Notas de álgebra linear

Max Jáuregui

27 de Julho de 2019

Conteúdo

1 Espaços e subespaços vetoriais

Um conjunto E é um *espaço vetorial* se estão definidas em E as operações de *adição* e multiplicação por um escalar (número real), que cumprem os seguintes axiomas:

1. para quaisquer $u, v \in E$ e $\alpha \in \mathbb{R}$, $u + v \in E$ e $\alpha u \in E$;
2. para quaisquer $u, v, w \in E$ e $\alpha, \beta \in \mathbb{R}$, $u + (v + w) = (u + v) + w$ e $\alpha(\beta u) = (\alpha\beta)u$;
3. para quaisquer $u, v \in E$, $u + v = v + u$;
4. existe $0 \in E$ tal que $u + 0 = 0$ para todo $u \in E$;
5. para qualquer $u \in E$ existe $-u \in E$ tal que $u + (-u) = 0$;
6. para qualquer $u \in E$, $1u = u$;
7. para quaisquer $u, v \in E$ e $\alpha, \beta \in \mathbb{R}$, $\alpha(u + v) = \alpha u + \alpha v$ e $(\alpha + \beta)u = \alpha u + \beta u$.

Os elementos de E são chamados de *vetores*.

Exemplos de espaços vetoriais são:

- $\mathbb{R}^n$: espaço euclidiano n-dimensional;

- $\mathbb{R}^X$: conjunto das funções $f : X \to Y$, definindo $(f+g)(x) = f(x) + g(x)$ e $(\alpha f)(x) = \alpha f(x)$;

- $M(m \times n)$: conjunto das matrizes $m \times n$.

Teorema 1.1. *Se u e v são vetores de um espaço vetorial E, então $u+v = u+w \Rightarrow v = w$.*

Corolário 1.2. *Se u e v são vetores de um espaço vetorial E, então*

1. $u + v = u \Rightarrow v = 0$;

2. $u + v = 0 \Rightarrow v = -u$;

3. $0u = 0$;

4. $(-1)u = -u$.

Um *subespaço* de um espaço vetorial E é um conjunto $F \subset E$ tal que

1. $0 \in F$;

2. para quaisquer $u, v \in F$, $u + v \in F$;

3. para quaisquer $\alpha \in \mathbb{R}$ e $u \in F$, $\alpha u \in F$.

Exemplos de subespaços são:

- A origem $\{0\}$;

- Retas que passam pela origem: fixado $v \in E$, $v \neq 0$,

$$F = \{\alpha v \in E : \alpha \in \mathbb{R}\};$$

- Hiperplanos que passam pela origem:

$$H = \{(x_1, \ldots, x_n) \in E : \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\}.$$

Como a interseção de uma família arbitrária de subespaços é claramente um subespaço, segue em particular que o conjunto das soluções do sistema linear

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \vdots \qquad\qquad & \ \ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

é um subespaço de $\mathbb{R}^n$;

Seja E um espaço vetorial. O *subespaço gerado* por um conjunto $X \subset E$ é o conjunto $S(X)$ de todas as combinações lineares $\alpha_1 x_1 + \cdots + \alpha_n x_n$, com $x_1, \ldots, x_n \in X$. Pode-se provar facilmente que $S(X)$ é o menor subespaço que contém X, no sentido que se $F \subset E$ é um subespaço que contém X, então $S(X) \subset F$. Se $S(X) = F$, diz-se que X é um *conjunto de geradores de F*.

Dado um espaço vetorial E, seja $v \in E$ com $v \neq 0$. O subespaço gerado por $\{v\}$ é a reta que contém v e passa pela origem.

O conjunto $\{(1,0), (0,1)\} \subset \mathbb{R}^2$ é claramente um conjunto de geradores de $\mathbb{R}^2$.

Se F_1 e F_2 são subespaços, pode-se verificar imediatamente que o subespaço gerado por $F_1 \cup F_2$ é o conjunto $F_1 + F_2 = \{u + v : u \in F_1, v \in F_2\}$. Quando $F_1 \cap F_2 = \{0\}$, escreve-se $F_1 \oplus F_2$ no lugar de $F_1 + F_2$ e diz-se que $F_1 \oplus F_2$ é a *soma direta* de F_1 e F_2.

Teorema 1.3. *Sejam F, F_1 e F_2 subespaços. As seguintes afirmações são equivalentes:*

1. *$F = F_1 \oplus F_2$;*

2. *Para cada $v \in F$ existem únicos $u_1 \in F_1$ e $u_2 \in F_2$ tais que $v = u_1 + u_2$.*

Demonstração. ($1 \Rightarrow 2$) Se existem $u_1, v_1 \in F_1$ e $u_2, v_2 \in F_2$ tais que $v = u_1 + u_2 = v_1 + v_2$, então $F_1 \ni u_1 - v_1 = v_2 - u_2 \in F_2$. Logo, devemos ter $u_1 = v_1$ e $u_2 = v_2$. ($2 \Rightarrow 1$) Se $v \in F_1 \cap F_2$, então $v = v + 0 = 0 + v$, o que implica que $v = 0$. $\square$

Podemos verificar facilmente que $\mathbb{R}^2$ é a soma direta das retas que passam pela origem $\{(\alpha, \alpha) \in \mathbb{R}^2 : \alpha \in \mathbb{R}\}$ e $\{(-\alpha, \alpha) \in \mathbb{R}^2 : \alpha \in \mathbb{R}\}$.

Seja E um espaço vetorial. A *reta* que passa pelos pontos $u, v \in E$ é o conjunto

$$R = \{(1 - \alpha)u + \alpha v : \alpha \in \mathbb{R}\}.$$

Uma *variedade afim* é um conjunto $V \subset E$ tal que a reta que passa por qualquer par de pontos de V está contida em V. Vemos dessa definição que todo subespaço é uma variedade afim. De fato toda variedade afim pode ser obtida transladando um subespaço de acordo com o seguinte

Teorema 1.4. *Se V é uma variedade afim, existe um único subespaço F tal que $V = x + F = \{x\} + F$ para qualquer $x \in V$.*

Demonstração. Fixado $x_0 \in V$, seja $F_0 = -x_0 + V$. Vemos que $0 \in F_0$. Dado $u \in F_0$, temos que $x_0 + u \in V$, o que implica que $\forall \alpha \in \mathbb{R}\,, x_0 + \alpha u = [\alpha(x_0 + u) + (1 - \alpha)x_0] \in V$. Logo, $\alpha u \in F_0$. Dados $u, v \in F$, temos que $[(x_0 + u)/2 + (x_0 + v)/2] \in V$. Isso implica que $(u + v)/2 \in F$ e, por conseguinte, $u + v \in F_0$. Portanto, F_0 é um subespaço. Se existe um subespaço $G_0 \subset E$ tal que $V = x_0 + G_0$, $u \in G_0 \Leftrightarrow x_0 + u \in V \Leftrightarrow u \in F_0$. Portanto, para cada $x \in V$ existe um único subespaço $F \subset E$ tal que $V = x + F$. Logo, dados $x, y \in V$, existem únicos subespaços $F, G \subset E$ tais que $V = x + F = y + G$. Vemos que $u \in F \Rightarrow x + u \in V \Rightarrow (-y + x) + u \in G$. Como $-y + x \in G$, segue que $u \in G$. Logo, $F \subset G$. Analogamente podemos provar que $G \subset F$. $\square$

O conjunto das soluções do sistema linear

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1}$$

é claramente uma variedade afim V de $\mathbb{R}^n$. Se $v = (x_1, \ldots, x_n) \in V$ e F é o subespaço das soluções do sistema homogêneo associado, pelo teorema 1.4, $V = v + F$ e, por conseguinte, toda solução do sistema (1) pode ser obtida somando a solução particular v com a solução geral do sistema homogêneo associado.

2 Bases

Seja E um espaço vetorial. Diz-se que um conjunto $X \subset E$ é *linearmente independente (L.I.)* quando nenhum vetor de X é combinação linear de vetores de X. Nesse caso também diz-se que os vetores de X são L.I.. Se $X = \{v\}$, põe-se por definição que X é L.I. quando $v \neq 0$. Vemos imediatamente que, se X é L.I., qualquer $Y \subset X$ é L.I..

O conjunto $\{(1, 0), (0, 1)\}$ é claramente um subconjunto *L.I.* de $\mathbb{R}^2$. O conjunto $\{x, x^2, x^3, \ldots\}$ é claramente um subconjunto L.I. do espaço vetorial dos polinômios em x.

Teorema 2.1. *Um conjunto X é L.I. se, e somente se,*

$$\forall x_1, \ldots, x_n \in X\,, \alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \Rightarrow \alpha_1 = \ldots = \alpha_n = 0\,.$$

Demonstração. ($\Rightarrow$) Se existem $x_1, \ldots, x_n \in X$ tais que $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ com $\alpha_n \neq 0$, então $x_n = -(\alpha_1/\alpha_n)x_1 - \cdots - (\alpha_{n-1}/\alpha_n)x_{n-1}$, o que implica que X não é L.I.. ($\Leftarrow$) Se X não é L.I., existem $v, x_1, \ldots, x_n \in X$ tais que $v = \alpha_1 x_1 + \cdots + \alpha_n x_n$, o que implica que $v - \alpha_1 x_1 - \cdots - \alpha_n x_n = 0$. □

Corolário 2.2. *Se $\{x_1, \ldots, x_n\}$ é L.I. e $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta_1 x_1 + \cdots + \beta_n x_n$, então $\alpha_i = \beta_i$, $i = 1, \ldots, n$.*

Teorema 2.3. *O conjunto $X = \{x_1, \ldots, x_n\}$ é L.I. se nenhum vetor é combinação linear dos anteriores.*

Demonstração. Se X não é L.I., existem $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ não todos nulos tais que $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$. Seja r o maior índice tal que $\alpha_r \neq 0$. Logo, $x_r = -(\alpha_1/\alpha_r)x_1 - \cdots - (\alpha_{r-1}/\alpha_r)x_{r-1}$. □

Seja E um espaço vetorial. Um conjunto $X \subset E$ é *linearmente dependente (L.D.)* quando não é L.I.. Vemos imediatamente que, se um subconjunto $Y \subset X$ é L.D., X é L.D.. Em particular, se $0 \in X$, X é L.D..

O conjunto $\{(1,0), (0,1), (1,1)\} \subset \mathbb{R}^2$ é claramente L.D..

Seja E um espaço vetorial. Uma *base* de E é um conjunto L.I. $B \subset E$ tal que $S(B) = E$. Se B é uma base de E e $v = \alpha_1 x_1 + \cdots + \alpha_n x_n$ com $x_1, \ldots, x_n \in B$, pelo corolário 2.2, os números $\alpha_1, \ldots, \alpha_n$ são únicos e são chamados de as *coordenadas* do vetor v na base B.

O conjunto $\{(0,1), (1,0)\} \subset \mathbb{R}^2$ é claramente uma base de $\mathbb{R}^2$, chamada de a *base canônica* de $\mathbb{R}^2$. O conjunto $\{1, x, x^2, \ldots\}$ é claramente uma base do espaço vetorial dos polinômios em x.

Lema 2.4. *O sistema linear homogêneo*

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \vdots \qquad\qquad & \ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{2}$$

tem uma solução não trivial quando $m < n$, ou seja, tem uma solução $(x_1, \ldots, x_n) \neq 0$.

