Prof. Dr. Maurício Zahn
UFPel

Álgebra linear I

TEXTO DE MENSAGEM...

Dedicamos este trabalho a ...

Prefácio

Este material foi elaborado durante o Primeiro Semestre letivo de 2016, para atender a Disciplina de Álgebra Linear I que ministrei para os cursos de Física e Matemática da UFPel. Estas notas de aula estão sendo escritas como um material de apoio para os estudantes, em conjunto com as lista de exercícios, e corresponde ao conteúdo desenvolvido na referida disciplina.

Maurício Zahn

Conteúdo

Capítulo 1

Matrizes e sistemas lineares

1.1 Matrizes

Definição 1.1 Chama-se *matriz* a uma tabela com m linhas e n colunas, constituída por números, chamados de elementos da matriz.

Uma matriz será identificada por uma letra maiúscula e um elemento dessa matriz será indicado pela letra minúscula correspondente, acompanhada de dois índices i e j, onde o primeiro índice indica a linha em que tal elemento se encontra e o segundo índice a coluna onde o mesmo se encontra. Dessa forma, uma matriz A com m linhas e n colunas costuma ser representada, simbolicamente, por

$$A = (a_{ij})_{m \times n}.$$

Assim, abrindo a notação matricial acima, escrevemos

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

Por exemplo, considerando a matriz $B_{3\times 2}$ abaixo

$$B = \begin{pmatrix} 2 & -1 \\ 4 & 0 \\ 3 & 7 \end{pmatrix},$$

temos que

- b_{11}, o elemento da linha 1 e coluna 1, vale 2;
- b_{12}, o elemento da linha 1, coluna 2, vale -1;
- etc.

Observe que a matriz B dada acima possui 3 linhas e duas colunas, por isso escrevemos $B_{3\times 2}$. Dada uma matriz $A_{m\times n}$, dizemos que que $m \times n$ é o *tamanho* ou a *ordem* da matriz em questão.

Quando $m = n$, ou seja quando o número de linhas é igual ao número de colunas, dizemos que a matriz é *quadrada*, e nesse caso os elementos a_{ii} formam a *diagonal da matriz*, também chamada de diagonal principal. Quando $m \neq n$, dizemos que a matriz é *retangular*.

1.1.1 Tipos especiais de matrizes

Nessa seção vamos apresentar os principais tipos de matrizes.

(a) **Matriz nula.** É uma matriz quadrada ou retangular, onde todas as entradas são nulas. Por exemplo,

$$0_{2\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{e} \quad 0_{2\times 4} \;\text{ e }\; 0_{2\times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Quando definirmos a soma de matrizes veremos que a matriz nula corresponde ao neutro aditivo.

(b) **Matriz diagonal.** É uma matriz quadrada onde $a_{ij} = 0$ se $i \neq j$. Exemplo:

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

Uma outra forma de denotar essa mesma matriz é $D = \operatorname{diag}(2, -1, 8)$.

(c) **Matriz identidade.** É uma matriz diagonal (e quadrada) definida por

$$I_3 = \operatorname{diag}(1,1,1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

No caso quando forma uma matriz $n \times n$, temos

$$I_n = \operatorname{diag}(\underbrace{1,1,1,...,1}_{n}) = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ .... & .... & .... & .... & .... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix}$$

Uma outra maneira de denotar a matriz identidade de ordem n é escrever $I = (\delta_{ij})_{n\times n}$, onde

$$\delta_{ij} = \begin{cases} 1, \text{ se } & i = j \\ 0, \text{ se } & i \neq j \end{cases}.$$

Quando definirmos os produto de matrizes, veremos que esssa matriz corresponde à unidade multiplicativa, i.e., é o neutro multiplicativo.

(d) **Matriz triangular inferior.** É a matriz quadrada $A = (a_{ij})$ tal que $a_{ij} = 0$, se $i < j$. Ou seja, é uma matriz onde acima da diagonal as entradas são todas iguais a zero. Por exemplo, a matriz

$$T = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 2 & 8 \end{pmatrix}$$

é uma matriz triangular inferior.

(e) **Matriz triangular superior.** É a matriz quadrada $A = (a_{ij})$ tal que $a_{ij} = 0$, se $i > j$. Ou seja, é uma matriz onde abaixo da diagonal as entradas são todas iguais a zero. Por exemplo, a matriz

$$T = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

é uma matriz triangular superior. Note também que a matriz identidade I_n é ao mesmo tempo triangular inferior e superior.

1.1.2 Operações com matrizes

No que segue, vamos definir a adição de matrizes, o produto de uma constante por uma matriz e o produto de matrizes. Tendo em vista que precisaremos recorrer à notação de somatório, por ser mais compacta, vamos definir inicialmente esse conceito.

Definição 1.2 Seja $F(n)$ uma expressão qualquer que depende de $n \in \mathbb{N}$. Definimos o *somatório* com k de 1 até um índice n fixado, dos $F(k)$ por

$$\sum_{k=1}^{n} F(k) = F(1) + F(2) + F(3) + ... + F(n).$$

Assim, se $F(n) = n^2$, temos, por exemplo, que

$$\sum_{k=1}^{5} F(k) = \sum_{k=1}^{5} k^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 1 + 4 + 9 + 16 + 25 = 55.$$

O somatório goza das seguintes propriedades: sejam $F(n)$ e $G(n)$ duas expressões que dependem de n e $\alpha \in \mathbb{R}$, temos

(a) $\displaystyle\sum_{k=1}^{n} \alpha = \alpha \cdot n.$

(b) $\displaystyle\sum_{k=1}^{n} \alpha \cdot F(K) = \alpha \sum_{k=1}^{n} F(k).$

(c) $\displaystyle\sum_{k=1}^{n} (F(k) + G(k)) = \sum_{k=1}^{n} F(k) + \sum_{k=1}^{n} G(k).$

Todas essas propriedades são muito simples de provar, bastando abrir a definição de somatório. Deixaremos a encargo do leitor.

Definição 1.3 Dadas duas matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$ de mesmo tamanho, e $k \in \mathbb{R}$, definimos a soma de matrizes e o produto de um escalar por uma matriz, respectivamente, por

$$A + B = (a_{ij})_{m\times n} + (b_{ij})_{m\times n} = (a_{ij} + b_{ij})_{m\times n},$$

e

$$k \cdot A = (k \cdot a_{ij})_{m \times n},$$

para $i \in \{1, ..., m\}$ e $j \in \{1, ..., n\}$.

Ou seja, somar duas matrizes de mesmo tamanho é o mesmo que somar dois vetores e multiplicar uma matriz por um número real é o mesmo que multiplicar um vetor por um escalar.

Assim, por exemplo, dadas as matrizes $A = \begin{pmatrix} 2 & 0 \\ -3 & 1 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & -4 \\ 2 & 2 \end{pmatrix}$, temos que

$$A + B = \begin{pmatrix} 2 & 0 \\ -3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & -4 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -4 \\ -1 & 3 \end{pmatrix},$$

e

$$-7A = -7 \begin{pmatrix} 2 & 0 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -14 & 0 \\ 21 & -7 \end{pmatrix}.$$

Definição 1.4 Dadas duas matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{n \times p}$, definimos o produto entre A e B, e escrevemos $A \cdot B$, à matriz $C = (c_{ij})_{m \times p}$, onde

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}.$$

Repare que para o produto $A \cdot B$ esteja bem definido é necessário e suficiente que o número de colunas da primeira matriz seja igual ao número de linhas da segunda. Isto permitirá efetuar um produto entre linha e coluna.

Vamos ser mais claros na definição acima: abrindo as matrizes, teremos

$$A \cdot B = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1p} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2p} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ b_{n1} & b_{n2} & b_{n3} & \dots & b_{np} \end{pmatrix} =$$

$$= \begin{pmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1p} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2p} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ c_{m1} & c_{m2} & c_{m3} & \dots & c_{mp} \end{pmatrix},$$

onde

- $c_{11} = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + ... + a_{1n} \cdot b_{n1} = \sum_{k=1}^{n} a_{1k} \cdot b_{k1}$;
- $c_{12} = a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + ... + a_{1n} \cdot b_{n2} = \sum_{k=1}^{n} a_{1k} \cdot b_{k2}$;
- etc.

De forma geral, para determinar o elemento c_{ij} do produto $C = A \cdot B$, olhamos para a linha i de A e a coluna j de B, e então efetuamos o produto do primeiro elemento da linha i com o primeiro elemento da coluna j, e somamos co o produto do segundo elemento da linha i com o segundo elemento da coluna j, e assim por diante, até somar com o produto do último elemento da linha i com o último elemento da coluna j, ou seja,

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}.$$

Por exemplo, dadas as matrizes $A_{2\times 3} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 0 \end{pmatrix}$ e $B_{3\times 2} = \begin{pmatrix} 2 & -1 \\ 0 & 4 \\ 1 & 1 \end{pmatrix}$, temos que

$$A \cdot B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 4 \\ 1 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 \cdot 2 + 2 \cdot 0 + 3 \cdot 1 & 1 \cdot (-1) + 2 \cdot 4 + 3 \cdot 1 \\ 2 \cdot 2 + (-1) \cdot 0 + 0 \cdot 1 & 2 \cdot (-1) + (-1) \cdot 4 + 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 4 & -6 \end{pmatrix}$$

Convém observar que, neste caso, $A_{2\times 3} \cdot B_{3\times 2}$ resulta numa matriz 2×2. Já o produto $B_{3\times 2} \cdot A_{2\times 3}$ resultará numa matriz 3×3. Ou seja, em geral o produto de matrizes não comuta, ou seja, em geral temos que

$$A \cdot B \neq B \cdot A.$$

De fato, pode acontecer que o produto $A \cdot B$ esteja definido, mas o produto $B \times A$ não esteja sequer definido. Por exemplo, se $A_{2\times 3}$ e $B_{3\times 4}$, temos que

$A \cdot B$ estará bem definido, mas $B \cdot A$ não tem sentido, verifique!

Outro fato importante a ser observado é que dada uma matriz $A_{m\times n}$, temos que I_n é neutro multiplicativo à direita de A e que I_m é neutro multiplicativo à esquerda de A, ou seja,

$$A_{m\times n} \cdot I_n = A \quad \text{e} \quad I_m \cdot A_{m\times n} = A.$$

Definição 1.5 Dizemos que duas matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$, de mesmo tamanho, são *iguais* se $a_{ij} = b_{ij}$, $\forall i \in \{1, 2, ..., m\}$ e $\forall j \in \{1, 2, ..., n\}$.

Ou seja, da mesma forma que estabelecemos uma igualdade de vetores (dois vetores são iguais quando as componentes de mesma posição são iguais), temos que duas matrizes são iguais quando seus elementos de mesma posição forem iguais.

Para verificar uma série de propriedades envolvendo igualdade de matrizes, mostramos a igualdade entre seus elementos de mesma posição.

Teorema 1.6 *(Propriedades algébricas das matrizes) Sejam A, B, C matrizes e $\alpha, \beta \in \mathbb{R}$. Então, valem as seguintes propriedades (onde elas estiverem bem definidas)*

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

02) $A + B = B + A$.

03) *existe uma matriz 0 tal que $A + 0 = 0 + A = A$ (matriz nula).*

04) *para toda matriz A, existe uma matriz B tal que $A + B = 0$. Denotamos $B = -A$.*

05) *existe uma matriz I tal que $A \cdot I = I \cdot A = A$. (matriz identidade)*

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

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

08) $A \cdot (B + C) = A \cdot B + A \cdot C$.

09) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

10) $\alpha(A \cdot B) = (\alpha A) \cdot B = A \cdot (\alpha B)$.

Observação 1.7 Note que a propriedade 04) nos motiva o conceito de *diferença entre matrizes.* Ou seja, dadas duas matrizes A e B, de mesmo tamanho, podemos definir a diferença entre A e B pondo

$$A - B = A + (-B),$$

que não deixa de ser parecido com a maneira que definimos a diferença de dois vetores. Mas cuidado, a matriz $-A$ não recebe o nome de "simétrica"! Matriz simétrica é um conceito bem diferente e o veremos na Definição 1.11.

Demonstração da Proposição. Faremos a prova de algumas apenas e deixaremos as demais para o leitor.

01) $A + (B + C) = (A + B) + C$:
Escreva $B + C = S$, onde $S = (s_{ij})$ é tal que

$$s_{ij} = b_{ij} + c_{ij}.$$

Dessa forma, escrevemos

$$A + (B + C) = A + S = T = (t_{ij}),$$

onde

$$t_{ij} = a_{ij} + s_{ij}.$$

Por outro lado, escreva $A + B = Z$, onde $Z = (z_{ij})$, tal que

$$z_{ij} = a_{ij} + b_{ij},$$

Dessa forma, escrevendo

$$(A + B) + C = Z + C = W = (w_{ij}),$$

onde

$$w_{ij} = z_{ij} + c_{ij} = (a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij}) = a_{ij} + s_{ij} = t_{ij}.$$

Assim, concluímos que $w_{ij} = t_{ij}$, $\forall i$, $\forall j$, ou seja, $W = T$, o que mostra que

$$(A + B) + C = Z + C = W = T = A + S = A + (B + C),$$

provando 01).

05) Existe uma matriz I tal que $A \cdot I = I \cdot A = A$ (matriz identidade):
De fato, tome $A = (a_{ij})_{n\times n}$ e considere $I = (\delta_{ij})_{n\times n}$. Então, escrevendo $A \cdot I = C$, temos que

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot \delta_{kj} = a_{i1} \cdot \delta_{1j} + a_{i2} \cdot \delta_{2j} + ... + a_{ij} \cdot \delta_{jj} + ... + a_{in} \cdot \delta_{n_j} = a_{ij},$$

pois $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$. Ou seja, concluímos que

$$c_{ij} = a_{ij}, \forall i, j \in \{1, 2, ..., n\},$$

ou seja, mostramos que

$$A \cdot I = C = A.$$

Analogamente mostramos que $I \cdot A = A$. Isso prova 05).

08) $A \cdot (B + C) = A \cdot B + A \cdot C$:
Dadas as matrizes $A_{m\times n}$, $B_{n\times p}$ e $C_{n\times p}$, escreva

$$S = B + C,$$

e denotando $S = (s_{ij})_{n\times p}$, onde

$$s_{ij} = b_{ij} + c_{ij},$$

temos

$$A \cdot (B + C) = A_{m\times n} \cdot S_{n\times p} = T_{m\times p},$$

onde

$$t_{ij} = \sum_{k=1}^{n} a_{ik} \cdot s_{kj}. \tag{1.1}$$

Por outro lado, denote

$$A \cdot B = P_{m\times p} \quad \text{e} \quad A \cdot C = Q_{m\times p},$$

onde
$$p_{ij} = \sum_{\ell=1}^{n} a_{i\ell} \cdot b_{\ell j} \quad \text{e} \quad q_{ij} = \sum_{\ell=1}^{n} a_{i\ell} \cdot c_{\ell j}.$$

Dessa forma, temos que

$$A \cdot B + A \cdot C = P + Q = R,$$

onde cada elemento r_{ij} dessa matriz é dado por

$$r_{ij} = p_{ij} + q_{ij} = \sum_{\ell=1}^{n} a_{i\ell} \cdot b_{\ell j} + \sum_{\ell=1}^{n} a_{i\ell} \cdot c_{\ell j} = \sum_{\ell=1}^{n} (a_{i\ell} \cdot b_{\ell j} + a_{i\ell} \cdot c_{\ell j}) =$$

$$= \sum_{\ell=1}^{n} a_{i\ell}(b_{\ell j} + c_{\ell j}) = \sum_{\ell=1}^{n} a_{i\ell} \cdot s_{\ell j} = t_{ij},$$

ou seja,

$$r_{ij} = t_{ij},\ \forall i \in \{1, 2, ..., m\},\ \forall j \in \{1, 2, ..., p\},$$

ou seja, mostramos que $R = T$, i.e.,

$$A \cdot B + A \cdot C = R = T = A(B + C).$$

$(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Dadas as matrizes $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{n\times p}$ e $C = (c_{ij})_{p\times q}$, mostraremos que
$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Escreva $M = (A \cdot B)_{m\times p}$ e $W = (B \cdot C)_{n\times q}$. Assim, temos que os elementos m_{ij} de M e w_{ij} de W são, respectivamente, dados por

$$m_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj} \quad \text{e} \quad w_{ij} = \sum_{\ell=1}^{p} b_{i\ell} \cdot c_{\ell j}. \tag{1.2}$$

Note que, denotando

$$(A \cdot B) \cdot C = (M)_{m\times p} \cdot (C)_{p\times q} = (F)_{m\times q},$$

usando (1.2) temos que cada elemento f_{ij} de F é dado por

$$f_{ij} = \sum_{\ell=1}^{p} m_{i\ell} \cdot c_{\ell j} = \sum_{\ell=1}^{p} \left(\sum_{k=1}^{n} a_{ik} \cdot b_{k\ell} \right) \cdot c_{\ell j} = \sum_{\ell=1}^{p} \sum_{k=1}^{n} (a_{ik} \cdot b_{k\ell} \cdot c_{\ell j}) =$$

$$= \sum_{\ell=1}^{p} \sum_{k=1}^{n} a_{ik} \cdot (b_{k\ell} \cdot c_{\ell j}) = \sum_{k=1}^{n} \sum_{\ell=1}^{p} a_{ik} \cdot (b_{k\ell} \cdot c_{\ell j}) = \sum_{k=1}^{n} a_{ik} \left(\sum_{\ell=1}^{p} b_{k\ell} \cdot c_{\ell j} \right) =$$

$$= \sum_{k=1}^{n} a_{ik} \cdot w_{kj}.$$

Por outro lado, denotando

$$A \cdot (B \cdot C) = A_{m \times n} \cdot W_{n \times q} = G_{m \times q},$$

onde

$$g_{ij} = \sum_{k=1}^{n} a_{ik} \cdot w_{kj} = f_{ij}.$$

Como $g_{ij} = f_{ij}$, $\forall i \in \{1, 2, ..., m\}$ e $\forall j \in \{1, 2, ..., q\}$, concluímos que as matrizes F e G são iguais, ou seja, que

$$(A \cdot B) \cdot C = F = G = A \cdot (B \cdot C).$$

$\square$

Obs. Na prova acima usamos uma propriedade "comutativa" para somatórios, ou seja, vale a propriedade

$$\sum_{i=1}^{n} \sum_{j=1}^{p} F(i, j) = \sum_{j=1}^{p} \sum_{i=1}^{n} F(i, j),$$

onde $F(i, j)$ é uma expressão qualquer que depende dos índices i e j.

Como um bom exercício, prove essa igualdade.

1.1.3 Matriz transposta

Definição 1.8 Dada uma matriz $A = (a_{ij})_{m\times n}$, definimos a *matriz transposta* de A , e denotamos por A^t, a matriz definida por

$$A^t = (a_{ji})_{n\times m}.$$

Em palavras, dada uma matriz A, a transposta de A, A^t, é a matriz na qual sua primeira linha de A^t corresponde à primeira coluna de A, e assim por diante, ou seja, linhas se transformam em colunas. Vejamos um exemplo.

Exemplo 1.9

$$A = \begin{pmatrix} 2 & -1 & 7 \\ 1 & 0 & 3 \end{pmatrix}_{2\times 3} \Rightarrow A^t = \begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 7 & 3 \end{pmatrix}_{3\times 2}$$

Proposição 1.10 *A transposição de matrizes goza das seguintes propriedades (onde elas estejam bem definidas):*

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

02) $(k\cdot A)^t = k\cdot A^t$, *onde* $k\in\mathbb{R}$.

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

04) $(A\cdot B)^t = B^t\cdot A^t$.

Demonstração. Faremos a prova de 01), 03) e 04) e deixaremos as prova de 02) para o leitor.

01) $(A^t)^t = A$.
Seja a_{ij} o elemento da linha i e coluna j de A.
Então a_{ji} é o elemento da linha j e coluna i de A^t
Por fim, transpondo novamente, concluímos que a_{ij} é o elemento da linha i e coluna j de $(A^t)^t$.
Portanto, concluímos que, $A = (A^t)^t$.

03) $(A+B)^t = A^t + B^t$.
Escreva $C = A + B$. Então, basta mostrar que

$$C^t = A^t + B^t.$$

Sendo $C = A + B$, temos que o elemento c_{ij} da linha i e coluna j de C é dado por

$$c_{ij} = a_{ij} + b_{ij}.$$

Logo, o elemento $c_{ji} = a_{ji} + b_{ji}$ é o elemento da linha j e coluna i de C^t. Do mesmo modo, a_{ji} é o elemento da linha j e coluna i de A^t e b_{ji} é o elemento da linha j e coluna i de B^t. Ou seja, concluímos que

$$C^t = A^t + B^t.$$

04) $(A \cdot B)^t = B^t \cdot A^t$.
Tome $A_{m\times n}$ e $B_{n\times p}$ e escreva $C_{m\times p} = A \cdot B$. Então, o elemento c_{ij} de C é dado por

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}.$$

Vamos mostrar que $C^t = B^t \cdot A^t$.

Denotando por c^t_{ij} o elemento na linha i e coluna j de C^t, temos que

$$c^t_{ij} = c_{ji} = \sum_{k=1}^{n} a_{jk} \cdot b_{ki} = \sum_{k=1}^{n} a^t_{kj} \cdot b^t_{ik} = \sum_{k=1}^{n} b^t_{ik} \cdot a^t_{kj}, \tag{1.3}$$

onde a^t_{kj} é elemento de A^t e b^t_{ik} é elemento de B^t.

Por outro lado, definindo $M = B^t \cdot A^t$, com elemento

$$m_{ij} = \sum_{k=1}^{n} b^t_{ik} \cdot a^t_{kj},$$

segue de (1.3) que $c^t_{ij} = m_{ij}$, ou seja, concluímos que

$$C^t = M = B^t \cdot A^t,$$

e como $C = A \cdot B$, segue o resultado.

$\square$

Definição 1.11 Dizemos que uma matriz A é *simétrica* quando $A^t = A$ e que A é *anti-simétrica* quando $A^t = -A$.

Por exemplo, é fácil ver que a matriz identidade I é simétrica, e que a matriz $A = \begin{pmatrix} 2 & -1 & 3 \\ -1 & 4 & 0 \\ 3 & 0 & 5 \end{pmatrix}$ também é simétrica. Já a matriz $B = \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 3 \\ 1 & -3 & 0 \end{pmatrix}$ é anti-simétrica.

Exercícios

1. Seja $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{pmatrix}$, encontre as matrizes AA^t e A^tA, se existirem.

2. Seja $A = \begin{pmatrix} 2 & x^2 \\ 2x-1 & 0 \end{pmatrix}$. Se $A^t = A$, encontre o valor de x.

3. Dadas duas matrizes simétricas $n \times n$ A e B.

   (a) Prove que $A + B$ é simétrica.

   (b) Prove que se $AB = BA$, então AB é simétrica.

1.1.4 Matrizes inversíveis

No estudo de funções foi dado um importante conceito de função inversa. Relembrando, considerando X e Y dois conjuntos não vazios, dada uma função $f : X \to Y$, dizemos que f é inversível se existir uma função $g : Y \to X$ tal que

$$f \circ g = id_X \quad \text{e} \quad g \circ f = id_Y,$$

onde $id_X : X \to X$, $id_X(x) = x$ e $id_Y : Y \to Y$, $id_Y(y) = y$ são, respectivamente, as funções identidades de X e de Y.

Nesse caso, dizemos que g é a inversa de f e denotamos $g = f^{-1}$. (e também, g será inversível e sua inversa, g^{-1}, será f).

Existe, no estudo de matrizes um tipo de matriz que desempenha um papel análogo ao da função inversa, chamada de matriz inversa. Do mesmo mosdo que nem todas as funções são inversíveis (somente àquelas que obedecem a definição acima), temos que nem todas as matrizes possuirão inversas. No caso de matrizes, a composição da definição acima será substituída pelo produto de matrizes e as funções identidades, que fazem o papel da unidade da composição de funções, serão substituídas pela equivalente unidade do produto de matrizes, ou seja, a matriz identidade I. Ou seja, isso nos inspira a definição que segue.

Definição 1.12 Dizemos que uma matriz quadrada A é *inversível* se existir uma matriz B de mesmo tamanho tal que

$$A \cdot B = I \quad \text{e} \quad B \cdot A = I.$$

Nesse caso dizemos que B é a inversa de A, e escrevemos $B = A^{-1}$.

É claro que, pela definição acima temos que B também será inversível e sua inversa será $B^{-1} = A$.

Note também na definição que uma *conditio sine qua non* para que uma matriz tenha chance de possuir uma inversa é que ela deve ser quadrada.

Do mesmo modo que a inversa de uma função, quando existe, é única, vale também a unicidade da matriz inversa, quando esta existir. Ou seja, temos o resultado:

Proposição 1.13 *Se A for uma matriz inversível com inversas B e C, então $B = C$.*

Demonstração. Sejam B e C ambas inversas de A. Assim, como B é inversa de A temos que

$$B \cdot A = I.$$

Multiplicando-se à direita por C, vem

$$(B \cdot A) \cdot C = I \cdot C = C,$$

e por associatividade, obtemos

$$B \cdot (A \cdot C) = C.$$

Mas como C também é inversa de A, temos que $A \cdot C = I$, e com isso, a igualdade acima fica

$$C = B \cdot (A \cdot C) = B \cdot I = B \Longrightarrow C = B,$$

e isso prova a unicidade de inversa, como queríamos.

□

Proposição 1.14 *Valem as seguintes propriedades para uma matriz inversível A:*

(a) $(A^{-1})^{-1} = A$.

(b) $(k \cdot A)^{-1} = k^{-1} \cdot A^{-1}$, *onde* $k \in \mathbb{R} \setminus \{0\}$.

Demonstração.
(a) Como por hipótese A é inversível, temos que existe inversa A^{-1} tal que

$$A \cdot A^{-1} = I \quad \text{e} \quad A^{-1} \cdot A = I.$$

Mas então, temos que A^{-1} também é inversível, com inversa $(A^{-1})^{-1} = A$.

(b) Basta observar que

$$(kA)(k^{-1}A^{-1}) = k^{-1}(kA)A^{-1} = (k^{-1}k)(A \cdot A^{-1}) = 1 \cdot I = I,$$

e

$$(k^{-1}A^{-1})(kA) = k(k^{-1}A^{-1})A = (kk^{-1})(A^{-1} \cdot A) = 1 \cdot I = I,$$

e então temos que kA é inversível, com inversa $(kA)^{-1} = k^{-1}A^{-1}$.

□

Proposição 1.15 *Sejam A e B duas matrizes quadradas inversíveis de mesmo tamanho. Então, o produto $A \cdot B$ também é inversível, com inversa*

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

Demonstração. De fato, como A e B são inversíveis com inversas dadas, respectivamente, por A^{-1} e B^{-1}, basta efetuar os produtos, usando a associatividade:

$$(A \cdot B)(B^{-1} \cdot A^{-1}) = A(B \cdot B^{-1})A^{-1} = A \cdot I \cdot A^{-1} = A \cdot A^{-1} = I,$$

e

$$(B^{-1} \cdot A^{-1})(A \cdot B) = B^{-1}(A^{-1} \cdot A)B = B^{-1} \cdot I \cdot B = B^{-1} \cdot B = I,$$

o que mostra que $A \cdot B$ també é inversível, com inversa $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

□

1.1.5 Potências de matrizes

Motivados pelos conceitos de produto de matrizes e de matriz inversível, podemos definir:

Definição 1.16 Seja A uma matriz quadrada. Definimos as potências de A pondo

$$A^0 = I$$

e

$$A^n = \underbrace{A \cdot A \cdot ... \cdot A}_{n \text{ fatores}}, \forall n \in \mathbb{N}.$$

Quando A for inversível, definimos as potências negativas de A por

$$A^{-n} = (A^{-1})^n = \underbrace{A^{-1} \cdot A^{-1} \cdot ... \cdot A^{-1}}_{n \text{ fatores}}.$$

Segue diretamente dessa definição a Proposição abaixo, cuja demonstração fica como exercício para o leitor:

Proposição 1.17 *Seja A uma matriz inversível. Então, para $n \in \mathbb{N}$ fixado, A^n é inversível, e*

$$(A^n)^{-1} = (A^{-1})^n.$$

Exercícios

1. Prove que a matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ é inversível se, e somente se, $ad - bc \neq 0$. Neste caso, mostre que a inversa é dada por

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2. Mostre que, se A for uma matriz invertível, então A^t também é invertível, com $(A^t)^{-1} = (A^{-1})^t$.

3. Suponha que A seja uma matriz inversível. Mostre que se $AB = AC$, então $B = C$. Dê um exemplo de uma matriz não nula A tal que $AB = AC$, mas $B \neq C$.

4. Se A e B são matrizes quadradas e A é inversível, verifique que

$$(A + B)A^{-1}(A - B) = (A - B)A^{-1}(A + B).$$

1.1.6 Matriz na forma escalonada reduzida por linhas

Definição 1.18 Dizemos que uma matriz A está na *forma escalonada reduzida por linhas* se:

(i) numa linha não totalmente constituída por zeros, o primeiro elemento não nulo, chamado de *pivô*, vale 1;

(ii) se existirem linhas totalmente nulas, elas estão localizadas na base da matriz;

(iii) em duas linhas quaisquer não nulas, o pivô da linha superior localiza-se mais à esquerda do pivô da linha inferior;

(iv) em cada coluna onde há um pivô, os demais elementos dessa coluna são zeros.

Por exemplo, a matriz $A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ está na forma escalonada reduzida por linhas.

Essa forma de representação será bastante útil na resolução de sistemas lineares.

Observação 1.19 Note que, da definição acima, se uma matriz quadrada está na forma escalonada reduzida por linhas, então obrigatoriamente ela possui linhas nulas na base ou é a matriz identidade.

1.2 Sistemas lineares

Nesta seção vamos apresentar o conceito de sistema linear, bem como uma importante técnica de resolução via matrizes.

Definição 1.20 Chama-se uma equação linear a n incógnitas $x_1, x_2, ..., x_n$ a equação dada por

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1,$$

onde $a_{11}, a_{12}, ..., a_{1n}, b_1 \in \mathbb{R}$ são constantes conhecidas. Quando $b_1 = 0$ tal equação chama-se homogênea.

No caso da equação homogênea, a mesma admite uma solução trivial: $x_1 = x_2 = ... = x_n = 0$. Outras soluções são chamadas de não-triviais.

Definição 1.21 Chama-se um *sistema linear* um conjunto de equações lineares da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m_2}x_2 + ... + a_{mn}x_n = b_n \end{cases} . \tag{1.4}$$

A definição acima é bem geral: trata-se de um sistema com m equações a n incógitas.

Quando $b_i = 0$ para $i = 1, 2, ..., n$ o sistema chama-se *homogêneo*.
Por exemplo, o sistemas

$$\begin{cases} x_1 + x_2 + 2x_3 = 9 \\ 2x_1 + 4x_2 - 3x_3 = 1 \\ 3x_1 + 6x_2 - 5x_3 = 0 \end{cases} \quad \text{e} \quad \begin{cases} 2x + y = 2 \\ x - y - z = 3 \end{cases}$$

são sistemas lineares, o primeiro com três equações a três incógnitas e o segundo com duas equações a três incógnitas.

Definição 1.22 Chama-se *solução* do sistema linear (1.4) de m equações a n incógnitas a n-nupla ordenada $(a_1, a_2, ..., a_n)$ tal que, quando $x_1 = a_1$, ... $x_n = a_n$ em (1.4) , obtemos uma identidade.

Em outras palavras, uma solução de um determinado sistema linear são os valores que substituídos nas incógnitas das equações que definem o sistema verificarem a todas as equações.

Existem várias técnicas para resolver um sistema linear. A mais simples, consiste em, ir operando adequadamente as equações do sistema com o intuito de ir eliminando incógnitas até obtermos uma equação numa só variável. Depois, ir "retrocedendo" nas operações realizadas para determinar todas as demais incógnitas, e com isso montar a solução. Vejamos alguns exemplos práticos.

Exemplo 1.23 Resolver o sistema linear

$$\begin{cases} 3x - y = 5 \\ x + y = 3 \end{cases}$$

Solução. Podemos eliminar y somando as duas equações. Assim, somando-as obtemos a equação

$$4x = 8 \Rightarrow x = 2.$$

Dessa forma, sendo $x = 2$, na equação $x + y = 3$ vamos obter $2 + y = 3$, e portanto $y = 1$. Assim, a solução do sistema dado é $(x, y) = (2, 1)$. Repare que, geometricamente, cada equação no plano cartesiano $\mathbb{R}^2$, corresponde a equação de uma reta. Então a solução do sistema acima nos diz que essas retas são concorrentes, ou seja, se interceptam num ponto do plano cartesiano, exatemente no ponto de coordenadas $(2, 1)$.

No entanto, sabemos que duas retas no plano podem ser também coincidentes ou paralelas. No primeiro caso, o sistema assumiria infinitas soluções, e é chamado de *indeterminado*, já no segundo caso, o sistema não teria soluções, e seria chamado de *impossível* ou *incompartível.*

Como exercício, observe que o sistema

$$\begin{cases} x + y = 1, \\ 2x + 2y = 2 \end{cases}$$

possui infinitas soluções, e podemos escrever isso, por exemplo, $x = t$, $y = 1 - t$, $\forall t \in \mathbb{R}$; e que o sistema

$$\begin{cases} x + y = 1 \\ 2x + 2y = -3 \end{cases}$$

não tem solução.

Dado um sistema linear

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ \quad .... \\ a_{m1}x_1 + a_{m_2}x_2 + ... + a_{mn}x_n = b_n \end{cases},$$

podemos associá-lo a uma matriz da forma

$$\begin{pmatrix} a_{11} & a_{12} & ..... & a_{1n} & b_1 \\ a_{21} & a_{22} & ..... & a_{2n} & b_2 \\ ..... & ..... & ..... & ..... & ..... \\ a_{m1} & a_{m2} & ..... & a_{mn} & b_n \end{pmatrix},$$

chamada de *matriz aumentada* do sistema.

Um dos principais objetivos do estudo de matrizes é resolver sistemas lineares. Uma maneira bastante útil de se resolver um sistema linear consiste em efetuar certas operações entre suas equações, com o intuito de obter outras equações equivalentes às originais. Tais operações, são chamadas de operações elementares e consistem em:

(a) multiplicar uma equação por uma constante não nula;

(b) trocar duas equações do sistema de posição;

(c) substituir uma equação pela soma dela com um múltiplo de outra equação do sistema dado.

Tais operações podem ser feitas sobre a matriz aumentada do sistema, ou seja, elas nos motivam definir:

Definição 1.24 Dada uma matriz A, definimos as *operações elementares sobre linhas* de A por:

(i) trocar duas linhas de posição;

(ii) multiplicar uma linha por uma constante não nula;

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

Simbolicamente, chamando ℓ_i e ℓ_j as linhas i e j, $i \neq j$, de uma matriz A, temos que as três operações elementares sobre linhas, respectivamente, são

(i) $\ell_i \leftrightarrow \ell_j$;

(ii) $\ell_i \hookrightarrow k \cdot \ell_i$;

(iii) $\ell_i \hookrightarrow \ell_i + k \cdot \ell_j$.

Isto posto, podemos resolver um exercício, utilizando as ideias acima.

Exemplo 1.25 Resolver o sistema linear $\begin{cases} x + 2y - 4z = -4 \\ 2x + 5y - 9z = -10 \\ 3x - 2y + 3z = 11 \end{cases}$.

