{i=1}^{m} \alpha_{ij} v_i. \tag{2.3}$$

Para provar que S é LD, devemos verificar que existem escalares $\lambda_1, \ldots, \lambda_n$ não todos nulos tais que

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = 0.$$

Substituindo (2.3), temos

$$\lambda_1 \left(\sum_{i=1}^{m} \alpha_{i1} v_i \right) + \cdots + \lambda_n \left(\sum_{i=1}^{m} \alpha_{in} v_i \right) = 0 \Leftrightarrow$$

$$(\lambda_1 \alpha_{11} + \cdots + \lambda_n \alpha_{1n}) v_1 + \cdots + (\lambda_1 \alpha_{m1} + \cdots + \lambda_n \alpha_{mn}) v_m = 0.$$

Como B é base, B é LI. Logo, devemos ter

$$\left\{ \begin{array}{c} \lambda_1 \alpha_{11} + \cdots + \lambda_n \alpha_{1n} = 0 \\ \vdots \\ \lambda_1 \alpha_{m1} + \cdots + \lambda_n \alpha_{mn} = 0 \end{array} \right. .$$

Procuramos uma solução não trivial para esse sistema linear, nas incógnitas $\lambda_1, \ldots, \lambda_n$. Como $m < n$, o sistema possui mais variáveis do que equações. Assim, pelo Teorema 1.3.2, esse sistema possui infinitas soluções, o que garante que S é LD. ■

Corolário 2.5.14 Se V é um espaço vetorial finitamente gerado, então toda base de V possui o mesmo número de vetores.

Demonstração. Como V é finitamente gerado, existe conjunto gerador finito, e este contém uma base pela Proposição 2.5.12. Assim, a Proposição 2.5.13 garante que toda base será um conjunto finito.

Suponha agora que $B_1 = \{v_1, \ldots, v_m\}$ e $B_2 = \{u_1, \ldots, u_n\}$ são duas bases de V. Como B_1 é base, e B_2 é LI, pela Proposição 2.5.13, $n \leq m$. Por outro lado, B_2 é base e B_1 é LI, assim a Proposição 2.5.13 implica $m \leq n$. Portanto, $m = n$. ■

Observação 2.5.15 Se um espaço vetorial V não é finitamente gerado, então toda base de V é um conjunto infinito. Neste caso, é possível provar que duas bases quaisquer têm sempre a mesma cardinalidade (o que não será feito aqui).

Definição 2.5.16 Seja V um $\mathbb{K}$-espaço vetorial. Se V é finitamente gerado, define-se a *dimensão* de V como o número de elementos uma base qualquer de V. Caso contrário, diremos apenas que a dimensão de V é *infinita*.

Denotaremos a dimensão de V por $\dim_{\mathbb{K}} V$, ou simplesmente $\dim V$, quando não houver perigo de confusão sobre o corpo de escalares considerado.

■ **Exemplo 2.5.17** $\dim\{0\} = 0$, pois $\emptyset$ é base de $\{0\}$. ■

■ **Exemplo 2.5.18** $\dim \mathbb{R}^n = n$ e $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$. Por exemplo, a base canônica (Exemplo 2.5.5) tem n vetores. ■

■ **Exemplo 2.5.19** Considere $\mathbb{C}^n$ como $\mathbb{R}$-espaço vetorial. Então, apenas os n vetores e_j da base canônica no Exemplo 2.5.5 não geram todo o espaço $\mathbb{C}^n$ como $\mathbb{R}$-espaço vetorial (geram apenas vetores com coordenadas reais). É preciso acrescentar mais vetores que vão gerar as entradas complexas.

Seja $f_j \in \mathbb{C}^n$ o vetor cuja j-ésima coordenada é igual ao número complexo i (tal que $i^2 = -1$) e as demais são nulas, para $j = 1, \ldots, n$. Deixaremos como exercício a prova de que

$$B = \{e_1, \ldots, e_n, f_1, \ldots, f_n\} \subseteq \mathbb{C}^n$$

é uma base do $\mathbb{R}$-espaço vetorial $\mathbb{C}^n$. Portanto, $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$. ■

■ **Exemplo 2.5.20** $\dim M_{m\times n}(\mathbb{K}) = mn$. Por exemplo, a base canônica (Exemplo 2.5.6) possui mn vetores. ■

■ **Exemplo 2.5.21** $\dim P_n(\mathbb{K}) = n+1$. Por exemplo, a base canônica (Exemplo 2.5.7) tem $n+1$ vetores. Já o espaço $P(\mathbb{K})$ tem dimensão infinita. ■

■ **Exemplo 2.5.22** Determine a dimensão do seguinte subespaço de $\mathbb{R}^3$:

$$U = \{(x,y,z) \in \mathbb{R}^3 : x + 2y - z = 0\}.$$

Solução. Todo vetor $(x,y,z) \in U$ satisfaz

$$x + 2y - z = 0 \Leftrightarrow z = x + 2y \Leftrightarrow (x,y,z) = (x,y,x+2y) = x(1,0,1) + y(0,1,2) \in [(1,0,1),(0,1,2)].$$

Portanto, $(x,y,z) \in U$ se, e somente se, $(x,y,z) \in [(1,0,1),(0,1,2)]$. Donde segue que

$$U = [(1,0,1),(0,1,2)].$$

Assim, $B = \{(1,0,1),(0,1,2)\}$ gera U.

Vamos verificar se B é LI. Como B tem 2 vetores, pela Observação 2.4.11, B é LD se, e somente se, existe $\alpha \in \mathbb{R}$ tal que

$$(1,0,1) = \alpha(0,1,2) = (0,\alpha,2\alpha) \quad \text{ou} \quad (0,1,2) = \alpha(1,0,1) = (\alpha,0,\alpha).$$

É fácil ver que nenhuma das duas opções acima pode ocorrer, portanto B é LI, logo é uma base de U. Então, $\dim U = 2$. ■

Assim como todo conjunto gerador contém uma base, também vale que todo conjunto LI está contido numa base, como descreve a proposição a seguir.

Proposição 2.5.23 Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$. Então, qualquer subconjunto LI $\{v_1, \ldots, v_k\} \subseteq V$, com $k < n$, pode ser completado para formar uma base para V, isto é, existem $n-k$ vetores $v_{k+1}, \ldots, v_n \in V$ tais que $B = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ é uma base de V.

Demonstração. Considere $C = \{u_1, \ldots, u_n\}$ uma base de V e $S = B \cup C = \{v_1, \ldots, v_k, u_1, \ldots, u_n\}$. Seja B um subconjunto de S que contém $\{v_1, \ldots, v_k\}$ e é LI, com o maior número possível de vetores. Observe que, como S é finito, é possível obter B: basta ir testando dentre todos os subconjuntos de S contendo $\{v_1, \ldots, v_k\}$ quais são LI e tomar um com o maior número de vetores.

Sem perda de generalidade, suponha

$$B = \{v_1, \ldots, v_k, u_1, \ldots, u_m\}.$$

Pela forma que determinamos B, para todo $j > m$, temos que $B \cup \{v_j\}$ é LD. Como B é LI, pela Proposição 2.4.10(v), $v_j \in [B]$, para todo $j > m$. Assim, $[B] = [S]$, e como S contém a base C de V, concluímos que $[B] = [S] = V$. Portanto, B é base de V.

Segue do Corolário 2.5.14 que $k + m = n$, donde poderíamos simplesmente denotar

$$\{u_1, \ldots, u_m\} = \{v_{k+1}, \ldots, v_n\}.$$

■

Proposição 2.5.24 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e $S \subseteq V$. Então, são válidas as seguintes afirmações.

(i) Se $\#S > n$, então S é LD.

(ii) Se $\#S < n$, então $[S] \neq V$.

(iii) Se $\#S = n$ e S é LI, então S é uma base de V.

(iv) Se $\#S = n$ e S gera V, então S é base de V.

(v) S é base de V se, e somente se, todo $v \in V$ é uma combinação linear de vetores de S e os escalares dessa combinação linear são únicos.

Demonstração. **(i)** Segue da Proposição 2.5.13.

(ii) Suponha, por absurdo, que $\#S = k < n$ e $[S] = V$. Então, pela Proposição 2.5.12, existe $B \subseteq S$ base de V, com $\#B \leq k < n$, o que contradiz o Corolário 2.5.14.

(iii) Pela Proposição 2.5.23, existe B base de V tal que $S \subseteq B$. Por ser subconjunto, $n = \#S \leq \#B$. Como toda base de V possui n vetores, deve ocorrer $\#B = n = \#S$. Assim,

$$S \subseteq B \ \text{ e } \ \#S = \#B \ \Rightarrow \ B = S \ \text{ é base de } V.$$

(iv) Pela Proposição 2.5.12, existe $B \subseteq S$ base de V. Por ser subconjunto, $\#B \leq \#S = n$. Como toda base de V possui n vetores, deve ocorrer $\#B = n = \#S$. Assim,

$$B \subseteq S \ \text{ e } \ \#B = \#S \ \Rightarrow \ B = S \ \text{ é base de } V.$$

(v) Suponha S base de V. Então, $\#S = n$, S gera V e S é LI. Seja $S = \{v_1, \dots, v_n\}$. Dado $v \in V$, segue que v pode ser escrito como combinação linear de vetores de S, pois S gera V. Provemos que os escalares são únicos. Sejam $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{K}$ tais que

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Então,

$$(\alpha_1 - \beta_1)v_1 + \cdots + (\alpha_n - \beta_n)v_n = 0 \overset{S \text{ é LI}}{\Rightarrow} \alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0 \Rightarrow \alpha_1 = \beta_1, \dots, \alpha_n = \beta_n.$$

Reciprocamente, suponha que todo $v \in V$ é uma combinação linear de vetores de S e os escalares dessa combinação linear são únicos. Pela primeira parte da frase anterior, temos que S gera V. Provemos que S é LI. Suponha que $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ e $v_1, \dots, v_k \in S$ são tais que

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0.$$

Assim, temos uma combinação linear do vetor nulo $0 \in V$ envolvendo os vetores $v_1, \dots, v_k \in S$. Mas, também temos a combinação linear trivial

$$0 = 0 \cdot v_1 + \cdots + 0 \cdot v_k.$$

Como os escalares são únicos, deve ocorrer $\alpha_1 = \cdots = \alpha_k = 0$. Portanto, S é LI.

■

2.6 Dimensão de subespaços e da soma de subespaços

Proposição 2.6.1 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e $W \subseteq V$ um subespaço vetorial. Então, são válidas as seguintes afirmações.

(i) W é finitamente gerado e $\dim W \leq n$.

(ii) Se $\dim W = n$, então $W = V$.

(iii) Existe um subespaço vetorial $U \subseteq V$ tal que $V = W \oplus U$.

Demonstração. **(i)** Se $W = \{0\}$, então $\emptyset$ é base de W e $\dim W = 0$.

Suponha $W \neq \{0\}$. Seja $w_1 \in W \setminus \{0\}$. Se $W = [w_1]$, então $\{w_1\}$ é uma base de W e $\dim W = 1$.

Caso contrário, existe $w_2 \in W \setminus [w_1]$, o que implica que $\{w_1, w_2\}$ é LI. Se $W = [w_1, w_2]$, então $\{w_1, w_2\}$ é uma base de W e $\dim W = 2$.

Caso contrário...

Continuamos o processo acima até encontrar $B = \{w_1, \dots, w_m\}$ tal que B é base de W. Esse processo deve terminar com $m \leq n$, pois caso contrário, teríamos um conjunto LI em V com mais de n vetores, o que é um absurdo pela Proposição 2.5.24.

(ii) Seja $B = \{w_1, \dots, w_n\}$ uma base de W. Então, $B \subseteq W \subseteq V$ é um subconjunto LI com $n = \dim V$ vetores. Pela Proposição 2.5.24, B é base de V. Portanto, $W = [B] = V$.

(iii) Tome $B_1 = \{w_1, \dots, w_k\}$ uma base de W. Pela Proposição 2.5.23, existe $B_2 = \{v_{k+1}, \dots, v_n\} \subseteq V$ tal que $B_1 \cup B_2$ é base de V. Considere $U = [B_2]$. Deixamos como exercício a verificação de que $V = W \oplus U$.

■

Observação 2.6.2 O subespaço $U \subseteq V$ tal que $V = W \oplus U$ da proposição anterior não é único. Por exemplo, se $V = \mathbb{R}^2$ e $W = [(1,0)]$, então $\mathbb{R}^2 = W \oplus [(0,1)]$, mas também $\mathbb{R}^2 = W \oplus [(1,1)]$.

■ **Exemplo 2.6.3** Considere $V = P_2(\mathbb{R})$ e $W = [t^2+1, t^2-1] \subseteq V$. Determine $U \subseteq V$ tal que $V = W \oplus U$.

Solução. Como $B_1 = \{t^2+1, t^2-1\}$ é LI (verifique), B_1 é base de W. Vamos utilizar a ideia na demonstração da Proposição 2.6.1. Sabendo que $\dim V = 3$, precisamos encontrar um vetor $v \in V$ tal que $B_1 \cup \{v\}$ é base de V. Testando $v = t$, vemos que $\{t^2+1, t^2-1, t\}$ é LI (verifique), e portanto é base de V, já que tem $3 = \dim V$ vetores. Assim, $U = [t]$ é tal que $V = W \oplus U$.

Na Seção 2.8 será visto um método mais efetivo para completar bases do que simplesmente "chutar" vetores e testar se são LI. ■

Proposição 2.6.4 Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$. Se U e W são subespaços vetoriais de V, então

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

Demonstração. Primeiramente lembramos que $U \cap W$ é subespaço vetorial de V. Suponha $U \cap W \neq \{0\}$ e seja $B = \{v_1, \ldots, v_k\}$ uma base de $U \cap W$ $(1 \leq k \leq n)$. Como $B \subseteq U \cap W$, B é um subconjunto LI de U e de W. Então, pela Proposição 2.5.23, existem $u_1, \ldots, u_l \in U$ e $w_1, \ldots, w_m \in W$ tais que

$$B_1 = \{v_1, \ldots, v_k, u_1, \ldots, u_l\} \text{ é base de } U \quad \text{e} \quad B_2 = \{v_1, \ldots, v_k, w_1, \ldots, w_m\} \text{ é base de } W.$$

Provemos que

$$C = B_1 \cup B_2 = \{v_1, \ldots, v_k, u_1, \ldots, u_l, w_1, \ldots, w_m\} \text{ é base de } U+W.$$

Primeiro, mostremos que C gera $U+W$. A inclusão $[C] \subseteq U+W$ é imediata. Para a inclusão contrária, considere $v = u + w \in U+W$, com $u \in U$ e $w \in W$. Como B_1 é base de U e B_2 é base de W, existem $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l, \gamma_1, \ldots, \gamma_k, \delta_1, \ldots, \delta_m \in \mathbb{K}$ tais que

$$u = \alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 u_1 + \cdots + \beta_l u_l$$

e

$$w = \gamma_1 v_1 + \cdots + \gamma_k v_k + \delta_1 w_1 + \cdots + \delta_m w_m.$$

Logo,

$$u + w = (\alpha_1 + \gamma_1) v_1 + \cdots + (\alpha_k + \gamma_k) v_k + \beta_1 u_1 + \cdots + \beta_l u_l + \delta_1 w_1 + \cdots + \delta_m w_m \in [C].$$

Agora, mostremos que C é LI. Dada uma combinação linear nula:

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 u_1 + \cdots + \beta_l u_l + \gamma_1 w_1 + \cdots + \gamma_m w_m = 0, \tag{2.4}$$

temos

$$\gamma_1 w_1 + \cdots + \gamma_m w_m = -\alpha_1 v_1 - \cdots - \alpha_k v_k - \beta_1 u_1 - \cdots - \beta_l u_l \in U \cap W.$$

Portanto, existem $\delta_1, \ldots, \delta_k \in \mathbb{K}$ tais que

$$\gamma_1 w_1 + \cdots + \gamma_m w_m = \delta_1 v_1 + \cdots + \delta_k v_k \Rightarrow \gamma_1 w_1 + \cdots + \gamma_m w_m - \delta_1 v_1 - \cdots - \delta_k v_k = 0.$$

Mas B_2 é LI, logo

$$\gamma_1 = \cdots = \gamma_m = \delta_1 = \cdots = \delta_k = 0.$$

Assim, (2.4) se torna

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 u_1 + \cdots + \beta_l u_l = 0.$$

Como B_1 é LI, segue que

$$\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_l = 0.$$

Portanto, C é LI.

Concluímos então que C é base de $U+W$. Daí,

$$\dim(U+V) = k+l+m = (k+l)+(k+m)-k = \dim U + \dim W - \dim(U \cap W).$$

Se $U \cap W = \{0\}$, deixamos como exercício a verificação de que a união de uma base de U com uma base de W será uma base de $U+W$, donde segue o resultado. ■

■ **Exemplo 2.6.5** Considere $V = \mathbb{R}^4$ e os subespaços

$$U = [(1,0,1,0),(0,1,0,0)] \quad \text{e} \quad W = \{(x,y,z,t) \in \mathbb{R}^4 : x+y=0\}.$$

Determine as dimensões de U, W e $U+W$. A soma $U+W$ é direta? Vale que $U+W = \mathbb{R}^4$?

Solução. O subespaço U é gerado por $B_1 = \{(1,0,1,0),(0,1,0,0)\}$. Não é difícil verificar que B_1 é LI, portanto é base de U e $\dim U = 2$.

Vamos agora determinar uma base de W. Temos

$$v = (x,y,z,t) \in W \iff y = -x \iff v = (x,-x,z,t) = x(1,-1,0,0) + z(0,0,1,0) + t(0,0,0,1).$$

Ou seja,

$$v \in W \iff v \in [(1,-1,0,0),(0,0,1,0),(0,0,0,1)].$$

Portanto,

$$W = [(1,-1,0,0),(0,0,1,0),(0,0,0,1)].$$

Não é difícil verificar que $B_2 = \{(1,-1,0,0),(0,0,1,0),(0,0,0,1)\}$ é LI, portanto é base de W e $\dim W = 3$.

Vamos agora determinar uma base de $U \cap W$. Os vetores em U são da forma

$$v = a(1,0,1,0) + b(0,1,0,0) = (a,b,a,0), \ \ a,b \in \mathbb{R}.$$

Os vetores em W são da forma

$$v = (x,-x,z,t), \ \ x,y,z,t \in \mathbb{R}.$$

Para que $v \in U \cap W$, deve ocorrer

$$t = 0, \ \ x = z = a, \ \ b = -x = -a \Rightarrow v = (a,-a,a,0) = a(1,-1,1,0) \in [(1,-1,1,0)].$$

Assim, $U \cap W \subseteq [(1,-1,1,0)]$. Deixamos como exercício a verificação que $(1,-1,1,0) \in U \cap W$, o que implica $[(1,-1,1,0)] \subseteq U \cap W$, pois $U \cap W$ é subespaço vetorial. Portanto, $U \cap W = [(1,-1,1,0)]$, $\{(1,-1,1,0)\}$ é uma base de $U \cap W$ e $\dim(U \cap W) = 1$.

Pela Proposição 2.6.4,

$$\dim(U+W) = 2+3-1 = 4.$$

Além disso, como $U+W$ é um subespaço de $\mathbb{R}^4$, com $\dim(U+W) = \dim \mathbb{R}^4$, temos que $U+W = \mathbb{R}^4$. Por fim, a soma $U+W$ não é direta, pois $U \cap W \neq \{0\}$. ■

2.7 Coordenadas e matriz de mudança de base

Antes de definir formalmente os conceitos de coordenadas de um vetor e matriz de mudança de base, faremos uma introdução das ideias utilizando o espaço vetorial $\mathbb{R}^n$.

Dado um vetor $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ e a base canônica $C = \{e_1, \dots, e_n\}$, ao escrever v como uma combinação linear dos vetores de C, obtemos

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Os escalares $\alpha_1, \dots, \alpha_n$, nesse caso, são chamados de *coordenadas* de v em relação à base C, e denotamos

$$[v]_C = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \text{ou} \quad [v]_C = (\alpha_1, \dots, \alpha_n)_C = (\alpha_1, \dots, \alpha_n).$$

Observe que, nesse caso, o vetor $[v]_C$ das coordenadas de v se confunde com o próprio vetor v, mas não é sempre assim. Em geral, dada uma base $B = \{v_1, \dots, v_n\}$, existem únicos escalares $\beta_1, \dots, \beta_n$ tais que $v = \beta_1 v_1 + \cdots + \beta_n v_n$, e então define-se as coordenadas de v em relação à base B por $[v]_B = (\beta_1, \dots, \beta_n)_B$.

Por exemplo, se $n = 3$ e $B = \{(1,1,0), (0,1,0), (0,0,3)\}$, dado $v = (a,b,c) \in \mathbb{R}^3$, temos

$$(a,b,c) = \beta_1(1,1,0) + \beta_2(0,1,0) + \beta_3(0,0,3) = (\beta_1, \beta_1 + \beta_2, 3\beta_3) \Leftrightarrow \beta_1 = a,\ \beta_2 = b - a,\ \beta_3 = \frac{c}{3}.$$

Assim,

$$[v]_B = \left(a, b-a, \frac{c}{3}\right)_B.$$

Além de conhecer as coordenadas, queremos também saber como transitar das coordenadas de um vetor em relação a uma base para outra. É aí que entra a *matriz de mudança de base*. No exemplo acima, observe que

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b-a \\ c/3 \end{bmatrix}. \tag{2.5}$$

A matriz

$$M_{BC} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

é a matriz de mudança de base de B para C, pois ela faz a transição de coordenadas de um vetor em relação à base B para as coordenadas em relação à base C. Ou seja, (2.5) se traduz em

$$[v]_C = M_{BC}\,[v]_B.$$

Como a matriz M_{BC} é inversível, temos

$$[v]_B = M_{BC}^{-1}\,[v]_C.$$

A matriz $M_{CB} = M_{BC}^{-1}$ será a matriz de mudança de base de C para B.

Observe que a ordem dos vetores em B ou C importa, pois se trocarmos os vetores de posição, o vetor de coordenadas e a matriz de mudança de base seriam diferentes.

Definição 2.7.1 Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$. Fixe $B = \{v_1, \ldots, v_n\}$ uma *base ordenada* de V, isto é, uma base com a ordem dos vetores fixada. Dado $v \in V$, os únicos escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, são chamados de *coordenadas* de v em relação à base B, e denotamos

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \text{ou} \quad [v]_B = (\alpha_1, \ldots, \alpha_n)_B = (\alpha_1, \ldots, \alpha_n).$$

■ **Exemplo 2.7.2** Considere $V = P_2(\mathbb{R})$, a base $B = \{1, 1+x, 1+x^2\}$ e $p(x) = 2 + 4x + x^2 \in V$. Determine $[p(x)]_B$.

Solução. Precisamos determinar os escalares que aparecem ao escrevermos v como combinação linear do vetores da base B:

$$2 + 4x + x^2 = \alpha_1 \cdot 1 + \alpha_2 (1+x) + \alpha_3 (1+x^2).$$

Igualando os coeficientes dos polinômios, não é difícil ver que $\alpha_1 = -3$, $\alpha_2 = 4$ e $\alpha_3 = 1$. Portanto,

$$[p(x)]_B = (-3, 4, -1)_B.$$

■

A grande vantagem de trabalhar com coordenadas ao invés dos próprios vetores é o ganho computacional. Observe que, independente do espaço vetorial e da cara dos vetores (n-uplas, polinômios, funções, matrizes...) o vetor de coordenadas é sempre um elemento de $\mathbb{K}^n$ (ou de $M_{n\times 1}(\mathbb{K})$) onde n é a dimensão do espaço. Além disso, propriedades interessantes relacionadas aos vetores podem ser transportadas para suas coordenadas, como ilustra a proposição a seguir.

Proposição 2.7.3 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e B uma base de V. Para todos $u, v \in V$ e $\lambda \in \mathbb{K}$, são válidas:

(i) $[u+v]_B = [u]_B + [v]_B$;

(ii) $[\lambda u]_B = \lambda [u]_B$;

(iii) $v = 0 \in V$ se, e somente se, $[v]_B = 0 \in \mathbb{K}^n$;

(iv) $\{u_1, \ldots, u_m\} \subseteq V$ é LI se, e somente se, $\{[u_1]_B, \ldots, [u_m]_B\}$ é LI em $\mathbb{K}^n$.

Demonstração. Exercício. ■

Agora, veremos como construir a matriz de mudança de base. Fixe V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e duas bases ordenadas $B = \{u_1, \ldots, u_n\}$ e $C = \{v_1, \ldots, v_n\}$.

Basicamente, matriz de mudança de base de B para C tem na j-ésima coluna as coordenadas do vetor u_j em relação à base C. Para construí-la, então, precisamos escrever cada u_j como combinação linear dos vetores de C, e então coletar os escalares que aparecem.

Para cada $j = 1, \ldots, n$, existem únicos escalares $a_{1j}, \ldots, a_{nj} \in \mathbb{K}$, tais que

$$u_j = a_{1j} v_1 + \cdots + a_{nj} v_n = \sum_{i=1}^{n} a_{ij} v_i.$$

Seja $w \in V$. Suponha que $[w]_B = (\alpha_1, \ldots, \alpha_n)_B$ e $[w]_C = (\beta_1, \ldots, \beta_n)_C$, isto é,

$$w = \alpha_1 u_1 + \cdots + \alpha_n u_n = \sum_{j=1}^{n} \alpha_j u_j \quad \text{e} \quad w = \beta_1 v_1 + \cdots + \beta_n v_n = \sum_{i=1}^{n} \beta_i v_i$$

Então,

$$w = \sum_{j=1}^{n} \alpha_j u_j = \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{n} a_{ij} v_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} \alpha_j \right) v_i = \sum_{i=1}^{n} (a_{i1}\alpha_1 + \cdots + a_{in}\alpha_n) v_i.$$

Como os escalares da combinação linear dos vetores de uma base são únicos, temos

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + \cdots + a_{1n}\alpha_n \\ \qquad\vdots \\ \beta_n = a_{n1}\alpha_1 + \cdots + a_{nn}\alpha_n \end{cases}.$$

Em notação matricial, segue que

$$\underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}}_{[w]_C} = \underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}}_{M_{BC}} \underbrace{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}}_{[w]_B}.$$

A matriz $M_{BC} = (a_{ij})_{n \times n}$ é chamada de *matriz mudança de base* de B para C. Pelo exposto anteriormente, essa matriz satisfaz

$$\boxed{[w]_C = M_{BC}[w]_B, \quad \text{para todo } w \in V.} \tag{2.6}$$

Além disso, se $M = (b_{ij}) \in M_n(\mathbb{K})$ satisfaz $[w]_C = M[w]_B$, para todo $w \in V$, substituindo $w = u_j$, $j = 1, \ldots, n$, obtém-se que $b_{ij} = a_{ij}$ para todos i, j (verifique). Ou seja, $M = M_{BC}$. Portanto, a matriz de mudança de base é a <u>única</u> que satisfaz (2.6).

Observação 2.7.4 A notação de matriz de mudança de base não é unanimidade na literatura! Fiquem atentos se forem consultar outros livros, pois às vezes até a definição é diferente, invertendo o sentido da mudança. O padrão adotado para o curso de MAT-27 é a notação aqui apresentada. Em caso de dúvidas, lembrem da relação (2.6).

Proposição 2.7.5 Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e duas bases B e C. Então, M_{BC} é inversível e a matriz de mudança de base de C para B é $M_{CB} = M_{BC}^{-1}$.

Demonstração. Como B e C são bases, pela Proposição 2.7.3,

$$[v]_B = 0 \Leftrightarrow v = 0 \Leftrightarrow [v]_C = 0.$$

Então, o sistema linear homogêneo

$$M_{BC}X = 0$$

só admite a solução trivial $X = 0$. Portanto, a matriz do sistema M_{BC} é inversível. Sendo assim, para cada $v \in V$, temos

$$[v]_C = M_{BC}[v]_B \Leftrightarrow M_{BC}^{-1}[v]_C = [v]_B.$$

Pela unicidade da matriz de mudança de base (propriedade (2.6) com B no lugar de C e C no lugar de B), segue que $M_{BC}^{-1} = M_{CB}$. ■

■ **Exemplo 2.7.6** Seja $V = P_1(\mathbb{R})$. Determine a matriz de mudança de base da base canônica $C = \{1, x\}$ para a base $B = \{1+x, 1-x\}$ e as coordenadas do vetor $p(x) = 2+3x$ em relação à base B.

Solução. Podemos obter a matriz M_{CB} diretamente, ou obter M_{BC} e invertê-la. Note que, obter a matriz de mudança de qualquer base para a base canônica é simples, uma vez que é fácil descrever as coordenadas de qualquer vetor em relação a base canônica. No caso dos vetores de B:

$$1+x = 1\cdot 1 + 1\cdot x \quad \text{e} \quad 1-x = 1\cdot 1 - 1\cdot x,$$

logo

$$M_{BC} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \overset{(1.5)}{\Rightarrow} M_{CB} = M_{BC}^{-1} = -\frac{1}{2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}.$$

Agora, podemos usar a matriz M_{CB} para determinar as coordenadas de $p(x)$ em relação à B:

$$[p(x)]_B = M_{CB}[p(x)]_C = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5/2 \\ -1/2 \end{bmatrix}.$$

■

■ **Exemplo 2.7.7** Sejam $V = \{p(x) \in P_3(\mathbb{R}) : p(1) + p(-1) = 0\}$ e $B = \{x, x^3, 1+x-x^2\}$ base de V. Sabendo que

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

é a matriz de mudança de base de B para uma base C, determine a base C.

Solução. Sabemos, pela definição, que na matriz M_{CB}, de mudança de base de C para B, as colunas descrevem as coordenadas dos vetores de C em relação à base B. Assim, se conhecermos $M_{CB} = M_{BC}^{-1}$, poderemos determinar C. Determinemos M_{BC}^{-1}:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & -1 & 3 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 & 0 & 1 \end{array}\right] \overset{(L_2-L_1\to L_2)}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{(L_3+2L_2\to L_3)}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -2 & 2 & 1 \end{array}\right] \overset{(-L_2\to L_2)}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 2 & 1 \end{array}\right]$$

$$\overset{(L_2+L_3\to L_2)}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -2 & 2 & 1 \end{array}\right] \overset{(L_1-2L_3\to L_1)}{\sim} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -4 & -2 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -2 & 2 & 1 \end{array}\right].$$

Portanto,

$$M_{CB} = M_{BC}^{-1} = \begin{bmatrix} 5 & -4 & -2 \\ -1 & 1 & 1 \\ -2 & 2 & 1 \end{bmatrix}.$$

Daí, se $C = \{p_1, p_2, p_3\}$, então

$$[p_1]_B = (5, -1, -2)_B, \ \ [p_2]_B = (-4, 1, 2)_B, \ \ [p_3]_B = (-2, 1, 1)_B.$$

Logo,

$$p_1 = 5\cdot x - 1\cdot x^3 - 2\cdot(1+x-x^2) = -2+3x+2x^2-x^3,$$
$$p_2 = -4\cdot x + 1\cdot x^3 + 2\cdot(1+x-x^2) = 2-2x-2x^2+x^3,$$
$$p_3 = -2\cdot x + 1\cdot x^3 + 1\cdot(1+x-x^2) = 1-x-x^2+x^3.$$

■

2.8 Método prático para determinar ou completar bases de subespaços

Vimos anteriormente que todo conjunto gerador contém uma base e que todo subconjunto LI pode ser completado para uma base. Esses dois problemas podem ser traduzidos utilizando subespaços vetoriais: dado um conjunto gerador de um subespaço, como determinar uma base deste subespaço, e, dada uma base de um subespaço, como completar essa base para formar uma base de todo o espaço. Nesta seção, descreveremos métodos práticos, utilizando coordenadas, para resolver esses dois problemas.

Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e fixe $B = \{v_1, \dots, v_n\}$ uma base ordenada de V.

Tome um subconjunto $S = \{u_1, \dots, u_m\} \subseteq V$. Vamos determinar uma base para o subespaço $W = [S]$. Escreva cada vetor u_i como combinação linear da base B:

$$\begin{cases} u_1 = a_{11}v_1 + \cdots + a_{1n}v_n \\ \quad\vdots \\ u_m = a_{m1}v_1 + \cdots + a_{mn}v_n \end{cases}.$$

Procedimento 2.8.1 **(i)** Escreva a matriz A cuja i-ésima linha é formada pelas coordenadas de u_i em relação à B:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

(ii) Aplique operações elementares nas linhas de A até obter uma matriz A' na forma escalonada (ou escalonada reduzida).

(iii) As linhas não nulas de A' serão as coordenadas, em relação à base B, de vetores $w_1, \dots, w_r \in W$, $(r \leq m)$, de modo de $[\{w_1, \dots, w_r\}] = [S] = W$ e $S' = \{w_1, \dots, w_r\}$ é LI, ou seja, S' é base de W.

Para verificar as afirmações do item (iii), primeiro observamos como se traduzem as operações elementares nas linhas de A em termos dos vetores de S:

- trocar a posição de duas linhas da matriz corresponde a trocar dois vetores de posição;
- multiplicar uma linha da matriz por um escalar diferente de zero corresponde a trocar um vetor u_i por um múltiplo por escalar (não nulo) αu_i;
- somar a uma linha da matriz um múltiplo escalar de outra linha corresponde a trocar um vetor u_i por $u_i + \alpha u_j$ $(i \neq j)$.

Deixamos como exercício a verificação de que realizar essas operações num conjunto gerador não alteram o subespaço gerado. Como, eliminar os vetores nulos também não altera o subespaço gerado, segue que $[\{w_1, \dots, w_r\}] = [S] = W$.

Por fim, como a matriz A' está na forma escalonada, não é difícil verificar que suas linhas não nulas são vetores LI em $\mathbb{K}^n$, donde segue, junto com a Proposição 2.7.3, que $S' = \{w_1, \dots, w_r\}$ é LI em W.

■ **Exemplo 2.8.2** Sejam $V = P_3(\mathbb{R})$ e $W = [p_1, p_2, p_3] \subseteq V$, onde $p_1(x) = 1 + 2x^2 - x^3$, $p_2(x) = 1 + 5x^2$, $p_3(x) = -6x^2 - 2x^3$. Determine uma base para W.

Solução. Fixemos a base canônica $B = \{1, x, x^2, x^3\}$ de V. Seja $S = \{p_1, p_2, p_3\}$. A matriz A, cujas linhas são as coordenadas dos vetores em S em relação à base B, é dada por

$$A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 1 & 0 & 5 & 0 \\ 0 & 0 & -6 & -2 \end{bmatrix}.$$

Escalonando, temos

$$A = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 1 & 0 & 5 & 0 \\ 0 & 0 & -6 & -2 \end{bmatrix} \overset{(L_2 - L_1 \to L_2)}{\sim} \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & -6 & -2 \end{bmatrix} \overset{(L_3 + 2L_2 \to L_3)}{\sim} \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A'.$$

A matriz A' está na forma escalonada. Eliminando a terceira linha que é nula, as duas primeiras correspondem aos vetores

$$q_1(x) = 1 + 2x^2 - x^3 \quad \text{e} \quad q_2(x) = 3x^2 + x^3.$$

Portanto, $S' = \{q_1, q_2\}$ é uma base de W. ■

Suponha que aplicamos o Procedimento 2.8.1, mas durante a etapa (ii) não foi feita nenhuma troca de linhas nas operações elementares. Neste caso, talvez as linhas nulas não estejam todas abaixo das linhas não nulas, então A' não estaria bem numa forma escalonada, mas quase. Se a i-ésima linha de A' é nula, então o vetor u_i é combinação linear de outros vetores de S (verifique), e portanto ele poderia ser eliminado do conjunto gerador. Assim, se quiséssemos uma base de W contida em S, bastaria considerar S' o conjunto dos vetores u_j tais que a j-ésima linha de A' é não nula.

Considere agora $R = \{w_1, \ldots, w_r\}$ um subconjunto *LI* de V ($r \leq n$). Assim, R é uma base do subespaço $W = [R]$. Descreveremos como completar R para formar uma base de V. As demonstrações serão deixadas para o leitor, embora as ideias já foram apresentadas anteriormente.

Procedimento 2.8.3 **(i)** Escreva a matriz A cuja i-ésima linha é formada pelas coordenadas de w_i em relação à B.

(ii) Aplique operações elementares nas linhas de A até obter uma matriz A' na forma escalonada (ou escalonada reduzida). Como R é LI, A' não terá linhas nulas.

(iii) Acrescente $n - r$ linhas na matriz A' de modo a obter uma matriz $M \in M_n(\mathbb{K})$ na forma escalonada (ou escalonada reduzida). Essas novas linhas serão as coordenadas, em relação à base B, de vetores $w_{r+1}, \ldots, w_n \in V$ tais que $C = \{w_1, \ldots, w_r, w_{r+1}, \ldots, w_n\}$ é base de V.

Mais ainda, se $U = [w_{r+1}, \ldots, w_n]$, então $V = W \oplus U$.

■ **Exemplo 2.8.4** Considere $V = P_3(\mathbb{R})$ e $W = [p_1, p_2, p_3] \subseteq V$, onde $p_1(x) = 1 + 2x^2 - x^3$, $p_2(x) = 1 + 5x^2$, $p_3(x) = -6x^2 - 2x^3$, como no Exemplo 2.8.2. Determine um subespaço $U \subseteq V$ tal que $V = W \oplus U$.

Solução. Fixemos a base canônica $B = \{1, x, x^2, x^3\}$ de V. Vimos no Exemplo 2.8.2 que, se A é a matriz cujas linhas são as coordenadas dos vetores de $S = \{p_1, p_2, p_3\}$ em relação à base B, então

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

está na forma escalonada e pode ser obtida de A via operações elementares em suas linhas, sem efetuar trocas de linhas. Assim, as duas primeiras linhas correspondem aos vetores $q_1(x) = 1 + 2x^2 - x^3$ e

$q_2(x) = 3x^2 + x^3$, que formam uma base $S' = \{q_1, q_2\}$ de W. Se preferir uma base de W contida em S, basta observar que, como no escalonamento de A para A' não foi feita troca de linhas e a terceira linha em A' é nula, então p_3 é combinação linear de p_1 e p_2, e, $R = \{p_1, p_2\}$ é base de W contida em S.

Para completar uma base de V, considere a matriz obtida de A' eliminando a terceira linha que é nula. Ficamos assim com uma matriz 2×4, onde podemos acrescentar 2 linhas de modo a obter a matriz $M \in M_4(\mathbb{R})$ na forma escalonada:

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

As linhas 2 e 4 acrescentadas correspondem aos vetores $q_3(x) = x$ e $q_4(x) = x^3$. Os conjuntos $\{q_1, q_2, q_3, q_4\}$ e $\{p_1, p_2, q_3, q_4\}$ são bases de V, e se $U = [q_3, q_4]$, então $V = W \oplus U$. ■

Capítulo 3

Espaços vetoriais com produto interno

No Capítulo 2, introduzimos o conceito de Espaço Vetorial e estudamos suas propriedades. Neste capítulo, introduziremos o conceito de *produto interno*, que é uma generalização do produto interno ou escalar visto em geometria analítica para $\mathbb{R}^2$ e $\mathbb{R}^3$, e permite introduzir noções geométricas de comprimento e ortogonalidade para vetores de um espaço vetorial qualquer.

A discussão apresentada nesse capítulo foi motivada pelos livros [Axl], [Coe], [Reg] e [Str].

3.1 Definições e propriedades

Relembrando da Geometria Analítica, em $\mathbb{R}^3$, o produto escalar ou produto interno entre os vetores $u = (x_1, y_1, z_1)$ e $v = (x_2, y_2, z_2)$ é dado por

$$u \cdot v = x_1 x_2 + y_1 y_2 + z_1 z_2,$$

ou seja, é a soma do produto entre as coordenadas de u e v. Além disso, o produto interno também estava associado ao ângulo θ entre os vetores u e v, através da relação

$$u \cdot v = ||u||\,||v|| \cos\theta.$$

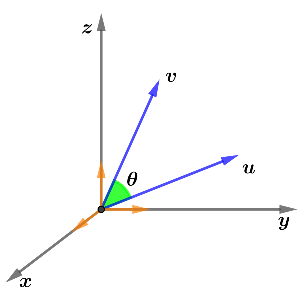


Figura 3.1: Ângulo θ entre os vetores u e v de $\mathbb{R}^3$.

A seguir, apresentamos a definição de produto interno para um espaço vetorial qualquer. Lembramos que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ e $\bar{z}$ denota o conjugado complexo do escalar z (ocorre $\bar{z} = z \Leftrightarrow z \in \mathbb{R}$).

Definição 3.1.1 Seja V um $\mathbb{K}$-espaço vetorial. Um *produto interno* sobre V é uma função $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{K}$ que satisfaz as propriedades:

(i) $\langle u+v,w\rangle = \langle u,w\rangle + \langle v,w\rangle$, para todos $u,v,w\in V$;

(ii) $\langle \lambda u,v\rangle = \lambda\langle u,v\rangle$, para todos $\lambda\in\mathbb{K}$, $u,v\in V$;

(iii) $\langle u,v\rangle = \overline{\langle v,u\rangle}$, para todos $u,v\in V$;

(iv) para todo $u\in V$, $\langle u,u\rangle\in\mathbb{R}$ e, se $u\neq 0$, $\langle u,u\rangle > 0$.

O par $(V,\langle\cdot,\cdot\rangle)$ é chamado de *espaço com produto interno* ou *espaço pré-Hilbert*.

Listaremos a seguir alguns exemplos de produto interno em diferentes espaços vetoriais. Um bom exercício para o leitor seria verificar que de fato eles satisfazem as quatro condições da Definição 3.1.1 sem olhar as demonstrações apresentadas.

■ **Exemplo 3.1.2** Sejam $u=(x_1,\dots,x_n)$ e $v=(y_1,\dots,y_n)$ vetores quaisquer de $\mathbb{K}^n$. A função

$$\langle x,y\rangle = \sum_{i=1}^n x_i\bar{y}_i$$

define um produto interno em $\mathbb{K}^n$. Este produto interno é chamado de *produto interno usual*, ou *canônico*, de $\mathbb{K}^n$.

Prova. Sejam $u=(x_1,\dots,x_n)$, $v=(y_1,\dots,y_n)$ e $w=(z_1,\dots,z_n)$ vetores quaisquer de $\mathbb{K}^n$ e λ um escalar. Segue das propriedades de números reais, ou complexos, que

(i) $\langle u+v,w\rangle = \sum_{i=1}^n (x_i+y_i)\bar{z}_i = \sum_{i=1}^n (x_i\bar{z}_i + y_i\bar{z}_i) = \sum_{i=1}^n x_i\bar{z}_i + \sum_{i=1}^n y_i\bar{z}_i = \langle u,w\rangle + \langle v,w\rangle$;

(ii) $\langle \lambda u,v\rangle = \sum_{i=1}^n (\lambda x_i)\bar{y}_i = \sum_{i=1}^n \lambda x_i\bar{y}_i = \lambda\sum_{i=1}^n x_i\bar{y}_i = \lambda\langle u,v\rangle$;

(iii) $\overline{\langle v,u\rangle} = \overline{\left(\sum_{i=1}^n y_i\bar{x}_i\right)} = \sum_{i=1}^n \overline{y_i\bar{x}_i} = \sum_{i=1}^n \bar{y}_i x_i = \sum_{i=1}^n x_i\bar{y}_i = \langle u,v\rangle$;

(iv) $\langle u,u\rangle = \sum_{i=1}^n x_i\bar{x}_i = \sum_{i=1}^n |x_i|^2 \in\mathbb{R}$. Além disso, se $u\neq 0$ então existe $x_i\neq 0$ e assim $\langle u,u\rangle > 0$.