Demonstração. Procedemos por indução no número de equações. Se temos uma equação $a_{11}x_1 + \cdots + a_{1n}x_n = 0$ com $n > 1$ e $a_{1n} \neq 0$, então, fixados $x_1, \ldots, x_{n-1} \in \mathbb{R}$,

$$\left(x_1, \ldots, x_{n-1}, -\frac{a_{11}}{a_{1n}}x_1 - \cdots - \frac{a_{1n-1}}{a_{1n}}x_{n-1}\right)$$

é uma solução não trivial. Suponhamos que o lema seja válido para $m-1$ equações e consideremos o sistema (2). Renomeando os coeficientes e as variáveis se necessário, podemos supor que $a_{mn} \neq 0$. Logo, da última equação de (2) obtemos que

$$x_n = -\frac{a_{m1}}{a_{mn}}x_1 - \cdots - \frac{a_{mn-1}}{a_{mn}}x_{n-1}\,. \tag{3}$$

Substituindo essa expressão nas $m-1$ equações restantes, obtemos um sistema de $m-1$ equações com $n-1 > m-1$ incógnitas, o qual, pela hipótese de indução, tem uma solução não trivial $(x_1, \ldots, x_{n-1})$. Logo, $(x_1, \ldots, x_{n-1}, x_n)$ com x_n dado por (3) é uma solução não trivial de (2). □

Teorema 2.5. *Seja E um espaço vetorial. Se $\{u_1, \ldots, u_n\}$ é um conjunto de geradores de E, então qualquer conjunto com mais de n vetores é L.D..*

Demonstração. Dados $v_1, \ldots, v_{n+1} \in E$, temos que $v_i = \alpha_{1i}u_1 + \cdots + \alpha_{ni}u_n$, $i = 1, \ldots, n+1$. Se $\beta_1 v_1 + \cdots + \beta_{n+1}v_{n+1} = 0$, então

$$\left(\sum_{i=1}^{n+1} \alpha_{1i}\beta_i\right) u_1 + \cdots + \left(\sum_{i=1}^{n+1} \alpha_{ni}\beta_i\right) u_n = 0\,,$$

o qual é satisfeito em particular quando

$$\begin{aligned} \alpha_{11}\beta_1 + \cdots + \alpha_{1n+1}\beta_{n+1} &= 0 \\ \vdots \qquad\qquad & \quad \vdots \\ \alpha_{n1}\beta_1 + \cdots + \alpha_{nn+1}\beta_{n+1} &= 0\,. \end{aligned}$$

Pelo lema 2.4, esse sistema tem uma solução não trivial $(\beta_1, \ldots, \beta_{n+1})$, o que implica que $\{v_1, \ldots, v_{n+1}\}$ é L.D.. □

Corolário 2.6. *Seja E um espaço vetorial. Se $\{x_1, \ldots, x_n\}$ é uma base de E, então toda base de E tem n elementos.*

Um espaço vetorial E tem *dimensão finita* quando existe uma base $\{u_1, \ldots, u_n\} \subset E$. O número n é chamado de *dimensão* de E e é denotado por $\dim E$. Se $E = \{0\}$, põe-se por definição que $\dim E = 0$. Quando um espaço vetorial não tem dimensão finita, diz-se que tem *dimensão infinita*.

Corolário 2.7. *Seja E um espaço vetorial com $\dim E = n$. Um conjunto $X \subset E$ de n elementos é L.I. se, e somente se, $S(X) = E$.*

Teorema 2.8. *Seja E um espaço vetorial com $\dim E = n$. Tem-se que*

1. *todo conjunto de geradores de E contém uma base;*

2. *todo conjunto L.I. de E pode ser estendido a uma base;*

3. *todo subespaço $F \subset E$ é de dimensão finita e $\dim F \leq n$;*

4. *se $F \subset E$ é um subespaço tal que $\dim F = n$, então $F = E$.*

Demonstração.

1. Seja X um conjunto de geradores de E e seja $B = \{x_1, \ldots, x_m\}$ um subconjunto de X com o número máximo de vetores L.I. ($m \leq n$). Se existe $u \in X$ tal que u não é combinação linear de vetores de B, então $B \cup \{u\}$ seria um conjunto L.I.. Contradição! Logo, todo vetor de X deve ser combinação linear de vetores de B. Como $X \subset S(B)$ e $S(X) = E$, segue que $S(B) = E$, ou seja, B é uma base de E.

2. Seja $\{v_1, \ldots, v_k\} \subset E$ um conjunto L.I.. Procuramos um vetor $v_{k+1} \in E$ que não seja combinação linear de $v_1, \ldots, v_k$. Se esse vetor não existe, $\{v_1, \ldots, v_k\}$ é um conjunto de geradores de E e, por conseguinte, uma base de E. Por outro lado, se tal vetor existe, procuramos por um vetor $v_{k+2} \in E$ que não seja combinação linear de $v_1, \ldots, v_k, v_{k+1}$. Seguindo esse procedimento, o qual tem fim devido a que existem no máximo n vetores L.I. em E, obtemos um conjunto $B = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_m\} \subset E$ que tem o número máximo de vetores L.I.. Logo, B é uma base de E.

3. Seja $B = \{v_1, \dots, v_m\} \subset F$ um conjunto com o número máximo de vetores L.I.. Logo, B é uma base de F. Como $B \subset E$ é L.I., deve-se ter $m \leq n$.

4. Seja $B = \{u_1, \dots, u_n\}$ uma base de F. Como $B \subset E$ é um conjunto L.I. que tem n elementos, $S(B) = E$, ou seja, B é uma base de E. Portanto, $F = E$. $\square$

O hiperplano $H = \{(x_1, \dots, x_n) \in \mathbb{R}^n : \alpha_1 x_1 + \cdots + \alpha_n x_n = 0\}$ tem dimensão $n-1$. Com efeito, assumindo que $\alpha_n \neq 0$, temos que

$$x_n = -\frac{\alpha_1}{\alpha_n} x_1 - \cdots - \frac{\alpha_{n-1}}{\alpha_n} x_{n-1}\,.$$

Logo, os vetores

$$(1, 0, \dots, 0, -\alpha_1/\alpha_n), (0, 1, \dots, 0, -\alpha_2/\alpha_n), \dots, (0, 0, \dots, 1, -\alpha_{n-1}/\alpha_n)$$

formam uma base de H.

Define-se a *dimensão* de uma variedade afim V como a dimensão do subespaço F tal que $V = x + F$ para qualquer $x \in V$.

3 Transformações lineares

Dados os espaços vetoriais E e F, uma *transformação linear* é uma aplicação $A : E \to F$ tal que $A(u+v) = Au + Av$ e $A(\alpha u) = \alpha A u$ para quaisquer $\alpha \in \mathbb{R}$ e $u, v \in E$. Se $F = E$, diz-se que A é um *operador linear.* Se $F = \mathbb{R}$, diz-se que A é um *funcional linear.*

O conjunto $\mathcal{L}(E, F)$ de todas as transformações lineares $A : E \to F$ é claramente um espaço vetorial. Se $F = E$, escrevemos $\mathcal{L}(E)$ no lugar de $\mathcal{L}(E, E)$. Se $F = \mathbb{R}$, escrevemos E^* no lugar de $\mathcal{L}(E, \mathbb{R})$, que é chamado de *espaço dual* de E.

Teorema 3.1. *Sejam E e F espaços vetoriais, $\mathcal{B}$ uma base de E e $f : \mathcal{B} \to F$ uma aplicação qualquer. Existe uma única transformação linear $A : E \to F$ tal que $Au = f(u)$ para todo $u \in \mathcal{B}$.*

Demonstração. Definimos a aplicação $A : E \to F$ pondo $Av = \alpha_1 f(u_1) + \cdots + \alpha_n f(u_n)$ quando $v = \alpha_1 u_1 + \cdots + \alpha_n u_n$ com $u_1, \dots, u_n \in \mathcal{B}$. Vamos provar que A é uma transformação linear. Dados $v, w \in E$, existem $u_1, \dots, u_n \in \mathcal{B}$ e $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{R}$, talvez não todos diferentes de zero, tais que $v = \alpha_1 u_1 + \cdots + \alpha_n u_n$ e $w = \beta_1 u_1 + \cdots + \beta_n u_n$. Logo,

$$\begin{aligned} A(u+v) &= A\left(\sum_{i=1}^{n} (\alpha_i + \beta_i) u_i\right) \\ &= \sum_{i=1}^{n} (\alpha_i + \beta_i) f(u_i) \\ &= \sum_{i=1}^{n} \alpha_i f(u_i) + \sum_{i=1}^{n} \beta_i f(u_i) \\ &= Av + Aw\,. \end{aligned}$$

Por outro lado, dado $\gamma \in \mathbb{R}$, temos que

$$A(\gamma v) = A\left(\sum_{i=1}^{n} (\gamma \alpha_i) u_i\right) = \gamma \sum_{i=1}^{n} \alpha_i f(u_i) = \gamma A v\,.$$

Vamos provar agora a unicidade de A. Seja $B : E \to F$ uma transformação linear tal que $B(u) = f(u)$ para todo $u \in \mathcal{B}$. Para qualquer $v \in E$ existem $u_1, \ldots, u_n \in \mathcal{B}$ tais que $v = \alpha_1 u_1 + \cdots + \alpha_n u_n$. Logo,

$$Bv = B\left(\sum_{i=1}^{n} \alpha_i u_i\right) = \sum_{i=1}^{n} \alpha_i f(u_i) = Av\,. \qquad \square$$

Se E é um espaço vetorial com $\dim E = 1$, toda operador linear $A : E \to E$ é da forma $A = \lambda I$, onde $\lambda \in \mathbb{R}$ e $I : E \to E$ é o *operador identidade* definido por $Iv = v$. Com efeito, escolhendo $u \in E$, $u \neq 0$, temos que $\{u\}$ é uma base de E. Pelo teorema 3.1, o operador A fica definido pela igualdade $Au = w$. Logo, dado $v \in E$, existe $\alpha \in \mathbb{R}$ tal que $v = \alpha u$ e, por conseguinte, $Av = \alpha w$. Como $w = \lambda u$ para um certo $\lambda \in \mathbb{R}$, $Av = \alpha\lambda u = \lambda v = \lambda I v$.

A rotação de vetores no plano é um exemplo de um operador linear. Com efeito, esse operador R está definido pelas equações $R(1, 0) = (\cos\theta, \operatorname{sen}\theta)$ e $R(0, 1) = (-\operatorname{sen}\theta, \cos\theta)$.

Se $C_0([a, b])$ é o espaço vetorial das funções contínuas $f : [a, b] \to \mathbb{R}$, a função $I : C_0([a, b]) \to \mathbb{R}$ definida por $I(f) = \int_a^b f(x)\,dx$ é um funcional linear.

Se $A : E \to F$ e $B : F \to G$ são transformações lineares, então a aplicação $BA : E \to G$, definida por $BAv = B(Av)$, é uma transformação linear, chamada de o *produto* de B e A. Com efeito, dados $u, v \in E$ e $\alpha \in \mathbb{R}$ temos que $BA(u + v) = B(Au + Av) = BAu + BAv$ e $BA(\alpha v) = B(\alpha Av) = \alpha BAv$.

Um operador $A : E \to E$ é dito *nilpotente* quando existe $k \in \mathbb{N}$ tal que $A^k = 0$. Por exemplo, o operador $A : \mathbb{R}^2 \to \mathbb{R}^2$ definido por $A(1, 0) = (0, 1)$ e $A(0, 1) = 0$ é nilpotente, pois $A^2 = 0$.

A *imagem* de uma transformação linear $A : E \to F$ é o conjunto $Im(A) = \{Av \in F : v \in E\}$. Vemos imediatamente que $Im(A)$ é um subespaço de F. Se $Im(A) = F$, diz-se que A é *sobrejetiva*.

Um funcional linear $A : E \to \mathbb{R}$ é nulo ou é sobrejetivo, pois os únicos subespaços de $\mathbb{R}$ são $\{0\}$ e o próprio $\mathbb{R}$.