Solução. A matriz aumentada do sistema é dada por

$$\begin{pmatrix} 1 & 2 & -4 & -4 \\ 2 & 5 & -9 & -10 \\ 3 & -2 & 3 & 11 \end{pmatrix}$$

Efetuando as operações elementares sobre linhas como nos esquemas a seguir, vamos obter (repare que vamos procurar construir uma "escada" onde os degraus serão pivôs, ou seja, faremos o primeiro elemento não nulo de cada linha da matriz ser igual a 1):

$$\begin{pmatrix} 1 & 2 & -4 & -4 \\ 2 & 5 & -9 & -10 \\ 3 & -2 & 3 & 11 \end{pmatrix} \xrightarrow{\ell_2 \hookrightarrow \ell_2 - 2\ell_1} \begin{pmatrix} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 3 & -2 & 3 & 11 \end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \ell_3 - 3\ell_1}$$

$$\begin{pmatrix} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 0 & -8 & 15 & 23 \end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \ell_3 + 8\ell_2} \begin{pmatrix} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 7 & 7 \end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \frac{1}{7}\ell_3}$$

$$\begin{pmatrix} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

que corresponde ao sistema

$$\begin{cases} x + 2y - 4z = -4 \\ \qquad y - z = -2 \\ \qquad\quad z = 1 \end{cases},$$

o qual, facilmente verificamos que $z = 1$, $y = -1$ e $x = 2$. Logo, a solução dos sistema dado é $(x, y, z) = (-2, -1, 1)$.

Observação 1.26 Poderíamos continuar o procedimento, anulando-se agora os termos acima dos pivôs, ou seja, continuar como segue:

$$\begin{pmatrix} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\ell_2 \hookrightarrow \ell_2 + 8\ell_3} \begin{pmatrix} 1 & 2 & -4 & -4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\ell_1 \hookrightarrow \ell_1 - 2\ell_2}$$

$$\begin{pmatrix} 1 & 0 & -4 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\ell_1 \hookrightarrow \ell_1 + 4\ell_3} \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

ou seja, deixamos a matriz na forma escalonada reduzida por linhas, o que corresponde a

$$\begin{cases} x = 2 \\ y = -1 \\ z = 1 \end{cases}$$

Convém definir:

Definição 1.27 Chama-se *método de Gauss-Jordan* o algoritmo efetuado em um sistema linear de forma a reduzí-lo para a forma de escada, ou seja, quando a matriz aumantada do referido sistema fica sob a forma de uma matriz triangular superior com o primeiro elemento não nulo (pivô) igual a 1.

Nesse caso, quando continuamos o procedimento com o intuito de zerar também acima dos pivôs, deixando a matriz na forma escalonada reduzida por linhas, o método chama-se *eliminação gaussiana*.

Exemplo 1.28 Resolver o sistema

$$\begin{cases} 2x + 5y - 8z = 4 \\ x + 2y - 3z = 1 \\ 3x + 8y - 13z = 7 \end{cases}$$

pelo método de eliminação gaussiana.

Solução. A matriz aumentada do sistema dado é

$$\begin{pmatrix} 2 & 5 & -8 & 4 \\ 1 & 2 & -3 & 1 \\ 3 & 8 & -13 & 7 \end{pmatrix}$$

Vamos efetuar as operações elementares sobre linhas como segue. Observe que como queremos pivôs iguais a 1, a primeira operação elementar a fazer será permutar as linhas 1 e 2 da matriz aumentada. Assim,

$$\begin{pmatrix} 2 & 5 & -8 & 4 \\ 1 & 2 & -3 & 1 \\ 3 & 8 & -13 & 7 \end{pmatrix} \xrightarrow{\ell_1 \leftrightarrow \ell_2} \begin{pmatrix} 1 & 2 & -3 & 1 \\ 2 & 5 & -8 & 4 \\ 3 & 8 & -13 & 7 \end{pmatrix} \xrightarrow{\ell_2 \hookrightarrow \ell_2 - 2\ell_1} \begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 2 \\ 3 & 8 & -13 & 7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 2 \\ 3 & 8 & -13 & 7 \end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \ell_3 - 3\ell_1} \begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 2 \\ 0 & 2 & -2 & 4 \end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \ell_3 - 2\ell_2} \begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Note que uma linha da matriz zerou. Isso indica que o sistema dado é indeterminado, ou seja, possui infinitas soluções. Continuando o processo para deixar a matriz aumentada na forma escalonada reduzida por linhas, obtemos

$$\begin{pmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\ell_1 \hookrightarrow \ell_1 - 2\ell_2} \begin{pmatrix} 1 & 0 & 1 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

que corresponde ao sistema

$$\begin{cases} x \quad\quad + z = -3 \\ \quad\quad y - 2z = 2 \end{cases},$$

e chamando $z = t$, obtemos $y = 2 + 2t$ e $x = -3 - t$, ou seja, as soluções do sistema dado são

$$(x, y, z) = (-3 - t, 2 + 2t, t),\ t \in \mathbb{R}.$$

Exercícios

1. Considere os sistemas lineares

$$\begin{cases} x + y + z = 1 \\ 2x + 2y + 2z = 4 \end{cases} \quad \text{e} \quad \begin{cases} x + y + z = 0 \\ 2x + 2y + 2z = 0 \end{cases}$$

   (a) Mostre que o primeiro sistema não possui soluções e escreva o que isso significa quanto aos planos representados por essas equações.

(b) Mostre que o segundo sistema tem uma infinidade de soluções e escreva o que isso significa quanto aos plannos representados por essas equações.

2. Resolva cada sistema linear abaixo, pelo método de eliminação de Gauss-Jordan:

(a) $\begin{cases} x_1 + x_2 + 2x_3 = 8 \\ -x_1 - 2x_2 + 3x_3 = 1 \\ 3x_1 - 7x_2 + 4x_3 = 10 \end{cases}$ (b) $\begin{cases} 2x_1 + 2x_2 + 2x_3 = 0 \\ -2x_1 + 5x_2 + 2x_3 = 1 \\ 8x_1 + x_2 + 4x_3 = -1 \end{cases}$

3. Verifique se o sistema homogêneo abaixo possui uma solução não nula:

$$\begin{cases} x + y - z = 0 \\ 2x + 4y - z = 0 \\ 3x + 2y + 2z = 0 \end{cases} .$$

4. Resolva cada sistema linear abaixo:

(a) $\begin{cases} x + 2y - 4z = -4 \\ 2x + 5y - 9z = -10 \\ 3x - 2y + 3z = 11 \end{cases}$ (b) $\begin{cases} x + 2y - 3z = -1 \\ -3x + y - 2z = -7 \\ 5x + 3y - 4z = 2 \end{cases}$

(c) $\begin{cases} x + 2y - 3z = 1 \\ 2x + 5y - 8z = 4 \\ 3x + 8y - 13z = 7 \end{cases}$

1.2.1 Matrizes elementares

Definição 1.29 Dizemos que uma matriz quadrada de ordem n é *elementar* quando ela resulta de apenas uma operação elementar sobre linhas da matriz identidade I_n.

Por exemplo, as matrizes

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{e} \quad \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

são matrizes elementares. Por quê?

A importância das matrizes elementares é que as mesmas são usadas para efetuar operações elementares sobre linhas de uma matriz via multiplicação de matrizes. Ou seja, temos o seguinte resultado.

Proposição 1.30 *Sejam $A_{m\times n}$ uma matriz e denote por $E_{m\times m}$ a matriz elementar obtida por uma operação elementar sobre linhas da matriz I_m. Então, o produto $E\cdot A$ é a matriz que resulta em efetuarmos a mesma operação elementar sobre linhas em A.*

Demonstração. Fica como exercício.

□

Esta proposição não será usada na prática, será apenas uma ferramenta para fins teóricos. Isto porque podemos, na prática, efetuar as operações elementares sobre linhas diretamente na matriz desejada, sem precisar recorrer a produtos de matrizes elementares.

Note que, ao efetuar uma operação elementar sobre linhas em uma matriz identidade I, determinando assim uma matriz elementar E, podemos aplicar em cima de E uma operação elementar "inversa" à anterior a fim de recuperar a matriz identidade. Por exemplo, se temos I_3 e multiplicamos a linha 3 por 5, vamos obter a matriz elementar

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Agora, se multiplicarmos a linha 3 de tal matriz por $\frac{1}{5}$ vamos recuperar I_3.

Em geral, para cada operação elementar sobre linhas existe uma operação elementar inversa, conforme a tabela:

Op. sobre linhas de I que produzem E	Op. sobre linhas de E que produzem I
$\ell_i \hookrightarrow k\cdot\ell_i$	$\ell_i \hookrightarrow \frac{1}{k}\cdot\ell_i$
$\ell_i \leftrightarrow \ell_j$	$\ell_i \leftrightarrow \ell_j$
$\ell_i \hookrightarrow \ell_i + k\cdot\ell_j$	$\ell_i \hookrightarrow \ell_i - k\cdot\ell_j$

Proposição 1.31 *Uma matriz elementar é invertível e a sua inversa é uma matriz elementar.*

Demonstração. Seja E uma matriz elementar, que é obtida por uma única operação elementar sobre linhas de I.

Seja E_0 a matriz elementar obtida de I ao se aplicar a operação elementar inversa àquela que produziu E. Pela Proposição 1.30, e observando que operações elementares inversas se cancelam entre si, determinamos que

$$E_0 \cdot E = I \quad \text{e} \quad E \cdot E_0 = I,$$

o que mostra que a matriz elementar E é inversível e além disso que sua inversa também é uma matriz elementar.

□

Teorema 1.32 *Seja A_n uma matriz quadrada. São equivalentes as afirmações:*

(a) *A forma escalonada reduzida por linhas de A_n é I_n.*

(b) *A_n pode ser expressa como o produto de matrizes elementares.*

(c) *A_n é inversível.*

Demonstração.
(a) $\Rightarrow$ (b): Suponha que a forma escalonada reduzida por linhas de A_n seja I_n. Logo, segue que existe uma sequência de operações elementares sobre linhas que reduz A a I_n. Pela Proposição 1.30 segue que cada operação elementar sobre linhas pode ser efetuada multiplicando-se à esquerda da matriz A uma matriz elementar. Dessa forma, temos que existem $E_1, E_2, ..., E_k$ matrizes elementares tais que

$$E_k \cdot ... \cdot E_2 \cdot E_1 \cdot A = I_n.$$

Como pela Proposição 1.31 cada uma das matrizes elementares é invertível, denotando por E_j^{-1} a inversa de cada E_j, $j = 1, 2, ..., k$, multiplicando a igualdade acima, à esquerda, sucessivamente por $E_k{-1}$, E_{k-1}^{-1}, ..., E_2^{-1}, E_1^{-1}, obtemos

$$(E_1^{-1} \cdot E_2{-1} \cdot ... \cdot E_k^{-1})(E_k \cdot ... \cdot E_2 \cdot E_1 \cdot A) = (E_1^{-1} \cdot E_2{-1} \cdot ... \cdot E_k^{-1})I_n,$$

o que, associando dois a dois adequadamente, obtemos

$$A = E_1^{-1} \cdot E_2{-1} \cdot ... \cdot E_k^{-1},$$

e como pela mesma Proposição 1.31 as matrizes da direita são todas matrizes elementares, segue que A é expressa como um produto de matrizes elementares, ou seja, vale (b).

(b) $\Rightarrow$ (c): Suponha que A_n pode ser expressa como o produto de matrizes elementares. Dessa forma, seja

$$A = E_1 \cdot E_2 \cdot ... \cdot E_k,$$

onde E_j é matriz elementar, para todo $j \in \{1, 2, ..., k\}$. Como toda matriz elementar é invertível (Proposição 1.31), e o produto de matrizes invertíveis é inversível, segue que A é inversível, provando (c).

(c) $\Rightarrow$ (a): Suponha A inversível. Seja M a matriz na forma escalonada reduzida por linhas de A. Precisamos mostrar que $M = I$. Sejam E_1,..., E_k matrizes elementares tais que

$$E_k \cdot ... \cdot E_2 \cdot E_1 \cdot A = M.$$

Como A e todos E_j's são invertíveis, segue que M, o seu produto, também é inversível.

Como M também é a forma escalonada reduzida por linhas, M não pode ter uma linha nula, por ser inversível. Logo, somos obrigados a concluir que $M = I_n$. Logo, vale (a).

$\square$

Definição 1.33 Dizemos que duas matrizes A e B são *equivalentes*, e escrevemos $A \sim B$, se existir uma sequência de operações elementares que reduz A a B.

Como cada operação elementar sobre linhas está associada a uma matriz elementar, pela Proposição 1.30 segue que

$$A \sim B \Leftrightarrow \exists \text{ matrizes elementares } E_1, ..., E_k \text{ tais que } B = E_k \cdot ... \cdot E_1 \cdot A.$$

Assim, conforme o Teorema 1.32, dizemos que uma matriz A é invertível se, e somente se, $A \sim I$.

De posse da Definição 1.33 e do Teorema 1.32, escrevemos

Teorema 1.34 *Sejam A e B duas matrizes quadradas de mesmo tamanho. São equivalentes:*

(a) $A \sim B$.

(b) existe uma matriz U inversível tal que $B = U \cdot A$.

(c) existe uma matriz V inversível tal que $A = V \cdot B$.

1.2.2 Algoritmo de inversão de matrizes

O Teorema 1.32 discutido na seção anterior, bem como demais resultados, nos fornece um método eficaz para obtermos a inversa de uma matriz quadrada A, quando esta existir.

De fato, suponha que $A \sim I$. Então existem E_1, E_2,..., E_k uma sequência de matrizes elementares tais que

$$E_k \cdot ... \cdot E_2 \cdot E_1 \cdot A = I. \tag{1.5}$$

Como cada matriz elementar E_j possui inversa E_j^{-1}, que também é elementar, segue da igualdade acima que

$$A = E_1^{-1} \cdot E_2^{-1} \cdot ... \cdot E_k^{-1}.$$

Supondo que A seja inversível, tomando a inversa da igualdade acima, otemos

$$A^{-1} = (E_1^{-1} \cdot E_2^{-1} \cdot ... \cdot E_k^{-1})^{-1} = (E_k^{-1})^{-1} \cdot ... \cdot (E_2^{-1})^{-1} \cdot (E_1^{-1})^{-1} = E_k \cdot ... \cdot E_2 \cdot E_1 \cdot I.$$

ou seja

$$A^{-1} = E_k \cdot ... \cdot E_2 \cdot E_1 \cdot I. \tag{1.6}$$

Observando (1.5) e (1.6) notamos que, ao efetuar uma sequência de operações elementares sobre linhas em A para reduzí-la a I, segue que efetuando a mesma sequência de operações elementares sobre linhas em I, na mesma ordem, obteremos A^{-1}. Ou seja, provamos o seguinte algoritmo de obtenção da inversa de uma metriz quadrada A:

Teorema 1.35 *(Algoritmo para obtenção de inversa) Dada uma matriz quadrada A. Se efetuarmos uma sequência de operações elementares sobre linhas que reduza A à matriz identidade I, então essa mesma sequência de operações elementares sobre linhas, na mesma ordem, reduz I à A^{-1}.*

Na prática, para determinar A^{-1}, costuma-se escrever uma matriz aumentada $(A|I)$, formada por dois grandes "blocos" A e I, e com isso, quando efetuarmos operações elementares sobre linhas para reduzir o primeiro bloco A em I, automaticamente o segundo bloco I será transformado em A^{-1}, ou seja, simbolicamente, temos

$$(A|I) \underbrace{\dashrightarrow}_{E_k \cdot \ldots \cdot E_2 \cdot E_1} (I|A^{-1})$$

Para ilustrar esse algoritmo, vejamos um exemplo prático:

Exemplo 1.36 Obtenha a inversa da matriz

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & -1 & 0 \end{pmatrix},$$

se existir.

Solução. A matriz aumentada é dada por

$$(A\,|\,I) = \begin{pmatrix} 2 & 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Efetuaremos operações elementares sobre linhas nessa matriz até conseguirmos obter I no primeiro bloco, daí teremos A^{-1} no segundo bloco. Assim,

$$\begin{pmatrix} 2 & 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\ell_1 \leftrightarrow \ell_2} \begin{pmatrix} 1 & 2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\ell_2 \hookrightarrow \ell_2 - 2\ell_1}$$

$$\begin{pmatrix}1 & 2 & 0 & 0 & 1 & 0\\ 0 & -2 & 1 & 1 & -2 & 0\\ 1 & -1 & 0 & 0 & 0 & 1\end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \ell_3 - \ell_1} \begin{pmatrix}1 & 2 & 0 & 0 & 1 & 0\\ 0 & -2 & 1 & 1 & -2 & 0\\ 0 & -3 & 0 & 0 & -1 & 1\end{pmatrix} \xrightarrow{\ell_2 \leftrightarrow \ell_3}$$

$$\begin{pmatrix}1 & 2 & 0 & 0 & 1 & 0\\ 0 & -3 & 0 & 0 & -1 & 1\\ 0 & -2 & 1 & 1 & -2 & 0\end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow -\frac{1}{3}\ell_3} \begin{pmatrix}1 & 2 & 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0 & \frac{1}{3} & -\frac{1}{3}\\ 0 & -2 & 1 & 1 & -2 & 0\end{pmatrix} \xrightarrow{\ell_3 \hookrightarrow \ell_3 + 2\ell_2}$$

$$\begin{pmatrix}1 & 2 & 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0 & \frac{1}{3} & -\frac{1}{3}\\ 0 & 0 & 1 & 1 & -\frac{4}{3} & -\frac{2}{3}\end{pmatrix} \xrightarrow{\ell_1 \hookrightarrow \ell_1 - 2\ell_2} \begin{pmatrix}1 & 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3}\\ 0 & 1 & 0 & 0 & \frac{1}{3} & -\frac{1}{3}\\ 0 & 0 & 1 & 1 & -\frac{4}{3} & -\frac{2}{3}\end{pmatrix},$$

logo, concluímos que

$$A^{-1} = \begin{pmatrix}0 & \frac{1}{3} & \frac{2}{3}\\ 0 & \frac{1}{3} & -\frac{1}{3}\\ 1 & -\frac{4}{3} & -\frac{2}{3}\end{pmatrix}.$$

Como um exercício, o leitor pode verificar que

$$A \cdot A^{-1} = I_3 \ \text{ e } \ A^{-1} \cdot A = I_3.$$

Observação 1.37 Se durante o processo acima no primeiro bloco onde a matriz A deve ser transformada em I_3, se alguma linha se tornasse nula, isso indicaria que a forma escalonada reduzida por linhas de A não é a identidade, ou seja, que A não seria inversível. Por exemplo, na tentativa de encontrar uma inversa para

$$B = \begin{pmatrix}2 & 2 & 1\\ 1 & 0 & 3\\ 3 & 2 & 4\end{pmatrix},$$

veremos que uma linha do bloco onde deve aparecer I_3 obteremos uma linha nula, o que indica que B não é inversível.

Exercícios

1. Seja A uma matriz 3×3. Encontre uma matriz B para a qual BA é a matriz que resulta de A permutando as duas primeiras linhas e depois multiplicando a terceira linha por seis.

2. Considere a matriz $A = \begin{pmatrix} 1 & 0 \\ -5 & 2 \end{pmatrix}$

   (a) Encontre matrizes elementares E_1 e E_2 tais que $E_2E_1A = I$.

   (b) Escreva A^{-1} como um produto de matrizes elementares.

3. Sejam e_1, e_2 e e_3, respectivamente, as operações sobre linhas:
   "Trocar as linhas ℓ_1 e ℓ_2"; "Substituir ℓ_3 por $7\ell_3$" e *" Substituir ℓ_2 por $-3\ell_1 + \ell_2$"*.
   Determine as matrizes quadradas elementares de ordem 3 correspondentes E_1, E_2 e E_3.

4. Reduza cada uma das matrizes abaixo à forma canônica escalonada reduzida por linhas.

   (a) $A = \begin{pmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 8 & 8 & -1 & 26 & 19 \end{pmatrix}$ (b) $B = \begin{pmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{pmatrix}$

5. Mostre que se $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & 1 \end{pmatrix}$ é uma matriz elementar, então $ab = 0$.

6. Sejam A, B e C matrizes $n \times n$ tais que $A = BC$. Prove que se B é invertível, então qualquer sequência de operações elementares sobre as linhas que reduz B a I_n, também reduz A a C.

7. Utilizando-se do algoritmo para obter inversas de matrizes, encontre a inversa A^{-1} de A em cada caso, se A for inversível.

   (a) $A = \begin{pmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{pmatrix}$ (b) $A = \begin{pmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & 4 \end{pmatrix}$

   (c) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$ (d) $A = \begin{pmatrix} -1 & 3 & -4 \\ 2 & 4 & 1 \\ -4 & 2 & 9 \end{pmatrix}$

8. Em que condições a matriz diagonal $A = \begin{pmatrix} a_1 & 0 & ... & 0 \\ 0 & a_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & a_n \end{pmatrix}$ é inversível e qual é sua inversa?

9. Considere o sistema linear a seguir:

$$\begin{cases} x + y + z = 6 \\ x - y + 2z = 2 \\ 2x + y - z = 3 \end{cases} \tag{1.7}$$

   (a) Resolva-o aplicando operações elementares sobre linhas.

   (b) Reescrevendo o sistema (1) na notação matricial

$$Ax = b\,, \tag{1.8}$$

   podemos encontrar o valor da matriz-solução x, multiplicando (2) à esquerda por A^{-1}, se esta inversa existir. Encontre A^{-1} e obtenha o valor de x dessa forma. Compare sua resposta com a obtida no item anterior.

Capítulo 2

Determinantes

Neste capítulo vamos definir o importante conceito de determinante de uma matriz quadrada, que será muito útil para estabelecer um outro critério para analisar se uma dada matriz quadrada é inversível ou não. O estudo de determinantes também é importante para resolver sistemas lineares, além de outras aplicações que serão estudadas num curso de Cálculo.

2.1 Conceito

Dada uma matriz quadrada A, vamos associá-la a um número real, o qual chamaremos de *determinante* da matriz.

Para motivar esse novo conceito, o qual servirá para nos dizer se tal matriz é inversível ou não, vamos considerar vários casos, do mais simples ao mais geral.

Caso 1. *Quando A tem tamanho* 1×1. Nesse caso, uma matriz $A = (a_{11})$ é formada apenas por um elemento, ou seja, é apenas um número real. Como um número real possui um inverso multiplicativo se ele for diferente de zero, essa matriz A será inversível se, e somente se, $a_{11} \neq 0$, ou seja, existirá uma matriz $A^{-1} = (\frac{1}{a_{11}})$ tal que

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

Dessa forma, podemos definir o determinante de A, denotado por $\det A$, pondo

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

e nesse caso temos que a matriz A será inversível se, e somente se $\det A \neq 0$.

Vamos mostrar que essa propriedade também valerá para matrizes $n \times n$.

Caso 2. *Quando A tem tamanho* 2×2. Nesse caso, a forma geral da matriz A é

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Vamos supor que $a_{11} \neq 0$. Pelo estudo do capítulo anterior, temos que A será inversível se, e somente se, ela for equivalente à matriz identidade I_2, ou seja, se, e só se, a forma escalonada reduzida por linhas de A resultar em I_2.

Dessa forma, como estamos supondo $a_{11} \neq 0$, podemos efetuar a seguinte operação elementar sobre linhas em A:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{\ell_2 \hookrightarrow \ell_2 - \frac{a_{21}}{a_{11}}\ell_1} \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} \cdot a_{12} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ 0 & \frac{a_{11} \cdot a_{22} - a_{21} \cdot a_{12}}{a_{11}} \end{pmatrix},$$

e disso, temos que, para a matriz A ser equivalente à matriz identidade I_2, obrigatoriamente devemos impor que

$$a_{11} \cdot a_{22} - a_{21} \cdot a_{12} \neq 0.$$

Assim, definimos o determinante da matriz $A_{2\times 2}$ por

$$\det A = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}.$$

Uma maneira simbólica de representar o determinante de uma matriz

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

de ordem 2, é escrever

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}.$$

Dessa forma, uma matriz $A_{2\times 2}$ será inversível se, e somente se, for equivalente à I_2 , o que é equivalente a dizer que $\det A \neq 0$.

Caso 03. *Quando A tem tamanho* 3×3. Novamente, A será inversível se, e somente se, $A \sim I_3$, onde

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Vamos considerar que $a_{11} \neq 0$ (se for igual a zero trocamos por uma linha onde o primeiro elemento seja diferente de zero), assim, efetuando as operações elementares sobre linhas

$$\ell_2 \hookrightarrow \ell_2 - \frac{a_{21}}{a_{11}} \cdot \ell_1 \ \text{ e } \ \ell_3 \hookrightarrow \ell_3 - \frac{a_{31}}{a_{11}} \cdot \ell_1,$$

vamos obter de A a matriz

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \frac{a_{11}\cdot a_{22} - a_{21}\cdot a_{12}}{a_{11}} & \frac{a_{11}\cdot a_{23} - a_{21}\cdot a_{13}}{a_{11}} \\ 0 & \frac{a_{11}\cdot a_{32} - a_{31}\cdot a_{12}}{a_{11}} & \frac{a_{11}\cdot a_{33} - a_{31}\cdot a_{13}}{a_{11}} \end{pmatrix}$$

Observe que A será inversível se, e somente se, nenhuma das linhas da matriz acima for nula, ou seja, se, e só se,

$$a_{11} \cdot \begin{vmatrix} \frac{a_{11}\cdot a_{22} - a_{21}\cdot a_{12}}{a_{11}} & \frac{a_{11}\cdot a_{23} - a_{21}\cdot a_{13}}{a_{11}} \\ \frac{a_{11}\cdot a_{32} - a_{31}\cdot a_{12}}{a_{11}} & \frac{a_{11}\cdot a_{33} - a_{31}\cdot a_{13}}{a_{11}} \end{vmatrix} \neq 0,$$

ou seja, se, e só se,

$$(a_{22} \cdot a_{11} - a_{21} \cdot a_{12})(a_{11} \cdot a_{33} - a_{31} \cdot a_{13}) -$$
$$- (a_{11} \cdot a_{23} - a_{21} \cdot a_{13})(a_{11} \cdot a_{32} - a_{31} \cdot a_{12}) \neq 0,$$

que equivale a

$$a_{11}^2 \cdot a_{22} \cdot a_{33} - a_{11} \cdot a_{13} \cdot a_{22} \cdot a_{31} - a_{11} \cdot a_{12} \cdot a_{21} \cdot a_{33} + a_{12} \cdot a_{13} \cdot a_{21} \cdot a_{31} -$$

$$-a_{11}^2 \cdot a_{23} \cdot a_{32} + a_{11} \cdot a_{12} \cdot a_{23} \cdot a_{31} + a_{11} \cdot a_{13} \cdot a_{21} \cdot a_{32} - a_{12} \cdot a_{13} \cdot a_{21} \cdot a_{31} \neq 0,$$

isto é,

$$a_{11}\left(a_{11} \cdot a_{22} \cdot a_{33} - a_{13} \cdot a_{22} \cdot a_{31} - a_{12} \cdot a_{21} \cdot a_{33} - a_{11} \cdot a_{23} \cdot a_{32} + \right.$$

$$\left. + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32}\right) \neq 0,$$

e como $a_{11} \neq 0$, segue do produto acima que

$$a_{11} \cdot a_{22} \cdot a_{33} - a_{13} \cdot a_{22} \cdot a_{31} - a_{12} \cdot a_{21} \cdot a_{33} - a_{11} \cdot a_{23} \cdot a_{32} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} \neq 0.$$

Dessa maneira, para uma matriz $A_{3\times 3}$, definimos

$$\det A = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - a_{13} \cdot a_{22} \cdot a_{31} - a_{12} \cdot a_{21} \cdot a_{33} - a_{11} \cdot a_{23} \cdot a_{32},$$

e temos que A será inversível se, e somente se, $\det A \neq 0$. Uma maneira simbólica de denotar o determinante de $A_{3\times 3}$ é escrever

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Caso geral. Para poder encontrar um caso geral para definir o determinante de uma matriz $A_{n\times n}$, vamos retomar os casos anteriores e fazer algumas considerações. Seja $A_{2\times 2}$ dada por

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Neste caso, podemos definir matrizes menores M_{11} e M_{12} por

$$M_{11} = (a_{22}) \;\text{ e }\; M_{12} = (a_{21}),$$

onde M_{11} denota a matriz que se obtém de A eleminando-se a linha 1 e a coluna 1 de A e M_{12} é a matriz que se obtém de A eliminando-se a primeira linha e a segunda coluna.

Dessa forma, o determinante de A pode ser escrito como

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21} = a_{11} \cdot \det M_{11} - a_{12} \cdot \det M_{12}.$$

Já no caso de uma matriz

$$A_{3\times 3} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

vamos definir

$$M_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}, \ M_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \ \text{ e } \ M_{13} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix},$$

onde

$$\det M_{11} = a_{22} \cdot a_{33} - a_{23} \cdot a_{32}, \quad \det M_{12} = a_{21} \cdot a_{33} - a_{23} \cdot a_{31} \quad \text{e}$$

$$\det M_{13} = a_{21} \cdot a_{32} - a_{22} \cdot a_{31}.$$

Assim,

$$\det A = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - a_{13} \cdot a_{22} \cdot a_{31} - a_{12} \cdot a_{21} \cdot a_{33} - a_{11} \cdot a_{23} \cdot a_{32}$$

$$= a_{11}(a_{22} \cdot a_{33} - a_{23} \cdot a_{32}) - a_{12}(a_{21} \cdot a_{33} - a_{23} \cdot a_{31}) + a_{13}(a_{21} \cdot a_{32} - a_{22} \cdot a_{31}) =$$

$$= a_{11} \cdot \det M_{11} - a_{12} \cdot \det M_{12} + a_{13} \cdot \det M_{13},$$

ou seja, um determinante de uma matriz de ordem 3 pode ser reduzido a uma combinação de três determinantes de ordem 2. Isto nos motiva o seguinte conceito:

Definição 2.1 Seja $A = (a_{ij})_{n\times n}$ uma matriz e M_{ij} uma matriz $(n-1)\times(n-1)$ obtida de A eliminando-se a linha i e a coluna j de A que contém o elemento a_{ij}. O determinante $\det M_{ij}$ é chamado de *menor* de a_{ij}. Definimos também o *cofator* A_{ij} de a_{ij} por

$$A_{ij} = (-1)^{i+j} \cdot \det M_{ij}.$$

Com essa definição, podemos finalmente generalizar a definição de determinante para uma matriz $n \times n$ na seguinte:

Definição 2.2 Dada uma matriz $A_{n\times n}$, definimos o *determinante* de A ao número real

$$\det A = \begin{cases} a_{11} & \text{se } \ n = 1 \\ a_{11} \cdot A_{11} + a_{12} \cdot A_{12} + ... + A_{1n} \cdot A_{1n} & \text{se } \ n > 1, \end{cases}$$

onde

$$A_{1j} = (-1)^{1+j} \cdot \det(M_{1j}),\ j = 1, 2, ..., n.$$

Vejamos um exemplo de cálculo de determinante. Considere a matriz

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 3 & 2 \\ 2 & 1 & 3 \end{pmatrix}.$$

Vamos calcular o seu determinante de duas formas: usando o conceito de $\det A$ para matriz 3×3 e usando a definição geral com cofatores A_{1j}. Assim, usando direto a regra para cálculo de determinante de uma matriz 3×3, obtemos

$$\det A = 1 \cdot 3 \cdot 3 + 2 \cdot 2 \cdot 2 + (-1) \cdot 1 \cdot 1 - (-1) \cdot 3 \cdot 2 - 2 \cdot 1 \cdot 1 - 2 \cdot 1 \cdot 3 =$$

$$= 9 + 8 - 1 + 6 - 2 - 6 = 14.$$

E, usando a definição geral envolvendo cofatores, obtemos,

$$\det A = 1 \cdot A_{11} + 2 \cdot A_{12} - 1 \cdot A_{13} =$$

$$= 1 \cdot (-1)^{1+1} \begin{vmatrix} 3 & 2 \\ 1 & 3 \end{vmatrix} + 2 \cdot (-1)^{1+2} \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} - 1 \cdot (-1)^{1+3} \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} =$$

$$= 1 \cdot (3 \cdot 3 - 2 \cdot 1) - 2(1 \cdot 3 - 2 \cdot 2) - 1(1 \cdot 1 - 3 \cdot 2) = 9 - 2 + 2 + 5 = 14,$$

o mesmo resultado obtido na primeira forma de cálculo.

Observe que, neste exemplo, o cálculo do determinante de uma matriz de ordem 3, usando a definição geral, é mais trabalhoso do que o cálculo direto pela regra para cálculo de determinante de uma matriz 3×3. No entanto, essa regra passa a ser útil quando temos uma matriz quadrada maior, como 4×4, 5×5, etc., pois nesses casos não temos uma regra mais prática para o cálculo a não ser a definição geral.

Observação 2.3 Na definição acima de determinante de uma matriz quadrada usamos a primeira linha para efetuar a expansão dos menores. No entanto, poderíamos usar outra linha ou até mesmo uma coluna para efetuar o cálculo do determinante. Como ilustração, vamos calcular o determinante da matriz $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 3 & 2 \\ 2 & 1 & 3 \end{pmatrix}$ dada como exemplo acima, expandindo os menores pela segunda coluna de A, e veremos que o determinante resultará também em 14:

$$\det A = 2 \cdot A_{12} + 3 \cdot A_{22} + 1 \cdot A_{32} =$$

$$= 2 \cdot (-1)^{1+2} \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 3 \cdot (-1)^{2+2} \begin{vmatrix} 1 & -1 \\ 2 & 3 \end{vmatrix} + 1 \cdot (-1)^{2+3} \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} =$$

$$= -2(1 \cdot 3 - 2 \cdot 2) + 3(1 \cdot 3 - (-1) \cdot 2) - 1(1 \cdot 2 - (-1) \cdot 1) = 2 + 15 - 3 = 14.$$

Exercícios

1. Calcule os seguintes determinantes:

(a) $\begin{vmatrix} 2 & -1 \\ 3 & 5 \end{vmatrix}$ (b) $\begin{vmatrix} 3 & -1 & 2 \\ 4 & -2 & -3 \\ 0 & 1 & 6 \end{vmatrix}$ (c) $\begin{vmatrix} 1 & -1 & 2 & -2 \\ -3 & 2 & -3 & 1 \\ 0 & 1 & 6 & 0 \\ 2 & 1 & 2 & -1 \end{vmatrix}$

2. Mostre que

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = (y - x)(z - x)(z - y).$$

2.2 Propriedades dos determinantes

Nesta seção vamos apresentar as principais propriedades dos determinantes.

Proposição 2.4 *Seja A uma matriz quadrada. Então*

$$\det(A^T) = \det(A).$$

Demonstração. Faremos a prova por indução sobre a ordem da matriz. Começamos com a base da indução $n = 2$:

(i) Quando $n = 2$, seja $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, então temos que

$$\det A = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

Por outro lado, temos que $A^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$, e disso segue que

$$\det A^T = a_{11} \cdot a_{22} - a_{12} \cdot a_{21} = \det A,$$

logo, vale a base da indução.