Se pensarmos em u e v como vetores coluna de $\mathbb{K}^n$, isto é, $u,v\in M_{n\times 1}(\mathbb{K})$, então podemos escrever

$$\langle u,v\rangle = x^\top\bar{y}.$$

■

■ **Exemplo 3.1.3** Se $A=(a_{ij}), B=(b_{ij})\in M_n(\mathbb{K})$, a função

$$\langle A,B\rangle = \sum_{i,j=1}^n a_{ij}\overline{b_{ij}} = \sum_{i=1}^n\sum_{j=1}^n a_{ij}\overline{b_{ij}}$$

define um produto interno em $M_n(\mathbb{K})$. Este produto interno é chamado de *produto interno usual*, ou canônico, de $M_n(\mathbb{K})$. Outra maneira de reescrever essa função, utilizando matriz transposta conjugada e a função traço, é

$$\langle A,B\rangle = \operatorname{tr}\left(A\overline{B}^\top\right) \quad \text{(verifique).}$$

Prova. Sejam $A,B,C \in M_n(\mathbb{K})$, $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times n}$, $C = (c_{ij})_{n\times n}$ e λ um escalar. Segue das propriedades de números reais, ou complexos, que

(i) $\langle A+B,C\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij}+b_{ij})\overline{c_{ij}} = \sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij}\overline{c_{ij}}+b_{ij}\overline{c_{ij}}) = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{c_{ij}}+\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}\overline{c_{ij}}$

$= \langle A,C\rangle + \langle B,C\rangle$;

(ii) $\langle \lambda A,B\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda a_{ij})\overline{b_{ij}} = \lambda\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{b_{ij}} = \lambda\langle A,B\rangle$;

(iii) $\overline{\langle B,A\rangle} = \overline{\sum_{i=1}^{n}\sum_{j=1}^{n}b_{ij}\overline{a_{ij}}} = \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{b_{ij}\overline{a_{ij}}} = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{b_{ij}} = \langle A,B\rangle$;

(iv) $\langle A,A\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\overline{a_{ij}} = \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 \in \mathbb{R}$. Além disso, se $A \neq 0$ então existe $a_{ij} \neq 0$ e assim $\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 > 0$. Ou seja, $\langle A,A\rangle > 0$.

Portanto, a função

$$\langle A,B\rangle = \sum_{i,j=1}^{n} a_{ij}\overline{b_{ij}} = \operatorname{tr}\left(A\overline{B}^{\top}\right)$$

define um produto interno em $M_n(\mathbb{K})$. ■

■ **Exemplo 3.1.4** Se $f,g \in \mathscr{C}([a,b],\mathbb{R})$, a função

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt$$

define um produto interno no espaço $\mathscr{C}([a,b],\mathbb{R})$. Este produto interno é chamado de *produto interno usual*, ou *canônico*, de $\mathscr{C}([a,b],\mathbb{R})$.

Prova. As condições (i) e (ii) da Definição 3.1.1 seguem das propriedades do Cálculo Integral e a condição (iii) é mais imediata. Para verificar a condição (iv), considere $f \in \mathscr{C}([a,b],\mathbb{R})$ não identicamente nula. Como f é contínua em $[a,b]$, $f^2(t) = f(t)f(t)$, também é contínua e não nula, logo, existe um intevarlo $[c,d] \subseteq [a,b]$ tal que $f^2(t) > 0$ para todo $t \in [c,d]$. Tomando $m > 0$ o valor mínimo de f^2 em $[c,d]$, temos

$$\langle f,f\rangle = \int_a^b f(t)f(t)dt = \int_a^b f^2(t)dt \geq \int_c^d f^2(t)\,dt \geq \int_c^d m\,dt = m(d-c) > 0.$$

■

■ **Exemplo 3.1.5** Seja $V = P_2(\mathbb{R})$. Dados $p,q \in V$, a função

$$\langle p,q\rangle = p(0)q(0)+p(1)q(1)+p(-2)q(-2)$$

define um produto interno V.

Prova. As condições (i), (ii) e (iii) da Definição 3.1.1 seguem facilmente das propriedades de números reais. Para verificar a condição (iv), dado $p \in V$, temos

$$\langle p,p\rangle = p(0)^2+p(1)^2+p(-2)^2 \geq 0$$

e $\langle p,p\rangle = 0$ se, e somente se,

$$p(0) = p(1) = p(-2) = 0.$$

Como os elementos em V são polinômios de grau no máximo 2, a única maneira de p ter 3 raízes distintas é se p for o polinômio identicamente nulo. Portanto, $\langle p,p\rangle = 0$ se, e somente se, $p = 0$. ∎

Listaremos a seguir, as propriedades mais imediatas de um produto interno. As demonstrações seguem dos itens da Definição 3.1.1, sendo um bom exercício tentar prová-las sem olhar a demonstração apresentada.

Proposição 3.1.6 — Propriedades básicas. Sejam V um $\mathbb{K}$-espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$, $u,v,w \in V$ e $\alpha \in \mathbb{K}$.

(i) $\langle 0,u\rangle = \langle u,0\rangle = 0$.

(ii) $\langle u,u\rangle = 0$ se, e somente se, $u = 0$.

(iii) $\langle u,\alpha v\rangle = \overline{\alpha}\langle u,v\rangle$.

(iv) $\langle u,v+w\rangle = \langle u,v\rangle + \langle u,w\rangle$.

(v) Dados $m,n \in \mathbb{N}$, $\alpha_i,\beta_j \in \mathbb{K}$, $u_i,v_j \in V$, $i = 1,\dots,m$, $j = 1,\dots,n$, temos

$$\left\langle \sum_{i=1}^{m} \alpha_i u_i, \sum_{j=1}^{n} \beta_j v_j \right\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\overline{\beta_j}\left\langle u_i,v_j\right\rangle.$$

Demonstração. **(i)** $\langle 0,u\rangle = \langle 0\cdot u,u\rangle = 0\langle u,u\rangle = 0$ e $\langle u,0\rangle = \overline{\langle 0,u\rangle} = \overline{0} = 0$.

(ii) $(\Rightarrow)$ Segue do item (iv) da Definição 3.1.1 que $\langle u,u\rangle > 0$ para todo $u \neq 0$. Logo, se $\langle u,u\rangle = 0$, então $u = 0$.

$(\Leftarrow)$ Se $u = 0$, então $\langle u,u\rangle = \langle 0,0\rangle = 0$ pelo item (i).

(iii) $\langle u,\alpha v\rangle = \overline{\langle \alpha v,u\rangle} = \overline{\alpha\langle v,u\rangle} = \overline{\alpha}\,\overline{\langle v,u\rangle} = \overline{\alpha}\langle u,v\rangle$.

(iv) $\langle u,v+w\rangle = \overline{\langle v+w,u\rangle} = \overline{\langle v,u\rangle + \langle w,u\rangle} = \overline{\langle v,u\rangle} + \overline{\langle w,u\rangle} = \langle u,v\rangle + \langle u,w\rangle$.

(v) Segue por indução sobre m e n, utilizando os itens (i) e (ii) da Definição 3.1.1 e os itens (iii) e (iv) desta proposição.

∎

Além do produto interno, o conceito de norma vetorial também foi visto no curso de Geometria Analítica e será generalizado para um espaço vetorial qualquer.

Definição 3.1.7 Seja V um $\mathbb{K}$-espaço vetorial. Uma *norma* em V é uma função $||\cdot|| : V \to \mathbb{R}$ que satisfaz as propriedades a seguir.

(i) $||v|| \geq 0$ para todo $v \in V$.

(ii) $||v|| = 0$ se, e somente se, $v = 0$.

(iii) $||\alpha v|| = |\alpha|\,||v||$ para todo escalar $\alpha \in \mathbb{K}$ e todo vetor $v \in V$.

(iv) $||u+v|| \leq ||u|| + ||v||$ para todos $u,v \in V$ (*desigualdade triangular*).

O par $(V, ||\cdot||)$ é chamado de *espaço vetorial normado*. Se $||v|| = 1$, dizemos que o vetor v está *normalizado*.

■ **Exemplo 3.1.8** Seja $u = (x_1, \dots, x_n)$ um vetor qualquer de $\mathbb{R}^n$. A função

$$||u||_\infty = \max_{1 \le i \le n} |x_i|$$

define uma norma em $\mathbb{R}^n$, chamada de *norma infinito* ou *norma do máximo*.

Prova. Sejam $u = (x_1, \dots, x_n)$ e $v = (y_1, \dots, y_n)$ vetores quaisquer de $\mathbb{R}^n$ e α um escalar. Segue das propriedades de números reais, que

(i) $||u||_\infty = \max\limits_{1 \le i \le n} |x_i| \ge 0$;

(ii) $||u||_\infty = \max\limits_{1 \le i \le n} |x_i| = 0$ se, e somente se, $v = 0$;

(iii) $||\alpha u||_\infty = \max\limits_{1 \le i \le n} |\alpha x_i| = \max\limits_{1 \le i \le n} |\alpha||x_i| = |\alpha| \max\limits_{1 \le i \le n} |x_i| = |\alpha|\,||u||_\infty$;

(iv) $||u+v|| = \max\limits_{1 \le i \le n} |x_i + y_i| = |x_k + y_k|$ para algum $k \in \{1, \dots, n\}$. Então,

$$||u+v|| = |x_k + y_k| \le |x_k| + |y_k| \le \max_{1 \le i \le n} |x_i| + \max_{1 \le j \le n} |y_j| = ||u|| + ||v||.$$

Logo, a função

$$||v||_\infty = \max_{1 \le i \le n} |x_i|$$

define uma norma em $\mathbb{R}^n$. ■

A partir de um produto interno, pode-se obter uma norma no espaço vetorial, como veremos a seguir.

Teorema 3.1.9 Seja V um $\mathbb{K}$-espaço com produto interno $\langle \cdot, \cdot \rangle$. A função

$$||v|| = \sqrt{\langle v, v \rangle}, \quad \text{para todo } v \in V,$$

define uma norma em V. Neste caso, dizemos que $||\cdot||$ é a *norma proveniente do produto interno*, ou, *norma induzida pelo produto interno*.

Demonstração. Sejam u e v vetores quaisquer de V e α um escalar. Como $\langle u, u \rangle \in \mathbb{R}$ e $\langle 0, 0 \rangle = 0$, segue das condições da Definição 3.1.1 que

(i) $||v|| = \sqrt{\langle v, v \rangle} \ge 0$;

(ii) $||v|| = \sqrt{\langle v, v \rangle} = 0$ se, e somente se, $v = 0$;

(iii) $||\alpha v|| = \sqrt{\langle \alpha v, \alpha v \rangle} = \sqrt{\alpha \overline{\alpha} \langle v, v \rangle} = \sqrt{|\alpha|^2 \langle v, v \rangle} = |\alpha|\,||v||$.

(iv) A desigualdade triangular para a norma induzida é um resultado importante e será enunciada a seguir.

■

Proposição 3.1.10 Seja V um $\mathbb{K}$-espaço com produto interno $\langle\cdot,\cdot\rangle$. Então, a função $||\cdot||: V \to \mathbb{R}$ dada por $||v|| = \sqrt{\langle v,v\rangle}$ satisfaz, para todos $u,v \in V$:

(i) $|\langle u,v\rangle| \leq ||u||||v||$. Essa propriedade é conhecida como *Desigualdade de Cauchy-Schwarz*. Além disso, a igualdade é verificada se, e somente se, os vetores u e v são LD.

(ii) $||u+v|| \leq ||u||+||v||$.

Demonstração. **(i)** Note que o resultado é válido se um dos vetores é o vetor nulo (ocorre a igualdade e u e v são LD). Vamos supor, então, que $u \neq 0$ e $v \neq 0$. Dado $\lambda \in \mathbb{K}$ e defina $w = v - \lambda u$. Como $u,v \in V$ segue que w também é um vetor de V. Então,

$$\begin{aligned} 0 \leq ||w||^2 = \langle w,w\rangle = \langle v-\lambda u, v-\lambda u\rangle &= \langle v,v\rangle - \langle \lambda u,v\rangle - \langle v,\lambda u\rangle + \langle \lambda u,\lambda u\rangle \\ &= ||v||^2 - \lambda\langle u,v\rangle - \bar{\lambda}\langle v,u\rangle + |\lambda|^2||u||^2. \end{aligned}$$

Em particular, para $\lambda = \dfrac{\overline{\langle u,v\rangle}}{||u||^2}$, temos

$$0 \leq ||w||^2 = ||v||^2 - \frac{|\langle u,v\rangle|^2}{||u||^2} - \frac{|\langle u,v\rangle|^2}{||u||^2} + \frac{|\langle u,v\rangle|^2}{||u||^4}||u||^2 = ||v||^2 - \frac{|\langle u,v\rangle|^2}{||u||^2},$$

ou seja,

$$\frac{|\langle u,v\rangle|^2}{||u||^2} \leq ||v||^2 \;\Rightarrow\; |\langle u,v\rangle| \leq ||u||||v||.$$

Além disso, a igualdade ocorre se, e somente se, $w = 0$, ou seja, se u e v são LD.

(ii) Observe que

$$||u+v||^2 = \langle u+v,u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} + \langle v,v\rangle = ||u+v||^2 = ||u||^2 + 2\,\mathrm{Re}(\langle u,v\rangle) + ||v||^2. \quad (3.1)$$

Como $\mathrm{Re}(z) \leq |z|$ para todo $z \in \mathbb{C}$, segue, junto com o item anterior, que

$$||u+v||^2 \leq ||u||^2 + 2|\langle u,v\rangle| + ||v||^2 \leq ||u||^2 + 2||u||||v|| + ||v||^2,$$

ou seja,

$$||u+v||^2 = (||u||+||v||)^2 \Leftrightarrow ||u+v|| \leq ||u||+||v||.$$

■

Os exemplos de normas a seguir decorrem do Teorema 3.1.9 e dos Exemplos 3.1.2, 3.1.3, 3.1.4.

■ **Exemplo 3.1.11** Seja $u = (x_1,\ldots,x_n)$ um vetor qualquer de $\mathbb{K}^n$. A função

$$||x||_2 = \sqrt{x_1\overline{x_1} + \cdots + x_n\overline{x_n}}$$

define uma norma em $\mathbb{K}^n$. Essa norma é conhecida como *norma usual*, ou *canônica*, ou *Euclidiana*, de $\mathbb{K}^n$. ■

■ **Exemplo 3.1.12** Se $A = (a_{ij}) \in M_n(\mathbb{K})$, a função

$$||A||_F = \left(\sum_{i,j=1}^{n} a_{ij}\overline{a_{ij}}\right)^{1/2} = \sqrt{\mathrm{tr}\left(A\overline{A}^{\top}\right)}$$

define uma norma em $M_n(\mathbb{K})$. Essa norma é conhecida como *norma de Frobenius*. ■

■ **Exemplo 3.1.13** Se $f \in \mathscr{C}([a,b],\mathbb{R})$, a função

$$||f|| = \left(\int_a^b f(x)^2\,dx\right)^{1/2}$$

define uma norma em $\mathscr{C}([a,b],\mathbb{R})$. Essa norma é conhecida como *norma usual*, ou *canônica*, de $\mathscr{C}([a,b],\mathbb{R})$. ■

Proposição 3.1.14 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $||\cdot||$ a norma induzida associada. Então, são válidas as identidades a seguir, para todos $u,v \in V$.

(i) *Identidades de Polarização*

(a) se $\mathbb{K}=\mathbb{R}$, $\langle u,v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2$;

(b) se $\mathbb{K}=\mathbb{C}$, $\langle u,v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2 + \frac{i}{4}||u+iv||^2 - \frac{i}{4}||u-iv||^2$, onde $i^2=-1$.

(ii) *Lei do Paralelogramo*
$||u+v||^2 + ||u-v||^2 = 2||u||^2 + 2||v||^2$.

Demonstração. **(i)** Dados $u,v \in V$, segue de (3.1) e conta análoga que

$$\begin{aligned} ||u+v||^2 &= ||u||^2 + 2\,\mathrm{Re}(\langle u,v\rangle) + ||v||^2 \\ ||u-v||^2 &= ||u||^2 - 2\,\mathrm{Re}(\langle u,v\rangle) + ||v||^2. \end{aligned} \tag{3.2}$$

Considere $\mathbb{K}=\mathbb{R}$. Subtraindo as equações (3.2), temos

$$||u+v||^2 - ||u-v||^2 = 4\,\mathrm{Re}(\langle u,v\rangle) = 4\langle u,v\rangle \Leftrightarrow \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2 = \langle u,v\rangle.$$

Se $\mathbb{K}=\mathbb{C}$, subtraindo as equações (3.2), temos

$$||u+v||^2 - ||u-v||^2 = 4\,\mathrm{Re}(\langle u,v\rangle).$$

Substituindo agora v por iv em (3.2), e usando que $\mathrm{Re}(-iz) = \mathrm{Im}(z)$ para todo $z \in \mathbb{C}$, temos

$$||u+iv||^2 = ||u||^2 + 2\,\mathrm{Re}(-i\langle u,v\rangle) + |i|\,||v||^2 = ||u||^2 + 2\,\mathrm{Im}(\langle u,v\rangle) + ||v||^2, \tag{3.3}$$

e, analogamente,

$$||u-iv||^2 = ||u||^2 + 2\,\mathrm{Re}(i\langle u,v\rangle) + |-i|\,||v||^2 = ||u||^2 - 2\,\mathrm{Im}(\langle u,v\rangle) + ||v||^2. \tag{3.4}$$

Subtraindo (3.4) de (3.3), obtemos

$$||u+iv||^2 - ||u-iv||^2 = 4\,\mathrm{Im}(\langle u,v\rangle) \Leftrightarrow \frac{1}{4}||u+iv||^2 - \frac{1}{4}||u-iv||^2 = \mathrm{Im}(\langle u,v\rangle).$$

Como $\langle u,v\rangle = \mathrm{Re}(\langle u,v\rangle) + i\,\mathrm{Im}(\langle u,v\rangle)$, segue que

$$\langle u,v\rangle = \frac{1}{4}||u+v||^2 - \frac{1}{4}||u-v||^2 + \frac{i}{4}||u+iv||^2 - \frac{i}{4}||u-iv||^2.$$

(ii) Somando as equações 3.2 obtemos

$$||u+v||^2 + ||u-v||^2 = 2||u||^2 + 2||v||^2.$$

■

Dado um produto interno num espaço vetorial, o Teorema 3.1.9 afirma que é possível definir uma norma no espaço a partir dele. Perguntas naturais que surgem são: dada uma norma num espaço vetorial, ela é a norma induzida por algum produto interno? Se sim, é possível recuperar o produto interno entre dois vetores quaisquer a partir da norma?

A resposta da segunda pergunta é sim, e para descrever o produto interno entre dois vetores quaisquer conhecendo-se a norma, utiliza-se as Identidades de Polarização.

A primeira pergunta não é simples de responder, foi preciso um extenso trabalho de muitos matemáticos, com destaque para Maurice René Fréchet e John von Neumann [Mey]. A conclusão obtida foi a seguinte.

Dada uma norma $||\cdot||$ em um espaço vetorial V, existe um produto interno em V que induz essa norma se, e somente se, a norma satisfaz a Lei do Paralelogramo para todos $u,v \in V$.

A demonstração desse fato utiliza ferramentas de Espaços Métricos que fogem do escopo deste curso, e por isso não será apresentada aqui.

Exercício 3.1 Verifique que a norma infinito em $\mathbb{R}^n$ não satisfaz a Lei do Paralelogramo e, portanto, não existe um produto interno associado a tal norma.

Se V é um $\mathbb{K}$-espaço vetorial normado, então podemos introduzir as noções geométricas de distância entre dois vetores e comprimento de um vetor.

Definição 3.1.15 A *distância* entre os vetores u e v é definida por $d(u,v) = ||u-v||$ e o *comprimento* de u por $||u||$.

Além disso, se V é um $\mathbb{R}$-espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$, podemos generalizar a noção de ângulo entre dois vetores, de $\mathbb{R}^2$ e $\mathbb{R}^3$ para V. Sejam $u,v \in V \setminus \{0\}$. Pela Desigualdade de Cauchy-Schwarz,

$$|\langle u,v\rangle| \leq ||u||||v|| \Leftrightarrow -||u||||v|| \leq \langle u,v\rangle \leq ||u||||v|| \Leftrightarrow -1 \leq \frac{\langle u,v\rangle}{||u||||v||} \leq 1.$$

Logo, existe único $\theta \in [0,\pi]$ tal que

$$\cos\theta = \frac{\langle u,v\rangle}{||u||||v||}. \tag{3.5}$$

Esse θ pode ser definido como o *ângulo* entre u e v. O ângulo entre o vetor nulo e qualquer outro é definido como sendo 0.

Observação 3.1.16 Neste curso, estaremos interessados nas normas provenientes de produto interno. Daqui por diante, toda vez que nos referirmos a norma em um espaço vetorial com produto interno, estaremos nos referindo a norma induzida por ele.

3.2 Ortogonalidade e Processo de Gram-Schmidt

A ideia de ortogonalidade entre vetores ou retas é essencial na geometria. Nesta seção, estenderemos essa noção para espaços vetoriais quaisquer com produto interno. Lembre que, em $\mathbb{R}^2$ ou $\mathbb{R}^3$ dois vetores u,v são ortogonais (ou perpendiculares) se $\langle u,v\rangle = 0$. Quando se tratam de espaços mais gerais, como o de matrizes ou funções, por exemplo, a visualização geométrica não será possível. No entanto, como veremos adiante, algumas propriedades interessantes continuam válidas.

Definição 3.2.1 Seja V um $\mathbb{K}$-espaço vetorial com produto interno $\langle \cdot, \cdot \rangle$. Dados $u, v \in V$, dizemos que u e v são *ortogonais*, e denotamos $u \perp v$, se $\langle u, v \rangle = 0$.
Um subconjunto $S \subseteq V$ é dito *ortogonal* se os seus elementos são dois a dois ortogonais, isto é $u \perp v$ para todos $u, v \in S$, $u \neq v$. Além disso, dizemos que S é *ortonormal* se for um conjunto ortogonal e $||u|| = 1$ para todo $u \in S$.

Observação 3.2.2 O vetor nulo é ortogonal a todos os vetores do espaço. Além disso, ele é o **único** vetor com essa propriedade, pois se $\langle u, v \rangle = 0$ para todo $v \in V$, em particular para $v = u$ tem-se $\langle u, u \rangle = 0 \Rightarrow u = 0$.

■ **Exemplo 3.2.3** Seja $V = \mathbb{R}^4$ com o produto interno usual. Determine se os vetores $u = (1, 0, 3, -2)$ e $v = (3, 5, -1, 0)$ são ortogonais.

Solução. Basta calcularmos o produto interno entre esses dois vetores,

$$\langle u, v \rangle = 1 \cdot 3 + 0 \cdot 5 - 3 \cdot 1 - 2 \cdot 0 = 0.$$

Como $\langle u, v \rangle = 0$, segue que u e v são ortogonais. ■

■ **Exemplo 3.2.4** Seja $V = \mathbb{C}^3$ com produto interno usual. Determine se os vetores $u = (-2, i, 1)$ e $v = (1, 0, i)$ são ortogonais.

Solução. Basta calcularmos o produto interno entre esses dois vetores,

$$\langle u, v \rangle = -2 \cdot \bar{1} + i \cdot \bar{0} + 1 \cdot \bar{i} = -2 - i.$$

Como $\langle u, v \rangle \neq 0$, segue que u e v não são ortogonais. ■

■ **Exemplo 3.2.5** Seja $V = \mathscr{C}([-\pi, \pi], \mathbb{R})$ com o produto interno usual. Determine se o conjunto $S = \{\operatorname{sen} x, \cos x\}$ é ortonormal.

Solução. Primeiramente vamos calcular o produto interno entre os vetores de S,

$$\langle \operatorname{sen} x, \cos x \rangle = \int_{-\pi}^{\pi} \operatorname{sen} x \cos x \, dx = \left[\frac{\operatorname{sen}^2 x}{2} \right]_{-\pi}^{\pi} = 0.$$

Como $\langle \operatorname{sen} x, \cos x \rangle = 0$, segue que $\operatorname{sen} x$ e $\cos x$ são ortogonais e, portanto, S é ortogonal. Além disso,

$$||\operatorname{sen} x||^2 = \langle \operatorname{sen} x, \operatorname{sen} x \rangle = \int_{-\pi}^{\pi} \operatorname{sen}^2 x \, dx = \left[\frac{x}{2} - \frac{\operatorname{sen}(2x)}{4} \right]_{-\pi}^{\pi} = \pi \neq 1,$$

ou seja, S não é um conjunto ortonormal. ■

■ **Exemplo 3.2.6** Seja $V = \mathscr{C}([0, 1], \mathbb{R})$ com produto interno usual. Determine o ângulo entre t e t^2.

Solução. Utilizando a definição em (3.5), precisamos calcular $||t||$, $||t^2||$ e o produto interno $\langle t, t^2 \rangle$. Assim,

$$||t||^2 = \int_0^1 t^2 \, dt = \left[\frac{t^3}{3} \right]_0^1 = \frac{1}{3} \Rightarrow ||t|| = \frac{1}{\sqrt{3}},$$

$$||t^2||^2 = \int_0^1 t^4 \, dt = \left[\frac{t^5}{5} \right]_0^1 = \frac{1}{5} \Rightarrow ||t^2|| = \frac{1}{\sqrt{5}},$$

$$\langle t, t^2 \rangle = \int_0^1 t^3 \, dt = \left[\frac{t^4}{4} \right]_0^1 = \frac{1}{4}.$$

Daí, $\cos\theta = \dfrac{\sqrt{15}}{4}$ e $\theta = \arccos\left(\dfrac{\sqrt{15}}{4}\right)$. ■

A seguir, veremos algumas propriedades de vetores ortogonais e conjuntos ortogonais.

Proposição 3.2.7 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $S \subseteq V$ um conjunto ortogonal que não contém o vetor nulo. Então S é um conjunto LI.

Demonstração. Sejam $v_1, \ldots, v_n \in S$, $\alpha_i \in \mathbb{K}$, $i = 1, \ldots, n$, tais que

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

Para cada $j \in \{1, \ldots, n\}$ temos que $\langle v_i, v_j \rangle = 0$, se $i \neq j$, e assim

$$0 = \langle \alpha_1 v_1 + \cdots + \alpha_n v_n, v_j \rangle = \alpha_j \langle v_j, v_j \rangle = \alpha_j ||v_j||^2.$$

Como $v_j \neq 0$, segue que $\alpha_j = 0$. Portanto, S é um conjunto L.I. ■

Definição 3.2.8 Uma *base ortogonal* para um espaço vetorial V é uma base para V que também é um conjunto ortogonal. E, uma *base ortonormal* para V é uma base para V que também é um conjunto ortonormal.

■ **Exemplo 3.2.9** As bases canônicas de $\mathbb{K}^n$ e $M_n(\mathbb{K})$ são bases ortonormais em relação aos produtos internos usuais desses espaços (verifique). ■

A proposição a seguir apresenta a primeira vantagem de se trabalhar com bases ortogonais.

Proposição 3.2.10 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno $B = \{v_1, \ldots, v_n\}$ uma base ortogonal para um subespaço W de V. Então, qualquer $w \in W$ é escrito como

$$w = \sum_{j=1}^{n} \frac{\langle w, v_j \rangle}{||v_j||^2} v_j. \tag{3.6}$$

Essa representação é chamada de *expansão de Fourier* para w.

Demonstração. Seja $w \in W$. Como B é base ortogonal de W, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que

$$w = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Além disso, para cada $j \in \{1, \ldots, n\}$,

$$\langle w, v_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle.$$

Como B é ortogonal, $\langle v_i, v_j \rangle = 0$ para todo $i \neq j$, logo

$$\langle w, v_j \rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle = \alpha_j \langle v_j, v_j \rangle = \alpha_j ||v_j||^2.$$

Como $v_j \neq 0$, segue que $\alpha_j = \dfrac{\langle w, v_j \rangle}{||v_j||^2}$, ou seja, $w = \displaystyle\sum_{j=1}^{n} \frac{\langle w, v_j \rangle}{||v_j||^2} v_j$. ■

Observação 3.2.11 As ordens dos produtos internos que aparecem na equação (3.6) são importantes se $\mathbb{K}=\mathbb{C}$.

A Proposição 3.2.10 mostra que as coordenadas de um vetor são facilmente determinadas quando trabalhamos com uma base $B=\{v_1,\ldots,v_n\}$ ortogonal:

$$[w]_B=\left(\frac{\langle w,v_1\rangle}{||v_1||^2}\,,\,\ldots\,,\,\frac{\langle w,v_n\rangle}{||v_n||^2}\right).$$

Além disso, se B é uma base ortonormal de V, dados $u=\sum_{i=1}^{n}\alpha_i v_i$ e $v=\sum_{j=1}^{n}\beta_j v_j$ em V, temos

$$\langle u,v\rangle=\left\langle\sum_{i=1}^{n}\alpha_i\,v_i,\sum_{j=1}^{n}\beta_j\,v_j\right\rangle=\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\overline{\beta_j}\langle v_i,v_j\rangle=\sum_{i=1}^{n}\alpha_i\overline{\beta_i}.$$

Ou seja, se $[u]_B=(\alpha_1,\ldots,\alpha_n)$ e $[v]_B=(\beta_1,\ldots,\beta_n)$, então o produto interno $\langle u,v\rangle$ em V é igual ao produto interno usual de $\mathbb{K}^n$ aplicado às coordenadas $\langle [u]_B,[v]_B\rangle$, e portanto é simples de se trabalhar do ponto de vista computacional.

Uma pergunta natural que surge é: todo espaço vetorial com produto interno possui base ortonormal?

A resposta é sim, e para obter bases ortonormais, utilizaremos o processo de Gram-Schmidt que será visto adiante. Antes, apresentamos um resultado que será útil nesse processo.

Proposição 3.2.12 Seja V um $\mathbb{K}$-espaço vetorial e $S=\{v_1,\ldots,v_n\}\subseteq V$ um subconjunto ortogonal. Então, dado $v\in V$, o vetor

$$u=v-\sum_{i=1}^{n}\frac{\langle v,v_i\rangle}{||v_i||^2}v_i \tag{3.7}$$

é ortogonal a todo vetor do subespaço gerado por S. Se S é ortonormal, então $u=v-\sum_{i=1}^{n}\langle v,v_i\rangle v_i$.

Antes de demonstrar esse resultado, faremos um exemplo em $\mathbb{R}^2$ para entender melhor o seu significado.

■ **Exemplo 3.2.13** Considere o espaço $\mathbb{R}^2$ com o produto interno usual. Sejam $v_1=(1,0)$ e $S=\{v_1\}$. Dado um vetor $v=(a,b)\in\mathbb{R}^2$, temos

$$u=v-\langle v,v_1\rangle v_1=(a,b)-\langle(a,b),(1,0)\rangle(1,0)=(a,b)-a\cdot(1,0)=(0,b),$$

ou seja, u só tem componente não nula na direção de $(0,1)$, que é ortogonal à v_1. ■

Demonstração da Proposição 3.2.12. Seja v um vetor qualquer de V. Defina

$$u=v-\sum_{i=1}^{n}\frac{\langle v,v_i\rangle}{||v_i||^2}v_i.$$

Queremos mostrar que u é ortogonal a qualquer vetor do subespaço gerado por S. Para isso, primeiramente vamos mostrar que u é ortogonal a qualquer vetor do conjunto S. Se $v_j\in S$, então

$$\langle u,v_j\rangle=\left\langle v-\sum_{i=1}^{n}\frac{\langle v,v_i\rangle}{||v_i||^2}v_i,v_j\right\rangle=\langle v,v_j\rangle-\sum_{i=1}^{n}\frac{\langle v,v_i\rangle}{||v_i||^2}\langle v_i,v_j\rangle.$$

Como S é ortogonal, $\langle v_i, v_j \rangle = 0$ para todo $i \neq j$ e segue que

$$\langle u, v_j \rangle = \langle v, v_j \rangle - \frac{\langle v, v_j \rangle}{||v_j||^2} ||v_j||^2 = 0,$$

ou seja, $u \perp v_j$, $j = 1, \ldots, n$. Logo, u é ortogonal a qualquer vetor de S.

Considere agora um vetor qualquer de $[S]$, digamos $w \in [S]$. Então, $w = \alpha_1 v_1 + \cdots + \alpha_n v_n$ e assim

$$\langle u, w \rangle = \left\langle u, \sum_{i=1}^{n} \alpha_j v_j \right\rangle = \sum_{i=1}^{n} \overline{\alpha}_j \langle u, v_i \rangle = 0.$$

Portanto, u é ortogonal ao subespaço gerado por S. ■

Processo de Gram-Schmidt

Seja $B = \{v_1, \ldots, v_n\}$ uma base qualquer para um espaço vetorial V com produto interno. Construiremos uma base **ortogonal** $\Gamma = \{u_1, \ldots, u_n\}$ para V, a partir dos vetores de B. Isso será feito de maneira indutiva, de modo que, para cada $k \in \{1, \ldots, n\}$, $\Gamma_k = \{u_1, \ldots, u_k\}$ será uma base ortogonal para $V_k = [v_1, \ldots, v_k]$.

Primeiro passo: como B é base, $v_1 \neq 0$, logo $u_1 = v_1$ é tal que $\Gamma_1 = \{u_1\}$ é ortogonal e $[v_1] = [u_1]$.

Ilustraremos explicitamente o passo para $k = 2$ para facilitar a compreensão. Queremos determinar u_2 ortogonal a u_1 tal que $[u_1, u_2] = [v_1, v_2]$. Considere

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{||u_1||^2} u_1.$$

Pela Proposição 3.2.12, $u_2 \perp u_1$. Além disso, $u_2 \neq 0$, pois, caso contrário, teríamos

$$v_2 = \frac{\langle v_2, u_1 \rangle}{||u_1||^2} u_1 \Rightarrow v_2 \in [u_1] = [v_1]$$

e $\{v_1, v_2\}$ seria LD, o que não é verdade. Portanto, $\Gamma_2 = \{u_1, u_2\}$ é um conjunto ortogonal de vetores não nulos, com $[u_1, u_2] \subseteq [u_1, v_2] = [v_1, v_2]$ e, portanto, é base ortogonal para $V_2 = [v_1, v_2]$.

Suponha $1 < k \leq n$ e que $\Gamma_{k-1} = \{u_1, \ldots, u_{k-1}\}$ seja uma base ortogonal para $V_{k-1} = [v_1, \ldots, v_{k-1}]$ (hipótese de indução). Determinaremos u_k tal que $\Gamma_k = \{u_1, \ldots, u_{k-1}, u_k\}$ é uma base ortogonal para V_k. Defina

$$u_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, u_i \rangle}{||u_i||^2} u_i.$$

Pela Proposição 3.2.12, u_k é ortogonal a todos os vetores de Γ_{k-1}. Além disso, $u_k \neq 0$, pois, caso contrário, chegaríamos que $v_k \in [u_1, \ldots, u_{k-1}] = [v_1, \ldots, v_{k-1}]$, absurdo pois $\{v_1, \ldots, v_k\}$ é LI. Portanto, $\Gamma_k = \{u_1, \ldots, u_k\}$ é base ortogonal para $V_k = [v_1, \ldots, v_k]$.

Pelo Princípio de Indução, $\Gamma_n = \{u_1, \ldots, u_n\}$ é base ortogonal para V. Para obter uma base ortonormal, basta normalizar os vetores de Γ_n: considerando $w_j = \dfrac{u_j}{||u_j||}$, $j = 1, \ldots, n$, então $\overline{\Gamma}_n = \{w_1, \ldots, w_n\}$ gera o mesmo espaço V_n e é ortonormal, logo é uma base ortonormal para V_n. Também, pode-se normalizar os vetores em cada etapa do processo.

Esse procedimento define o chamado *Processo de Ortonormalização de Gram-Schmidt*, e demonstra o teorema a seguir.

Teorema 3.2.14 Todo espaço vetorial de dimensão finita $n \geq 1$ com produto interno possui base ortonormal.

Observação 3.2.15 Se, no processo de Gram-Schmidt, iniciássemos com um conjunto gerador $S = \{v_1, \ldots, v_m\}$ que não é LI (portanto não é base), poderíamos aplicar da mesma maneira o procedimento, com a diferença que obteríamos alguns vetores $w_j = 0$. Assim, no final seria obtido um conjunto gerador ortogonal, onde o número de vetores não nulos é exatamente a dimensão de $[S]$.

■ **Exemplo 3.2.16** Seja $V = \mathbb{R}^4$ com produto interno usual, encontre uma base ortonormal para $W = [(1,1,1,1),(1,1,2,1),(1,1,1,2)]$.

Solução. Sejam $v_1 = (1,1,1,1)$, $v_2 = (1,1,2,1)$ e $v_3 = (1,1,1,2)$. O conjunto $B = \{v_1, v_2, v_3\}$ é LI (verifique), logo é uma base para W. Para encontrar uma base ortonormal para W basta aplicarmos o processo de Gram-Schmidt para B. Temos,

(i) $w_1 = (1,1,1,1)$, $||w_1|| = \sqrt{4} = 2$ e $u_1 = \dfrac{w_1}{||w_1||} = \dfrac{1}{2}(1,1,1,1)$.

(ii) $w_2 = v_2 - \langle v_2, u_1 \rangle u_1$ e $u_2 = \dfrac{w_2}{||w_2||}$.

Calculando os valores necessários:

$$\langle v_2, u_1 \rangle = \frac{5}{2},$$

$$w_2 = (1,1,2,1) - \frac{5}{2}\frac{1}{2}(1,1,1,1) = \frac{1}{4}(-1,-1,3,-1),$$

$$||w_2|| = \frac{\sqrt{12}}{4} = \frac{\sqrt{3}}{2},$$

concluímos que $u_2 = \dfrac{1}{2\sqrt{3}}(-1,-1,3,-1)$.

(iii) $w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$ e $u_3 = \dfrac{w_3}{||w_3||}$.

Calculando os valores necessários:

$$\langle v_3, u_1 \rangle = \frac{5}{2},$$

$$\langle v_3, u_2 \rangle = \frac{-1}{2\sqrt{3}},$$

$$w_3 = (1,1,1,2) - \frac{5}{2}\frac{1}{2}(1,1,1,1) + \frac{1}{2\sqrt{3}}\frac{1}{2\sqrt{3}}(-1,-1,3,-1) = \tfrac{1}{3}(-1,-1,0,2),$$

$$||v_3|| = \frac{\sqrt{6}}{3},$$

concluímos que $u_3 = \dfrac{1}{\sqrt{6}}(-1,-1,0,2)$.

Logo,

$$C = \left\{ \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right), \left(-\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, \frac{3}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}} \right), \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, 0, \frac{2}{\sqrt{6}} \right) \right\}$$

é base ortonormal para W.

■

■ **Exemplo 3.2.17** Seja $B = \{1, x\}$. Encontre uma base ortonormal para $[B]$, considerando o produto interno usual de $\mathscr{C}([-1,1],\mathbb{R})$.

Solução. Sejam $v_1(x) = 1$ e $v_2(x) = x$. Para encontrar uma base ortonormal para $[B]$ basta aplicarmos o processo de Gram-Schmidt. Assim,

(i) $w_1(x) = 1$, $||w_1||^2 = \int_{-1}^{1} 1dx = 2 \Rightarrow u_1(x) = \dfrac{w_1(x)}{||w_1||} = \dfrac{1}{\sqrt{2}}$.

(ii) $w_2 = v_2 - \langle v_2, u_1 \rangle u_1$ e $u_2 = \dfrac{w_2}{||w_2||}$.

Calculando os valores necessários:

$$\langle v_2, u_1 \rangle = \int_{-1}^{1} x \cdot \frac{1}{\sqrt{2}}\, dx = \left[\frac{x^2}{2\sqrt{2}} \right]_{-1}^{1} = 0,$$

$$w_2(x) = x,$$

$$||w_2||^2 = \int_{-1}^{1} x^2 dx = \left[\frac{x^3}{3} \right]_{-1}^{1} = \frac{2}{3},$$

concluímos que $u_2(x) = \dfrac{\sqrt{3}}{\sqrt{2}}\, x$.

Logo, $C = \left\{ \dfrac{1}{\sqrt{2}}, \dfrac{\sqrt{3}x}{\sqrt{2}} \right\}$ é uma base ortonormal para $[B]$. ■

■ **Exemplo 3.2.18 — Fatoração QR.** Sejam $A \in M_{m\times n}(\mathbb{R})$ uma matriz cujas colunas são os vetores coluna $x_1, \ldots, x_n \in \mathbb{R}^m$ e $R(A)$ o espaço coluna de A (veja Exemplo 2.3.6). Pelo processo de ortonormalização de Gram-Schmidt, podemos encontrar $u_1, \ldots, u_n \in \mathbb{R}^m$ ortogonais tais que $R(A) = [u_1, \ldots, u_n]$ e $[u_1, \ldots, u_k] = [x_1, \ldots, x_k]$ para qualquer $k \in \{1, \ldots, n\}$. Ou seja, para cada k, existem escalares $r_{1k}, \ldots, r_{kk} \in \mathbb{R}$ tais que

$$x_k = r_{1k}u_1 + \cdots + r_{kk}u_k + 0u_{k+1} + \cdots + 0u_n.$$

Definindo

$$Q = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix} \in M_{m\times n}(\mathbb{R}) \text{ e } R = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ 0 & r_{22} & \ldots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & r_{nn} \end{bmatrix} \in M_n(\mathbb{R})$$

podemos escrever $A = QR$, com R uma matriz triangular superior e Q uma matriz cujas colunas são ortogonais. Essa fatoração é bastante utilizada em matemática aplicada. Além disso, se $\dim(R(A)) = n$, então R é não singular. ■

Quando trabalhamos com bases ortonormais, a matriz de mudança de base também possui uma característica interessante, relacionada com a definição a seguir.

Definição 3.2.19 Seja $A \in M_n(\mathbb{K})$. Dizemos que A é uma matriz *unitária* se $A\overline{A}^\top = \overline{A}^\top A = I_n$, isto é, $A^{-1} = \overline{A}^\top$. Quando $\mathbb{K} = \mathbb{R}$, onde neste caso $AA^\top = A^\top A = I_n$, dizemos que A é *ortogonal*.

A proposição a seguir traz uma caracterização de matriz unitária em termos de suas linhas ou colunas.