Teorema 3.2. *Uma transformação linear $A : E \to F$ é sobrejetiva se, e somente se, todo conjunto de geradores de E é transformado em um conjunto de geradores de F.*

Demonstração. ($\Rightarrow$) Dado $w \in F$, existe $v \in E$ tal que $Av = w$. Se $X \subset E$ é um conjunto de geradores de E, existem $u_1, \ldots, u_n \in X$ tais que $v = \alpha_1 u_1 + \cdots + \alpha_n u_n$. Logo, $w = \alpha_1 Au_1 + \cdots + \alpha_n Au_n$. Portanto, $\{Au \in F : u \in X\}$ é um conjunto de geradores de F. ($\Leftarrow$) Dado $w \in F$, existem $u_1, \ldots, u_n \in E$ tais que $w = \alpha_1 Au_1 + \cdots + \alpha_n Au_n = A(\alpha_1 u_1 + \cdots + \alpha_n u_n)$. Logo, A é sobrejetiva. $\square$

Corolário 3.3. *Se $A : E \to F$ é uma transformação linear e $\dim E < \infty$, então $\dim Im(A) \leq \dim E$.*

Uma transformação linear $B : F \to E$ é dita uma *inversa à direita* de uma transformação linear $A : E \to F$ quando $AB : F \to F$ é o operador identidade, ou seja, $ABv = v$ para todo $v \in F$.

Teorema 3.4. *Seja F um espaço vetorial de dimensão finita. Uma transformação linear $A : E \to F$ é sobrejetiva se, e somente se, possui uma inversa à direita.*

Demonstração. ($\Rightarrow$) Seja X um conjunto de geradores de E. Pelo teorema 3.2, o conjunto $\{Au : u \in X\}$ é um conjunto de geradores de F. Como $\dim F < \infty$, existem $u_1, \ldots, u_n \in X$

tais que $\{Au_1, \ldots, Au_n\}$ é uma base de F. Definimos então uma transformação linear $B : F \to E$ pondo $B(Au_i) = u_i$, $i = 1, \ldots, n$. Dado $v \in F$, temos que $v = \alpha_1 Au_1 + \cdots + \alpha_n Au_n$ e $ABv = A(\alpha_1 u_1 + \cdots + \alpha_n u_n) = v$. Portanto, B é uma inversa à direita de A. $(\Leftarrow)$ Seja $B : F \to E$ uma inversa à direita de A. Dado $w \in F$, seja $Bw = v$. Dessa maneira, existe $v \in E$ tal que $Av = ABw = w$, ou seja, A é sobrejetiva. $\square$

O funcional $I : C^0([a, b]) \to \mathbb{R}$ definido por $I(f) = \int_a^b f(x)\,dx$ é sobrejetivo. Uma inversa à direita de I é a transformação linear $J : \mathbb{R} \to C^0([a, b])$ definida por $J(1) = 1/(b - a)$.

O *núcleo* de uma transformação linear $A : E \to F$ é o conjunto $N(A) = \{v \in E : Av = 0\}$. Assim como a imagem de A, podemos verificar imediatamente que $N(A)$ é um subespaço de E.

Uma transformação linear $A : E \to F$ é *injetiva* se, dados $u, v \in E$, $Au = Av \Rightarrow u = v$. Por exemplo, a transformação linear $A : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $A(1, 0) = (1, 0, 0)$ e $A(0, 1) = (0, 1, 0)$ é injetiva.

Teorema 3.5. *Uma transformação linear $A : E \to F$ é injetiva se, e somente se, $N(A) = \{0\}$.*

Demonstração. $(\Rightarrow)$ Dado $v \in E$, se $Av = 0 = A0$, então $v = 0$. $(\Leftarrow)$ Dados $u, v \in E$, $Au = Av \Rightarrow A(u - v) = 0 \Rightarrow u - v \in N(A) \Rightarrow u - v = 0$. $\square$

Teorema 3.6. *Uma transformação linear $A : E \to F$ é injetiva se, e somente se, todo subconjunto L.I. de E é transformado em um subconjunto L.I. de F.*

Demonstração. $(\Rightarrow)$ Dado um conjunto L.I. $X \subset E$, seja $Y = \{Au : u \in X\}$. Se $v_1, \ldots, v_n \in Y$ e $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$, existem $u_1, \ldots, u_n \in X$ tais que $v_i = Au_i$, $i = 1, \ldots, n$. Logo, $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \Rightarrow \alpha_1 Au_1 + \cdots + \alpha_n Au_n = A(\alpha_1 u_1 + \cdots + \alpha_n u_n) = 0 \Rightarrow \alpha_1 u_1 + \cdots + \alpha_n u_n = 0 \Rightarrow \alpha_1 = \ldots = \alpha_n = 0$. $(\Leftarrow)$ Para qualquer $v \in E$ com $v \neq 0$ temos que $Av \neq 0$. Logo, $N(A) = \{0\}$ e, pelo teorema 3.5, A é injetiva. $\square$

Corolário 3.7. *Sejam E e F espaços vetoriais de dimensão finita. Se existe uma transformação linear injetiva $A : E \to F$, então $\dim E \leq \dim F$.*

Segue do corolário 3.7 que não existe transformação linear injetiva $A : \mathbb{R}^3 \to \mathbb{R}^2$.

Teorema 3.8. *Seja $A : E \to F$ uma transformação linear. Para cada $b \in Im(A)$, o conjunto $V = \{x \in E : Ax = b\}$ é uma variedade afim de E paralela a $N(A)$.*

Demonstração. Verifica-se facilmente que V é uma variedade afim. Dado $x_0 \in V$, vemos que $x \in V \Leftrightarrow x - x_0 \in N(A) \Leftrightarrow x \in x_0 + N(A)$. Portanto, $V = x_0 + N(A)$. $\square$

Uma transformação linear $B : F \to E$ é uma *inversa à esquerda* de uma transformação linear $A : E \to F$ se $BA : E \to E$ é o operador identidade.

Teorema 3.9. *Sejam E e F espaços vetoriais de dimensão finita. Uma transformação linear $A : E \to F$ é injetiva se, e somente se, possui uma inversa à esquerda.*

Demonstração. $(\Rightarrow)$ Seja $\{u_1, \ldots, u_n\}$ uma base de E. Pelo teorema 3.6, o conjunto $\{Au_1, \ldots, Au_n\}$ é L.I.. Logo, pode ser estendido a uma base $\{Au_1, \ldots, Au_n, v_1, \ldots, v_m\}$ de F. Definimos então a transformação linear $B : F \to E$ pondo $B(Au_i) = u_i$ para $i = 1, \ldots, n$ e $Bv_j = u_1$ para $j = 1, \ldots, m$. Dado $v \in E$, temos que $v = \alpha_1 + \cdots + \alpha_n v_n$ e $Av = \alpha_1 Au_1 + \cdots + \alpha_n Au_n$. Logo, $BAv = \alpha_1 u_1 + \cdots + \alpha_n u_n = v$. Dessa maneira, B é uma inversa à esquerda de A. $(\Leftarrow)$ Se $B : F \to E$ é uma inversa à esquerda de A, dados $u, v \in E$, $Au = Av \Rightarrow u = BAu = BAv = v$. Portanto, A é injetiva. $\square$

A transformação linear $A : \mathbb{R} \to \mathbb{R}^2$ definida por $A(1) = (1, 1)$ é claramente injetiva. Uma inversa à esquerda de A é o funcional linear $B : \mathbb{R}^2 \to \mathbb{R}$ definido por $B(1, 1) = 1$ e $B(-1, 1) = 0$.

Uma transformação linear $A : E \to F$ é *invertível* quando possui uma inversa à esquerda e uma inversa à direita. Nesse caso, se $B, C : F \to E$ são respectivamente uma inversa à esquerda e uma inversa à direita de A, então $B = B(AC) = (BA)C = C$, ou seja, A possui uma única transformação linear inversa, chamada de a *inversa* de A e denotada usualmente por A^{-1}.

Uma transformação linear $A : E \to F$ é chamada de um *isomorfismo* quando é injetiva e sobrejetiva. Nesse caso, diz-se também que E e F são *isomorfos*. Em virtude dos teoremas 3.2 e 3.6, uma transformação linear $A : E \to F$ é um isomorfismo se, e somente se, transforma bases de E em bases de F. Além disso, se E e F têm dimensão finita, segue dos dos corolários 3.3 e 3.7 que $\dim E = \dim F$. Nesse caso, pelos teoremas 3.4 e 3.9, temos também que uma transformação linear $A : E \to F$ é um isomorfismo se, e somente se, é invertível.

Se E e F são espaços vetoriais de dimensão finita com $\dim E = \dim F$, então eles são isomorfos. Com efeito, se $\{u_1, \dots, u_n\}$ é uma base de E e $\{v_1, \dots, v_n\}$ é uma base de F, basta definir uma transformação linear $A : E \to F$ pondo $Au_i = v_i$, $i = 1, \dots, n$. Como A transforma uma base de E em uma base de F, A é um isomorfismo.

O espaço vetorial dos polinômios em x de grau $\leq n$ é isomorfo a $\mathbb{R}^{n+1}$, pois ambos têm a mesma dimensão.

Teorema 3.10 (Teorema do núcleo e da imagem)**.** *Sejam E e F espaços vetoriais de dimensão finita. Se $A : E \to F$ é uma transformação linear, então* $\dim E = \dim N(A) + \dim Im(A)$.

Demonstração. Sejam $\{Au_1, \dots, Au_m\}$ uma base de $Im(A)$ e $\{v_1, \dots, v_n\}$ uma base de $N(A)$. O conjunto $\mathcal{B} = \{u_1, \dots, u_m, v_1, \dots, v_n\}$ é L.I., pois $\alpha_1 u_1 + \dots + \alpha_m u_m + \beta_1 v_1 + \dots + \beta_n v_n = 0 \Rightarrow \alpha_1 Au_1 + \dots + \alpha_m Au_m + 0 = 0 \Rightarrow \alpha_1 = \ldots = \alpha_m = 0$ e, por conseguinte, $\alpha_1 u_1 + \dots + \alpha_m u_m + \beta_1 v_1 + \dots + \beta_n v_n = 0 \Rightarrow \beta_1 v_1 + \dots + \beta_n v_n = 0 \Rightarrow \beta_1 = \ldots = \beta_n = 0$. Dado $v \in E$, existem $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ tais que $Av = \alpha_1 Au_1 + \dots + \alpha_m Au_m$. Logo, $A(v - \alpha_1 u_1 - \dots - \alpha_m u_m) = 0$, o que implica que $v - \alpha_1 u_1 - \dots - \alpha_m u_m \in N(A)$. Consequentemente, existem $\beta_1, \dots, \beta_n \in \mathbb{R}$ tais que $v - \alpha_1 u_1 - \dots - \alpha_m u_m = \beta_1 v_1 + \dots + \beta_n v_n$, ou seja, $v = \alpha_1 u_1 + \dots + \alpha_m u_m + \beta_1 v_1 + \dots + \beta_n v_n$. Portanto $\mathcal{B}$ é uma base de E. □

Corolário 3.11. *Sejam E e F espaços vetoriais de dimensão finita com* $\dim E = \dim F$. *Uma transformação linear $A : E \to F$ é injetiva se, e somente se, é sobrejetiva.*

O corolário 3.11 em geral é falso para espaços vetoriais de dimensão infinita. Por exemplo, consideremos os operadores lineares $A, B : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ definidos por $A(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ e $B(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$. Vemos que A é injetivo mas não é sobrejetivo e que B é sobrejetivo mas não é injetivo.

4 Soma direta e projeção

Sejam E_1 e E_2 espaços vetoriais. O produto cartesiano $E_1 \times E_2$ é um espaço vetorial se definimos $(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$ e $\alpha(u_1, v_1) = (\alpha u_1, \alpha v_1)$ para quaisquer $u_1, v_1 \in E_1$, $u_2, v_2 \in E_2$ e $\alpha \in \mathbb{R}$.