(ii) Suponha que a Proposição seja verdadeira para todas as matrizes de tamanho $k \times k$, ou seja, que

$$\det M_{k\times k} = \det M^T_{k\times k},\ \forall M_{k\times k}.$$

Seja $A_{(k+1)\times(k+1)}$. Então, calculando o determinante de A pela expansão dos menores pela primeira linha de A, obtemos

$$\det A = \sum_{j=1}^{k+1} (-1)^{1+j} \cdot a_{1j} \cdot \det(M_{1j}),$$

e como M_{1j} é de tamanho $k \times k$, para $j = 1, 2, ..., (k+1)$, segue pela hipótese de indução que

$$\det(M_{1j}) = \det(M^T_{1j}),$$

e daí temos

$$\det A = \sum_{j=1}^{k+1} (-1)^{1+j} \cdot a_{1j} \cdot \det(M^T_{1j}) = \det A^T,$$

pois o somando dessa igualdade corresponde à expansão por menores de $\det A^T$, usando a primeira coluna de A^T.

Logo, por (i) e (ii) a prova segue por indução matemática.

□

Proposição 2.5 *Se A for uma matriz triangular (superior ou inferior), então o determinante de A será simplesmente o produto dos elementos da diagonal principal de A.*

Demonstração. Observe que a transposta de uma matriz triangular inferior é uma matriz triangular superior, e vice-versa. Assim, vamos fazer a prova usando apenas matriz triangular inferior, pois a Proposição anterior garantirá o mesmo resultado para a triangular superior.

Assim, seja $A_{n\times n}$ uma matriz triangular inferior. Faremos a prova por indução sobre a ordem da matriz.

(i) Vejamos o caso $n = 2$ (base da indução). Ou seja, dada $A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix}$, temos que

$$\det A = a_{11} \cdot a_{22}.$$

Logo, vale a base da indução.

(ii) Suponha o resultado válido para toda matriz triangular inferior de ordem $k \times k$. Seja A uma matriz de ordem $(k+1)\times(k+1)$. Então, calculando o determinante de A via determinantes menores a partir da primeira linha de A, obtemos

$$\det A = \sum_{j=1}^{k+1} (-1)^{1+j} \cdot a_{1j} \cdot \det(M_{1j}) = a_{11} \cdot \det(M_{11}),$$

visto que os demais somandos são iguais a zero.

Como M_{11} é também uma matriz triangular inferior de ordem $k \times k$, pela hipótese da indução segue que

$$\det M_{11} = a_{22} \cdot a_{33} \cdot \ldots \cdot a_{(k+1),(k+1)},$$

e daí concluímos que

$$\det A = a_{11} \cdot \det(M_{11}) = a_{11} \cdot a_{22} \cdot ... \cdot a_{(k+1),(k+1)}.$$

Portanto, por (i) e (ii) o resultado segue por indução.

$\square$

Proposição 2.6 *Seja A uma matriz quadrada de ordem n. Então;*

(a) *se uma linha de A for nula, então* $\det A = 0$.

(b) *se A possuir duas linhas ou duas colunas iguais, então* $\det A = 0$.

Demonstração. Fica como exercício para o leitor, bastando usar a expansão em cofatores.

$\square$

No que segue, apresentaremos o que ocorre com o determinante de uma matriz quadrada A quando efetuarmos uma operação elementar sobre linhas em A. Antes, porém, enunciemos um lema básico.

Lema 2.7 *Seja $A_{n\times n}$ uma matriz quadrada. Se $A_{\ell k}$ for o cofator de $a_{\ell k}$, então*

$$a_{i1} \cdot A_{\ell 1} + a_{i2} \cdot A_{\ell 2} + ... + a_{in} \cdot A_{\ell n} = \begin{cases} \det A & , \quad se \quad i = \ell \\ 0 & , \quad se \quad i \neq \ell \end{cases}.$$

Demonstração. Se $i = \ell$ temos exatamente a definição de determinante de A. Vamos então considerar o caso $i \neq \ell$. Neste caso, se na matriz $A_{n\times n}$ substituirmos a linha ℓ pela linha i, vamos obter uma matriz $\tilde{A}$ que possui duas linhas iguais, a saber, as linhas i e ℓ:

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & ..... & a_{1n} \\ ..... & ..... & ..... & ..... \\ a_{i1} & a_{i2} & ..... & a_{in} \\ ..... & ..... & ..... & ..... \\ a_{i1} & a_{i2} & ..... & a_{in} \\ ..... & ..... & ..... & ..... \\ a_{n1} & a_{n2} & ..... & a_{nn} \end{pmatrix}$$

Assim, como $\tilde{A}$ possui duas linhas iguais, pela Proposição 2.6-(b) segue que $\det \tilde{A} = 0$, e então, em particular, desenvolvendo o determinante pelos cofatores na linha ℓ, temos

$$0 = \det \tilde{A} = a_{i1} \cdot \tilde{A}_{\ell 1} + ... + a_{in} \cdot \tilde{A}_{\ell n},$$

e observando que $\tilde{A}_{ik} = A_{ik}$, para $k = 1, 2, ..., n$, concluímos que (lembre que $i \neq \ell$)

$$a_{i1} \cdot A_{\ell 1} + ... + a_{in} \cdot A_{\ell n} = a_{i1} \cdot \tilde{A}_{\ell 1} + ... + a_{in} \cdot \tilde{A}_{\ell n} = 0,$$

como desejado.

□

Assim, podemos enunciar a proposição que encerra os resultados que acontecem com o determinante de uma matriz quando efetuamos uma das três operações sobre linhas em A:

Proposição 2.8 *Seja $A_{n\times n}$ uma matriz quadrada cujo determinante é* $\det A$. *Então, valem as propriedades:*

(a) *Se trocarmos duas linhas quaisquer de A, segue que o determinante de A muda de sinal. Simbolicamente, se $\ell_i \leftrightarrow \ell_j$, então o determinante dessa nova matriz será* $-\det A$.

(b) *Se multiplicarmos uma linha de A por uma constante $k \neq 0$, então o determinante dessa nova matriz será* $k \cdot \det A$.

(c) *Se substituirmos uma linha de A pela soma dela mesma mais um múltiplo de outra linha, o determinante dessa nova matriz continuará sendo* $\det A$. *Simbolicamente, se $\ell_i \hookrightarrow \ell_i + k \cdot \ell_j$, onde $k \neq 0$, então o determinante dessa nova matriz ainda será* $\det A$.

Demonstração.
(a) Esse é o caso mais difícil de verificar e vamos fazer apenas o caso $A = (a_{ij})_{2\times 2}$. Assim, a operação elementar $\ell_1 \leftrightarrow \ell_2$ associa-se à matriz elementar $E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, e então como $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, temos que $\det A = a_{11} \cdot a_{22} -$

$a_{12} \cdot a_{21}$. Note que trocar de posição as linhas 1 e 2 corresponde a multiplicar A à esquerda por E, ou seja, teremos

$$EA = \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix},$$

donde segue que

$$\det(EA) = a_{12} \cdot a_{21} - a_{11} \cdot a_{22} = -\det(A).$$

Note também que $\det(E) = -1$, e então concluímos também que

$$\det(EA) = -1 \cdot \det(A) = \det(E) \cdot \det(A). \tag{2.1}$$

Se trabalharmos com matrizes 3×3 também concluiremos os mesmos resultados, em geral, o caso $n \times n$.

(b) Dada $A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}$.

Fixemos uma linha i_0 qualquer e multipliquemos a mesma por $k \neq 0$, ou seja, faremos a seguinte operação elementar sobre linhas em A: $\ell_{i_o} \hookrightarrow k \cdot \ell_{i_0}$, que corresponde à matriz elementar

$$E = (e_{ij})_{n \times n}, \quad \text{onde} \quad \begin{cases} e_{i_0 i_0} = k \\ e_{ij} = 1, \ \text{se} \quad i = j \ \text{ e } \ i \neq i_0 \\ e_{ij} = 0, \ \text{se} \quad i \neq j \end{cases}.$$

Note que E é uma matriz diagonal diag $(1, 1, ..., 1, k, 1, ...1)$, que assume o valor k na posição $i_0 i_0$, logo, é uma matriz triangular, donde segue que $\det(E) = k$.

A ação dessa operação elementar na linha i_0 de A corresponde a multiplicar A à esquerda por E, ou seja, corresponde á matriz

$$EA = \begin{pmatrix} a_{11} & .... & a_{1n} \\ ... & ... & ... \\ ka_{i_0 1} & ... & ka_{i_0 n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}.$$

Calculando o $\det(EA)$ expandindo por cofatores na i_0-ésima linha, obtemos

$$\det(EA) = k \cdot a_{i_0 1}(EA)_{i_0 1} + ... + k \cdot a_{i_0 n}(EA)_{i_0 n} = k \cdot \left(\sum_{p=1}^{n} a_{i_0 p}(EA)_{i_0 p}\right),$$

e como $(EA)_{i_0 j} = (-1)^{i_0+j} \cdot \det(M_{i_0 j}) = A_{i_0 j}$, segue que

$$\det(EA) = k \cdot \left(\sum_{p=1}^{n} a_{i_0 p}(A)_{i_0 p}\right) = k \cdot \det(A),$$

o que conclui (b).

Note que, neste caso, temos também que

$$\det(EA) = k \cdot \det(A) = \det(E) \cdot \det(A). \tag{2.2}$$

(c) Novamente, dada a matriz Dada $A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}$, vamos efetuar a seguinte operação elementar sobre linhas em A: $\ell_i \hookrightarrow \ell_i + k \cdot \ell_j$. Assim, a matriz elementar E correspondente a tal operação sobre linhas corresponde a uma matriz triangular (superior ou inferior, dependendo se $i > j$ ou $i < j$). De qualquer modo, os elementos da diagonal serão todos 1 e daí, segue que

$$\det(E) = 1.$$

Do mesmo modo, a ação dessa mesma operação elementar sobre linhas na matriz A corresponde a multiplicar A à equerda pela matriz elementar E, ou seja,

$$EA = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{i1} + k \cdot a_{j1} & ... & a_{in} + k \cdot a_{jn} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}.$$

Dessa forma, calculando o determinante de EA desenvolvendo por cofatores na linha i, vamos obter

$$\det(EA) = \sum_{p=1}^{n} (a_{ip} + k \cdot a_{jp})(EA)_{ip},$$

onde $(EA)_{ip}$ denota o cofator do elemento $(ea)_{ip}$ da matriz EA. Como

$$(EA)_{ip} = (-1)^{i+p}\det(M_{ip}) = A_{ip},$$

segue que, com auxílio do Lema 2.7, obtemos

$$\det(EA) = \sum_{p=1}^{n}(a_{ip} + k \cdot a_{jp})A_{ip} = \sum_{p=1}^{n} a_{ip}A_{ip} + k \cdot \sum_{p=1}^{n} a_{jp}A_{ip} = \det(A) + k \cdot 0,$$

ou seja, $\det(EA) = \det(A)$, provando (c). Note que, neste caso também temos a propriedade

$$\det(EA) = 1 \cdot \det(A) = \det(E) \cdot \det(A). \tag{2.3}$$

□

Observando a prova da Proposição anterior em (2.1), (2.2) e (2.3), percebemos que, dada uma matriz quadrada A e uma matriz elementar E qualquer, de mesmo tamanho, então

$$\det(E \cdot A) = \det(E) \cdot \det(A). \tag{2.4}$$

Essa propriedade nos inspira o seguinte resultado geral:

Proposição 2.9 *Sejam A e B duas matrizes quadradas de mesmo tamanho. Então*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Demonstração. Vamos considerar dois casos:
(i) Se A for inversível[1] - Então, nesse caso, temos que a forma escalonada reduzida por linhas de A é I_n, ou seja, existe uma sequência finita de operações elementares que reduzem A em I, e como a cada operação elementar está associada a uma matriz elementar, segue que existem $E_1, ..., E_k$ matrizes elementares tais que

$$A = E_k \cdot ... \cdot E_1 \cdot I.$$

Assim, de (2.4), segue que

$$\det(A \cdot B) = \det((E_k \cdot ... \cdot E_1) \cdot B) = \det(E_k \cdot (E_{k-1} \cdot ... \cdot E_1 \cdot B)) =$$

[1] Poderíamos supor B inversível e concluiríamos o mesmo.

$$= \det(E_k) \cdot \det(E_{k-1} \cdot ... \cdot E_1 \cdot B) = ... = \det(E_k) \cdot \det(E_{k-1}) \cdot ... \cdot \det(E_1) \cdot \det(B) =$$

$$= \det(E_k) \cdot ... \cdot \det(E_3) \cdot \det(E_2 \cdot E_1) \cdot \det(B) = ...$$

$$... = \det(E_k \cdot ... \cdot E_1) \cdot \det(B) = \det(A) \cdot (B).$$

(ii) Se A não é inversível[2] - Neste caso, a forma escalonada reduzida por linhas de A possui pelo menos uma linha nula, e daí $\det(A) = 0$. Do mesmo modo, o produto $A \cdot B$ resultará numa matriz com uma linha também nula (verifique!) e daí segue que $\det(A \cdot B) = 0$. Ou seja, temos nesse caso que

$$\det(A \cdot B) = 0 = \det(A) = \det(A) \cdot \det(B).$$

Portanto, independente do caso, segue o resultado.

$\square$

Exercícios

1. Sem calcular o determinante, justifique por que $x = 0$ e $x = 2$ satisfazem a equação
$$\begin{vmatrix} x^2 & x & 2 \\ 2 & 1 & 1 \\ 0 & 0 & -5 \end{vmatrix} = 0.$$

2. Sejam A e B duas matrizes $n \times n$. Sabemos que o produto de matrizes, em geral, não comuta, ou seja, em geral tem-se que $A \cdot B \neq B \cdot A$. Isso vale também para $\det(A \cdot B)$ e $\det(B \cdot A)$? Justifique.

3. Se $\det M \neq 0$ e $MN = MP$, mostre que $N = P$.

4. Seja A uma matriz inversível com inversa A^{-1}. Mostre que
$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

5. Sejam A uma matriz $n \times n$ e α um escalar. Mostre que
$$\det(\alpha \cdot A) = \alpha^n \cdot \det(A).$$

[2] Se supormos que B não é inversível segue o mesmo resultado.

6. Se A é uma matriz 3×3 tal que $\det A = 8$, determine o valor de $\det(3A)$, justificando sua resposta.

7. Uma outra maneira de definir que duas matrizes A e B são *semelhantes* é se existir uma matriz inversível P tal que $B = P^{-1} \cdot A \cdot P$. Mostre que se A e B são semelhantes, então $\det A = \det B$.

8. Se A for uma matriz idempotente, ou seja, tal que $A^2 = A$, quanto vale $\det A$?

9. Dizemos que uma matriz quadrada A é *ortogonal* se $A \cdot A^T = I$. Dessa forma, se A for uma matriz ortogonal, mostre que $\det A = \pm 1$.

10. Se $A^T B^T = \begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix}$ e $\det B = 2$, quanto vale $\det A$?

11. Se $A \cdot M = M \cdot B$, onde M é inversível, mostre que, para qualquer escalar λ, vale
$$\det(A - \lambda I) = \det(B - \lambda I).$$

2.3 Matriz adjunta e a regra de Cramer

Nesta seção vamos definir um tipo especial de matriz, chamada de matriz adjunta, à qual será crucial para desenvolvermos uma outra técnica para obter a inversa de uma matriz inversível A, bem como provar um Teorema importante de resolução de sistemas lineares, conhecido por *regra de Cramer*.

Definição 2.10 Dada uma matriz quadrada $A = (a_{ij})_{n \times n}$, definimos a *matriz adjunta* de A, e denotamos por $\operatorname{adj} A$, a matriz

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

onde $A_{ij} = (-1)^{i+j} \cdot \det(M_{ij})$ é o cofator do elemento a_{ij} da matriz A.

É importante observar que a matriz adjunta de uma matriz quadrada A, é definida como sendo uma matriz dos cofatores da matriz A, mas perceba que dispomos os respectivos cofatores A_{ij} na ordem transposta ao elemento a_{ij} respectivo da matriz A, ou seja, o cofator A_{ij} de a_{ij} corresponde ao elemento da linha j e coluna i de $\operatorname{adj} A$. Isso foi definido propositalmente dessa maneira para cumprir a seguinte propriedade:

$$A \cdot \operatorname{adj} A = (\det A) \cdot I. \tag{2.5}$$

De fato, basta efetuar os cálculos, observando o Lema 2.7:

$$A \cdot \operatorname{adj} A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ..... & ..... & ..... & ..... \\ a_{n1} & a_{n2} & ... & a_{1n} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & ... & A_{n1} \\ A_{12} & A_{22} & ... & A_{n2} \\ ..... & ..... & ..... & ..... \\ A_{1n} & A_{2n} & ... & A_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{k=1}^{n} a_{1k} \cdot A_{1k} & \sum_{k=1}^{n} a_{1k} \cdot A_{2k} & ... & \sum_{k=1}^{n} a_{1k} \cdot A_{nk} \\ \sum_{k=1}^{n} a_{2k} \cdot A_{1k} & \sum_{k=1}^{n} a_{2k} \cdot A_{2k} & ... & \sum_{k=1}^{n} a_{2k} \cdot A_{nk} \\ ..... & ..... & ... & ..... \\ \sum_{k=1}^{n} a_{nk} \cdot A_{1k} & \sum_{k=1}^{n} a_{nk} \cdot A_{2k} & ... & \sum_{k=1}^{n} a_{nk} \cdot A_{nk} \end{pmatrix} =$$

$$= \begin{pmatrix} \det A & 0 & ... & 0 \\ 0 & \det A & ... & 0 \\ ..... & ..... & ... & ..... \\ 0 & 0 & ... & \det A \end{pmatrix} = (\det A) \cdot I.$$

Assim, de (2.5), supondo A inversível, existe A^{-1} inversa de A tal que, multiplicando (2.5) à esquerda por A^{-1} vamos obter

$$A^{-1} \cdot A \cdot \operatorname{adj} A = A^{-1} \cdot (\det A) I = (\det A) \cdot A^{-1},$$

ou seja,

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A.$$

Ou seja, acabamos de provar o seguinte resultado:

Proposição 2.11 *Se A for uma matriz inversível, então sua inversa A^{-1} é dada por*

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A.$$

Dessa forma, dispomos de dois procedimentos para determinar a inversa de uma matriz quadrada A, se esta existir: podemos utilizar o algoritmo das operação elementares sobre linhas descrito no capítulo anterior, ou podemos obter através do cálculo da matriz adjunta de A, conforme a Proposição acima.

Teorema 2.12 *(Regra de Cramer) Seja o sistema linear $A \cdot x = b$, onde $A = (a_{ij})_{n\times n}$ corresponde à matriz dos coeficientes, $x = (x_1 \quad x_2 \quad .... \quad x_n)^T$ corresponde à matriz das incógnitas x_i e $b = (b_1 \quad b_2 \quad .... \quad b_n)^T$ corresponde à matriz dos termos independentes, se o sistema possui solução única, então ela é dada por*

$$x_i = \frac{\det A_i}{\det A},$$

onde $\det A_i$ denota o determinante da matriz A_i obtida de A substituindo-se a coluna i pela matriz-coluna b.

Demonstração. Suponha $Ax = b$ o sistema linear na hipóteses do Teorema. Como existe solução única, segue que A é inversível. Assim, existe A^{-1} inversa de A tal que, multiplicando a equação matricial $Ax = b$ à esquerda por A^{-1}, obtemos

$$A^{-1} \cdot Ax = A^{-1} \cdot b,$$

ou seja, determinamos a solução

$$x = A^{-1} \cdot b,$$

o que, pela Proposição 2.11 segue que a solução x do sistema é dada por

$$x = A^{-1} \cdot b = \frac{1}{\det A}(\operatorname{adj} A)b,$$

onde

$$(\text{adj}\,A)b = \begin{pmatrix} A_{11} & A_{21} & ... & A_{n1} \\ A_{12} & A_{22} & ... & A_{n2} \\ ..... & ..... & ..... & ..... \\ A_{1n} & A_{2n} & ... & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ ... \\ b_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} b_k A_{k1} \\ \sum_{k=1}^{n} b_k A_{k2} \\ ... \\ \sum_{k=1}^{n} b_k A_{kn} \end{pmatrix}$$

Note que, definindo A_i por

$$A_i = \begin{pmatrix} A_{11} & ... & A_{i-1,1} & b_1 & A_{i+1,1} & ... & A_{n1} \\ A_{12} & ... & A_{i-1,2} & b_2 & A_{i+1,2} & ... & A_{n2} \\ ..... & ..... & ..... & ..... & ..... & ..... & ..... \\ A_{1n} & ... & A_{i-1,n} & b_n & A_{i+1,n} & ... & A_{nn} \end{pmatrix},$$

temos que

$$\det A_i = b_1 \cdot A_{1i} + b_2 \cdot A_{2i} + ... + b_n \cdot A_{ni} = \sum_{k=1}^{n} b_k \cdot A_{ki},$$

e portanto,

$$x_i = \frac{1}{\det A} \sum_{k=1}^{n} b_k \cdot A_{ki} = \frac{\det A_i}{\det A}.$$

□

Exercícios

1. Dada a matriz

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

Determine a matriz inversa A^{-1} de duas formas: (a) usando o algoritmo de obtenção da inversa via operações elementares sobre linhas; (b) usando a matriz adjunta.

2. Dada a matriz $A = \begin{pmatrix} 2 & 1 & -3 \\ 0 & 2 & 1 \\ 5 & 1 & 3 \end{pmatrix}$, calcule:

   (a) $\operatorname{adj} A$.

   (b) $\det A$.

   (c) A^{-1}.

3. Calcule a inversa da matriz $M = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, através da matriz adjunta.

4. Use a regra de Cramer para resolver os sistemas lineares abaixo:

   (a) $\begin{cases} 2x + y - 3z = 2 \\ x + y - z = 0 \\ -x - y + 2z = -1 \end{cases}$ (b) $\begin{cases} x + 2y - z = 1 \\ 2x + y - 4z = -3 \\ x - y + z = 2 \end{cases}$

5. Determine o valor de $m \in \mathbb{R}$ para que o sistema

$$\begin{cases} x + y + mz = 1 \\ x - y + z = 0 \\ mx + y = 1 \end{cases}$$

   possua solução única.

Capítulo 3

Espaços vetoriais

Neste capítulo estudaremos o importante conceito de espaço vetorial, bem como suas principais propriedades e consequências, tais como o conceito de subespaço vetorial e mudança de base. O conceito de espaço vetorial é de extrema importância na Matemática e serve de ferramenta em estudos mais avançados, como Equações Diferenciais, Sistemas Dinâmicos e também em Análise Funcional, por exemplo. Os elementos de um espaço vetorial serão chamados de vetores. Dessa maneira, nos vem imediatamente em mente o espaço $\mathbb{R}^3$ e o estudo de vetores feito na Disciplina de Geometria Analítica. No entanto, o espaço $\mathbb{R}^3$ é um caso particular de espaço vetorial e mostraremos que existem muitos outros tipos de espaços vetoriais especiais.

3.1 Espaços vetoriais e exemplos

Provavelmente o leitor deva ter visto esse conceito mediante operações com vetores em $\mathbb{R}^3$ num curso de Geometria Analítica. No entanto, naquele curso o conceito era exatamente para os vetores do $\mathbb{R}^3$. Vamos, novamente definir um espaço vetorial, embora pareça ser o mesmo conceito, alertamos o leitor de que não estaremos lidando, necessariamente com vetores do $\mathbb{R}^3$, mas qualquer conjunto não vazio V que satisfaça a definição dada a seguir.

Definição 3.1 Chama-se um *espaço vetorial* real todo conjunto não vazio V, munido de duas operações: uma adição

$$+ : V \times V \to V$$

$$(u, v) \mapsto u + v \in V$$

e uma multiplicação por um escalar

$$\cdot : \mathbb{R} \times V \to V$$

$$(\alpha, u) \mapsto \alpha \cdot u \in V,$$

tal que cumprem as seguintes propriedades: dados $u, v, w \in V$ e $\alpha, \beta \in \mathbb{R}$, temos

A1. $u + (v + w) = (u + v) + w$ (associatividade);

A2. $u + v = v + u$ (comutatividade);

A3. $\exists 0 \in V$ tal que $u + 0 = u$ (existência do neutro aditivo em V);

A4. $\forall u \in V$, $\exists - u \in V$ tal que $u + (-u) = 0$;

M1. $\alpha(u + v) = \alpha \cdot u + \alpha \cdot v$ (distributividade);

M2. $(\alpha + \beta)u = \alpha \cdot u + \beta \cdot u$ (distributividade);

M3. $\alpha(\beta u) = (\alpha\beta)u$;

M4. $1 \cdot u = u$.

Dado um espaço vetorial real V, os elementos $v \in V$ são chamados de *vetores* de V.

Vamos trabalhar nesse curso com espaços vetoriais reais, ou seja, espaços vetoriais onde os escalares são números reais. No entanto, poderíamos considerar os escalares sendo números complexos, e então estaríamos trabalhando com espaços vetoriais complexos, ou, mais geralmente, poderíamos trabalhar com espaços vetoriais num *corpo* $\mathbb{K}$, mas este tópico foge de um primeiro curso de Álgebra linear.

A seguir apresentamos alguns exemplos de espaços vetoriais.

Exemplo 1. Como estudado na Geometria analítica, Os conjuntos de vetores $\mathbb{R}^2$ e $\mathbb{R}^3$, munidos da adição usual de vetores e o produto usual de um escalar por um vetor formam espaços vetoriais. Não é difícil mostrar as oito propriedades da definição de espaço vetorial para vetores do $\mathbb{R}^3$, por exemplo.

Mais geralmente, para $n \geq 1$, o conjunto $\mathbb{R}^n$ das n-uplas ordenadas, munido das operações

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

$$(x, y) \mapsto x + y,$$

onde $x = (x_1, x_2, ..., x_n)$ e $y = (y_1, y_2, ..., y_n)$, e daí $x+y = (x_1+y_1, ..., x_n+y_n)$, e

$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

$$(\alpha, x) \mapsto \alpha \cdot x,$$

onde $\alpha \cdot x = \alpha(x_1, ..., x_n) = (\alpha x_1, ..., \alpha x_n)$, é um espaço vetorial real.

Exemplo 2. O conjunto $M(n, n) = \{A = (a_{ij})_{n\times n} \ : a_{ij} \in \mathbb{R}, 1 \leq i, j \leq n\}$ de todas as matrizes $n \times n$, munido das operações de adição usual de matrizes e produto de um escalar por uma matriz é um espaço vetorial. De fato, basta visitar o Teorema 1.6, itens 01, 02, 03, 04, 06 e 07 do Teorema, e observar que, para qualquer matriz $A \in M(n, n)$ e para quaisquer $\alpha, \beta \in \mathbb{R}$, valem:

$$(\alpha\beta)A = \alpha(\beta A) \ \text{ e } \ 1 \cdot A = A.$$

Neste caso, os vetores de tal espaço são as matrizes $n \times n$.

Exemplo 3. Defina o conjunto $C([a, b])$ de todas as funções contínuas em $[a, b]$ com valores em $\mathbb{R}$, ou seja,

$$C([a, b]) = \{f : [a, b] \to \mathbb{R} \ : \ f \text{ é contínua em } [a, b]\},$$

munido com as operações usuais de adição de funções e multiplicação de uma constante por uma função, ou seja,

$$+ : C([a,b]) \times C([a,b]) \to C([a,b])$$

$$(f,g) \mapsto f+g,$$

onde, $\forall x \in [a,b]$, $(f+g)(x) = f(x)+g(x)$, que é uma função contínua, pois soma de funções contínuas é uma função contínua; e

$$\cdot : \mathbb{R} \times C([a,b]) \to C([a,b])$$

$$(\alpha, f) \mapsto \alpha \cdot f,$$

onde $\forall x \in [a,b]$, $(\alpha \cdot f)(x) = \alpha \cdot f(x)$, que é uma função contínua, pois o produto de uma constante por uma função contínua é uma função contínua.

Note que $C([a,b]) \neq \emptyset$, pois a função identicamente nula $0(x) = 0$, $\forall x \in [a,b]$ é contínua.

Vamos mostrar que $C([a,b])$ é um espaço vetorial. De fato, para isso, precisamos verificar todas as oito propriedades da definição de espaço vetorial: dados $f, g, h \in C([a,b])$ e $\alpha, \beta \in \mathbb{R}$, temos:

A1. $(f+g)+h = f+(g+h)$: De fato, dado $x \in [a,b]$, temos que

$$((f+g)+h)(x) = (f+g)(x)+h(x) = (f(x)+g(x))+h(x) =$$

$$= f(x)+(g(x)+h(x)) = f(x)+(g+h)(x) = (f+(g+h))(x),$$

e como essa igualdade é verdadeira $\forall x \in [a,b]$, concluímos que

$$(f+g)+h = f+(g+h).$$

A2. $f+g = g+f$: De fato, dado $x \in [a,b]$, temos que

$$(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x),$$

e como essa igualdade é verdadeira $\forall x \in [a,b]$, concluímos que

$$(f+g)+h = f+(g+h).$$

A3. $\exists 0 \in C([a,b])$ tal que, $\forall f \in C([a,b])$, temos $f + 0 = f$:
De fato, considere a função identicamente nula $0(x) = 0$, que é contínua, logo, $\forall x \in [a,b]$, temos

$$(f + 0)(x) = f(x) + 0(x) = f(x) + 0 = f(x),$$

ou seja, $f + 0 = f$.

A4. $\forall f \in C([a,b])$, $\exists -f \in C([a,b])$ tal que $f + (-f) = 0$, onde 0 é a função identicamente nula: De fato, dada $f \in [a,b]$, seja $-f = -1 \cdot f$, que também é contínua, pois é o produto de uma constante, no caso -1, pela função contínua f, ou seja, $-f \in C([a,b])$, e daí, para todos $x \in [a,b]$ teremos

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) + (-f(x)) = f(x) - f(x) = 0,$$

$\forall x \in [a,b]$, ou seja, $f + (-f) = 0$.

M1. $\alpha(f + g) = \alpha \cdot f + \alpha \cdot g$: De fato, dados $\alpha \in \mathbb{R}$ e $f, g \in C([a,b])$, temos, $\forall x \in [a,b]$, que

$$(\alpha(f + g))(x) = \alpha((f + g)(x)) = \alpha(f(x) + g(x)) = \alpha \cdot f(x) + \alpha \cdot g(x) =$$

$$= (\alpha \cdot f)(x) + (\alpha \cdot g)(x) = (\alpha \cdot f + \alpha \cdot g)(x),$$

ou seja, $\alpha(f + g) = \alpha \cdot f + \alpha \cdot g$.

M2. $(\alpha + \beta)f = \alpha \cdot f + \beta \cdot f$: Análogo ao anterior, fica como exercício.

M3. $(\alpha\beta)f = \alpha(\beta f)$: De fato, dados $\alpha \in \mathbb{R}$ e $f \in C([a,b])$, temos $\forall x \in [a,b]$, que

$$((\alpha\beta)f)(x) = (\alpha\beta)(f(x)) = \alpha(\beta f(x)) = \alpha((\beta f)(x)) = (\alpha(\beta f))(x),$$

ou seja, $(\alpha\beta)f = \alpha(\beta f)$.

M4. $1 \cdot f = f$: Este é óbvio, pois $\forall x \in [a,b]$ $(1 \cdot f)(x) = 1 \cdot (f(x)) = f(x)$.

Portanto, $C([a,b])$ é um espaço vetorial, o espaço vetorial das funções contínuas em $[a,b]$. Nesse caso, as funções contínuas em $[a,b]$ são os vetores desse espaço.

Exemplo 4. Fixado $n \in \mathbb{N}$, seja $P_n = \{a_n x^n + a_{n-1}x^{n-1} + ... + a_0 \, : \, a_i \in \mathbb{R},\, i = 0, 1, 2, ..., n\}$ o conjunto dos polinômios de grau menor ou igual a n, munido da adição usual de polinômios e produto de um escalar por un polinômio. Não é difícil verificar as oito propriedades da definição para comprovar que P_n é um espaço vetorial. Fica como exercício para o leitor.

Exemplo 5. Seja $D = \{f : (a,b) \to \mathbb{R} \, : \, f \text{ é derivável em } (a,b)\}$, munido da adição usual de funções com o produto de um escalar por uma função. Como valem as regras de derivação

$$(f+g)' = f' + g' \quad \text{e} \quad (k \cdot f)' = k \cdot f',$$

onde $f, g \in D$ e $k \in \mathbb{R}$, não é difícil mostrar que D é um espaço vetorial.

Exercícios

1. Seja $\mathbb{R}^\infty$ o conjunto de todas as sequências infinitas de números reais, ou seja,

   $$\mathbb{R}^\infty = \{x = (x_1, x_2, ..., x_n, ...) \, : \, x_i \in \mathbb{R}\},$$

   munido da operações de adição

   $$(x_1, x_2, ..., x_n, ...) + (y_1, y_2, ..., y_n, ...) = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n, ...),$$

   e multiplicação por escalar

   $$\alpha(x_1, x_2, ..., x_n, ...) = (\alpha x_1, \alpha x_2, ..., \alpha x_n, ...).$$

   Mostre que $\mathbb{R}^\infty$ munido das operações acima é um espaço vetorial.

2. Em $\mathbb{R}^n$ defina as operações

   $$\vec{u} \oplus \vec{v} = \vec{u} - \vec{v} \quad \text{e} \quad \alpha \odot \vec{u} = -\alpha \cdot \vec{u}.$$

   Quais axiomas de espaço vetorial são satisfeitos para $(\mathbb{R}^n, \oplus, \odot)$?

3. Verifique se são espaços vetoriais os seguintes conjuntos:

   (a) o $\mathbb{R}^2$ munido da adição usual e a multiplicação $\alpha(x, y) = (\alpha x, 0)$.

   (b) o $\mathbb{R}^2$ munido da adição $(x_1, y_1) + (x_2, y_2) = (x_1 + 2x_2, y_1 + 2y_2)$ e a multiplicação por escalar usual.

   (c) o $\mathbb{R}^2$ munido da adição $(x_1, y_1) + (x_2, y_2) = (y_1 + y_2, x_1 + x_2)$ e a multiplicação por escalar usual.

4. Dados os espaços vetoriais V_1 e V_2, considere o conjunto $E = V_1 \times V_2$, cujos elementos são os pares ordenados $v = (v_1, v_2)$, onde $v_1 \in V_1$ e $v_2 \in V_2$. Defina operações que tornem V um espaço vetorial.

3.2 Subespaços vetoriais

Nesta seção vamos estudar o conceito de subespaço vetorial de um espaço vetorial, que na verdade trata-se de um espaço vetorial "menor" dentro do espaço dado, que preserva as mesmas operações do espaço "maior".

Definição 3.2 Seja V um espaço vetorial e considere $W \subset V$ um subconjunto não vazio de V. Dizemos que W é um *subespaço vetorial* de V se cumprir as seguintes condições: para quaisquer $u, v \in W$ e $\alpha \in \mathbb{R}$, valem,

(a) $u + v \in W$;

(b) $\alpha \cdot u \in W$.

Note que, dado um espaço vetorial V qualquer, um subconjunto W de V automaticamente herdará todas as oito propriedades de espaço vetorial, isso porque todo elemento de W é um elemento de V e por essa razão as propriedades estarão válidas. No entanto, para que W possa ter uma estrutura de espaço vetorial, resta verificar se a adição e a multiplicação por escalar estão fechadas em W, ou seja, se somarmos dois vetores em W o vetor resultante deve continuar em W, e o mesmo para o produto de um escalar por um elemento de W. E exatamente essas duas exigências estão encerradas na definição acima de subespaço vetorial.