Proposição 3.2.20 Seja $A \in M_n(\mathbb{K})$. São equivalentes:

(i) A é unitária;

(ii) as colunas de A formam um conjunto ortonormal em $\mathbb{K}^n$ (considerando produto interno usual);

(iii) as linhas de A formam um conjunto ortonormal em $\mathbb{K}^n$ (considerando produto interno usual).

Demonstração. Escreva $A = (a_{ij})$, $u_i = (a_{i1}, \ldots, a_{in})$ a i-ésima linha de A e $v_j = (a_{1j}, \ldots, a_{nj})^\top$ a j-ésima coluna de A. Observe que

$$A\overline{A}^\top = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{n1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{bmatrix} = \begin{bmatrix} \langle u_1, u_1 \rangle & \langle u_1, u_2 \rangle & \cdots & \langle u_1, u_n \rangle \\ \langle u_2, u_1 \rangle & \langle u_2, u_2 \rangle & \cdots & \langle u_2, u_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_n, u_1 \rangle & \langle u_n, u_2 \rangle & \cdots & \langle u_n, u_n \rangle \end{bmatrix}$$

e,

$$\overline{A}^\top A = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{n1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_2, v_1 \rangle & \cdots & \langle v_n, v_1 \rangle \\ \langle v_1, v_2 \rangle & \langle v_2, v_2 \rangle & \cdots & \langle v_n, v_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_1, v_n \rangle & \langle v_2, v_n \rangle & \cdots & \langle v_n, v_n \rangle \end{bmatrix}.$$

Portanto, considerando $\delta_{ii} = 1$ e $\delta_{ij} = 0$ para todo $i \neq j$, temos

$$\langle u_i, u_j \rangle = \delta_{ij} \text{ para todos } i, j = 1, \ldots, n \Leftrightarrow A\overline{A}^\top = I_n,$$

e

$$\langle v_i, v_j \rangle = \delta_{ij} \text{ para todos } i, j = 1, \ldots, n \Leftrightarrow \overline{A}^\top A = I_n.$$

■

Proposição 3.2.21 Seja V um $\mathbb{K}$-espaço vetorial com produto interno. Sejam $B = \{v_1, \ldots, v_n\}$ e $C = \{u_1, \ldots, u_n\}$ bases ortonormais para V. Então, a matriz de mudança de bases M_{BC} é unitária.

Demonstração. Suponha

$$M_{BC} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Então,

$$[v_1]_C = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad [v_2]_C = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \quad \ldots, \quad [v_n]_C = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix}.$$

Ou seja, para cada $j = 1, \ldots, n$, temos que

$$v_j = a_{1j}u_1 + a_{2j}u_2 + \cdots + a_{nj}u_n = \sum_{i=1}^{n} a_{ij}u_i.$$

Como B é base ortonormal, segue que

$$\delta_{ij} = \langle v_j, v_i \rangle = \left\langle \sum_{k=1}^{n} a_{kj}u_k, \sum_{l=1}^{n} a_{li}u_l \right\rangle = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kj}\overline{a_{li}} \langle u_k, u_l \rangle = \sum_{k=1}^{n} a_{kj}\overline{a_{ki}}, \tag{3.8}$$

onde

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}.$$

Por outro lado, como $M_{BC} = (a_{ij})$, temos $\overline{M_{BC}}^{\top} = (b_{ij})$ com $b_{ij} = \overline{a_{ji}}$. Logo

$$(\overline{M_{BC}}^{\top} M_{BC})_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj} = \sum_{k=1}^{n} \overline{a_{ki}}a_{kj} \overset{(3.8)}{=} \delta_{ij},$$

ou seja, $\overline{M_{BC}}^{\top} M_{BC} = I_n$ e, portanto, $\overline{M_{BC}}^{\top} = M_{BC}^{-1}$. ■

3.3 Complemento ortogonal

Definição 3.3.1 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $S \subseteq V$ um subconjunto de V. O *conjunto ortogonal* de S é o definido

$$S^{\perp} = \{v \in V : \langle v, u \rangle = 0, \ \forall\, u \in S\}.$$

Observação 3.3.2 **(i)** Também podemos escrever $S^{\perp} = \{v \in V : \langle u, v \rangle = 0, \ \forall\, u \in S\}$;

(ii) Se $S = \{0\}$ então $S^{\perp} = V$;

Exercício 3.2 Prove as afirmações a seguir.

(i) $S^{\perp}$ é subespaço vetorial de V (mesmo que S não o seja).

(ii) Se S contém uma base para V, então $S^{\perp} = \{0\}$.

Proposição 3.3.3 Considere V um $\mathbb{K}$-espaço vetorial com produto interno. Sejam $W \subseteq V$ um subespaço e $B = \{w_1, \ldots, w_k\}$ um conjunto gerador para W. Então, $v \in W^{\perp}$ se, e somente se, $\langle v, w_i \rangle = 0$ para todo $i = 1, \ldots, k$.

Demonstração. $(\Rightarrow)$ Seja $v \in W^{\perp}$. Como $W^{\perp} = \{v \in V : \langle v, w \rangle = 0, \ \forall\, w \in W\}$, segue que $\langle v, w_i \rangle = 0$, para todo $i = 1, \ldots, k$.

($\Leftarrow$) Seja $w \in W$ um vetor qualquer. Como B gera W, existem escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tais que

$$w = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_k w_k.$$

Se $v \in V$ é tal que $\langle v, w_i \rangle = 0$, para todo $i = 1, \ldots, k$, então

$$\langle w, v \rangle = \left\langle \sum_{j=1}^{k} \alpha_j w_j, v \right\rangle = \sum_{j=1}^{k} \alpha_j \langle w_j, v \rangle = 0$$

e, portanto, $v \in W^{\perp}$. ■

■ **Exemplo 3.3.4** Sejam $V = \mathbb{R}^4$ com produto interno usual e $W = [(1,0,1,1),(1,1,0,1)]$. Determine uma base para $W^{\perp}$.

Solução. Temos que $W^{\perp} = \{(x,y,z,t) \in \mathbb{R}^4 : \langle (x,y,z,t), w \rangle = 0,\ \forall\, w \in W\}$. Como $B = \{w_1, w_2\} = \{(1,0,1,1),(1,1,0,1)\}$ é base de W, basta procurarmos pelos vetores $v \in \mathbb{R}^4$ tais que $\langle v, w_i \rangle = 0$, $i = 1,2$. Se $v = (x,y,z,t) \in W^{\perp}$, então

$$0 = \langle v, w_1 \rangle = \langle (x,y,z,t),(1,0,1,1) \rangle = x + z + t \;\Rightarrow\; x = -z - t,$$

$$0 = \langle v, w_2 \rangle = \langle (x,y,z,t),(1,1,0,1) \rangle = x + y + t \;\Rightarrow\; y = -x - t = z + t - t = z$$

Logo,

$$v = (x,y,z,t) = (-z-t, z, z, t) = z(-1,1,1,0) + t(-1,0,0,1) \in [(-1,1,1,0),(-1,0,0,1)].$$

Ou seja, $W^{\perp} \subseteq [(-1,1,1,0),(-1,0,0,1)]$. Não é difícil ver que $(-1,1,1,0),(-1,0,0,1)$ são ortogonais a w_1 e w_2, portanto $(-1,1,1,0),(-1,0,0,1) \in W^{\perp}$, donde segue que $[(-1,1,1,0),(-1,0,0,1)] \subseteq W^{\perp}$ (pois $W^{\perp}$ é subespaço). Logo, $W^{\perp} = [(-1,1,1,0),(-1,0,0,1)]$ e $C = \{(-1,1,1,0),(-1,0,0,1)\}$ é uma base para $W^{\perp}$. ■

■ **Exemplo 3.3.5** Considere $V = M_{2\times 2}(\mathbb{R})$ com produto interno usual e

$$W = \left\{ \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in V : x + 2y + 3z + 4w = 0 \right\}.$$

Determine uma base para $W^{\perp}$.

Solução. Se $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \in W$, então, $x = -2y - 3z - 4w$. Assim, $A \in W$ se, e somente se,

$$A = \begin{bmatrix} -2y-3z-4w & y \\ z & w \end{bmatrix} = y \begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix} + z \begin{bmatrix} -3 & 0 \\ 1 & 0 \end{bmatrix} + w \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix} \in \left[\begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix} \right].$$

Logo, $W = [A_1, A_2, A_3] = \left[\begin{bmatrix} -2 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix} \right]$.

Agora, se $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in W^{\perp}$ então $\langle B, A_i \rangle = 0$, $i = 1,2,3$. Ou seja,

$$0 = \langle B, A_1 \rangle = -2a + b \;\Rightarrow\; b = 2a,$$

$$0 = \langle B, A_2 \rangle = -3a + c \;\Rightarrow\; c = 3a,$$

$$0 = \langle B, A_3 \rangle = -4a + d \;\Rightarrow\; d = 4a.$$

Logo, $W^{\perp} = \left[\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right]$ e $C = \left\{ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right\}$ é base para $W^{\perp}$. ■

Proposição 3.3.6 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$, com produto interno e $W \subset V$ um subespaço vetorial. Então $V = W \oplus W^{\perp}$.

Demonstração. Primeiramente vamos mostrar que $V = W + W^{\perp}$. Seja $B = \{w_1, \dots, w_k\}$ uma base ortonormal para W. É claro que $W + W^{\perp} \subseteq V$. Por outro lado, se $v \in V$, segue da Proposição 3.2.12 que

$$u = v - \sum_{i=1}^{k} \langle v, w_i \rangle w_i \in W^{\perp}.$$

Logo, $v = \sum_{i=1}^{k} \langle v, w_i \rangle w_i + u \in W + W^{\perp}$ e, portanto, $V = W + W^{\perp}$.

Agora, se $w \in W \cap W^{\perp}$ temos que $w \in W^{\perp}$. Ou seja, $\langle w, v \rangle = 0$ para todo vetor $v \in W$. Em particular, $\langle w, w \rangle = 0$, o que implica em $w = 0$. Logo, $W \cap W^{\perp} = \{0\}$ e, portanto, $V = W \oplus W^{\perp}$. ■

Devido a Proposição 3.3.6, se $S = W$ é um subespaço vetorial, então o conjunto ortogonal $S^{\perp}$ leva um nome especial, definido a seguir.

Definição 3.3.7 Se V é um espaço vetorial com produto interno e $W \subseteq V$ é um subespaço vetorial, então $W^{\perp}$ é chamado de *complemento ortogonal* de W.

O próximo resultado é uma consequência imediata das Proposições 3.3.6 e 2.6.4.

Corolário 3.3.8 Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$, com produto interno. Se $W \subset V$ é um subespaço vetorial, então, $\dim V = \dim W + \dim W^{\perp}$.

Exercício 3.3 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita com produto interno. Se $W \subset V$ é um subespaço, prove que $(W^{\perp})^{\perp} = W$.

3.4 Projeção ortogonal

Antes de introduzirmos o conceito de projeção ortogonal num espaço vetorial com produto interno, vamos relembrar o seu significado em $\mathbb{R}^2$ e $\mathbb{R}^3$, provavelmente visto em Geometria Analítica.

Em $\mathbb{R}^2$, considere uma reta r passando pela origem e $P = (a, b) \in \mathbb{R}^2$ um ponto que não pertence a r. Para calcular a distância de P a r, consideramos o vetor $v = \overrightarrow{OP} = (a, b)$. A projeção ortogonal de v em r (ou em qualquer vetor paralelo à r) é o vetor w, que é paralelo a reta r e tal que $v - w$ é ortogonal à r. Neste caso, $||v - w||$ é a distância de P a r, e é a menor dentre todas as distâncias de P aos pontos de r. Veja a Figura 3.2.

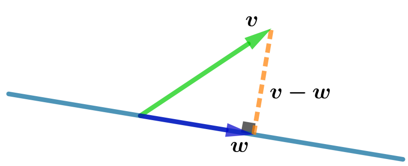


Figura 3.2: Projeção ortogonal de $v \in \mathbb{R}^2$ na reta gerada por w.

Em linguagem de álgebra linear, r é um subespaço vetorial W, v é um vetor que não pertence a W e

$w \in W$ é tal que $||v-w|| = \min\{||v-u|| : u \in W\}$. Também poderia se considerar $v \in W$ (ou $P \in r$), mas nesse caso $w = v$ seria a projeção ortogonal de v em W e $||v-w|| = 0$.

Em $\mathbb{R}^3$ a situação é análoga. Se W é um plano em $\mathbb{R}^3$ passando pela origem e $v = (a,b,c) \in \mathbb{R}^3$, a projeção ortogonal de v em W é o vetor $w \in W$ tal que $v-w$ é ortogonal ao plano W e $||v-w|| = \min\{||v-u|| : u \in W\}$ é a menor distância entre o ponto $P = (a,b,c)$ e um ponto do plano W. Observe que o complemento ortogonal $W^\perp$ neste caso é uma reta (passando pela origem) ortogonal a W e paralela ao vetor $v-w$, isto é $v-w \in W^\perp$. Veja a Figura 3.3.

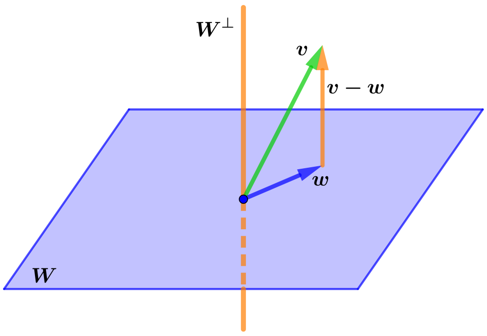


Figura 3.3: Projeção ortogonal de $v \in \mathbb{R}^3$ no subespaço W.

Essas ideias serão generalizadas para um espaço vetorial V com produto interno.

Definição 3.4.1 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno, $W \subseteq V$ um subespaço vetorial e $v \in V$. Dizemos que $w \in W$ é *projeção ortogonal de v sobre W*, e denotamos $w = \text{proj}_W(v)$, se $v-w$ for ortogonal a todo vetor de W, isto é, se $v-w \in W^\perp$.

O resultado a seguir garante a existência e unicidade da projeção ortogonal sobre subespaços de dimensão finita.

Proposição 3.4.2 Seja V um $\mathbb{K}$-espaço vetorial com produto interno e $W \subseteq V$ um subespaço vetorial de V com dimensão finita. Dado $v \in V$ existe um único $w \in W$ tal que $v-w \in W^\perp$.

Demonstração. Existência: Seja $B = \{w_1, \ldots, w_k\}$ uma base ortonormal para W. Segue da Proposição 3.2.12 que, dado $v \in V$,

$$u = v - \sum_{i=1}^{k} \langle v, w_i \rangle w_i \ \in W^\perp.$$

Logo, $w = \sum_{i=1}^{k} \langle v, w_i \rangle w_i \ \in W$ satisfaz $v-w \in W^\perp$ e, portanto, w é uma projeção ortogonal de v em W.

Unicidade: Suponha que $w, \bar{w} \in W$ sejam projeções ortogonais de v em W, ou seja,

$$v-w \in W^\perp \ \text{ e } \ v-\bar{w} \in W^\perp.$$

Note que,

$$||w-\bar{w}||^2 = \langle w-\bar{w}, w-\bar{w} \rangle = \langle w-\bar{w}, w-v+v-\bar{w} \rangle = \langle w-\bar{w}, w-v \rangle + \langle w-\bar{w}, v-\bar{w} \rangle.$$

Como $w-\bar{w} \in W$, $w-v \in W^\perp$ e $v-\bar{w} \in W^\perp$, segue que $||w-\bar{w}||^2 = 0$ e, portanto, $w = \bar{w}$. ■

Observação 3.4.3 Na demonstração da proposição anterior apareceu uma fórmula para a projeção ortogonal, a saber,

$$\text{proj}_W(v) = \sum_{i=1}^{k} \langle v, w_i \rangle w_i,$$

desde que $B = \{w_1, \ldots, w_k\}$ seja uma base ortonormal para W. Caso $B = \{u_1, \ldots, u_k\}$ seja uma base ortogonal para W (mas não necessariamente ortonormal), então

$$\text{proj}_W(v) = \sum_{i=1}^{k} \frac{\langle v, u_i \rangle}{||u_i||^2} u_i.$$

Observação 3.4.4 Note que, no Processo de Ortogonalização de Gram-Schmidt, em cada etapa $k = 2, \ldots, n$, o vetor u_k é exatamente $v_k - \text{proj}_{V_{k-1}}(v_k) \in V_{k-1}^{\perp}$.

■ **Exemplo 3.4.5** Seja $V = \mathbb{R}^4$ com o produto interno usual. Determine a projeção ortogonal de $u = (1, 5, -1, 12)$ no espaço solução do sistema linear homogêneo

$$\begin{cases} x_1 + x_2 + x_4 = 0 \\ 2x_1 + x_2 + x_3 = 0 \end{cases}.$$

Solução. Seja S o conjunto solução do sistema dado. Note que $S = [(-1, 1, 1, 0), (1, -2, 0, 1)]$. Porém, os dois vetores não são ortogonais. Ortogonalizando, obtemos a base ortogonal $B = \{u_1, u_2\} = \{(-1, 1, 1, 0), (0, -1, 1, 1)\}$ (verifique). Segue que,

$$\text{proj}_S(u) = \frac{\langle u, u_1 \rangle}{||u_1||^2} u_1 + \frac{\langle u, u_2 \rangle}{||u_2||^2} u_2 = (-1, -1, 3, 2).$$

■

■ **Exemplo 3.4.6** Considere $V = P_2(\mathbb{R})$ com o produto interno

$$\langle p, q \rangle = p(-2)q(-2) + p(0)q(0) + p(1)q(1).$$

Sejam $p(x) = x^2 - 4$ e $W = [x - 1]$. Determine $\text{proj}_W(p)$.

Solução. Note que $B = \{x - 1\}$ é uma base ortogonal para W. Então,

$$\text{proj}_W(p) = \frac{\langle p, x - 1 \rangle}{||x - 1||^2}.$$

Como

$$||x - 1||^2 = \langle x - 1, x - 1 \rangle = (-3)(-3) + (-1)(-1) + 0 = 10$$

$$\langle p, x - 1 \rangle = 0 + (-4)(-1) + 0 = 4$$

segue que

$$\text{proj}_W(p) = \frac{4}{10}(x - 1) = \frac{2(x - 1)}{5}.$$

■

■ **Exemplo 3.4.7** Seja $V = \mathscr{C}([0, 1], \mathbb{R})$ com produto interno usual. Determine a projeção ortogonal de e^x em $W = [1, x]$.

Solução. Como

$$\langle 1, x \rangle = \int_0^1 1 \cdot x \, dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2} \neq 0$$

seque que $\{1, x\}$ não é uma base ortogonal para W. Ortogonalizando essa base, obtemos $B = \left\{1, x - \frac{1}{2}\right\}$. Assim,

$$\text{proj}_W(e^x) = \frac{\langle e^x, 1\rangle}{||1||^2} 1 + \frac{\langle e^x, x - \frac{1}{2}\rangle}{||x - \frac{1}{2}||^2}\left(x - \frac{1}{2}\right) = (e-1) + 6(3-e)\left(x - \frac{1}{2}\right),$$

ou seja,

$$\text{proj}_W(e^x) = 4e - 10 + (18 - 6e)x.$$

■

■ **Exemplo 3.4.8** Sejam $V = \mathbb{R}^3$ com produto interno usual e $W = [(1,2,-1)]$. Determine a projeçao ortogonal do elemento $u = (4,5,2) \in \mathbb{R}^3$ sobre o subespaço W e sobre o subespaço $W^\perp$.

Solução. A projeção ortogonal de u sobre W é dada por

$$\tilde{u} = \text{proj}_W(u) = \frac{\langle u, w\rangle}{||w||^2} w = (2, 4, -2).$$

Além disso, a projeção ortogonal $\tilde{u}$ é o único vetor de W tal que $u - \tilde{u} \in W^\perp$. Assim como, a projeção ortogonal de u em $W^\perp$ é o único vetor $\bar{u}$ de $W^\perp$ tal que $u - \bar{u} \in (W^\perp)^\perp = W$. Como $u - \tilde{u} \in W^\perp$ e $u - (u - \tilde{u}) = \tilde{u} \in W$, seque que $\bar{u} = u - \tilde{u} = \text{proj}_{W^\perp}(u)$. Logo,

$$\text{proj}_{W^\perp}(u) = u - \tilde{u} = (4,5,2) - (2,4,-2) = (2,1,4).$$

■

O resultado a seguir afirma que a $\text{proj}_W(v)$ é a *melhor aproximação* de $v \in V$ por um vetor em W.

Proposição 3.4.9 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno, $W \subseteq V$ um subespaço vetorial de V e $v \in V$. Dado $\bar{w} \in W$, são equivalentes:

(i) $v - \bar{w} \in W^\perp$;

(ii) $||v - \bar{w}|| < ||v - w||$ para todo $w \in W \setminus \{\bar{w}\}$.

Demonstração. **(i)** $\Rightarrow$ **(ii)** Seja $\bar{w} \in W$ tal que $v - \bar{w} \in W^\perp$. Para $w \in W$, $w \neq \bar{w}$, temos que

$$||v - w||^2 = ||v - \bar{w} + \bar{w} - w||^2 = \langle v - \bar{w} + \bar{w} - w, v - \bar{w} + \bar{w} - w\rangle$$

$$= \langle v - \bar{w}, v - \bar{w}\rangle + \langle v - \bar{w}, \bar{w} - w\rangle + \langle \bar{w} - w, v - \bar{w}\rangle + \langle \bar{w} - w, \bar{w} - w\rangle$$

$$= ||v - \bar{w}||^2 + ||\bar{w} - w||^2 > ||v - \bar{w}||^2,$$

pois $\bar{w} - w \in W \setminus \{0\}$, donde $||\bar{w} - w|| > 0$.

(ii) $\Rightarrow$ **(i)** Seja $\bar{w} \in W$ tal que $||v - \bar{w}|| < ||v - w||$ para todo $w \in W \setminus \{\bar{w}\}$. Suponha, por absurdo, que $v - \bar{w} \notin W^\perp$. Então, existe $u \in W$ tal que $\langle u, v - \bar{w}\rangle \neq 0$. Note que $u \neq 0$. Considere $S = [u, \bar{w}]$. Observe que S é um subespaço de W com $1 \leq \dim S \leq 2$. Tome $\bar{s} \in S$ a projeção ortogonal de v em S, ou seja, $v - \bar{s} \in S^\perp$. Temos $\bar{s} \neq \bar{w}$, pois $\langle v - \bar{w}, u\rangle \neq 0$ e $\langle v - \bar{s}, u\rangle = 0$.

Pela parte (i) $\Rightarrow$ (ii) já demonstrada, $||v-\bar{s}|| < ||v-s||$ para todo $s \in S \setminus \{\bar{s}\}$. Ou seja,

$$||v-\bar{s}|| < ||v-\bar{w}||, \text{ pois } \bar{w} \in S,$$

$$||v-\bar{w}|| < ||v-\bar{s}||, \text{ pois } \bar{s} \in W,$$

o que é uma contradição. Portanto, $v-\bar{w} \in W^{\perp}$. ■

■ **Exemplo 3.4.10** [Coe] Considere $V = P_3(\mathbb{R})$ com o produto interno $\langle p,q\rangle = \int_0^1 p(x)q(x)\,dx$. Determine o polinômio de grau 1 que melhor se aproxima de $p(x) = x^3$.

Solução. Queremos $\bar{q}(x) = a+bx \in P_1(\mathbb{R})$ tal que $||p-\bar{q}|| < ||p-q||$ para todo $q \in P_1(\mathbb{R}) \setminus \{\bar{q}\}$. Pela Proposição 3.4.9, $\bar{q} = \text{proj}_{P_1}(p)$. Como $B = \{1, \sqrt{3}(2x-1)\}$ é base ortonormal de $P_1(\mathbb{R})$ (verifique), segue que

$$\bar{q}(x) = \text{proj}_{P_1(\mathbb{R})}(x^3) = \langle x^3, 1\rangle + \langle x^3, \sqrt{3}(2x-1)\rangle\,\sqrt{3}\,(2x-1) = \frac{1}{4} + \frac{9}{20}(2x-1) = \frac{9x}{10} - \frac{1}{5}.$$

Veja a ilustração de p e $\bar{q}$ na Figura 3.4, considerando o intervalo $[0,1]$.

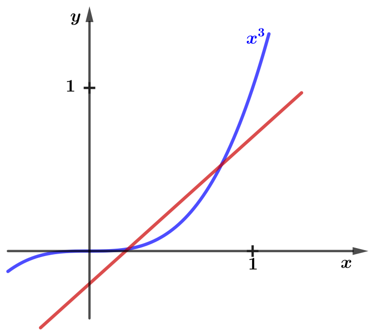


Figura 3.4: Projeção de x^3 em $P_1(\mathbb{R})$ para $x \in [0,1]$.

■

3.5 O Problema de Quadrados Mínimos

O sistema linear

$$\begin{cases} 2x = b_1 \\ 3x = b_2 \\ 4x = b_3 \end{cases} \tag{3.9}$$

possui solução apenas quando $(b_1,b_2,b_3) \in [(2,3,4)]$. Se $(b_1,b_2,b_3) \notin [(2,3,4)]$ esse sistema não possui solução. Entretanto, podemos determinar uma "solução aproximada", isto é, um vetor que minimize o erro cometido.

Definindo $a = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ e $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ podemos escrever (3.9) na forma $ax = b$. Uma solução aproximada pode ser obtida, por exemplo, minimizando a norma do erro cometido. Por conveniência, trabalharemos com a norma Euclidiana ao quadrado, assim

$$E^2(x) = \|ax-b\|^2 = (2x-b_1)^2 + (3x-b_2)^2 + (4x-b_3)^2.$$

Em Cálculo I foi visto que os candidatos a minimizadores de uma função são os seus pontos críticos. Nesse caso, os pontos críticos são aqueles cuja derivada é igual a zero. Como

$$\frac{d}{dx}E^2(x) = 2\cdot 2\,(2x-b_1) + 2\cdot 3(3x-b_2) + 2\cdot 4(4x-b_3),$$

temos

$$\frac{d}{dx}E^2(x) = 0 \quad \Leftrightarrow \quad (2^2+3^2+4^2)x = 2b_1+3b_2+4b_3,$$

ou seja,

$$x = \frac{2b_1+3b_2+4b_3}{2^2+3^2+4^2} = \frac{\langle a,b\rangle}{\|a\|^2}$$

é o único ponto crítico desse problema. Como $\dfrac{d^2}{dx^2}E^2(x) = 2(2^2+3^2+4^2) > 0$, segue que o ponto encontrado é um minimizador do problema.

No caso de um sistema linear com mais de uma variável a situação é bem parecida. Considere

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{3.10}$$

um sistema linear incompatível (ou impossível), com coeficientes em $\mathbb{K}$. Assim, não existe $x\in\mathbb{K}^n$ que seja solução de (3.10). Queremos encontrar uma solução aproximada, $\tilde{x}\in\mathbb{K}^n$, que minimize o erro cometido.

Denotando por $A=(a_{ij})\in M_{m\times n}(\mathbb{K})$ a matriz do sistema (3.10) e $b=(b_i)\in M_{m\times 1}(\mathbb{K})$, o problema de encontrar $\tilde{x}\in\mathbb{K}^n$ tal que $A\tilde{x}$ esteja o mais próximo possível de b, isto é,

$$||A\tilde{x}-b|| < ||Ax-b|| \quad \forall\, x\in\mathbb{K}^n\setminus\{\tilde{x}\},$$

é chamado de *Problema de Quadrados Mínimos*.

Sejam $A_1,\ldots,A_n\in\mathbb{K}^m$ as colunas de A, isto é, $A_i=[a_{1i},a_{2i},\ldots,a_{mi}]^\top$. Se $y=(y_j)\in M_{n\times 1}(\mathbb{K})$ então, pela Proposição 1.1.2**(xiv)**

$$Ay = y_1A_1 + y_2A_2 + \cdots + y_nA_n$$

ou seja, Ay é uma combinação linear das colunas de A. Logo, $Ay\in R(A)=[A_1,A_2,\ldots,A_n]\subseteq\mathbb{K}^m$.

Como queremos encontrar $\tilde{x}$ tal que $A\tilde{x}$ esteja o mais próximo possível de b, basta calcularmos a projeção ortogonal de b sobre o espaço coluna $R(A)$. Note que, dizer que o sistema (3.10) é incompatível, é equivalente a dizer que $b\notin R(A)$. Veja a Figura 3.5.

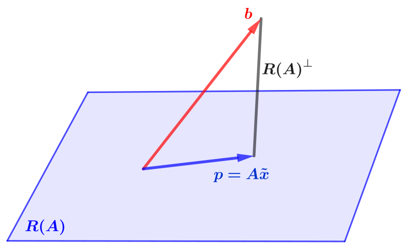


Figura 3.5: Projeção ortogonal de b no espaço coluna de A.

Pela Proposição 3.4.9, se $p = A\tilde{x}$ é a projeção ortogonal de b em $R(A)$, então $b - p \in R(A)^{\perp} = [A_1, A_2, \ldots, A_n]^{\perp}$, ou seja,

$$\langle b - A\tilde{x}, A_i\rangle = 0\,,\ \ \forall\, i = 1, \ldots, n,$$

ou ainda, $\langle b, A_i\rangle = \langle A\tilde{x}, A_i\rangle$, para todo $i = 1, \ldots, n$. Obtemos assim o sistema linear

$$\begin{cases} \langle A_1, A_1\rangle\tilde{x}_1 + \langle A_2, A_1\rangle\tilde{x}_2 + \cdots + \langle A_n, A_1\rangle\tilde{x}_n = \langle b, A_1\rangle \\ \langle A_1, A_2\rangle\tilde{x}_1 + \langle A_2, A_2\rangle\tilde{x}_2 + \cdots + \langle A_n, A_2\rangle\tilde{x}_n = \langle b, A_2\rangle \\ \qquad\qquad\qquad\qquad \vdots \\ \langle A_1, A_n\rangle\tilde{x}_1 + \langle A_2, A_n\rangle\tilde{x}_2 + \cdots + \langle A_n, A_n\rangle\tilde{x}_n = \langle b, A_n\rangle \end{cases} \tag{3.11}$$

que sempre admite solução, pois a projeção ortogonal existe. Usando o produto interno usual de $\mathbb{K}^m$, observe que $\left(\overline{A}^{\top} A\right)_{ij} = \langle A_j, A_i\rangle$, assim podemos reescrever (3.11) na forma matricial

$$\overline{A}^{\top} A\tilde{x} = \overline{A}^{\top} b.$$

Resolvendo esse sistema, obtém-se a melhor aproximação de uma "solução" do sistema $Ax = b$.

Definição 3.5.1 O sistema $\overline{A}^{\top} A\tilde{x} = \overline{A}^{\top} b$ obtido acima é chamado de *sistema de equações normais*.

O *método dos quadrados mínimos* é utilizado quando temos um sistema de equações $Ax = b$ que não possui solução. Neste caso, dizemos que $\tilde{x}$ é uma "solução" se a diferença $||A\tilde{x} - b||^2$ é a menor possível. Essa "solução" é chamada de *solução de quadrados mínimos*.

Como visto anteriormente, a solução de quadrados mínimos $\tilde{x}$ deve satisfazer as equações normais $\overline{A}^{\top} Ax = \overline{A}^{\top} b$. Se as colunas da matriz A forem LIs, a matriz $\overline{A}^{\top} A$ será não singular, e a solução de quadrados mínimos é dada unicamente por $\tilde{x} = \left(\overline{A}^{\top} A\right)^{-1} \overline{A}^{\top} b$. Se $\{A_1, \ldots, A_n\}$ for LD, a melhor aproximação poderá ser escrita de várias maneiras diferentes como combinação linear de $A_1, \ldots, A_n$. Costuma-se escolher aquela de menor norma.

■ **Exemplo 3.5.2** Considere o sistema

$$\begin{cases} \qquad\quad x_1 - x_4 = \ \ 0 \\ -x_2 + x_3 + 2x_4 = -2 \\ x_1 - x_2 + x_3 + x_4 = -1 \end{cases}.$$

Deixamos como exercício a verificação de que esse sistema é impossível. Na forma matricial, esse sistema é dado por $Ax = b$ onde

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 2 \\ 1 & -1 & 1 & 1 \end{bmatrix} \in M_{3\times 4}(\mathbb{R}),\ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in M_{4\times 1}(\mathbb{R})\ \text{ e }\ b = \begin{bmatrix} 0 \\ -2 \\ -1 \end{bmatrix} \in M_{3\times 1}(\mathbb{R}).$$

Buscaremos uma solução para o problema de quadrados mínimos, utilizando as equações normais. Temos,

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

e

$$A^\top b = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -3 \\ -5 \end{bmatrix}.$$

O sistema de equações normais é dado por

$$\begin{bmatrix} 2 & -1 & 1 & 0 \\ -1 & 2 & -2 & -3 \\ 1 & -2 & 2 & 3 \\ 0 & -3 & 3 & 6 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \\ \tilde{x}_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -3 \\ -5 \end{bmatrix}$$

cuja solução geral é $\tilde{x} = \begin{bmatrix} \alpha + 1/3 \\ 4\alpha + 10/3 \\ 2\alpha + 5/3 \\ \alpha \end{bmatrix}$, $\alpha \in \mathbb{R}$. Neste caso, obtém-se infinitas soluções, pois as colunas de A são LD. Não há contradição com a unicidade da projeção ortogonal, pois

$$A\tilde{x} = \begin{bmatrix} 1/3 \\ -5/3 \\ -4/3 \end{bmatrix}, \quad \text{para todo } \alpha \in \mathbb{R},$$

é a (única) projeção ortogonal de b em $R(A)$. Mas a maneira de escrever $\begin{bmatrix} 1/3 \\ -5/3 \\ -4/3 \end{bmatrix}$ como combinação linear das colunas de A que não é única.

Uma melhor aproximação de uma "solução" do sistema $Ax = b$, por exemplo, é $x = \begin{bmatrix} 1/3 \\ 10/3 \\ 5/3 \\ 0 \end{bmatrix}$. ■

Uma solução de quadrados mínimos pode ser utilizada para descrever uma tendência nos dados, que poderá ser usada para fazer previsões.

■ **Exemplo 3.5.3** O custo para produzir um certo livro é dado por $c(t) = a + bt$, com a sendo o custo de edição e tipografia e b o custo de impressão e encadernação (para cada livro). Determine a função custo quando foi observado que:

# livros	10	20	50	100
custo (R$)	1200,00	1500,00	2000,00	2700,00

Solução. Estamos interessados em resolver o sistema linear

$$\begin{cases} \text{a+10b = 1200} \\ \text{a+20b = 1500} \\ \text{a+50b = 2000} \\ \text{a+100b = 2700} \end{cases}$$

ou, na forma matricial, $Ax = b$ com

$$A = \begin{bmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 50 \\ 1 & 100 \end{bmatrix} \in M_{4\times 2}(\mathbb{R}), \; x = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 \;\text{ e }\; b = \begin{bmatrix} 1200 \\ 1500 \\ 2000 \\ 2700 \end{bmatrix} \in \mathbb{R}^4.$$

Como esse sistema linear não admite solução (verifique), vamos buscar uma solução para o problema de quadrados mínimos

$$\begin{array}{ll} \text{Minimizar} & ||Ax-b|| \\ x \in \mathbb{R}^2 & \end{array}$$

utilizando as equaçẽs normais. Temos

$$A^\top A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 10 & 20 & 50 & 100 \end{bmatrix} \begin{bmatrix} 1 & 10 \\ 1 & 20 \\ 1 & 50 \\ 1 & 100 \end{bmatrix} = \begin{bmatrix} 4 & 180 \\ 180 & 13000 \end{bmatrix}$$

e

$$A^\top b = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 10 & 20 & 50 & 100 \end{bmatrix} \begin{bmatrix} 1200 \\ 1500 \\ 2000 \\ 2700 \end{bmatrix} = \begin{bmatrix} 7400 \\ 412000 \end{bmatrix}.$$

O sistema de equações normais é dado por

$$\begin{bmatrix} 4 & 180 \\ 180 & 13000 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 7400 \\ 412000 \end{bmatrix}$$

cuja solução é $\tilde{x} = \begin{bmatrix} 1124,5 \\ 16,12 \end{bmatrix}$. Ou seja, uma aproximação para a função custo é

$$c(t) = 1124,5 + 16,12t.$$

■

O caso geral está descrito no próximo exemplo.

■ **Exemplo 3.5.4 — Ajuste Linear.** Dados m pontos não colineares $(t_1, y_1), \ldots, (t_m, y_m)$, encontre a reta $r : y(t) = a + bt$ que melhor ajusta esses pontos.

Solução. Se todos os pontos pertencessem à reta r, teríamos

$$\begin{array}{lcl} (t_1, y_1) \in r & \Rightarrow & a + bt_1 = y_1 \\ (t_2, y_2) \in r & \Rightarrow & a + bt_2 = y_2 \\ & \vdots & \\ (t_m, y_m) \in r & \Rightarrow & a + bt_m = y_m \end{array}$$

obtendo assim o sistema linear $Ax = b$, com

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \in M_{m\times 2}(\mathbb{R}), \ \ x = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 \ \text{ e } \ b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m.$$

O sistema de equações normais torna-se

$$\begin{bmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m t_i y_i \end{bmatrix}.$$

■

■ **Exemplo 3.5.5 — Ajuste Polinomial.** Dados m pares de números $(t_1, y_1), \dots, (t_m, y_m)$, encontre o polinômio de grau n que melhor ajusta y_i como função de t_i.

Solução. Seja $p(t) = a_0 + a_1 t + \cdots + a_n t^n$ o polinômio procurado. Se todos os pontos pertencessem ao gráfico dessa função polinomial, teríamos

$$\begin{aligned} a_0 + a_1 t_1 + \cdots + a_n t_1^n &= y_1 \\ a_0 + a_1 t_2 + \cdots + a_n t_2^n &= y_2 \\ &\vdots \\ a_0 + a_1 t_m + \cdots + a_n t_m^n &= y_m \end{aligned}$$

obtendo assim o sistema linear $Ax = b$, com

$$A = \begin{bmatrix} 1 & t_1 & \cdots & t_1^n \\ 1 & t_2 & \cdots & t_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & \cdots & t_m^n \end{bmatrix} \in M_{m\times(n+1)}(\mathbb{R}), \quad x = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^{n+1} \text{ e } b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m.$$

Para encontrar os coeficientes $a_0, \dots, a_n$, basta resolver o sistema de equações normais

$$\begin{bmatrix} m & \sum_{i=1}^m t_i & \cdots & \sum_{i=1}^m t_i^n \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 & \cdots & \sum_{i=1}^m t_i^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m t_i^n & \sum_{i=1}^m t_i^{n+1} & \cdots & \sum_{i=1}^m t_i^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m t_i y_i \\ \vdots \\ \sum_{i=1}^m t_i^n y_i \end{bmatrix}.$$

■

A Figura 3.6 ilustra um exemplo com 10 pontos e polinômios de diferentes graus. Intuitivamente podemos pensar que quanto maior o grau do polinômio, melhor será a aproximação obtida. O que não é verdade. Ao resolvermos esse problema, dificuldades númericas começam a aparecer e a aproximação deixa de ser precisa, veja a Figura 3.7. Não discutiremos aqui os fenômentos que levam a esses problemas, o estudante interessado encontrará maiores informações em um curso de Álgebra Linear Numérica ou Computacional.

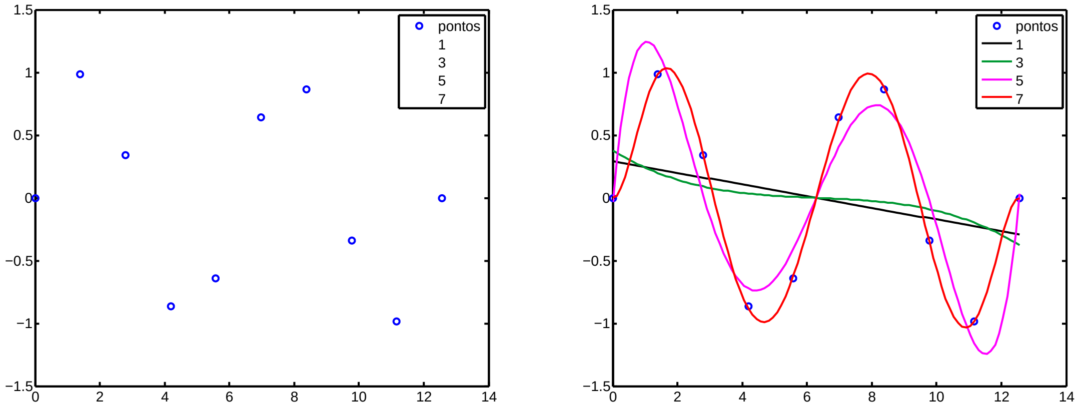


Figura 3.6: Ajuste polinomial usando polinômios de grau 1 (em preto), grau 3 (em verde), grau 5 (em rosa) e grau 7 (em vermelho).

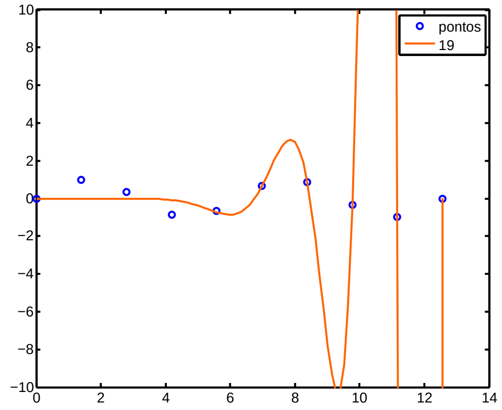


Figura 3.7: Ajuste polinomial usando um polinômio de grau 19.

No exemplo anterior usamos um polinômio para ajustar um conjunto de dados. Mas poderíamos ter utilizado qualquer função.

Dados um conjunto de pontos $(t_1, y_1), \ldots, (t_m, y_m)$ e um conjunto de funções $f_1(t), \ldots, f_n(t)$, podemos pensar em encontrar a melhor aproximação

$$y = \sum_{j=1}^{n} x_j f_j(t).$$

Para tanto, basta utilizarmos a formulação do Problema de Quadrados Mínimos vista nos exemplos anteriores.

Capítulo 4

Transformações lineares

Neste capítulo, estudaremos as funções entre espaços vetoriais. Assim como no cálculo estuda-se as funções contínuas, deriváveis, integráveis, aqui em Álgebra Linear não nos interessam funções quaisquer, mas sim aquelas que "preservam" as operações de espaço vetorial: soma e produto por escalar. Essas funções serão chamadas de transformações lineares. Num primeiro momento, estudaremos propriedades gerais de transformações lineares e veremos como relacionar espaços vetoriais distintos. Em seguida, o foco serão funções de um espaço vetorial nele mesmo, que são chamadas de operadores lineares.

A discussão apresentada nesse capítulo foi motivada pelos livros [Axl], [Coe] e [Nic].