Sejam F_1 e F_2 subespaços de um espaço vetorial E de dimensão finita tais que $E = F_1 \oplus F_2$. Se $\{u_1, \ldots, u_m\}$ e $\{v_1, \ldots, v_n\}$ são bases de F_1 e F_2 respectivamente, vemos imediatamente que $\{(u_1, 0), \ldots, (u_m, 0), (0, v_1), \ldots, (0, v_n)\}$ é uma base de $F_1 \times F_2$. Logo, a transformação linear $A : F_1 \times F_2 \to E$ definida por $A(u_i, 0) = u_i$, $i = 1, \ldots, m$, e $A(0, v_j) = v_j$, $j = 1, \ldots, n$, é um isomorfismo, pois $\{u_1, \ldots, u_m, v_1, \ldots, v_n\}$ é uma base de E.

Teorema 4.1. *Sejam E_1 e E_2 espaços vetoriais de dimensão finita. Tem-se que* $\dim E_1 + \dim E_2 = \dim(E_1 \cap E_2) + \dim(E_1 + E_2)$.

Demonstração. A transformação linear $A : E_1 \times E_2 \to E_1 + E_2$ definida por $A(v_1, v_2) = v_1 + v_2$ é claramente sobrejetiva. Vemos que $A(v_1, v_2) = 0 \Rightarrow v_1 = -v_2 \Rightarrow N(A) = \{(v, -v) \in E_1 \times E_2 : v \in E_1\}$. Considerando o isomorfismo $B : E_1 \cap E_2 \to N(A)$ definido por $B(v) = (v, -v)$, temos que $\dim N(A) = \dim(E_1 \cap E_2)$. Portanto, pelo teorema do núcleo e da imagem, $\dim E_1 + \dim E_2 = \dim(E_1 \times E_2) = \dim(E_1 \cap E_2) + \dim(E_1 + E_2)$. □

Sejam F_1, F_2 subespaços de um espaço vetorial E tais que $E = F_1 \oplus F_2$. Define-se a *projeção de E sobre F_1 paralelamente a F_2* como o operador linear $P : E \to E$ definido por $Pv = u_1$, onde $v = u_1 + u_2$, com $u_1 \in F_1$ e $u_2 \in F_2$. Vemos imediatamente que $Im(P) = F_1$ e $N(P) = F_2$.

Um operador linear $A : E \to E$ é *idempotente* quando $A^2 = A$.

Teorema 4.2. *Um operador linear $P : E \to E$ é a projeção de E sobre $Im(P)$ paralelamente a $N(P)$ se, e somente se, $P^2 = P$.*

Demonstração. ($\Rightarrow$) Dado $w \in E$, existem únicos $u \in Im(P)$ e $v \in N(P)$ tais que $w = u + v$ e $Pw = u$. Logo, $P^2 w = P(u + 0) = u = Pw$. Portanto $P^2 = P$. ($\Leftarrow$) Dado $v \in E$, temos que $v = Pv + (v - Pv)$. Como $P(v - Pv) = Pv - P^2 v = 0$, temos que $v - Pv \in N(P)$. Se $v \in Im(P) \cap N(P)$, existe $u \in E$ tal que $Pu = v$ e $Pv = 0$, ou seja, $P^2 u = Pu = v = 0$. Logo, $E = Im(P) \oplus N(P)$. Isso implica que, dado $w \in E$, existem únicos $z \in Im(P)$ e $v \in N(P)$ tais que $w = z + v$. Como existe $u \in E$ tal que $Pu = z$, segue que $Pw = Pz = P^2 u = Pu = z$. Portanto, P é uma projeção de E sobre $Im(P)$ paralelamente a $N(P)$. □

Uma *involução* é um operador linear $S : E \to E$ tal que $S^2 v = v$ para todo $v \in E$. Por exemplo, o operador $S : \mathbb{R}^2 \to \mathbb{R}^2$ definido por $S(1, 0) = (1, 0)$ e $S(0, 1) = (0, -1)$ é uma involução que reflete todo vetor do plano em torno do eixo das abscissas.

Teorema 4.3. *Seja $S : E \to E$ uma involução. Os conjuntos $F_1 = \{v \in E : Sv = v\}$ e $F_2 = \{v \in E : Sv = -v\}$ são subespaços tais que $E = F_1 \oplus F_2$. Para todo $v = u_1 + u_2$, com $u_1 \in F_1$ e $u_2 \in F_2$, tem-se que $Sv = u_1 - u_2$. Além disso, se $I : E \to E$ é o operador identidade, então $P = (S + I)/2$ é a projeção de E sobre F_1 paralelamente a F_2.*

Demonstração. F_1 e F_2 são claramente subespaços. Dado $v \in E$, temos que $v = (v + Sv)/2 + (v - Sv)/2$, onde $(v + Sv)/2 \in F_1$ e $(v - Sv)/2 \in F_2$. Além disso, $v \in F_1 \cap F_2$ implica que $Sv = v$ e $Sv = -v$. Logo, $v = 0$ e, por conseguinte, $E = F_1 \oplus F_2$. Dado $v \in E$, existem então únicos $u_1 \in F_1$ e $u_2 \in F_2$ tais que $v = u_1 + u_2$. Logo, $Sv = u_1 - u_2$. Finalmente, se $P = (S + I)/2$, então $P^2 = (S^2 + 2S + I)/4 = (S + I)/2$ e, como podemos ver facilmente, $Im(P) = F_1$ e $N(P) = F_2$. Portanto, P é a projeção de E sobre F_1 paralelamente a F_2. □

5 A matriz de uma transformação linear

Dados os espaços vetoriais E e F de dimensão finita, seja $A : E \to F$ uma transformação linear. Se $\mathcal{V} = \{u_1, \ldots, u_n\}$ e $\mathcal{W} = \{v_1, \ldots, v_m\}$ são bases de E e de F respectivamente, o vetor Au_j pode ser expressado como uma combinação linear de vetores de $\mathcal{W}$. Dessa maneira,

$$Au_j = \sum_{i=1}^{n} a_{ij} v_i \,, \quad j = 1, \ldots, n \tag{4}$$

Os coeficientes a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, podem ser interpretados como os termos de uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$, chamada da *matriz da transformação linear* A. Vemos que, fixadas as bases $\mathcal{V}$ e $\mathcal{W}$, a cada transformação linear $A : E \to F$ lhe corresponde uma única matriz $\mathbf{a} \in \mathcal{M}(m \times n)$. Por outro lado, dada uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$, a equação (4) define uma única transformação linear $A : E \to F$, em virtude do teorema 3.1, desde que as bases $\mathcal{V}$ e $\mathcal{W}$ estejam fixadas. Mais precisamente, podemos provar facilmente que a transformação linear $\varphi : \mathcal{L}(E, F) \to \mathcal{M}(m \times n)$ definida por $\varphi(A) = \mathbf{a}$ é um isomorfismo. Em particular, segue disso que $\dim \mathcal{L}(E, F) = mn$.

A matriz de uma transformação linear $A : \mathbb{R}^n \to \mathbb{R}^m$ é entendida como sendo a matriz de A em relação às bases canônicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, a menos que se indique o contrário. Por exemplo, a matriz do operador linear $R : \mathbb{R}^2 \to \mathbb{R}^2$ definido por $R(1, 0) = (\cos\theta, \operatorname{sen}\theta)$ e $R(0, 1) = (-\operatorname{sen}\theta, \cos\theta)$ é

$$\mathbf{r} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} .$$

Teorema 5.1. *Dados os espaços vetoriais E, F e G de dimensão finita, sejam $A : E \to F$ e $B : F \to G$ transformações lineares. Fixadas as bases $\mathcal{U} = \{u_1, \ldots, u_n\}$, $\mathcal{V} = \{v_1, \ldots, v_p\}$ e $\mathcal{W} = \{w_1, \ldots, w_m\}$ de E, F e G, sejam $\mathbf{a}$ e $\mathbf{b}$ as matrizes de A e de B respectivamente. Se $\mathbf{c}$ é a matriz de BA, então $c_{ij} = \sum_{k=1}^{p} b_{ik} a_{kj}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. A matriz $\mathbf{c}$ é chamada do* produto das matrizes *$\mathbf{b}$ e $\mathbf{a}$ e é denotada por $\mathbf{c} = \mathbf{ba}$.*

Demonstração. Temos que

$$\begin{aligned} Au_j &= \sum_{k=1}^{p} a_{kj} v_k \,, \quad j = 1, \ldots, n \,, \\ Bv_k &= \sum_{i=1}^{m} b_{ik} w_i \,, \quad k = 1, \ldots, p \,, \\ \text{e} \quad ABu_j &= \sum_{i=1}^{m} c_{ij} w_i \,, \quad j = 1, \ldots, n \,. \end{aligned}$$

Dos dois primeiros grupos de equações temos que

$$BAu_j = \sum_{k=1}^{p} a_{kj} Bv_k = \sum_{k=1}^{p} \sum_{i=1}^{m} a_{kj} b_{ik} w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{p} b_{ik} a_{kj} \right) w_i \,.$$

Logo, usando o terceiro grupo de equações, temos que

$$\sum_{i=1}^{m} c_{ij} w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{p} b_{ik} a_{kj} \right) w_i \,.$$

Como $\mathcal{W}$ é um conjunto L.I., segue que $c_{ij} = \sum_{k=1}^{p} b_{ik} a_{kj}$. □

As seguintes propriedades de matrizes são consequências diretas das propriedades sobre transformações lineares:

1. Dados $\mathbf{a} \in \mathcal{M}(m \times n)$, $\mathbf{b} \in \mathcal{M}(n \times p)$ e $\mathbf{c} \in \mathcal{M}(p \times q)$, tem-se que $\mathbf{a}(\mathbf{bc}) = (\mathbf{ab})\mathbf{c}$.

2. Dados $\mathbf{a} \in \mathcal{M}(m \times n)$ e $\mathbf{b}, \mathbf{c} \in \mathcal{M}(n \times p)$, tem-se que $\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac}$.

3. Dados $\mathbf{a}, \mathbf{b} \in \mathcal{M}(m \times n)$ e $\mathbf{c} \in \mathcal{M}(n \times p)$, tem-se que $(\mathbf{a} + \mathbf{b})\mathbf{c} = \mathbf{ac} + \mathbf{bc}$.

4. Os termos da *matriz identidade* $\mathbf{I}_n \in \mathcal{M}(n \times n)$ são $I_{ij} = \delta_{ij}$, $i, j \in \{1, \dots, n\}$, onde δ_{ij} é o *símbolo de Kronecker*, definido por $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$.

5. Dado $\mathbf{a} \in \mathcal{M}(m \times n)$, tem-se que $\mathbf{a}\mathbf{I}_n = \mathbf{a}$ e $\mathbf{I}_m\mathbf{a} = \mathbf{a}$.

6. Uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$ tem uma *inversa à esquerda* se existe $\mathbf{b} \in \mathcal{M}(n \times m)$ tal que $\mathbf{ba} = \mathbf{I}_n$. A matriz $\mathbf{a}$ tem uma inversa à esquerda se, e somente se, os vetores-coluna de $\mathbf{a}$ são L.I.

7. Uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$ tem uma *inversa à direita* se existe $\mathbf{b} \in \mathcal{M}(n \times m)$ tal que $\mathbf{ab} = \mathbf{I}_m$. A matriz $\mathbf{a}$ tem uma inversa à direita se, e somente se, os vetores-coluna de $\mathbf{a}$ geram $\mathbb{R}^m$.

8. Uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$ é *invertível* se ela tem uma inversa à esquerda $\mathbf{b} \in \mathcal{M}(n \times m)$ e uma inversa à direita $\mathbf{c} \in \mathcal{M}(n \times m)$. Nesse caso tem-se que $\mathbf{b} = \mathbf{c} = \mathbf{a}^{-1}$, chamada da *inversa* de $\mathbf{a}$, e que $m = n$, ou seja, a matriz $\mathbf{a}$ é *quadrada*.

9. Uma matriz $\mathbf{a} \in \mathcal{M}(n \times n)$ possui uma inversa à esquerda se, e somente se, possui uma inversa à direita. Nesse caso, ambas inversas são iguais a $\mathbf{a}^{-1}$.