Observação 3.3 Convém notar que se $W \subset V$ é um subespaço vetorial de V, obrigatoriamente segue que o neutro aditivo 0 de V também pertence a W. Isto se justifica pela propriedade (b) da definição de subespaço, tomando $\alpha = 0$. Assim, se $0 \notin W$, segue que W não é um subespaço vetorial de V.

Abaixo apresentamos alguns exemplos e contra-exemplos de subespaço vetorial.

Exemplo 1. Todo espaço vetorial V possui pelo menos dois subespaços, chamados de *subespaços triviais*: o subespaço $\{0\}$ formado apenas pelo vetor nulo e todo o espaço V, subespaço de si mesmo.

Exemplo 2. Seja $V = \mathbb{R}^2$ com as operações de adição de vetores e multiplicação de escalar por vetor usuais. Defina $W \subset V$ por

$$W = \{(x, y) \in \mathbb{R}^2 \, : \, y = 2x\}.$$

Ou seja, estamos considerando como espaço vetorial V o plano cartesiano e tomamos como W o subconjunto do plano formado pelos pontos da reta $y = 2x$. Note que $W \neq \emptyset$, pois $(0, 0) \in W$, ou seja e reta $y = 2x$ que define o subconjunto W passa pela origem do $\mathbb{R}^2$.

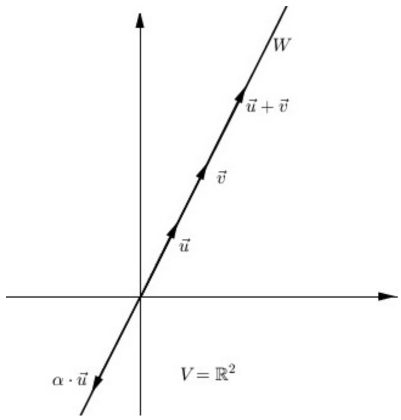
Afirmamos que W é um subespaço vetorial de V. De fato, dados $\vec{u} = (a, 2a)$ e $\vec{v} = (b, 2b)$ elementos quaisquer de W e $\alpha \in \mathbb{R}$ um escalar qualquer, segue que

(a) $\vec{u} + \vec{v} = (a, 2a) + (b, 2b) = (a + b, 2a + 2b) = (a + b, 2(a + b)) \in W$;

(b) $\alpha \cdot \vec{u} = \alpha \cdot (a, 2a) = (\alpha \cdot a, \alpha \cdot 2b) = (\alpha a, 2(\alpha a)) \in W$.

Logo, W é um subespaço de V.

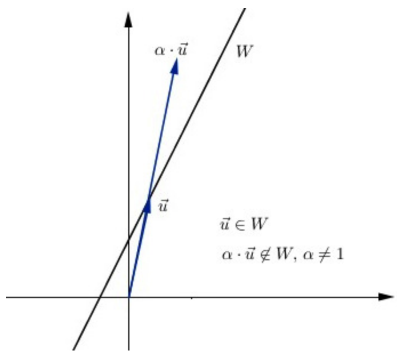
Podemos dar uma interpretação geometrica para esse exemplo. Observe o desenho abaixo, onde temos que W é o conjunto de pontoa do $\mathbb{R}^2$ sobre a reta $y = 2x$, uma reta que passa pela origem do plano $\mathbb{R}^2 = V$. Como estudado na Geometria analítica, sabemos que existe uma correspondência biunívoca entre ponto e vetor, segue que um vetor pertencerá a um conjunto se a “ponta da seta" estiver no conjunto. Dessa forma, dados $\vec{u}, \vec{v} \in W$ e $\alpha \in \mathbb{R}$, vemos que $\vec{u} + \vec{v} \in W$ e que $\alpha \cdot \vec{u} \in W$, e portanto, $W \subset V$ é um subespaço vetorial de V (no desenho tomamos $\alpha < 0$).



Exemplo 3. No mesmo contexto do exemplo anterior, considere $V = \mathbb{R}^2$ e W subconjunto de V dado por $W = \{(x, y) \in \mathbb{R}^2 \, : \, y = 2x + 1\}$.
Neste caso temos que W não é um subespaço vetorial de V, pois, por exemplo, dado $\vec{u} = (a, 2a + 1)$ um vetor qualquer em W e $\alpha \neq 1$ um escalar, temos que

$$\alpha \cdot \vec{u} = \alpha(a, 2a + 1) = (\alpha a, 2\alpha a + \alpha) \notin W,$$

ou seja, o produto de um escalar por um vetor de W não fica em W. O mesmo aconteceria se somássemos dois elementos quaisquer de W: verificaríamos que a soma também estaria fora de W. Repare, neste exemplo que W é uma reta que não passa pela origem, logo, o vetor nulo não pertencendo a W já nos mostra que W não pode ser subespaço de V. Veja a ilustação abaixo.



Exemplo 4. Sendo $V = \mathbb{R}^2$ com as operações de adição de vetores e multiplicação de escalar por vetor usuais. Defina $W = \{(x, y) \in \mathbb{R}^2 \, : \, y = x^2\}$, ou seja W é o conjunto de todos os pontos do plano cartesiano que estão sobre a parábola $y = x^2$. Não é difícil verificar que W não é um subespaço vetorial de V. Deixaremos os detalhes para o leitor, inclusive recomendamos fazer um desenho para expressar V e W, destacando os vetores $\vec{u}, \vec{v} \in W$, $\vec{u}+\vec{v}$ e também $\alpha\vec{u}$.

Repare neste contra-exemplo que a origem $\vec{0} = (0, 0)$ pertence a W, mas mesmo assim W não é subespaço de V. Isso porque o neutro pertencer a W não é garantia de que W seja subespaço de V, sabemos apenas que se o neutro não pertencer a W, então W não é subespaço de V.

Exemplo 5. Seja $V = \mathbb{R}^3$ com as operações de adição de vetores e multiplicação de escalar por vetor usuais. Considere

$$W = \{(x, y, z) \in \mathbb{R}^3 \, : \, y = 0\}.$$

Note que o subconjunto W corresponde ao plano xz de equação $y = 0$, perpendicular ao plano horizontal xy. Vamos mostrar que W é um subespaço vetorial de $V = \mathbb{R}^3$. De fato, dados $\vec{u} = (u_1, 0, u_3)$ e $\vec{v} = (v_1, 0, v_3)$ elementos de W e α um escalar real, temos que

(a) $\vec{u} + \vec{v} = (u_1, 0, u_3) + (v_1, 0, v_3) = (u_1 + v_1, 0, u_3 + v_3) \in W$;

(b) $\alpha\vec{u} = \alpha(u_1, 0, u_3) = (\alpha u_1, 0, \alpha u_3) \in W$.

Logo, W é um subespaço vetorial de V. Faça um desenho para ilustrar.

Exemplo 6. Seja $V = M(3,3)$ o espaço vetorial das matrizes 3×3 com entradas reais, munido da adição de matrizes e o produto de escalar por matriz usuais. Defina $W \subset V$ o conjunto de todas as matrizes triangulares superiores 3×3. Como a soma de matrizes triangulares superiores é uma matriz triangular superior e o produto de um escalar por uma matriz triangular superior é ainda uma matriz triangular superior, segue que W é um subespaço vetorial de V.

Exemplo 7. Seja $V = P_n$ o espaço vetorial de todos os polinômios de grau menor ou igual a n, com as operações usuais de adição de polinômios e o produto de um número real por um polinômio. Defina

$$W = \{p \in P_n \,:\, p(0) = 0\}.$$

Primeiramente, note que $W \neq \emptyset$ pois o polinômio identicamente nulo $p(x) \equiv 0$ é tal que $p(0) = 0$, logo, pertence a W e então W está bem definido.

Mostremos que W é subespaço vetorial de P. Dados $p_1, p_2 \in W$ e $\alpha \in \mathbb{R}$. Então:

(a) $p_1 + p_2 \in W$, pois $(p_1 + p_2)(0) = p_1(0) + p_2(0) = 0$;

(b) $\alpha \cdot p_1 \in W$, pois $\alpha \cdot p_1(0) = \alpha \cdot 0 = 0$.

Portanto, W é um subespaço vetorial de $V = P_n$.

Exemplo 8. Seja $V = C([a,b])$ o espaço das funções contínuas em $[a,b]$, munido de suas operações usuais de soma de funções e produto de um escalar por uma função, c.f. estudado anteriormente. Defina

$$C^1([a,b]) = \{f : [a,b] \to \mathbb{R} \text{ tal que } f' : [a,b] \to \mathbb{R} \text{ é contínua}\}.$$

Primeiramente notamos que $C^1([a,b]) \neq \emptyset$ pois a função identicamente nula pertence a tal conjunto.

Vamos mostrar também que $C^1([a,b]) \subset C([a,b])$. De fato, dado $g \in C^1([a,b])$, segue que $g' : [a,b] \to \mathbb{R}$ é contínua. Portanto, do Cálculo integral segue que g' é integrável e então podemos definir $G : [a,b] \to \mathbb{R}$ por

$$G(x) = \int_0^x g'(t)\,dt,$$

que é conínua. Pelo Teorema Fundamental do Cálculo segue que

$$G(x) = g(x) - g(0),$$

e então $g(x) = G(x) + g(0)$, que será, portanto, contínua. Logo, concluímos que $g \in C([a,b])$, e daí $C^1([a,b]) \subset C([a,b])$.

Por fim, mostremos que $C^1([a,b])$ é um subespaço vetorial de $C([a,b])$. De fato, dados $f, g \in C^1([a,b])$, segue que f' e g' são contínuas em $[a,b]$, e como a soma de funções contínuas é uma função contínua, concluímos que $f' + g'$ é contínua em $[a,b]$, ou seja,

$$f' + g' \in C^1([a,b]). \tag{3.1}$$

Do mesmo modo, dado $\alpha \in \mathbb{R}$ temos $\alpha f'$ é contínua em $[a,b]$, pois o produto de uma constante por uma função contínua é uma função contínua, ou seja, concluímos que

$$\alpha \cdot f' \in C^1([a,b]). \tag{3.2}$$

Portanto, por (3.1) e (3.2) concluímos que $C^1([a,b])$ é um subespaço vetorial de $C([a,b])$.

No que segue apresentaremos dois resultados importantes sobre subespaços vetoriais. Apenas devido à familiaridade com vetores em $\mathbb{R}^3$ vamos trabalhar com a notação vetorial $\vec{v}$ para denotar um vetor v de um espaço vetorial V, mas frisamos que se V for um espaço vetorial mais geral, como por exemplo $C([a,b])$, os vetores $\vec{v}$ serão funções contínuas de $[a,b]$ em $\mathbb{R}$ e, nesse caso, $\vec{v} = f$, para f contínua em $[a,b]$.

Proposição 3.4 *Sejam W_1 e W_2 dois subespaços vetoriais de um espaço vetorial V. Então o conjunto $W_1 \cap W_2$ também é um subespaço vetorial de V.*

Demonstração. Sejam W_1 e W_2 subespaços de V. Defina o conjunto

$$W_1 \cap W_2 = \{\vec{u} : \vec{u} \in W_1 \ \text{ e } \ \vec{u} \in W_2\}.$$

Primeiramente observamos que $W_1 \cap W_2 \neq \emptyset$ pois $\vec{0} \in W_1$ e $\vec{0} \in W_2$. Logo, o conjunto $W_1 \cap W_2$ está bem definido.

Dados $\vec{u}, \vec{v} \in W_1 \cap W_2$ e $\alpha \in \mathbb{R}$. Logo, $\vec{u}, \vec{v} \in W_1$ e $\vec{u}, \vec{v} \in W_2$.

Assim, temos que

(i) $\vec{u} + \vec{v} \in W_1$ e $\alpha\vec{u} \in W_1$, pois W_1 é subespaço vetorial de V.

(ii) $\vec{u} + \vec{v} \in W_2$ e $\alpha\vec{u} \in W_2$, pois W_2 é subespaço vetorial de V.

Portanto, de (i) e (ii) temos que $\vec{u}+\vec{v} \in W_1 \cap W_2$ e $\alpha\vec{u} \in W_1 \cap W_2$, provando que $W_1 \cap W_2$ é um subespaço vetorial de V.

□

Proposição 3.5 *Sejam W_1 e W_2 dois subespaços vetoriais de um espaço vetorial V. Então o conjunto*

$$W_1 + W_2 = \{\vec{w}_1 + \vec{w}_2 \ : \ \vec{w}_1 \in W_1 \ \ e \ \ \vec{w}_2 \in W_2\}$$

é um subespaço vetorial de V.

Demonstração. Note que $W_1 + W_2 \neq \emptyset$ pois $\vec{0} \in W_1$ e $\vec{0} \in W_2$ e é tal que

$$\vec{0} = \vec{0} + \vec{0} \in W_1 + W_2.$$

Logo, $W_1 + W_2$ está bem definida.

Mostremos que $W_1 + W_2$ é um subespaço vetorial de V.
Dados $\vec{u}, \vec{v} \in W_1 + W_2$ e $\alpha \in \mathbb{R}$. Então existem $\vec{u}_1$, $\vec{v}_1 \in W_1$ e $\vec{u}_2$, $\vec{v}_2 \in W_2$ tais que

$$\vec{u} = \vec{u}_1 + \vec{u}_2 \ \text{ e } \ \ \vec{v} = \vec{v}_1 + \vec{v}_2.$$

Então

$$\vec{u} + \vec{v} = (\vec{u}_1 + \vec{u}_2) + (\vec{v}_1 + \vec{v}_2) = \vec{u}_1 + (\vec{u}_2 + \vec{v}_1) + \vec{v}_2 =$$

$$= \vec{u}_1 + (\vec{v}_1 + \vec{u}_2) + \vec{v}_2 = (\vec{u}_1 + \vec{v}_1) + (\vec{u}_2 + \vec{v}_2) \in W_1 + W_2,$$

pois $\vec{u}_1 + \vec{v}_1 \in W_1$ e $\vec{u}_2) + \vec{v}_2 \in W_2$, e

$$\alpha \cdot \vec{u} = \alpha(\vec{u}_1 + \vec{u}_2) = \alpha\vec{u}_1 + \alpha\vec{u}_2 \in W_1 + W_2,$$

pois $\alpha\vec{u}_1 \in W_1$ e $\alpha\vec{u}_2 \in W_2$, e isso conclui a prova da Proposição.

□

Observação 3.6 Quando $W_1 \cap W_2 = \{0\}$, então o espaço vetorial $W_1 + W_2$ chama-se *soma direta de* W_1 *com* W_2 e nesse caso denotamos por $W_1 \oplus W_2$.

Exercícios

1. Mostre que os seguntes subconjuntos do $\mathbb{R}^4$ são subespaços vetoriais:

   (a) $W = \{(x, y, z, t) \in \mathbb{R}^4 \;:\; x + y = 0 \;\text{ e }\; z - t = 0\}$.

   (b) $S = \{(x, y, z, t) \in \mathbb{R}^4 \;:\; 2x + y - t = 0 \;\text{ e }\; z = 0\}$.

2. Seja $V = \mathbb{R}^2$ e considere o subconjunto

   $$W = \{(x, y) \in \mathbb{R}^2 \;: xy \leq 0\}.$$

   Desenhe $W \subset \mathbb{R}^2$ e verifique se W é um subespaço vetorial, justificando sua resposta.

3. Seja V o espaço vetorial real de todas as funções de $\mathbb{R}$ em $\mathbb{R}$. Quais dos seguintes conjuntos de funções são subespaços de V?

   (a) de todas as funções f tais que $f(x^2) = f(x)^2$.

   (b) de todas as funções f tais que $f(0) = f(1)$.

   (c) de todas as funções f tais que $f(-1) = 0$.

   (d) de todas as funções contínuas.

4. Seja V o espaço vetorial de todas as funções de $\mathbb{R}$ em $\mathbb{R}$. Seja V_p o subconjunto de todas as funções pares, i.e., tais que $f(-x) = f(x)$; e seja V_i o subconjunto de todas as funções ímpares, i.e., tais que $f(-x) = -f(x)$. Prove que V_p e V_i são subespaços de V.

3.3 Vetores linearmente independentes e linearmente dependentes

Como já foi exposto acima, apenas por familiaridade com a Geometria Analítica, para os resultados gerais vamos denotar um vetor de um espaço vetorial V com uma seta, ou seja, $\vec{v}$ denotará um vetor em V. Mas, como sempre, dependerá do contexto do espaço vetorial estudado. No Exemplo 3 que será dado adiante os vetores serão funções reais e então não há necessidade em denotar um vetor com uma seta, por exemplo.

O conceito de dependência e independência linear é de extrema importância na Álgebra linear, pois desenvolveremos na Seção seguinte o importante conceito de base de um espaço vetorial nessa seção, e esse conceito depende de independência linear. Antes, porém, precisamos definir combinação linear de vetores.

Definição 3.7 Sejam $\vec{v}_1$, $\vec{v}_2$,..., $\vec{v}_n$ n vetores de um espaço vetorial V e α_1, α_2, ..., α_n escalares. Dizemos que o vetor

$$\vec{v} = \alpha_1 \cdot \vec{v}_1 + \alpha_2 \cdot \vec{v}_2 + ... + \alpha_n \cdot \vec{v}_n$$

é uma *combinação linear* dos vetores $\vec{v}_1$, $\vec{v}_2$,..., $\vec{v}_n$.

Por exemplo, considerando o espaço vetorial $V = \mathbb{R}^2$, temos que o vetor $\vec{u} = (-1, 5)$ é uma combinação linear dos vetores $\vec{v}_1 = (1, 1)$ e $\vec{v}_2 = (2, -1)$, pois

$$(-1, 5) = 3 \cdot (1, 1) - 2 \cdot (2, -1).$$

Uma vez fixados os vetores $\vec{v}_1$, ..., $\vec{v}_n$ de um espaço vetorial V, podemos definir o conjunto de todas as combinações lineares desses vetores, e denotaremos por $[\vec{v}_1, ..., \vec{v}_n]$. Temos pois, o seguinte resultado:

Proposição 3.8 *Fixados os vetores $\vec{v}_1$, ..., $\vec{v}_n$ de um espaço vetorial V, o conjunto $W = [\vec{v}_1, \ ..., \ \vec{v}_n]$ de todas as combinações lineares de tais vetores é um subespaço vetorial de V.*

Demonstração. Sejam $\vec{u}$ e $\vec{v}$ vetores em $W = [\vec{v}_1, \ ..., \ \vec{v}_n]$. Então, existem constantes $\alpha_i, \beta_i \in \mathbb{R}$, com $i = 1, 2, ..., n$ tais que

$$\vec{u} = \alpha_1 \cdot \vec{v}_1 + ... + \alpha_n \cdot \vec{v}_n \ \text{ e } \ \vec{v} = \beta_1 \cdot \vec{v}_1 + ... + \beta_n \cdot \vec{v}_n.$$

Assim, temos que

$$\begin{aligned} \vec{u} + \vec{v} &= (\alpha_1 \cdot \vec{v}_1 + ... + \alpha_n \cdot \vec{v}_n) + (\beta_1 \cdot \vec{v}_1 + ... + \beta_n \cdot \vec{v}_n) = \\ &= (\alpha_1 + \beta_1)\vec{v}_1 + ... + (\alpha_n + \beta_1)\vec{v}_n \in W; \end{aligned} \tag{3.3}$$

e, para qualquer $\gamma \in \mathbb{R}$, vale

$$\gamma \cdot \vec{u} = \gamma(\alpha_1 \cdot \vec{v}_1 + ... + \alpha_n \cdot \vec{v}_n) = (\gamma\alpha_1) \cdot \vec{v}_1 + ... + (\gamma\alpha_n) \cdot \vec{v}_n \in W. \tag{3.4}$$

Logo, de (3.3) e (3.4) segue o resultado.

$\square$

O resultado acima nos fornece um importante conceito:

Definição 3.9 O subespaço vetorial $W = [\vec{v}_1, ..., \vec{v}_n]$ de V acima definido chama-se *subespaço gerado* pelos vetores $\vec{v}_1, ..., \vec{v}_n$.

Proposição 3.10 *O subespaço gerado $W = [\vec{v}_1, ..., \vec{v}_n]$ por n vetores de um espaço vetorial V é o menor subespaço que contém os vetores $\vec{v}_1, ..., \vec{v}_n$.*

Demonstração. Seja W_1 outro subespaço vetorial de V tal que contenha os vetores $\vec{v}_1, ..., \vec{v}_n$. Para provar a Proposição é suficiente mostrar que $W \subset W_1$. De fato, seja $\vec{v} \in W$, então existem escalares $\alpha_1, ..., \alpha_n$ tais que

$$\vec{v} = \alpha_1 \cdot \vec{v}_1 + ... + \alpha_n \cdot \vec{v}_n.$$

Mas como $\vec{v}_1$,..., $\vec{v}_n \in W_1$ e W_1 é subespaço de V, segue que $\alpha_i\vec{v}_i \in W_1$, para $i = 1, 2, ..., n$; e ainda, como W_1 é subespaço de V, temos que

$$\alpha_1 \cdot \vec{v}_1 + ... + \alpha_n \cdot \vec{v}_n \in W_1,$$

ou seja, $\vec{v} \in W_1$. Portanto, mostramos que $W \subset W_1$, ou seja, o subespaço $W = [\vec{v}_1, ..., \vec{v}_n]$ gerado pelos vetores $\vec{v}_1, ..., \vec{v}_n$ é o menor subespaço de V que os contém.

□

Por fim, apresentamos o importante conceito de dependeência e independência linear.

Definição 3.11 Seja V um espaço vetorial e $\vec{v}_1, ..., \vec{v}_n$ vetores de V. Dizemos que esses vetores são *linearmente independentes* (abreviadamente, L.I.), se

$$\alpha_1 \cdot \vec{v}_1 + \alpha_2 \cdot \vec{v}_2 + ... + \alpha_n \cdot \vec{v}_0 = \vec{0},$$

se, e somente se, $\alpha = \alpha_2 = ... = \alpha_n = 0$. No caso onde pelo menos um dos α_i seja diferente de zero, dizemos que esses vetores são *linearmente dependentes* (abreviadamente, L.D.)

Exemplo 1. Considerando $V = \mathbb{R}^3$, temos que os vetores $\vec{u} = (2, 4, -8)$ e $\vec{v} = (3, 6, -12)$ são linearmente dependentes (L.D.), pois

$$\alpha(2, 4, -8) + \beta(3, 6, -12) = (0, 0, 0)$$

se, e somente se,

$$\begin{cases} 2\alpha + 3\beta = 0 \\ 4\alpha + 6\beta = 0 \\ -8\alpha - 12\beta = 0 \end{cases}$$

Note que a terceira equação do sistema linear homogêneo acima é -4 vezes a primeira equação e a segunda é 2 vezes a primeira equação, portanto, resta apenas

$$2\alpha + 3\beta = 0 \Rightarrow \alpha = -\frac{3\beta}{2}.$$

Assim, por exemplo, para $\beta = 2$ temos $\alpha = -3$, e são tais que

$$-3 \cdot \vec{u} + 2 \cdot \vec{v} = -3 \cdot (2, 4, -8) + 2 \cdot (3, 6, -12) = (-6, -12, 24) + (6, 12, -24) = \vec{0},$$

ou seja tal combinação linear resulta no vetor nulo sem que os escalares sejam todos nulos. Portanto, os vetores $\vec{u}$ e $\vec{v}$ dados são L.D.

Exemplo 2. Considerando $V = \mathbb{R}^3$, sejam os vetores $\vec{i} = (1, 0, 0)$; $\vec{j} = (0, 1, 0)$ e $\vec{k} = (0, 0, 1)$, mostremos que os mesmos são L.I.

De fato, se α, β e γ são escalares tais que

$$\alpha\vec{i} + \beta\vec{j} + \gamma\vec{k} = \vec{0},$$

obtemos

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

ou seja, $\alpha = \beta = \gamma = 0$, donde segue que o conjunto de vetores $\{\vec{i},\ \vec{j},\ \vec{k}\}$ é L.I.

Exemplo 3. Seja V o espaço vetorial das funções de $\mathbb{R}$ em $\mathbb{R}$ munido das operações usuais de soma de funções e produto de um escalar por uma função. Note que o vetor nulo nesse espaço é a função identicamente nula 0. Considere nesse espaço os vetores $f = e^x$, $g = 1 + e^{2x}$ e $h = 2$. Afirmamos que esses três vetores são L.I. em V pois, sendo $\alpha, \beta, \gamma \in \mathbb{R}$ escalares, temos que

$$\alpha f + \beta g + \gamma h = 0 \Leftrightarrow \alpha e^x + \beta(1 + e^{2x}) + \gamma \cdot 2 = 0,$$

e como $e^x > 0$, $1 + e^{2x} > 0$ e $2 > 0$, $\forall x$, somos obrigados a concluir que $\alpha = \beta = \gamma = 0$, ou seja, os vetores f, g e h são L.I. em V.

Teorema 3.12 *Um conjunto de vetores $\{\vec{v}_1, ..., \vec{v}_n\}$ de um espaço vetorial V é L.D. se, e somente se, um desses vetores for combinação linear dos demais.*

Demonstração. Primeiramente, suponha que o conjunto $\{\vec{v}_1, ..., \vec{v}_n\}$ de um espaço vetorial V seja L.D. Assim, dado um conjunto de escalares $\{\alpha_1, ..., \alpha_n\}$, segue queexiste pelo menos um escalar $\alpha_j \neq 0$ tal que a combinação linear a seguir seja o vetor nulo:

$$\alpha_1\vec{v}_1 + ... + \alpha_j\vec{v}_j + \alpha_n\vec{v}_n = \vec{0}.$$

Disso segue que

$$\alpha_j\vec{v}_j = -\alpha_1\vec{v}_1 - ... - \alpha_{j-1}\vec{v}_{j-1} - \alpha_{j+1}\vec{v}_{j+1} - ... - \alpha_n\vec{v}_n,$$

e, portanto, dividindo por $\alpha_j \neq 0$ obtemos

$$\vec{v}_j = -\frac{\alpha_1}{\alpha_1}\vec{v}_1 - ... - \frac{\alpha_n}{\alpha_j}\vec{v}_n,$$

ou seja, mostramos que um dos vetores do conjunto $\{\vec{v}_1, ..., \vec{v}_n\}$ é combinação linear dos demais.

Reciprocamente, suponha que

$$v_j = \beta_1 \vec{v_1} + ... + \beta_{j-1}\vec{v}_{j-1} + \beta_{j+1}\vec{v}_{j+1} + ... + \beta_n \vec{v}_n$$

é uma combinação linear dos demais vetores do conjunto $\{\vec{v}_1, ..., \vec{v}_n\}$. Vamos mostrar que tal conjunto é L.D. De fato, da igualdade acima é imediato que

$$\beta_1 \vec{v_1} + ... + \beta_{j-1}\vec{v}_{j-1} - 1 \cdot \vec{v}_j \beta_{j+1}\vec{v}_{j+1} + ... + \beta_n \vec{v}_n,$$

e portanto, o conjunto $\{\vec{v}_1, ..., \vec{v}_n\}$ de vetores de V é L.D. Isso completa a prova do Teorema.

$\square$

Exercícios

1. Escreva a matriz $A = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$ como uma combinação linear das matrizes

$$M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{e} \quad P = \begin{pmatrix} 0 & 2 \\ 0 & -1 \end{pmatrix}.$$

2. Considere o espaccco vetorial $P_2 = \{at^2 + bt + c \, : a, b, c \in \mathbb{R}\}$ e os vetores $p_1 = t^2 - 2t + 1$, $p_2 = t + 2$ e $p_3 = 2t^2 - t$.

   (a) Escreva o vetor $p = 5t^2 - 5t + 7$ como combinação linear de p_1, p_2 e p_3.

   (b) Determine uma condição para a, b e c de modo que o vetor $at^2 + bt + c$ seja uma combinação linear de p_2 e p_3.

3. Seja $V = C(\mathbb{R})$ o espaço vetorial das funções contínuas de $\mathbb{R}$ em $\mathbb{R}$, e considere os vetores $f = \cos^2 x$ e $g = \operatorname{sen}^2 x$. Quais dos seguintes vetores pertencem a $[f, g]$?

   (a) $\cos 2x$ (b) $3 + x^2$ (c) 1 (d) $\operatorname{sen} x$ (e) 0

4. Mostre que, se u, v e w são vetores LI, então $u+v$, $u+w$ e $v+w$ também são LI.

5. Considere $V[a,b]$ como sendo o espaço de todas as funções reais de uma variável real t de $[a,b]$ em $\mathbb{R}$. Mostrar que os seguintes pares de vetores a seguir são LI:
(a) t, t^2. (b) te^t, e^{2t}. (c) $\sin t, \cos t$. (d) $\cos t, \cos 3t$.

6. Defina a *média* $u \star v$ entre os vetores de um espaço vetorial V pondo $u \star v = \frac{1}{2}u + \frac{1}{2}v$. Prove que $(u \star v) \star w = u \star (v \star w)$ se, e somente se, $u = w$.

3.4 Base de um espaço vetorial

Quando estudamos a Geometria Analítica vimos que, por exemplo, em $\mathbb{R}^3$, qualquer vetor $\vec{u} = (u_1, u_2, u_3)$ pode ser escrito como uma combinação linear dos vetores $\vec{i} = (1,0,0)$, $\vec{j} = (0,1,0)$ e $\vec{k} = (0,0,1)$. De fato, basta observar que

$$\vec{u} = (u_1, u_2, u_3) = (u_1, 0, 0) + (0, u_2, 0) + (0, 0, u_3) =$$

$$= u_1(1,0,0) + u_2(0,1,0) + u_3(0,0,1) = u_1\vec{i} + u_2\vec{j} + u_3\vec{k}.$$

Na Geometria Analítica definimos que o conjunto de vetores $\{\vec{i}, \vec{j}, \vec{k}\}$ era uma *base* para o $\mathbb{R}^3$, pois todo vetor desse espaço pode ser escrito como uma combinação linear desses três vetores. Mais ainda, tal base foi chamada de *base canônica* do $\mathbb{R}^3$.

Além disso, vimos na seção anterior que o conjunto de vetores $\{\vec{i}, \vec{j}, \vec{k}\}$ é L.I. no espaço vetorial $\mathbb{R}^3$ (veja o Exemplo 2 da Seção anterior).

Isso nos inspira definir num sentido mais geral o que vem a ser uma base de um espaço vetorial V:

Definição 3.13 Seja V um espaço vetorial. Dizemos que um conjunto $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ de vetores de V é uma *base* para V se:

(a) o conjunto $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ for L.I. ;

(b) $[\vec{v}_1, \vec{v}_2, ..., \vec{v}_n] = V$.

A seguir apresentamos alguns exemplos.

Exemplo 1. Seja $V = \mathbb{R}^2$ e considere os vetores $\vec{v}_1 = (1,1)$ e $\vec{v}_2 = (0,1)$ de V. Note que

$$\alpha\vec{v_1} + \beta\vec{v_2} = \vec{0} \Leftrightarrow \alpha(1,1) + \beta(0,1) = (0,0),$$

e portanto, obtemos

$$\begin{cases} \alpha = 0 \\ \alpha + \beta = 0 \Rightarrow \beta = 0 \end{cases},$$

i.e., $\alpha = \beta = 0$, e disso segue que o conjunto de vetores $\{\vec{v}_1, \vec{v}_2\}$ é L.I.

Resta mostrar que o subespaço gerado por $\vec{v}_1$ e $\vec{v}_2$ é todo o $\mathbb{R}^2$, ou seja, que

$$[\vec{v}_1, \vec{v}_2, ..., \vec{v}_n] = \mathbb{R}^2.$$

De fato, seja $\vec{v} = (x, y)$ um vetor qualquer em $V = \mathbb{R}^2$. Assim, se supormos que

$$\vec{v} = \alpha\vec{v}_1 + \beta\vec{v}_2,$$

segue que

$$(x, y) = \alpha(1,1) + \beta(0,1) = (\alpha, \alpha + \beta),$$

e disso segue que

$$\alpha = x \;\text{ e }\; \alpha + \beta = y \Rightarrow x + \beta = y \Rightarrow \beta = y - x.$$

Assim, concluímos que, $\forall(x, y) \in \mathbb{R}^2$, podemos escrever

$$(x, y) = x(1,1) + (y - x)(0,1) = x\vec{v}_1 + (y - x)\vec{v}_2,$$

donde segue que $[\vec{v}_1, \vec{v}_2, ..., \vec{v}_n] = \mathbb{R}^2$.

Portanto, os vetores $\vec{v}_1 = (1,1)$ e $\vec{v}_2 = (0,1)$ formam uma base para o $\mathbb{R}^2$.

Outras bases para o $\mathbb{R}^2$ são, por exemplo, o conjunto de vetores $\{(0,1); (1,0)\}$, chamada de base canônica do $\mathbb{R}^2$, e também o conjunto de vetores $\{(3,-1); (2,0)\}$

e ficam como exercício para o leitor. Mas o conjunto de vetores $\{(1, -3); (-2, 6)\}$ não forma uma base para o $\mathbb{R}^2$. Por quê?

Exemplo 2. Os espaços de funções possuem bases compostas por um número infinito enumerável de vetores. Por exemplo, seja P o espaço vetorial de todos os polinômios de qualquer grau, munido das operações usuais. Nesse caso, não é difícil mostrar que, por exemplo, o conjunto infinito de vetores $\{1,\ x,\ x^2,\ x^3,\ ...\}$ forma uma base para P. Verifique!

Exemplo 3. Seja $\mathbb{Q}(\sqrt{2})$ o conjunto definido por

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}\ :\ a, b \in \mathbb{Q}\}.$$

Munindo esse conjunto com a adição

$$(a + b\sqrt{2}) + (p + q\sqrt{2}) = (a + p) + (b + q)\sqrt{2},$$

onde $a, b, p, q \in \mathbb{Q}$, e a multiplicação por escalar $\alpha \in \mathbb{Q}$,

$$\alpha(a + b\sqrt{2}) = \alpha a + \alpha b\sqrt{2},$$

não é difícil mostrar que $\mathbb{Q}(\sqrt{2})$ é um espaço vetorial sobre $\mathbb{Q}$ (não sobre os reais, pois os escalares α agora são tomados em $\mathbb{Q}$).

Isto posto, não é difícil mostrar que o conjunto $\{1,\ \sqrt{2}\}$ de vetores de $\mathbb{Q}(\sqrt{2})$ é uma base para $\mathbb{Q}(\sqrt{2})$. Fica como exercício para o leitor.

Teorema 3.14 *Seja $\{\vec{v}_1, ..., \vec{v}_n\}$ um conjunto de vetores não nulos de um espaço vetorial V que geram V. Então, podemos extrair uma base para V desse conjunto de vetores.*

Demonstração. Se o conjunto $\{\vec{v}_1, ..., \vec{v}_n\}$ for L.I., tal conjunto já é uma base para V e o Teorema está provado.

Suponhamos então que o conjunto $\{\vec{v}_1, ..., \vec{v}_n\}$ seja L.D. A ideia consiste então em conseguir "tirar" os vetores desse conjunto que nos atrapalham no que diz respeito à independência linear e que "não estragam" a geração de todo o V.

Assim, se tal conjunto for L.D., então pelo Teorema 3.12 segue que pelo menos um dos vetores do conjunto acima é combinação linear dos demais.

Sem perda de generalidade, assuma que $\vec{v}_n$ seja esse vetor. Então,

$$\vec{v}_n = \alpha_1\vec{v}_1 + ... + \alpha_{n-1}\vec{v}_{n-1}.$$

Logo, o conjunto $\{\vec{v}_1, ..., \vec{v}_{n-1}\}$ ainda gera V, pois sendo $\vec{v}_n$ uma combinação linear dos demais, este é desnecessário para a geração de V.