4.1 Conceitos básicos

Definição 4.1.1 Sejam U e V espaços vetoriais sobre o mesmo corpo $\mathbb{K}$. Uma função $T : U \to V$ é uma *transformação linear* se satisfaz:

(i) $T(u_1 + u_2) = T(u_1) + T(u_2)$, para todos $u_1, u_2 \in U$;

(ii) $T(\lambda u) = \lambda T(u)$, para todos $u \in U, \lambda \in \mathbb{K}$.

Se $U = V$, $T : U \to U$ também é chamada de *operador linear* em V.

Não é difícil verificar que, $T : U \to V$ é transformação linear se, e somente se,

$$\boxed{T(\lambda u_1 + u_2) = \lambda T(u_1) + T(u_2), \text{ para todos } u_1, u_2 \in U,\ \lambda \in \mathbb{K}.}$$

Observação 4.1.2 Na definição anterior, apesar de usar a mesma notação, a soma em $u_1 + u_2$ não é a mesma de $T(u_1) + T(u_2)$, se os espaços vetoriais U e V são distintos. O mesmo vale para a multiplicação por escalar. Ficará claro pelo contexto qual é a operação em cada caso.

■ **Exemplo 4.1.3** Sejam U e V dois $\mathbb{K}$-espaços vetoriais quaisquer. A função nula $T : U \to V$ definida por $T(u) = 0$, para todo $u \in U$, é uma transformação linear (verifique).

A função identidade $\mathrm{Id} : U \to U$, definida por $\mathrm{Id}(u) = u$, para todo $u \in U$, também é uma transformação linear. ■

■ **Exemplo 4.1.4** A função $T:\mathbb{K}^3\to M_2(\mathbb{K})$ dada por $T(a,b,c)=\begin{bmatrix} a+2b & 0\\ 0 & 3c-b\end{bmatrix}$ é uma transformação linear.

Prova. Dados $u_1=(a_1,b_1,c_1), u_2=(a_2,b_2,c_2)\in\mathbb{K}^3$, temos

$$T(u_1+u_2)=\begin{bmatrix} a_1+a_2+2(b_1+b_2) & 0\\ 0 & 3(c_1+c_2)-(b_1-b_2)\end{bmatrix}$$

$$=\begin{bmatrix} a_1+2b_1 & 0\\ 0 & 3c_1-b_1\end{bmatrix}+\begin{bmatrix} a_2+2b_2 & 0\\ 0 & 3c_2-b_2\end{bmatrix}=T(u_1)+T(u_2).$$

Também, dados $u=(a,b,c)\in\mathbb{K}^3$ e $\lambda\in\mathbb{K}$, temos

$$T(\lambda u)=\begin{bmatrix} \lambda a+2\lambda b & 0\\ 0 & 3\lambda c-\lambda b\end{bmatrix}=\begin{bmatrix} \lambda(a+2b) & 0\\ 0 & \lambda(3c-b)\end{bmatrix}=\lambda\begin{bmatrix} a+2b & 0\\ 0 & 3c-b\end{bmatrix}=\lambda T(u).$$

■

■ **Exemplo 4.1.5** O operador derivação $D:P(\mathbb{R})\to P(\mathbb{R})$ dado por

$$D(a_0+a_1x+\cdots+a_nx^n)=(a_0+a_1x+\cdots+a_nx^n)'=a_1+2a_2x+\cdots+na_nx^{n-1}$$

é linear.

Prova. Sabemos do Cálculo que, se $f,g:\mathbb{R}\to\mathbb{R}$ são funções deriváveis (o que é válido para funções polinomiais), então $(f+g)'=f'+g'$ e $(\lambda f)'=\lambda f'$ para todo $\lambda\in\mathbb{R}$. Portanto, dados $p,q\in P(\mathbb{R})$ e $\lambda\in\mathbb{R}$, temos

$$D(\lambda p+q)=\lambda p'+q'=\lambda D(p)+D(q).$$

Logo, D é um operador linear em $P(\mathbb{R})$.

O operador derivação também poderia ser definido no espaço mais geral das funções $f:\mathbb{R}\to\mathbb{R}$ deriváveis. ■

■ **Exemplo 4.1.6** Segue da Proposição 1.1.4 que a função $T:M_n(\mathbb{K})\to\mathbb{K}$ dada por $T(A)=\operatorname{tr}(A)$ é uma transformação linear. ■

■ **Exemplo 4.1.7** A função $T:\mathbb{R}^2\to\mathbb{R}$ dada por $T(x,y)=x^2+2y$, não é uma transformação linear, pois, por exemplo, $T(1,0)=1$, mas $T(2,0)=4\neq 2=2T(1,0)$. ■

■ **Exemplo 4.1.8** Considere $V=\mathbb{R}^{\mathbb{N}}$ o espaço de sequências. A função $T:V\to V$ dada por $T(x_1,x_2,x_3,\dots)=(x_2,x_3,x_4,\dots)$ é um transformação linear em V (verifique). ■

■ **Exemplo 4.1.9** Seja $a\in\mathbb{K}$ fixo. A função $T_a:\mathbb{K}\to\mathbb{K}$ definida por $T(x)=ax$ para todo $x\in\mathbb{K}$ é uma transformação linear (verifique).

Por outro lado, supondo $T:\mathbb{K}\to\mathbb{K}$ uma transformação linear, para todo $x\in\mathbb{K}$ temos

$$T(x)=T(x\cdot 1)=xT(1).$$

Tomando $a=T(1)$, segue que $T=T_a$. Portanto, todos os operadores lineares em $\mathbb{K}$ são da forma T_a, $a\in\mathbb{K}$. ■

De uma modo mais geral, temos o seguinte exemplo.

■ **Exemplo 4.1.10** Fixe $A \in M_{m\times n}(\mathbb{K})$. Considere a função $T_A : \mathbb{K}^n \to \mathbb{K}^m$ dada por $T_A(x) = Ax$. Devido as propriedades algébricas de matrizes (Proposição 1.1.2), T_A é uma transformação linear.

Reciprocamente, se $T : \mathbb{K}^n \to \mathbb{K}^m$ é uma transformação linear, então existe $A \in M_{m\times n}(\mathbb{K})$ tal que $T = T_A$, isto é, $T(x) = Ax$ para todo $x \in \mathbb{K}^m$.

Prova. Considere $B = \{e_1, e_2 \ldots, e_n\}$ a base canônica de $\mathbb{K}^n$. Escreva

$$T(e_1) = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(e_2) = \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots \quad , \quad T(e_n) = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Assim, a matriz $A = (a_{ij}) \in M_{m\times n}(\mathbb{K})$ tem como j-ésima coluna $T(e_j)$. Dado $x = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, temos $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$, logo

$$\begin{aligned} T(x) &= T(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1T(e_1) + x_2T(e_2) + \cdots + x_nT(e_n) \\ &= x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = Ax, \end{aligned}$$

onde a última igualdade segue da Proposição 1.1.2**(xiv)**. Além disso, concluímos que $\operatorname{Im} T = R(A)$ é o espaço coluna de A (Exemplo 2.3.6). ■

■ **Exemplo 4.1.11 — Espaço das transformações lineares.** Sejam U e V espaços vetoriais sobre $\mathbb{K}$. Se $S : U \to V$ e $T : U \to V$ são transformações lineares e $\lambda \in \mathbb{K}$, então as funções:

- $S + T : U \to V$ definida por $(S+T)(u) = S(u) + T(u)$ para todo $u \in U$;
- $\lambda S : U \to V$ definida por $(\lambda S)(u) = \lambda S(u)$ para todo $u \in U$;

também são transformações lineares (verifique). Assim, se $\mathscr{L}(U,V)$ é o conjunto de todas as transformações lineares de U em V, temos operações soma e multiplicação por escalar definidas nesse conjunto. Segue das propriedades de soma e multiplicação por escalar em V, que, com essas operações, $\mathscr{L}(U,V)$ é um $\mathbb{K}$-espaço vetorial, onde o vetor nulo é a transformação linear nula. ■

A proposição a seguir descreve algumas propriedades mais imediatas da definição de transformação linear. Sua demonstração será deixada como exercício para o leitor.

Proposição 4.1.12 — Propridades básicas. Sejam U e V espaços vetoriais sobre $\mathbb{K}$ e $T : U \to V$ uma transformação linear. Então,

(i) $T(0_U) = 0_V$

(ii) $T(-u) = -T(u)$, para todo $u \in U$

(iii) Dado $k \in \mathbb{N}$, $T\left(\sum_{i=1}^{k} \alpha_i u_i\right) = \sum_{i=1}^{k} \alpha_i T(u_i)$, para todos $\alpha_i \in \mathbb{K}$, $u_i \in U$.

■ **Exemplo 4.1.13** A função $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x,y) = (x+1, y+1)$ não é uma transformação linear, pois $T(0,0) = (1,1) \neq (0,0)$. ■

Teorema 4.1.14 Sejam U e V $\mathbb{K}$-espaços vetoriais. Se $B = \{u_1, \ldots, u_n\}$ é base de U e $\{v_1, \ldots, v_n\} \subseteq V$ é um subconjunto qualquer, então existe uma única transformação linear $T : U \to V$ tal que $T(u_i) = v_i$, $i = 1, \ldots, n$.

Demonstração. Primeiro, mostremos a existência de T. Seja $u \in U$. Como B é base, existem únicos $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ tais que

$$u = \alpha_1 u_1 + \cdots + \alpha_n u_n.$$

Defina $T : U \to V$ por

$$T(u) = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Segue da unicidade dos escalares $\alpha_1, \ldots, \alpha_n$, que cada $u \in U$ possui uma única imagem $T(u)$, ou seja, T está bem definida. Além disso, pela definição de T, $T(u_i) = v_i$ para todo $i = 1, \ldots, n$ (neste caso $\alpha_i = 1$ e $\alpha_j = 0$ para $i \neq j$).

Mostremos que T é uma transformação linear. Sejam $u = \sum_{i=1}^{n} \alpha_i u_i$, $w = \sum_{i=1}^{n} \beta_i u_i \in U$ (com $\alpha_i, \beta_i \in \mathbb{K}$) e $\lambda \in \mathbb{K}$. Então,

$$\lambda u + w = \sum_{i=1}^{n} \lambda \alpha_i u_i + \sum_{i=1}^{n} \beta_i u_i = \sum_{i=1}^{n} (\lambda \alpha_i + \beta_i) u_i.$$

Logo, pela definição de T,

$$T(\lambda u + w) = \sum_{i=1}^{n} (\lambda \alpha_i + \beta_i) v_i = \lambda \sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{n} \beta_i v_i = \lambda T(u) + T(w).$$

Portanto, T é uma transformação linear que satisfaz $T(u_i) = v_i$ para todo $i = 1, \ldots, n$.

Provaremos agora a unicidade de T. Suponha que $S : U \to V$ é uma transformação linear que satisfaz $S(u_i) = v_i$ para todo $i = 1, \ldots, n$. Dado $u = \sum_{i=1}^{n} \alpha_i u_i \in U$, como S é linear, temos

$$S(u) = S\left(\sum_{i=1}^{n} \alpha_i u_i\right) = \sum_{i=1}^{n} \alpha_i S(u_i) = \sum_{i=1}^{n} \alpha_i v_i = T(u).$$

Como $u \in U$ é arbitrário, segue que $S = T$. ∎

Observação 4.1.15 Devido ao Teorema 4.1.14, se U tem dimensão finita, toda transformação linear $T : U \to V$ fica unicamente determinada ao conhecermos a imagem de T numa base $\{u_1, \ldots, u_n\}$ de U. Muitas vezes, ao apresentar uma transformação linear, falaremos: seja T a transformação linear tal que $T(u_i) = v_i$. Neste caso, considere a única extensão linear possível, de $T|_B : B \to V$ para $T : U \to V$, dada pelo teorema anterior.

■ **Exemplo 4.1.16** Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ uma transformação linear tal que $T(1,1) = (2,1,2)$ e $T(0,1) = (1,1,2)$. Determine explicitamente $T(x,y)$.

Solução. Primeiramente, observe que $u_1 = (1,1)$ e $u_2 = (0,1)$ formam uma base de $\mathbb{R}^2$. Para descrever $T(x,y)$, precisamos escrever o vetor genérico $(x,y) \in \mathbb{R}^2$ como combinação linear da base $\{u_1, u_2\}$. Temos,

$$(x,y) = x(1,1) + (y-x)(0,1) \qquad \text{para todos } x, y \in \mathbb{R}.$$

Logo,

$$T(x,y) = xT(1,1) + (y-x)T(0,1) = x(2,1,2) + (y-x)(1,1,2) = (x+y, y, 2y).$$

■

■ **Exemplo 4.1.17** Sabendo que T é uma transformação linear, complete o elemento que deve estar no lugar da interrogação

$$\begin{aligned} T:\ P_2(\mathbb{R}) &\longrightarrow P_3(\mathbb{R}) \\ 1+x+x^2 &\longmapsto 2 \\ 1-2x &\longmapsto x^2-3 \\ x^2-2x &\longmapsto x^3-x \\ x+1 &\longmapsto ? \end{aligned}$$

Solução. Precisamos escrever $x+1$ como combinação linear de $1+x+x^2, 1-2x$ e x^2-2x. Escrevendo

$$1+x=\alpha(1+x+x^2)+\beta(1-2x)+\gamma(x^2-2x)=(\alpha+\beta)\cdot 1+(\alpha-2\beta-2\gamma)x+(\alpha+\gamma)x^2,$$

obtemos o sistema linear

$$\begin{cases} \alpha+\beta=1 \\ \alpha-2\beta-2\gamma=1 \\ \alpha+\gamma=0 \end{cases},$$

cuja solução é $\alpha=\frac{3}{5}$, $\beta=\frac{2}{5}$ e $\gamma=-\frac{3}{5}$. Portanto,

$$T(x+1)=\frac{3}{5}T(1+x+x^2)+\frac{2}{5}T(1-2x)-\frac{3}{5}T(x^2-2x)=\frac{3}{5}\cdot 2+\frac{2}{5}(x^2-3)-\frac{3}{5}(x^3-x),$$

ou seja,

$$?=T(x+1)=\frac{1}{5}(-3x^3+2x^2+3x).$$

■

Introduzimos a seguir elementos importantes relacionados às transformações lineares.

Definição 4.1.18 Sejam U e V espaços vetoriais sobre $\mathbb{K}$ e $T: U \to V$ uma transformação linear. O conjunto

$$\{u \in U : T(u)=0\}$$

é chamado de *núcleo* de T e será denotado por $\ker T$ ou $\operatorname{Nuc} T$. O conjunto

$$\{v \in V : v=T(u) \text{ para algum } u \in U\}$$

é chamado de *imagem* de T e será denotado por $\operatorname{Im} T$.

Proposição 4.1.19 Sejam U e V espaços vetoriais sobre $\mathbb{K}$ e $T: U \to V$ uma transformação linear. Então,

(i) $\ker T$ é subespaço vetorial de U;

(ii) $\operatorname{Im} T$ é subespaço vetorial de V;

(iii) T é injetora se, e somente se, $\ker T=\{0\}$;

(iv) se B é base de U, então $T(B)=\{T(u): u \in B\}$ gera $\operatorname{Im} T$.

Demonstração. **(i)** Primeiro observe que $0 \in \ker T$ já que, sendo T transformação linear, ocorre $T(0)=0$. Agora, dados $\lambda \in \mathbb{K}$ e $u_1, u_2 \in \ker T$, temos $T(u_1)=0=T(u_2)$ e

$$T(u_1+u_2)=T(u_1)+T(u_2)=0+0=0 \ \Rightarrow\ u_1+u_2 \in \ker T$$

e,
$$T(\lambda u_1) = \lambda T(u_1) = \lambda \cdot 0 = 0 \;\Rightarrow\; \lambda u_1 \in \ker T.$$
Portanto, $\ker T$ é subespaço de U.

(ii) Observe que $0 = T(0) \in \operatorname{Im} T$. Sejam $v_1 = T(u_1), v_2 = T(u_2) \in \operatorname{Im} T$, com $u_1, u_2 \in U$, e $\lambda \in \mathbb{K}$. Então,
$$v_1 + v_2 = T(u_1) + T(u_2) = T(u_1 + u_2) \in \operatorname{Im} T$$
e
$$\lambda v_1 = \lambda T(u_1) = T(\lambda u_1) \in \operatorname{Im} T.$$
Portanto, $\operatorname{Im} T$ é subespaço de V.

(iii) Suponha que T é injetora. Mostremos que $\ker T = \{0\}$. A inclusão $\{0\} \subseteq \ker T$ ocorre sempre. Seja, agora, $u \in \ker T$. Então, $T(u) = 0 = T(0)$. Como T é injetora, segue que $u = 0$. Portanto, $\ker T = \{0\}$.

Reciprocamente, suponha $\ker T = \{0\}$. Mostremos que T é injetora. Sejam $u_1, u_2 \in U$ tais que $T(u_1) = T(u_2)$. Então,
$$T(u_1 - u_2) = T(u_1) - T(u_2) = 0 - 0 = 0 \;\Rightarrow\; u_1 - u_2 \in \ker T = \{0\} \;\Rightarrow\; u_1 - u_2 = 0 \;\Rightarrow\; u_1 = u_2.$$

(iv) Como $\operatorname{Im} T$ é subespaço e $T(B) \subseteq \operatorname{Im} T$, segue que $[T(B)] \subseteq \operatorname{Im} T$. Mostremos, agora, que $\operatorname{Im} T \subseteq [T(B)]$. Seja $v \in \operatorname{Im} T$. Então, existe $u \in U$ tal que $T(u) = v$. Sendo B uma base de U, existem $u_1, \ldots, u_n \in B$, $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tais que
$$u = \lambda_1 u_1 + \cdots + \lambda_n u_n.$$
Daí,
$$T(u) = T(\lambda_1 u_1 + \cdots + \lambda_n u_n) = \lambda_1 T(u_1) + \cdots + \lambda_n T(u_n) \in [T(B)].$$
Portanto, $[T(B)] = \operatorname{Im} T$.

■

■ **Exemplo 4.1.20** Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x,y,z) = (x-y, z, x+z-y)$. Determine $\ker T$ e $\operatorname{Im} T$.

Solução. Começamos com o núcleo de T:
$$T(x,y,z) = 0 \;\Leftrightarrow\; (x-y, z, x+z-y) = (0,0,0) \;\Leftrightarrow\; x-y = z = x+z-y = 0 \;\Leftrightarrow\; x = y \text{ e } z = 0.$$
Assim, $(x,y,z) \in \ker T$ se, e somente se,
$$(x,y,z) = (x,x,0) = x(1,1,0) \in [(1,1,0)].$$
Portanto, $\ker T = [(1,1,0)]$. Para a imagem de T, considere a base $B = \{e_1, e_2, e_3\}$ canônica de $\mathbb{R}^3$. Pela Proposição 4.1.19, $T(B)$ gera $\operatorname{Im} T$. Temos
$$T(e_1) = (1,0,1), \quad T(e_2) = (-1,0,-1), \quad T(e_3) = (0,1,1).$$
Logo, $\operatorname{Im} T = [(1,0,1), (-1,0,-1), (0,1,1)]$. Como $(-1,0,-1) = -(1,0,1)$, podemos eliminar esse vetor do conjunto gerador sem prejuízo. Assim,
$$\operatorname{Im} T = [(1,0,1), (0,1,1)].$$
Esse último conjunto gerador é LI, portanto forma uma base de $\operatorname{Im} T$. ■

■ **Exemplo 4.1.21** Considere a transformação linear $T:\mathbb{C}^2\to\mathbb{R}^3$ dada por $T(a+bi,c+di)=(2a+c+d,a+b+c,-2b-c+d)$, onde $a,b,c,d\in\mathbb{R}$ e $\mathbb{C}^2$ é visto como $\mathbb{R}$-espaço vetorial. Determine $\ker T$ e $\operatorname{Im} T$.

Solução. Começaremos por $\ker T$. Temos

$$T(a+bi,c+di)=0 \Leftrightarrow (2a+c+d,a+b+c,-2b-c+d)=(0,0,0) \Leftrightarrow \begin{cases} 2a+c+d=0 \\ a+b+c=0 \\ -2b-c+d=0 \end{cases}.$$

A resolução desse sistema linear homogêneo com 3 equações e 4 variáveis será deixada como exercício. O conjunto solução pode ser expressado na forma

$$\begin{aligned}\{(a,b,-a-b,b-a):a,b\in\mathbb{R}\}&=\{a(1,0,-1,-1)+b(0,1,-1,1):a,b\in\mathbb{R}\}\\ &=[(1,0,-1,-1),(0,1,-1,1)].\end{aligned}$$

Portanto,

$$\ker T=[(1,-1-i),(i,-1+i)].$$

Para obter um conjunto gerador para $\operatorname{Im} T$, encontraremos $T(B)$, onde $B=\{(1,0),(i,0),(0,1),(0,i)\}$ é a base canônica de $\mathbb{C}^2$ como $\mathbb{R}$-espaço vetorial. Temos,

$$T(1,0)=(2,1,0),\quad T(i,0)=(0,1,-2),\quad T(0,1)=(1,1,-1),\quad T(0,i)=(1,0,1).$$

Daí,

$$\operatorname{Im} T=[(2,1,0),(0,1,-2),(1,1,-1),(1,0,1)]. \tag{4.1}$$

Essa descrição já determina o subespaço $\operatorname{Im} T$. Se quiséssemos encontrar uma base para $\operatorname{Im} T$, poderíamos, por exemplo, utilizar as ideias da Seção 2.8. O Teorema do Núcleo e Imagem, que será apresentado a seguir, garante que $\dim(\operatorname{Im} T)=2$. Portanto, são necessários apenas 2 dos 4 vetores em (4.1) para gerar $\operatorname{Im} T$. ■

Teorema 4.1.22 — do núcleo e imagem. Sejam U e V espaços vetoriais sobre $\mathbb{K}$, com $\dim(U)=n$ (finita) e $T:U\to V$ uma transformação linear. Então,

$$\dim U=\dim(\ker T)+\dim(\operatorname{Im} T).$$

Demonstração. Considere, inicialmente, $\ker T\neq\{0\}\neq\operatorname{Im} T$. Sejam $B_1=\{u_1,\dots,u_k\}$ uma base de $\ker T$ e $B_2=\{T(w_1),\dots,T(w_m)\}$ uma base de $\operatorname{Im} T$. Mostremos que $B=\{u_1,\dots,u_k,w_1,\dots,w_m\}$ é base de U. Para mostrar que B é LI, considere $\alpha_1,\dots,\alpha_k,\beta_1,\dots,\beta_m\in\mathbb{K}$ tais que

$$\alpha_1u_1+\cdots+\alpha_ku_k+\beta_1w_1+\cdots+\beta_mw_m=0.$$

Assim,

$$T(\alpha_1u_1+\cdots+\alpha_ku_k+\beta_1w_1+\cdots+\beta_mw_m)=T(0)=0.$$

Como $\alpha_1u_1+\cdots+\alpha_ku_k\in\ker T$, temos $T(\alpha_1u_1+\cdots+\alpha_ku_k)=0$. Logo

$$T(\alpha_1u_1+\cdots+\alpha_ku_k+\beta_1w_1+\cdots+\beta_mw_m)=\beta_1T(w_1)+\cdots+\beta_mT(w_m)=0.$$

Mas B_2 é base, portanto é LI e segue que

$$\beta_1=\cdots=\beta_m=0.$$

Daí,

$$\alpha_1 u_1 + \cdots + \alpha_k u_k = \alpha_1 u_1 + \cdots + \alpha_k u_k + \beta_1 w_1 + \cdots + \beta_m w_m = 0.$$

Como B_1 é base, é LI, logo segue que

$$\alpha_1 = \cdots = \alpha_k = 0.$$

Portanto, B é LI.

Mostremos agora que $[B] = U$. A inclusão $[B] \subseteq U$ é imediata. Para provar a inclusão contrária, seja $u \in U$. Como $T(u) \in \operatorname{Im} T = [B_2]$, existem $\beta_1, \ldots, \beta_m \in \mathbb{K}$ tais que

$$T(u) = \beta_1 T(w_1) + \cdots + \beta_m T(w_m) = T(\beta_1 w_1 + \cdots + \beta_m w_m).$$

Assim,

$$T(u - (\beta_1 w_1 + \cdots + \beta_m w_m)) = T(u) - T(\beta_1 w_1 + \cdots + \beta_m w_m) = 0 \Rightarrow u - (\beta_1 w_1 + \cdots + \beta_m w_m) \in \ker T.$$

Mas $\ker T = [B_1]$, logo, existem $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tais que

$$u - (\beta_1 w_1 + \cdots + \beta_m w_m) = \alpha_1 u_1 + \cdots + \alpha_k u_k \Rightarrow u = \alpha_1 u_1 + \cdots + \alpha_k u_k + \beta_1 w_1 + \cdots + \beta_m w_m \in [B].$$

Portanto B gera U, logo é base de U, e assim concluímos o teorema.

Se $\ker T = \{0\}$, prova-se analogamente que, se $B_2 = \{T(w_1), \ldots, T(w_m)\}$ é uma base de $\operatorname{Im} T$, então $B = \{w_1, \ldots, w_m\}$ é base de U.

Se $\operatorname{Im} T = \{0\}$, então $\dim(\operatorname{Im} T) = 0$ e $T(u) = 0$ para todo $u \in U$, donde segue que $\ker T = U$. ■

Observação 4.1.23 Se U e V possuem dimensão finita, segue do Teorema 4.1.22 que:

(i) se $\dim U > \dim V$, então não existe $T : U \to V$ linear e injetora.

(ii) se $\dim U < \dim V$, então não existe $T : U \to V$ linear e sobrejetora.

■ **Exemplo 4.1.24** Considere a (única) transformação linear $T : \mathbb{R}^3 \to P_3(\mathbb{R})$ tal que $T(1,0,1) = 2 + x^2 + x^3$, $T(0,1,0) = 1 + x^2$, $T(0,0,1) = x^2 - x^3$. Determine $\dim(\ker T)$ e $\dim(\operatorname{Im} T)$.

Solução. É fácil ver que $B = \{(1,0,1),(0,1,0),(0,0,1)\}$ é uma base de $\mathbb{R}^3$. Portanto, $T(B) = \{2 + x^2 + x^3, 1 + x^2, x^2 - x^3\}$ gera $\operatorname{Im} T$. Vejamos se $T(B)$ é LI. Colocando as coordenadas de seus vetores em relação à base canônica de $P_3(\mathbb{R})$ numa matriz (já trocando de posição as duas primeiras linhas) e escalonando, obtemos

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \overset{(L_2 - 2L_1 \leftarrow L_2)}{\sim} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \overset{(L_3 + L_2 \leftarrow L_3)}{\sim} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ou seja, $T(B)$ é LD e o terceiro vetor $x^2 - x^3$ pode ser obtido como combinação linear dos outros dois. Portanto, uma base para $\operatorname{Im} T$ é $\{1 + x^2, 2 + x^2 + x^3\}$ e $\dim(\operatorname{Im} T) = 2$. Pelo Teorema do Núcleo e Imagem,

$$3 = \dim U = \dim(\ker T) + \dim(\operatorname{Im} T) = \dim(\ker T) + 2 \Rightarrow \dim(\ker T) = 1.$$

■

4.2 Isomorfismos

Nesta seção, introduziremos o conceito de isomorfismo entre espaços vetoriais. Dois espaços isomorfos, do ponto de vista da álgebra linear, são essencialmente o mesmo espaço, pois compartilham as mesmas propriedades de espaços vetoriais, o que muda é apenas a aparência dos seus elementos.

Definição 4.2.1 Sejam U e V espaços vetoriais sobre o mesmo $\mathbb{K}$.

(i) Dizemos que uma transformação linear $T : U \to V$ é um *isomorfismo* se T for bijetora. Se $U = V$, um isomorfismo $T : U \to U$ também é chamado de *automorfismo*.

(ii) Dizemos que U e V são *isomorfos* se existe um isomorfismo $T : U \to V$. Neste caso, denotamos $U \cong V$.

■ **Exemplo 4.2.2** A função identidade $\mathrm{Id} : U \to U$, definida por $\mathrm{Id}(u) = u$, para todo $u \in U$, é um isomorfismo. ■

■ **Exemplo 4.2.3** A transformação linear $T : \mathbb{R}^2 \to P_1(\mathbb{R})$ dada por $T(a,b) = a + (a+b)x$ é um isomorfimo.

Prova. Observe que

$$T(a,b) = 0 \Leftrightarrow a = 0 \text{ e } a+b = 0 \Leftrightarrow a = b = 0 \Leftrightarrow (a,b) = (0,0).$$

Portanto, $\ker T = \{0\}$, o que implica T injetora. Pelo Teorema do Núcleo e Imagem,

$$2 = \dim \mathbb{R}^2 = \dim(\ker T) + \dim(\operatorname{Im} T) = 0 + \dim(\operatorname{Im} T).$$

Logo, $\dim(\operatorname{Im} T) = 2 = \dim P_1(\mathbb{R})$. Portanto $\operatorname{Im} T = P_1(\mathbb{R})$, o que implica T sobrejetora.

Assim, T é um isomorfismo e $\mathbb{R}^2 \cong P_1(\mathbb{R})$. ■

■ **Exemplo 4.2.4** Para todo $n \in \mathbb{N}$, $P_n(\mathbb{K}) \cong \mathbb{K}^{n+1}$.

Prova. Deixaremos como exercício a verificação de que a função $T : P_n(\mathbb{K}) \to \mathbb{K}^{n+1}$, definida por

$$T(a_0 + a_1 x + \cdots + a_n x^n) = (a_0, a_1, \ldots, a_n),$$

é um isomorfismo. ■

Observe nos exemplos anteriores que os espaços isomorfos possuem a mesma dimensão. Isso não é uma coincidência! Veremos nesta seção que dois espaços vetoriais são isomorfos se, e somente se, eles possuem a mesma dimensão.

Proposição 4.2.5 Se $T : U \to V$ é uma transformação linear injetora, então T leva todo conjunto LI em U num conjunto LI em V.

Demonstração. Seja $S \subseteq U$ um conjunto LI. Mostremos que $T(S) = \{T(u) : u \in S\}$ é LI em V. Suponha que $u_1, \ldots, u_m \in U$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ são tais que

$$\alpha_1 T(u_1) + \cdots + \alpha_n T(u_n) = 0.$$

Então, como T é linear,

$$T(\alpha_1 u_1 + \cdots + \alpha_n u_n) = 0 \Rightarrow \alpha_1 u_1 + \cdots + \alpha_n u_n \in \ker T = \{0\},$$

pois T é injetora. Assim, $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$, e como S é LI, segue que

$$\alpha_1 = \cdots = \alpha_n = 0.$$

■

Proposição 4.2.6 Se $U \cong V$, então $\dim U = \dim V$.

Demonstração. Como $U \cong V$, existe $T : U \to V$ isomorfismo.

Se $\dim U = \infty$, então, existe $S \subseteq U$, LI e infinito. Como T é injetora, $T(S)$ é infinito e, pela Proposição 4.2.5, $T(S)$ é LI. Logo $\dim V = \infty$.

Suponha agora $\dim U$ finita. Pelo Teorema do Núcleo e Imagem,

$$\dim U = \dim(\ker T) + \dim(\operatorname{Im} T) = 0 + \dim(\operatorname{Im} T) = \dim(\operatorname{Im} T).$$

Mas, por hipótese, T é sobrejetora. Logo $\operatorname{Im} T = V$, donde segue que $\dim U = \dim(\operatorname{Im} T) = \dim V$. ■

Proposição 4.2.7 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de mesma dimensão finita n e $T : U \to V$ uma transformação linear. São equivalentes:

(i) T é injetora;

(ii) T é sobrejetora;

(iii) T é isomorfismo.

Demonstração. **[(i) ⇒ (ii)]** T injetora $\Rightarrow \ker T = \{0\} \Rightarrow \dim(\ker T) = 0$. Pelo Teorema do Núcleo e Imagem, $n = \dim U = \dim(\operatorname{Im} T)$. Logo, $\operatorname{Im} T$ é um subespaço de V com a mesma dimensão de V, portanto $\operatorname{Im} T = V$.

[(ii) ⇒ (iii)] T sobrejetora $\Rightarrow \operatorname{Im} T = V$. Logo, pelo Teorema do Núcleo e Imagem,

$$n = \dim U = \dim(\ker T) + \dim(\operatorname{Im} T) = \dim(\ker T) + n$$

donde segue que $\dim(\ker T) = 0$, o que implica $\ker T = \{0\}$. Portanto, T é injetora, logo é isomorfismo.

[(iii) ⇒ (i)] Óbvio. ■

Observação 4.2.8 A Proposição 4.2.7 não é verdadeira se $\dim U = \infty$. Por exemplo, o operador derivação $D : P(\mathbb{R}) \to P(\mathbb{R})$ não é injetor, pois $0 \neq 1 \in \ker D$, mas D é sobrejetor (verifique).

Proposição 4.2.9 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de mesma dimensão finita $n \geq 1$. São equivalentes:

(i) T é isomorfismo;

(ii) T leva qualquer base de U numa base de V;

(iii) existe uma base B de U tal que $T(B)$ é base de V.

Demonstração. **[(i) $\Rightarrow$ (ii)]** Seja $B = \{u_1, \dots, u_n\}$ uma base qualquer de V. Como T é injetora, $\#T(B) = n$ e, pela Proposição 4.2.5, $T(B)$ é LI. Assim, $T(B) \subseteq V$ é um suconjunto LI com $n = \dim V$ elementos, portanto $T(B)$ é base de V.

[(ii) $\Rightarrow$ (iii)] Óbvio.

[(iii) $\Rightarrow$ (i)] Seja B uma base de U tal que $T(B)$ é base de V. Então, $[T(B)] = V$. Por outro lado, pela Proposição 4.1.19, $[T(B)] = \operatorname{Im} T$. Logo, $\operatorname{Im} T = V$ e T é sobrejetora. O resultado agora segue da Proposição 4.2.7. ■

■ **Exemplo 4.2.10** Verifique se a transformação linear $T : M_2(\mathbb{R}) \to \mathbb{R}^4$ definida por

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (2b - c + d, a + c + 3d, 2b + 4c + d, 3c)$$

é um isomorfismo.

Solução. Considere $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ a base canônica de $M_2(\mathbb{R})$. Temos,

$$T(E_{11}) = (0,1,0,0),\ \ T(E_{12}) = (2,0,2,0),\ \ T(E_{21}) = (-1,1,4,3),\ \ T(E_{22}) = (1,3,1,0).$$

Observe que

$$2T(E_{22}) - 6T(E_{11}) = T(E_{12}),$$

portanto $T(B)$ é LD, logo não é base de $\mathbb{R}^4$. Pela Proposição 4.2.9, T não é isomorfismo. ■

Teorema 4.2.11 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de dimensão finita. Então, $U \cong V$ se, e somente se, $\dim U = \dim V$.

Demonstração. A implicação $U \cong V \ \Rightarrow\ \dim U = \dim V$ segue da Proposição 4.2.6 e é válida mesmo para dimensão infinita.

Vamos provar a recíproca. Se $\dim U = \dim V = 0$, então ambos são espaços com apenas 1 vetor (o nulo) e a única função que pode ser definida entre eles é um isomorfismo. Suponha $\dim U = \dim V = n \geq 1$. Sejam $B = \{u_1, \dots, u_n\}$ uma base de U e $C = \{v_1, \dots, v_n\}$ uma base de V. Considere $T : U \to V$ a única transformação linear tal que $T(u_i) = v_i$, $i = 1, \dots, n$. Em particular, $T(B) = C$ é base de V, e portanto T leva uma base de U numa base de V. Pela Proposição 4.2.9, T é isomorfismo. ■

Proposição 4.2.12 Sejam U e V $\mathbb{K}$-espaços vetoriais e $T : U \to V$ um isomorfismo. Então, $T^{-1} : U \to V$ também é um isomorfismo.

Demonstração. Como $T : U \to V$ é isomorfismo, T é bijetora, e portanto inversível. Por propriedades de funções, a função inversa $T^{-1} : V \to U$ também é bijetora. Resta mostrar que T^{-1} é uma transformação linear. De fato, sejam $v_1, v_2 \in V$ e $\lambda \in \mathbb{K}$. Vamos mostrar que $T^{-1}(\lambda v_1 + v_2) = \lambda T^{-1}(v_1) + T^{-1}(v_2)$. Sejam $u_1, u_2 \in U$ tais que $T(u_1) = v_1$ e $T(u_2) = v_2$. Então, usando que T é linear,

$$T^{-1}(\lambda v_1 + v_2) = T^{-1}(\lambda T(u_1) + T(u_2)) = T^{-1}(T(\lambda u_1 + u_2)) = T^{-1} \circ T(\lambda u_1 + u_2)$$

$$= \lambda u_1 + u_2 = \lambda T^{-1}(v_1) + T^{-1}(v_2).$$

■

4.3 Matriz de uma transformação linear

Dados dois espaços vetoriais de dimensão finita e uma transformação linear entre eles, introduziremos o conceito de matriz dessa transformação. Em posse disso, será possível descrever essa transformação como multiplicação de vetores por uma matriz fixa, assim como no Exemplo 4.1.10. Além disso, veremos com traduzir propriedades de uma transformação linear em termos de propriedades matriciais, o que trará um ganho computacional.

Sejam U e V dois $\mathbb{K}$-espaços vetoriais de dimensão finita, com $\dim U = n \geq 1$ e $\dim V = m \geq 1$, e $T : U \to V$ uma transformação linear.

Fixe $B = \{u_1, \dots, u_n\}$ uma base ordenada para U e $C = \{v_1, \dots, v_m\}$ uma base ordenada para V. Para cada $j = 1, \dots, n$, escreva $T(u_j)$ como combinação linear dos vetores de C:

$$T(u_j) = a_{1j}v_1 + \cdots + a_{mj}v_m = \sum_{i=1}^{m} a_{ij}v_i. \tag{4.2}$$

Lembre que os escalares a_{ij} são únicos.

A matriz $(a_{ij}) \in M_{m\times n}(\mathbb{K})$ é chamada de *matriz da transformação linear T em relação às bases B e C* e será denotada por $[T]_{BC}$. No caso de $U = V$ e $B = C$ denotamos $[T]_{BC}$ simplesmente por $[T]_B$. Note que, a j-ésima coluna de $[T]_{BC}$ é exatamente $[T(u_j)]_C$ (coordenadas de $T(u_j)$ na base C).

Seja $u \in U$. Suponha que $[u]_B = (\alpha_1, \dots, \alpha_n)$ e $[T(u)]_C = (\beta_1, \dots, \beta_m)$, isto é,

$$u = \alpha_1 u_1 + \cdots + \alpha_n u_n = \sum_{j=1}^{n} \alpha_j u_j \quad \text{e} \quad T(u) = \beta_1 v_1 + \cdots + \beta_m v_m = \sum_{i=1}^{m} \beta_i v_i.$$

Por outro lado, como T é linear, temos

$$T(u) = T\left(\sum_{j=1}^{n} \alpha_j u_j\right) = \sum_{j=1}^{n} \alpha_j T(u_j) \overset{(4.2)}{=} \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{m} a_{ij}v_i\right) = \sum_{j=1}^{n}\sum_{i=1}^{m} (\alpha_j a_{ij})v_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\alpha_j\right) v_i.$$

Como os escalares da combinação linear dos vetores de uma base são únicos, temos

$$\beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j = a_{i1}\alpha_1 + \cdots + a_{in}\alpha_n, \quad \text{para todo } i = 1, \dots, m.$$

Em notação matricial, segue que

$$\underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}}_{[T(u)]_C} = \underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}_{[T]_{BC}} \underbrace{\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}}_{[u]_B}.$$

Ou seja, a matriz da transformação linear satisfaz

$$\boxed{[T(u)]_C = [T]_{BC}\,[u]_B, \quad \text{para todo } u \in U.} \tag{4.3}$$

Além disso, se $M = (b_{ij}) \in M_{m\times n}(\mathbb{K})$ satisfaz $[T(u)]_C = M[u]_B$, para todo $u \in U$, substituindo $u = u_j$, $j = 1, \dots, n$, obtém-se que $b_{ij} = a_{ij}$ para todos i, j (verifique). Ou seja, $M = [T]_{BC}$. Portanto, a matriz da transformação T em relação às bases B e C é a <u>única</u> matriz que satisfaz (4.3).

■ **Exemplo 4.3.1** Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ a transformação linear dada por

$$T(x,y) = (-2y, 4x+y, 5y-3x).$$

Considerando B e C as bases canônicas de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente, temos,

$$T(1,0) = (0,4,-3) = (0,4,-3)_C \quad \text{e} \quad T(0,1) = (-2,1,5) = (-2,1,5)_C.$$

Portanto,

$$[T]_{BC} = \begin{bmatrix} 0 & -2 \\ 4 & 1 \\ -3 & 5 \end{bmatrix}.$$

Note que, se considerarmos os elementos de $\mathbb{R}^2$ e $\mathbb{R}^3$ como vetores colunas, temos $T(v) = [T]_{BC} \cdot v$ para todo $v \in \mathbb{R}^2$, ou seja, $T = T_A$, com $A = [T]_{BC}$ (Exemplo 4.1.10). Isso ocorre porque estamos considerando as bases canônicas, e daí os vetores se confundem com as próprias coordenadas.

Considere agora as bases $B = \{(1,1),(2,0)\}$ e $C = \{(1,0,-1),(0,0,4),(0,2,1)\}$. Então,

$$T(1,1) = (-2,5,2) \quad \text{e} \quad T(2,0) = (0,8,-6).$$

Agora,

$$(-2,5,2) = -2(1,0,-1) - \frac{5}{8}(0,0,4) + \frac{5}{2}(0,2,1) \Rightarrow [T(1,1)]_C = \left(-2, -\frac{5}{8}, \frac{5}{2}\right),$$

e

$$(0,8,-6) = 0(1,0,-1) - \frac{5}{2}(0,0,4) + 4(0,2,1) \Rightarrow [T(2,0)]_C = \left(0, -\frac{5}{2}, 4\right)$$

Portanto,

$$[T]_{BC} = \begin{bmatrix} -2 & 0 \\ -5/8 & -5/2 \\ 5/2 & 4 \end{bmatrix}.$$

Neste caso, não é verdade que $T(v) = [T]_{BC} \cdot v$ para todo $v \in \mathbb{R}^2$, mas sim a relação dada em (4.3). ■

■ **Exemplo 4.3.2** Seja $A \in M_{m\times n}(\mathbb{K})$. Considere a transformação linear $T_A : \mathbb{K}^n \to \mathbb{K}^m$ dada por $T_A(x) = Ax$ (Exemplo 4.1.10). Se B é a base canônica de $\mathbb{K}^n$ e C é a base canônica de $\mathbb{K}^m$, então

$$[T_A]_{BC} = A.$$

Prova. Basta observar que, se A_j é a j-ésima coluna de A, então para cada $e_j \in B$,

$$[T(e_j)]_C = T(e_j) = Ae_j = A_j.$$

Lembre que, em relação à base canônica C, todo vetor de $\mathbb{K}^m$ é igual ao seu vetor de coordenadas. ■

■ **Exemplo 4.3.3** Considere a (única) transformação linear $T : \mathbb{R}^3 \to P_3(\mathbb{R})$ tal que $T(1,0,1) = 2+x^2+x^3$, $T(0,1,0) = 1+x^2$, $T(0,0,1) = x^2-x^3$, dada no Exemplo 4.1.24. Se B e C são as bases canônicas de $\mathbb{R}^3$ e $P_3(\mathbb{R})$, respectivamente, determine $[T]_{BC}$.