Sejam $\mathcal{V} = \{v_1, \dots, v_n\}$ e $\mathcal{V}' = \{v'_1, \dots, v'_n\}$ bases de um espaço vetorial E e $\mathcal{W} = \{w_1, \dots, w_m\}$ e $\mathcal{W}' = \{w'_1, \dots, w'_m\}$ bases de um espaço vetorial F. Sejam também $\mathbf{a}$ e $\mathbf{a}'$ as matrizes de uma transformação linear $A : E \to F$ quando se consideram os pares de bases $(\mathcal{V}, \mathcal{W})$ e $(\mathcal{V}', \mathcal{W}')$ respectivamente. Temos então que $Av_j = \sum_{i=1}^m a_{ij} w_i$ e $Av'_j = \sum_{i=1}^m a'_{ij} w'_i$, $j = 1, \dots, n$. Considerando os isomorfismos $P : E \to E$ e $Q : F \to F$ definidos por $Pv_j = v'_j$, $j = 1, \dots, n$ e $Qw_i = w'_i$, $i = 1, \dots, m$, temos que

$$\sum_{k=1}^m (ap)_{kj} w_k = APv_j = Av'_j = \sum_{i=1}^m a'_{ij} w'_i = \sum_{i=1}^m \sum_{k=1}^m a'_{ij} q_{ki} w_k \,,$$

o que implica que $(ap)_{kj} = \sum_{i=1}^m a'_{ij} q_{ki} = (qa)_{kj}$, ou seja $\mathbf{ap} = \mathbf{qa}'$. Portanto, $\mathbf{a}' = \mathbf{q}^{-1}\mathbf{ap}$. A matriz $\mathbf{p}$ é chamada de uma *matriz de passagem* da base $\mathcal{V}$ para a base $\mathcal{V}'$. Vemos que as coordenadas do j-ésimo vetor-coluna de $\mathbf{p}$ são as coordenadas do vetor v'_j na base $\mathcal{V}$. Observações análogas valem para a matriz de passagem $\mathbf{q}$.

Seja $A : \mathbb{R}^2 \to \mathbb{R}^2$ o operador linear definido por $A(1,1) = (1,1)$ e $A(-1,1) = (0,0)$. Logo, a matriz de A na base $\mathcal{B} = \{(1,1), (-1,1)\}$ é

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

Podemos verificar facilmente que $A(1,0) = (1/2, 1/2)$ e $A(0,1) = (1/2, 1/2)$. Logo, a matriz de A na base canônica é

$$\mathbf{a}' = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

A matriz $\mathbf{a}'$ pode ser obtida mais dificilmente usando a matriz de passagem

$$\mathbf{p} = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

da base $\mathcal{B}$ para a base canônica e a relação $\mathbf{a}' = \mathbf{p}^{-1}\mathbf{a}\mathbf{p}$.

Sejam E e F espaços vetoriais de dimensão finita. O *posto* de uma transformação linear $A : E \to F$ é a dimensão da sua imagem. Se $\mathbf{a}$ é a matriz de A, então o posto de A é a dimensão do subespaço gerado pelos vetores-coluna de $\mathbf{a}$.

Dada uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$, define-se o *posto segundo colunas* de $\mathbf{a}$ como a dimensão do subespaço de $\mathbb{R}^m$ gerado pelos vetores-coluna de $\mathbf{a}$. Define-se também o *posto segundo linhas* de $\mathbf{a}$ como a dimensão do subespaço de $\mathbb{R}^n$ gerado pelos vetores-linha de $\mathbf{a}$. Embora o subespaço gerado pelos vetores-coluna de $\mathbf{a}$ e o gerado pelos vetores-linha de $\mathbf{a}$ sejam em geral diferentes, tem-se o seguinte:

Teorema 5.2. *O posto segundo linhas de uma matriz* $\mathbf{a} \in \mathcal{M}(m \times n)$ *é igual a seu posto segundo colunas.*

Demonstração. Seja $\{u_1, \ldots, u_r\} \subset \mathbb{R}^m$ uma base do subespaço gerado pelos vetores-coluna $v_1, \ldots, v_n$ de $\mathbf{a}$. Logo, $v_j = \sum_{k=1}^r \alpha_{jk} u_k$, $j = 1, \ldots, n$. Considerando que $u_k = (u_{k1}, \ldots, u_{km})$, $k = 1, \ldots, r$, temos que $a_{ij} = \sum_{k=1}^r \alpha_{jk} u_{ki}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Logo, o i-ésimo vetor-linha de $\mathbf{a}$ é

$$w_i = \left(\sum_{k=1}^r \alpha_{1k} u_{ki}, \ldots, \sum_{k=1}^r \alpha_{nk} u_{ki} \right) = \sum_{k=1}^r u_{ki} z_k \,,$$

onde $z_k = (\alpha_{1k}, \ldots, \alpha_{nk})$, $k = 1, \ldots, r$. Isso implica que os vetores $z_1, \ldots, z_r$ e $w_1, \ldots, w_m$ geram o mesmo subespaço de $\mathbb{R}^n$. Dessa maneira, o posto segundo linhas de $\mathbf{a}$ é $\leq r$. Fazendo um procedimento análogo começando com uma base do subespaço gerado pelas colunas de $\mathbf{a}$, vamos concluir que o posto segundo colunas de $\mathbf{a}$ não excede o posto segundo linhas de $\mathbf{a}$. Portanto, ambos postos devem ser iguais. $\square$

Dado um espaço vetorial E de dimensão n, sejam $f_1, \ldots, f_m : E \to \mathbb{R}$ funcionais lineares não-nulos. Queremos determinar a dimensão do subespaço $F = N(f_1) \cap \ldots \cap N(f_m)$. Se $\{u_1, \ldots, u_n\}$ é uma base de E, então $N(f_i) = \{(x_1, \ldots, x_n) \in E : \sum_{j=1}^n f_i(u_j) x_j = 0\}$, $i = 1, \ldots, m$. Segue imediatamente disso que F é isomorfo ao núcleo da transformação linear $A : \mathbb{R}^n \to \mathbb{R}^m$ cuja matriz é

$$\mathbf{a} = \begin{bmatrix} f_1(u_1) & \ldots & f_1(u_n) \\ \vdots & & \vdots \\ f_m(u_1) & \ldots & f_m(u_n) \end{bmatrix}.$$

Dessa maneira, pelo teorema do núcleo e da imagem, $\dim F = \dim N(A) = n - \dim Im(A)$. Por outro lado, o subespaço gerado pelos vetores-linha de $\mathbf{a}$ é claramente isomorfo ao subespaço de E^* gerado pelos funcionais $f_1, \ldots, f_m$. Se r é a dimensão desse subespaço, pelo teorema 5.2, temos que $\dim Im(A) = r$. Portanto, $\dim F = n - r$.

Seja E um espaço vetorial de dimensão n. Vamos mostrar um método prático para determinar uma base do subespaço gerado por um conjunto $\{v_1, \ldots, v_m\} \subset E$. Esse método está baseado nos seguintes fatos de verificação imediata:

1. para quaisquer $i, j \in \{1, \ldots, n\}$,

$$S(\{v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n\}) = S(\{v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n\})\,;$$

2. para quaisquer $\alpha \in \mathbb{R}$ e $i, j \in \{1, \ldots, n\}$,

$$S(\{v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n\}) = S(\{v_1, \ldots, v_i + \alpha v_j, \ldots, v_j, \ldots, v_n\})\,.$$

Prosseguimos agora com a descrição do método:

1. Construímos uma matriz $\mathbf{a}$ cujos vetores-linha são os vetores $v_1, \ldots, v_m$;

2. Caso seja necessário, trocamos convenientemente as linhas da matriz $\mathbf{a}$ de tal forma que obtemos uma matriz $\mathbf{a}'$ cuja primeira coluna não-nula tem o primeiro termo diferente de zero. Por exemplo,

$$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 3 & 0 & 2 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 2 \end{bmatrix}.$$

3. Para $i = 2, \ldots, m$, adicionamos à i-ésima linha de $\mathbf{a}'$ um múltiplo da primeira linha de $\mathbf{a}'$ de tal forma que obtemos uma matriz $\mathbf{a}''$ cuja primeira coluna não-nula tem o i-ésimo termo igual a zero. Por exemplo,

$$\begin{bmatrix} 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 2 \end{bmatrix} \xrightarrow{L_3 - 3L_1} \begin{bmatrix} 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -7 & -1 \end{bmatrix}$$

4. Obtendo uma matriz $\mathbf{a}''$, desconsideramos a primeira linha e a primeira coluna não-nula. Dessa maneira, a segunda coluna não-nula de $\mathbf{a}''$ é considerada como a primeira coluna não-nula e repetimos os passos 2 e 3. Por exemplo,

$$\begin{bmatrix} 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -7 & -1 \end{bmatrix} \xrightarrow{L_3 + (7/2)L_2} \begin{bmatrix} 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

5. O processo para quando obtemos uma *matriz escalonada*, que é uma matriz tal que o primeiro termo não-nulo da i-esima linha está à esquerda do primeiro termo não-nulo das linhas subsequentes e as linhas nulas, caso existam, se encontram abaixo das linhas não nulas. Por exemplo,

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

é uma matriz escalonada.

6. Os vetores-linha não-nulos da matriz escalonada obtida formam uma base de $S(\{v_1, \ldots, v_m\})$, pois são claramente L.I..

Os passos 2, 3, 4 e 5 formam o chamado *processo de eliminação* ou *método de Gauss*. As modificações realizadas nos passos 2 e 3 são chamadas de *operações elementares*.

A *transposta* de uma matriz $\mathbf{a} \in \mathcal{M}(m \times n)$ é a matriz $\mathbf{a}^T \in \mathcal{M}(n \times m)$ tal que $a^T_{ji} = a_{ij}$, $j = 1, \ldots, n$, $i = 1, \ldots, m$. Dada uma transformação linear $A : E \to F$ entre espaços vetoriais de dimensão finita, o processo de eliminação pode ser aplicado para encontrarmos uma base de $Im(A)$. Nesse caso, se $\mathbf{a}$ é a matriz de A, devemos aplicar o processo de eliminação à matriz $\mathbf{a}^T$. Logo, os vetores-linha da matriz escalonada obtida formarão uma base de $Im(A)$. Se só estamos interessados no posto de A, podemos aplicar o processo de eliminação diretamente à matriz $\mathbf{a}$, em virtude do teorema 5.2. Nesse caso, o posto de A será o número de linhas não-nulas da matriz escalonada obtida.

Um sistema linear

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad &\quad \vdots \\ a_{mn}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{5}$$

pode ser escrito matricialmente como $\mathbf{ax} = \mathbf{b}$, onde

$$\mathbf{a} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Esse sistema tem solução se, e somente se, o vetor $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ pertence à imagem da transformação linear $A : \mathbb{R}^n \to \mathbb{R}^m$, cuja matriz é $\mathbf{a}$. Logo, o sistema (5) tem solução se, e somente se, o subespaço gerado por b e os vetores-coluna de $\mathbf{a}$ é igual a $Im(A)$. Além disso, nesse caso, a solução é única se, e somente se, A é injetiva (ver teorema 3.8).

Definindo a matriz aumentada

$$[\mathbf{a}, \mathbf{b}] = \begin{bmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{bmatrix},$$

temos o seguinte:

1. O sistema (5) não tem solução se, e somente se, o posto de $[\mathbf{a}, \mathbf{b}]$ é maior do que o posto de $\mathbf{a}$, ou seja, ao aplicarmos o processo de eliminação a $[\mathbf{a}, \mathbf{b}]$, o número de linhas não-nulas da matriz escalonada obtida $[\mathbf{a}', \mathbf{b}']$ é maior do que o número de linhas não-nulas da matriz escalonada $\mathbf{a}'$.