Se esse conjunto já for L.I., então o Teorema já está provado; mas se for L.D., então novamente pelo Teorema 3.12 segue que um deles é combinação linear dos demais, e portanto, desnecessário para a geração de V. Sem perda de generalidade, assuma que seja $\vec{v}_{n-1}$ esse vetor.

Assim, temos que tal vetor pode ser descartado da lista anterior de tal modo que o conjunto $\{\vec{v}_1, ..., \vec{v}_{n-1}\}$ ainda gera V. Se tal conjunto já for L.I., então o Teorema está provado, e caso contŕario, novamente poderemos escrever um dos vetores restantes como combinação linear dos demais, sendo este, portanto, descessário. Ou seja, seguindo esse raciocínio um número finito de vezes, chegaremos a um conjunto de vetores $\{\vec{v}_1, ..., \vec{v}_k\}$ que ainda gera V e é L.I, ou seja, uma base de V.

□

Teorema 3.15 *Seja V um espaço vetorial gerado por um número finito de vetores $\vec{v}_1, ..., \vec{v}_n$. Então, qualquer conjunto com mais de n vetores será L.D.*

Demonstração. Suponha que $V = [\vec{v}_1, ..., \vec{v}_n]$. Então, pelo Teorema anterior segue que podemos extrair dessa coleção uma base para V. Seja $\{\vec{v}_1, ..., \vec{v}_k\}$ tal base, com $k \leq n$. Sejam $\vec{w}_1$, ..., $\vec{w}_m$ m vetores de V, onde $m > n$. Vamos mostrar que a coleção $\{\vec{w}_1, ..., \vec{w}_m\}$ é L.D. Como tais vetores pertencem a V e

$\{\vec{v}_1, ..., \vec{v}_k\}$ é uma base, segue que existem escalares $a_{ij} \in \mathbb{R}$ tais que

$$\vec{w}_1 = a_{11}\vec{v}_1 + a_{12}\vec{v}_2 + ... + a_{1k}\vec{v}_k$$

$$\vec{w}_2 = a_{21}\vec{v}_1 + a_{22}\vec{v}_2 + ... + a_{2k}\vec{v}_k$$

$$\vdots$$

$$\vec{w}_m = a_{m1}\vec{v}_1 + a_{m2}\vec{v}_2 + ... + a_{mk}\vec{v}_k$$

Sejam $\alpha_1, ..., \alpha_m$ escalares tais que

$$\alpha_1\vec{w}_1 + ... + \alpha_m\vec{w}_m = \vec{0}.$$

Assim, levando as m iguadades acima para esta combinação linear, vamos obter

$$\alpha_1(a_{11}\vec{v}_1 + ... + a_{1k}\vec{v}_k) + ... + \alpha_m(a_{m1}\vec{v}_1 + ... + a_{mk}\vec{v}_k) = \vec{0},$$

ou seja,

$$(\alpha_1 a_{11} + \alpha_2 a_{21} + ... + \alpha_m a_{m1})\vec{v}_1 + ... + (\alpha_1 a_{1k} + ... + \alpha_m a_{mk})\vec{v}_k = \vec{0},$$

e como os vetores $\vec{v}_1$,..., $\vec{v}_k$ são L.I., obtemos

$$\begin{cases} a_{11}\alpha_1 + a_{21}\alpha_2 + ... + a_{m1}\alpha_m = 0 \\ \quad\vdots \\ a_{1k}\alpha_1 + a_{2k}\alpha_2 + ... + a_{mk}\alpha_m = 0 \end{cases},$$

um sistema linear homogêneo com k equações a m incógnitas α_1, α_2, ..., α_m. Como por construção $k \leq n < m$, segue que tal sistema homogêneo admite solução não trivial, e portanto, teremos algum $\alpha_j \neq 0$, e com isso segue que os vetores $\vec{w}_1$, $\vec{w}_2$, ..., $\vec{w}_m$ são L.D.

$\square$

Uma importante consequência do Teorema acima é seguinte resultado:

Corolário 3.16 *Qualquer base de um dado espaço vetorial V possui a mesma quantidade de vetores.*

Demonstração. Sejam $\{\vec{v}_1, ..., \vec{v}_n\}$ e $\{\vec{w}_1, ..., \vec{w}_m\}$ duas bases quaisquer de V. Vamos mostrar que $m = n$.

De fato, como $\vec{v}_1$, ..., $\vec{v}_n$ geram V e $\vec{w}_1$, ..., $\vec{w}_m$ são L.I., segue do Teorema anterior que $m \leq n$.

Do mesmo modo, como $\vec{w}_1$, ..., $\vec{w}_m$ geram V e $\vec{v}_1$, ..., $\vec{v}_n$ são L.I., segue do Teorema anterior que $n \leq m$.

Ou seja, $m = n$.

$\square$

Observando o resultado acima, que nos diz que qualquer base de um certo espaço vetorial tem sempre um mesmo número de elementos, é conveniente definir esse número comum como a *dimensão* do espaço vetorial dado. Ou seja, temos o seguinte conceito:

Definição 3.17 Chama-se *dimensão* de um espaço vetorial V o número de elementos da sua base.

Por exemplo, considere o espaço vetorial $\mathbb{R}^3$. Sabemos que $\dim \mathbb{R}^3 = 3$, pois qualquer base de $\mathbb{R}^3$ será formada por 3 vetores. Seja $W = \{(x, y, z) \in \mathbb{R}^3 : x = z \text{ e } y = -z\}$. É fácil ver que W é subespaço vetorial de V e deixamos isso para o leitor. Vamos determinar nesse exemplo qual seria a dimensão do subespaço W. Note que W pode ser escrito como

$$W = \{(z, -z, z) : z \in \mathbb{R}\} = \{z(1, -1, 1) : z \in \mathbb{R}\},$$

ou seja, W consiste no conjunto dos vetores múltiplos de $\vec{u} = (1, -1, 1)$. Ou seja, o espaço gerado por $\vec{u}$, denotado por $[(1, -1, 1)]$. É claro que $[(1, -1, 1)] = W$ e que o conjunto $\{(1, -1, 1)\}$ é L.I.[1]

[1]De fato, qualquer conjunto com um único vetor não nulo será L.I., pois, sendo $\vec{u} \neq \vec{0}$, então $\alpha\vec{u} = \vec{0}$ se, e somente se, $\alpha = 0$.

Portanto, segue que $\{(1, -1, 1)\}$ é uma base para W, e como tal base é formada por apenas um vetor, concluímos que $\dim W = 1$.

Geometricamente, observe que o conjunto W consiste numa reta passando pela origem do $\mathbb{R}^3$, possuindo $\vec{u} = (1, -1, 1)$ como vetor diretor, e uma reta realmente possui "dimensão um". Sugerimos o leitor fazer um desenho para ilustrar isso.

Mais geralmente, para $n \geq 1$ fixado, o espaço $\mathbb{R}^4$ possuirá dimensão n.

Considere $M(m, n)$ o espaço vetorial das matrizes de tamanho $m \times n$, munido das operações usuais de adição de matrizes e multiplicação de um escalar por uma matriz. Uma base para tal espaço seria o conjunto das matrizes onde todas as entradas são nulas, exceto uma, sempre em posições diferentes. Uma tal base possuirá, então, $m \cdot n$ vetores. Por exemplo,

$$\left\{\begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 0 \end{pmatrix}, ..., \begin{pmatrix} 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{pmatrix}\right\}$$

é uma base para $M(m, n)$. Neste caso, temos que $\dim M(m, n) = m \cdot n$.

Um terceiro exemplo interessante seria considerar P_n o espaço vetorial dos polinômios de grau menor ou igual a n. Nesse caso, como $\{1, x, x^2, ..., x^n\}$ é uma base para P_n (Verifique!), segue que $\dim P_n = n + 1$.

O próximo resultado estabelece uma forma de se obter uma base para um espaço vetorial V, conhecendo-se uma coleção de vetores L.I.

Teorema 3.18 *(Completamento) Qualquer conjunto de vetores L.I. de um espaço vetorial V pode ser completado de modo a formar uma base para V.*

Demonstração. Seja $\{\vec{v}_1, ..., \vec{v}_k\}$ um conjunto de vetores L.I. de V, e considere que $\dim V = n$, onde $n \geq k$.

Se $[\vec{v}_1, ..., \vec{v}_k] = V$, então o conjunto $\{\vec{v}_1, ..., \vec{v}_k\}$ será uma base para V, e nesse caso segue que $k = n$, c.f. a Definição 3.17.

Por outro lado, se $[\vec{v}_1, ..., \vec{v}_k] \neq V$, então existe pelo menos um vetor de V que não é combinação linear dos vetores $\vec{v}_1$,..., $\vec{v}_k$. Denotemos tal vetor por $\vec{v}_{k+1}$. Assim, temos que $\vec{v}_{k+1} \in V$, mas $\vec{v}_{k+1} \notin [\vec{v}_1, ..., \vec{v}_k]$.

Logo, segue que $\{\vec{v}_1, ..., \vec{v}_k, \vec{v}_{k+1}\}$ ainda é L.I., pois sendo $\{\vec{v}_1, ..., \vec{v}_k\}$ L.I., segue que

$$\alpha_1\vec{v}_1 + ... + \alpha_k\vec{v}_k = \vec{0} \Leftrightarrow \alpha_i = 0, \forall i \in \{1, 2, ..., k\}$$

Portanto, com tais α's temos que

$$\alpha_1\vec{v}_1 + ... + \alpha_k\vec{v}_k + \alpha_{k+1}\vec{v}_{k+1} = \vec{0} \Leftrightarrow \alpha_i = 0, \forall i \in \{1, 2, ..., k, k+1\}.$$

Assim, sendo $\{\vec{v}_1, ..., \vec{v}_{k+1}\}$ L.I., temos que, se $[\vec{v}_1, ..., \vec{v}_{k+1}] = V$, então o conjunto $\{\vec{v}_1, ..., \vec{v}_{k+1}\}$ já é uma base para V, mas se $[\vec{v}_1, ..., \vec{v}_{k+1}] \neq V$, segue que existe pelo menos um vetor de V que não é combinação linear dos vetores $\vec{v}_1$,..., $\vec{v}_{k+1}$, o qual vamos denotar por $\vec{v}_{k+2}$.

Assim, repetindo o raciocínio acima, temos que o conjunto $\{\vec{v}_1, ..., \vec{v}_{k+1}, \vec{v}_{k+2}\}$ também será L.I. e com isso segue que se $[\vec{v}_1, ..., \vec{v}_{k+2}] = V$, então o conjunto $\{\vec{v}_1, ..., \vec{v}_{k+2}\}$ já será uma base para V; e caso contrário existirá um vetor $\vec{v}_{k+3} \in V$ tal que $\vec{v}_{k+3} \notin [\vec{v}_1, ..., \vec{v}_{k+2}]$. Repetindo-se o processo um número finito de vezes chegaremos em um ponto onde a coleção $\{\vec{v}_1, ..., \vec{v}_k, ..., \vec{v}_\ell\}$ será finalmente L.I., com $[\vec{v}_1, ..., \vec{v}_\ell] = V$, sendo assim uma base para V.

□

Teorema 3.19 *(Teorema da dimensão) Sejam U e W dois subespaços de um espaço vetorial V. Se $\dim V < \infty$, então*

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

Demonstração. Seja $\beta_1 = \{z_1, ..., z_k\}$ uma base para $U \cap W$.

Note então que $\dim(U \cap W) = k$.

Como β_1 é L.I. em U e também em W, segue pelo Teorema 3.18 do completamento que existem $\vec{u}_1, ..., \vec{u}_\ell \in U$ e $\vec{w}_1, ..., \vec{w}_m \in W$ tais que

$$\beta_2 = \{\vec{z}_1, ..., \vec{z}_k, \vec{u}_1, ..., \vec{u}_\ell\} \text{ é base para } U,$$

e

$$\beta_3 = \{\vec{z}_1, ..., \vec{z}_k, \vec{w}_1, ..., \vec{w}_m\} \text{ é base para } W,$$

e notamos disso que $\dim U = k + \ell$ e $\dim W = k + m$.

Afirmammos que o conjunto $\beta = \{\vec{z}_1, ..., \vec{z}_k, \vec{u}_1, ..., \vec{u}_\ell, \vec{w}_1, ..., \vec{w}_m\}$ é uma base para o espaço $U + W$. De fato, basta mostrar (a) e (b) abaixo:

(a) $[\beta] = U + W$. De fato, seja $\vec{x} \in U + W$. Então, temos existem $\vec{u} \in U$ e $\vec{w} \in W$ tais queque $\vec{x} = \vec{u} + \vec{w}$.
Como β_2 é base para U e β_3 é base para W, podemos escrever

$$\vec{u} = \alpha_1 \vec{z}_1 + ... + \alpha_k \vec{z}_k + \beta_1 \vec{u}_1 + ... + \beta_\ell \vec{u}_\ell,$$

e

$$\vec{w} = a_1 \vec{z}_1 + ... + a_k \vec{z}_k + b_1 \vec{w}_1 + ... + b_m \vec{w}_m.$$

Assim,

$$\vec{x} = \vec{u} + \vec{w} = (\alpha_1 + a_1) z_1 + ... + (\alpha_k + a_k) z_k + \beta_1 \vec{u}_1 + ... + \beta_\ell \vec{u}_\ell + b_1 \vec{w}_1 + ... + b_m \vec{w}_m,$$

logo, vale (a).

(b) O conjunto β é L.I. De fato, seja

$$\alpha_1 \vec{z}_1 + ... + \alpha_k \vec{z}_k + \beta_1 \vec{u}_1 + ... + \beta_\ell \vec{u}_\ell + \gamma_1 \vec{w}_1 + ... + \gamma_m \vec{w}_m = \vec{0}. \quad (3.5)$$

Então, escrevendo

$$\underbrace{\alpha_1 \vec{z}_1 + ... + \alpha_k \vec{z}_k + \beta_1 \vec{u}_1 + ... + \beta_\ell \vec{u}_\ell}_{\in U} = \underbrace{-\gamma_1 \vec{w}_1 - ... - \gamma_m \vec{w}_m}_{\in W},$$

e então segue que $\gamma_1\vec{w}_1 - ... - \gamma_m\vec{w}_m \in U \cap W$, e com isso segue que existem escalares $c_1, ..., c_k \in \mathbb{R}$ tais que

$$-\gamma_1\vec{w}_1 - ... - \gamma_m\vec{w}_m = c_1\vec{z}_1 + ... + c_k\vec{z}_k,$$

e portanto,

$$c_1\vec{z}_1 + ... + c_k\vec{z}_k + \gamma_1\vec{w}_1 + ... + \gamma_m\vec{w}_m = \vec{0};$$

e como a base β_3 de W é L.I., segue que

$$c_1 = ... = c_k = \gamma_1 = ... = \gamma_m = 0.$$

Assim, sendo os $\gamma_j = 0$, segue que (3.5) fica escrito como

$$\alpha_1\vec{z}_1 + ... + \alpha_k\vec{z}_k + \beta_1\vec{u}_1 + ... + \beta_\ell\vec{u}_\ell = \vec{0}.$$

Como a base β_2 de U é L.I., segue que

$$\alpha_1 = ... = \alpha_k = \beta_1 = ... = \beta_\ell = 0.$$

Portanto, o conjunto β é de fato L.I., valendo também o item (b).

De (a) e (b) segue que β é uma base para $U + W$. Nesse caso, temos também que $\dim(U + W) = k + \ell + m$.

Assim, por construção temos que $\dim U \cap W = k$, $\dim U = k + \ell$, $\dim W = m + k$ e $\dim(U + W) = k + \ell + m$, e portanto,

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

□

Teorema 3.20 *Seja $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ uma base para o espaço vetorial V. Então, todo vetor $v \in V$ é escrito de maneira única como uma combinação linear dos vetores da base.*

Demonstração. Seja $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ uma base de V e seja $\vec{v} \in V$. Como $V = [\vec{v}_1, ..., \vec{v}_n]$, segue que existem $\alpha_1, ..., \alpha_n \in \mathbb{R}$ tais que

$$\vec{v} = \alpha_1\vec{v}_1 + ... + \alpha_n\vec{v}_n. \tag{3.6}$$

Mostremos a unicidade. De fato, se existirem $\beta_1, ..., \beta_n \in \mathbb{R}$ tais que

$$\vec{v} = \beta_1\vec{v}_1 + ... + \beta_n\vec{v}_n, \tag{3.7}$$

então, fazendo (3.6) $-$ (3.7) vamos obter

$$(\alpha_1 - \beta_1)\vec{v}_1 + ... + (\alpha_n - \beta_n)\vec{v}_n = \vec{0},$$

e como o conjunto $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ é L.I., segue que $\alpha_i = \beta_1$, para $i = 1, 2, ..., n$.
$\square$

Exercícios

1. Quais dos seguintes conjuntos formam uma base para $\mathbb{R}^3$?
   (a) $\{(1,1,-1); (2,-1,0); (3,2,0)\}$ (b) $\{(1,0,1); (0,-1,2); (-2,1,-4)\}$
   (c) $\{(2,1,-1); (-1,0,1); (0,0,1)\}$ (d) $\{(1,2,3); (4,1,2)\}$

2. Mostrar que os vetores $\overrightarrow{u} = (1,1,1)$, $\overrightarrow{v} = (1,2,3)$, $\overrightarrow{w} = (3,0,2)$ e $\overrightarrow{s} = (2,-1,1)$ geram o $\mathbb{R}^3$ e encontrar uma base dentre estes vetores.

3. Mostre que $\mathbb{R}^3 = [(1,1,1);\ (1,1,0);\ (0,1,1)]$.

4. Mostre que $P_3 = [x^2 + x^3;\ x;\ 2x^2 + 1;\ 3]$.

5. Encontre os valores de $a \in \mathbb{R}$ para que o conjunto

   $$\beta = \{(a,1,0); (1,a,1); (0,1,a)\}$$

   seja uma base para $\mathbb{R}^3$.

6. Mostre que os polinômios $1-t^3$, $(1-t^2)$, $1-t$ e 1 geram o espaço vetorial dos polinômios de grau ≤ 3.

7. (a) Um certo espaço vetorial V é gerado por cinco vetores LI. O que se pode dizer sobre a dimensão de V?

   (b) Um certo espaço vetorial V é gerado por cinco vetores LD. O que se pode dizer sobre a dimensão de V?

8. Seja $V = \mathbb{R}^3$ e o conjunto $\beta = \{(0,1,1); (1,1,0); (1,2,1)\} \subset \mathbb{R}^3$.

(a) Mostre que β não é uma base para $\mathbb{R}^3$.

(b) Determine uma base para $\mathbb{R}^3$ que possua dois elementos de β.

9. No espaço vetorial $\mathbb{R}^3$ consideremos os seguintes subespaços: $S = [(1,-1,2);(2,1,1)]$; $T = [(0,1,-1);(1,2,1)]$; $U = \{(x,y,z) : x+y = 4x - z = 0\}$ e $V = \{(x,y,z) : 3x - y - z = 0\}$. Determine as dimensões de $S, T, U, V, S+T$ e $S \cap T$.

10. Qual é a dimensão do espaço das matrizes 2×2 diagonais?

3.5 Mudança de base

Na seção anterior estudamos o importante conceito de base de um espaço vetorial, onde qualquer vetor de tal espaço fica escrito como uma combinação linear dos vetores da referida base, de maneira única (Teorema 3.20). Vimos também que um espaço vetorial V possui várias bases diferentes, e assim, dependendo do problema, representar os vetores de um certo espaço numa certa base pode vir a ser mais vantajoso do que representá-lo em outra base. Assim, precisamos desenvolver uma técnica para converter um dado vetor em um espaço de uma base para outra, ou seja, precisamos aprender a efetuar uma *mudança de base.*

3.5.1 Coordenadas de um vetor

Definição 3.21 Sejam $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ uma base de V e $\vec{v} \in V$ tal que $\vec{v} = \alpha_1\vec{v}_1 + ... + \alpha_n\vec{v}_n$. Dizemos que os escalares α_i são as *coordenadas* do vetor $\vec{v}$ na base β, e denotaremos por

$$[\vec{v}]_\beta = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

A matriz coluna acima representada, que comporta as coordenadas do vetor $\vec{v}$ na base β chama-se *matriz das coordenadas do vetor $\vec{v}$ na base β.*

Retornando ao Exemplo 1 da seção anterior, onde consideramos $V = \mathbb{R}^2$, vimos que o conjunto de vetores $\beta = \{\vec{v}_1, \vec{v}_2\}$, onde $\vec{v}_1 = (1,1)$ e $\vec{v}_2 = (0,1)$, forma uma base para o $\mathbb{R}^2$. Isto porque eles são L.I. e $[\vec{v}_1, \vec{v}_2] = \mathbb{R}^2$. Assim, dado $\vec{w} = (x,y) \in \mathbb{R}^2$ um vetor qualquer de $\mathbb{R}^2$, podemos escrever

$$\vec{w} = (x,y) = x(1,1) + (y-x)(0,1) = x \cdot \vec{v}_1 + (y-x) \cdot \vec{v}_2.$$

Dessa forma, dizemos que x e $y-x$ são as coordenadas do vetor (x,y) na base $\{\vec{v}_1, \vec{v}_2\}$, e escrevemos

$$[(x,y)]_\beta = \begin{bmatrix} x \\ y-x \end{bmatrix}$$

Já na base $\{\vec{i}, \vec{j}\}$, onde $\vec{i} = (1,0)$ e $\vec{j} = (0,1)$, temos que $\vec{w} = (x,y) \in \mathbb{R}^2$ é tal que

$$\vec{w} = (x,y) = x(1,0) + y(0,1) = x \cdot \vec{i} + y \cdot \vec{j};$$

ou seja, nesse caso temos que as coordenadas do vetor $\vec{w}$ coincidem com os escalares da combinação linear dos vetores $\vec{i}$ e $\vec{j}$. Devido a essa coincidência dizemos que $\{\vec{i}, \vec{j}\}$ chama-se *base canônica* do $\mathbb{R}^2$. Quando estamos na base canônica escreveremos simplesmente as coordenadas de (x,y) ba base canônica por $[(x,y)] = \begin{bmatrix} x \\ y \end{bmatrix}$.

Um fato importante que devemos notar consiste em perceber que os elementos da matriz das coordenadas ficam dispostos de acordo com a ordem em que os elementos da base β aparecem. Por exemplo, se $\beta = \{(0,1), (1,0)\}$ (base canônica do $\mathbb{R}^2$), e $\beta_1 = \{(1,0), (0,1)\}$, note que os vetores que compõe ambas as bases são os mesmos, mas não estão dispostos na mesma ordem. Assim, temos que $[(3,-2)]_\beta = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ mas $[(3,-2)]_{\beta_1} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

Por isso, vamos considerar de agora em diante uma base ser sempre *ordenada*, ou seja, os vetores da base estão ordenados na ordem em que já aparecem.

3.5.2 Mudança de base

Nesta seção mostraremos como efetuar a mudança de uma base para outra, de um dado vetor.

Sejam $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ e $\beta_1 = \{\vec{w}_1, ..., \vec{w}_n\}$ duas bases quaisquer de um espaço vetorial V (bases ordenadas, c.f. comentado na seção anterior). Dado $\vec{v} \in V$, temos que em cada uma das bases $\vec{v}$ possuirá uma representação única, a saber:

$$\vec{v} = x_1\vec{v}_1 + ... + x_n\vec{v}_n \ \text{(na base } \beta),$$

$$\vec{v} = y_1\vec{w}_1 + ... + y_n\vec{w}_n \ \text{(na base } \beta_1).$$

Associando as coordenadas de $\vec{v}$ nas bases β e β_1, respectivamente, temos

$$[\vec{v}]_\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad [\vec{v}]_{\beta_1} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Como $\{\vec{v}_1, ..., \vec{v}_n\}$ é base para V, segue que os vetores $\vec{w}_i \in V$ devem ser escritos também como combinação linear dos $\vec{v}_j$, ou seja,

$$\vec{w}_1 = a_{11}\vec{v}_1 + a_{21}\vec{v}_2 + ... + a_{n1}\vec{v}_n$$

$$\vec{w}_2 = a_{12}\vec{v}_1 + a_{22}\vec{v}_2 + ... + a_{n2}\vec{v}_n$$

$$\vdots$$

$$\vec{w}_n = a_{1n}\vec{v}_1 + a_{2n}\vec{v}_2 + ... + a_{nn}\vec{v}_n$$

Assim, como $\vec{v} = y_1\vec{w}_1 + ... + y_n\vec{w}_n$, escrevemos

$$\vec{v} = y_1(a_{11}\vec{v}_1 + a_{21}\vec{v}_2 + ... + a_{n1}\vec{v}_n) + ... + y_n(a_{1n}\vec{v}_1 + a_{2n}\vec{v}_2 + ... + a_{nn}\vec{v}_n) =$$

$$= (a_{11}y_1 + ... + a_{1n}y_n)\vec{v}_1 + (a_{21}y_1 + ... + a_{2n}y_n)\vec{v}_2 + ... + (a_{n1}y_1 + ... + a_{nn}y_n)\vec{v}_n,$$

e como $\vec{v} = x_1\vec{v}_1 + ... + x_n\vec{v}_n$ e as coordenadas de um vetor numa base são únicas (Teorema 3.20), segue que

$$x_1 = a_{11}y_1 + ... + a_{1n}y_n$$

$$x_2 = a_{21}y_1 + ... + a_{2n}y_n$$

$$\vdots$$

$$x_n = a_{n1}y_1 + ... + a_{nn}y_n,$$

o que, em notação matricial, fica:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Denotando $[I]_{\beta}^{\beta_1} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$, temos

$$[\vec{v}]_{\beta} = [I]_{\beta}^{\beta_1} \cdot [\vec{v}]_{\beta_1},$$

onde $[I]_{\beta}^{\beta_1}$ recebe o nome de *matriz de mudança da base β_1 para a base β.*

Exercícios

1. Mostre que cada conjunto a seguir é uma base para o $\mathbb{R}^2$. Em seguida, determine as coordenadas do vetor $\overrightarrow{u} = (6, 2)$ em relação a cada uma das bases dadas.
   (a) $\alpha = \{(3, 0); (0, 3)\}$ (b) $\beta = \{(1, 2); (2, 1)\}$
   (c) $\gamma = \{(1, 0); (0, 1)\}$ (d) $\delta = \{(0, 1); (1, 0)\}$

2. Quais são as coordenadas de $\overrightarrow{u} = (1, 0, 0)$ em relação à base
   $$\beta = \{(1, 1, 1); (-1, 1, 0); (1, 0, -1)\}$$
   do $\mathbb{R}^3$?

3. Determinar as coordenadas do vetor $u = (2, 1, 4) \in \mathbb{R}^3$ em relação às bases:
   (a) canônica; (b) $\beta = \{(1, 1, 1); (1, 0, 1); (1, 0, -1)\}$.

4. Mostre que os vetores $v_1 = (2, 6, 3)$, $v_2 = (1, 5, 4)$ e $v_3 = (-2, 1, 7)$ formam uma base do $\mathbb{R}^3$. Expresse o vetor $v = (3, 7, 1)$ como uma combinação linear de v_1, v_2 e v_3. Quais são as coordenadas de v em relação à base $\{v_1, v_2, v_3\}$?

5. Determinar as coordenadas do vetor $\overrightarrow{u} = (x, y, z)$ em relação a cada base do $\mathbb{R}^3$ dada.

   (a) $\beta = \{(1, 1, -1); (1, -1, 1); (-1, 1, 1)\}$

   (b) $\beta = \{(1, 0, 0); (1, 2, 1); (0, 5, 2)\}$

6. Ache a matriz de mudança de base da base $\beta = \{(1, 1, 0); (0, 1, 0); (0, 0, 3)\}$ para a base canônica do $\mathbb{R}^3$.

7. No espaço $\mathbb{R}^2$ consideremos as bases $\beta = \{e_1, e_2, e_3\}$ (canônica) e $\gamma = \{g_1, g_2, g_3\}$ relacionadas da seguinte maneira:

$$g_1 = e_1 + e_3$$

$$g_2 = 2e_1 + e_2 + e_3$$

$$g_3 = e_1 + e_2 + e_3.$$

   Determinar as matrizes de mudança de base de β para γ e de γ para β.

Capítulo 4

Transformações lineares

No capítulo anterior estudamos vários tipos de espaços vetoriais, seus subespaços, bem como propriedades peculiares. Neste capítulo definiremos um tipo especial de função entre dois espaços vetoriais $f : V \to W$ tal que preserva a estrutura de espaço vetorial, bem como propriedades importantes.

4.1 Transformação linear

Definição 4.1 Sejam V e W dois espaços vetoriais. Dizemos que que uma função $T : V \to W$ é uma *transformação linear* se, para quaisquer $\vec{u}, \vec{v} \in V$ e para qualquer $\alpha \in \mathbb{R}$ cumprirem as propriedades

(a) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$;

(b) $T(\alpha\vec{u}) = \alpha T(\vec{u})$.

Dada $T : V \to W$ uma transformação (linear ou não), o espaço vetorial V é chamado de domínio da transformação e o espaço vetorial W de contradomínio (ou codomínio) da transformação. O conjunto dos vetores $T(V) \subset W$ chama-se imagem da transformação. Sobre a imagem de uma transformação linear daremos uma atenção especial mais adiante.

É importante notar as operações de adição que estão envolvidas em (a): note que, ao escrever $T(\vec{u}+\vec{v}) = T(\vec{u})+T(\vec{v})$, a adição $\vec{u}+\vec{v}$ é a adição definida em V, já a adição $T(\vec{u}) + T(\vec{v})$ é a adição definida em W.

Observação. Quando $V = W$, chamamos a transformação linear $T : V \to V$ de *operador linear*. Quando $W = \mathbb{R}$ a transformação linear $T : V \to \mathbb{R}$ chama-se *funcional linear*.

A seguir apresentamos alguns exemplos (e contraexemplos) de transformações lineares.

Exemplo 1. A aplicação $T : \mathbb{R}^2 \to \mathbb{R}$ definida por $T(x, y) = 2x - 3y$ é uma transformação linear que manda vetores do espaço vetorial $\mathbb{R}^2$ para o espaço vetorial $\mathbb{R}$. De fato, dados $\vec{u} = (a, b)$ e $\vec{v} = (m, n) \in \mathbb{R}^2$ e $\alpha \in \mathbb{R}$, temos que

(a) $T(\vec{u} + \vec{v}) = T((a, b),(m, n)) = T((a + m, b + n)) = 2(a + m) - 3(b + n) =$

$$= (2a - 3b) + (2m - 3n) = T((a, b)) + T((m, n)) = T(\vec{u}) + T(\vec{v}),$$

(b) $T(\alpha\vec{u}) = T(\alpha(a, b)) = T(\alpha a, \alpha b) = 2\alpha a - 3\alpha b = \alpha(2a - 3b) =$

$$= \alpha T((a, b)) = \alpha T(\vec{u}).$$

Exemplo 2. Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $T(x, y) = (x + y, -y, x + 2y)$. Afirmamos que T é uma transformação linear do espaço vetorial $\mathbb{R}^2$ para o espaço vetorial $\mathbb{R}^3$. De fato, dados $\vec{u} = (a, b)$ e $\vec{v} = (m, n)$ vetores em $\mathbb{R}^2$ e $\alpha \in \mathbb{R}$, temos que

(a) $T(\vec{u}+\vec{v}) = T((a+m, b+n)) = (a+m+b+n, -(b+n), a+m+2(b+n)) =$

$$= (a + b, -b, a + 2b) + (m + n, -n, m + 2n) = T((a, b)) + T((m, n)) =$$

$$= T(\vec{u}) + T(\vec{v}).$$

(b) $T(\alpha\vec{u}) = T(\alpha(a, b)) = T((\alpha a, \alpha b)) = (\alpha a + \alpha b, -\alpha b, \alpha a + 2\alpha b) =$

$$= \alpha(a + b, -b, a + 2b) = \alpha T((a, b)) = \alpha T(\vec{u}).$$

Exemplo 3. A aplicação $T_1 : \mathbb{R} \to \mathbb{R}$ dada por $T_1(x) = 2x$ é uma transformação linear (verifique!). No entanto, $T_2 : \mathbb{R} \to \mathbb{R}$ dada por $T_2(x) = 2x + 1$ não é linear, a constante 1 "estraga" a linearidade. De fato, por exemplo, note que,

$$T_1(3) = 2 \cdot 3 + 1 = 7,$$

e no entanto, por exemplo,

$$T_1(3) = T_1(2 + 1),$$

mas

$$T_1(2) + T_1(1) = 2 \cdot 2 + 1 + 2 \cdot 1 + 1 = 8 \neq 7 = T_1(3).$$

Exemplo 4. Seja V o espaço vetorial das funções deriváveis em (a, b). Então, a aplicação $D : V \to C([a, b])$, dada por

$$D(f)(x) = f'(x),$$

ou seja, a aplicação D é a derivação. Como valem as regras de derivação

$$D(f + g)(x) = (f + g)'(x) = (f' + g')(x) = f'(x) + g'(x) = D(f)(x) + D(g)(x),$$

e, para $\alpha \in \mathbb{R}$, vale

$$D(\alpha f)(x) = (\alpha f)'(x) = (\alpha f')(x) = \alpha(f'(x)) = \alpha D(f)(x),$$

segue que o operador D acima definido é uma transformação linear.

Exemplo 5. Defina $T : C([a, b]) \to \mathbb{R}$ pondo

$$T(f) = \int_a^b f(x)dx.$$

Note que T é uma transformação linear, pois dados $f, g \in C([a, b])$ e $\alpha \in \mathbb{R}$, do Cálculo segue que

$$T(f + g) = \int_a^b (f + g)(x)dx = \int_a^b (f(x) + g(x))dx =$$

$$= \int_a^b f(x)dx + \int_a^b g(x)dx = T(f) + T(g),$$

e
$$T(\alpha f) = \int_a^b (\alpha f)(x)dx = \int_a^b \alpha f(x)dx = \alpha \int_a^b f(x)dx = \alpha T(f).$$

Exemplo 6. A aplicação $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x - y, x + y^2)$ não é linear. Justifique!

Exemplo 7. Seja $T : M(2, 2) \to \mathbb{R}^3$ dada por

$$T(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}) = (a_{11} + a_{12}, a_{21}, -a_{22}).$$

Afirmamos que T é uma transformação linear do espaço vetorial $M(2, 2)$ das matrizes quadradas de ordem 2×2 no espaço vetorial $\mathbb{R}^3$, verifique!