Solução. Temos,

$$T(1,0,0) = T(1,0,1) - T(0,0,1) = 2+x^2+x^3-x^2+x^3 = 2+2x^3,$$

e $T(0,1,0)$ e $T(0,0,1)$ já foram dados. Logo,

$$[T]_{BC} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix}.$$

■

■ **Exemplo 4.3.4** Seja U um espaço vetorial de dimensão finita. Considere a transformação identidade $\mathrm{Id} : U \to U$, que é dada por $\mathrm{Id}(u) = u$ para todo $u \in U$. Dadas duas bases B e C de U, a matriz dessa transformação $[\mathrm{Id}]_{BC}$ é exatamente a matriz de mudança de base M_{BC} de B para C (verifique). ■

Fixadas as bases B e C dos espaços vetoriais U e V, respectivamente, para qualquer transformação linear $T : U \to V$ está associada uma única matriz $[T]_{BC}$. O resultado a seguir enuncia a recíproca desse fato: dada uma matriz qualquer, pode-se definir unicamente a transformação linear associada.

Proposição 4.3.5 Sejam U e V dois espaços vetoriais de dimensão finita sobre $\mathbb{K}$, com $\dim U = n$ e $\dim V = m$. Dadas B e C bases de U e V, respectivamente, e uma matriz $A \in M_{m\times n}(\mathbb{K})$, existe uma única transformação linear $T : U \to V$ tal que $[T]_{BC} = A$.

Demonstração. Escreva $B = \{u_1, \dots, u_n\}$, $C = \{v_1, \dots, v_m\}$ e

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

Considere $T : U \to V$ a única transformação linear tal que

$$T(u_j) = a_{1j}v_1 + \cdots + a_{mj}v_m, \quad j = 1, \dots n.$$

Segue da definição de matriz de transformação que $[T]_{BC} = A$. ■

Proposição 4.3.6 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de dimensão finita e B e C bases de U e V, respectivamente. Dadas $S, T : U \to V$ transformações lineares e $\lambda \in \mathbb{K}$, tem-se

$$[S+T]_{BC} = [S]_{BC} + [T]_{BC} \quad \text{e} \quad [\lambda S]_{BC} = \lambda [S]_{BC}.$$

Demonstração. Exercício. ■

A proposição a seguir nos fornece a dimensão do espaço das transformações lineares $\mathscr{L}(U,V)$ (Exemplo 4.1.11), no caso de U e V serem espaços de dimensão finita.

Proposição 4.3.7 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de dimensão finita, digamos $\dim U = n$ e $\dim V = m$. Então, $\mathscr{L}(U,V) \cong M_{m\times n}(\mathbb{K})$. Em particular, $\dim \mathscr{L}(U,V) = mn$.

Demonstração. Fixe B uma base ordenada de U e C uma base ordenada de V. Considere a função $\varphi : \mathscr{L}(U,V) \to M_{m\times n}(\mathbb{K})$ definida por

$$\varphi(T) = [T]_{BC} \quad \text{para toda } T \in \mathscr{L}(U,V).$$

Primeiramente, note que φ está bem definida, pois dada uma transformação linear $T : U \to V$, $[T]_{BC} \in M_{m\times n}(\mathbb{K})$ é única. Segue da Proposição 4.3.6 que φ é uma transformação linear. Além disso, pela Proposição 4.3.5, φ é sobrejetora.

Resta provar que φ é injetora, ou equivalentemente, $\ker\varphi = \{0\}$.

Seja $T \in \ker\varphi$. Então, $[T]_{BC}$ é a matriz nula. Escrevendo $B = \{u_1,\dots,u_n\}$, para cada $j = 1,\dots,n$, a j-ésima coluna de $[T]_{BC}$ é $[T(u_j)]_C = 0$, donde segue que $T(u_j) = 0$. Como $T(B)$ gera $\operatorname{Im} T$, segue que $\operatorname{Im} T = \{0\}$, logo $T = 0$.

Portanto, φ é um isomorfismo. ■

Dados U,V,W espaços vetoriais sobre o mesmo $\mathbb{K}$, e, $S : U \to V$ e $T : V \to W$ transformações lineares, não é difícil verificar que a função composta $T \circ S : U \to W$ é uma transformação linear. A proposição a seguir descreve como a matriz da composta se relaciona com as matrizes de S e T.

Proposição 4.3.8 Sejam U,V,W $\mathbb{K}$-espaços vetoriais de dimensão finita, e, B, C, D bases de U,V,W, respectivamente. Sejam $S : U \to V$ e $T : V \to W$ duas transformações lineares, então

$$[T \circ S]_{BD} = [T]_{CD}\,[S]_{BC}.$$

Demonstração. Seja $u \in U$. Aplicando (4.3) duas vezes, temos

$$[T]_{CD}[S]_{BC}[u]_B = [T]_{CD}[S(u)]_C = [T(S(u))]_D = [T \circ S(u)]_D.$$

Assim, a matriz $M = [T]_{CD}[S]_{BC}$ satisfaz $[T \circ S(u)]_D = M[u]_B$ para todo $u \in U$. Pela unicidade dessa propriedade, segue que

$$M = [T]_{CD}[S]_{BC} = [T \circ S]_{BD}.$$

■

Corolário 4.3.9 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de mesma dimensão finita $n \geq 1$, B e C bases de U e V, respectivamente. Uma transformação linear $T : U \to V$ é um isomorfismo se, e somente se, $[T]_{BC}$ for não singular.
Nesse caso,

$$\left[T^{-1}\right]_{CB} = \left([T]_{BC}\right)^{-1}.$$

Demonstração. Suponha que T é isomorfismo. Então, T é inversível e $T^{-1} \circ T = \mathrm{Id}_U$. Pela Proposição 4.3.8, temos

$$[T^{-1}]_{CB}[T]_{BC} = [T^{-1} \circ T]_B = [\mathrm{Id}_U]_B = I \quad \text{(matriz identidade)}.$$

Portanto, $[T]_{BC}$ é inversível e $\left([T]_{BC}\right)^{-1} = [T^{-1}]_{CB}$.

Reciprocamente, suponha $[T]_{BC}$ inversível. Então, existe uma matriz $A \in M_n(\mathbb{K})$ tal que

$$A[T]_{BC} = [T]_{BC}A = I.$$

Mostremos que $\ker T = \{0\}$. Seja $u \in U$ tal que $T(u) = 0$. Então

$$[T]_{BC}[u]_B = [T(u)]_C = 0 \;\Rightarrow\; A[T]_{BC}[u]_B = A \cdot 0 = 0 \;\Rightarrow\; I[u]_B = 0 \;\Rightarrow\; [u]_B = 0 \;\Rightarrow\; u = 0.$$

Logo, T é injetora. Como U e V têm a mesma dimensão, segue que T é isomorfismo. ■

■ **Exemplo 4.3.10** Considere a transformação linear $T : \mathbb{R}^2 \to P_1(\mathbb{R})$ dada por $T(a,b) = (a+2b) + (4a-b)x$. Mostre que T é um isomorfismo e determine T^{-1}.

Solução. Considere $B = \{e_1, e_2\} \subseteq \mathbb{R}^2$ e $C = \{1, x\} \subseteq P_1(\mathbb{R})$ as bases canônicas. Temos,

$$T(e_1) = 1 + 4x \quad \text{e} \quad T(e_2) = 2 - x \ \Rightarrow \ [T]_{BC} = \begin{bmatrix} 1 & 2 \\ 4 & -1 \end{bmatrix}.$$

Note que $\det[T]_{BC} = -9 \neq 0$, logo $[T]_{BC}$ é não singular. Pelo Corolário 4.3.9, T é isomorfismo e

$$\left[T^{-1}\right]_{CB} = \left([T]_{BC}\right)^{-1} = -\frac{1}{9}\begin{bmatrix} -1 & -2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} 1/9 & 2/9 \\ 4/9 & -1/9 \end{bmatrix}.$$

A partir da matriz $\left[T^{-1}\right]_{CB}$ é possível determinar $T^{-1}(a+bx)$ explicitamente. Temos,

$$[T^{-1}(a+bx)]_B = \left[T^{-1}\right]_{CB}[a+bx]_C = \begin{bmatrix} 1/9 & 2/9 \\ 4/9 & -1/9 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} (a+2b)/9 \\ (4a-b)/9 \end{bmatrix}.$$

Logo,

$$T^{-1}(a+bx) = \frac{a+2b}{9}e_1 + \frac{4a-b}{9}e_2 = \left(\frac{a+2b}{9}, \frac{4a-b}{9}\right).$$

■

Para finalizar essa seção, apresentamos a relação entre matrizes de um mesmo operador em relação à bases distintas.

Definição 4.3.11 Duas matrizes $A, B \in M_n(\mathbb{K})$ são ditas *semelhantes* se existe uma matriz $P \in M_n(\mathbb{K})$ não singular tal que

$$B = P^{-1}AP.$$

Proposição 4.3.12 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e $T : V \to V$ um operador linear. Se B e C são duas bases de V, então

$$[T]_B = M_{CB}\,[T]_C\,M_{BC} = M_{BC}^{-1}\,[T]_C\,M_{BC}.$$

Em particular, as matrizes $[T]_B$ e $[T]_C$ são semelhantes.

Demonstração. Vamos considerar a composição de T com a transformação identidade $\mathrm{Id} : U \to U$, alternando as bases B para o domínio e C para o contradomínio, e vice-versa. Pela Proposição 4.3.8, temos

$$[T]_B = [\mathrm{Id} \circ T \circ \mathrm{Id}]_B = [\mathrm{Id} \circ T]_{CB}\,[\mathrm{Id}]_{BC} = [\mathrm{Id}]_{CB}\,[T]_C\,[\mathrm{Id}]_{BC}.$$

Pelo Exemplo 4.3.4, $[\mathrm{Id}]_{BC} = M_{BC}$ e $[\mathrm{Id}]_{CB} = M_{CB} = M_{BC}^{-1}$, donde segue o resultado. ■

4.4 Posto

Definição 4.4.1 Dada uma transformação linear $T : U \to V$, o posto de T é definido por $\dim(\mathrm{Im}\, T)$.

Relacionaremos o posto de uma transformação linear entre espaços de dimensão finita, com o posto de sua matriz. Lembre-se que o espaço coluna de uma matriz, $R(A)$, foi definido no Exemplo 2.3.6.

Definição 4.4.2 Seja $A \in M_{m\times n}(\mathbb{K})$.

(i) O *posto coluna* de A é definido como sendo a dimensão do subespaço $R(A)$ de $\mathbb{K}^m$, gerado pelas colunas de A.

(ii) O *posto linha* de A é definido como sendo dimensão do subespaço $R(A^\top)$ de $\mathbb{K}^n$, gerado pelas linhas de A.

Ou seja, o posto coluna de A é o número de colunas LI que A possui. E o posto linha de A é o número de linhas LI.

■ **Exemplo 4.4.3** Determine o posto linha e o posto coluna da matriz $A = \begin{bmatrix} 4 & 0 & 1 & 8 \\ 3 & 5 & 2 & 6 \end{bmatrix}$.

Solução. Como as linhas não são múltiplas escalares, elas são LI. Portanto o posto linha de A é igual a 2.

Em relação ao posto coluna, como as colunas de A são vetores em $\mathbb{R}^2$, a dimensão de $R(A)$ é no máximo 2. Não é difícil verificar que, por exemplo, as duas primeiras colunas de A são LI. Portanto o posto coluna de A também é 2. ■

No exemplo anterior, o posto linha e o posto coluna da matriz A são iguais. Isso não é coincidência, como será provado adiante.

Observação 4.4.4 Como explicado no Procedimento 2.8.1, se forem feitas operações elementares nas linhas de uma matriz A, o espaço linha da nova matriz A' é o mesmo espaço linha de A. Em particular, esses espaços possuem a mesma dimensão. Além disso, se fizermos operações elementares sobre as linhas de A e obtivermos uma matriz R escalonada (ou escalonada reduzida), o posto linha de A será exatamente o número de linhas não nulas de R.

■ **Exemplo 4.4.5** Determine o posto linha da matriz

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 \\ 2 & 4 & 1 & 0 & 2 \\ 0 & 0 & 3 & 0 & 12 \\ 4 & 8 & -1 & 1 & -6 \end{bmatrix}.$$

Solução. Escalonaremos a matriz dada. Aplicando as operações elementares:

$$L_2 - 2L_1 \to L_2, \quad \frac{1}{3}L_3 \to L_3, \quad L_4 - 4L_1 \to L_4,$$

obtemos

$$\begin{bmatrix} 1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & -1 & 1 & -2 \end{bmatrix}.$$

Aplicando agora as operações elementares:

$$L_3 - L_2 \to L_3, \quad L_4 + L_2 \to L_4,$$

obtemos

$$\begin{bmatrix} 1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

Trocando as duas últimas linhas de posição, obtemos uma forma escalonada R (que é reduzida) da matriz A. Como R possui 3 linhas não nulas, o posto linha de R, que é igual ao posto linha de A, é 3. ∎

Proposição 4.4.6 Sejam U e V dois $\mathbb{K}$-espaços vetoriais de dimensão finita, B e C bases de U e V, respectivamente. Dada uma transformação linear $T : U \to V$, o posto de T é igual ao posto coluna de $[T]_{BC}$.

Demonstração. Se $B = \{u_1, \ldots, u_n\}$, então $\operatorname{Im} T = [T(u_1), \ldots, T(u_n)]$. Logo, o posto de T, é igual a dimensão do subespaço gerado por $T(u_1), \ldots, T(u_n)$. Por outro lado, as colunas de $[T]_{BC}$ são exatamente $[T(u_1)]_C, \ldots, [T(u_n)]_C$ e o posto coluna de $[T]_{BC}$ é a dimensão do subespaço gerado por esses vetores. Segue da Proposição 2.7.3 que $T(u_j) \in [\{T(u_1), \ldots, T(u_n)\} \setminus \{T(u_j)\}]$ se, e somente se, $[T(u_j)]_C \in [\{[T(u_1)]_C, \ldots, [T(u_n)]_C\} \setminus \{[T(u_j)]_C\}]$. Daí, não é difícil concluir que as dimensões de $\operatorname{Im} T$ e $R([T]_{BC})$ são iguais. ∎

Lema 4.4.7 Sejam $A \in M_{m\times n}(\mathbb{K})$, $B \in M_{m\times p}(\mathbb{K})$ e $C \in M_{p\times n}(\mathbb{K})$ tais que $A = BC$. Então, o posto coluna de A é menor ou igual ao posto coluna de B.

Demonstração. Considere as transformações lineares $T_A : \mathbb{K}^n \to \mathbb{K}^m$, $T_B : \mathbb{K}^p \to \mathbb{K}^n$ e $T_C : \mathbb{K}^n \to \mathbb{K}^p$. Pelo Exemplo 4.3.2, as matrizes de T_A, T_B, T_C em relação às bases canônicas são exatamente A, B, C, respectivamente. Pela Proposição 4.3.8 junto com a hipótese $A = BC$, temos que $T_A = T_B \circ T_C$. Agora, não é difícil verificar que $\dim(\operatorname{Im} T_A) \leq \dim(\operatorname{Im} T_B)$, pois $\operatorname{Im} T_A = \operatorname{Im}(T_B|_{\operatorname{Im} T_C})$. Por fim, segue da Proposição 4.4.6 que o posto coluna de A é menor ou igual ao posto coluna de B. ∎

Teorema 4.4.8 — Teorema do posto. Seja $A \in M_{m\times n}(\mathbb{K})$. Então, o posto linha de A é igual ao posto coluna de A.

Demonstração. Seja $p \leq n$ o posto coluna de A e $q \leq m$ o posto linha de A.

Considere $A_j \in M_{m\times 1}(\mathbb{K})$ a j-ésima coluna de A. Então, o espaço coluna $R(A) = [A_1, \ldots, A_n]$ tem dimensão p. Considere $\{B_1, \ldots, B_p\} \subseteq M_{m\times 1}(\mathbb{K})$ uma base de $R(A)$. Assim, para cada $j = 1, \ldots, n$, existem $\alpha_{1j}, \ldots, \alpha_{pj} \in \mathbb{K}$ tais que

$$A_j = \alpha_{1j}B_1 + \cdots + \alpha_{pj}B_p. \tag{4.4}$$

Considere $B \in M_{m\times p}(\mathbb{K})$ a matriz cujas colunas são $B_1, \ldots, B_p$ e $C = (\alpha_{kj}) \in M_{p\times n}(\mathbb{K})$, onde os α_{kj} são dados por (4.4). Segue de (4.4) que

$$A_j = B\begin{bmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{pj} \end{bmatrix} \text{ para todo } j = 1, \ldots, n \;\Rightarrow\; A = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix} = B\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{p1} & \cdots & \alpha_{pn} \end{bmatrix} = BC.$$

Além disso, $A^\top = C^\top B^\top$. Pelo Lema 4.4.7, temos

$$q = \text{posto linha de } A = \text{posto coluna de } A^\top \leq \text{posto coluna de } C^\top = \text{posto linha de } C \leq p,$$

pois C possui p linhas. Assim, provamos que $q \leq p$.

Repetindo todo o argumento anterior, iniciando com $A^\top$ e com o seu posto coluna (que igual ao posto linha de A), prova-se que $p \leq q$, donde segue o resultado desejado. ∎

De agora em diante, denotaremos por $\operatorname{posto}(A)$ o posto linha ou coluna de A.

Exercício 4.1 Sejam $T: U \to V$ e $S: V \to W$ transformações lineares, com U, V e W de dimensão finita. Prove as afirmações a seguir.

(i) $\text{posto}(S \circ T) \leq \text{posto}(S)$.

(ii) $\text{posto}(S \circ T) \leq \text{posto}(T)$.

(iii) Se S é injetiva, então $\text{posto}(S \circ T) = \text{posto }(T)$.

(iv) Se T é sobrejetiva, então $\text{posto}(S \circ T) = \text{posto }(S)$.

(v) Dê um exemplo em que $\text{posto}(T) = \text{posto }(S) > \text{posto}(S \circ T)$.

4.5 Autovalores, autovetores, operadores diagonalizáveis

Iniciaremos essa seção com um problema prático, para servir de motivação e introduzir as primeiras ideias. Trata-se de um exemplo clássico, conhecido como Sistema Predador-Presa.

Considere uma população de coelhos e uma população de raposas convivendo num mesmo habitat. Os coelhos (presas) se reproduzem rapidamente e as raposas (predadores) comem os coelhos para sobreviver. A taxa de crescimento da população de coelhos depende da sua taxa de reprodução e da taxa de eliminação pelas raposas (que comem os coelhos). A taxa de crescimento das raposas depende do número de raposas e da quantidade de coelhos que elas tem para comer. Suponha que no ano inicial, $n = 0$, haviam 30 coelhos e 20 raposas, e que, a cada ano que passa, o número de coelhos $c(n+1)$ e o número de raposas $r(n+1)$ são dados em função dos quantidades do ano anterior pelas equações

$$\begin{cases} c(n+1) = 4c(n) - 2r(n) \\ r(n+1) = c(n) + r(n) \end{cases} . \tag{4.5}$$

Queremos descobrir qual é o número de coelhos e de raposas depois de 50 anos, isto é, $c(50)$ e $r(50)$.

Observe que, se $p(n) = \begin{bmatrix} c(n) \\ r(n) \end{bmatrix}$ é o vetor população depois de n anos, então $p(0) = \begin{bmatrix} 30 \\ 20 \end{bmatrix}$ e segue de (4.5) que

$$p(n+1) = \begin{bmatrix} c(n+1) \\ r(n+1) \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c(n) \\ r(n) \end{bmatrix}.$$

Assim, se $A = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$, temos

$$p(1) = Ap(0), \quad p(2) = Ap(1) = A^2 p(0), \quad \dots \quad , \quad p(n+1) = Ap(n) = \cdots = A^{n+1} p(0).$$

Logo, $p(50) = A^{50} p(0)$. Calcular A^{50} não é uma tarefa fácil, vejamos as primeiras potências:

$$A^2 = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 14 & -10 \\ 5 & -1 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 14 & -10 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 46 & -38 \\ 19 & -11 \end{bmatrix}.$$

Precisamos de uma maneira mais esperta de fazer essa conta. A grande sacada, que adiante vocês aprenderão como obter, é observar que os vetores $u = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$ e $v = \begin{bmatrix} 20 \\ 10 \end{bmatrix}$ satisfazem

$$Au = \begin{bmatrix} 20 \\ 20 \end{bmatrix} = 2u \quad \text{e} \quad Av = \begin{bmatrix} 60 \\ 30 \end{bmatrix} = 3v,$$

respectivamente. Ou seja, multiplicar A por u e v resulta num múltiplo escalar desses vetores, não muda a direção deles. Tais vetores são chamados de *autovetores* de A, e os escalares 2 e 3 são chamados de *autovalores* de A. Neste caso,

$$A^2u = A(2u) = 2(Au) = 2^2u, \ \ldots \ , \quad A^nu = 2^nu,$$

e, analogamente,

$$A^nv = 3^nv.$$

Assim, é mais simples calcular potências de A multiplicada pelos autovetores. Então, escrevamos $p(0)$ como combinação linear de u e v:

$$p(0) = \begin{bmatrix} 30 \\ 20 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix} + \begin{bmatrix} 20 \\ 10 \end{bmatrix} = u + v.$$

Agora,

$$p(50) = A^{50}p(0) = A^{50}(u+v) = A^{50}u + A^{50}v = 2^{50}u + 3^{50}v = \begin{bmatrix} 2^{50}\cdot 10 + 3^{50}\cdot 20 \\ 2^{50}\cdot 10 + 3^{50}\cdot 10 \end{bmatrix}.$$

Portanto, $c(50) = 2^{50}\cdot 10 + 3^{50}\cdot 20$ e $r(50) = 2^{50}\cdot 10 + 3^{50}\cdot 10$.

Uma outra consequência da existência de u e v, é que se P é a matriz cujas colunas são u e v, isto é, se $P = \begin{bmatrix} 10 & 20 \\ 10 & 10 \end{bmatrix}$, e D é a matriz diagonal com os autovalores 2 e 3, ou seja, $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, então

$$A = PDP^{-1} \quad \text{(verifique)}.$$

Logo, A é semelhante a uma matriz diagonal. Neste caso, dizemos que A é *diagonalizável*. Usando a relação acima, fica simples de se calcular potências de A, uma vez que

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PD^2P^{-1}, \ \ldots \ , \quad A^n = PD^nP^{-1},$$

e potência de matriz diagonal também é simples:

$$D^2 = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}, \ \ldots \ , \quad D^n = \begin{bmatrix} 2^n & 0 \\ 0 & 3^n \end{bmatrix}.$$

Agora formalizaremos os conceitos apresentados nesse exemplo, para o caso de matrizes e operadores lineares. Iniciaremos no contexto de operadores.

Definição 4.5.1 Sejam V um $\mathbb{K}$-espaço vetorial e $T : V \to V$ um operador linear.

(i) Um escalar $\lambda \in \mathbb{K}$ é um *autovalor* de T se existe $v \in V \setminus \{0\}$ com $T(v) = \lambda v$.

(ii) Se $v \in V \setminus \{0\}$ satisfaz $T(v) = \lambda v$, para algum $\lambda \in \mathbb{K}$, dizemos que v é um *autovetor* de T, associado à λ.

Denote por $\mathrm{Aut}_T(\lambda)$ o seguinte subconjunto de V:

$$\mathrm{Aut}_T(\lambda) = \{v \in V : T(v) = \lambda v\}.$$

Esse conjunto é sempre um subespaço vetorial de V:

- $0 \in \mathrm{Aut}_T(\lambda)$ pois $T(0) = 0 = \lambda \cdot 0$.

- Se $u, v \in \mathrm{Aut}_T(\lambda)$, então $u+v \in \mathrm{Aut}_T(\lambda)$, pois $T(u+v) = T(u)+T(v) = \lambda u + \lambda v = \lambda(u+v)$.
- Se $u \in \mathrm{Aut}_T(\lambda)$ e $\alpha \in \mathbb{K}$, então $\alpha u \in \mathrm{Aut}_T(\lambda)$, pois $T(\alpha u) = \alpha T(u) = \alpha(\lambda u) = \lambda(\alpha u)$.

Também, não é difícil verificar que

$$\mathrm{Aut}_T(\lambda) = \ker(T - \lambda\,\mathrm{Id}),$$

onde $\mathrm{Id} : V \to V$ é o operador identidade.

Além disso, λ é um autovalor de V se, e somente se, $\mathrm{Aut}_T(\lambda) \neq \{0\}$. Observe também que $\mathrm{Aut}_T(0) = \ker T$, portanto 0 é autovalor de T se, e somente se, T não é injetora.

Definição 4.5.2 $\mathrm{Aut}_T(\lambda)$ é chamado de *autoespaço* de T associado a λ.

Exemplo 4.5.3 Considere $T : \mathbb{K}^2 \to \mathbb{K}^2$ definido por $T(x,y) = (-y,x)$.

Procuremos os possíveis autovalores de T, considerando os dois casos $\mathbb{K} = \mathbb{C}$ e $\mathbb{K} = \mathbb{R}$.

Temos,

$$T(x,y) = \lambda(x,y) \Leftrightarrow (-y,x) = (\lambda x, \lambda y) \Leftrightarrow -y = \lambda x \text{ e } x = \lambda y \Rightarrow -y = \lambda^2 y \Leftrightarrow (\lambda^2+1)y = 0.$$

Se $y = 0$, então $x = \lambda y = 0$, e assim (x,y) não seria autovetor. Supondo $y \neq 0$, concluímos que

$$\lambda^2 + 1 = 0.$$

Como λ é um escalar em $\mathbb{K}$, no caso $\mathbb{K} = \mathbb{R}$ a equação anterior não possui solução. Nesse caso, T não possui autovalor (nem autovetor). No caso $\mathbb{K} = \mathbb{C}$, temos duas soluções: $\lambda = i$ e $\lambda = -i$.

Para $\lambda = i$, $x = iy$ e os vetores não nulos da forma (iy,y) são autovetores de T associados à i, ou seja,

$$\mathrm{Aut}_T(i) = \{(iy,y) : y \in \mathbb{C}\} = [(i,1)].$$

Para $\lambda = -i$, $x = -iy$ e os vetores não nulos da forma $(-iy,y)$ são autovetores de T associados à $-i$, ou seja,

$$\mathrm{Aut}_T(-i) = \{(-iy,y) : y \in \mathbb{C}\} = [(-i,1)].$$

No caso $\mathbb{K} = \mathbb{R}$, observe que T é o operador rotação de 90° no sentido anti-horário, donde é fácil perceber que não existe autovetor de T, pois nenhum vetor não nulo será levado num múltiplo escalar de si mesmo. ■

4.5.1 Propriedades de autovalores e autovetores

Nesta subseção veremos algumas propriedades envolvendo autovalores e autovetores de operadores lineares.

Proposição 4.5.4 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$, B uma base de V, $\lambda \in \mathbb{K}$ e $T : V \to V$ um operador linear. São equivalentes:

(i) λ é autovalor de T;

(ii) $T - \lambda\,\mathrm{Id}$ não é injetor;

(iii) $T - \lambda\,\mathrm{Id}$ não é inversível;

(iv) a matriz $[T - \lambda\,\mathrm{Id}]_B$ é singular;

(v) $\det([T-\lambda\,\mathrm{Id}]_B) = \det([T]_B - \lambda I_n) = 0$, onde I_n é a matriz identidade de ordem n.

Demonstração. **[(i) ⇔ (ii)]** λ é autovalor de T ⇔ existe $v \in V\setminus\{0\}$ tal que $T(v) = \lambda v$ ⇔ existe $v \in V\setminus\{0\}$ tal que $(T-\lambda\,\mathrm{Id})(v) = 0$ ⇔ $\ker(T-\lambda\,\mathrm{Id}) \neq \{0\}$ ⇔ $T-\lambda\,\mathrm{Id}$ não é injetor.

[(ii) ⇔ (iii)] Segue da Proposição 4.2.7.

[(iii) ⇔ (iv)] Segue do Corolário 4.3.9.

[(iv) ⇔ (v)] Por propriedades de matrizes, $[T-\lambda\,\mathrm{Id}]_B$ é singular ⇔ $\det([T-\lambda\,\mathrm{Id}]_B) = 0$. Por fim, pela Proposição 4.3.6,

$$[T-\lambda\,\mathrm{Id}]_B = [T]_B - \lambda[\mathrm{Id}]_B = [T]_B - \lambda I_n.$$

■

Pela proposição anterior, λ é um autovalor de T se, e somente se, λ é raiz do polinômio

$$p_T(x) = \det([T - x\,\mathrm{Id}]_B) = \det([T]_B - xI_n).$$

Este polinômio é chamado de *polinômio característico de* T.

Note que, $p_T(x)$ não depende da base B fixada, pois se C é uma outra base de V, então pela Proposição 4.3.12, temos

$$\det([T-x\mathrm{Id}]_B) = \det(M_{BC}^{-1}[T-x\mathrm{Id}]_C M_{BC}) = \det([T-x\mathrm{Id}]_C)\det(M_{BC}^{-1})\det(M_{BC}) = \det([T-x\mathrm{Id}]_C).$$

Além disso, a matriz $[T]_B - xI_n$ possui polinômios de grau 1 com coeficientes em $\mathbb{K}$ nas entradas da diagonal principal e constantes em $\mathbb{K}$ nas demais entradas, ou seja $p_T(x) = \det([T]_B - xI_n)$ é um polinômio em $P(\mathbb{K})$ de grau exatamente $n = \dim V$, cujo coeficiente líder é $(-1)^n$.

Também, se $\mathbb{K} = \mathbb{C}$, $p_T(x)$ sempre tem raízes em $\mathbb{K}$, donde segue que T sempre tem autovalor.

■ **Exemplo 4.5.5** Considere $T:\mathbb{R}^3 \to \mathbb{R}^3$ dado por $T(a,b,c) = (2a-2b-c, -2b+c, 4a-2b-3c)$. Determine, se existirem, todos os autovalores e autovetores de T.

Solução. Para determinar o polinômio característico de T, precisamos da matriz de T em alguma base. Fixe $B = \{e_1,e_2,e_3\}$ a base canônica de $\mathbb{R}^3$. O leitor facilmente verificará que

$$[T]_B = \begin{bmatrix} 2 & -2 & -1 \\ 0 & -2 & 1 \\ 4 & -2 & -3 \end{bmatrix}.$$

Assim,

$$p_T(x) = \det([T]_B - xI_3) = \det\begin{bmatrix} 2-x & -2 & -1 \\ 0 & -2-x & 1 \\ 4 & -2 & -3-x \end{bmatrix} = -x^3 - 3x^2 - 2x = -x(x+1)(x+2),$$

cujas raízes são $0, -1, -2$, que são os autovalores de T. Determinemos agora os autoespaços associados, e consequentemente os autovetores.

- $\mathrm{Aut}_T(-2) = \ker(T+2\,\mathrm{Id})$. Trabalhando com a matriz do operador, para $v = (a,b,c) \in \mathbb{R}^3$ temos

$$v \in \ker(T+2\mathrm{Id}) \Leftrightarrow (T+2\mathrm{Id})(v) = 0 \Leftrightarrow ([T]_B + 2I_3)[v]_B = [T+2\mathrm{Id}]_B[v]_B = [(T+2\mathrm{Id})(v)]_B = 0$$

$$\Leftrightarrow \begin{bmatrix} 4 & -2 & -1 \\ 0 & 0 & 1 \\ 4 & -2 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} 4a-2b-c = 0 \\ c = 0 \\ 4a-2b-c = 0 \end{cases} \Leftrightarrow c = 0 \text{ e } b = 2a.$$

Logo, $v \in \mathrm{Aut}_T(-2)$ se, e somente se,

$$v = (a, 2a, 0) = a(1,2,0) \in [(1,2,0)].$$

Portanto, $\mathrm{Aut}_T(-2) = [(1,2,0)]$.

- $\mathrm{Aut}_T(-1) = \ker(T + \mathrm{Id})$. Para $v = (a,b,c) \in \mathbb{R}^3$ temos

$$v \in \ker(T+\mathrm{Id}) \Leftrightarrow ([T]_B + I_3)[v]_B = 0 \Leftrightarrow \begin{bmatrix} 3 & -2 & -1 \\ 0 & -1 & 1 \\ 4 & -2 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} 3a - 2b - c = 0 \\ -b + c = 0 \\ 4a - 2b - 2c = 0 \end{cases} \Leftrightarrow a = b = c.$$

Logo, $v \in \mathrm{Aut}_T(-1)$ se, e somente se,

$$v = (a,a,a) = a(1,1,1) \in [(1,1,1)].$$

Portanto, $\mathrm{Aut}_T(-1) = [(1,1,1)]$.

- $\mathrm{Aut}_T(0) = \ker T$. Para $v = (a,b,c) \in \mathbb{R}^3$ temos

$$v \in \ker T \Leftrightarrow [T]_B[v]_B = 0 \Leftrightarrow \begin{bmatrix} 2 & -2 & -1 \\ 0 & -2 & 1 \\ 4 & -2 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} 2a - 2b - c = 0 \\ -2b + c = 0 \\ 4a - 2b - 3c = 0 \end{cases} \Leftrightarrow a = c = 2b.$$

Logo, $v \in \mathrm{Aut}_T(0)$ se, e somente se,

$$v = (2b, b, 2b) = b(2,1,2) \in [(2,1,2)].$$

Portanto, $\mathrm{Aut}_T(0) = [(2,1,2)]$.

■

■ **Exemplo 4.5.6** Considere $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ dado por $T(a+bx+cx^2) = 2a+b+c+(b-a)x+(a+c)x^2$. Determine, se existirem, todos os autovalores e autovetores de T.

Solução. Fixe $B = \{1, x, x^2\}$ a base canônica de $V = P_2(\mathbb{R})$. O leitor facilmente verificará que

$$[T]_B = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Assim,

$$p_T(x) = \det([T]_B - xI_3) = \det \begin{bmatrix} 2-x & 1 & 1 \\ -1 & 1-x & 0 \\ 1 & 0 & 1-x \end{bmatrix} = (2-x)(1-x)^2,$$

cujas raízes são 1 e 2, que são os autovalores de T. Determinemos agora os autoespaços associados, e consequentemente os autovetores.

- $\mathrm{Aut}_T(1) = \ker(T - \mathrm{Id})$. Trabalhando com a matriz do operador, para $v = a + bx + cx^2 \in V$ temos

$$v \in \ker(T - \mathrm{Id}) \Leftrightarrow (T - \mathrm{Id})(v) = 0 \Leftrightarrow ([T]_B - I_3)[v]_B = [T - \mathrm{Id}]_B[v]_B = [(T - \mathrm{Id})(v)]_B = 0$$

$$\Leftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} a + b + c = 0 \\ -a = 0 \\ a = 0 \end{cases} \Leftrightarrow a = 0 \text{ e } b = -c.$$

Logo, $v \in \mathrm{Aut}_T(1)$ se, e somente se,

$$v = -cx + cx^2 = c(-x + x^2) \in [-x + x^2].$$

Portanto, $\mathrm{Aut}_T(1) = [-x + x^2]$.

- $\mathrm{Aut}_T(2) = \ker(T - 2\,\mathrm{Id})$. Para $v = a + bx + cx^2 \in V$ temos

$$v \in \ker(T - 2\,\mathrm{Id}) \Leftrightarrow ([T]_B - 2I_3)[v]_B = 0 \Leftrightarrow \begin{bmatrix} 0 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} b + c = 0 \\ -a - b = 0 \\ a - c = 0 \end{cases} \Leftrightarrow a = c = -b.$$

Logo, $v \in \mathrm{Aut}_T(2)$ se, e somente se,

$$v = -b + bx - bx^2 = b(-1 + x - x^2) \in [-1 + x - x^2].$$

Portanto, $\mathrm{Aut}_T(2) = [-1 + x - x^2]$.

■

■ **Exemplo 4.5.7** Considere $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear tal que sua matriz na base $B = \{(1,0,0),(1,1,0),(0,1,1)\}$ é dada por

$$[T]_B = \begin{bmatrix} 5 & 0 & -3 \\ 0 & -1 & 0 \\ 6 & 0 & -4 \end{bmatrix}.$$

Determine, se existirem, todos os autovalores e autovetores de T.

Solução. A partir da matriz de T fornecida, podemos determinar o polinômio característico de T:

$$p_T(x) = \det([T]_B - xI_3) = \det \begin{bmatrix} 5 - x & 0 & -3 \\ 0 & -1 - x & 0 \\ 6 & 0 & -4 - x \end{bmatrix} = -(x+1)^2(x-2),$$

cujas raízes são -1 e 2, que são os autovalores de T. Determinemos agora os autoespaços associados, e consequentemente os autovetores.

- $\mathrm{Aut}_T(-1) = \ker(T + \mathrm{Id})$. Para $v \in \mathbb{R}^3$ com $[v]_B = (a,b,c)$ temos

$$v \in \ker(T + \mathrm{Id}) \Leftrightarrow ([T]_B + I_3)[v]_B = 0 \Leftrightarrow \begin{bmatrix} 6 & 0 & -3 \\ 0 & 0 & 0 \\ 6 & 0 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \left\{ \begin{array}{l} 6a-3c=0 \\ 6a-3c=0 \end{array} \right. \Leftrightarrow c=2a.$$

Logo, $v \in \mathrm{Aut}_T(-1)$ se, e somente se,

$$[v]_B=(a,b,2a)_B=a(1,0,2)_B+b(0,1,0)_B \in [(1,0,2)_B,(0,1,0)_B].$$

Explicitamente,

$$(1,0,2)_B=(1,0,0)+2(0,1,1)=(1,2,2) \quad \text{e} \quad (0,1,0)_B=(1,1,0).$$

Portanto, $\mathrm{Aut}_T(-1)=[(1,2,2),(1,1,0)]$.

- $\mathrm{Aut}_T(2)=\ker(T-2\,\mathrm{Id})$. Para $v \in \mathbb{R}^3$ com $[v]_B=(a,b,c)$ temos

$$v \in \ker(T+\mathrm{Id}) \Leftrightarrow ([T]_B+I_3)[v]_B=0 \Leftrightarrow \begin{bmatrix} 3 & 0 & -3 \\ 0 & -2 & 0 \\ 6 & 0 & -6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \left\{ \begin{array}{r} 3a-3c=0 \\ -2b=0 \\ 6a-6c=0 \end{array} \right. \Leftrightarrow b=0 \text{ e } c=a.$$

Logo, $v \in \mathrm{Aut}_T(2)$ se, e somente se,

$$[v]_B=(a,0,a)_B=a(1,0,1)_B \in [(1,0,1)_B].$$

Explicitamente,

$$(1,0,1)_B=(1,0,0)+(0,1,1)=(1,1,1).$$

Portanto, $\mathrm{Aut}_T(2)=[(1,1,1)]$.

■

■ **Exemplo 4.5.8** Considere $T:\mathbb{C}^2 \to \mathbb{C}^2$ a transformação linear tal que $T(1,i)=(1,2+i)$ e $T(0,1)=(0,1)$. Determine, se existirem, todos os autovalores e autovetores de T.

Solução. Observe que $B=\{(1,i),(0,1)\}$ é base de $\mathbb{C}^2$. Temos,

$$(1,0)=(1,i)-i(0,1) \Rightarrow T(1,0)=T(1,i)-iT(0,1)=(1,2+i)-i(0,1)=(1,2).$$

Considere a base canônica $C=\{(1,0),(0,1)\}$ de $\mathbb{C}^2$. Assim,

$$[T]_C=\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \Rightarrow p_T(x)=\det\begin{bmatrix} 1-x & 0 \\ 2 & 1-x \end{bmatrix}=(1-x)^2.$$

Assim, o único autovalor de T é 1. Se $v=(a,b)\in\mathbb{C}^2$, então

$$v \in \mathrm{Aut}_T(1)=\ker(T-\mathrm{Id}) \Leftrightarrow ([T]_C-I_2)[v]_C=0 \Leftrightarrow \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow 2a=0 \Leftrightarrow a=0.$$

Logo, $v \in \mathrm{Aut}_T(1)$ se, e somente se,

$$v=(0,b)=b(0,1)\in[(0,1)].$$

Portanto, $\mathrm{Aut}_T(1)=[(0,1)]$.

■

Os exemplos anteriores ilustraram algumas situações que podem ocorrer em relação aos autovalores e autovetores de um operador linear.

Note que, para os λ encontrados, o espaço solução do sistema homogêneo $([T]_B - \lambda I_n)[v]_B = 0$ sempre tem dimensão maior que zero. Isso deve mesmo ocorrer, pois uma vez que λ é raiz do polinômio característico, λ é autovalor, e portanto existe autovetor associado, isto é, existe $v \in V \setminus \{0\}$ tal que $T(v) = \lambda v \Leftrightarrow (T - \lambda\,\mathrm{Id})(v) = 0$.

Um outro fato que pode ser observado nos exemplos é que autovetores associados à autovalores distintos são LI. Como veremos a seguir, isso não é uma simples coincidência!

Proposição 4.5.9 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $T : V \to V$ um operador linear. Se $\lambda_1, \ldots, \lambda_m$, $m \geq 2$ são autovalores de T dois a dois distintos, então a soma

$$\mathrm{Aut}_T(\lambda_1) + \cdots + \mathrm{Aut}_T(\lambda_m)$$

é direta. Em particular, se B_j é base de $\mathrm{Aut}_T(\lambda_j)$, então $B_1 \cup \cdots \cup B_m$ é LI.

Demonstração. Mostraremos que dados $v_j \in \mathrm{Aut}_T(\lambda_j)$, $j = 1, \ldots, m$, com

$$v_1 + \cdots + v_m = 0 \;\Rightarrow\; v_j = 0 \;\text{ para todo } j = 1, \ldots, m. \tag{4.6}$$

Note que isso é equivalente a dizer que todo vetor na soma se decompõe de maneira única, pois $u_1 + \cdots + u_m = v_1 + \cdots + v_m$ se, e somente se, $(u_1 - v_1) + \cdots + (u_m - v_m) = 0$.