2. O sistema (5) tem uma única solução se, e somente se, o posto de $[\mathbf{a}, \mathbf{b}]$ e o posto de $\mathbf{a}$ são iguais a n, ou seja, ao aplicarmos o processo de eliminação a $[\mathbf{a}, \mathbf{b}]$, o número de linhas não-nulas da matriz escalonada obtida $[\mathbf{a}', \mathbf{b}']$ e o número de linhas não-nulas da matriz escalonada $\mathbf{a}'$ são iguais a n. Em particular, para isso acontecer é necessário ter $m \geq n$.

3. O sistema (5) tem infinitas soluções se, e somente se, o posto de $[\mathbf{a}, \mathbf{b}]$ e o posto de $\mathbf{a}$ são iguais a $r < n$, ou seja, ao aplicarmos o processo de eliminação a $[\mathbf{a}, \mathbf{b}]$, o número de linhas não-nulas da matriz escalonada obtida $[\mathbf{a}', \mathbf{b}']$ e o número de linhas não-nulas da matriz escalonada $\mathbf{a}'$ são iguais a $r < n$.

Em termos das equações do sistema (5), as operações elementares efetuadas sobre $[\mathbf{a}, \mathbf{b}]$ para obter $[\mathbf{a}', \mathbf{b}']$ se traduzem em

- alterar a ordem das equações;
- somar um múltiplo de uma equação a outra equação.

Dessa maneira, se o sistema (5) tem solução, a matriz aumentada $[\mathbf{a}', \mathbf{b}']$ corresponde a um sistema linear $\mathbf{a}'\mathbf{x} = \mathbf{b}'$ que tem as mesmas soluções que o sistema (5). No entanto, como $[\mathbf{a}', \mathbf{b}']$ é escalonada, o sistema $\mathbf{a}'\mathbf{x} = \mathbf{b}'$ pode ser resolvido facilmente de baixo para cima.

Vamos resolver o sistema linear

$$\begin{aligned} x + y + z &= 3 \\ x - y + 2z &= 5 \\ 3x + 2y &= 5\,. \end{aligned} \tag{6}$$

A matriz aumentada associada a esse sistema é

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 2 & 5 \\ 3 & 2 & 0 & 5 \end{bmatrix}.$$

Logo,

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 2 & 5 \\ 3 & 2 & 0 & 5 \end{bmatrix} \xrightarrow{L_2 - L_1, L_3 - 3L_1} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 1 & 2 \\ 0 & -1 & -3 & -4 \end{bmatrix} \xrightarrow{L_3 - L_2/2} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 1 & 2 \\ 0 & 0 & -7/2 & -5 \end{bmatrix}$$

Vemos que a única solução do sistema (6) é solução do sistema linear

$$\begin{aligned} x + y + z &= 3 \\ -2y + z &= 2 \\ -7z/2 &= 5\,. \end{aligned}$$

De onde obtemos que $z = -10/7$, $y = -12/7$ e $x = 43/7$.

O processo de eliminação pode ser estendido considerando mais uma operação elementar que consiste em multiplicar uma linha por um número real diferente de zero. Esse procedimento é conhecido como *método de Gauss-Jordan*, cuja aplicação fornece um método prático para calcular a inversa de uma matriz quadrada $\mathbf{a} \in \mathcal{M}(n \times n)$, a qual existe se, e somente se, o posto de $\mathbf{a}$ é igual a n. De fato, encontrar a inversa de $\mathbf{a}$ equivale a resolver n sistemas de n equações lineares, os quais podem ser escritos de forma compacta como $\mathbf{a}\mathbf{a}^{-1} = \mathbf{I}_n$, onde os termos de $\mathbf{a}^{-1}$ são as incógnitas. O método prático para encontrar $\mathbf{a}^{-1}$ é considerar a matriz aumentada $[\mathbf{a}, \mathbf{I}_n]$ e aplicar o método de Gauss-Jordan com o objetivo de obtermos a matriz $[\mathbf{I}_n, \mathbf{b}]$. Dessa maneira, a matriz $\mathbf{a}^{-1}$, deve satisfazer a condição $\mathbf{I}_n\mathbf{a}^{-1} = \mathbf{b}$, ou seja, $\mathbf{a}^{-1} = \mathbf{b}$.

Vamos encontrar a inversa da matriz

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

Temos que

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{L_2-L_1, L_3-3L_1} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 & 1 & 0 \\ 0 & -1 & -3 & -3 & 0 & 1 \end{bmatrix}$$
$$\xrightarrow{L_3-L_2/2} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -7/2 & -5/2 & -1/2 & 1 \end{bmatrix}$$
$$\xrightarrow{(-1/2)L_2, (-2/7)L_3} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1/2 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1 & 5/7 & 1/7 & -2/7 \end{bmatrix}$$
$$\xrightarrow{L_1-L_2} \begin{bmatrix} 1 & 0 & 3/2 & 1/2 & 1/2 & 0 \\ 0 & 1 & -1/2 & 1/2 & -1/2 & 0 \\ 0 & 0 & 1 & 5/7 & 1/7 & -2/7 \end{bmatrix}$$
$$\xrightarrow{L_1-3L_3/2, L_2+L_3/2} \begin{bmatrix} 1 & 0 & 0 & -4/7 & 2/7 & 3/7 \\ 0 & 1 & 0 & 6/7 & -3/7 & -1/7 \\ 0 & 0 & 1 & 5/7 & 1/7 & -2/7 \end{bmatrix}.$$

Portanto,

$$\mathbf{a}^{-1} = \begin{bmatrix} -4/7 & 2/7 & 3/7 \\ 6/7 & -3/7 & -1/7 \\ 5/7 & 1/7 & -2/7 \end{bmatrix}.$$

6 Produto interno

Um *produto interno* em um espaço vetorial E é uma função $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ que tem as seguintes propriedades:

1. dados $u, v, w \in E$, tem-se que $\langle u, v+w \rangle = \langle u, v \rangle + \langle u, w \rangle$ e $\langle u+w, v \rangle = \langle u, v \rangle + \langle w, v \rangle$;
2. dados $u, v \in E$ e $\alpha \in \mathbb{R}$, tem-se que $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle = \langle u, \alpha v \rangle$;
3. para quaisquer $u, v \in E$, $\langle u, v \rangle = \langle v, u \rangle$;
4. para qualquer $v \in E$ com $v \neq 0$, tem-se que $\langle v, v \rangle > 0$.

Teorema 6.1. *Seja E um espaço vetorial munido de um produto interno. Logo,*

1. *para qualquer $v \in E$, $\langle v, 0 \rangle = 0$;*
2. *se $u \in E$ é tal que $\langle u, v \rangle = 0$ para todo $v \in E$, então $u = 0$;*
3. *se $u, v \in E$ são tais que $\langle u, w \rangle = \langle v, w \rangle$, então $u = v$.*

Demonstração.

1. $\langle v, 0 \rangle = \langle v, 0 + 0 \rangle = 2\langle v, 0 \rangle$. Logo, $\langle v, 0 \rangle = 0$.
2. se fosse $u \neq 0$, teríamos que $\langle u, v \rangle > 0$ quando $v = u$. Contradição! Logo, devemos ter $u = 0$.
3. Segue diretamente do item 2. □

Seja E um espaço vetorial munido de um produto interno. A *norma* de um vetor $v \in E$ é definida por $|v| = \sqrt{\langle v, v\rangle}$. Se $v \in E$ é tal que $|v| = 1$, diz-se que v é um *vetor unitário*.

No espaço euclidiano $\mathbb{R}^n$, define-se o *produto interno canônico* de dois vetores $u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n) \in \mathbb{R}^n$, expressos na base canônica de $\mathbb{R}^n$, por $\langle u, v\rangle = x_1y_1 + \cdots + x_ny_n$.

Dados $u, v \in \mathbb{R}^2$, vamos mostrar que $\langle u, v\rangle = |u||v|\cos\theta$, onde θ é o ângulo formado pelos vetores u e v. Se um dos vetores é igual a 0, a igualdade é verdadeira. Consideremos então que $u \neq 0$ e $v \neq 0$. Logo, temos que $u = |u|(\cos\alpha, \sin\alpha)$ e $v = |v|(\operatorname{sen}\beta, \cos\beta)$, onde α e β são os ângulos que os vetores u e v formam com o vetor $(1, 0)$ respectivamente. Logo $\langle u, v\rangle = |u||v|(\cos\alpha\cos\beta + \sin\alpha\sin\beta) = |u||v|\cos(\beta - \alpha) = |u||v|\cos\theta$.

A função $\langle\cdot,\cdot\rangle : C^0([a, b]) \times C^0([a, b]) \to \mathbb{R}$ definida por $\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$, é um produto interno no espaço vetorial $C^0([a, b])$. Com efeito, só temos que provar que, dada a função $f \in C^0([a, b])$ não-nula, $\langle f, f\rangle > 0$, pois as condições restantes que caracterizam um produto interno são claramente satisfeitas. Como $f^2 \in C^0([a, b])$ e $[f(x)]^2 > 0$ para todo $x \in [a, b]$, existe $c \in [a, b]$ tal que $[f(x)]^2 \geq [f(c)]^2 > 0$. Logo, $\langle f, f\rangle \geq [f(c)]^2(b - a) > 0$.

Se E é um espaço vetorial de dimensão finita e $\mathcal{B} = \{u_1, \ldots, u_n\}$ é uma base de E, então a função $\langle\cdot,\cdot\rangle : E \times E \to \mathbb{R}$ definida por $\langle v, w\rangle = \alpha_1\beta_1 + \cdots + \alpha_n\beta_n$, onde $v = \alpha_1u_1 + \cdots + \alpha_nu_n$ e $w = \beta_1u_1 + \cdots \beta_nu_n$, é claramente um produto interno. Logo, todo espaço vetorial de dimensão finita pode ser considerado como um espaço vetorial munido de um produto interno.

Seja E um espaço vetorial munido de um produto interno. Diz-se que dois vetores $u, v \in E$ são *ortogonais*, e escreve-se $u \perp v$, quando $\langle u, v\rangle = 0$. Um conjunto $X \subset E$ é dito *ortogonal* se qualquer par de vetores de X são ortogonais. Se cada vetor de X é unitário, diz-se ainda que X é um *conjunto ortonormal*. Nesse caso, se X é uma base de E, diz-se que X é uma *base ortonormal*.

Teorema 6.2 (Teorema de Pitágoras)**.** *Seja E um espaço vetorial munido de um produto interno. Se $u, v \in E$ são tais que $u \perp v$, então $|u + v|^2 = |u|^2 + |v|^2$.*

Teorema 6.3. *Seja E um espaço vetorial munido de um produto interno. Se $X \subset E$ é um conjunto ortogonal que não contém o vetor nulo, então é L.I..*

Demonstração. Sejam $v_1, , \ldots, v_n \in X$ tais que $\alpha_1v_1 + \cdots + \alpha_nv_n = 0$. Logo, $\langle v_i, \alpha_1v_1 + \cdots + \alpha_nv_n\rangle = \alpha_i|v_i|^2 = 0$, $i = 1, \ldots, n$, o que implica que $\alpha_i = 0$. Portanto, X é L.I.. $\square$

Seja E um espaço vetorial munido de um produto interno. Dados $u, v \in E$ com $u \neq 0$, define-se a *projeção ortogonal de v sobre o eixo que contém u* por

$$\operatorname{pr}_u(v) = \frac{\langle v, u\rangle}{|u|^2}u\,.$$

O nome é devido a que o vetor $w = v - \operatorname{pr}_u(v)$ é tal que $\langle w, u\rangle = 0$, ou seja, $w \perp u$.