Proposição 4.2 *Seja $T : V \to W$ uma transformação linear do espaço vetorial V no espaço vetorial W. Valem as propriedades:*

(a) $T(0_V) = 0_W$, *onde 0_V denota o neutro aditivo de V e 0_W denota o neutro aditivo de W. Ou seja, uma transformação linear de V em W manda o neutro aditivo de V no neutro aditivo de W.*

(b) *Para qualquer $\vec{v} \in V$, $T(-\vec{v}) = -T(\vec{v})$.*

(c) *Para quaisquer $\vec{u}, \vec{v} \in V$, $T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v})$.*

Demonstração. (a) Basta notar que

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V),$$

onde a última igualdade é devida à linearidade da T. Como $T(0_V)$ é um vetor em W, existe o vetor oposto $-T(0_V)$ tal que, somando-o na igualdade acima, vamos obter

$$-T(0_V) + T(0_V) = -T(0_V) + T(0_V) + T(0_V),$$

e daí
$$0_W = 0_W + T(0_V) = T(0_V),$$

onde a última igualdade é deviada ao fato de 0_W ser o neutro aditivo de W, ou seja, concluímos que
$$T(0_V) = 0_W.$$

(b) Dado $\vec{v} \in V$, basta utilizar a propriedade (b) da definição de transformção linear considerando $\alpha = -1$:

$$T(-\vec{v}) = T(-1\vec{v}) = -1T(\vec{v}) = -T(\vec{v}).$$

(c) Dados $\vec{u}, \vec{v} \in V$, basta usar a linearidade da T e o item anterior:

$$T(\vec{u} - \vec{v}) = T(\vec{u} + (-\vec{v})) = T(\vec{u}) + T(-\vec{v}) = T(\vec{u}) - T(\vec{v}).$$

□

Proposição 4.3 *Sejam V e W dois espaços vetoriais, $\{\vec{v}_1, ..., \vec{v}_n\}$ uma base de V e $\vec{w}_1, ..., \vec{w}_n \in W$. Então, existe uma única transformação linear $T : V \to W$ tal que $T(\vec{v}_1) = \vec{w}_1$, $T(\vec{v}_2) = \vec{w}_2$, ..., $T(\vec{v}_n) = \vec{w}_n$.*

Demonstração. Como $\{\vec{v}_1, ..., \vec{v}_n\}$ é uma base de V, segue que existem únicos escalares $\alpha_1, ..., \alpha_n \in \mathbb{R}$ tais que

$$\vec{v} = \alpha_1\vec{v}_1 + ... + \alpha_n\vec{v}_n.$$

Defina $T : V \to W$ pondo

$$T(\vec{v}) = T(\sum_{i=1}^{n} \alpha_i\vec{v}_i) \stackrel{\text{def.}}{=} \sum_{i=1}^{n} \alpha_i\vec{w}_i.$$

Devido à unicidade dos α_i's, segue que T está bem definido, e além disso, notamos que $T(\vec{v}_i) = \vec{w}_i$, para $i = 1, 2, ..., n$.

Afirmamos que T assim definida é linear. De fato, para todo $\vec{u}, \vec{v} \in V$, temos que

$$\vec{u} = \sum_{i=1}^{n} \beta_i\vec{v}_i \quad \text{e} \quad \vec{v} = \sum_{i=1}^{n} \gamma_i\vec{v}_i,$$

e daí

$$T(\vec{u} + \vec{v}) = T(\sum_{i=1}^{n} \beta_i\vec{v}_i + \sum_{i=1}^{n} \gamma_i\vec{v}_i) = T(\sum_{i=1}^{n}(\beta_i + \gamma_i)\vec{v}_i) \stackrel{\text{def.}}{=} \sum_{i=1}^{n}(\beta_i + \gamma_i)\vec{w}_i =$$

$$= \sum_{i=1}^{n} \beta_i \vec{w}_i + \sum_{i=1}^{n} \gamma_i \vec{w}_i \overset{\text{def.}}{=} T(\sum_{i=1}^{n} \beta_i \vec{v}_i) + T(\sum_{i=1}^{n} \gamma_i \vec{v}_i) = T(\vec{u}) + T(\vec{v}),$$

e para $\alpha \in \mathbb{R}$, temos

$$T(\alpha \vec{u}) = T(\alpha \sum_{i=1}^{n} \beta_i \vec{v}_i) = T(\sum_{i=1}^{n} (\alpha\beta_i)\vec{v}_i) \overset{\text{def.}}{=} \sum_{i=1}^{n} (\alpha\beta_i)\vec{w}_i) =$$

$$= \alpha \sum_{i=1}^{n} \beta_i \vec{w}_i \overset{\text{def.}}{=} \alpha T(\sum_{i=1}^{n} \beta_i \vec{v}_i) = \alpha T(\vec{v}).$$

Por fim, mostremos a unicidade da T. Seja $S : V \to W$ transformação linear tal que $S(\vec{v}_i) = \vec{w}_i$. Assim, dado $\vec{v} = \sum_{i=1}^{n} \alpha_i \vec{v}_i \in V$, temos que

$$S(\vec{v}) = S(\sum_{i=1}^{n} \alpha_i \vec{v}_i) = \sum_{i=1}^{n} \alpha_i \vec{w}_i = T(\vec{v}).$$

□

Abaixo vejamos um exemplo de aplicação.

Exemplo. Determine a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(1, 0) = (1, 2, -1)$ e $T(0, 1) = (2, 2, 3)$.

Solução. Primeiramante, perceba que $\{(1, 0), (0, 1)\}$ é uma base do $\mathbb{R}^2$ (pois é um conjunto L.I. de dois vetores e a dimensão do $\mathbb{R}^2$ é 2). Assim esrevemos a combinação linear de um vetor (x, y) qualquer do $\mathbb{R}^2$ pondo

$$(x, y) = x(1, 0) + y(0, 1).$$

Como a T a ser determinada deve ser linear, segue que

$$T(x, y) = T(x(1, 0) + y(0, 1)) = x \cdot T(1, 0) + y \cdot T(0, 1).$$

Pelas condições impostas no problema, vem

$$T(x, y) = x(1, 2, -1) + y(2, 2, 3) = (x, 2x, -x, 2y, 2y, 3y) =$$

$$= (x + 2y, 2x + 2y, -x + 3y),$$

e tal T é única de acordo com a Proposição acima.

Exercícios

1. Mostre que a aplicação $F : \mathbb{R}^3 \to \mathbb{R}$ dada por $F(x, y, z) = 2x - 3y + z$ é uma transformação linear.

2. Mostre que a aplicação $T : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $T(x, y) = (x+1, 2y, x+y)$ não é uma transformação linear.

3. Seja V o espaço vetorial das matrizes $n \times n$ sobre $\mathbb{R}$. Seja B um elemento de V e considere a aplicação $\varphi_B : V \to V$ dada por

$$\varphi_B(A) = A \cdot B - B \cdot A.$$

   Mostre que φ_B é uma transformação linear.

4. Ache a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T(1, 0, 0) = (2, 0)$; $T(0, 1, 0) = (1, 1)$ e $T(0, 0, 1) = (0, -1)$. Em seguida, obtenha $\vec{v} \in \mathbb{R}^3$ tal que $T(\vec{v}) = (3, 2)$.

5. Ache a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(1, 2) = (2, 3)$ e $T(0, 1) = (1, 4)$.

6. Seja $T : V \to V$ uma transformação linear e C o conjunto dos vetores de V que são deixados fixos por T, ou seja,

$$C = \{\vec{v} \in V \ : \ T(\vec{v}) = \vec{v}\}.$$

   Mostre que C é um subespaço vetorial de V.

7. (Sel. Mestrado UFRGS/2007/2) Obtenha todas as transformações lineares $L : \mathbb{R}^3 \to \mathbb{R}^3$ tais que $L(u) = 3u$, $L(v) = 3v$ e $L(w) = 3w$, onde $u = (1, 0, 0)$, $v = (1, 1, 0)$ e $w = (0, 1, 1)$.

4.2 Operações com transformações lineares

Do mesmo modo que é feito no estudo de funções, definimos as operações de adição, subtração e composição de transformações lineares como segue.

Definição 4.4 Dadas $T_1, T_2 : V \to W$ duas transformações lineares. Definimos as aplicações $T_1 + T_2$ e $T_1 - T_2$, de V em W, respectivamente, por: para todo $\vec{v} \in V$,

$$(T_1 + T_2)(\vec{v}) = T_1(\vec{v}) + T_2(\vec{v}) \quad \text{e} \quad (T_1 - T_2)(\vec{v}) = T_1(\vec{v}) - T_2(\vec{v}).$$

é fácil ver que $T_1 + T_2$ e $T_1 - T_2$ são também transformações lineares e deixamos o leitor verificar estes fatos.

Exemplo. Dados $T_1, T_2 : \mathbb{R}^3 \to \mathbb{R}^2$, respectivamente, por

$$T_1(x, y, z) = (x + z, -x + 2y + 3z) \quad \text{e} \quad T_2(x, y, z) = (4x + 2y - 2z, 2y),$$

temos que $T_1 + T_2 : \mathbb{R}^3 \to \mathbb{R}^2$ é dada por

$$(T_1 + T_2)(x, y, z) = T_1(x, y, z) + T_2(x, y, z) =$$

$$= (x + z, -x + 2y + 3z) + (4x + 2y - 2z, 2y) =$$

$$= (5x + 2y - z, -x + 4y + 3z),$$

e $T_1 - T_2 : \mathbb{R}^3 \to \mathbb{R}^2$ é dada por

$$(T_1 - T_2)(x, y, z) = T_1(x, y, z) - T_2(x, y, z) =$$

$$= (x + z, -x + 2y + 3z) - (4x + 2y - 2z, 2y) =$$

$$= (-3x - 2y + z, -x + 3z).$$

Definição 4.5 Dadas as transformações lineares $T_1 : U \to V$ e $T_2 : V \to W$ (note que o contradomínio de T_1 é o domínio de T_2), definimos a *composta* $T_2 \circ T_1 : U \to W$, por, $\forall \vec{u} \in U$,

$$(T_2 \circ T_1)(\vec{u}) = T_2(T_1(\vec{u})).$$

Exemplo. Sejam $T_1 : \mathbb{R}^2 \to M(2, 2)$ e $T_2 : M(2, 2) \to \mathbb{R}^3$, dadas, respectivamente, por

$$T_1(x, y) = \begin{bmatrix} x + y & x \\ 2x - 3y & 2y \end{bmatrix} \quad \text{e} \quad T_2(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}) = (a_{11}, a_{11} + a_{22}, 2a_{12} - a_{21}).$$

Então $T_2 \circ T_1 : \mathbb{R}^2 \to \mathbb{R}^3$ será dada por

$$(T_2 \circ T_1)(x, y) = T_2(T_1(x, y)) = T_2(\begin{bmatrix} x+y & x \\ 2x-3y & 2y \end{bmatrix}) =$$
$$= (x + y, x + 3y, 3y).$$

A linearidade não se perde com a composição, conforme a Proposição abaixo:

Proposição 4.6 *Sejam $T_1 : U \to V$ e $T_2 : V \to W$ transformações lineares. Então $T_2 \circ T_1 : U \to W$ também é uma transformação linear.*

Demonstração. Dados $\vec{u}, \vec{v} \in U$ e $\alpha \in \mathbb{R}$, segue pela linearidade de T_1 e depois pela linearidade de T_2 que

$$(T_2 \circ T_1)(\vec{u} + \vec{v}) = T_2(T_1(\vec{u} + \vec{v})) = T_2(T_1(\vec{u}) + T_1(\vec{v})) =$$
$$= T_2(T_1(\vec{u})) + T_2(T_1(\vec{v})) = (T_2 \circ T_1)(\vec{u}) + (T_2 \circ T_1)(\vec{v}),$$

e

$$(T_2 \circ T_1)(\alpha\vec{u}) = T_2(T_1(\alpha\vec{v})) = T_2(\alpha T_1(\vec{v})) = \alpha T_2(T_1(\vec{v})) = \alpha(T_2 \circ T_1)(\vec{v}),$$

e portanto, $T_2 \circ T_1$ é uma transformação linear.

□

No que segue, vamos definir um conjunto especial.

Definição 4.7 Dados V e W dois espaços vetoriais. Definimos o conjunto

$$\mathcal{L}(V, W) = \{T : V \to W \ : \ T \text{ é transformação linear}\}.$$

Ou seja, fixados dois espaços vetoriais V e W, $\mathcal{L}(V, W)$ denota o conjunto de todas as transformações lineares de V em W. A partir da operação de adição de transformações lineares apresentada no início dessa seção e o produto de um escalar por uma transformação, segue o seguinte resultado:

Proposição 4.8 *Dados V e W dois espaços vetoriais. O conjunto $\mathcal{L}(V, W)$ munido da adição*

$$+ : \mathcal{L}(V, W) \times \mathcal{L}(V, W) \to \mathcal{L}(V, W),$$

$$(T_1, T_2) \mapsto T_1 + T_2,$$

e o produto por escalar

$$\cdot : \mathbb{R} \times \mathcal{L}(V, W) \to \mathcal{L}(V, W),$$

$$(\alpha, T_1) \mapsto \alpha \cdot T_1$$

é um espaço vetorial, denotado espaço vetorial de todas as transformações lineares de V em W.

Demonstração. Devem ser verificadas todas as oito propriedades da definição de espaço vetorial. Deixamos para o leitor.

□

Observação. Quando $V = W$, escrevemos $\mathcal{L}(V, V) = \mathcal{L}(V)$.

Exercícios

1. Sendo F, G e $H \in \mathcal{L}(\mathbb{R}^2)$ definidos por $F(x, y) = (x, 2y)$, $G(x, y) = (y, x+y)$ e $H(x, y) = (0, x)$, determinar $F + H$, $F \circ G$, $G \circ (H + F)$, $G \circ F$ e $H \circ F$.

2. Sejam S e T os operadores lineares em $\mathbb{R}^2$ definidos por $S(x, y) = (0, x)$ e $T(x, y) = (x, 0)$. Mostre que $S \circ T = 0$, mas que $T \circ S \neq 0$. Mostre também que $T^2 = T$.

4.3 Núcleo e imagem de uma transformação

Nesta seção apresentaremos os importantes conceitos de núcleo e imagem de uma transformação linear, bem como suas principais propriedades.

Definição 4.9 Seja $T : V \to W$ uma transformação linear. Definimos o *núcleo* de T, e denotamos por $\ker T$, o conjunto

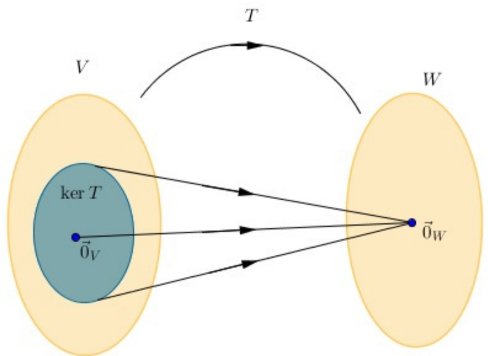
$$\ker T = \{\vec{v} \in V \ : \ T(\vec{v}) = \vec{0}_W\},$$

ou seja, o conjunto de todos os vetores de V que são levados pela T ao vetor nulo de W.

Pela definição acima notamos que $\ker T \subset V$. Uma outra notação para o núcleo de T é $\mathcal{N}(T)$.

Note que o núcleo de T está bem definido, pois de acordo com a Proposição 4.2, item (a), segue que para qualquer transformação linear $T : V \to W$, tem-se que $T(\vec{0}_V) = \vec{0}_W$, logo, $\vec{0}_V \in \ker T$, ou seja, o núcleo de uma transformação nunca é vazio, pois contém pelo menos o vetor nulo de V.

Abaixo temos uma ilustração clássica para a definição de $\ker T$.



O resultado a seguir mostra que núcleo de uma transformação linear é mais do que um subconjunto de V, ele tem estrutura de espação vetorial.

Proposição 4.10 *Seja $T : V \to W$ uma transformação linear entre os espaços vetoriais V e W. Então, $\ker T$ é um subespaço vetorial de V.*

Demonstração. Dados $\vec{u}, \vec{v} \in \ker T$ e $\alpha \in \mathbb{R}$. Precisamos mostrar que $\vec{u}+\vec{v} \in \ker T$ e que $\alpha\vec{u} \in \ker T$. De fato, pela linearidade da T segue que

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) = \vec{0} + \vec{0} = \vec{0},$$

e portanto concluímos que $\vec{u} + \vec{v} \in \ker T$, e

$$T(\alpha\vec{u}) = \alpha T(\vec{u}) = \alpha \cdot \vec{0} = \vec{0},$$

e portanto $\alpha\vec{u} \in \ker T$. Isso conclui a prova da Proposição.

$\square$

A seguir apresentamos alguns exemplos.

Exemplo 1. Determine o núcleo e a dimensão do núcleo de $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x + 2y, x - y)$.

Solução. Neste caso,

$$\ker T = \{(x, y) \in \mathbb{R}^2 \, : \, T(x, y) = (0, 0)\} = \{(x, y) \in \mathbb{R}^2 \, : \, (x+2y, x-y) = (0, 0)\},$$

e portanto o núcleo de T será a solução do sistema linear homogêneo

$$\begin{cases} x + 2y = 0 \\ x - y = 0 \end{cases},$$

cuja solução é apenas a trivial $(x, y) = (0, 0)$.

Portanto, neste caso, $\ker T = \{(0, 0)\}$, e disso $\dim \ker T = 0$.

Exemplo 2. Determine o núcleo e a dimensão do núcleo de $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $T(x, y, z) = (x + 2y - z, y + z)$.

Solução. Temos que

$$\ker T = \{(x, y, z) \in \mathbb{R}^3 \, : \, T(x, y, z) = (0, 0)\},$$

o que segue que o nócleo de T consiste na solução do sistema homogêneo

$$\begin{cases} x + 2y + z = 0 \\ y + z = 0 \end{cases},$$

donde segue que $y = -z$ e daí $x + 2(-z) + z = 0$ implica em $x = z$. Portanto, temos que

$$\ker T = \{(x, y, z) \in \mathbb{R}^3 \, : \, x = z \text{ e } \; y = -z\} = \{(z, -z, z) \, : \, z \in \mathbb{R}\} =$$

$$= \{z(1, -1, 1) \, : \, z \in \mathbb{R}\} = [(1, -1, 1)],$$

ou seja, $\ker T$ é o subespaço gerado pelo vetor $(1, -1, 1)$. Como um único vetor não nulo é L.I., segue que $\dim \ker T = 1$.

Neste caso podemos dar uma interpretação geométrica para $\ker T$: o núcleo de T é um subespaço vetorial do $\mathbb{R}^3$, e consiste na reta que passa pela origem e que possui o vetor $(1, -1, 1)$ como vetor diretor (faça um desenho).

Exemplo 3. Determine o núcleo e a dimensão do núcleo de $T : \mathbb{R}^3 \to \mathbb{R}$ dada por $T(x, y, z) = x + 2y - z$.

Solução. Temos que

$$\ker T = \{(x, y, z) \in \mathbb{R}^3 \, : \, T(x, y, z) = 0\} =$$

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

$$= \{x(1, 0, 1) + y(0, 1, 2) \, : \, x, y \in \mathbb{R}\} = [(1, 0, 1), (0, 1, 2)],$$

ou seja $\ker T$ é o subespaço gerado pelos vetores $(1, 0, 1)$ e $(0, 1, 2)$. Como tais vetores são L.I. (verifique!), segue que $\dim \ker T = 2$.

Neste exemplo também podemos dar uma interpretação geométrica para $\ker T$: consiste no plano do $\mathbb{R}^3$ passando pela origem e contendo os vetores $(1, 0, 1)$ e $(0, 1, 2)$. Faça um desenho.

Exemplo 4. Sendo P_n o espaço vetorial dos polinômios de grau menor ou igual a n, considere a transformação linear $D : P_n \to P_n$ dada por $P_n(p(x)) = p'(x)$, ou seja, a transformação linear é a derivação de polinômio. Então o núcleo de D será dado por

$$\ker D = \{p(x) \in P_n \, : \, D(p(x)) = 0\} = \{p(x) = a \, : \, a \in \mathbb{R}\},$$

ou seja, o núcleo de D corresponde ao subespaço vetorial de todos os polinômios constantes, visto que a derivada de uma função constante é a função nula.

No que segue apresentamos o conceito de injetividade em transformações lineares, o que, na verdade, consiste exatamente no conceito de função injetiva estudado no Cálculo.

Definição 4.11 Dizemos que uma transformação linear $T : V \to W$ é *injetiva* se, e somente se, para todo $\vec{u}, \vec{v} \in V$, tais que $\vec{u} \neq \vec{v}$, implicar em $T(\vec{u}) \neq T(\vec{v})$.

Em palavras, dizemos que uma transformação linear é injetiva se, e só se, domínios diferentes acarretarem em imagens também diferentes.

Equivalentemente, podemos definir a injetividade da seguinte maneira: $T : V \to W$ é injetiva se, e somente se, $\forall \vec{u}, \vec{v} \in V$ tal que $T(\vec{u}) = T(\vec{v})$, implicar em $\vec{u} = \vec{v}$.

Convém observar que uma transformação linear sempre manda neutro no neutro. Assim, se outro vetor não nulo for para o neutro mediante a T, ou seja, se $\ker T$ for um conjunto "maior" do que somente o zero, a injetividade da T já fica comprometida, i.e., T já não seria injetiva.

Isto posto, em consonância com o conceito de núcleo de uma transformação linear, temos o seguinte resultado:

Teorema 4.12 *Seja $T : V \to W$ uma transformação linear. Então, T é injetiva se, e somente se,* $\ker T = \{\vec{0}\}$.

Demonstração. Seja $T : V \to W$ uma transformação linear. Suponha que T é injetiva. Por absurdo, suponha que $\ker T \neq \{\vec{0}\}$. Logo, segue que existe $\vec{v} \neq \vec{0}$ tal que $\vec{v} \in \ker T$.

Assim, temos $T(\vec{v}) = \vec{0} = T(\vec{0})$, mas $\vec{v} \neq \vec{0}$, o que é um absurdo pois T é injetiva. Portanto, $\ker T = \{\vec{0}\}$.

Reciprocamente, suponha que $\ker T = \{\vec{0}\}$. Vamos mostrar que T é injetiva. Dados $\vec{u}, \vec{v} \in V$ tais que $T(\vec{u}) = T(\vec{v})$. Precisamos mostrar que $\vec{u} = \vec{v}$.

De fato, como $T(\vec{u}) = T(\vec{v})$, segue que $T(\vec{u}) - T(\vec{v}) = \vec{0}$, e pela linearidade da T vem que $T(\vec{u} - \vec{v}) = \vec{0}$, e portanto $\vec{u} - \vec{v} \in \ker T$. Como $\ker T = \{\vec{0}\}$, concluímos que $\vec{u} - \vec{v} = \vec{0}$, ou seja, $\vec{u} = \vec{v}$.

□

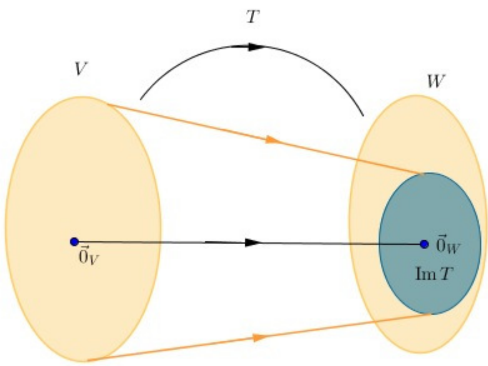
Com base nessa Proposição temos que a transformação T do Exemplo 1 dado acima é injetiva, já as transformações dos Exemplos 2 a 4 não são injetivas.

Definição 4.13 Seja $T : V \to W$ uma transformação linear. Definimos a *imagem* de T, e denotamos por $\operatorname{Im}(T)$ ou $T(V)$, o conjunto

$$\operatorname{Im}(T) = \{\vec{w} \in W \, : \, \exists \vec{v} \in V \text{ tal que } T(\vec{v}) = \vec{w}\}.$$

Pela própria definição de imagem de uma transformação linear temos que $\operatorname{Im}(T) \subset W$. Alé disso, $\operatorname{Im}(T) \neq \emptyset$ pois $\vec{0}_W = T(\vec{0}_V) \in \operatorname{Im}(T)$, e portanto, a imagem de T está bem definida.

Na figura abaixo apresentamos uma ilustração do conceito de imagem de uma transformação linear.



Abaixo mostramos um importante resultado que nos mostra que a imagem de uma transformação linear é mais do que um conjunto, i.e., $\operatorname{Im}(T)$ tem uma estrutura de espaço vetorial.

Proposição 4.14 *Seja $T : V \to W$ transformação linear. Então* $\mathrm{Im}(T)$ *é um subespaço vetorial de W.*

Demonstração. Dados $\vec{u}, \vec{v} \in \mathrm{Im}(T)$ e $\alpha \in \mathbb{R}$, temos que existem $\vec{x}, \vec{y} \in V$ tais que $T(\vec{x}) = \vec{u}$ e $T(\vec{y}) = \vec{v}$. Assim, usando a linearidade da T, temos

$$\vec{u} + \vec{v} = T(\vec{x}) + T(\vec{y}) = T(\vec{x} + \vec{y}) \in \mathrm{Im}(T),$$

pois $\vec{x} + \vec{y} \in V$, e

$$\alpha\vec{u} = \alpha T(\vec{x}) = T(\alpha\,\vec{x}) \in \mathrm{Im}(T),$$

pois $\alpha\vec{x} \in V$.

$\square$

Vejamos um exemplo. Considere $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $T(x, y, z) = (x + 2y - z, y + z)$, dada no Exemplo 2 acima. Naquele exemplo, vimos que $\dim \ker T = 1$, pois determinamos que $\ker T = [(1, -1, 1)]$. Vamos determinar a imagem de T e sua dimensão. Para isso, basta observar que

$$\mathrm{Im}(T) = \{(x + 2y - z, y + z) \,:\, x, y, z \in \mathbb{R}\} = \{x(1, 0) + y(2, 1) + z(-1, 1)\} =$$

$$= [(1, 0), (2, 1), (-1, 1)],$$

o que nos leva a crer, à primeira vista, que a dimensão da imagem de T seria 3. No entanto, como sabemos que trata-se de uma coleção de vetores do $\mathbb{R}^2$, certamente pelo menos um deles é combinação linear dos demais, ou seja, são L.D. Não é difícil ver que $(2, 1) = 3(1, 0) + 1(-1, 1)$, e disso, podemos descartar o vetor $(2, 1)$ e daí $[(1, 0), (-1, 1)] = \mathrm{Im}\,T$ e já é L.I. (Verifique!). Logo, concluímos que $\dim \mathrm{Im}\,T = 2$.

Nesse exemplo, observe que $\dim \ker T + \dim \mathrm{Im}\,T = 1 + 2 = 3 = \dim \mathbb{R}^3$. Tal resultado não é uma simples coincidência, como podemos observar no resultado abaixo.

Teorema 4.15 *Seja $T : V \to W$ uma transformação linear. Então,*

$$\dim \ker T + \dim \mathrm{Im}\,T = \dim V.$$

Demonstração. Seja $\{\vec{v}_1, ..., \vec{v}_k\}$ uma base de $\ker T$. Logo, $\dim \ker T = k$.

Como $\ker T$ é um subespaço vetorial de V, segue que a coleção $\{\vec{v}_1, ..., \vec{v}_k\}$ é L.I. em V.

Assim, segue pelo Teorema 3.18 do completamento que podemos completar esta coleção de modo a se obter uma base para V.

Dessa forma, seja o seguinte completamento $\{\vec{v}_1, ..., \vec{v}_k, \vec{w}_1, ..., \vec{w}_m\}$ uma base para V. Disso segue que $\dim V = k + m$.

Logo, temos que

$$\dim V = k + m = \dim \ker T + m.$$

Resta mostrar que $\dim \operatorname{Im} T = m$.

Afirmamos que o conjunto $\{T(\vec{w}_1), ..., T(\vec{w}_m)\}$ é uma base para $\operatorname{Im} T$. Isso será suficiente para provar o Teorema. Vamos à prova dessa afirmação:

(a) $[T(\vec{w}_1), ..., T(\vec{w}_m)] = \operatorname{Im} T$.

De fato, seja $\vec{w} \in W$. Então temos que $\exists \vec{v} \in V$ tal que $T(\vec{v}) = \vec{w}$.

Como $\{\vec{v}_1, ..., \vec{v}_k, \vec{w}_1, ..., \vec{w}_m\}$ uma base para V, segue que existem escalares $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_m \in \mathbb{R}$ tais que

$$\vec{v} = \alpha_1 \vec{v}_1 + ... + \alpha_k \vec{v}_k + \beta_1 \vec{w}_1 + ... + \beta_m \vec{w}_m.$$

Aplicando a T, e considerando a sua linearidade, vem

$$\vec{w} = T(\vec{v}) = \alpha_1 T(\vec{v}_1) + ... + \alpha_k T(\vec{v}_k) + \beta_1 T(\vec{w}_1) + ... + \beta_m T(\vec{w}_m).$$

Como para cada $i = 1, 2, ..., k$ tem-se que $\vec{v}_i \in \ker T$, segue que $\sum_{i=1}^{k} \alpha_i T(\vec{v}_i) = \vec{0}$, e daí a igualdade acima resulta em

$$\vec{w} = \beta_1 T(\vec{w}_1) + ... + \beta_m T(\vec{w}_m) = T(\sum_{j=1}^{m} \beta_j \vec{w}_j) \in \operatorname{Im} T.$$

Logo, vale **(a)**, o que indica que o subespaço gerado pelos vetores $T(\vec{w}_1)$, ..., $T(\vec{w}_m)$ é $\operatorname{Im} T$.

(b) $\{T(\vec{w}_1), ..., T(\vec{w}_m)\}$ é L.I.

De fato, sejam $\gamma_1, ..., \gamma_m \in \mathbb{R}$ tais que

$$\gamma_1 T(\vec{w}_1) + ... + \gamma_m T(\vec{w}_m) = \vec{0}.$$

Pela linearidade da T segue que

$$T(\gamma_1 \vec{w}_1 + ... + \gamma_m \vec{w}_m) = \vec{0}.$$

Logo, segue que $\gamma_1 \vec{w}_1 + ... + \gamma_m \vec{w}_m \in \ker T$. Como $\{\vec{v}_1, ..., \vec{v}_k\}$ é a base de $\ker T$, segue que existem escalares $a_1, ..., a_k$ reais tais que

$$\gamma_1 \vec{w}_1 + ... + \gamma_m \vec{w}_m = a_1 \vec{v}_1 + ... + a_k \vec{v}_k,$$

ou seja,

$$a_1 \vec{v}_1 + ... + a_k \vec{v}_k - \gamma_1 \vec{w}_1 - ... - \gamma_m \vec{w}_m = \vec{0}.$$

e como esta combinação linear está em termos da base de V, tais vetores são L.I., e disso a igualdade acima implica em todos os escalares serem nulos. Em particular, $a_1 = a_2 = ... = a_k = 0$, como queríamos. Logo vale **(b)**.

De **(a)** e **(b)** segue que $\{T(\vec{w}_1), ..., T(\vec{w}_m)\}$ é uma base para $\operatorname{Im} T$, donde segue que $\dim \operatorname{Im} T = m$.

□

Duas consequências imediatas do Teorema acima são os Corolários que seguem, os quais as demonstrações deixaremos a encrgo do leitor. Antes, porém, vejamos o conceito de transformação sobrejetiva.

Definição 4.16 Dizemos que uma transformação linear $T : V \to W$ é *sobrejetiva* se $T(V) = W$, ou seja, se a imagem Im T de V mediante T for todo o contradomínio W.

Uma outra maneira de definir a sobrejetividade da $T : V \to W$ é mosttrar que $\forall \vec{w} \in W$, $\exists \vec{v} \in V$ tal que $T(\vec{v}) = \vec{w}$.

Corolário 4.17 *Seja $T : V \to W$ uma transformação linear. Se $\dim V = \dim W$ então T é injetiva se, e somente se, T é sobrejetiva.*

Corolário 4.18 *Seja $T : V \to W$ uma transformação linear injetiva. Se $\dim V = \dim W$, então T leva base em base, i.e., se $\{\vec{v}_1, ..., \vec{v}_k\}$ é uma base de V, então $\{T(\vec{v}_1), ..., T(\vec{v}_k)\}$ será uma base de W.*

Exercícios

1. Seja $G : \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear dada por
$$G(x, y, z) = (x + 2y - z, y + z, x + y - 2z).$$
Encontre uma base e a dimensão do núcleo da G.

2. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ a aplicação linear dada por
$$T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z).$$
Determine uma base e a dimensão para o núcleo e para a imagem de T.

3. Sejam F e G operadores lineares de um espaço vetorial V. Mostre que $\ker G \subset \ker(F \circ G)$. Dê um exemplo onde vale a igualdade.

4. Sejam $F \in \mathcal{L}(U, V)$ e $G \in \mathcal{L}(V, W)$ tais que $\ker F = \{0\}$ e $\ker G = \{0\}$. Prove que $\ker(G \circ F) = \{0\}$.

5. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x, y, z) = (z, x - y, -z)$
   (a) Determine uma base do núcleo de T.
   (b) Dê a dimensão da imagem de T.

(c) T é sobrejetora? Justifique.

(d) Faça um desenho para $\ker T$ e $\operatorname{Im} T$.

6. Ache uma aplicação linear $T : \mathbb{R}^3 \to \mathbb{R}^4$ cuja imagem é gerada por $(1, 2, 0, -4)$ e $(2, 0, -1, -3)$.

7. (Sel. Mestrado UFSM/2013/2) Considere o $\mathbb{R}$-espaço vetorial $P_2(\mathbb{R})$, dos polinômios de graus menor ou igual a 2, e a transformação linear $T : \mathbb{R}^2 \to P_2(\mathbb{R})$ determinada por $T(1, 0) = 1 - t$ e $T(0, 1) = 1 - t^2$.

   (a) Encontre o núcleo e a imagem da transformação T.

   (b) Esta transformação é injetiva, sobrejetiva, bijetiva? Justifique sua resposta.

8. (Sel. Mestrado UFRGS/2011/2) Sejam U e V dois espaços vetoriais de dimensão finita sobre $\mathbb{R}$ e $T : U \to V$ uma transformação linear. Considere $S = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_k\} \subset U$.

   (a) Mostre que se $\{T(\vec{u}_1), T(\vec{u}_2), ..., T(\vec{u}_k)\}$ é linearmente independente em V, então S é linearmente independente em U.

   (b) Mostre que se T é injetora e S é linearmente independente em U, então $\{T(\vec{u}_1), T(\vec{u}_2), ..., T(\vec{u}_k)\}$ é linearmente independente em V.

9. (Sel. Mestrado UFRGS/2008/1) Seja P_n o espaço vetorial dos polinômios de grau menor ou igual a n:

$$P_n = \{f(t) = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n,\ t \in \mathbb{R}, a_i \in \mathbb{R}\}.$$

   (a) Obtenha uma base e a dimensão de P_4.

   (b) Mostre que P_3 é um subespaço vetorial de P_4.

   (c) Considere o operador linear $D : P_4 \to P_4$ definido como $Df(x) = \frac{df}{dx}(x)$. Obtenha a matriz da transformação linear D numa base conveniente de P_4.

   (d) O operador D do item (c) é injetivo? É sobrejetivo?

4.4 Isomorfismos e transformações inversas

Quando uma transformação linear T cumprir as Definições 4.11 e 4.16, obtemos a seguinte:

Definição 4.19 Dizemos que uma transformação linear $T : V \to W$ é *bijetiva* quando for injetiva e sobrejetiva.

O conceito acima recebe uma atenção especial na Álgebra linear para definir o conceito de isomorfismo entre espaços vetoriais, conceito esse que nos assegura que dois espaços vetoriais, num certo sentido, se comportam de maneira igual.

Definição 4.20 Dizemos que uma transformação linear $T : V \to W$ é um *isomorfismo* quando T for bijetiva. Neste caso, dizemos que os espaços vetoriais V e W são ditos *isomorfos*, e escrevemos $V \sim W$.

A seguir apresentamos alguns exemplos de espaços vetoriais isomorfos.

Exemplo 1. $\mathbb{R}^3 \sim P_2$, onde $P_2 = \{p(t) = a_0 + a_1 t + a_2 t^2 \,:\, a_0, a_1, a_2 \in \mathbb{R}\}$ é o espaço vetorial dos polinômios de graus menor ou igual a 2.