Faremos por indução em $m \geq 2$. Suponha, inicialmente, $m = 2$ e $v_1 + v_2 = 0$, com $v_j \in \mathrm{Aut}_T(\lambda_j)$, $j = 1, 2$. Então

$$\lambda_1 v_1 + \lambda_2 v_2 = T(v_1) + T(v_2) = T(v_1 + v_2) = T(0) = 0. \tag{4.7}$$

Além disso,

$$\lambda_2 v_1 + \lambda_2 v_2 = \lambda_2(v_1 + v_2) = 0. \tag{4.8}$$

Subtraindo (4.7) e (4.8), obtemos

$$\lambda_1 v_1 - \lambda_2 v_1 = 0 \;\Rightarrow (\lambda_1 - \lambda_2)v_1 = 0.$$

Por hipótese, $\lambda_1 - \lambda_2 \neq 0$, donde segue que $v_1 = 0$, e daí $v_2 = 0$.

Suponha (4.6) válida para $m = k \geq 2$ (hipótese de indução). Utilizaremos a mesma ideia do caso $m = 2$ para mostrar para $m = k + 1$. Temos,

$$\lambda_1 v_1 + \cdots + \lambda_k v_k + \lambda_{k+1} v_{k+1} = T(v_1) + \cdots + T(v_{k+1}) = T(v_1 + \cdots + v_{k+1}) = T(0) = 0. \tag{4.9}$$

Também,

$$\lambda_{k+1}(v_1 + \cdots + v_k + v_{k+1}) = 0. \tag{4.10}$$

Subtraindo (4.9) e (4.10), obtemos

$$(\lambda_1 - \lambda_{k+1})v_1 + \cdots + (\lambda_k - \lambda_{k+1})v_k = 0.$$

Para cada $j = 1, \ldots, k$, temos que $(\lambda_j - \lambda_{k+1})v_j \in \mathrm{Aut}_T(\lambda_j)$. Logo, pela hipótese de indução,

$$(\lambda_j - \lambda_{k+1})v_j = 0 \;\text{ para todo } j = 1, \ldots, k.$$

Como, por hipótese, $\lambda_j - \lambda_{k+1} \neq 0$ para todo $j < k$, segue que $v_j = 0$ para todo $j = 1, \ldots, k$. Portanto,

$$v_{k+1} = v_1 + \cdots + v_k + v_{k+1} = 0.$$

■

Corolário 4.5.10 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $T : V \to V$ um operador linear. Se $\lambda_1, \ldots, \lambda_m$ são autovalores de T dois a dois distintos e $v_j \in V$ é um autovetor associado à λ_j, $j = 1, \ldots, m$, então $\{v_1, \ldots, v_m\}$ é LI.

Demonstração. Suponha $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ tais que

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0.$$

Como $\alpha_j v_j \in \mathrm{Aut}_T(\lambda_j)$, pela Proposição 4.5.9, $\alpha_j v_j = 0$ para todo $j = 1, \ldots, m$. Mas, cada $v_j \neq 0$, pois são autovetores, logo $\alpha_j = 0$ para todo $j = 1, \ldots, m$. ■

4.5.2 Operadores diagonalizáveis e potências de matrizes

Nesta subseção, vamos apresentar o conceito diagonalização e determinar condições necessárias e suficientes para que um operador seja diagonalizável.

Definição 4.5.11 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $T : V \to V$ um operador linear. Dizemos que T é *diagonalizável* se existe uma base B de V tal que $[T]_B$ é uma matriz diagonal.

Note que, a matriz $[T]_B$ é diagonal se, e somente se, todos os vetores da base B são autovetores de T. Assim, dizer que T é diagonalizável é equivalente a dizer que existe uma base de V formada por autovetores de T.

■ **Exemplo 4.5.12** O operador T definido no Exemplo 4.5.3 é diagonalizável, pois $B = \{(i,1), (-i,1)\}$ é uma base de $\mathbb{C}^2$ formada por autovetores de T. Além disso,

$$[T]_B = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

■

■ **Exemplo 4.5.13** O operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido no Exemplo 4.5.5 é diagonalizável, pois $D = \{(1,2,0), (1,1,1), (2,1,2)\} \subseteq \mathbb{R}^3$ é LI, e portanto é uma base de $\mathbb{R}^3$. Assim, $\mathbb{R}^3$ possui uma base formada por autovetores de T, e

$$[T]_D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

■

■ **Exemplo 4.5.14** O operador $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ definido no Exemplo 4.5.6 não é diagonalizável, pois todos os autovetores de T pertencem à $\mathrm{Aut}_T(1)$ ou $\mathrm{Aut}_T(2)$. Como

$$\dim \mathrm{Aut}_T(1) = 1 = \dim \mathrm{Aut}_T(2),$$

é possível conseguir, no máximo, 2 autovetores LI em V e $\dim V = 3$. Portanto, V não possui uma base formada por autovetores de T. ■

■ **Exemplo 4.5.15** O operador $T : \mathbb{R}^3 \to \mathbb{R}^3$ definido no Exemplo 4.5.7 é diagonalizável, pois $D = \{(1,2,2), (1,1,0), (1,1,1)\} \subseteq \mathbb{R}^3$ é LI (verifique), D é uma base de $\mathbb{R}^3$. Assim, $\mathbb{R}^3$ possui uma base formada por autovetores de T. Além disso,

$$[T]_D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

■

■ **Exemplo 4.5.16** O operador $T:\mathbb{C}^2\to\mathbb{C}^2$ definido no Exemplo 4.5.8 não é diagonalizável, pois todos os autovetores de T pertencem à $\mathrm{Aut}_T(1)$. Como $\dim \mathrm{Aut}_T(1)=1$, é possível conseguir, no máximo, 1 autovetor LI em $\mathbb{C}^2$. Portanto, $\mathbb{C}^2$ não possui uma base formada por autovetores de T. ■

Teorema 4.5.17 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $T:V\to V$ um operador linear. Se $\lambda_1,\dots,\lambda_m$ são todos os autovalores distintos de T, então T é diagonalizável se, e somente se,

$$\dim V=\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_m)).$$

Neste caso, $V=\mathrm{Aut}_T(\lambda_1)\oplus\cdots\oplus\mathrm{Aut}_T(\lambda_m)$.

Demonstração. Sejam $\lambda_1,\dots,\lambda_m$ todos os autovalores distintos de T. Pela Proposição 4.5.9,

$$\dim(\mathrm{Aut}_T(\lambda_1)+\cdots+\mathrm{Aut}_T(\lambda_m))=\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_m)).$$

Suponha que T é diagonalizável. Então, existe uma base de V formada por autovalores de T. Daí, todo $v\in V$ se escreve como combinação linear de autovetores, ou seja, como soma de vetores em $\mathrm{Aut}_T(\lambda_1)+\cdots+\mathrm{Aut}_T(\lambda_m)$. Portanto,

$$V=\mathrm{Aut}_T(\lambda_1)+\cdots+\mathrm{Aut}_T(\lambda_m)\ \Rightarrow \dim V=\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_m)).$$

Reciprocamente, suponha

$$\dim V=\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_m))=\dim(\mathrm{Aut}_T(\lambda_1)+\cdots+\mathrm{Aut}_T(\lambda_m)).$$

Tomando B_j base de $\mathrm{Aut}_T(\lambda_j)$, pela Proposição 4.5.9, $B=B_1\cup\cdots\cup B_m$ é LI, e possui uma quantidade de vetores igual a dimensão de V, portanto é uma base de V, formada por autovetores de T. Logo, T é diagonalizável.

■

Corolário 4.5.18 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n\geq 1$ e $T:V\to V$ um operador linear. Se T possui n autovalores distintos, então T é diagonalizável.

Demonstração. Sejam $\lambda_1,\dots,\lambda_n$ todos os autovalores distintos de T. Como $\dim(\mathrm{Aut}_T(\lambda_j))\geq 1$, segue que

$$\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_n))\geq n.$$

Por outro lado, pela Proposição 4.5.9,

$$\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_n))=\dim(\mathrm{Aut}_T(\lambda_1)+\cdots+\mathrm{Aut}_T(\lambda_n))\leq \dim V=n,$$

pois $\mathrm{Aut}_T(\lambda_1)+\cdots+\mathrm{Aut}_T(\lambda_n)$ é um subespaço de V. Logo,

$$\dim(\mathrm{Aut}_T(\lambda_1))+\cdots+\dim(\mathrm{Aut}_T(\lambda_n))=n,$$

e o resultado segue do Teorema 4.5.17. ■

Definição 4.5.19 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita, $T:V\to V$ um operador linear e $\lambda\in\mathbb{K}$ um autovalor de T. Suponha $p_T(x)=(x-\lambda)^m q(x)$ com $m\geq 1$ e $q(\lambda)\neq 0$. Definimos $m_a(\lambda)=m$ a *multiplicidade algébrica* de λ e $m_g(\lambda)=\dim(\mathrm{Aut}_T(\lambda))$ a *multiplicidade geométrica* de λ.

Note que a multiplicidade algébrica do autovalor λ é exatamente a multiplicidade como raiz do polinômio característico, que é o maior natural $m\geq 1$ tal que $(x-\lambda)^m$ divide $p_T(x)$.

Proposição 4.5.20 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita, $T : V \to V$ um operador linear e $\lambda \in \mathbb{K}$ um autovalor de T. Então, $m_g(\lambda) \le m_a(\lambda)$.

Demonstração. Sejam $n = \dim V$ e $k = m_g(\lambda) = \dim(\mathrm{Aut}_T(\lambda)) \le n$. Considere $B' = \{w_1, \dots, w_k\}$ uma base de $\mathrm{Aut}_T(\lambda)$.

Se $k = n$, então B' é uma base de V formada por autovetores associados à λ. Daí, não é difícil ver que $p_T(x) = \det([T]_{B'} - xI_n) = (\lambda - x)^n = (-1)^n(x-\lambda)^n$, donde segue que $m_a(\lambda) = n = m_g(\lambda)$.

Suponha agora $k < n$. Completando B' para uma base de V, existem $w_{k+1}, \dots, w_n \in V$ tais que $B = \{w_1, \dots, w_k, w_{k+1}, \dots, w_n\}$ é base de V. Como $T(w_j) = \lambda w_j$ para todo $j = 1, \dots, k$, temos

$$[T]_B = \left[\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & A_1 \\ & & \lambda & \\ \hline & 0 & & A_2 \end{array}\right], \quad \text{onde } A_1 \in M_{k\times(n-k)}(\mathbb{K}) \text{ e } A_2 \in M_{(n-k)\times(n-k)}(\mathbb{K}).$$

Expandindo $\det([T]_B - xI_n)$ em cofatores na primeira coluna, em seguida na primeira coluna da matriz $(n-1)\times(n-1)$ restante, e assim sucessivamente k vezes, obtemos

$$p_T(x) = \det([T]_B - xI_n) = (\lambda - x)^k \det(A_2 - xI_{n-k}) = (-1)^k(x-\lambda)^k \det(A_2 - xI_{n-k}).$$

Como $m_a(\lambda)$ é o maior natural l tal que $(x-\lambda)^l$ divide $p_T(x)$, segue que $m_g(\lambda) = k \le m_a(\lambda)$. ■

■ **Exemplo 4.5.21** Determine as multiplicidades algébrica e geométrica dos autovalores do operador linear $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$, cuja matriz em relação à base

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é dada por

$$[T]_B = \begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix}.$$

Solução. Os autovalores de T são as raízes do polinômio característico:

$$p_T(x) = \det([T]_B - xI_4) = \det \begin{bmatrix} 1-x & 0 & 0 & -3 \\ 0 & -2-x & 0 & 0 \\ 0 & 0 & -2-x & 0 \\ -3 & 0 & 0 & 1-x \end{bmatrix}$$

$$= (1-x)\det \begin{bmatrix} -2-x & 0 & 0 \\ 0 & -2-x & 0 \\ 0 & 0 & 1-x \end{bmatrix} + 3\det \begin{bmatrix} 0 & -2-x & 0 \\ 0 & 0 & -2-x \\ -3 & 0 & 0 \end{bmatrix}$$

$$= (1-x)^2(-2-x)^2 - 9(2+x)^2 = (2+x)^2(x^2-2x-8) = (2+x)^3(x-4).$$

Logo, os autovalores de T são $\lambda = -2$ e $\lambda = 4$, com multiplicidade algébrica $m_a(-2) = 3$ e $m_a(4) = 1$.

Agora, para $\lambda = -2, 4$, temos $1 \leq m_g(\lambda) = \dim(\mathrm{Aut}_T(\lambda))$. Como $m_g(\lambda) \leq m_a(\lambda)$, segue que

$$m_g(4) = 1 \quad \text{e} \quad 1 \leq m_g(-2) \leq 3.$$

Determinemos $m_g(-2) = \dim(\mathrm{Aut}_T(-2)) = \dim(\ker(T + 2\,\mathrm{Id}))$. Para $A \in M_2(\mathbb{R})$ com $[A]_B = (a,b,c,d)$, temos

$$A \in \ker(T + 2\,\mathrm{Id}) \Leftrightarrow ([T]_B + 2I_4)[A]_B = 0 \Leftrightarrow \begin{bmatrix} 3 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -3 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow 3a - 3d = 0 \Leftrightarrow d = a.$$

Logo, $A \in \mathrm{Aut}_T(-2)$ se, e somente se,

$$[A]_B = (a,b,c,a)_B = a(1,0,0,1)_B + b(0,1,0,0)_B + c(0,0,1,0)_B \in [(1,0,0,1)_B, (0,1,0,0)_B, (0,0,1,0)_B].$$

Explicitamente,

$$(1,0,0,1)_B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad (0,1,0,0)_B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad (0,0,1,0)_B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},$$

que são três matrizes LI. Portanto, $\mathrm{Aut}_T(-2) = \left[\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right]$ e

$$m_g(-2) = \dim(\mathrm{Aut}_T(-2)) = 3.$$

■

Teorema 4.5.22 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e $T : V \to V$ um operador linear. Se $\lambda_1, \ldots, \lambda_m$, $m \geq 1$, são todos os autovalores distintos de T, então T é diagonalizável se, e somente se, existem $k_1, \ldots, k_m \in \mathbb{N}$ tais que

$$p_T(x) = (-1)^n (x - \lambda_1)^{k_1} \ldots (x - \lambda_m)^{k_m}, \quad \text{com } m_g(\lambda_j) = m_a(\lambda_j) = k_j \text{ para todo } j = 1, \ldots, m.$$

Neste caso, $k_1 + \cdots + k_m = n$.

Demonstração. Suponha que T é diagonalizável. Seja $k_j = m_a(\lambda_j)$, $j = 1, \ldots, m$. Pelo Teorema 4.5.17 e Proposição 4.5.20,

$$n = \dim V = m_g(\lambda_1) + \cdots + m_g(\lambda_m) \leq k_1 + \cdots + k_m.$$

Por outro lado, $(x - \lambda_j)^{k_j}$ divide $p_T(x)$ e $\mathrm{grau}(p_T(x)) = n$, logo $k_1 + \cdots + k_m \leq n$. Portanto, $k_1 + \cdots + k_m = n$, e o resultado segue.

Reciprocamente, suponha que existam $k_1, \ldots, k_m \in \mathbb{N}$ tais que

$$p_T(x) = (-1)^n (x - \lambda_1)^{k_1} \ldots (x - \lambda_m)^{k_m}, \quad \text{com } m_g(\lambda_j) = m_a(\lambda_j) = k_j \text{ para todo } j = 1, \ldots, m.$$

Daí, $k_j = m_a(\lambda_j) = m_g(\lambda_j)$, $j = 1, \ldots, m$, e $k_1 + \cdots + k_m = \mathrm{grau}(p_T(x)) = n$. Pelo Teorema 4.5.17, T é diagonalizável. ■

Recomendamos ao leitor que volte aos exemplos apresentados nesta seção, determine quais são as multiplicidades algébrica e geométrica dos autovalores e veja como determinar se o operador é diagonalizável a partir do Teorema 4.5.22.

Apresentaremos agora as definições dessa seção no contexto de matrizes.

Definição 4.5.23 Seja $A \in M_n(\mathbb{K})$.

(i) Dizemos que $\lambda \in \mathbb{K}$ é um *autovalor* de A se existe $v \in \mathbb{K}^n \setminus \{0\}$ tal que $Av = \lambda v$.

(ii) Se $v \in \mathbb{K}^n \setminus \{0\}$ satisfaz $Av = \lambda v$, para algum $\lambda \in \mathbb{K}$, dizemos que v é um *autovetor* de A, associado à λ.

(iii) Dizemos que A é *diagonalizável* se A é semelhante à uma matriz diagonal.

(iv) O polinômio $p_A(x) = \det(A - xI_n)$ é chamado de *polinômio característico* de A.

Não repetiremos os resultados dessa seção para o caso de matrizes. A ponte entre os contextos de operadores lineares e matrizes sempre pode ser feita da maneira a seguir.

$\star$ Dada uma matriz $A \in M_n(\mathbb{K})$, define-se o operador $T_A : \mathbb{K}^n \to \mathbb{K}^n$ por $T_A(x) = Ax$. A matriz desse operador em relação à base canônica C de $\mathbb{K}^n$ é $[T]_C = A$ (Exemplo 4.3.2). Então, as definições e propriedades apresentadas para T, se traduzem para A considerando multiplicação de A por vetores em $\mathbb{K}^n$.

$\star$ Dado um operador linear $T : V \to V$, com $\dim V = n$, fixada uma base B de V, temos a matriz $A = [T]_B \in M_n(\mathbb{K})$. Por (4.3), $[T(v)]_B = A[v]_B$, assim, essencialmente T é T_A, visto que podemos trabalhar com coordenadas ao invés dos próprios vetores para obter informações relacionadas a dependência linear, conjunto gerador, dimensão, etc.

Aplicação: potências de matrizes

■ **Exemplo 4.5.24** Suponha que tenhamos uma população inicial de 200 indivíduos saudáveis, mas ocorre uma epidemia grave. Os indivíduos podem ser classificadas como saudáveis, doentes ou mortos. Por causa da epidemia, a cada ano, 60% dos indivíduos saudáveis ficam doentes, apenas 30% se mantêm saudável, e 10% dos indivíduos saudáveis morrem. Sabemos também que a cada ano 60% dos doentes morrem, 20% voltam a ser saudáveis e 20% permanecem doentes. Assumindo que todos os indivíduos mortos permanecem mortos, determine quantos indivíduos são saudáveis, doentes ou mortos depois de k anos. Quantos anos levará até que apenas 10% (20 dos 200 indivíduos iniciais) da população ainda esteja viva?

Solução. Defina as variáveis:

$x_1(k)$: número de indivíduos saudáveis depois de k anos;

$x_2(k)$: número de indivíduos doentes depois de k anos;

$x_3(k)$: número de indivíduos mortos.

Pelos dados apresentados, obtemos as seguintes equações:

$x_1(k+1) = 0.3x_1(k) + 0.2x_2(k)$;

$x_2(k+1) = 0.6x_1(k) + 0.2x_2(k)$;

$x_3(k+1) = 0.1x_1(k) + 0.6x_2(k) + x_3(k)$.

Assim, se

$$x_k = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

podemos reescrever as equações na forma matricial $x_k = Ax_{k-1} = A^k x_0$, onde

$$A = \begin{bmatrix} 0.3 & 0.2 & 0 \\ 0.6 & 0.2 & 0 \\ 0.1 & 0.6 & 1 \end{bmatrix} \quad \text{e} \quad x_0 = \begin{bmatrix} 200 \\ 0 \\ 0 \end{bmatrix}.$$

Para estudar o número de indivíduos saudáveis, doentes e mortos ao longo dos anos, vamos diagonalizar a matriz A, pois assim ficará mais simples de calcular suas potências. Iniciaremos determinando os autovalores de A, que são as raízes do polinômio característico:

$$p_A(x) = \det \begin{bmatrix} 0.3-x & 0.2 & 0 \\ 0.6 & 0.2-x & 0 \\ 0.1 & 0.6 & 1-x \end{bmatrix} = -(x+0.1)(x-0.6)(x-1).$$

Portanto, os autovalores de A são $\lambda_1 = -0.1$, $\lambda_2 = 0.6$ e $\lambda_3 = 1.0$. Como A possui três autovalores distintos (quantidade igual a ordem da matriz A), segue que A é diagonalizável, semelhante a matriz

$$D = \begin{bmatrix} -0.1 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}.$$

Determinemos agora os autovetores de A:

$\lambda_1 = -0.1$: se $v_1 = (a,b,c)^\top$ é autovetor de A associado à λ_1 então $(A+0.1I_3)v = 0$, logo

$$\begin{bmatrix} 0.4 & 0.2 & 0 \\ 0.6 & 0.3 & 0 \\ 0.1 & 0.6 & 1.1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow b = -2a \text{ e } c = a \Leftrightarrow v_1 = a \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \ a \in \mathbb{R} \setminus \{0\}.$$

Escolha $v_1 = (1,-2,1)^\top$.

$\lambda_2 = 0.6$: se $v_2 = (a,b,c)^\top$ é autovetor de A associado à λ_2 então $(A-0.6I_3)v = 0$, logo

$$\begin{bmatrix} -0.3 & 0.2 & 0 \\ 0.6 & -0.4 & 0 \\ 0.1 & 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow b = \frac{3a}{2} \text{ e } c = -\frac{5a}{2} \Leftrightarrow v_2 = a \begin{bmatrix} 1 \\ 3/2 \\ -5/2 \end{bmatrix}, \ a \in \mathbb{R} \setminus \{0\}.$$

Escolha $v_2 = (-2,-3,5)^\top$.

$\lambda_3 = 1$: se $v_3 = (a,b,c)^\top$ é autovetor de A associado à λ_3 então $(A-I_3)v = 0$, logo

$$\begin{bmatrix} -0.7 & 0.2 & 0 \\ 0.6 & -0.8 & 0 \\ 0.1 & 0.6 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow a = b = 0 \Leftrightarrow v_3 = c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \ c \in \mathbb{R} \setminus \{0\}.$$

Escolha $v_3 = (0,0,1)^\top$.

Considerando a base $B = \{v_1, v_2, v_3\}$ de $\mathbb{R}^3$ formada pelos autovetores de A, temos que, dado qualquer $x_0 \in \mathbb{R}^3$, se $[x_0]_B = (a_1, a_2, a_3)_B$, isto é, se $x_0 = a_1v_1 + a_2v_2 + a_3v_3$, segue que

$$A^k x_0 = A^k(a_1v_1 + a_2v_2 + a_3v_3) = a_1\lambda_1^k v_1 + a_2\lambda_2^k v_2 + a_3\lambda_3^k v_3.$$

Como, por hipótese, $x_0 = (200,0,0)^\top$, temos que

$$x_0 = \frac{600}{7} v_1 - \frac{400}{7} v_2 + 200 v_3.$$

Logo,

$$x_k = \frac{600}{7}(-0.1)^k \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} - \frac{400}{7}(0.6)^k \begin{bmatrix} -2 \\ -3 \\ 5 \end{bmatrix} + 200\,(1)^k \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

ou seja, após k anos, teremos:

$(-1)^k \dfrac{600}{7 \cdot 10^k} + \dfrac{800}{7}\left(\dfrac{6}{10}\right)^k$ indivíduos saudáveis,

$(-1)^{k+1} \dfrac{1200}{7 \cdot 10^k} + \dfrac{1200}{7}\left(\dfrac{6}{10}\right)^k$ indivíduos doentes, e

$(-1)^k \dfrac{600}{7 \cdot 10^k} - \dfrac{2000}{7}\left(\dfrac{6}{10}\right)^k + 200$ indivíduos mortos.

Além disso, quando $k \to \infty$, a população tende a ser extinta.

Uma outra maneira de calcular as potências de A, seria escrevendo $A = PDP^{-1}$, onde $P \in M_3(\mathbb{R})$ é não singular. Para determinar P, considere $T_A : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T_A(x) = Ax$. A matriz desse operador em relação à base canônica C de $\mathbb{R}^n$ é $[T]_C = A$, e, em relação à base de autovetores B é $[T]_B = D$. Assim,

$$A = [T]_C = M_{BC}\,[T]_B\,M_{CB} = M_{BC}\,D\,M_{BC}^{-1}.$$

Ou seja, tome $P = M_{BC}$ a matriz de mudança de base de B para C. As colunas dessa matriz são formadas pelas coordenadas de v_j em relação à base canônica, que são os próprios vetores v_j. Logo,

$$P = \begin{bmatrix} 1 & -2 & 0 \\ -2 & -3 & 0 \\ 1 & 5 & 1 \end{bmatrix}.$$

Daí, basta calcular a inversa P^{-1}, e concluir

$$A = PDP^{-1} \;\Rightarrow\; A^k = PD^kP^{-1} = P \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{bmatrix} P^{-1}.$$

■

Capítulo 5

Transformações lineares entre espaços com produto interno

No Capítulo 4 foram estudadas as transformações lineares, que são as funções entre espaços vetoriais que preservam a estrutura a soma e a multiplicação por escalar. Neste capítulo, estudaremos as transformações lineares entre espaços com produto interno, obtendo ferramentas e resultados adicionais, que também serão traduzidos para o contexto de matrizes. Iniciaremos com as isometrias, que são as transformações lineares que preservam produto interno e assim, também preservam a geometria dos espaços envolvidos: norma, a distância e o ângulo entre os vetores. Em seguida, estaremos interessados nos operadores lineares, onde o principal objetivo será deduzir um dos teoremas mais importantes da Álgebra Linear: o Teorema Espectral.

A discussão apresentada nesse capítulo foi motivada pelos livros [Axl], [Coe], [Nic], [Reg] e [Str].

5.1 Isometrias

Definição 5.1.1 Sejam U e V dois $\mathbb{K}$-espaços vetoriais com produto interno. Dizemos que uma transformação linear $T: U \to V$ *preserva produto interno* se $\langle T(u), T(v)\rangle = \langle u, v\rangle$, para todos $u, v \in U$. Um *isomorfismo entre espaços com produto interno* é um isomorfismo que preserva produto interno.

Observação 5.1.2 Note que as funções que definem um produto interno em U e V podem ser diferentes. Dessa forma, na definição anterior, $\langle T(u), T(v)\rangle$ considera o produto interno de V, enquanto que $\langle u, v\rangle$ considera o produto interno de U.

Veremos a seguir alguns exemplos de transformações lineares que preservam produto interno.

■ **Exemplo 5.1.3** Considere $V = \mathbb{R}^3$ com o produto interno $\langle u, v\rangle = x_1x_2 + 4y_1y_2 + 9z_1z_2$, $u = (x_1, y_1, z_1)$ e $v = (x_2, y_2, z_2)$; e $W = \{A \in M_3(\mathbb{R}) : A^\top = A\}$ com o produto interno $\langle A, B\rangle = \frac{1}{2}\operatorname{tr}(AB^\top)$. Mostre que a transformação linear $T: V \to W$ definida por

$$T(x,y,z) = \begin{bmatrix} 0 & x & 2y \\ x & 0 & 3z \\ 2y & 3z & 0 \end{bmatrix}$$

preserva produto interno.

Solução. Dados $u = (x_1, y_1, z_1)$ e $v = (x_2, y_2, z_2)$ vetores de V, temos

$$\langle u, v\rangle = x_1x_2 + 4y_1y_2 + 9z_1z_2.$$

Por outro lado,

$$\langle T(u), T(v)\rangle = \frac{1}{2}\mathrm{tr}\left(\begin{bmatrix} 0 & x_1 & 2y_1 \\ x_1 & 0 & 3z_1 \\ 2y_1 & 3z_1 & 0 \end{bmatrix}\begin{bmatrix} 0 & x_2 & 2y_2 \\ x_2 & 0 & 3z_2 \\ 2y_2 & 3z_2 & 0 \end{bmatrix}\right) = x_1x_2 + 4y_1y_2 + 9z_1z_2.$$

Portanto, T preserva produto interno. ■

■ **Exemplo 5.1.4** Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $B = \{v_1, \ldots, v_n\}$ uma base ortonormal para V. Considere $\mathbb{K}^n$ com o produto interno usual. Mostre que a função $S : V \to \mathbb{K}^n$ definida por $S(v) = [v]_B$ é uma transformação linear que preserva produto interno.

Solução. Deixamos como exercício a verificação de que S é uma transformação linear. Sejam $u = \alpha_1 v_1 + \cdots + \alpha_n v_n$ e $v = \beta_1 v_1 + \cdots + \beta_n v_n$. Como B é uma base ortonormal, sabemos que

$$\langle u, v\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta}_i$$

independentemente do produto interno de V. Por outro lado,

$$\langle S(u), S(v)\rangle = \langle(\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n)\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta}_i.$$

Logo, S preserva produto interno. ■

Proposição 5.1.5 Se $T : U \to V$ é uma transformação linear que preserva produto interno, então T é injetora.

Demonstração. Sejam $u, v \in V$ tais $T(u) = T(v)$. Como T preserva produto interno, temos que

$$||u - v||^2 = \langle u - v, u - v\rangle = \langle T(u - v), T(u - v)\rangle = \langle T(u) - T(v), T(u) - T(v)\rangle = 0$$

ou seja, $u = v$. Logo, T é injetora. ■

Proposição 5.1.6 Sejam U e V dois espaços vetoriais com produto interno, $\dim U = \dim V$ finita, e $T : U \to V$ uma transformação linear. São equivalentes:

(i) T preserva produto interno;

(ii) T é um isomorfismo entre espaços com produto interno;

(iii) T leva toda base ortonormal de U em base ortonormal de V;

(iv) T leva alguma base ortonormal de U em uma base ortonormal de V.

Demonstração. **[(i) $\Rightarrow$ (ii)]** Segue das Proposições 5.1.5 e 4.2.7 que T é uma bijeção. Como T preserva produto interno, T é um isomorfismo entre espaços com produto interno.

[(ii) ⇒ (iii)] Seja $B = \{u_1,\dots,u_n\}$ uma base ortonormal de U. Mostremos que $C = \{T(u_1),\dots,T(u_n)\}$ é base ortonormal de V. Como T preserva produto interno, $\langle T(u_i),T(u_j)\rangle = \langle u_i,u_j\rangle = 0$ para todo $i \neq j$, e, $||T(u_i)|| = ||u_i|| = 1$ para todo $i = 1,\dots,n$. Portanto, C é base ortonormal.

[(iii) ⇒ (iv)] Imediato.

[(iv) ⇒ (i)] Seja $B = \{u_1,\dots,u_n\}$ uma base ortonormal de U tal que $C = \{T(u_1),\dots,T(u_n)\}$ é base ortonormal de V. Dados $u,v \in U$, existem escalares $\alpha_i,\beta_i \in \mathbb{K}$, $i = 1,\dots,n$, tais que $u = \alpha_1u_1+\cdots+\alpha_nu_n$ e $v = \beta_1u_1+\cdots+\beta_nu_n$. Então,

$$\langle T(u),T(v)\rangle = \left\langle \sum_{i=1}^{n}\alpha_iT(u_i), \sum_{j=1}^{n}\beta_jT(u_j)\right\rangle = \sum_{i=1}^{n}\alpha_i\overline{\beta}_i.$$

Por outro lado,

$$\langle u,v\rangle = \left\langle \sum_{i=1}^{n}\alpha_iu_i, \sum_{j=1}^{n}\beta_ju_j\right\rangle = \sum_{i=1}^{n}\alpha_i\overline{\beta}_i.$$

Logo, $\langle T(u),T(v)\rangle = \langle u,v\rangle$, donde segue que T preserva produto interno. ■

Exercício 5.1 Sejam U e V espaços com produto interno e $T : U \to V$ uma função tal que $\langle T(u),T(v)\rangle = \langle u,v\rangle$ para todos $u,v \in U$. Mostre que T é transformação linear.

■ **Exemplo 5.1.7** Sejam $U = \mathscr{C}([0,1],\mathbb{R})$ com produto interno

$$\langle f,g\rangle = \int_0^1 x^2f(x)g(x)\,dx,$$

e $V = \mathscr{C}([0,1],\mathbb{R})$ com produto interno usual. A transformação linear $T : U \to V$ dada por $T(f(x)) = xf(x)$, preserva produto interno mas não é um isomorfismo [Coe].

Prova. Note que o Teorema 5.1.6 não pode ser aplicado, pois $\dim(U) = \dim(V)$ mas não é finita. Sejam $f,g \in U$, então

$$\langle T(f),T(g)\rangle = \int_0^1 T(f(x))T(g(x))dx = \int_0^1 x^2f(x)g(x)dx = \langle f,g\rangle.$$

Logo, T preserva produto interno. Por outro lado, a função $g(x) = 1 \in V$ mas $g \notin \operatorname{Im}T$. Portanto, T não é sobrejetora. ■

Definição 5.1.8 Sejam U e V dois $\mathbb{K}$-espaços vetoriais normados. Uma transformação linear $T : U \to V$ que satisfaz $||T(u)|| = ||u||$ para todo $u \in U$ é chamado de *isometria*.

Proposição 5.1.9 Uma transformação linear $T : U \to V$ preserva produto interno se, e somente se, T é isometria.

Demonstração. Basta usar as Identidades de Polarização. ■

■ **Exemplo 5.1.10** Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a rotação de ângulo θ (sentido anti-horário). Mostre que T é uma isometria em $\mathbb{R}^2$ com produto interno usual.

Prova. Seja $B = \{e_1,e_2\}$ a base canônica de $\mathbb{R}^2$. A Figura 5.1 exemplifica a ação do operador T na base B. Note que,

$$T(e_1) = \cos(\theta)e_1 + \operatorname{sen}(\theta)e_2$$

$$T(e_2) = \cos\left(\frac{\pi}{2}+\theta\right)e_1 + \operatorname{sen}\left(\frac{\pi}{2}+\theta\right)e_2 = -\operatorname{sen}(\theta)e_1 + \cos(\theta)e_2.$$

Logo, se (x,y) é um vetor qualquer de $\mathbb{R}^2$, temos

$$\begin{aligned} T(x,y) &= xT(e_1) + yT(e_2) = x(\cos(\theta)e_1 + \operatorname{sen}(\theta)e_2) + y(-\operatorname{sen}(\theta)e_1 + \cos(\theta)e_2) \\ &= (x\cos\theta - y\operatorname{sen}\theta, x\operatorname{sen}\theta + y\cos\theta). \end{aligned}$$

Assim, considerando a norma usual de $\mathbb{R}^2$ (proveniente do produto interno usual),

$$||T(x,y)||^2 = (x\cos\theta - y\operatorname{sen}\theta)^2 + (x\operatorname{sen}\theta + y\cos\theta)^2 = x^2 + y^2 = ||(x,y)||^2.$$

Portanto, T é uma isometria.

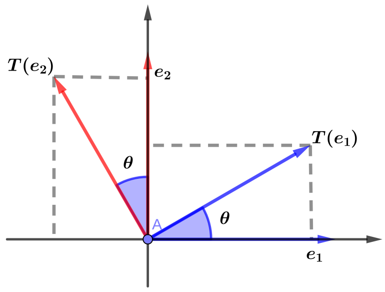


Figura 5.1: Rotação de ângulo θ aplicada nos vetores da base B.

■

■ **Exemplo 5.1.11** Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x,y) = \frac{1}{5}(4x+3y, 3x-4y)$. Considerando $\mathbb{R}^2$ com produto interno usual, mostre que T preserva produto interno.

Prova. Vamos usar a equivalência: T preserva produto interno se, e somente se, T leva alguma base ortonormal de $\mathbb{R}^2$ em uma base ortonormal de $\mathbb{R}^2$.

Seja $B = \{e_1, e_2\}$ a base canônica de $\mathbb{R}^2$. Então,

$$T(e_1) = \frac{1}{5}(4,3) \text{ e } T(e_2) = \frac{1}{5}(3,-4).$$

Como B é uma base ortonormal e os vetores $\frac{1}{5}(4,3)$, $\frac{1}{5}(3,-4)$ também são ortonormais, segue que T preserva produto interno. ■

■ **Exemplo 5.1.12** Determine os valores positivos a, b, c, d, e para que o operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ dado por $T(x,y,z) = (2ax - 2ay + az, bx + by, cx - dy - ez)$ seja uma isometria em $\mathbb{R}^3$ com o produto interno usual.

Solução. A base canônica $\{e_1, e_2, e_3\}$ de $\mathbb{R}^3$ é ortonormal em relação ao produto interno usual. Logo, $\{T(e_1), T(e_2), T(e_3)\}$ também deve ser um conjunto ortonormal. Ou seja,

$$0 = \langle T(e_1), T(e_2)\rangle = \langle (2a,b,c), (-2a,b,-d)\rangle = -4a^2 + b^2 - cd$$

$$0 = \langle T(e_1, T(e_3))\rangle = \langle (2a,b,c), (a,0,-e)\rangle = 2a^2 - ce$$

$$0 = \langle T(e_2), T(e_3)\rangle = \langle (-2a,b,-d), (a,0,-e)\rangle = -2a^2 + de$$

$$1 = ||T(e_1)||^2 = ||(2a,b,c)||^2 = 4a^2 + b^2 + c^2$$

$$1 = ||T(e_2)||^2 = ||(-2a, b, -d)||^2 = 4a^2 + b^2 + d^2$$
$$1 = ||T(e_3)||^2 = ||(a, 0, e)||^2 = a^2 + e^2$$

Resolvendo o sistema não linear acima, obtemos (verifique)

$$a = \frac{1}{3},\ b = \frac{\pm\sqrt{2}}{2},\ c = d = \frac{\sqrt{2}}{6},\ e = \frac{2\sqrt{2}}{3}.$$

■

O próximo resultado descreve uma propriedade interessante da matriz de uma isometria em relação à bases ortonormais.

Proposição 5.1.13 Sejam U e V dois $\mathbb{K}$-espaços vetoriais, de mesma dimensão finita n, com produto interno, B e C bases ortonormais de U e V, respectivamente, e $T : U \to V$ uma transformação linear. Então, T é uma isometria se, e somente se, $[T]_{BC}$ é uma matriz unitária.

Demonstração. Sejam $B = \{u_1, \ldots, u_n\}$ e $C = \{v_1, \ldots, v_n\}$ e $[T]_{BC} = (a_{ij}) \in M_n(\mathbb{K})$. Daí, para cada $j = 1, \ldots, n$, $T(u_j) = \sum_{i=1}^{n} a_{ij} v_i$. Note que, para $i, j = 1, \ldots, n$,

$$\left(\overline{[T]}_{BC}^{\top}[T]_{BC}\right)_{ij} = \sum_{k=1}^{n} \overline{a_{ki}} a_{kj}.$$

Além disso, como C é ortonormal, para $i, j = 1, \ldots, n$,

$$\langle T(u_i), T(u_j)\rangle = \left\langle \sum_{k=1}^{n} a_{ki} v_k, \sum_{r=1}^{n} a_{rj} v_r \right\rangle = \sum_{k=1}^{n}\sum_{r=1}^{n} a_{ki}\overline{a_{rj}}\langle v_k, v_r\rangle = \sum_{k=1}^{n} a_{ik}\overline{a_{kj}} = \left(\overline{[T]}_{BC}^{\top}[T]_{BC}\right)_{ij}. \quad (5.1)$$

($\Rightarrow$) Como T é isometria, logo T preserva produto interno, e B é ortonormal, temos

$$\delta_{ij} = \langle u_i, u_j\rangle = \langle T(u_i), T(u_j)\rangle \overset{(5.1)}{=} \left(\overline{[T]}_{BC}^{\top}[T]_{BC}\right)_{ij},$$

em que $\delta_{ii} = 1$ e $\delta_{ij} = 0$ para todo $i \neq j$. Assim, $\left(\overline{[T]}_{BC}^{\top}[T]_{BC}\right)_{ij} = \delta_{ij}$, donde segue que $\overline{[T]}_{BC}^{\top}[T]_{BC} = I_n$. Portanto, $[T]_{BC}$ é uma matriz unitária.

($\Leftarrow$) Como T é unitária, para $i, j = 1, \ldots, n$,

$$\delta_{ij} = \left(\overline{[T]}_{BC}^{\top}[T]_{BC}\right)_{ij} \overset{(5.1)}{=} \langle T(u_i), T(u_j)\rangle.$$

Logo, $T(B)$ é ortonormal. Portanto, segue da Proposição 5.1.6 que T preserva produto interno, donde segue que T é uma isometria.

■

■ **Exemplo 5.1.14** Verifique se o operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz em relação à base canônica, C, é igual a

$$[T]_C = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix},$$

é uma isometria.

Solução. Como a base canônica é ortonormal e $[T]_C$ possui colunas ortonormais, segue que $[T]_C$ é unitária. Portanto, T é uma isometria. ■

O resultado a seguir estabelece a relação entre uma isometria e os seus autovalores.

Proposição 5.1.15 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $T : V \to V$ um operador linear.

(i) Se T é uma isometria e λ é autovalor de T, então $|\lambda| = 1$.

(ii) Se $\{v_1, \dots, v_n\}$ é uma base ortonormal de V tal que $T(v_i) = \lambda_i v_i$ e $|\lambda_i| = 1$ para todo $i = 1, \dots, n$, então T é uma isometria.

Demonstração. **(i)** Seja λ um autovalor de T. Então, existe $v \in V$, $v \neq 0$, tal que $T(v) = \lambda v$. Como T é uma isometria,

$$||v||^2 = ||T(v)||^2 = \langle T(v), T(v) \rangle = \langle \lambda v, \lambda v \rangle = |\lambda|^2 ||v||^2 \quad \Rightarrow \quad |\lambda| = 1.$$

(ii) Como $\{v_1, \dots, v_n\}$ é base ortonormal de V, basta mostrarmos que $\{T(v_1), \dots, T(v_n)\}$ também é base ortonormal de V. De fato, para $i, j \in \{1, \dots, n\}$, temos que

$$\langle T(v_i), T(v_j) \rangle = \langle \lambda_i v_i, \lambda_j v_j \rangle = \lambda_i \overline{\lambda}_j \langle v_i, v_j \rangle = 0,$$

se $i \neq j$, e

$$||T(v_i)|| = ||\lambda_i v_i|| = |\lambda_i| ||v_i|| = 1.$$

Portanto, $\{T(v_1), \dots, T(v_n)\}$ é base ortonormal de V e T é uma isometria.

■

Exercício 5.2 Considere $\mathbb{R}^2$ com produto interno usual. Seja $R : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão em relação à reta $y = 2x$. Descreva R explicitamente, ou seja, $R(x, y)$, e verifique se R é uma isometria.

5.2 O operador adjunto

A definição de operador adjunto será introduzida nesta seção, após um teorema que garantirá sua existência e unicidade em dimensão finita. Esse operador terá um papel importante no próximo objetivo, que é o de responder a pergunta: dado um operador linear $T : V \to V$, sob quais condições existe base ortonormal de V formada por autovetores de T? Além de responder essa pergunta, veremos as propriedades adicionais que conseguimos obter a partir disso, para operadores e matrizes.

Teorema 5.2.1 Sejam U e V dois $\mathbb{K}$-espaços vetoriais com produto interno, U com dimensão finita, e $T : U \to V$ uma transformação linear. Então, existe uma única transformação linear $T^* : V \to U$ tal que $\langle T(u), v \rangle = \langle u, T^*(v) \rangle$ para todos $u \in U$, $v \in V$.