Teorema 6.4 (Desigualdade de Schwarz)**.** *Seja E um espaço vetorial munido de um produto interno. Se $u, v \in E$, então $|\langle u, v\rangle| \leq |u||v|$. Além disso, a igualdade ocorre se, e somente se, v é um múltiplo de u.*

Demonstração. Se $u = 0$, OK. Se $u \neq 0$, o vetor $w = v - \operatorname{pr}_u(v)$ é ortogonal a u. Logo, pelo teorema de Pitágoras, $|v|^2 = |\operatorname{pr}_u(v)|^2 + |w|^2$. Logo, $|v|^2 \geq |\operatorname{pr}_u(v)|^2$, o que implica que $|u||v| \geq |\langle u, v\rangle|$. Além disso, vemos que a igualdade acontece se, e somente se, $w = 0$, ou seja, quando v é um múltiplo de u. $\square$

Corolário 6.5 (Desigualdade triangular). *Seja E um espaço vetorial munido de um produto interno. Se $u, v \in E$, então $|u+v| \leq |u| + |v|$, onde a igualdade ocorre se, e somente se, existe $\alpha \geq 0$ tal que $v = \alpha u$.*

Seja E um espaço métrico munido de um produto interno. Define-se a *distância* entre dois vetores $u, v \in E$ por $d(u,v) = |u - v|$. Vemos que $d(u,v) > 0$ se $u \neq v$, $d(u,v) = d(v,u)$ e $d(u,v) \leq d(u,w) + d(w,v)$ para todo $w \in E$.

Seja E um espaço vetorial de dimensão finita munido de um produto interno. Vamos descrever o *processo de ortonormalização de Gram-Schmidt* para obter uma base ortonormal $\{u_1, \ldots, u_n\} \subset E$ a partir de uma base qualquer $\{v_1, \ldots, v_n\} \subset E$ tal que $u_i \in S(\{v_1, \ldots, v_i\})$, $i = 1, \ldots, n$. Primeiramente pomos $w_1 = v_1$. Supondo definidos $w_1, \ldots, w_m \in E$, ortogonais dois a dois, definimos

$$w_{m+1} = v_{m+1} - \sum_{i=1}^{m} \mathrm{pr}_{w_i}(v_{m+1})\,.$$

Vemos que $\langle w_j, w_{m+1} \rangle = \langle w_j, v_{m+1} \rangle - \langle w_j, \mathrm{pr}_{w_j}(v_{m+1}) \rangle = 0$, $j = 1, \ldots, m$. Além disso, $w_{m+1} \neq 0$, pois $v_{m+1} \notin S(\{w_1, \ldots, w_m\}) = S(\{v_1, \ldots, v_m\})$. Dessa maneira, obtemos um conjunto ortogonal $\{w_1, \ldots, w_n\} \subset E$. Finalmente, pondo $u_i = w_i/|w_i|$, $i = 1, \ldots, n$, obtemos uma base ortonormal $\{u_1, \ldots, u_n\} \subset E$. Observamos que, se $v_1, \ldots, v_r$, $r < n$, são vetores ortonormais, então a base ortonormal $\{u_1, \ldots, u_n\}$, obtida pelo processo de Gram-Schmidt, é tal que $u_1 = v_1, \ldots, u_r = v_r$. Portanto, toda base ortonormal de um subespaço $F \subset E$ pode ser estendida a uma base ortonormal de E.

Seja E um espaço vetorial munido de um produto interno. Se $\{u_1, \ldots, u_m\}$ é uma base ortonormal de um subespaço $F \subset E$, define-se a *projeção ortogonal de um vetor $v \in E$ sobre F* por

$$\mathrm{pr}_F(v) = \sum_{i=1}^{m} \mathrm{pr}_{u_i}(v)\,.$$

O vetor $\mathrm{pr}_F(v)$ não depende da escolha da base ortonormal de F. Com efeito, se $\{w_1, \ldots, w_m\}$ é outra base ortonormal de F, então

$$\sum_{i=1}^{m} \mathrm{pr}_{w_i}(v) = \sum_{i=1}^{m} \langle v, w_i \rangle w_i = \sum_{i=1}^{m} \sum_{j=1}^{m} \langle v, w_i \rangle \langle w_i, u_j \rangle u_j\,,$$

pois, para qualquer $z \in F$, $z = \sum_{j=1}^{m} \langle z, u_j \rangle u_j$. Logo,

$$\sum_{i=1}^{m} \mathrm{pr}_{w_i}(v) = \sum_{j=1}^{m} \left\langle v, \sum_{i=1}^{m} \langle u_j, w_i \rangle w_i \right\rangle u_j = \sum_{j=1}^{m} \langle v, u_j \rangle u_j = \mathrm{pr}_F(v)\,.$$

O vetor $v - \mathrm{pr}_F(v)$ é ortogonal a todo vetor da base ortonormal $\{u_1, \ldots, u_n\}$ e, por conseguinte, é ortogonal a todo vetor de F. Além disso, dado $z \in F$, vemos que $v - z = [v - \mathrm{pr}_F(v)] + [\mathrm{pr}_F(v) - z]$. Como $\mathrm{pr}_F(v) - z \in F$, segue que $v - \mathrm{pr}_F(v) \perp \mathrm{pr}_F(v) - z$. Logo, pelo teorema de Pitágoras, temos que $|v - z|^2 = |v - \mathrm{pr}_F(v)|^2 + |\,\mathrm{pr}_F(v) - z|^2$ e, por conseguinte, $|v - z| \geq |v - \mathrm{pr}_F(v)|$. Portanto, $\mathrm{pr}_F(v)$ é o vetor de F que está a menor distância de v.

Seja $\mathcal{B} = \{u_1, \ldots, u_n\}$ uma base ortonormal de um espaço vetorial E, munido de um produto interno. Se $v, w \in E$, vemos que

$$\langle v, w \rangle = \left\langle \sum_{i=1}^{n} \langle v, u_i \rangle u_i, \sum_{j=1}^{n} \langle w, u_j \rangle u_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle v, u_i \rangle \langle w, u_j \rangle \langle u_i, u_j \rangle = \sum_{i=1}^{n} \langle v, u_i \rangle \langle u_i, w \rangle\,,$$

que é análogo ao produto interno canônico em $\mathbb{R}^n$. Se $\{v_1, \ldots, v_n\}$ é uma base de E, não necessariamente ortonormal, pode-se definir um produto interno $[\cdot,\cdot] : E \times E \to \mathbb{R}$ pondo

$$[w, z] = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} \alpha_i \beta_j \,,$$

quando $w = \alpha_1 v_1 + \cdots + \alpha_n v_n$, $z = \beta_1 v_1 + \cdots + \beta_n v_n$, onde $\mathbf{g}$ é uma matriz *positiva*, ou seja, $\mathbf{g}^T = \mathbf{g}$ e $\mathbf{x}^T \mathbf{g} \mathbf{x} > 0$ para qualquer matriz $\mathbf{x} \in \mathcal{M}(n \times 1)$ não-nula.

7 A adjunta

Seja E um espaço vetorial munido de um produto interno. Para cada $u \in E$, a função $f : E \to \mathbb{R}$ definida por $f(v) = \langle u, v \rangle$ é claramente um funcional linear.

Teorema 7.1. *Seja E um espaço vetorial de dimensão finita munido de um produto interno. Se $f : E \to \mathbb{R}$ é um funcional linear, existe um único $w \in E$ tal que $f(v) = \langle w, v \rangle$ para todo $v \in E$.*

Demonstração. Seja $\{u_1, \ldots, u_n\}$ uma base ortonormal de E. Dado $v \in E$, temos que $v = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_n \rangle u_n$. Logo,

$$f(v) = \sum_{i=1}^{n} \langle v, u_i \rangle f(u_i) = \left\langle v, \sum_{i=1}^{n} f(u_i) u_i \right\rangle .$$

Dessa maneira, existe $w = f(u_1)u_1 + \cdots + f(u_n)u_n \in E$ tal que $f(v) = \langle w, v \rangle$. Se existe $w' \in E$ tal que $f(v) = \langle w', v \rangle$ para todo $v \in E$, então $\langle w', v \rangle = \langle w, v \rangle$ para todo $v \in E$, o que implica que $w' = w$. □

Sejam E e F espaços vetoriais de dimensão finita munidos de produto interno. A *adjunta* de uma transformação linear $A : E \to F$ é uma aplicação $A^* : F \to E$ tal que $\langle Au, v \rangle = \langle u, A^* v \rangle$ para quaisquer $u \in E$ e $v \in F$. Em virtude do teorema 7.1, para cada $v \in F$, existe um único $w \in F$ tal que o funcional linear $f : E \to F$ definido por $f(u) = \langle Au, v \rangle$ é tal que $f(u) = \langle u, w \rangle$ para todo $u \in E$. Logo, nesse caso, devemos ter $A^* v = w$. Portanto, a adjunta de A existe e, como pode-se verificar facilmente, é única. Além disso, A^* é uma transformação linear, pois, dados $u \in E$, $v, w \in F$ e $\alpha \in \mathbb{R}$, temos que $\langle u, A^*(v + w) \rangle = \langle Au, v + w \rangle = \langle Au, v \rangle + \langle Au, w \rangle = \langle u, A^* v + A^* w \rangle$ e $\langle u, A^*(\alpha v) \rangle = \langle Au, \alpha v \rangle = \alpha \langle Au, v \rangle = \langle u, \alpha A^* v \rangle$.

Teorema 7.2. *Sejam E e F espaços vetoriais de dimensão finita, munidos de produto interno. Se $\mathcal{V} = \{u_1, \ldots, u_n\}$ e $\mathcal{W} = \{v_1, \ldots, v_m\}$ são bases ortonormais de E e F respectivamente e* $\mathbf{a}$ *é a matriz de uma transformação linear $A : E \to F$, então a matriz de A^* é* $\mathbf{a}^T$.

Demonstração. Como $A^* v_j = \sum_{i=1}^{n} a^*_{ij} u_i$, $j = 1, \ldots, m$, vemos que $a^*_{ij} = \langle u_i, A^* v_j \rangle = \langle A u_i, v_j \rangle$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Por outro lado, $A u_i = \sum_{j=1}^{m} a_{ji} v_j$, $i = 1, \ldots, n$, de onde segue que $a_{ji} = \langle v_j, A u_i \rangle = a^*_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Portanto, $\mathbf{a}^* = \mathbf{a}^T$. □

Corolário 7.3. *Seja $A : E \to F$ uma transformação linear entre espaços vetoriais de dimensão finita, munidos de produto interno. Logo, o posto de A^* é igual ao posto de A.*

Teorema 7.4. *Sejam E, F e G espaços vetoriais de dimensão finita, munidos de produto interno. Logo,*

1. *se $I : E \to E$ é o operador identidade, então $I^* = I$;*
2. *dados $A, B \in \mathcal{L}(E, F)$, tem-se que $(A + B)^* = A^* + B^*$;*
3. *dados $A \in \mathcal{L}(E, F)$ e $\alpha \in \mathbb{R}$, tem-se que $(\alpha A)^* = \alpha A^*$;*
4. *dados $A \in \mathcal{L}(E, F)$ e $B \in \mathcal{L}(F, G)$, tem-se que $(BA)^* = A^* B^*$;*
5. *dado $A \in \mathcal{L}(E, F)$, tem-se que $A^{**} = A$;*
6. *$A \in \mathcal{L}(E, F)$ é injetiva se, e somente se, A^* é sobrejetiva;*
7. *$A \in \mathcal{L}(E, F)$ é um isomorfismo se, e somente se, A^* é também um isomorfismo; tem-se ainda que $(A^*)^{-1} = (A^{-1})^*$.*

Demonstração.

6. ($\Rightarrow$) Se A é injetiva, A tem uma inversa à esquerda $B : F \to E$. Logo, $BA = I_E$ e, pelos itens 1 e 4, temos que $A^* B^* = I_E$. Dessa maneira, A^* tem uma inversa à direita, o que implica que é sobrejetiva. ($\Leftarrow$) Análogo. $\square$

Corolário 7.5.