De fato, defina a aplicação $T : \mathbb{R}^3 \to P_2$ pondo

$$T(a_0, a_1, a_2) = a_0 + a_1 t + a_2 t^2.$$

É fácil ver que T é linear e deixamos esse detalhe para o leitor.

(i) T é injetiva: basta mostrar que $\ker T = \{0\}$, onde $0 = 0 + 0t + 0t^2$ denota o polinômio identicamente nulo. De fato,

$$\ker T = \{(a_0, a_1, a_2) \in \mathbb{R}^3 \,:\, T(a_0, a_1, a_2) = 0\} =$$

$$= \{(a_0, a_1, a_2) \in \mathbb{R}^3 \,:\, a_0 + a_1 t + a_2 t^2 = 0 + 0t + 0t^2\} = \{(0, 0, 0)\}.$$

Logo, pelo Teorema 4.12 segue que T é injetiva.

(ii) T é sobrejetiva: dado $p(t) = a + bt + ct^2$, segue que $\exists\, \vec{v} \in \mathbb{R}^3$ tal que $T(\vec{v}) = p(t)$. De fato, basta tomar $\vec{v} = (a, b, c)$.

Logo, de (i) e (ii) segue que T é bijetiva, e portanto, um isomorfismo, donde segue que $\mathbb{R}^3 \sim P_2$.

Exemplo 2. $\mathbb{C} \sim \mathbb{R}^2$, onde $\mathbb{C}$ denota o espaço vetorial dos números complexos munido da adição ususal de números complexos

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$

e a multiplicação usual por escalar

$$\alpha(a + bi) = \alpha a + \alpha bi.$$

De fato, defina $T : \mathbb{C} \to \mathbb{R}^2$ por

$$T(a + bi) = (a, b).$$

É fácil ver que T é linear e deixamos para o leitor.

(i) T é injetiva: de fato,

$$\ker T = \{z = a + bi \in \mathbb{C} \,:\, T(z) = (0, 0)\} =$$

$$= \{a + bi \,:\, (a, b) = (0, 0)\} = \{(0, 0)\},$$

logo, T é injetiva.

(ii) T é sobrejetiva: de fato, para todo $(x, y) \in \mathbb{R}^2$, $\exists z \in \mathbb{C}$ tal que $T(z) = (x, y)$. Basta tomar $z = x + yi$.

Logo, de (i) e (ii) segue que T é bijetiva, e portanto um isomorfismo entre $\mathbb{C}$ e $\mathbb{R}^2$, donde segue que $\mathbb{C} \sim \mathbb{R}^2$.

Exemplo 3. $M(2, 2) \sim \mathbb{R}^4$, onde $M(2, 2)$ denota o espaço vetorial das matrizes 2×2 com entradas reais, munido das suas operações usuais. De fato, basta definir $T : M(2,2) \to \mathbb{R}^4$ por

$$T\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = (a_{11}, a_{12}, a_{21}, a_{22}).$$

É fácil ver que T é linear e bijetiva, donde segue o isomorfismo. Deixamos para o leitor completar os detalhes.

Observe que se V e W são espaços vetoriais isomorfos, i.e., se $V \sim W$, então $\dim V = \dim W$. Veja, por exemplo, os três exemplos dados acima. Isso é devido ao Teorema 4.15. De fato, sendo $T : V \to W$ um isomorfismo, em particular T é injetivo, e disso $\ker T = \{0\}$, e com isso segue que $\dim \operatorname{Im} T = \dim V$, e como T em particular também é sobrejetiva, segue que $\operatorname{Im} T = W$, logo, $\dim V = \dim W$.

Sendo $T : V \to W$ isomorfismo (linear e bijetiva), a mesma admite uma transformação inversa $T^{-1} : W \to V$, que também será linear. De fato, a bijetividade da T garante a existência de uma transformação inversa, bastando apenas verificar que a linearidade continuará existindo em T^{-1}. Antes disso, vamos definir melhor este conceito de inversa, do mesmo modo que se faz em Cálculo. Antes, porém, é importante ressaltar a seguinte notação: sendo H um espaço vetorial qualquer, definimos a transformação linear $I_H : H \to H$, $I_H(\vec{u}) = \vec{u}$ a transformação linear identidade em H. É fácil ver que I_H é linear e deixamos a encargo do leitor. Isto posto, definimos:

Definição 4.21 Dizemos que uma transformação linear $T : V \to W$ é *inversível* se existir uma transformação $G : W \to V$ tal que

$$T \circ G : V \to V,\ (T \circ G)(\vec{v}) = I_V(\vec{v}) = \vec{v}$$

e

$$G \circ T : W \to W,\ (G \circ W)(\vec{w}) = I_W(\vec{w}) = \vec{w},$$

Nesse caso, dizemos que G é a transformação inversa da T e escrevemos $G = T^{-1}$ (e vice-versa).

Vejamos que T^{-1} continua linear. Considere $T : V \to W$ linear. Assim, dados $\vec{u}, \vec{v} \in V$ e $\alpha \in \mathbb{R}$ temos que

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) \;\text{ e }\; T(\alpha\vec{u}) = \alpha T(\vec{u}).$$

Disso, segue que existem $\vec{z}, \vec{w} \in W$ tais que $T(\vec{u}) = \vec{z}$ e $T(\vec{v}) = \vec{w}$ (e daí $\vec{z} = T^{-1}(\vec{z})$ e $\vec{v} = T^{-1}(\vec{w})$). Logo, $T^{-1} : W \to V$ será tal que (usando a linearidade da T)

$$T^{-1}(\vec{z} + \vec{w}) = T^{-1}(T(\vec{u}) + T(\vec{v})) = T^{-1}(T(\vec{u} + \vec{v})) =$$

$$= (T^{-1} \circ T)(\vec{u} + \vec{v}) = \vec{u} + \vec{v} = T^{-1}(\vec{z}) + T^{-1}(\vec{w}),$$

e

$$T^{-1}(\alpha\vec{z}) = T^{-1}(\alpha T(\vec{u})) = T^{-1}(T(\alpha\vec{u})) =$$

$$= (T^{-1} \circ T)(\alpha\vec{u}) = \alpha\vec{u} = \alpha T^{-1}(\vec{z}),$$

o que mostra a linearidade da T^{-1}.

Observando o Corolário 4.18 temos que uma transformação linear bijetiva manda base em base. Assim, uma maneira de se obter a inversa de uma transformação linear inversível é seguindo o seguinte procedimento apresentado no exemplo abaixo.

Exemplo. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ (neste caso $V = W = \mathbb{R}^3$) dada por

$$T(x, y, z) = (x + y - z, x + z, 2y - z).$$

Essa transformação T é sem dúvida linear. Vamos calcular o seu núcleo para verificar se é injetiva:

$$\ker T = \{(x, y, z) \in \mathbb{R}^3 \,:\, T(x, y, z) = (0, 0, 0)\},$$

ou seja, $\ker T$ corresponde à solução do sistema linear homogêneo

$$\begin{cases} x + y - z = 0 \\ x + z = 0 \\ 2y - z = 0 \end{cases},$$

que é apenas a terna ordenada $(0,0,0)$. Portanto, $\ker T = \{(0,0,0)\}$, donde segue que $T:\mathbb{R}^3 \to \mathbb{R}^3$ é injetiva. De acordo com o Corolário 4.17, T é também sobrejetiva. Portanto, T é bijetiva.

Logo, T é um isomorfismo e disso segue que existe $T^{-1}:\mathbb{R}^3 \to \mathbb{R}^3$. Como pelo Corolário 4.18, sendo T bijetiva, ela manda base para base, segue que, consifderando a base ordenada canônica de $V = \mathbb{R}^3$ $\{(1,0,0),(0,1,0),(0,0,1)\}$, temos que pela $T : V \to W$, obtemos

$$T(1,0,0) = (1,1,0);\ T(0,1,0) = (1,0,2);\ T(0,0,1) = (-1,1,-1);$$

que produzirá $\{(1,1,0),(1,0,2),(-1,1,-1)\}$ pela T uma base para $W = \mathbb{R}^3$.

Assim, para qualquer $(x,y,z) \in W = \mathbb{R}^3$, podemos escrever

$$(x,y,z) = \alpha(1,1,0) + \beta(1,0,2) + \gamma(-1,1,-1), \tag{4.1}$$

e assim obtemos o sistema linear nas incógnitas α, β e γ

$$\begin{cases} \alpha + \beta - \gamma = x \\ \alpha + \gamma = y \\ 2\beta - \gamma = z \end{cases},$$

o que nos fornece $\alpha = \frac{2x+y-z}{3}$, $\beta = \frac{y+2z-x}{3}$ e $\gamma = \frac{2y+z-2x}{3}$, e com isso, (4.1) fica

$$(x,y,z) = \frac{2x+y-z}{3}(1,1,0) + \frac{y+2z-x}{3}(1,0,2) + \frac{2y+z-2x}{3}(-1,1,-1).$$

Aplicando T^{-1}, que deverá ser linear, vem

$$T^{-1}(x,y,z) = \frac{2x+y-z}{3}T^{-1}(1,1,0) + \frac{y+2z-x}{3}T^{-1}(1,0,2) + \\ + \frac{2y+z-2x}{3}T^{-1}(-1,1,-1).$$

Como $T^{-1}:\mathbb{R}^3 \to \mathbb{R}^3$ será tal que

$$T^{-1}(1,1,0) = (1,0,0);\ T^{-1}(1,0,2) = (0,1,0) \ \text{ e } \ T^{-1}(-1,1,-1) = (0,0,1),$$

segue que

$$T^{-1}(x,y,z) = \frac{2x+y-z}{3}(1,0,0) + \frac{y+2z-x}{3}(0,1,0) + \frac{2y+z-2x}{3}(0,0,1),$$

ou seja, determinamos $T^{-1} : \mathbb{R}^3 \to \mathbb{R}^3$ por

$$T^{-1}(x,y,z) = (\frac{2x+y-z}{3}, \frac{y+2z-x}{3}, \frac{2y+z-2x}{3}).$$

De fato, o leitor pode verificar que $T \circ T^{-1}(x,y,z) = (x,y,z)$ e que $T^{-1} \circ T(x,y,z) = (x,y,z)$, o que mostra através da definição que tal T^{-1} é de fato a inversa de T.

Exercícios

1. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear dado por

$$T(x,y,z) = (2x,\, 4x-y,\, 2x+3y-z).$$

   (a) Determine o núcleo e a imagem de T e suas dimensões.

   (b) Calcule $T^2 = T \circ T$.

   (c) Mostre que T é invertível.

   (d) Determine a transformação inversa T^{-1}. Quais são seus núcleo e imagem?

4.5 Matriz de uma transformação linear

Nesta seção vamos apresentar a representação matricial de uma transformação linear $T : V \to W$. Veremos que cada transformação linear estará associada a uma matriz, onde a formação dessa matriz dependerá das bases de V e de W.

Seja $T : V \to W$ uma transformação linear, com $\dim V = n$ e $\dim W = m$. Considere $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ e $\gamma = \{\vec{w}_1, ..., \vec{w}_m\}$, respectivamente, as bases de V e de W. Dessa maneira, temos que os vetores $T(\vec{v}_1)$, ..., $T(\vec{v}_n)$ de W serão escritos como uma combinação linear dos vetores da base γ de W, ou seja,, existe únicos escalares a_{ij} tais que

$$T(\vec{v}_1) = a_{11}\vec{w}_1 + a_{21}\vec{w}_2 + ... + a_{m1}\vec{w}_m$$

$$T(\vec{v}_2) = a_{12}\vec{w}_1 + a_{22}\vec{w}_2 + ... + a_{m2}\vec{w}_m$$

$$\vdots$$

$$T(\vec{v}_n) = a_{1n}\vec{w}_1 + a_{2n}\vec{w}_2 + ... + a_{mn}\vec{w}_m$$

Devido à unicidade dos escalares a_{ij} determinados pelos equacionamentos acima, a matriz transposta dos coeficientes das combinações lineares acima definida será unicamente determinada, em dependência das base β e γ. Isso nos motiva apresentar uma definição matricial para uma transformação linear:

Definição 4.22 A matriz $A = (a_{ij})_{m\times n}$ definida pelos coeficientes determinados acima chama-se *matriz de T em relação às bases β e γ*, e denotamos por $[T]^{\beta}_{\gamma}$.

Abaixo apresentamos alguns exemplos.

Exemplo 1. Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por

$$T(x, y, z) = (3x + z, 2x - y + z).$$

Considerando $\beta = \{(1,1,1), (1,1,0), (1,0,0)\}$ e $\gamma = \{(1,3), (1,4)\}$, respectivamente, as bases do $\mathbb{R}^3$ e do $\mathbb{R}^2$, determinar a matriz da transformação $[T]^{\alpha}_{\gamma}$.

Solução. Como $T(1,1,1) = (4,2)$; $T(1,1,0) = (3,1)$ e $T(1,0,0) = (3,2)$, segue que escrevendo tais vetores como combinação linear dos vetores da base γ, vamos obter

$$(4,2) = a_{11}(1,3) + a_{21}(1,4)$$

$$(3,1) = a_{12}(1,3) + a_{22}(1,4)$$

$$(3,2) = a_{13}(1,3) + a_{23}(1,4)$$

Cada igualdade acima dará um sistema 2×2 e obteremos os escalares $a_{11} = 14$; $a_{21} = 10$; $a_{12} = 11$; $a_{22} = -8$; $a_{13} = 10$ e $a_{23} = -7$. Assim, a matriz da transformação T nas bases β e γ será

$$[T]_\gamma^\beta = \begin{bmatrix} 14 & 11 & 10 \\ -10 & -8 & -7 \end{bmatrix}.$$

Exemplo 2. Considere a mesma transformação do exemplo anterior, ou seja, $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $T(x, y, z) = (3x + z, 2x - y + z)$, mas as bases sendo as canônicas, i.e., $\beta = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ e $\gamma = \{(1, 0), (0, 1)\}$. Nesse caso, temos

$$T(1, 0, 0) = (3, 2) = 3(1, 0) + 2(0, 1)$$

$$T(0, 1, 0) = (0, -1) = 0(1, 0) - 1(0, 1)$$

$$T(0, 0, 1) = (1, 1) = 1(1, 0) + 1(0, 1),$$

e portanto, a matriz da transformação nessas bases será

$$[T]_\gamma^\beta = \begin{bmatrix} 3 & 0 & 1 \\ 2 & -1 & 1 \end{bmatrix}.$$

Repare que quando ambas as bases são canônicas, a matriz da transformação linear é obtida da T imediatamente, pois os coeficientes das coordenadas da imagem correspondem aos coeficientes da matriz procurada (Verifique!)

O próximo resultado nos mostra que, considerando as representações matriciais de transformações lineares, a matriz resultante da composição da mesmas corresponde ao produto das matrizes associadas.

Teorema 4.23 *Sejam $T_1 : V \to W$ e $T_2 : W \to U$ transformações lineares e α, β e γ, respectivamente, as bases de V, W e U. Então, a matriz associada à transformação linear composta $T_2 \circ T_1 : V \to U$ será*

$$[T_2 \circ T_1]_\gamma^\alpha = [T_2]_\gamma^\beta \cdot [T_1]_\beta^\alpha.$$

Demonstração. Sejam $\alpha = \{\vec{v}_1, ..., \vec{v}_k\}$, $\beta = \{\vec{w}_1, ..., \vec{w}_\ell\}$ e $\gamma = \{\vec{u}_1, ..., \vec{u}_m\}$, respectivamente, bases de V, W e U.

Assim, como $T(\vec{v}_1)$, ..., $T(\vec{v}_k) \in W$, segue que existem escalares $a_{ij} \in \mathbb{R}$ tais que

$$T_1(\vec{v}_j) = \sum_{i=1}^{\ell} a_{ij}\vec{w}_i, \text{ para } j = 1, 2, ..., k.$$

Do mesmo modo, como $T_2(\vec{w}_1)$, ..., $T_2(\vec{w}_m) \in U$, segue que existem escalares $b_{tp} \in \mathbb{R}$ tais que

$$T_2(\vec{w}_p) = \sum_{t=1}^{m} b_{tp}\vec{u}_t, \text{ para } p = 1, 2, ..., \ell.$$

Dessa forma, segue que

$$(b_{tp})_{m\times\ell} \cdot (a_{ij})_{\ell\times k} = [T_2]_\gamma^\beta \cdot [T_1]_\beta^\alpha, \tag{4.2}$$

uma matriz de ordem $m \times k$.

Por outro lado,

$$(T_2 \circ T_1)(\vec{v}_j) = T_2(T_1(\vec{v}_j)) = T_2(\sum_{i=1}^{\ell} a_{ij}\vec{w}_i) = \sum_{i=1}^{\ell} a_{ij}T_2(\vec{w}_i) =$$

$$= \sum_{i=1}^{\ell} a_{ij} \sum_{t=1}^{m} b_{ti}\vec{u}_t = \sum_{i=1}^{\ell}\sum_{t=1}^{m}(b_{ti}a_{ij})\vec{u}_t,$$

e daí segue que

$$[T_2 \circ T_1]_\beta^\alpha = (b_{ti}a_{ij})_{m\times k},$$

e com (4.2) temos

$$[T_2 \circ T_1]_\beta^\alpha = (b_{ti}a_{ij})_{m\times k} = (b_{ti})_{m\times\ell} \cdot (a_{ij})_{\ell\times k} = [T_2]_\gamma^\beta \cdot [T_1]_\beta^\alpha,$$

como queríamos mostrar.

$\square$

Vejamos um exemplo simples. Considere $T_1 : \mathbb{R}^3 \to \mathbb{R}^2$ e $T_2 : \mathbb{R}^2 \to \mathbb{R}^3$, dadas, respectivamente, por

$$T_1(x, y, z) = (2x + y - z, y + z) \text{ e } T_2(x, y) = (x, x + y, x - y).$$

Vamos determinar a matrizes dessas representações nas bases canônicas. É imediato que elas são dadas por

$$[T_1] = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{e} \quad [T_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Assim,

$$[T_2 \circ T_1] = [T_2] \cdot [T_1] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 2 & 0 \\ 2 & 0 & -2 \end{bmatrix}.$$

Ou seja, obtemos, nas bases canônicas, $T_2 \circ T_1 : \mathbb{R}^3 \to \mathbb{R}^3$,

$$(T_2 \circ T_1)(x, y, z) = (2x + y - z, 2x + 2y, 2x - 2z).$$

Exercícios

1. Sejam $\alpha = \{(1, -1), (0, 2)\}$ e $\beta = \{(1, 0, -1), (0, 1, 2), (1, 2, 0)\}$ bases de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente, e

$$[T]^\alpha_\beta = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}.$$

   (a) Ache T.

   (b) Se $S(x, y) = (2y, x - y, x)$, ache $[S]^\alpha_\beta$.

   (c) Ache uma base γ tal que $[T]^\alpha_\gamma = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$.

2. Se $[R] = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ e $[S] = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 1 \end{bmatrix}$, ache $R \circ S$.

3. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ uma transformação linear cuja matriz com relação à base canônica seja
$$\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}.$$
   (a) Determine $T(x, y, z)$.
   (b) Qual é a matriz do operador T com relação à base
$$\beta = \{(-1, 1, 0), (1, -1, 1), (0, 1, -1)\}\,?$$
   (c) O operador T é invertível? Justifique.

4. Determine a representação matricial de cada um dos seguintes operadores do $\mathbb{R}^2$ em relação às bases indicadas:
   (a) $F(x, y) = (2x, 3y - x)$ e base canônica.
   (b) $F(x, y) = (3x - 4y, x + 5y)$ e base $\beta = \{(1, 2),\ (2, 3)\}$.

5. Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ a transformação linear dada por
$$T(x, y, z) = (x + y, 2x - y + z).$$
   (a) Determine o núcleo de T, uma base para ele e sua dimensão.
   (b) Determine a imagem de T, uma base para ele e a sua dimensão.
   (c) Considerando as bases
$$\beta = \{(1, 1, 1), (-1, 0, 0), (0, 2, 0)\} \text{ e } \quad \gamma = \{(2, -1), (-1, 0)\}$$
   do $\mathbb{R}^3$ e do $\mathbb{R}^2$, respectivamente, determine a matriz da transformação $[T]^\beta_\gamma$.

6. (Sel. Mestrado UFSM/2015/2) Seja P_2 o espaço vetorial dos polinômios de grau menor ou igual a dois.
   (a) Verifique que $\beta = \{1 + t, -1 + t, t^2\}$ é uma base de P_2.
   (b) Verifique que $T : P_2 \to P_2$ definido por $T(f(t)) = \frac{df(t)}{dt} - f(t)$ é um operador linear.
   (c) Encontre a matriz que representa T com relação à base β.

Capítulo 5

Autovalores e autovetores

Neste capítulo estudaremos o conceito de autovetor e autovalor associados a uma transformação linear.

5.1 Conceito

Seja $T : V \to V$ um operador linear (ou seja, uma transformação linear de V em si mesmo). Perguntamos: quais são os vetores $\vec{v} \in V$ que são deixados invariantes mediante T, ou seja, quais são os pontos fixos de T, i.e., os vetores $\vec{v} \in V$ tais que $T(\vec{v}) = \vec{v}$? Mais geralmente, e o que norteará este capítulo, estamos interessados em determinar todos aqueles vetores, se existirem, tais que mandam pela T o vetor $\vec{v} \in V$ a um múltiplo de si mesmo. Definimos, então:

Definição 5.1 *Seja $T : V \to V$ um operador linear. Dizemos que $\vec{v} \in V$, $\vec{v} \neq \vec{0}$ é um autovetor* associado a T se existir um $\lambda \in \mathbb{R}$ tal que $T(\vec{v}) = \lambda\vec{v}$. Nesse caso, o escalar $\lambda \in \mathbb{R}$ chama-se *autovalor* associado ao autovetor T.

Por exemplo, seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (3x, 3y)$.

Nesse caso, vamos verificar se existe $\lambda \in \mathbb{R}$ tal que $T(\vec{v}) = \lambda\vec{v}$. Assim,

$$T(x, y) = \lambda(x, y) \Leftrightarrow (3x, 3y) = (\lambda x, \lambda y),$$

e disso obtemos o sistema

$$\begin{cases} 3x = \lambda x \\ 3y = \lambda y \end{cases}$$

que corresponde ao sistema linear homogêneo

$$\begin{cases} (3-\lambda)x = 0 \\ (3-\lambda)y = 0 \end{cases}.$$

O mesmo possuirá uma solução não trivial quando seu determinante for zero (regra de Cramer). Isto porque $\Delta x = \Delta y = 0$. Então, para possuir infinitas soluções, $\Delta = 0$, ou seja,

$$\begin{vmatrix} 3-\lambda & 0 \\ 0 & 3-\lambda \end{vmatrix} = 0 \Leftrightarrow (3-\lambda)^2 = 0 \Leftrightarrow \lambda = 3,$$

ou seja, $\lambda = 3$ é um autovalor. Vamos determinar os autovetores associados a tal autovalor. De fato, quando $\lambda = 3$, temos

$$T(x,y) = 3(x,y) \Leftrightarrow (3x,3y) = (3x,3y),\ \forall (x,y) \in \mathbb{R}^2,$$

ou seja, $\lambda = 3$ é autovalor para todo vetor $(x,y) \in \mathbb{R}^2$.

Vejamos um segundo exemplo. Considere $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$T(x,y,z) = (4x+z, 2x+3y+2z, x+4z).$$

Primeiramente, para determinar os autovalores de T, equacionamos

$$T(x,y,z) = \lambda(x,y,z),$$

o que nos fornece o sistema

$$\begin{cases} 4x + z = \lambda x \\ 2x + 3y + 2z = \lambda y \\ x + 4z = \lambda z \end{cases},$$

ou seja, obtemos o sistema linear homogêneo

$$\begin{cases} (4-\lambda)x + z = 0 \\ 2x + (3-\lambda)y + 2z = 0 \\ x + (4-\lambda)z = 0 \end{cases} .$$

Obviamente a solução trivial $(0,0,0)$ não nos interessa. Como pela regra de Cramer tal sistema possuirá infinitas soluções quando o determinante associado à matriz dos coeficientes for nula, $\Delta = 0$, será esse o caso de nosso interesse, ou seja, quando

$$\begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 0 & 4-\lambda \end{vmatrix} = 0.$$

Da igualdade acima vamos determinar uma equação de terceiro grau na incógnita λ, cujas raízes fornecerão os autovalores procurados. Com eles, determinamos em seguida os autovetores associados.

Na próxima seção voltaremos a este exemplo, resolvendo-o por completo. Vamos apresentar, primeiramente, a técnica geral que encerra o procedimento de busca dos autovalores e autovetores.

5.2 Procedimento para obter autovalores e autovetores

Seja $T : V \to V$ uma transformação linear de V em V. Considere a matriz $[T]$ da transformação na base canônica de V: (observe que tal matriz $[T]$ será quadrada).

$$[T] = \begin{bmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n_1} & ... & a_{nn} \end{bmatrix}.$$

Assim, queremos verificar se existe $\vec{v} = (x_1, ..., x_n) \in V$, com $\vec{v} \neq \vec{0}$, tal que $T(\vec{v}) = \lambda\vec{v}$, para algum $\lambda \in \mathbb{R}$.

Assumindo a igualdade acima, matricialmente, escrevemos

$$[T] \cdot \vec{v} = \lambda \vec{v} \Leftrightarrow [T] \cdot \vec{v} - \lambda \vec{v} = \vec{0} \Leftrightarrow$$

$$\Leftrightarrow ([T] - \lambda I) \cdot \vec{v} = \vec{0},$$

ou seja,

$$\left(\begin{bmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{bmatrix} - \begin{bmatrix} \lambda & ... & 0 \\ ... & ... & ... \\ 0 & 0 & \lambda \end{bmatrix} \right) \cdot \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ ... \\ 0 \end{bmatrix},$$

o que nos fornece

$$\begin{bmatrix} a_{11} - \lambda & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} - \lambda \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ ... \\ 0 \end{bmatrix}. \tag{5.1}$$

Como o sistema homogêneo acima possuirá solução não trivial (que nos interessa) se, e somente se, $\det([T] - \lambda I) = 0$, temos que os autovalores associados a T são determinados pelos zeros do polinômio $p(\lambda)$ definido por

$$p(\lambda) = \det([T] - \lambda I),$$

chamado de *polinômio característico*.

Assim, $p(\lambda) = 0$, chamada de *equação característica*, nos fornece os autovalores de T, ou seja, os autovalores de T são obtidos de

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} - \lambda & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n_1} & a_{n_2} & ... & a_{nn} - \lambda \end{vmatrix} = 0.$$

Sejam $\lambda_1, ..., \lambda_n$ os autovalores obtidos da equação acima. Para determinar os autovetores associados, precisamos resolver o sistema (5.1), para cada autovalor $\lambda = \lambda_j$, $j = 1, 2, ..., n$ encontrado.

Voltemos ao exemplo da seção anterior, vamos determinar os autovalores e autovetores de $T : \mathbb{R}^3 \to \mathbb{R}^3$ dado por

$$T(x, y, z) = (4x + z, 2x + 3y + 2z, x + 4z).$$

Para isso, a matriz associada à T, nas bases canônicas, é

$$[T] = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}.$$

Assim, o polinômio característico será dado por

$$p(\lambda) = \det([T] - \lambda I) = \begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 0 & 4-\lambda \end{vmatrix}.$$

Logo, os autovalores serão as raízes de $p(\lambda) = 0$, ou seja, quando

$$\begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 0 & 4-\lambda \end{vmatrix} = 0,$$

o que nos fornece

$$(3-\lambda)((4-\lambda)^2 - 1) = 0,$$

donde segue que obtemos dois autovalores $\lambda_1 = \lambda_2 = 3$ e $\lambda_3 = 5$. (um deles tem *multiplicidade* 2).

Para determinar os autovetores associados a $\lambda_1 = 3$ e $\lambda_3 = 5$, resolvemos, aplicando a igualdade (5.1) para $\lambda_1 = 3$ e depois para $\lambda_3 = 5$:

- $\lambda_1 = 3$: Nesse caso, temos

$$\begin{bmatrix} 4-3 & 0 & 1 \\ 2 & 3-3 & 2 \\ 1 & 0 & 4-3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

o que nos fornece o sistema linear

$$\begin{cases} x + z = 0 \\ 2x + 2z = 0 \\ x + z = 0 \end{cases},$$

que corresponde apenas à equação $x + z = 0$, e disso $z = -x$. Portanto, temos que todo vetor (x, y, z) do $\mathbb{R}^3$ tal que $z = -x$ será um autovator associado ao autovalor $\lambda_1 = 3$, i.e., $(x, y, -x)$, com $x, y, \in \mathbb{R}$, é autovalor associado ao autovetor $\lambda = 3$.

- $\lambda_3 = 5$: Nesse caso, temos

$$\begin{bmatrix} 4-5 & 0 & 1 \\ 2 & 3-5 & 2 \\ 1 & 0 & 4-5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

o que nos fornece o sistema linear

$$\begin{cases} -x + z = 0 \\ 2x - 2y + 2z = 0 \\ x - z = 0 \end{cases},$$

que corresponde ao sistema linear

$$\begin{cases} x - y + z = 0 \\ x - z = 0 \end{cases}.$$

Disso, segue pela segunda equação que $x = z$. Logo, a primeira equação nos fornecerá $y = 2z$. Portanto, todo vetor (x, y, z) do $\mathbb{R}^3$ tal que $x = z$ e $y = 2x$ será um autovetor associado ao autovalor $\lambda = 5$. Ou seja, para cada $z \in \mathbb{R}$, $(z, 2z, z)$ é autovetor associado ao autovalor $\lambda = 5$.

No que segue, apresentamos uma propriedade referente a autovalores e autovetores.

Proposição 5.2 *Seja $T : V \to V$ um operador linear. Se $\vec{v} \in V$ é autovetor de T associado ao autovalor λ, então, para qualquer constante $k \neq 0$, temos que $k\vec{v}$ também será um autovetor de T associado ao mesmo autovalor λ.*

Demonstração. De fato, se $\vec{v} \in V$ é autovetor de T associado ao autovalor λ, então, temos que

$$T(\vec{v}) = \lambda\vec{v}.$$

Dado $k \neq 0$, observamos que, pela linearidade de T, vem

$$T(k\vec{v}) = kT(\vec{v}) = k(\lambda\vec{v}) = \lambda(k\vec{v}),$$

donde segue o resultado.

□

Voltando ao estudo de matrizes, dizemos que duas matrizes A e B, de mesmo tamanho, são ditas *semelhantes*, se existir uma matriz inversível M tal que $B = M^{-1}AM$. Isto posto, temos o seguinte resultado.

Proposição 5.3 *Se A e B forem duas matrizes semelhantes, então elas possuem o mesmo polinômio característico, e portanto, possuem mesmos autovalores.*

Demonstração. De fato, se A e B são semelhantes, segue que existe uma matriz inversível M tal que $B = M^{-1}AM$. Sejam $p_A(\lambda)$ e $p_B(\lambda)$, respectivamente, o polinômio característico de A e de B. Assim,

$$p_B(\lambda) = \det(B - \lambda I) = \det(M^{-1}AM - \lambda I) =$$

$$= \det(M^{-1}AM - \lambda M^{-1}M) = \det(M^{-1}AM - M^{-1}(\lambda I)M) =$$

$$\det(M^{-1}(A - \lambda I)M) = \det M^{-1} \cdot \det(A - \lambda I) \cdot \det M =$$

$$= \det(A - \lambda I) = p_A(\lambda).$$

Assim, possuindo mesmo polinômio característico, segue que A e B possuirão mesmos autovalores.

□

Exercícios

1. Determine os autovalores e autovetores associados a cada transformação linear abaixo:

   (a) $T : \mathbb{R}2 \to \mathbb{R}^2$, $T(x, y) = (x + 2y, -x + 4y)$.

   (b) $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (x + y, x - y + 2z, 2x + y - z)$.

(c) $T : P_2 \to P_2$, $T(at^2 + bt + c) = at^2 + ct + b$.

(d) $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(x, y, z) = (x + y + z, 2y + z, 2y + 3z)$.

2. Determine os autovalores e autovetores das seguintes matrizes:

(a) $A = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}$ (b) $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$ (c) $A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix}$

(d) $A = \begin{pmatrix} 1 & 3 & -3 \\ 0 & 4 & 0 \\ -3 & 3 & 1 \end{pmatrix}$ (e) $A = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{pmatrix}$

3. Se um operador linear $T : V \to V$ é tal que $\lambda = 0$ é um autovalor associado a T, mostre que o operador T não é inversível.

4. Mostre que uma matriz A e a sua transposta A^t possuem mesmos autovalores.

Capítulo 6

Espaços com produto interno

Neste capítulo vamos introduzir o conceito de produto interno e suas principais propriedades.

6.1 Produto interno

Definição 6.1 Seja V um espaço vetorial. Dizemos que uma aplicação

$$\langle .,. \rangle : V \times V \to \mathbb{R},$$

$$(\vec{u}, \vec{v}) \mapsto \langle \vec{u}, \vec{v} \rangle \in \mathbb{R}$$

é um *produto interno* em V se satisfizer as seguintes propriedades: para todo $\vec{u}, \vec{v}, \vec{w} \in V$ e para todo $\lambda \in \mathbb{R}$ tivermos

(a) $\langle \vec{u}, \vec{u} \rangle \geq 0$ e $\langle \vec{u}, \vec{u} \rangle = 0 \Leftrightarrow \vec{u} = \vec{0}$. (positividade)

(b) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$. (simetria)

(c) $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$. (bilinearidade)

(d) $\langle \lambda\vec{u}, \vec{v} \rangle = \lambda\langle \vec{u}, \vec{v} \rangle$. (bilinearidade)

Note que de (d) segue o seguinte resultado: $\langle \vec{0}, \vec{v} \rangle = 0$. De fato, basta observar que

$$\langle \vec{0}, \vec{v} \rangle = \langle 0 \cdot \vec{u}, \vec{v} \rangle = 0 \langle \vec{u}, \vec{v} \rangle = 0.$$

A seguir apresentamos alguns exemplos de espaços com produto interno.

Exemplo 1. Seja $V = \mathbb{R}^3$. Dados $\vec{u} = (u_1, u_2, u_3)$ e $\vec{v} = (v_1, v_2, v_3)$, considere a aplicação $\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$, ou seja, o produto escalar usual de vetores em $\mathbb{R}^3$, c.f. o estudo em Geometria Analítica. Ou seja, vamos mostrar que o produto escalar usual do $\mathbb{R}^3$ é um produto interno. De fato, dados $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ e $\vec{w} = (w_1, w_2, w_3)$ vetores e $\lambda \in \mathbb{R}$, temos:

(a) $\langle \vec{u}, \vec{u} \rangle \geq 0$ pois $\langle \vec{u}, \vec{u} \rangle = u_1 u_1 + u_2 u_2 + u_3 u_3 = u_1^2 + u_2^2 + u_3^2 \geq 0$, e além disso, $\langle \vec{u}, \vec{u} \rangle = 0$ se, e somente se, $u_1^2 + u_2^2 + u_3^2 = 0$, ou seja, se, e somente se $(u_1, u_2, u_3) = (0, 0, 0)$.

(b) $\langle \vec{u}, \vec{v} \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3 = v_1 u_1 + v_2 u_2 + v_3 u_3 = \langle \vec{v}, \vec{u} \rangle$.