Demonstração. Considere $B = \{u_1, \dots, u_n\}$ uma base ortonormal para U. Vamos definir $T^* : V \to U$. Seja $v \in V$. Dado $u \in U$, como B é ortonormal, segue da Proposição 3.2.10 que

$$u = \sum_{i=1}^{n} \langle u, u_i \rangle u_i.$$

Daí,

$$\langle T(u),v\rangle = \left\langle T\left(\sum_{i=1}^{n}\langle u,u_i\rangle u_i\right),v\right\rangle = \left\langle \sum_{i=1}^{n}\langle u,u_i\rangle T(u_i),v\right\rangle = \sum_{i=1}^{n}\langle u,u_i\rangle\langle T(u_i),v\rangle.$$

Como $\langle T(u_i),v\rangle \in \mathbb{K}$, $i = 1,\ldots,n$, temos

$$\sum_{i=1}^{n}\langle u,u_i\rangle\langle T(u_i),v\rangle = \sum_{i=1}^{n}\langle u,\overline{\langle T(u_i),v\rangle}u_i\rangle = \left\langle u,\sum_{i=1}^{n}\overline{\langle T(u_i),v\rangle}u_i\right\rangle.$$

Defina $T^*(v) = \sum_{i=1}^{n}\overline{\langle T(u_i),v\rangle}u_i$. Pelo exposto anteriormente,

$$\langle T(u),v\rangle = \langle u,T^*(v)\rangle \quad \text{para todo } u \in U. \tag{5.2}$$

Se $\bar{u} \in U$ é outro vetor tal que

$$\langle T(u),v\rangle = \langle u,\bar{u}\rangle \quad \text{para todo } u \in U,$$

então,

$$\langle u,\bar{u}\rangle = \langle T(u),v\rangle = \langle u,T^*(v)\rangle\ \forall\, u\in U \Rightarrow \langle u,\bar{u}-T^*(v)\rangle = 0\ \forall\, u\in U \Rightarrow \bar{u}-T^*(v) = 0 \Rightarrow u = T^*(v).$$

Assim, $T^*(v)$ é o único vetor em U satisfazendo (5.2), donde seque que a função $T^* : V \to U$ fica bem definida. Por construção, temos

$$\langle T(u),v\rangle = \langle u,T^*(v)\rangle \quad \text{para todos} \ \ u \in U, v \in V. \tag{5.3}$$

Mostremos que T^* é transformação linear. Sejam $v_1,v_2 \in V$ e $\lambda \in \mathbb{K}$. Então, para todo $u \in U$, temos

$$\begin{aligned}
\langle u,T^*(v_1+\lambda v_2)\rangle = \langle T(u),v_1+\lambda v_2\rangle = \langle T(u),v_1\rangle + \overline{\lambda}\langle T(u),v_2\rangle = \langle u,T^*(v_1)\rangle + \overline{\lambda}\langle u,T^*(v_2)\rangle \\
= \langle u,T^*(v_1)\rangle + \langle u,\lambda T^*(v_2)\rangle = \langle u,T^*(v_1)+\lambda T^*(v_2)\rangle \\
\Rightarrow \langle u,T^*(v_1+\lambda v_2) - T^*(v_1) - \lambda T^*(v_2)\rangle = 0 \quad \text{para todo} \ \ u \in U \\
\Rightarrow T^*(v_1+\lambda v_2) - T^*(v_1) - \lambda T^*(v_2) = 0 \\
\Rightarrow T^*(v_1+\lambda v_2) = T^*(v_1) + \lambda T^*(v_2).
\end{aligned}$$

Portanto, T^* é uma transformação linear que satisfaz (5.3). Para a unicidade, suponha que $S : V \to U$ é uma transformação linear que satisfaz

$$\langle u,S(v)\rangle = \langle T(u),v\rangle \overset{(5.3)}{=} \langle u,T^*(v)\rangle \quad \text{para todos} \ \ u \in U, v \in V.$$

Então, para todos $u \in U, v \in V$,

$$\langle u,T^*(v) - S(v)\rangle = 0 \Rightarrow T^*(v) = S(v).$$

Logo, $S = T^*$. ■

Definição 5.2.2 Sejam U e V dois $\mathbb{K}$-espaços vetoriais com produto interno e $T : U \to V$ uma transformação linear. Dizemos que T possui *adjunto* se existe uma transformação linear $T^* : V \to U$ tal que $\langle T(u),v\rangle = \langle u,T^*(v)\rangle$ para todos $u \in U$ e $v \in V$. Neste caso, T^* é chamado de *adjunto* de T.

Observação 5.2.3 **(i)** O adjunto, quanto existe, depende dos produtos internos de U e V e é único. A demonstração da "unicidade" no Teorema 5.2.1 se aplica em geral.

(ii) Pelo Teorema 5.2.1, se U tem dimensão finita, sempre existe o adjunto $T^*: V \to U$ de $T: U \to V$.

■ **Exemplo 5.2.4** Seja $T: \mathbb{R}^3 \to \mathbb{R}^2$ a transformação linear dada por $T(x_1,x_2,x_3) = (x_3, x_1 - 2x_2)$. Determine o adjunto de T.

Solução. Seja $T^*: \mathbb{R}^2 \to \mathbb{R}^3$ o adjunto de T, então

$$\langle u, T^*(v)\rangle = \langle T(u), v\rangle, \text{ para todos } u \in \mathbb{R}^3 \text{ e } v \in \mathbb{R}^2.$$

Tome $v = (y_1,y_2) \in \mathbb{R}^2$, então para todo $u = (x_1,x_2,x_3) \in \mathbb{R}^3$ temos que

$$\langle (x_1,x_2,x_3), T^*(y_1,y_2)\rangle = \langle T(x_1,x_2,x_3),(y_1,y_2)\rangle = \langle (x_3,x_1-2x_2),(y_1,y_2)\rangle = y_1x_3 + y_2(x_1-2x_2)$$

$$= x_1y_2 + -2x_2y_2 + x_3y_1 = \langle (x_1,x_2,x_3),(y_2,-2y_2,y_1)\rangle.$$

Segue que

$$\langle (x_1,x_2,x_3), T^*(y_1,y_2)\rangle = \langle (x_1,x_2,x_3),(y_2,-2y_2,y_1)\rangle,$$

ou seja,

$$\langle (x_1,x_2,x_3), T^*(y_1,y_2) - (y_2,-2y_2,y_1)\rangle = 0$$

para todo $(x_1,x_2,x_3) \in \mathbb{R}^3$. Portanto,

$$T^*(y_1,y_2) = (y_2,-2y_2,y_1).$$

■

O próximo exemplo é uma tradução livre do Exemplo 7.4 de [Axl].

■ **Exemplo 5.2.5** Dados $u_0 \in U$ e $v_0 \in V$ fixos, defina $T: U \to V$ por $T(u) = \langle u, u_0\rangle v_0$. Determine T^*.

Solução. Seja $T^*: V \to U$ o adjunto de T, então

$$\langle u, T^*(v)\rangle = \langle T(u), v\rangle, \text{ para todos } u \in U \text{ e } v \in V.$$

Tome $u \in U$, então para todo $v \in V$ temos que

$$\langle u, T^*(v)\rangle = \langle T(u), v\rangle = \langle \langle u,u_0\rangle v_0, v\rangle = \langle u,u_0\rangle\langle v_0,v\rangle = \langle u, \overline{\langle v_0,v\rangle} u_0\rangle.$$

Seque que,

$$\langle u, T^*(v)\rangle = \langle u, \langle v,v_0\rangle u_0\rangle,$$

ou seja,

$$\langle u, T^*(v) - \langle v,v_0\rangle u_0\rangle = 0$$

para todo $u \in U$. Portanto,

$$T^*(v) = \langle v,v_0\rangle u_0.$$

■

■ **Exemplo 5.2.6** Considere $V = M_2(\mathbb{R})$ com produto interno usual $\langle A,B\rangle = \operatorname{tr}\left(A\overline{B}^{\top}\right) = \operatorname{tr}\left(AB^{\top}\right)$, $M \in V$ uma matriz fixa e $T : V \to V$ a transformação linear definida por $T(A) = MA - AM$. Determine T^*.

Solução. Para todas $A, B \in V$, temos

$$\begin{aligned}\langle T(A),B\rangle = \langle MA-AM,B\rangle = \operatorname{tr}(B^{\top}(MA-AM)) &= \operatorname{tr}(B^{\top}MA) - \operatorname{tr}(B^{\top}AM) = \operatorname{tr}(B^{\top}MA) - \operatorname{tr}(MB^{\top}A)\\ &= \operatorname{tr}(B^{\top}MA - MB^{\top}A) = \operatorname{tr}((B^{\top}M - MB^{\top})A) = \langle A, M^{\top}B - BM^{\top}\rangle.\end{aligned}$$

Portanto,

$$T^*(B) = M^{\top}B - BM^{\top}.$$

■

Nos exemplos anteriores foi fácil determinar $T^*(v)$ diretamente pela relação: $\langle T(u),v\rangle = \langle u,T^*(v)\rangle$ para todos $u \in U$, $v \in V$. Quando essa identificação não for imediata, podemos utilizar a fórmula descrita a seguir, que é uma consequência da demonstração do Teorema 5.2.1.

Proposição 5.2.7 Sejam U e V dois $\mathbb{K}$-espaços vetoriais com produto interno, U com dimensão finita, e $T : U \to V$ uma transformação linear. Se $B = \{u_1, \ldots, u_n\}$ é uma base ortonormal para U, então o adjunto $T^* : V \to U$ é dado por

$$T^*(v) = \sum_{i=1}^{n} \overline{\langle T(u_i), v\rangle} u_i, \quad \text{para todo } v \in V.$$

A seguir, apresentamos algumas propriedades algébricas dos adjuntos.

Proposição 5.2.8 Considere U e V dois $\mathbb{K}$-espaços vetoriais com produto interno. Sejam $S, T : U \to V$ transformações lineares que admitem adjuntos e $\lambda \in \mathbb{K}$. Então:

(i) $S+T$ admite adjunto e $(S+T)^* = S^* + T^*$;

(ii) λT admite adjunto e $(\lambda T)^* = \overline{\lambda} T^*$;

(iii) se $U = V$, $T \circ S$ admite adjunto e $(T \circ S)^* = S^* \circ T^*$;

(iv) T^* admite adjunto e $(T^*)^* = T$.

Demonstração. **(i)** Para todos $u \in U$ e $v \in V$, temos

$$\begin{aligned}\langle (S+T)(u),v\rangle &= \langle S(u)+T(u),v\rangle = \langle S(u),v\rangle + \langle T(u),v\rangle\\ &= \langle u,S^*(v)\rangle + \langle u,T^*(v)\rangle = \langle u,(S^*+T^*)(v)\rangle.\end{aligned}$$

Portanto, $(S+T)^* = S^* + T^*$.

(ii) Para todos $u \in U$ e $v \in V$, temos

$$\langle (\lambda T)(u),v\rangle = \langle \lambda T(u),v\rangle = \lambda\langle T(u),v\rangle = \lambda\langle u,T^*(v)\rangle = \langle u,\overline{\lambda}T^*(v)\rangle = \langle u,(\overline{\lambda}T^*)(v)\rangle.$$

Portanto, $(\lambda T)^* = \overline{\lambda} T^*$.

(iii) Para todos $u, v \in U = V$, temos

$$\langle (T\circ S)(u),v\rangle = \langle T(S(u)),v\rangle = \langle S(u),T^*(v)\rangle = \langle u,S^*(T^*(v))\rangle = \langle u,(S^*\circ T^*)(v)\rangle.$$

Portanto, $(T \circ S)^* = S^* \circ T^*$.

(iv) Para todos $u \in U$ e $v \in V$, temos

$$\langle T^*(v), u\rangle = \overline{\langle u, T^*(v)\rangle} = \overline{\langle T(u), v\rangle} = \langle v, T(u)\rangle.$$

Portanto, $(T^*)^* = T$.

■

O próximo resultado estabelece a relação entre as matrizes de T e T^* em relação a bases ortonormais.

Proposição 5.2.9 Sejam U e V espaços vetoriais de dimensão finita com produto interno e $T : U \to V$ uma transformação linear. Se B é base ortonormal de U e C é base ortonormal de V, então

$$[T^*]_{CB} = \overline{[T]}_{BC}^{\top}.$$

Demonstração. Sejam $B = \{u_1, \ldots, u_n\}$ e $C = \{v_1, \ldots, v_m\}$ bases ortonormais de U e V, respectivamente. Então,

$$[T]_{BC} = \begin{bmatrix} | & | & & | \\ [T(u_1)]_C & [T(u_2)]_C & \cdots & [T(u_n)]_C \\ | & | & & | \end{bmatrix}.$$

Além disso, se $[T]_{BC} = (a_{ij})_{m\times n}$, então, para cada $j = 1, \ldots, n$,

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i,$$

daí, para cada $j = 1, \ldots, n$,

$$\langle T(u_j), v_i\rangle = \left\langle \sum_{k=1}^{m} a_{kj} v_k, v_i \right\rangle = \sum_{k=1}^{m} a_{kj} \langle v_k, v_i\rangle = a_{ij}\langle v_i, v_i\rangle = a_{ij}||v_i||^2 = a_{ij}.$$

Logo,

$$T(u_j) = \sum_{i=1}^{m} \langle T(u_j), v_i\rangle v_i,$$

ou seja, $([T]_{BC})_{ij} = \langle T(u_j), v_i\rangle$.

Por outro lado, pela Proposição 5.2.7, para cada $j = 1, \ldots, n$,

$$T^*(v_j) = \sum_{i=1}^{n} \overline{\langle T(u_i), v_j\rangle} u_i,$$

ou seja, $([T^*]_{CB})_{ij} = \overline{\langle T(u_i), v_j\rangle} = (\overline{[T]_{BC}})_{ji}$. Portanto, $[T^*]_{CB} = \overline{[T]}_{BC}^{\top}$. ■

■ **Exemplo 5.2.10** Sejam $\mathbb{R}^2$ com produto interno usual e $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear defina por $T(x, y) = (x + y, x)$. Determine T^*.

Solução. Seja $B = \{e_1, e_2\}$ a base canônica de $\mathbb{R}^2$. Como B é ortonormal no produto interno usual, $T(e_1) = (1, 1)$ e $T(e_2) = (1, 0)$, pela Proposição 5.2.7,

$$T^*(x, y) = \sum_{i=1}^{2} \langle T(e_i), (x, y)\rangle e_i = (x + y)e_1 + xe_2 = (x + y, x) = T(x, y).$$

Uma outra maneira, seria encontrar

$$[T]_B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

e como B é ortonormal no produto interno usual, pela Proposição 5.2.9, temos

$$[T^*]_B = \overline{[T]}_B^\top \overset{\mathbb{K}=\mathbb{R}}{=} [T]_B^\top = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = [T]_B.$$

Portanto, $T^* = T$. ■

■ **Exemplo 5.2.11** Sejam $\mathbb{C}^2$ com produto interno usual e $T : \mathbb{C}^2 \to \mathbb{C}^2$ a transformação linear definida por $T(x_1, x_2) = (x_1 + 2x_2, x_2 - ix_1)$. Determine T^* e as matrizes $[T]_{BC}$ e $[T^*]_{CB}$ para $B = \{e_1, e_2\}$, a base canônica de $\mathbb{C}^3$, e $C = \{e_1, 2e_2\}$.

Solução. Temos $T(e_1) = (1, -i)$ e $T(e_2) = (2, 1)$, logo, como B é ortonormal, segue da Proposição 5.2.7 que

$$\begin{aligned} T^*(x_1, x_2) &= \sum_{i=1}^{2} \overline{\langle T(e_i), (x_1, x_2)\rangle} e_i \\ &= \overline{\langle (1, -i), (x_1, x_2)\rangle} e_1 + \overline{\langle (2, 1), (x_1, x_2)\rangle} e_2 \\ &= \overline{(\overline{x_1} + \overline{x_2}(-i))} e_1 + \overline{(2\overline{x_1} + \overline{x_2})} e_2 \\ &= (x_1 + ix_2, 2x_1 + x_2). \end{aligned}$$

Além disso,

$$[T]_{BC} = \begin{bmatrix} 1 & 2 \\ -i/2 & 1/2 \end{bmatrix} \text{ e } [T^*]_{CB} = \begin{bmatrix} 1 & 2i \\ 2 & 2 \end{bmatrix}.$$

Note que $[T^*]_{CB} \neq \overline{[T]}_{BC}^\top$, mas isso ocorreu pois a base C não é ortonormal. ■

■ **Exemplo 5.2.12** Considere $\mathbb{R}^3$ com produto interno usual, B a base canônica de $\mathbb{R}^3$ e $T : \mathbb{R}^3 \to \mathbb{R}^3$ transformação linear tal que

$$[T]_B = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Determine T^*.

Solução. Como B é ortonormal no produto interno usual, temos

$$[T^*]_B = \overline{[T]}_B^\top \overset{\mathbb{K}=\mathbb{R}}{=} [T]_B^\top = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}.$$

Daí,

$$T^*(e_1) = (2, 1, -1), \quad T^*(e_2) = (0, 1, 1), \quad T^*(e_3) = (-1, 0, 1).$$

Logo,

$$T^*(x, y, z) = xT^*(e_1) + yT^*(e_2) + zT^*(e_3) = (2x - z, x + y, y + z - x).$$

■

Definição 5.2.13 Seja V um $\mathbb{K}$-espaço vetorial com produto interno e $T : V \to V$ um operador linear. Dizemos que T é *unitário* se admite adjunto e $T \circ T^* = T^* \circ T = \mathrm{Id}_V$, isto é, T é isomorfismo e $T^{-1} = T^*$.

Proposição 5.2.14 Seja V um $\mathbb{K}$-espaço vetorial com produto interno e $T : V \to V$ um operador linear. Então, T é unitário se, e somente se, T é um isomorfismo entre espaços com produto interno.

Demonstração. ($\Rightarrow$) Supondo T unitário, segue que T é isomorfismo e $T^{-1} = T^*$. Resta mostrar que T preserva produto interno. Dados $u, v \in V$, temos

$$\langle T(u), T(v)\rangle = \langle u, T^*(T(v))\rangle = \langle u, (T^* \circ T)(v)\rangle = \langle u, \mathrm{Id}_V(v)\rangle = \langle u, v\rangle.$$

($\Leftarrow$) Temos que T é isomorfismo e preserva produto interno. Logo, para todos $u, v \in V$, temos

$$\langle T(u), v\rangle = \langle T(u), (T \circ T^{-1})(v)\rangle = \langle T(u), T(T^{-1}(v))\rangle = \langle u, T^{-1}(v)\rangle.$$

Portanto, $T^{-1} = T^*$. ■

Corolário 5.2.15 Sejam V um $\mathbb{K}$-espaços vetorial de dimensão finita com produto interno, $T : V \to V$ um operador linear e B uma base ortonormal de V. Então, T é unitário se, e somente se, $[T]_B$ é uma matriz unitária.

Demonstração. ($\Rightarrow$) Como T é unitário, segue que $T^{-1} = T^*$. Assim, pelo Corolário 4.3.9 e Proposição 5.2.9,

$$[T]_B^{-1} = [T^{-1}]_B = [T^*]_B = \overline{[T]_B}^{\top},$$

logo $[T]_B$ é unitária.

($\Leftarrow$) Como $[T]_B$ é uma matriz unitária, ela é não singular e $[T]_B^{-1} = \overline{[T]_B}^{\top}$. Novamente, pelo Corolário 4.3.9 e Proposição 5.2.9 segue T é isomorfismo e $[T^{-1}]_B = [T^*]_B$, donde concluímos que $T^{-1} = T^*$. ■

■ **Exemplo 5.2.16** Descreva todas as matrizes $A \in M_2(\mathbb{R})$ ortogonais (unitárias).

Solução. Se $A \in M_2(\mathbb{R})$ ortogonal, isto é, A é não singular e $A^{-1} = A^\top$. Assim,

$$1 = \det(I_2) = \det(A^\top A) = \det(A^\top)\det A = (\det A)^2 \;\Rightarrow\; \det A = \pm 1.$$

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

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{e} \quad A^{-1} = A^\top = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

temos duas possibilidades:

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{com} \quad a^2 + b^2 = 1 \quad (\text{se } \det A = 1);$$

ou

$$A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \quad \text{com} \quad a^2 + b^2 = 1 \quad (\text{se } \det A = -1).$$

Como $a^2 + b^2 = 1$, existe único $\theta \in [0, 2\pi)$ tal que $a = \cos\theta$ e $b = \operatorname{sen}\theta$, logo

$$A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}.$$

No primeiro caso, A é a matriz de rotação por θ no sentido anti-horário, e no segundo caso A é semelhante à matriz $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (verifique) e portanto é uma reflexão. ■

5.3 Operadores autoadjuntos

No restante desta seção, estaremos interessados em estudar propriedades de operadores lineares $T: V \to V$ e seus adjuntos $T^*: V \to V$. Por exemplo, quando eles comutam (serão chamados de normais) ou coincidem (serão chamados de autoadjuntos), e quais consequências podemos obter em cada caso. Tais propriedades nos levarão ao importante Teorema Espectral, que dará condições necessárias e suficientes para que o seja ortogonalmente diagonalizável.

Definição 5.3.1 Seja $T: V \to V$ um operador linear em um espaço com produto interno. Dizemos que T é *autoadjunto* (ou *hermitiano*) se T admite adjunto e, além disso, $T^* = T$. Em outras palavras, T é autoadjunto se, e somente se, $\langle T(u), v\rangle = \langle u, T(v)\rangle$, para todos $u, v \in V$.
Se $\mathbb{K} = \mathbb{R}$, também dizemos que T é *simétrico*.

■ **Exemplo 5.3.2** O operador $T: \mathbb{R}^2 \to \mathbb{R}^2$ visto no Exemplo 5.2.10 é autoadjunto. ■

Proposição 5.3.3 Seja $T: V \to V$ um operador linear em um $\mathbb{K}$-espaço vetorial V com produto interno e dimensão finita. São equivalentes:

(i) T é autoadjunto;

(ii) $\overline{[T]}_B^{\top} = [T]_B$ para toda base ortonormal B de V;

(iii) existe uma base ortonormal B de V tal que $\overline{[T]}_B^{\top} = [T]_B$.

Demonstração. Segue da Proposição 5.2.9. ■

■ **Exemplo 5.3.4** Sejam $\mathbb{C}^2$ com produto interno usual e T o operador linear em $\mathbb{C}^2$ cuja matriz em relação à base canônica B é

$$[T]_B = \begin{bmatrix} 2 & b \\ 3 & ic \end{bmatrix}.$$

Determine para quais valores de b e c reais temos que T é autoadjunto.

Solução. Como a base canônica de $\mathbb{C}^2$ é uma base ortonormal em relação ao produto interno usual, podemos usar a Proposição 5.3.3. Assim,

$$\overline{[T]}_B^{\top} = \begin{bmatrix} 2 & 3 \\ b & -ic \end{bmatrix} = \begin{bmatrix} 2 & b \\ 3 & ic \end{bmatrix} \quad \Rightarrow \quad b = 3 \ \text{ e } \ c = 0.$$

■

Observação 5.3.5 Se $T: V \to V$ é autoadjunto, $\dim V = n$, B é uma base ortonormal de V e $[T]_B = (a_{ij})$, então $a_{ij} = \overline{a_{ji}}$ para todos $i, j = 1, \ldots, n$. Em particular, os elementos da diagonal principal de $[T]_B$ são números reais.

Como visto anteriormente, dado um operador linear $T: V \to V$, basta conhecermos o comportamento de T em uma base de V para obtermos toda a informação sobre T. A seguir veremos algo semelhante para determinar se o operador T é autoadjunto.

Proposição 5.3.6 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno, $B = \{v_1, \ldots, v_n\}$ uma base de V e $T : V \to V$ um operador linear. Se $\langle T(v_i), v_j\rangle = \langle v_i, T(v_j)\rangle$, para todos $i, j = 1, \ldots, n$, então T é autoadjunto.

Demonstração. Sejam $u, v \in V$ vetores quaisquer. Então, existem escalares α_i, β_i, $i = 1, \ldots, n$, tais que

$$u = \sum_{i=1}^{n} \alpha_i v_i \ \text{ e } \ v = \sum_{i=1}^{n} \beta_i v_i.$$

Logo,

$$\begin{aligned}\langle T(u), v\rangle &= \left\langle \sum_{i=1}^{n} \alpha_i T(v_i), \sum_{j=1}^{n} \beta_j v_j \right\rangle = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \overline{\beta}_j \langle T(v_i), v_j\rangle = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n} \overline{\beta}_j \langle v_i, T(v_j)\rangle \\ &= \left\langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{j=1}^{n} \beta_j T(v_j) \right\rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, T\left(\sum_{j=1}^{n} \beta_j v_j\right) \right\rangle = \langle u, T(v)\rangle,\end{aligned}$$

e, portanto, T é autoadjunto. ■

■ **Exemplo 5.3.7** Considere $\mathbb{R}^2$ com produto interno usual. Verifique se os operadores lineares definidos a seguir são autoadjuntos.

(i) $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(1,2) = (3,1)$ e $T(-1,1) = (0,-1)$.

(ii) $S : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $S(1,2) = (-1,4)$ e $S(-1,1) = (-4,1)$.

Solução. Note que $B = \{(1,2), (-1,1)\}$ é base de $\mathbb{R}^2$. Além disso, como $\mathbb{K} = \mathbb{R}$, $\langle T(v), v\rangle = \overline{\langle v, T(v)\rangle} = \langle v, T(v)\rangle$ para todo $v \in V$.

(i) Como

$$\langle T(1,2), (-1,1)\rangle = \langle (3,1), (-1,1)\rangle = -2,$$
$$\langle (1,2), T(-1,1)\rangle = \langle (1,2), (0,-1)\rangle = -2,$$

segue que T é autoadjunto.

(ii) Como

$$\langle S(1,2), (-1,1)\rangle = \langle (-1,4), (-1,1)\rangle = 5$$

e

$$\langle (1,2), S(-1,1)\rangle = \langle (1,2), (-4,1)\rangle = -2$$

segue que S não é autoadjunto.

■

Na sequência veremos algumas propriedades de operadores adjuntos em um espaço vetorial complexo.

Proposição 5.3.8 Sejam V um $\mathbb{C}$-espaço vetorial com produto interno e $T : V \to V$ um operador linear. São equivalentes:

(i) $T = 0$;

(ii) $\langle T(u), u\rangle = 0$ para todo $u \in V$;

(iii) $\langle T(u), v\rangle = 0$ para todos $u, v \in V$.

Demonstração. **[(i) $\Rightarrow$ (ii)]** Imediado.

[(ii) $\Rightarrow$ (iii)] Sejam $u, v \in V$. Quaisquer que sejam $\alpha, \beta \in \mathbb{C}$, tomando $w = \alpha u + \beta v$, temos

$$\begin{aligned} 0 = \langle T(w), w\rangle = \langle T(\alpha u + \beta v), \alpha u + \beta v\rangle &= \langle \alpha T(u) + \beta T(v), \alpha u + \beta v\rangle \\ = \alpha\overline{\alpha}\langle T(u), u\rangle + \alpha\overline{\beta}\langle T(u), v\rangle &+ \beta\overline{\alpha}\langle T(v), u\rangle + \beta\overline{\beta}\langle T(v), v\rangle \\ &= \alpha\overline{\beta}\langle T(u), v\rangle + \beta\overline{\alpha}\langle T(v), u\rangle. \end{aligned}$$

Em particular, a igualdade

$$\alpha\overline{\beta}\langle T(u), v\rangle + \beta\overline{\alpha}\langle T(v), u\rangle = 0$$

é valida para $\alpha = \beta = 1$ e para $\alpha = i$ e $\beta = 1$, ou seja,

$$\begin{cases} \langle T(u), v\rangle + \langle T(v), u\rangle = 0 \\ i\langle T(u), v\rangle - i\langle T(v), u\rangle = 0 \end{cases}.$$

Resolvendo o sistema acima concluímos que $\langle T(u), v\rangle = 0$ (verifique).

[(iii) $\Rightarrow$ (i)] Temos que $\langle T(u), v\rangle = 0$, para todos $u, v \in V$. Em particular, para $v = T(u)$,

$$\langle T(u), T(u)\rangle = 0 \quad \Rightarrow \quad T(u) = 0$$

para todo $u \in V$, ou seja, $T = 0$. ■

Observação 5.3.9 A Proposição 5.3.8 não é verdadeira se V for um $\mathbb{R}$-espaço vetorial. Veja o exemplo a seguir.

■ **Exemplo 5.3.10** Considere o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ definido por $T(x, y) = (-y, x)$. Claramente $T \neq 0$. Mas,

$$\langle T(x, y), (x, y)\rangle = \langle(-y, x), (x, y)\rangle = -yx + xy = 0 \quad \text{para todo } (x, y) \in \mathbb{R}^2.$$

■

Proposição 5.3.11 Seja V um $\mathbb{K}$-espaço vetorial com produto interno e $T : V \to V$ um operador autoadjunto. Se $\langle T(v), v\rangle = 0$ para todo $v \in V$, então $T = 0$.

Demonstração. O caso $\mathbb{K} = \mathbb{C}$ segue da Proposição 5.3.8 (mesmo que T não seja autoadjunto). Suponha $\mathbb{K} = \mathbb{R}$. Então, como $T = T^*$, para todos $u, v \in V$ temos

$$\langle T(u), v\rangle = \langle u, T(v)\rangle = \langle T(v), u\rangle.$$

Daí, não é difícil verificar que

$$\langle T(u+v), u+v\rangle - \langle T(u-v), u-v\rangle = 4\langle T(u), v\rangle.$$

Assim, para todos $u, v \in V$,

$$\langle T(u), v\rangle = \frac{1}{4}\langle T(u+v), u+v\rangle - \frac{1}{4}\langle T(u-v), u-v\rangle = 0 - 0 = 0.$$

Tomando $v = T(u)$, obtemos $\langle T(u), T(u)\rangle = 0$ donde segue que $T(u) = 0$ para todo $u \in V$. ■

Proposição 5.3.12 Seja V um $\mathbb{C}$-espaço vetorial com produto interno e $T : V \to V$ um operador linear que admite adjunto. Então, T é autoajunto se, e somente se, $\langle T(v), v\rangle \in \mathbb{R}$ para todo $v \in V$.

Demonstração. $(\Rightarrow)$ Seja $v \in V$. Como T é autoadjunto temos

$$\langle T(v), v\rangle = \langle v, T(v)\rangle = \overline{\langle T(v), v\rangle} \quad \Rightarrow \quad \langle T(v), v\rangle \in \mathbb{R}.$$

$(\Leftarrow)$ Note que,

$$\langle v, T^*(v)\rangle = \langle T(v), v\rangle = \langle v, T(v)\rangle,$$

ou seja, $\langle v, T^*(v) - T(v)\rangle = \langle v, (T^* - T)(v)\rangle = 0$ para todo $v \in V$. Segue da Proposição 5.3.8 que $T^* - T = 0$. Portanto, T é autoadjunto. ■

O resultado a seguir estabelece uma propriedades importantes sobre os autovalores e autovetores de operadores autoadjuntos.

Proposição 5.3.13 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $T : V \to V$ um operador linear autoadjunto.

(i) Se $\lambda \in \mathbb{K}$ é um autovalor de T, então $\lambda \in \mathbb{R}$.

(ii) Se $v_1, v_2 \in V$ são autovetores de T associados à autovalores distintos, então v_1 e v_2 são ortogonais.

Demonstração. **(i)** Seja $v \in V$ um autovetor de T associado à λ. Então, como T é autoadjunto,

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle \;\Rightarrow\; (\lambda - \overline{\lambda})\langle v, v\rangle = 0.$$

Como $v \neq 0$, segue que $\lambda = \overline{\lambda}$, ou seja, $\lambda \in \mathbb{R}$.

(ii) Sejam $v_1, v_2 \in V$ autovetores de T associados aos autovalores λ_1 e λ_2, respectivamente, com $\lambda_1 \neq \lambda_2$. Então, como T é autoadjunto,

$$\lambda_1\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \langle T(v_1), v_2\rangle = \langle v_1, T(v_2)\rangle = \langle v_1, \lambda_2 v_2\rangle = \lambda_2\langle v_1, v_2\rangle$$

ou seja,

$$(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0.$$

Como $\lambda_1 \neq \lambda_2$, segue que $\langle v_1, v_2\rangle = 0$.

■

5.4 Diagonalização de operadores autoadjuntos

Na Seção 4.5 estudamos as condições que garantem que um operador linear seja diagonalizável. Nesta seção, veremos que todos os operadores autoadjuntos são diagonalizáveis, e, mais ainda, que existe uma base ortonormal de autovetores. Como consequência, seguirá que todas as matrizes reais simétricas ou matrizes complexas hermitianas são diagonalizáveis, fato esse que é essencial em muitas aplicações.

Proposição 5.4.1 Seja V um $\mathbb{K}$-espaço vetorial com produto interno e dimensão finita $n \geq 1$. Se $T : V \to V$ é um operador linear autoadjunto, então, T possui autovetor.

Demonstração. Se $\mathbb{K} = \mathbb{C}$, então o polinômio característico $p_T(x)$ possui raiz (pois $\mathbb{C}$ é algebricamente fechado), ou seja, T possui autovalor, donde segue que existe autovetor.

Suponha $\mathbb{K} = \mathbb{R}$. Sejam $B = \{v_1, \ldots, v_n\}$ uma base ortonormal de V e $A = [T]_B \in M_n(\mathbb{R})$. Como $T = T^*$, a matriz A é simétrica. Considere o $\mathbb{C}$-espaço $\mathbb{C}^n$ com produto interno usual. Sejam C a base canônica de $\mathbb{C}^n$ e $S : \mathbb{C}^n \to \mathbb{C}^n$ o operador linear definido por $S(v) = Av$ para todo $v \in \mathbb{C}^n$ (visto como vetor coluna). Note que $[S]_C = A$. Como A é real e simétrica, $[S^*]_C = \overline{A}^\top = A$. Ou seja, S também é um operador autoadjunto. Mas S está definido em um espaço vetorial sobre $\mathbb{C}$, logo $p_S(x)$ possui raiz, que é autovalor de S. Decorre da Proposição 5.3.13 que os autovalores de S são todos reais. Note que,

$$p_S(x) = \det([S]_C - xI_n) = \det(A - xI_n) = p_T(x),$$

ou seja, as raízes do polinômio característico de T são todas reais. Logo, T possui autovalor, e portanto possui autovetor. ∎

Definição 5.4.2 Um operador linear $T : V \to V$ é dito *ortogonalmente diagonalizável* se V possui uma base ortonormal formada por autovetores de T.

O próximo resultado garante que todo operador autoadjunto é ortogonalmente diagonalizável.

Proposição 5.4.3 Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita $n \geq 1$ e $T : V \to V$ um operador linear. Se T é autoadjunto, então V possui uma base ortonormal formada por autovetores de T.

Demonstração. Pela Proposição 5.4.1 existe $v_1 \in V \setminus \{0\}$ autovetor de T. Faremos a prova por indução sobre n.

Se $n = 1$, então $\left\{ \frac{v_1}{||v_1||} \right\}$ é uma base ortonormal de V formada por autovetores de T.

Suponha $n > 1$ e o resultado válido para $n - 1$. Sejam $W = [v_1]$ e $\lambda_1 \in \mathbb{K}$ tal que $T(v_1) = \lambda_1 v_1$. Dado $u \in W^\perp$, como $T^* = T$, temos

$$\langle v_1, T(u) \rangle = \langle v_1, T^*(u) \rangle = \langle T(v_1), u \rangle = \langle \lambda_1 v_1, u \rangle = \lambda_1 \langle v_1, u \rangle = 0.$$

Portanto, $T(u) \in W^\perp$. Assim, a restrição $T|_{W^\perp} : W^\perp \to W^\perp$ é um operador linear em $W^\perp$. Como $V = W \oplus W^\perp$, segue que $\dim W^\perp = n - 1$, e pela hipótese de indução, existe base $\{v_2, \ldots, v_n\}$ ortonormal de $W^\perp$ formada por autovetores de T. Por fim, $B = \left\{ \frac{v_1}{||v_1||}, v_2, \ldots, v_n \right\}$ é uma base ortonormal de V formada por autovetores de T. ∎

O teorema a seguir determina que a condição de ser autoajunto é necessária e suficiente para que um operador linear sobre um espaço vetorial <u>real</u> seja ortogonalmente diagonalizável.

Teorema 5.4.4 — Espectral Real. Sejam V um $\mathbb{R}$-espaço vetorial de dimensão finita $n \geq 1$ e $T : V \to V$ um operador linear. Então, T é autoadjunto se, e somente se, V possui uma base ortonormal formada por autovetores de T.

Demonstração. ($\Rightarrow$) Segue da Proposição 5.4.3.

($\Leftarrow$) Seja $B = \{v_1, \dots, v_n\}$ uma base ortonormal de V formada por autovetores de T. Se λ_i é o autovalor de T associado ao autovetor v_i, então

$$[T]_B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Como $[T]_B$ é simétrica e B é ortonormal, segue que T é autoadjunto. ∎

■ **Exemplo 5.4.5** Considere $\mathbb{R}^4$ com o produto interno usual. Seja $T : \mathbb{R}^4 \to \mathbb{R}^4$ uma transformação linear tal que $\ker T = [(2,1,1,0),(1,-1,-1,0)]$, $(0,3,-3,0) \in \ker(T-I)$ e 2 é um autovalor de T associado ao autovetor $(0,0,0,2)$. Mostre que T é autoadjunto.

Prova. Como $\ker T = [(2,1,1,0),(1,-1,-1,0)]$ segue que $\lambda = 0$ é autovalor de T e $m_g(0) = 2$. Como $(0,3,-3,0) \in \ker(T-I)$, segue que $\lambda = 1$ é um autovalor de T e $(0,3,-3,0)$ é um autovetor associado. Além disso, $\lambda = 2$ é autovalor associado ao autovetor $(0,0,0,2)$. Note que $B = \{(2,1,1,0),(1,-1,-1,0),(0,3,-3,0),(0,0,0,2)\}$ é um conjunto ortogonal (e portanto LI) em $\mathbb{R}^4$. Logo, existe uma base ortonormal para $\mathbb{R}^4$ formada por autovetores de T (basta normalizar os vetores de B). Portanto, T é autoadjunto. ■

Corolário 5.4.6 — Decomposição Espectral Real. Dada $A \in M_n(\mathbb{R})$ uma matriz simétrica, existem matrizes $Q, \Lambda \in M_n(\mathbb{R})$ com Q ortogonal, formada pelos autovetores ortonormais de A, e Λ diagonal, formada pelos autovalores de A, tais que $A = Q\Lambda Q^\top$.

Demonstração. Sejam C a base canônica de $\mathbb{R}^n$ e $T : \mathbb{R}^n \to \mathbb{R}^n$ o operador linear tal que $[T]_C = A$. Como C é ortonormal e A é simétrica, segue que T é autoadjunto. Logo, existe $B = \{v_1, \dots, v_n\}$ base ortonormal de $\mathbb{R}^n$ formada por autovetores de T. Considere $\lambda_i \in \mathbb{R}$, $i = 1, \dots, n$, o autovetor de T associado a v_i. Defina

$$Q = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} \in M_n(\mathbb{R}) \text{ e } \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \in M_n(\mathbb{R}).$$

Pela Proposição 3.2.20, a matriz Q é ortogonal, isto é, $Q^\top = Q^{-1}$. Agora,

$$AQ = \begin{bmatrix} | & | & & | \\ Av_1 & Av_2 & \cdots & Av_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = Q\Lambda,$$

ou seja, $A = Q\Lambda Q^\top$. ∎

Observação 5.4.7 Se $A = Q\Lambda Q^\top$, então

$$A = Q\Lambda Q^\top = \lambda_1 v_1 v_1^\top + \lambda_2 v_2 v_2^\top + \cdots + \lambda_n v_n v_n^\top.$$

Ou seja, podemos decompor A em termos de seu espectro (conjunto de autovalores).

■ **Exemplo 5.4.8** Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x,y) = (3x+y, x+3y)$. Verifique se existe uma base de $\mathbb{R}^2$ formada por autovetores ortonormais de T. Se existir, encontre a decomposição espectral da matriz de T em relação à base canônica.

Solução. Seja $B = \{(1,0),(0,1)\}$ a base canônica de $\mathbb{R}^2$. Como $T(1,0) = (3,1)$ e $T(0,1) = (1,3)$, segue que

$$[T]_B = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Como B é ortonormal e $[T]_B$ é simétrica, T é autoadjunto. Logo, existe uma base ortonormal de $\mathbb{R}^2$ formada por autovetores de T. Vamos calcular os autovalores e autovetores de T. Temos

$$p_T(x) = \det \begin{bmatrix} 3-x & 1 \\ 1 & 3-x \end{bmatrix} = (3-x)^2 - 1 = x^2 - 6x + 8 = (x-2)(x-4)$$

ou seja, $\lambda_1 = 2$ e $\lambda_2 = 4$ são os autovalores de T.

Seja $v_1 = (a,b)$ um autovetor de T associado à $\lambda_1 = 2$. Então,

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \quad \Leftrightarrow \quad v_1 \in [(1,-1)],$$

ou seja, $\mathrm{Aut}_T(2) = [(1,-1)]$.

Seja $v_2 = (a,b)$ autovetor de T associado à $\lambda_2 = 4$, então

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \quad \Leftrightarrow \quad v_2 \in [(1,1)]$$

ou seja, $\mathrm{Aut}_T(4) = [(1,1)]$.

Como esperado, v_1 e v_2 são ortogonais. Normalizando v_1 e v_2, obtemos $C = \left\{ \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) \right\}$ base ortonormal de $\mathbb{R}^2$ formada por autovetores de T. Além disso,

$$[T]_B = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

é a decomposição espectral de $[T]_B$. ■

5.5 Operadores normais e diagonalização

Dando continuidade à identificação de operadores diagonalizáveis, nessa seção veremos que todos os operadores normais são diagonalizáveis em espaços complexos.

Definição 5.5.1 Um operador linear $T : V \to V$ num $\mathbb{K}$-espaço V com produto interno é dito *normal* se existir o operador adjunto T^* e, além disso, $T \circ T^* = T^* \circ T$.
Uma matriz $A \in M_n(\mathbb{K})$ é chamada de *normal* se $A\overline{A}^\top = \overline{A}^\top A$.

Observação 5.5.2 **(i)** Todo operador autoadjunto (e toda matriz hermitiana) é normal.

(ii) Se V tem dimensão finita, segue da Proposição 5.2.9 que $T : V \to V$ é um operador normal se, e somente se, dada uma base B ortonormal de V, a matriz $[T]_B$ é normal.

■ **Exemplo 5.5.3** Seja C a base canônica de $\mathbb{C}^2$. Verifique se o operador linear $T : \mathbb{C}^2 \to \mathbb{C}^2$ tal que $[T]_C = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$ é normal.

Solução. Note que

$$[T \circ T^*]_C = [T]_C[T^*]_C = [T]_C\overline{[T]}_C^\top = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

e

$$[T^* \circ T]_C = [T^*]_C[T]_C = \overline{[T]}_C^\top[T]_C = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}.$$

Logo, T é normal. ■

O resultado a seguir justifica o porquê do nome “normal” e fornece mais um critério para a verificação dessa condição.