1. $\mathbf{I}^T = \mathbf{I}$;
2. *dados $\mathbf{a}, \mathbf{b} \in \mathcal{M}(m \times n)$, tem-se que $(\mathbf{a} + \mathbf{b})^T = \mathbf{a}^T + \mathbf{b}^T$;*
3. *dados $\mathbf{a} \in \mathcal{M}(m \times n)$ e $\alpha \in R$, tem-se $(\alpha \mathbf{a})^T = \alpha \mathbf{a}^T$;*
4. *dados $\mathbf{a} \in \mathcal{M}(m \times n)$ e $\mathbf{b} \in \mathcal{M}(n \times p)$, tem-se que $(\mathbf{ba})^T = \mathbf{a}^T \mathbf{b}^T$;*
5. *dado $\mathbf{a} \in \mathcal{M}(m \times n)$, tem-se que $(\mathbf{a}^T)^T = \mathbf{a}$;*
6. *uma matriz quadrada $\mathbf{a}$ é invertível se, e somente se, $\mathbf{a}^T$ é invertível; tem-se ainda que $(\mathbf{a}^T)^{-1} = (\mathbf{a}^{-1})^T$.*

Seja E um espaço vetorial munido de um produto interno. Define-se o *complemento ortogonal* de um conjunto $X \subset E$ como o conjunto $X^\perp = \{v \in E : \forall x \in X, \langle v, x \rangle = 0\}$. Vemos imediatamente que $X^\perp$ é um subespaço de E. Além disso, se $u \in E$ é uma combinação linear de vetores de X, então $v \in X^\perp \Rightarrow \langle v, u \rangle = 0 \Rightarrow v \in S(X)^\perp$. Como a recíproca é claramente verdadeira, temos que $S(X)^\perp = X^\perp$.

Teorema 7.6. *Seja E um espaço vetorial de dimensão finita munido de um produto interno. Para qualquer subespaço $F \subset E$, tem-se que $E = F \oplus F^\perp$.*

Demonstração. Seja $\{u_1, \dots, u_m\}$ uma base ortonormal de F. O operador linear $P : E \to E$ definido por $P(v) = \mathrm{pr}_F(v) = \sum_{i=1}^m \langle u_i, v \rangle u_i$ é uma projeção. Além disso, $Im(P) = F$ e $N(P) = F^\perp$. Portanto, $E = F \oplus F^\perp$. $\square$

Corolário 7.7. *Seja E um espaço vetorial de dimensão finita munido de um produto interno. Para qualquer subespaço $F \subset E$, tem-se que* $\dim E = \dim F + \dim F^\perp$.

Corolário 7.8. *Seja E um espaço vetorial de dimensão finita munido de um produto interno. Para qualquer subespaço $F \subset E$, tem-se que $(F^{\perp})^{\perp} = F$.*

Teorema 7.9. *Sejam E e F espaços vetoriais de dimensão finita munidos de produto interno. Dada a transformação linear $A : E \to F$, tem-se que*

1. $N(A)^{\perp} = Im(A^*)$;
2. $Im(A)^{\perp} = N(A^*)$;
3. $N(A^*)^{\perp} = Im(A)$;
4. $Im(A^*)^{\perp} = N(A)$.

Demonstração.

1. Vemos que $u \in Im(A^*) \Rightarrow \exists w \in F : A^*w = u \Rightarrow \forall v \in N(A), \langle v, u\rangle = \langle Av, w\rangle = 0 \Rightarrow u \in N(A)^{\perp}$. Logo, $Im(A^*) \subset N(A)^{\perp}$. Como $\dim N(A)^{\perp} = \dim E - \dim N(A) = \dim Im(A) = \dim Im(A^*)$, segue que $Im(A^*) = N(A)^{\perp}$.

2-4. Podem ser obtidos a partir do item 1, usando o fato de que $A^{**} = A$ e o corolário 7.8.
□

Corolário 7.10. *O sistema linear*

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\quad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_n \end{aligned}$$

tem solução se, e somente se, o vetor $(b_1, \ldots, b_n) \in \mathbb{R}^n$ é ortogonal a toda solução do sistema linear

$$\begin{aligned} a_{11}y_1 + \cdots + a_{m1}y_m &= 0 \\ \vdots \qquad\quad & \quad \vdots \\ a_{1n}y_1 + \cdots + a_{nm}y_m &= 0\,. \end{aligned}$$

8 Subespaços invariantes

Seja E um espaço vetorial. Diz-se que um subespaço $F \subset E$ é *invariante* por um operador linear $A : E \to E$ quando $A(F) \subset F$. Por exemplo, $N(A)$ e $Im(A)$ são exemplos de subespaços invariantes por A. Um vetor $v \in E$, $v \neq 0$, é dito um *autovetor* de A quando existe $\lambda \in \mathbb{R}$ tal que $Av = \lambda v$. Nesse caso, λ é chamado de um *autovalor* de A. Vemos que dizer que existe um subespaço de dimensão 1 invariante por A equivale a dizer que A tem um autovetor.

Seja $p(x) = a_0 + a_1x + \cdots + a_nx^n$ um polinômio com $a_n > 0$. Diz-se que p é *mônico* quando $a_n = 1$. Pelo teorema fundamental da álgebra, o polinômio p tem n raízes complexas $r_1, \ldots, r_n$ e, dessa maneira, $p(x) = (x - r_1)\cdots(x - r_n)$. Se queremos trabalhar só com números reais, então alguns pares de fatores na decomposição de $p(x)$ formarão *polinômios irredutíveis* de grau 2, ou seja, que não têm raízes reais. Dessa maneira, todo polinômio não-constante pode ser escrito como o produto de polinômios mônicos irredutíveis de grau 1 ou 2.

Se $A : E \to E$ é um operador linear em um espaço de dimensão finita e $p(x) = a_0 + a_1x + \cdots + a_nx^n$, definimos $p(A) = a_0 + a_1A \cdots + a_nA^n$.

Lema 8.1. *Seja E um espaço vetorial de dimensão finita. Dado um operador linear $A : E \to E$, existem um polinômio mônico irredutível p, de grau 1 ou 2, e um vetor $v \in E$ não-nulo tais que $p(A)v = 0$.*

Demonstração. Se $\dim E = n$, então o espaço vetorial $\mathcal{L}(E)$ tem dimensão n^2, pois é isomorfo ao espaço $\mathcal{M}(n \times n)$. Logo, o conjunto $\{I, A, \ldots, A^{n^2}\}$ é L.D.. Logo, existem $\alpha_0, \alpha_1, \ldots, \alpha_{n^2} \in \mathbb{R}$ não todos nulos tais que $\alpha_0 I + \alpha_1 A + \cdots + \alpha_{n^2} A^{n^2} = 0$. Seja m o maior índice para o qual $\alpha_m \neq 0$ e definamos o polinômio $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_m x^m$. Pelo teorema fundamental da álgebra, existem polinômios mônicos irredutíveis $q_1, \ldots, q_r$, $r < m$, de grau 1 ou 2 tais que $p(x) = q_1(x) \cdots q_r(x)$. Logo, $q_1(A) \cdots q_r(A) = p(A) = 0$. Isso implica que pelo menos um dos operadores $q_i(A)$, $i = 1, \ldots, r$, não deve ser invertível e, por conseguinte, não deve ser injetivo. Logo, existem um polinômio mônico irredutível $q(x)$ de grau 1 ou 2 e um vetor $v \in E$, $v \neq 0$, tais que $q(A)v = 0$. □

Teorema 8.2. *Dado um espaço vetorial de dimensão finita E, seja $A : E \to E$ um operador linear. Existem subespaços de dimensão 1 ou 2 invariantes por A.*

Demonstração. Segue do lema 8.1 que existem um polinômio mônico irredutível p, de grau 1 ou 2, e um vetor $v \in E$ não-nulo tais que $p(A)v = 0$. Se p tem grau 1, então existe $\lambda \in \mathbb{R}$ tal que $p(x) = -\lambda + x$. Logo, $p(A)v = 0 \Rightarrow Av = \lambda v$, ou seja, v é um autovetor de A. Se p tem grau 2, existem $\alpha_0, \alpha_1 \in \mathbb{R}$ tais que $p(x) = \alpha_0 + \alpha_1 x + x^2$ e $\alpha_1^2 - 4\alpha_0 < 0$. Logo, $p(A)v = 0 \Rightarrow A^2 v = -\alpha_0 v - \alpha_1 Av$, o que implica que $A^2 v \in S(\{v, Av\})$. Dessa maneira, $S(\{v, Av\})$ é invariante por A. Por outro lado, $\{v, Av\}$ é um conjunto L.I., pois caso contrário existiria $\lambda \in \mathbb{R}$ tal que $Av = \lambda v$. Isso implicaria que $p(A)v = (\alpha_0 + a_1\lambda + \lambda^2)v = 0$, ou seja, $\alpha_0 + a_1\lambda + \lambda^2 = 0$. Porém, isso é impossível para todo $\lambda \in \mathbb{R}$. Portanto, $S(\{v, Av\})$ tem dimensão 2. □

Teorema 8.3. *A autovalores diferentes de um operador linear $A : E \to E$ correspondem autovetores L.I..*

Demonstração. Sejam $v_1, v_2 \in E$ autovetores de A correspondentes a autovalores diferentes $\lambda_1, \lambda_2 \in \mathbb{R}$. Se $\alpha_1 v_1 + \alpha_2 v_2 = 0$, então $\alpha_1 Av_1 + \alpha_2 Av_2 = 0$ e, por conseguinte, $\alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 = 0$. Multiplicando a primeira equação por λ_2, obtemos que $\alpha_1\lambda_2 v_1 + \alpha_2\lambda_2 v_2 = 0$. Logo, restando a terceira equação, vamos ter que $\alpha_1(\lambda_1 - \lambda_2)v_1 = 0$, o que implica que $\alpha_1 = 0$ e, por conseguinte, $\alpha_2 = 0$. Suponhamos que, se $v_1, \ldots, v_n \in E$ são autovetores de A que correspondem a autovalores diferentes $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, o conjunto $\{v_1, \ldots, v_n\}$ é L.I.. Se $v_{n+1} \in E$ é um autovetor de A que corresponde a um autovalor $\lambda_{n+1} \notin \{\lambda_1, \ldots, \lambda_n\}$ e $\alpha_1 v_1 + \cdots + \alpha_n v_n + \alpha_{n+1} v_{n+1} = 0$, então $\alpha_1\lambda_1 v_1 + \cdots + \alpha_n\lambda_n v_n + \alpha_{n+1}\lambda_{n+1} v_{n+1} = 0$. Multiplicando a primeira equação por λ_{n+1} e restando a segunda equação à equação resultante, obtemos que $\alpha_1(\lambda_1 - \lambda_{n+1})v_1 + \cdots + \alpha_n(\lambda_n - \lambda_{n+1})v_n = 0$. Logo, $\alpha_1 = \ldots = \alpha_n = 0$, pois $\{v_1, \ldots, v_n\}$ é L.I., e isso implica que $\alpha_{n+1} = 0$. □

Corolário 8.4. *Seja E um espaço vetorial dimensão n. Se um operador linear $A : E \to E$ tem n autovalores diferentes, então existe uma base $\{v_1, \ldots, v_n\}$ de E formada pelos autovetores de A. A matriz de A em relação a essa base é* diagonal*, ou seja, $a_{ij} = 0$ se $i \neq j$.*

Dado um operador linear $A : E \to E$, tem-se que $\lambda \in \mathbb{R}$ é um autovalor de A se, e somente se, existe $v \in E$, $v \neq 0$, tal que $(A - \lambda I_E)v = 0$. Em outras palavras λ é um autovalor de A se, e somente se, $A - \lambda I_E$ é um operador linear não-invertível. Se $\dim E = 2$, dada uma base $\{u, v\}$ de E, existem $a, b, c, d \in \mathbb{R}$ tais que $A(u) = au + bv$ e $A(v) = cu + dv$.