(c) $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$: De fato, basta usar da distributividade do produto escalar usual:

$$\langle \vec{u}, \vec{v} + \vec{w} \rangle = \vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w} = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle.$$

(d) $\langle \lambda \vec{u}, \vec{v} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle$: De fato,

$$\langle \lambda \vec{u}, \vec{v} \rangle = \langle (\lambda u_1, \lambda u_2, \lambda u_3), (v_1, v_2, v_3) \rangle = \lambda u_1 v_1 + \lambda u_2 v_2 + \lambda u_3 v_3 =$$

$$= \lambda (u_1 v_1 + u_2 v_2 + u_3 v_3) = \lambda \langle \vec{u}, \vec{v} \rangle.$$

Exemplo 2. Seja $V = C([a, b])$ o espaço vetorial das funções contínuas em $[a, b]$. Defina a aplicação $\langle , . , \rangle : V \times V \to \mathbb{R}$ por

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx.$$

Afirmamos que tal aplicação define um produto interno em $V = C([a, b])$. De fato, dados $f, g, h \in V$ e $\lambda \in \mathbb{R}$, temos que

(a) $\langle f, f\rangle \geq 0$ pois

$$\langle f, f\rangle = \int_a^b f(x)f(x)dx = \int_a^b (f(x))^2 dx \geq 0,$$

e $\langle f, f\rangle = 0$ se, e somente se, $\int_a^b (f(x))^2 dx = 0$, se, e somente se, $f = 0$ (isto porque f é contínua).

(b) $\langle f, g\rangle = \langle g, f\rangle$, pois

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle g, f\rangle.$$

(c) $\langle f, g+h\rangle = \langle f, g\rangle + \langle f, h\rangle$, pois

$$\langle f, g+h\rangle = \int_a^b f(x)(g(x)+h(x))dx = \int_a^b [f(x)g(x) + f(x)h(x)]dx =$$

$$= \int_a^b f(x)g(x)dx + \int_a^b f(x)h(x)dx = \langle f, g\rangle + \langle f, h\rangle.$$

(d) $\langle \lambda f, g\rangle = \lambda\langle f, g\rangle$, pois

$$\langle \lambda f, g\rangle = \int_a^b \lambda f(x)g(x)dx = \lambda \int_a^b f(x)g(x)dx = \lambda\langle f, g\rangle.$$

Exercícios

1. Mostre que a aplicação $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ dada por

$$\langle (x_1, x_2), (y_1, y_2)\rangle = x_2y_2 + 4x_1y_1$$

é um produto interno.

2. Dados $\vec{u} = (u_1, u_2)$ e $\vec{v} = (v_1, v_2)$, verifique se a expressão abaixo define um produto interno em $\mathbb{R}^2$:

$$\langle \vec{u}, \vec{v}\rangle = u_1v_1 - u_1v_2 - v_2u_1 + 3u_2v_2.$$

3. Sendo $\vec{u} = (u_1, u_2)$ e $\vec{v} = (v_1, v_2)$ vetores em $\mathbb{R}^2$, defina a aplicação

$$\langle \vec{u}, \vec{v}\rangle = \frac{u_1v_1}{a^2} + \frac{u_2v_2}{b^2},$$

onde $a, b \in \mathbb{R} \setminus \{0\}$ estão fixados. Prove que tal aplicação define um produto interno.

6.2 Ortogonalidade

O próximo conceito generaliza a ideia de perpendicularismo de vetores.

Definição 6.2 Dizemos que dois vetores $\vec{u}$ e $\vec{v}$ em um espaço vetorial V são *ortogonais,* e escrevemos $\vec{u} \perp \vec{v}$ se, e somente se, $\langle \vec{u}, \vec{v} \rangle = 0$.

Por exemplo, em $\mathbb{R}^3$ temos que os vetores da base canônica são dois a dois ortogonais, considerando o produto interno sendo o produto escalar usual da Geometria Analítica.

Um outro exemplo mais interessante é considerar $V = C([0, 2\pi])$, munido do produto interno definido por

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx.$$

Nesse caso, tomando como vetores as funções $f(x) = \operatorname{sen} x$ e $g(x) = \cos x$, temos que

$$\langle f, g \rangle = \int_0^{2\pi} \operatorname{sen} x \cos x dx = \int_0^{2\pi} (\operatorname{sen} x)^1 (\cos x dx) = \left. \frac{\operatorname{sen}^2 x}{2} \right|_0^{2\pi} =$$

$$= \frac{1}{2}(\operatorname{sen}^2 2\pi - \operatorname{sen}^2 0) = 0,$$

logo $f(x) = \operatorname{sen} x$ e $g(x) = \cos x$ são ortogonais em V, com o produto interno acima definido.

Abaixo encerramos algumas propriedades sobre ortogonalidade de vetores.

Proposição 6.3 *Dados $\vec{u}, \vec{v}, \vec{w}$ em um espaço vetorial V munido de um produto interno $\langle ., . \rangle$, valem as seguintes propriedades de ortogonalidade de vetores:*

(a) *Para qualquer $\vec{u} \in V$, $\vec{u} \perp \vec{0}$.*

(b) *Se $\vec{u} \perp \vec{v}$, então $\vec{v} \perp \vec{u}$.*

(c) *Se $\vec{u} \perp \vec{v}$ e $\vec{u} \perp \vec{w}$, então $\vec{u} \perp (\vec{v} + \vec{w})$.*

Demonstração. A prova é trivial e fica como exercício para o leitor.

Proposição 6.4 *Seja $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ um conjunto de vetores não nulos dois a dois ortogonais. Então, tal conjunto é L.I.*

Demonstração. Seja $\{\vec{v}_1, ..., \vec{v}_n\}$ conjunto de vetores não nulos, dois a dois ortogonais, ou seja, tais vetores cumprem a propriedade

$$\langle \vec{v}_i, \vec{v}_j \rangle = 0, \quad \text{se} \quad i \neq j$$

(pois $\langle \vec{v}_j, \vec{v}_j \rangle > 0$ visto que $\vec{v}_j \neq 0$). Sejam $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}$ tais que

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + ... + \alpha_n \vec{v}_n = \vec{0}. \tag{6.1}$$

Dessa forma, para cada $j \in \{1, 2, ..., n\}$ fixado, temos que

$$0 = \langle \vec{0}, \vec{v}_j \rangle = \langle \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + ... + \alpha_n \vec{v}_n, \vec{v}_j \rangle.$$

Usando das propriedades da bilinearidade da definição de produto interno, vamos obter

$$\alpha_1 \langle \vec{v}_1, \vec{v}_j \rangle + \alpha_2 \langle \vec{v}_2, \vec{v}_j \rangle + ... + \alpha_j \langle \vec{v}_j, \vec{v}_j \rangle + ... + \alpha_n \langle \vec{v}_n, \vec{v}_j \rangle = 0,$$

e como $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ para $i \neq j$, restará apenas

$$\alpha_j \langle \vec{v}_j, \vec{v}_j \rangle = 0,$$

onde $\langle \vec{v}_j, \vec{v}_j \rangle \neq 0$ pois $\vec{v}_j \neq 0$, donde segue que $\alpha_j = 0$.

Como o procedimento acima vale para todo $j \in \{1, 2, ..., n\}$, concluímos que $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$, o que, combinado com (6.1), segue que o conjunto $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ é L.I.

□

Exercícios

1. Dados os vetores $\vec{u} = (1, 1, 1)$, $\vec{v} = (1, 2, -3)$ e $\vec{w} = (1, -4, 3)$ no espaço vetorial $\mathbb{R}^3$, munido com o produto interno usual, verifique quais pares são ortogonais.

2. Mostre que as funções $f(x) = \operatorname{sen} x$ e $g(x) = \cos x$ no espaço vetorial $C([-\pi, \pi])$ com produto interno definido por $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$ são ortogonais.

6.2.1 Ortogonal de um conjunto

Definição 6.5 Seja S um subconjunto não vazio de um espaço vetorial V. Definimos o ortogonal de S, e denotamos por $S^{\perp}$, o conjunto

$$S^{\perp} = \{\vec{v} \in V \,:\, \langle \vec{v}, \vec{s} \rangle = 0, \,\forall \vec{s} \in S\}.$$

Ou seja, dado S um conjunto não vazio, em um espaço vetorial V, o ortogonal de S é o conjunto dos vetores de V que são ortogonais a todos os vetores de $S \neq \emptyset$, destacando $\vec{s} \in S$.

Note que $S^{\perp}$ é sempre não vazio, pois dado S um vetor qualquer de S, pela Proposição 6.3 segue que $\vec{0} \perp \vec{s}$, logo, $\vec{0} \in S^{\perp}$. Portanto, temos que o ortogonal de S está bem definido.

Um resultado interessante que destacamos consiste em que, mesmo que S não possua estrutura de espaço vetorial, segue que $S^{\perp}$ será um subespaço vetorial de V. Ou seja, temos o resultado que segue.

Proposição 6.6 *Seja S um subconjunto não vazio de um espaço vetorial V. Então o ortogonal de S é um subespaço vetorial de V.*

Demonstração. De fato, dados $\vec{u}, \vec{v} \in S^{\perp}$ e $\alpha, \in \mathbb{R}$, segue que, $\forall \vec{s} \in S$, temos que

$$\langle \alpha\vec{u}, \vec{s} \rangle = \alpha\langle \vec{u}, \vec{s} \rangle = \alpha \cdot 0 = 0,$$

logo, $\alpha\vec{u} \in S^{\perp}$. Além disso,

$$\langle \vec{u} + \vec{v}, \vec{s} \rangle = \langle \vec{u}, \vec{s} \rangle + \langle \vec{v}, \vec{s} \rangle = 0 + 0 = 0,$$

e portanto $\vec{u} + \vec{v} \in S^{\perp}$. Isso conclui a prova da Proposição.

□

6.3 Ortogonalização de Gram-Schmidt

Nesta seção, dada uma base β de um espaço vetorial, vamos determinar um procedimento para, a partir de tal base, determinar uma base β' que seja ortogonal. Tal procedimento é chamado de *processo de ortogornalização de Gram-Schmidt.*

Definição 6.7 Dizemos que uma base $\gamma = \{\vec{v}_1, ..., \vec{v}_n\}$ de um espaço vetorial V com um produto interno $\langle\, , \rangle$ é *ortogonal* se seus vetores forem dois a dois ortogonais, ou seja quando

$$\langle \vec{v}_i, \vec{v}_j \rangle = \begin{cases} 1 \ , \ \text{se} \ \ i = j \\ 0 \ , \ \text{se} \ \ i \neq j \end{cases} .$$

Por exemplo, em $V = \mathbb{R}^3$ munido do produto interno sendo o produto escalar, temos que a base canônica $\beta = \{(1,0,0),(0,1,0),(0,0,1)\}$ é ortogonal, pois seus vetores são todos dois a dois ortogonais.

Seja $\gamma = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ uma base ortogonal de um espaço vetorial V. Assim, dado $\vec{u} \in V$, segue que $\vec{u} = \alpha_1\vec{v}_1 + ... + \alpha_n\vec{v}_n$, e devido à ortogonalidade da base, podemos escrever

$$\langle \vec{u}, \vec{v}_j \rangle = \langle \alpha_1\vec{v}_1 + ... + \alpha_n\vec{v}_n, \vec{v}_j \rangle =$$

$$= \alpha_1\langle \vec{v}_1, \vec{v}_j \rangle + ... + \alpha_j\langle \vec{v}_j, \vec{v}_j \rangle + ... + \alpha_n\langle \vec{v}_n, \vec{v}_j \rangle = \alpha_j\langle \vec{v}_j, \vec{v}_j \rangle,$$

pois $\langle \vec{v}_k, \vec{v}_j \rangle = 0$ se $k \neq j$. Ou seja, obtemos uma maneira de determinar os escalares α_j, $j = 1, 2, ..., n$:

$$\alpha_j = \frac{\langle \vec{u}, \vec{v}_j \rangle}{\langle \vec{v}_j, \vec{v}_j \rangle}. \tag{6.2}$$

Isto posto, vamos à dedução do processo de ortogonalização.

Seja $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ uma base qualquer de um espaço vetorial V. Vamos determinar, a partir de β uma base que seja ortogonal. Vamos denotar tal base por $\beta' = \{\vec{v'}_1, ..., \vec{v'}_n\}$, onde tais vetores devem ser determinados a partir dos vetores de β, e deverão ser dois a dois ortogonais. Vamos considerar o primeiro vetor sendo igual ao primeiro vetor da base dada, ou seja, $\vec{v'}_1 = \vec{v}_1$. Observe a ilustração abaixo.

DESENHO...

Assim, para determinar $\vec{v'}_2$ temos que determinar o escalar α_1 tal que $\vec{v}_2 = \alpha_1\vec{v'}_1 + \vec{v'}_2$, ou seja

$$\vec{v'}_2 = \vec{v}_2 - \alpha_1\vec{v'}_1.$$

Por (6.2) obtemos

$$\vec{v'}_2 = \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{v'}_1\rangle}{\langle \vec{v'}_1, \vec{v'}_1\rangle} \cdot \vec{v'}_1.$$

Note que $\vec{v'}_2$ e $\vec{v'}_1$ são de fato ortogonais, pois

$$\langle \vec{v'}_2, \vec{v'}_1\rangle = \langle \vec{v}_2 - \alpha_1\vec{v'}_1, \vec{v'}_1\rangle = \langle \vec{v}_2, \vec{v'}_1\rangle - \alpha_1\langle \vec{v'}_1, \vec{v'}_1\rangle =$$

$$= \langle \vec{v}_2, \vec{v'}_1\rangle - \frac{\langle \vec{v}_2, \vec{v'}_1\rangle}{\langle \vec{v'}_1, \vec{v'}_1\rangle} \cdot \langle \vec{v'}_1, \vec{v'}_1\rangle = 0.$$

Seguindo o raciocínio, determinamos $\vec{v'}_3$ a partir dos vetores $\vec{v'}_1$ e $\vec{v'}_2$ determinados anteriormente. Ou seja, precisamos determinar $\alpha_1, \alpha_2 \in \mathbb{R}$ tais que

$$\vec{v'}_3 = \vec{v}_3 - \alpha_2\vec{v'}_2 - \alpha_1\vec{v'}_1,$$

de tal modo que $\langle \vec{v'}_3, \vec{v'}_1\rangle = \langle \vec{v'}_3, \vec{v'}_2\rangle = 0$. Felizmente, a escolha adequada para tais escalares é considerar α_1 o mesmo determinado acima e α_2 tendo a mesma configuração, ou seja, ambos são determinados conforme (6.2):

$$\alpha_1 = \frac{\langle \vec{v}_2, \vec{v'}_1\rangle}{\langle \vec{v'}_1, \vec{v'}_1\rangle} \quad \text{e} \quad \alpha_2 = \frac{\langle \vec{v}_3, \vec{v'}_2\rangle}{\langle \vec{v'}_2, \vec{v'}_2\rangle},$$

e com isso, obtemos

$$\vec{v'}_3 = \vec{v}_3 - \frac{\langle \vec{v}_3, \vec{v'}_2\rangle}{\langle \vec{v'}_2, \vec{v'}_2\rangle} \cdot \vec{v'}_2 - \frac{\langle \vec{v}_2, \vec{v'}_1\rangle}{\langle \vec{v'}_1, \vec{v'}_1\rangle} \cdot \vec{v'}_1.$$

Não é difícil verificar, e por essa razão deixaremos para o leitor, que

$$\langle \vec{v'}_3, \vec{v'}_2\rangle = \langle \vec{v'}_3, \vec{v'}_1\rangle = 0.$$

Seguindo por indução, concluímos que, no passo k, $1 \leq k \leq n$, o vetor $\vec{v'}_k$ da base ortogonal procurada será da forma

$$\vec{v'}_k = \vec{v}_k - \sum_{j=2}^{k} \frac{\langle \vec{v}_j, \vec{v'}_{j-1} \rangle}{\langle \vec{v'}_{j-1}, \vec{v'}_{j-1} \rangle} \cdot \vec{v'}_{j-1},$$

onde $\langle \vec{v'}_k, \vec{v'}_j \rangle = 0$, para $j = 1, 2, ..., k-1$.

Convém obsrevar ainda que, de acordo com a Proposição 6.4, temos que o conjunto $\beta' = \{\vec{v'}_1, ..., \vec{v'}_n\}$ é L.I., e como dim $V = n$, segue que β' realmente é uma base para V, e ortogonal, por construção.

Ou seja, acabamos de provar o Teorema:

Teorema 6.8 *(Processo de ortogonalização de Gram-Schmidt) Seja V um espaço vetorial munido de um produto interno $\langle\, , \,\rangle$ e $\beta = \{\vec{v}_1, ..., \vec{v}_n\}$ uma base qualquer de V. Então, podemos construir uma base ortogonal $\beta' = \{\vec{v'}_1, ..., \vec{v'}_n\}$ para V a partir da base β dada pondo*

$$\vec{v'}_1 = \vec{v}_1,$$

e para $k = 2, 3, ..., n$,

$$\vec{v'}_k = \vec{v}_k - \sum_{j=2}^{k} \frac{\langle \vec{v}_j, \vec{v'}_{j-1} \rangle}{\langle \vec{v'}_{j-1}, \vec{v'}_{j-1} \rangle} \cdot \vec{v'}_{j-1}.$$

Observação 6.9 Os vetores da forma

$$\alpha_j \vec{v'}_j = \frac{\langle \vec{v}_{j+1}, \vec{v'}_j \rangle}{\langle \vec{v'}_j, \vec{v'}_j \rangle} \cdot \vec{v'}_j$$

são chamados de projeção do vetor $\vec{v}_{j+1}$ sobre o vetor $\vec{v'}_j$, e denotamos por $\text{proj}_{\vec{v'}_j} \vec{v}_{j+1}$. Mais geralmente, dados dois vetores $\vec{v}, \vec{w}$ em um espaço vetorial V com produto interno, definimos o vetor *projeção* de $\vec{v}$ sobre $\vec{w}$, por

$$\text{proj}_{\vec{w}} \vec{v} = \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle} \cdot \vec{w}.$$

Repare que no caso de $V = \mathbb{R}^3$ e o produto interno ser o produto escalar, a fórmula projeção de um vetor $\vec{v}$ sobre um vetor $\vec{w}$ coincide com a fórmula estudada na Geometria Analítica, ou seja,

$$\mathrm{proj}_{\vec{w}}\vec{v} = \frac{\vec{v} \cdot \vec{w}}{|\vec{w}|^2} \cdot \vec{w},$$

onde $|\cdot|$ denota o módulo de um vetor do $\mathbb{R}^3$ estudado na Geometria Analítica.

6.4 Norma

O próximo conceito generaliza a noção de módulo de um vetor.

Definição 6.10 Seja V um espaço vetorial com um produto interno $\langle ., . \rangle$. Definimos a *norma* como sendo a aplicação $|| \cdot || : V \to [0, +\infty)$ dada por

$$||\vec{v}|| = \sqrt{\langle \vec{v}, \vec{v} \rangle}.$$

Dessa forma, dizemos que, se V for um espaço vetorial munido de um produto interno $\langle ., . \rangle$, a norma $|| \cdot ||$ é *induzida* pelo produto interno como feito na definição acima.

Um espaço vetorial V que possui norma é chamado de *espaço vetorial normado*. Da mesma definição de norma, temos que todo espaço vetorial com produto interno é normado, pois pode-se induzir uma norma pelo produto interno bastando fazer $||\vec{v}|| = \sqrt{\langle \vec{v}, \vec{v} \rangle}, \forall \vec{v} \in V$.

Como afirmamos antes da definição acima, o conceito de norma generaliza a noção de módulo de um vetor. Observe o primeiro exemplo dado a seguir.

Exemplo 1. Considere $V = \mathbb{R}^3$, munido do produto interno usual $\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v}$, ou seja, o produto interno é o produto escalar usual de vetores. Dessa forma, o módulo de um vetor $\vec{v} = (v_1, v_2, v_3)$ corresponde à norma do mesmo, pois nesse caso

$$||\vec{v}|| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{v_1 \cdot v_1 + v_2 \cdot v_2 + v_3 \cdot v_3} = \sqrt{v_1^2 + v_2^2 + v_3^2} = |\vec{v}|,$$

onde $|\vec{v}|$ denota o módulo do vetor $\vec{v}$.

Exemplo 2. Seja $V = C[a,b]$ o espaço vetorial de todas as funções contínuas em $[a,b]$. Já vimos que podemos definir o seguinte produto interno em V:

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Logo, desse produto interno induzimos a norma

$$||f|| = \sqrt{\langle f, f\rangle} = \left(\int_a^b f(x)^2 dx\right)^{\frac{1}{2}}.$$

Assim, por exemplo, em $C([0,1])$ temos que a norma de $f(x) = 2x$ será

$$||f|| = \sqrt{\langle f, f\rangle} = \left(\int_0^1 (2x)^2 dx\right)^{\frac{1}{2}} = \left(\int_0^1 4x^2 dx\right)^{\frac{1}{2}} =$$

$$= \left(\left.\frac{4x^3}{3}\right|_0^1\right)^{\frac{1}{2}} = \sqrt{\frac{4}{3} - 0} = \frac{2}{\sqrt{3}}.$$

A norma de um vetor satisfaz certas propriedades importantes, elencadas no Teorema abaixo.

Teorema 6.11 *Seja V um espaço vetorial com produto interno munido da norma $||\cdot||$ que provém do produto interno. Então para quaisquer $\vec{u}, \vec{v} \in V$ e para qualquer $\lambda \in \mathbb{R}$, valem as propriedades:*

(a) $||\vec{u}|| \geq 0$ *e* $||\vec{u}|| = 0$ *se, e somente se,* $\vec{u} = \vec{0}$.

(b) $||\lambda \cdot \vec{u}|| = |\lambda| \cdot ||\vec{u}||$.

(c) $|\langle \vec{u}, \vec{v}\rangle| \leq ||\vec{u}|| \cdot ||\vec{v}||$ *(desigualdade de Cauchy-Schwarz)*

(d) $||\vec{u} + \vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$ *(desigualdade triangular)*

Antes da Prova do Teorema acima, observe que convém comparar todas essas propriedades com as correspondentes estudadas na Geometria Analítica no $\mathbb{R}^3$, onde o produto interno deve ser substituído pelo produto escalar e a norma pelo módulo do vetor. assim, por exemplo, a propriedade (a) nos diz que um vetor do $\mathbb{R}^3$ possui "tamanho" positivo e somente o vetor nulo possui tamanho nulo; já (b) refere-se ao que acontece com o módulo de um vetor de multiplicarmos o mesmo por um $\lambda \in \mathbb{R}$; (d) é a desigualdade triangular usual entre vetores (podemos desenhar um triângulo onde os dois lados menores são os vetores dados e o terceiro representa a sua soma). Vamos à prova do Teorema.

Demonstração do Teorema. Dados $\vec{u}, \vec{v} \in V$ e $\lambda \in \mathbb{R}$. Como a norma provém do produto interno, temos que

(a) $||\vec{u}|| = \sqrt{\langle \vec{u}, \vec{u} \rangle} \geq 0$, e $||\vec{u}|| = 0 \Leftrightarrow \sqrt{\langle \vec{u}, \vec{u} \rangle} = 0 \Leftrightarrow \langle \vec{u}, \vec{u} \rangle = 0 \Leftrightarrow \vec{u} = 0$.

(b) $||\lambda \cdot \vec{u}|| = \sqrt{\langle \lambda \cdot \vec{u}, \lambda \cdot \vec{u} \rangle} = \sqrt{\lambda \cdot \langle \vec{u}, \lambda \cdot \vec{u} \rangle} = \sqrt{\lambda \cdot \langle \lambda \cdot \vec{u}, \vec{u} \rangle} = \sqrt{\lambda^2 \cdot \langle \vec{u}, \vec{u} \rangle} =$

$$= \sqrt{\lambda^2}\sqrt{\langle \vec{u}, \vec{u} \rangle} = |\lambda| \cdot ||\vec{u}||.$$

(c) $|\langle \vec{u}, \vec{v} \rangle| \leq ||\vec{u}|| \cdot ||\vec{v}||$ (conhecida como desigualdade de Cauchy-Schwarz):
Dado $t \in \mathbb{R}$, usando da bilinearidade do produto intereno, podemos escrever

$$0 \leq \langle t\vec{u} + \vec{v}, t\vec{u} + \vec{v} \rangle = \langle t\vec{u} + \vec{v}, t\vec{u} \rangle + \langle t\vec{u} + \vec{v}, \vec{v} \rangle =$$

$$= \langle t\vec{u}, t\vec{u} \rangle + \langle \vec{v}, t\vec{u} \rangle + \langle t\vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle =$$

$$= t^2 \langle \vec{u}, \vec{u} \rangle + t\langle \vec{u}, \vec{v} \rangle + t\langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{v} \rangle =$$

$$= t^2 \langle \vec{u}, \vec{u} \rangle + 2t\langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle,$$

ou seja, obtemos que

$$t^2 \langle \vec{u}, \vec{u} \rangle + 2\langle \vec{u}, \vec{v} \rangle t + \langle \vec{v}, \vec{v} \rangle \geq 0,$$

que é uma inequação de segundo grau na variável t. Considerando a função $f(t) = t^2 \langle \vec{u}, \vec{u} \rangle + 2\langle \vec{u}, \vec{v} \rangle t + \langle \vec{v}, \vec{v} \rangle$, temos que seu gráfico é uma parábola, e como $f(t) \geq 0$, segue que a mesma não intercepta o eixo horizontal t. Logo, para isso acontecer, o discriminante $\Delta = b^2 - 4ac$ da equação dever ser não positivo, ou seja,

$$\Delta \leq 0,$$

o que equivale dizer que

$$(2\langle\vec{u},\vec{v}\rangle)^2 - 4\cdot\langle\vec{u},\vec{u}\rangle\cdot\langle\vec{v},\vec{v}\rangle \leq 0,$$

e daí

$$(\langle\vec{u},\vec{v}\rangle)^2 \leq \cdot\langle\vec{u},\vec{u}\rangle\cdot\langle\vec{v},\vec{v}\rangle.$$

Extraindo a raiz quadrada, e notando que $\langle\vec{u},\vec{u}\rangle, \langle\vec{v},\vec{v}\rangle \geq 0$, vem

$$\sqrt{(\langle\vec{u},\vec{v}\rangle)^2} \leq \sqrt{\langle\vec{u},\vec{u}\rangle\cdot\langle\vec{v},\vec{v}\rangle} = \sqrt{\langle\vec{u},\vec{u}\rangle}\cdot\sqrt{\langle\vec{v},\vec{v}\rangle},$$

ou seja,

$$|\langle\vec{u},\vec{v}\rangle| \leq ||\vec{u}||\cdot||\vec{v}||,$$

como queríamos mostrar.

(d) $||\vec{u}+\vec{v}|| \leq ||\vec{u}|| + ||\vec{v}||$ (desigualdade triangular): Basta notar que

$$||\vec{u}+\vec{v}||^2 = \langle\vec{u}+\vec{v},\vec{u}+\vec{v}\rangle = \langle\vec{u}+\vec{v},\vec{u}\rangle + \langle\vec{u}+\vec{v},\vec{v}\rangle =$$

$$= \langle\vec{u},\vec{u}\rangle + \langle\vec{v},\vec{u}\rangle + \langle\vec{u},\vec{v}\rangle + \langle\vec{v},\vec{v}\rangle = \langle\vec{u},\vec{u}\rangle + 2\cdot\langle\vec{u},\vec{v}\rangle + \langle\vec{v},\vec{v}\rangle =$$

$$= ||\vec{u}||^2 + 2\cdot\langle\vec{u},\vec{v}\rangle + ||\vec{v}||^2,$$

e como $\langle\vec{u},\vec{v}\rangle \leq |\langle\vec{u},\vec{v}\rangle|$, obtemos

$$||\vec{u}+\vec{v}||^2 \leq ||\vec{u}||^2 + 2|\langle\vec{u},\vec{v}\rangle| + ||\vec{v}||^2.$$

Pela desigualdade de cauchy-Schwarz, vamos obter

$$||\vec{u}+\vec{v}||^2 \leq ||\vec{u}||^2 + 2|\langle\vec{u},\vec{v}\rangle| + ||\vec{v}||^2 \leq ||\vec{u}||^2 + 2||\vec{u}||\cdot||\vec{v}|| + ||\vec{v}||^2,$$

ou seja,

$$||\vec{u}+\vec{v}||^2 \leq (||\vec{u}|| + ||\vec{v}||)^2,$$

e extraindo a raiz quadrada, segue que

$$||\vec{u}+\vec{v}|| \leq |\vec{u}|| + ||\vec{v}||\,, \quad \text{como queríamos mostrar.}$$

$\square$

Abaixo apresentamos um exemplo de aplicação da desigualdade de Cauchy-Schwarz.

Exemplo. Use a desigualdade de Cauchy-Schwarz em $\mathbb{R}^3$ para mostrar que dados valores reais positivos a_1, a_2 e a_3, vale

$$(a_1 + a_2 + a_3)\left(\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}\right) \geq 9.$$

Solução. Como $a_1, a_2, a_3 > 0$, vamos considerar os seguintes vetores do $\mathbb{R}^3$

$$\vec{u} = (\sqrt{a_1}, \sqrt{a_2}, \sqrt{a_3}) \text{ e } \vec{v} = (\frac{1}{\sqrt{a_1}}, \frac{1}{\sqrt{a_2}}, \frac{1}{\sqrt{a_3}}).$$

Assim, considerando a norma $||\cdot||$ de um vetor como sendo o proveniente do produto interno (no caso, escalar) usual do $\mathbb{R}^3$, obteremos

$$||\vec{u}|| = \sqrt{a_1 + a_2 + a_3} \text{ e } ||\vec{v}|| = \sqrt{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}}.$$

Pela desigualdade de Cauchy-Schwarz $|\langle \vec{u}, \vec{v} \rangle| \leq ||\vec{u}|| \cdot ||\vec{v}||$, obtemos

$$\sqrt{a_1} \cdot \frac{1}{\sqrt{a_1}} + \sqrt{a_2} \cdot \frac{1}{\sqrt{a_2}} + \sqrt{a_3} \cdot \frac{1}{\sqrt{a_3}} \leq \sqrt{a_1 + a_2 + a_3} \cdot \sqrt{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}},$$

ou seja,

$$3 \leq \sqrt{a_1 + a_2 + a_3} \cdot \sqrt{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}}.$$

Por fim, elevando ambos os membros aos quadrado obtemos

$$(a_1 + a_2 + a_3)\left(\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3}\right) \geq 9.$$

Observação 6.12 É importante registrar nesse ponto que o conceito de norma em um espaço vetorial é mais geral. De fato, podem existir espaços vetoriais que são normados sem a existência de produto interno, mas isso não é estudado em um curso de Álgebra Linear. Nesse caso, uma norma $||\cdot|| : V \to [0, +\infty)$ é toda a aplicação que satisfizer as propriedades (a), (b) e (d) do Teorema 6.11. Fica apenas como curiosidade esta observação.

Observação 6.13 Observamos também que podemos definir vários tipos de produto interno em um espaço vetorial V. Dessa forma, existirão, consequentemente, vários tipos de normas no mesmo espaço. Porém, todas serão equivalentes num certo sentido, mas isso também foge de um primeiro curso de Álgebra Linear.

Exercícios

1. Para cada vetor $\vec{v}$ do $\mathbb{R}^3$ dado a seguir, determine as normas que provém do produto interno usual.

   (a) $\vec{v} = (1, -2, 0)$ (b) $\vec{v} = (2, 1, -3)$ (c) $\vec{v} = (2, 0, 0)$

2. Seja $C([a, b])$ o espaço vetorial de todas as funções $f : [a, b] \to \mathbb{R}$ contínuas. Mostre que a aplicação $\langle f, g\rangle = \int_a^b f(x)g(x)\, dx$ é um produto interno. Induza daí uma norma $||.||$ e calcule $||f||$, sendo $f(x) = x - 1$.

3. Considere $f(t) = 3t - 5$ e $g(t) = t^2$ no espaço dos polinômios $P(t)$ munido com o produto interno $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$.

   (a) Calcule $\langle f, g\rangle$.

   (b) Calcule $||f||$ e $||g||$.

   (c) Usando os itens anteriores, verifique o Teorema de Cauchy-Schwarz.

4. (Sel. Mestrado UFSM/2010/1) Seja V o espaço vetorial real dado pelas funções contínuas no intervalo $[0, 2\pi]$.

   (a) Mostre que $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)dx$ é um produto interno em V.

   (b) Exiba dois vetores não nulos ortogonais em relação ao produto interno dado no item (a).

   (c) Considere $[f, g] = \int_0^{2\pi} f(x)g(2\pi - x)dx$. Então $[f, g]$ define um produto interno em V?

5. Dados $a, b, c > 0$, use a desigualdade de Cauchy-Schwarz para provar que

$$\sqrt{\frac{a+b}{a+b+c}} + \sqrt{\frac{a+c}{a+b+c}} + \sqrt{\frac{b+c}{a+b+c}} \leq \sqrt{6}.$$

6. Sejam f, g funções contínuas em $[0, 1]$. Prove que
$$\left[\int_0^1 f(t)\cdot g(t)dt\right]^2 \leq \left(\int_0^1 f(t)^2dt\right)\cdot\left(\int_0^1 g(t)^2dt\right).$$

7. (Sel. Mestrado UFRGS/2007/2) Se $a_1, a_2, ..., a_N$ são números reais, $N \in \mathbb{N}$, então prove que
$$\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} a_i\right)^2 \leq \sum_{i=1}^{N} a_i^2.$$

8. Mostre que em um espaço vetorial V com um produto interno vale a lei do paralelogramo:
$$||\vec{u}+\vec{v}||^2 + ||\vec{u}-\vec{v}||^2 = 2||\vec{u}||^2 + 2||\vec{v}||^2,$$
onde $||\cdot||$ é a norma induzida do produto interno.

9. Seja V um espaço vetorial com um produto interno. Para quaisquer vetores $\vec{u}, \vec{v} \in V$, prove que os vetores $||\vec{u}||\vec{v} + ||\vec{v}||\vec{u}$ e $||\vec{u}||\vec{v} - ||\vec{v}||\vec{u}$ são ortogonais.

10. Considere a aplicação $\langle ., . \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ definida por
$$\langle \vec{u}, \vec{v} \rangle = \frac{1}{9}u_1 \cdot v_1 + \frac{1}{4}u_2 \cdot v_2$$
   (a) Mostre que $\langle ., . \rangle$ define um produto interno em $\mathbb{R}^2$.
   (b) Induza uma norma e calcule a norma do vetor $\vec{u} = (2, -1)$.
   (c) Desenhe a bola unitária $B_{\mathbb{R}^2} = \{\vec{u} \in \mathbb{R}^2 : ||\vec{u}|| < 1\}$, onde $||\cdot||$ é a norma induzida do produto interno acima.

11. Prove o Teorema de Pitágoras: Seja $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ um conjunto de vetores em um espaço vetorial V munido de um produto interno, cujos vetores $\vec{v}_1$, ..., $\vec{v}_n$ são dois a dois ortogonais. Então
$$||\vec{v}_1 + \vec{v}_2 + ... + \vec{v}_n||^2 = ||\vec{v}_1||^2 + ||\vec{v}_2||^2 + ... + ||\vec{v}_n||^2.$$
De acordo com o Teorema acima, podemos afirmar que
$$\int_0^{2\pi} (\operatorname{sen} t + \cos t)^2 dt = \int_0^{2\pi} \operatorname{sen}^2 t\, dt + \int_0^{2\pi} \cos^2 t\, dt.$$
Explique.

Índice Remissivo