Proposição 5.5.4 Seja $T : V \to V$ um operador linear em um espaço com produto interno. Então, T é um operador normal se, e somente se, $||T(v)|| = ||T^*(v)||$ para todo $v \in V$.

Demonstração. $(\Rightarrow)$ Suponha que T é normal. Dado $v \in V$, temos

$$||T(v)||^2 = \langle T(v), T(v)\rangle = \langle v, T^* \circ T(v)\rangle = \langle v, T \circ T^*(v)\rangle = \overline{\langle T \circ T^*(v), v\rangle}.$$

Como $\langle T(v), T(v)\rangle \in \mathbb{R}$ (mesmo que $\mathbb{K} = \mathbb{C}$), segue que

$$\overline{\langle T \circ T^*(v), v\rangle} = \langle T \circ T^*(v), v\rangle = \langle T^*(v), T^*(v)\rangle = ||T^*(v)||^2.$$

Portanto, $||T(v)|| = ||T^*(v)||$.

$(\Leftarrow)$ Dado $v \in V$, temos

$$\begin{aligned} ||T(v)|| = ||T^*(v)|| \quad &\Rightarrow \quad \langle T(v), T(v)\rangle = \langle T^*(v), T^*(v)\rangle \\ &\Rightarrow \quad \langle v, T^* \circ T(v)\rangle = \langle v, T \circ T^*(v)\rangle \\ &\Rightarrow \quad \langle v, (T^* \circ T - T \circ T^*)(v)\rangle = 0, \ \ \forall v \in V. \end{aligned}$$

Note que o operador $T^* \circ T - T \circ T^*$ é autoajunto, logo, pela Proposição 5.3.11, $T^* \circ T = T \circ T^*$. Portanto, T é um operador normal. ■

O próximo resultado apresenta algumas propriedades dos autovalores e autovetores de um operador normal.

Proposição 5.5.5 Sejam V um $\mathbb{K}$-espaço vetorial com produto interno e $T : V \to V$ um operador normal.

(i) Se $T(v) = \lambda v$, com $\lambda \in \mathbb{K}$ e $v \neq 0$, então $T^*(v) = \overline{\lambda} v$.

(ii) Se $T(v_1) = \lambda_1 v_1$ e $T(v_2) = \lambda_2 v_2$, com $v_1, v_2 \neq 0$ e $\lambda_1 \neq \lambda_2$ $(\lambda_1, \lambda_2 \in \mathbb{K})$, então $v_1 \perp v_2$.

Demonstração. **(i)** Mostremos que $T-\lambda\,\mathrm{Id}$ também é normal. Temos,

$$(T-\lambda\,\mathrm{Id})\circ(T-\lambda\,\mathrm{Id})^* = (T-\lambda\,\mathrm{Id})\circ(T^*-\overline{\lambda}\,\mathrm{Id}) = T\circ T^* - \overline{\lambda}T - \lambda T^* + \lambda\overline{\lambda}\,\mathrm{Id},$$

e, por outro lado,

$$(T-\lambda\,\mathrm{Id})^*\circ(T-\lambda\,\mathrm{Id}) = T^*\circ T - \lambda T^* - \overline{\lambda}T + \lambda\overline{\lambda}\,\mathrm{Id}.$$

Como T é normal e $\lambda\overline{\lambda}=\overline{\lambda}\lambda$, segue que

$$(T-\lambda\,\mathrm{Id})\circ(T-\lambda\,\mathrm{Id})^* = (T-\lambda\,\mathrm{Id})^*\circ(T-\lambda\,\mathrm{Id}).$$

Assim, pela Proposição 5.5.5,

$$0 = ||(T-\lambda\,\mathrm{Id})(v)|| = ||(T-\lambda\,\mathrm{Id})^*(v)|| = ||(T^*-\overline{\lambda}\,\mathrm{Id})(v)||$$

ou seja, $(T^*-\overline{\lambda}\,\mathrm{Id})(v)=0$, logo $T^*(v)=\overline{\lambda}v$.

(ii) Sejam $\lambda_1\neq\lambda_2$ autovalores de T associados aos autovetores v_1 e v_2, respectivamente. Então,

$$\lambda_1\langle v_1,v_2\rangle = \langle\lambda_1 v_1,v_2\rangle = \langle T(v_1),v_2\rangle = \langle v_1,T^*(v_2)\rangle = \langle v_1,\overline{\lambda_2}v_2\rangle = \lambda_2\langle v_1,v_2\rangle$$

isto é, $(\lambda_1-\lambda_2)\langle v_1,v_2\rangle=0$. Como $\lambda_1\neq\lambda_2$, segue que $\langle v_1,v_2\rangle=0$. ■

Teorema 5.5.6 — Espectral Complexo. Sejam V um $\mathbb{C}$-espaço vetorial de dimensão finita $n\geq 1$ e $T:V\to V$ um operador linear. Então, T é normal se, e somente se, V possui uma base ortonormal formada por autovetores de T.

Demonstração. ($\Rightarrow$) O polinômio característico $p_T(x)$ possui raiz (pois $\mathbb{C}$ é algebricamente fechado), ou seja, T possui autovalor, donde segue que existe $v_1\in V\setminus\{0\}$ autovetor de T. Faremos a prova por indução sobre n.

Se $n=1$, então $\left\{\frac{v_1}{||v_1||}\right\}$ é uma base ortonormal de V formada por autovetores de T.

Suponha $n>1$ e o resultado válido para $n-1$. Sejam $W=[v_1]$ e $\lambda_1\in\mathbb{K}$ tal que $T(v_1)=\lambda_1 v_1$. Como T é normal, pela Proposição 5.5.5, $T^*(v_1)=\overline{\lambda_1}v_1$. Assim, dado $u\in W^\perp$, temos

$$\langle T(u),v_1\rangle = \langle u,T^*(v_1)\rangle = \langle u,\overline{\lambda_1}v_1\rangle = \lambda_1\langle u,v_1\rangle = 0.$$

Portanto, $T(u)\in W^\perp$. Logo, a restrição $T|_{W^\perp}:W^\perp\to W^\perp$ é um operador linear em $W^\perp$. Como $V=W\oplus W^\perp$, segue que $\dim W^\perp = n-1$, e pela hipótese de indução, existe base $\{v_2,\ldots,v_n\}$ ortonormal de $W^\perp$ formada por autovetores de T. Por fim, $B=\left\{\frac{v_1}{||v_1||},v_2,\ldots,v_n\right\}$ é uma base ortonormal de V formada por autovetores de T.

($\Leftarrow$) Seja $B=\{v_1,\ldots,v_n\}$ uma base ortonormal de V formada por autovetores de T. Se λ_i é o autovalor de T associado ao autovetor v_i, então

$$[T]_B = \begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix} \Rightarrow [T^*]_B = \overline{[T]}_B^\top = \begin{bmatrix}\overline{\lambda_1} & 0 & \cdots & 0\\ 0 & \overline{\lambda_2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \overline{\lambda_n}\end{bmatrix}.$$

Logo,

$$[T \circ T^*]_B = [T]_B\,[T^*]_B = \begin{bmatrix} \lambda_1\overline{\lambda_1} & 0 & \cdots & 0 \\ 0 & \lambda_2\overline{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\overline{\lambda_n} \end{bmatrix} = [T^*]_B\,[T]_B = [T^* \circ T]_B.$$

Portanto, $T \circ T^* = T^* \circ T$, ou seja, T é normal. ■

Corolário 5.5.7 — Decomposição Espectral Complexa. Se $N \in M_n(\mathbb{C})$ é uma matriz normal, então existem matrizes $U, \Lambda \in M_n(\mathbb{C})$, com U unitária, formada pelos autovetores de N, e Λ diagonal, formada pelos autovalores de N, tais que $N = U\Lambda\overline{U}^\top$.

■ **Exemplo 5.5.8** Determine a decomposição espectral da matriz $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$.

Solução. Note que a matriz A é normal, pois

$$AA^\top = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 2 \\ 2 & 5 & 2 \\ 2 & 2 & 5 \end{bmatrix}$$

e

$$A^\top A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 2 \\ 2 & 5 & 2 \\ 2 & 2 & 5 \end{bmatrix}.$$

Logo, existe uma base de $\mathbb{C}^3$ formada por autovetores ortonormais de A.

(i) Autovalores de A:

$$p_A(x) = \det \begin{bmatrix} 1-x & 2 & 0 \\ 0 & 1-x & 2 \\ 2 & 0 & 1-x \end{bmatrix} = (1-x)^3 + 8 = (x-3)(-x^2-3) = -(x-3)(x-\sqrt{3}i)(x+\sqrt{3}i).$$

Ou seja, $\lambda_1 = 3$, $\lambda_2 = \sqrt{3}i$ e $\lambda_3 = 0\sqrt{3}i$ são autovalores de A.

(ii) Autovetores de A:

$\lambda_1 = 3$: se $v = (a,b,c) \in \mathbb{C}^3$ é autovetor de A associado à $\lambda_1 = 3$, então

$$\begin{bmatrix} -2 & 2 & 0 \\ 0 & -2 & 2 \\ 2 & 0 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \;\Rightarrow\; a = b = c,$$

ou seja, $v = a(1,1,1)$. Logo, $\mathrm{Aut}_A(3) = [(1,1,1)]$.

$\lambda_2 = \sqrt{3}i$: se $v = (a,b,c) \in \mathbb{C}^3$ é autovetor de A associado à $\lambda_2 = \sqrt{3}i$, então

$$\begin{bmatrix} 1-\sqrt{3}i & 2 & 0 \\ 0 & 1-\sqrt{3}i & 2 \\ 2 & 0 & 1-\sqrt{3}i \end{bmatrix} = 0 \;\Rightarrow\; \begin{cases} b = \frac{1}{2}(-1+\sqrt{3}i)a \\ c = \frac{1}{4}(-1+\sqrt{3}i)^2 a = \frac{1}{2}(-1-\sqrt{3}i)a \end{cases},$$

ou seja, $v = \dfrac{a}{2}(-2, 1-\sqrt{3}i, 1+\sqrt{3}i)$. Logo, $\mathrm{Aut}_A(\sqrt{3}i) = [(-2, 1-\sqrt{3}i, 1+\sqrt{3}i)]$.

$\lambda_3 = -\sqrt{3}i$: se $v = (a,b,c) \in \mathbb{C}^3$ é autovetor de A associado à $\lambda_3 = -\sqrt{3}i$, então

$$\begin{bmatrix} 1+\sqrt{3}i & 2 & 0 \\ 0 & 1+\sqrt{3}i & 2 \\ 2 & 0 & 1+\sqrt{3}i \end{bmatrix} = 0 \ \Rightarrow \ \begin{cases} b = -\frac{1}{2}(1+\sqrt{3}i)a \\ c = -\frac{1}{4}(1+\sqrt{3}i)^2 a = -\frac{1}{2}(1-\sqrt{3}i)a \end{cases},$$

ou seja, $v = -\dfrac{a}{2}(-2, 1+\sqrt{3}i, 1-\sqrt{3}i)$. Logo, $\mathrm{Aut}_A(-\sqrt{3}i) = [(-2, 1+\sqrt{3}i, 1-\sqrt{3}i)]$.

Como esperado, os autovetores associados a autovalores distintos são ortogonais. Para montar a matriz U basta normalizá-los. Assim,

$$A = U\Lambda\overline{U}^{\top} = \frac{1}{12}\begin{bmatrix} 2 & -2 & -2 \\ 2 & 1-\sqrt{3}i & 1+\sqrt{3}i \\ 2 & 1+\sqrt{3}i & 1-\sqrt{3}i \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 \\ 0 & \sqrt{3}i & 0 \\ 0 & 0 & -\sqrt{3}i \end{bmatrix}\begin{bmatrix} 2 & 2 & 2 \\ -2 & 1+\sqrt{3}i & 1-\sqrt{3}i \\ -2 & 1-\sqrt{3}i & 1+\sqrt{3}i \end{bmatrix}.$$

$\blacksquare$

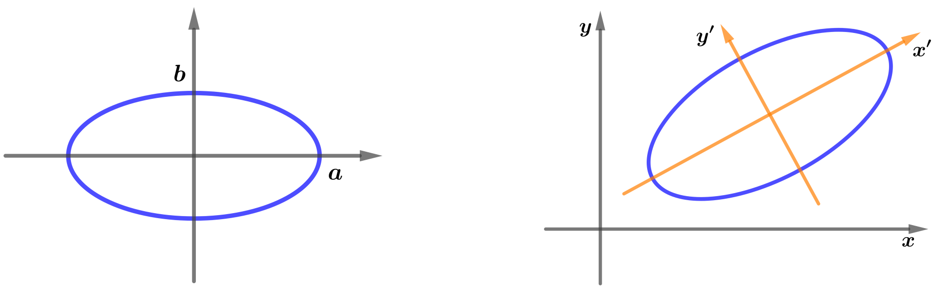
Capítulo 6

Formas quadráticas e aplicações

Em Geometria Analítica são estudadas as cônicas, que são curvas planas, descritas por equações polinomiais de grau 2 em 2 variáveis. Por exemplo, dados $a, b \in \mathbb{R} \setminus \{0\}$, a equação

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

descreve uma elipse em $\mathbb{R}^2$. Esse tipo de equação, onde não aparecem termos "mistos" xy, é simples de reconhecer e esboçar. No entanto, no caso geral, essa tarefa não é tão simples, e faz-se necessário uma "mudança de coordenadas" adequada. Veja a Figura 6.1.



(a) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (b) $\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1$

Figura 6.1: Elipses em $\mathbb{R}^2$

A elipse na Figura 6.1(b) teria uma equação bem mais complicada nas variáveis x e y, onde não seria possível visualizar rapidamente qual seria a cônica representada.

Neste capítulo, veremos como utilizar as ferramentas de Álgebra Linear vistas anteriormente para simplificar equações e facilitar o reconhecimento de cônicas ($\mathbb{R}^2$) e quádricas ($\mathbb{R}^3$). Faremos mais detalhadamente do caso das quádricas, mas as mesmas ideias se aplicarão no caso de cônicas.

A discussão apresentada nesse capítulo foi motivada pelos livros [Coe], [Lar] e [Lay].

6.1 Formas quadráticas

Uma *função quadrática* em $\mathbb{R}^2$ é uma função cuja expressão se trata de um polinômio de grau 2 (quadrática) em 2 variáveis (por ser definida em $\mathbb{R}^2$), ou seja, é uma função $p : \mathbb{R}^2 \to \mathbb{R}$ dada por

$p(x_1,x_2) = ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f$, com $a,b,c,d,e,f \in \mathbb{R}$. Ou ainda, na forma matricial, podemos escrever

$$p(x_1,x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + f.$$

A parte com os termos de grau 2 de $p(x_1,x_2)$:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + bx_2^2 + cx_1x_2,$$

é a parte quadrática de p, e define o que chamamos de forma quadrática, neste caso em $\mathbb{R}^2$. Essa definição será generalizada a seguir.

Definição 6.1.1 Uma *forma quadrática* em $\mathbb{R}^n$ é uma função $q : \mathbb{R}^n \to \mathbb{R}$ definida por $q(x) = x^\top Ax$, em que $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ e $A \in M_n(\mathbb{R})$ é uma matriz simétrica. A matriz A que define a forma quadrática é chamada de *matriz da forma quadrática*.

■ **Exemplo 6.1.2** Calcule a expressão da forma quadrática definida a partir da matriz $A = \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix}$.

Solução. Seja $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, então

$$q(x_1,x_2) = x^\top Ax = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 4x_1x_2 + 3x_2^2.$$

■

■ **Exemplo 6.1.3** Encontre a matriz da forma quadrática $q(x_1,x_2,x_3) = 3x_1^2 + 2x_2^2 + x_3^2 - x_1x_2 + 6x_2x_3$.

Solução. Note que,

$$q(x_1,x_2,x_3) = 3x_1^2 + 2x_2^2 + x_3^2 - x_1x_2 + 6x_2x_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 3 & -1/2 & 0 \\ -1/2 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Logo, a matriz da forma quadrática é $A = \begin{bmatrix} 3 & -1/2 & 0 \\ -1/2 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix}$. ■

Observação 6.1.4
- **(i)** Se a matriz A é diagonal, a forma quadrática $q(x) = x^\top Ax$ não contém termos mistos (x_ix_j com $i \neq j$).
- **(ii)** Dada uma função polinomial $p(x_1,\ldots,x_n)$ nas variáveis $x_1,\ldots,x_n$, que só contem termos de grau 2, existe única matriz $A \in M_n(\mathbb{R})$ simétrica tal que $p(x_1,\ldots,x_n) = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.
- **(iii)** É possível definir forma quadrática $q(x) = x^\top Ax$ para uma matriz $A \in M_n(\mathbb{R})$ qualquer. Porém, neste caso, a matriz A da representação de $q(x) = x^\top Ax$ não é única. Consideraremos apenas

matrizes simétricas, em razão também da Decomposição Espectral, que terá papel importante.

Mudança de variável

Se $x \in \mathbb{R}^n$ representa um vetor de variáveis em $\mathbb{R}^n$, uma *mudança de variável* ou *mudança de coordenadas* é uma mudança da forma $x = Qy$ (ou $y = Q^{-1}x$) em que Q é uma matriz não singular. Nesse caso, $y \in \mathbb{R}^n$ são as coordenadas do vetor x em relação à base de $\mathbb{R}^n$ definida pelas colunas de Q.

Substituindo a mudança na forma quadrática:

$$q(x) = x^\top Ax = (Qy)^\top A(Qy) = y^\top(Q^\top AQ)y = \tilde{q}(y)$$

e a matriz da "nova" forma quadrática, isto é, da forma quadrática na nova variável y, é $\tilde{A} = Q^\top AQ$.

Como a matriz A que define a forma quadrática é simétrica, segue da Decomposição Espectral que existe uma matriz Q ortogonal tal que $A = Q\Lambda Q^\top$, ou seja, $Q^\top AQ = \Lambda$. Assim, ao fazer a mudança $x = Qy$, é possível eliminar os termos mistos da equação da forma quadrática, pois $\tilde{q}(y) = y^\top \Lambda y$ com Λ diagonal.

Definição 6.1.5 Uma *forma quadrática diagonal* em $\mathbb{R}^n$ é uma função $q : \mathbb{R}^n \to \mathbb{R}$ definida por $q(x) = x^\top Ax$, em que $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ e $A \in M_n(\mathbb{R})$ é uma matriz diagonal.

■ **Exemplo 6.1.6** Determine uma mudança de variável por uma matriz ortogonal que reduz a forma quadrática $q(x_1,x_2) = 2x_1^2 + 4x_1x_2 + 5x_2^2$ em uma forma diagonal.

Solução. Temos que determinar uma matriz ortogonal Q para a mudança de variável $x = Qy$ que transforme a matriz A da forma quadrática q em uma matriz diagonal. Como A é uma matriz simétrica, basta encontrar os autovalores e autovetores ortonormais de A.

Note que $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$, seus autovalores são $\lambda_1 = 1$ e $\lambda_2 = 6$ e $\mathrm{Aut}_A(1) = [(2,-1)]$ e $\mathrm{Aut}_A(6) = [(1,2)]$ (verifique). Os autovetores associados aos dois autovalores distintos já são automaticamente ortogonais (pois A é simétrica), então basta normalizar os autovetores para obter a matriz Q da Decomposição Espectral de A: temos $A = Q\Lambda Q^\top$, em que

$$Q = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}.$$

Com a mudança de variável $y = Q^\top x$, obtemos $\tilde{q}(y_1,y_2) = y_1^2 + 6y_2^2$. ■

6.2 Quádricas

Definição 6.2.1 Uma *quádrica* em $\mathbb{R}^3$ é uma superfície formada pelos $(x,y,z) \in \mathbb{R}^3$ que satisfazem uma equação da forma

$$ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz + rx + sy + tz + k = 0$$

com $a,b,c,d,e,f,r,s,t,k \in \mathbb{R}$ e $a^2 + b^2 + c^2 \neq 0$.

A seguir, apresentamos as possíveis superfícies quádricas e suas equações na forma reduzida. Iniciamos com os casos chamados de *não degenerados*.

Elipsoide

Quádrica de equação

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

com $a, b, c \in \mathbb{R} \setminus \{0\}$. Se $a = b = c$, o elipsóide é uma esfera.

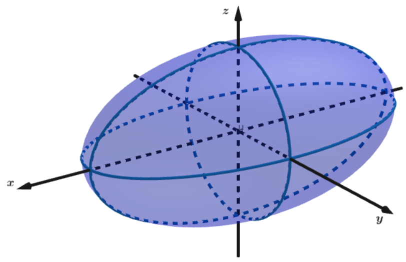


Figura 6.2: Elipsoide

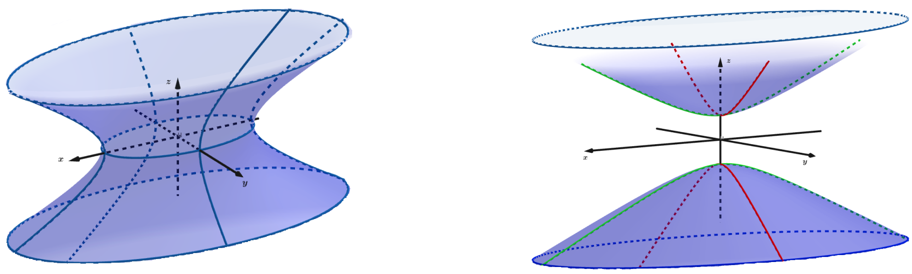
As interseções dessa superfície com os planos $x = k$, $y = k$ ou $z = k$ são elipses se $|k| < a$, $|k| < b$ ou $|k| < c$, respectivamente.

Hiperboloide

Quádrica de equação

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = d,$$

com $a, b, c, d \in \mathbb{R} \setminus \{0\}$. O caso $d > 0$ representa um *hiperboloide de uma folha* e o caso $d < 0$ um *hiperboloide de duas folhas*. Veja a Figura 6.3.



(a) Hiperboloide de uma folha. (b) Hiperboloide de duas folhas.

Figura 6.3: Hiperboloide

Se $d > 0$, as interseções desse hiperboloide de uma folha com os planos $x = k$ ou $y = k$ são hipérboles e com os planos $z = k$ são elipses.

Se $d < 0$, as interseções desse hiperboloide de duas folhas com os planos $x = k$ ou $y = k$ são hipérboles, e com os planos $z = k$ são elipses se $k > \sqrt{|d|c^2}$ ou $k < -\sqrt{|d|c^2}$ ou um conjunto vazio se $-\sqrt{|d|c^2} < k < \sqrt{|d|c^2}$.

Cone

Quádrica de equação

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0,$$

com $a,b,c \in \mathbb{R} \setminus \{0\}$.

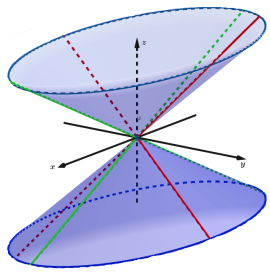


Figura 6.4: Cone

As interseções dessa superfície com os planos $x = k$ ou $y = k$ são hipérboles se $k \neq 0$ ou um par de retas concorrentes se $k = 0$. As interseções com os planos $z = k$ são elipses se $k \neq 0$ ou apenas um ponto se $k = 0$.

Paraboloide

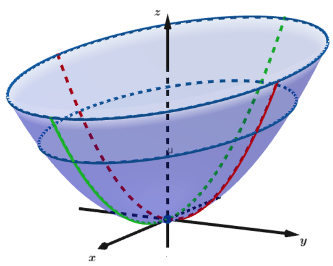
Sejam com $a,b,c \in \mathbb{R} \setminus \{0\}$. A equação

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z}{c}$$

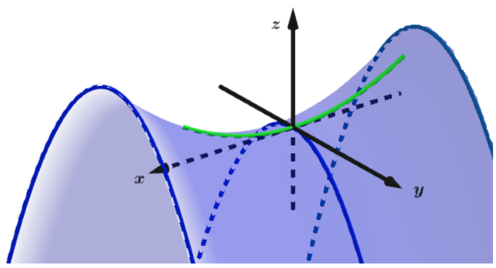
representa um *paraboloide elíptico*, e a equação

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c}$$

representa um *paraboloide hiperbólico* (ou *sela*). Veja a Figura 6.5.



(a) Paraboloide elíptico



(b) Paraboloide hiperbólico

Figura 6.5: Paraboloides

As interseções do paraboloide elíptico com os planos $x = k$ ou $y = k$ são parábolas e com os planos $z = k$ são elipses se $kc > 0$ ou um conjunto vazio se $kc < 0$.

As interseções do paraboloide hiperbólico com os planos $x = k$ ou $y = k$ são parábolas e as interseções com os planos $z = k$ são hipérboles se $k \neq 0$ ou um par de retas concorrentes se $k = 0$.

Os próximos casos são chamados de *degenerados*.

Cilindros

Sejam com $a,b,c \in \mathbb{R} \setminus \{0\}$. A quádrica de equação

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = c$$

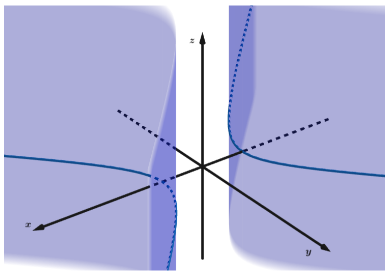
representa um *cilindro hiperbólico*, a equação

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=c, \quad c>0$$

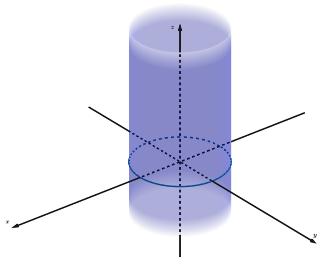
um *cilindro elíptico* e a equação

$$ax^2=by$$

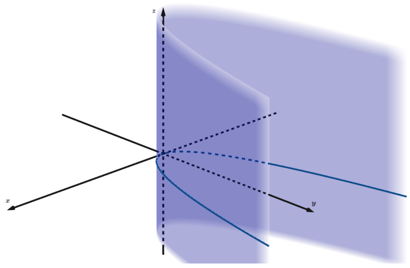
um *cilindro parabólico*. Veja a Figura 6.6.



(a) Cilindro hiperbólico



(b) Cilindro elíptico



(c) Cilindro parabólico

Figura 6.6: Cilindros

Outros casos degenerados

Também podem ocorrer como quádricas: um par de planos (concorrentes ou paralelos), um plano, uma reta, um ponto ou um conjunto vazio. Por exemplo:

$$x^2+y^2+z^2=0 \quad \text{(um ponto)}; \qquad x^2+y^2=0 \quad \text{(uma reta)}; \qquad x^2=0 \quad \text{(um plano)};$$

$$x^2-y^2=0 \quad \text{(um par de planos concorrentes)}; \qquad x^2=1 \quad \text{(um par de planos paralelos)};$$

$$x^2+y^2+z^2=-1 \quad \text{(conjunto vazio)}.$$

6.3 Reconhecimento de quádricas

Apresentaremos agora as ideias que serão utilizadas para reconhecer uma quádrica dada na sua forma geral

$$ax^2+by^2+cz^2+2dxy+2exz+2fyz+rx+sy+tz+k=0, \tag{6.1}$$

com $a,b,c,d,e,f,r,s,t,k\in\mathbb{R}$ e $a^2+b^2+c^2\neq 0$.

Primeiro, consideramos a forma quadrática em $\mathbb{R}^3$ definida pelos termos de grau 2 em (6.1):

$$q(x,y,z)=ax^2+by^2+cz^2+2dxy+2exz+2fyz=\begin{bmatrix}x & y & z\end{bmatrix}\begin{bmatrix}a & d & e\\ d & b & f\\ e & f & c\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}. \tag{6.2}$$

Como $A = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \in M_3(\mathbb{R})$ é simétrica, existem $Q, \Lambda \in M_3(\mathbb{R})$, Q ortogonal, Λ diagonal tais que $\Lambda = Q^\top A Q$. Assim, se $v = (x,y,z)^\top \in \mathbb{R}^3$ e $v' = (x',y',z')^\top = Q^\top v \Leftrightarrow v = Qv'$, temos

$$q(v) = v^\top A v = (Qv')^\top A(Qv') = v'^\top (Q^\top A Q)v' = v'^\top \Lambda v' = \tilde{q}(v'), \tag{6.3}$$

onde $\tilde{q}(v')$ é uma forma diagonal, ou seja, não contém termos mistos.

Note que, se $C = \{e_1, e_2, e_3\}$ é a base canônica de $\mathbb{R}^3$ e $B = \{v_1, v_2, v_3\}$ é uma base ortonomal de autovetores de A, com $Av_j = \lambda_j v_j$, então tomando Q a matriz cuja j-ésima coluna é v_j, temos

$$v = (x,y,z) = [v]_C, \quad v' = (x',y',z') = [v]_B, \quad Q = M_{BC}, \quad \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3).$$

Ou seja, em (6.3) estamos reescrevendo (6.2), dada em termos das coordenadas em relação à base canônica, para

$$\tilde{q}(x',y',z') = \lambda_1 (x')^2 + \lambda_2 (y')^2 + \lambda_3 (z')^2$$

em termos das coordenadas em relação à base ortonormal B. Geometricamente, ambas as equações $q(x,y,z) = 0$ e $\tilde{q}(x',y',z') = 0$ representam o mesmo lugar geométrico, porém a primeira está descrita no sistema de coordenadas (O, e_1, e_2, e_3) e a segunda no sistema de coordenadas (O, v_1, v_2, v_3).

Após realizar a mudança de coordenadas $v = Qv'$ em (6.2) para eliminar os termos mistos, precisamos fazer a mesma mudança nos termos de grau 1 (parte linear) de (6.1) para obter a nova equação para a quádrica dada toda em termos de x', y', z':

$$\lambda_1 (x')^2 + \lambda_2 (y')^2 + \lambda_3 (z')^2 + r'x' + s'y' + t'z' + k = 0. \tag{6.4}$$

Por fim, completamos os quadrados (6.4), a fim de obter uma equação na forma reduzida, como as descritas na Seção 6.2. Daí, faremos uma nova mudança da forma

$$x'' = x' - \alpha, \qquad y'' = y' - \beta, \qquad z'' = z' - \gamma,$$

que geometricamente corresponde a uma translação da origem de $O = (0,0,0)$ para $O' = (\alpha, \beta, \gamma)$ no sistema de coordenadas.

■ **Exemplo 6.3.1** Classifique a quádrica de equação

$$5x^2 + 8y^2 + 5z^2 - 4xy - 8xz - 4yz + \frac{2}{3}x + \frac{1}{3}y + \frac{2}{3}z + \frac{1}{3} = 0.$$

Descreva quais foram as mudanças de coordenadas empregadas para simplificar sua equação e faça um esboço da quádrica no novo sistema de coordenadas.

Solução. Iniciaremos com a parte quadrática da equação, a fim de eliminar seus termos mistos. Como descrito anteriormente, faremos a mudança

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = Q \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Qv'$$

com Q formada pelos autovetores ortonormais da matriz da forma quadrática associada.

(i) Forma quadrática associada:

$$q(x,y,z) = 5x^2 + 8y^2 + 5z^2 - 4xy - 8xz - 4yz = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 5 & -2 & -4 \\ -2 & 8 & -2 \\ -4 & -2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

(ii) Autovalores da matriz da forma quadrática:

$$\det\begin{bmatrix}5-x & -2 & -4\\ -2 & 8-x & -2\\ -4 & -2 & 5-x\end{bmatrix} = (5-x)^2(8-x)-16-16-16(8-x)-4(5-x)-4(5-x)$$

$$= -x^3+18x^2-81x = -x(x^2-18\lambda+81) = -x(x-9)^2.$$

Ou seja, $\lambda_1 = 9$, $\lambda_2 = 9$ e $\lambda_3 = 0$.

(iii) Autovetores ortonormais da matriz da forma quadrática

- $\lambda_1 = 9$: se $u_1 = (a,b,c)$ é autovetor associado à $\lambda_1 = 9$, então

$$\begin{bmatrix}-4 & -2 & -4\\ -2 & -1 & -2\\ -4 & -2 & -4\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix} = 0 \ \Rightarrow\ b = -2a-2c$$

isto é, $u_1 = (a,-2a-2b,b)$. Assim, podemos escolher $u_1 = (1,0,-1)$ e $u_2 = (1,-2,0)$, por exemplo. Note que u_1 e u_2 não são ortogonais. Utilizando o processo de ortogonalização de Gram-Schmidt obtemos:

$\langle u_1,u_2\rangle = 1$

$w_2 = u_2 - \langle u_1,u_2\rangle \dfrac{u_1}{||u_1||^2} = (1,-2,0)-\frac{1}{2}(1,0,-1) = \left(\frac{1}{2},-2,\frac{1}{2}\right).$

- $\lambda_3 = 0$: se $u = (a,b,c)$ é autovetor associado à $\lambda = 0$, então

$$\begin{bmatrix}5 & -2 & -4\\ -2 & 8 & -2\\ -4 & -2 & 5\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix} = 0 \ \Rightarrow\ \begin{bmatrix}5 & -2 & -4\\ 0 & 2 & -1\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix} = 0 \ \Rightarrow\ a = c = 2b$$

assim, $u_3 = (2,1,2)$, por exemplo.

Logo,

$$u_1 = \frac{1}{\sqrt{2}}(1,0,-1)\ ,\ u_2 = \frac{\sqrt{2}}{6}(1,-4,1)\ ,\ u_3 = \frac{1}{3}(2,1,2)\ \text{ e }\ Q = \begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/6 & 2/3\\ 0 & -2\sqrt{2}/3 & 1/3\\ -\sqrt{2}/2 & \sqrt{2}/6 & 2/3\end{bmatrix}.$$

A mudança de variável $v = Qv'$ transforma a forma quadrática $q(x,y,z)$ na forma diagonal $9(x')^2+9(y')^2$. Agora, precisamos efetuar a mudança na parte linear. Temos,

$$\begin{bmatrix}x\\ y\\ z\end{bmatrix} = \begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/6 & 2/3\\ 0 & -2\sqrt{2}/3 & 1/3\\ -\sqrt{2}/2 & \sqrt{2}/6 & 2/3\end{bmatrix}\begin{bmatrix}x'\\ y'\\ z'\end{bmatrix}$$

logo,

$$\frac{2}{3}x+\frac{2}{3}y+2z = \frac{2}{3}\left(\frac{\sqrt{2}}{2}x'+\frac{\sqrt{2}}{6}y'+\frac{2}{3}z'\right)+\frac{2}{3}\left(-\frac{2\sqrt{2}}{3}y'+\frac{1}{3}z'\right)+2\left(-\frac{\sqrt{2}}{2}x'+\frac{\sqrt{2}}{6}y'+\frac{2}{3}z'\right)$$

$$= -\frac{2\sqrt{2}}{3}x'+\frac{10}{9}z'.$$

Assim, a equação da quádrica nas variáveis x', y', z', que corresponde ao sistema de coordenadas (O, u_1, u_2, u_3), torna-se

$$9(x')^2 + 9(y')^2 - \frac{2\sqrt{2}}{3}x' + \frac{10}{9}z' + \frac{1}{3} = 0.$$

Completando o quadrado em x', temos

$$9\left(x' - \frac{\sqrt{2}}{27}\right)^2 + 9(y')^2 + \frac{10}{9}z' + \frac{1}{3} - \frac{2}{81} = 0 \Leftrightarrow 9\left(x' - \frac{\sqrt{2}}{27}\right)^2 + 9(y')^2 + \frac{10}{9}\left(z' + \frac{5}{18}\right) = 0.$$

Por fim, fazendo a mudança

$$x'' = x' - \frac{\sqrt{2}}{27}, \qquad y'' = y', \qquad z'' = z' + \frac{5}{18},$$

que corresponde a translação da origem de $O = (0,0,0)$ para $O' = \left(\frac{\sqrt{2}}{27}, 0, -\frac{5}{18}\right)$ no sistema de coordenadas, obtemos a equação

$$9(x'')^2 + 9(y'')^2 + \frac{10}{9}z'' = 0 \Leftrightarrow 9(x'')^2 + 9(y'')^2 = -\frac{10}{9}z''.$$

Note que, as interseções dessa superfície com os planos $x'' = k$ ou $y'' = k$ são parábolas e com os planos $z = k$ são circunferências se $k < 0$ ou um conjunto vazio se $k > 0$. Logo, a quádrica é um parabolóide elíptico. A Figura 6.7 mostra um esboço da superfície no sistema de coordenadas (O', u_1, u_2, u_3).

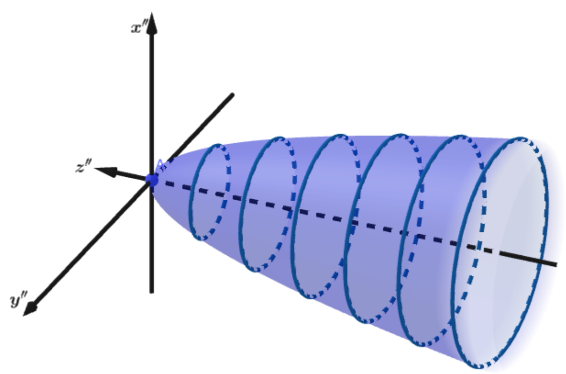


Figura 6.7: Esboço da quádrica dada no Exemplo 6.3.1

■

■ **Exemplo 6.3.2** Classifique a quádrica de equação

$$x^2 - y^2 + z^2 + 2xy - 2yz + 2xz = -1.$$

Solução. Iniciaremos com a parte quadrática da equação, a fim de eliminar seus termos mistos. Como descrito anteriormente, faremos a mudança

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = Q \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Qv'$$

com Q formada pelos autovetores ortonormais da matriz da forma quadrática associada.

(i) Forma quadrática associada:

$$q(x,y,z) = x^2 - y^2 + z^2 + 2xy - 2yz + 2xz = \begin{bmatrix} x & y & z \end{bmatrix} A \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

onde

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

(ii) Autovalores da matriz da forma quadrática:

$$\det \begin{bmatrix} 1-x & 1 & 1 \\ 1 & -1-x & -1 \\ 1 & -1 & 1-x \end{bmatrix} = -(1-x)^2(x+1) + 3x - 3 = -(1-x)[(x+1)(1-x)+3]$$

$$= -(1-x)(-x^2+4) = -(1-x)(x-2)(x+2)$$

Logo, $\lambda_1 = 1$, $\lambda_2 = 2$ e $\lambda_3 = -2$ são os autovalores de A.

Assim, se $\Lambda = \operatorname{diag}(1,2,-2)$ e Q é a matriz cujas colunas são os autovetores ortonormais de A, associados à 1, 2 e -2, respectivamente, temos

$$v^\top A v = -1 \quad \Leftrightarrow \quad v^\top Q \Lambda Q^\top v = -1 \quad \Leftrightarrow \quad (v')^\top \Lambda v' = -1.$$

Portanto, no novo sistema de coordenadas, a equação da quádrica será

$$(x')^2 + 2(y')^2 - 2(z')^2 = -1.$$

Note que, as interseções dessa superfície com os planos $x' = k$ ou $y' = k$ são hipérboles, e com os planos $z' = k$ são elipses se $k > \sqrt{2}/2$ ou $k < -\sqrt{2}/2$ ou um conjunto vazio se $-\sqrt{2}/2 < k < \sqrt{2}/2$. Portanto, a quádrica representa um hiperboloide de duas folhas.

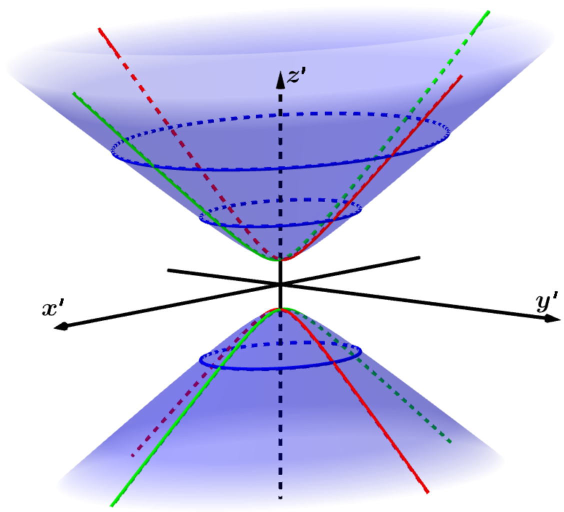


Figura 6.8: Esboço da quádrica dada no Exemplo 6.3.2

■

Referências Bibliográficas

[Axl] S. Axler, *Linear algebra done right*. 3ed., Springer, 2015.

[Coe] F. Coelho, M. Lourenço, *Um curso de álgebra linear*. 2ed., Edusp, 2004.

[Lar] R. Larson, D. Falvo, *Elementary linear algebra*. 6ed., Brooks Cole, 2008.

[Lay] D. Lay, *Linear algebra and Its applications*. 5ed., Pearson, 2014.

[Mey] C. Meyer, *Matrix analysis and applied linear algebra*. SIAM, 2010.

[Nic] W. Nicholson, *Linear algebra with applications*. 7ed., McGraw-Hill, 2013.

[Reg] R. Santos, *Álgebra linear e aplicações.* UFMG, 2013.

[Str] G. Strang, *Linear algebra and Its applications.* Cengage Learning, 2015.

Índice Remissivo

FOLHA DE REGISTRO DO DOCUMENTO

1. CLASSIFICAÇÃO/TIPO	2. DATA	3. REGISTRO N°	4. N° DE PÁGINAS
MD	06 de abril de 2023	DCTA/ITA/MD-002/2023	145

5. TÍTULO E SUBTÍTULO:

MAT 27 - Álgebra linear.

6. AUTOR(ES):

Fernanda de Andrade Pereira; Tiara Martini dos Santos

7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES):

Instituto Tecnológico de Aeronáutica – ITA / IEF

8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR:

Álgebra linear, Espaços vetoriais, Produto interno, Transformações lineares, Formas quadráticas, Quádricas.

9.PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO:

Álgebra linear; Sistemas lineares; Espaços vetoriais; Transformações lineares; Formas quadráticas; Matemática.

10. APRESENTAÇÃO: **X Nacional** **Internacional**

Departamento de Matemática; São José dos Campos, SP, Brasil; Abril de 2023; Curso Fundamental, Álgebra Linear, 1º edição, 2023.

11. RESUMO:

Esse texto foi elaborado a partir das notas de aula das autoras para o curso de MAT-27 - Álgebra Linear, disciplina obrigatória aos cursos de engenharia do ITA, com objetivo de servir como livro-texto para a mesma. Esse material aborda todo o conteúdo que normalmente é dado nas disciplinas de Álgebra Linear em cursos de graduação nas áreas de exatas. Nele são considerados espaços vetoriais reais e complexos, com foco nos espaços de dimensão finita.

12. GRAU DE SIGILO:

(X) OSTENSIVO **( ) RESERVADO** **( ) SECRETO**