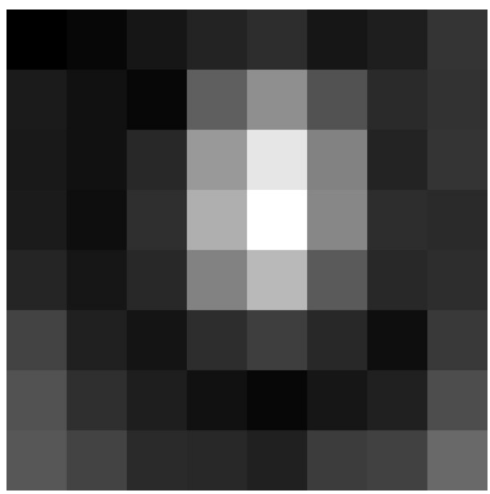
ÁLGEBRA LINEAR ESSENCIAL



Ronaldo Freire de Lima

Sumário

1

Matrizes

Matrizes são, primordialmente, tabelas numéricas sobre as quais se definem certas operações algébricas. Tais operações concedem ao conjunto das matrizes (de um mesmo tipo) a estrutura de espaço vetorial, que, por sua vez, constitui o conceito fundamental da Álgebra Linear. Além disso, as matrizes se aplicam ao estudo dos sistemas lineares (conforme veremos neste capítulo), bem como desempenham um papel decisivo no estudo das transformações lineares, as quais são justamente as funções estudadas na Álgebra Linear. Dessa forma, esses objetos permeiam toda a teoria que estudaremos, razão pela qual os introduziremos aqui, um tanto brevemente. Nosso intuito, na verdade, é o de considerar apenas os aspectos mais relevantes acerca de matrizes, os quais serão indispensáveis à nossa abordagem aos temas dos capítulos posteriores.

1.1 Matrizes e sua Álgebra

Consideraremos conhecido e estabelecido o conjunto dos números reais, $\mathbb{R}$, juntamente com sua estrutura de corpo, a qual é formada por esse conjunto e suas operações usuais de adição e multiplicação. Essas operações, como se sabe, têm as propriedades fundamentais de comutatividade, associatividade e existência de elemento neutro e inverso. Além disso, a multiplicação é distributiva com respeito a adição.

Sejam m e n dois números inteiros positivos. Uma *matriz* (*real*) do *tipo* $m \times n$ (lê-se: m por n) é uma função A que, a cada par de inteiros $(i, j) \in$

$\{1, \ldots, m\} \times \{1, \ldots, n\}$, associa um número real $a_{ij} = A(i, j)$.

Considerando-se os nossos propósitos, uma matriz será melhor representada como uma tabela disposta em m linhas e n colunas, em que a i-ésima linha é formada pelos reais $a_{i1}, \ldots, a_{in}$, e a j-ésima coluna é formada pelos reais $a_{1j}, \ldots, a_{mj}$. Assim, escreveremos:

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

Os reais a_{ij} que compõem uma matriz A são ditos as suas *entradas*. Indicaremos, também, uma matriz A do tipo $m \times n$ e com entradas a_{ij} por

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

ou, simplesmente, $A = (a_{ij})$. Quando $m = n$, a matriz A é dita *quadrada* de *ordem* n. O conjunto formado por suas entradas a_{ii} é dito, então, a *diagonal* de A. Uma matriz cujas entradas são todas iguais a zero chama-se *nula* e será denotada por $\mathbf{0}$ (independentemente de seu tipo).

A matriz $A = (a_{ij})_{3 \times 2}$, em que $a_{ij} = i + j$, por exemplo, é:

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

Dadas matrizes do mesmo tipo, $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{m \times n}$, juntamente com um número real λ, definem-se a *adição* $A + B$ e o *produto por escalar* λA por:

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

ou seja,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \quad \text{e} \quad \lambda A = \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix}.$$

Quando $C = A + B$, diz-se que C é a *soma* de A e B. Nesse contexto, os números reais são também chamados de *escalares* e, como no caso da multiplicação entre números reais, adota-se a notação $-A$ para o produto $(-1)A$.

Considerando-se matrizes A, B e C, todas do mesmo tipo, e $\lambda, \mu \in \mathbb{R}$, verificam-se facilmente as seguintes igualdades:

- $A + B = B + A$ (comutatividade da adição);
- $A + (B + C) = (A + B) + C$ (associatividade da adição);
- $A + \mathbf{0} = A$ (elemento neutro da adição);
- $A + (-A) = \mathbf{0}$ (existência do *oposto*);
- $\lambda(\mu A) = (\lambda\mu)A$ (associatividade da multiplicação por escalar);
- $1.A = A$ (elemento neutro da multiplicação por escalar).

Além disso, a adição e multiplicação por escalar são distributivas, uma com respeito à outra, isto é:

- $\lambda(A + B) = \lambda A + \lambda B$;
- $(\lambda + \mu)A = \lambda A + \mu A$.

Passemos, agora, à multiplicação de matrizes. Para tanto, consideremos $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{n\times p}$, e observemos que o número de colunas de A coincide com o número de linhas de B. Nesse caso, definimos o *produto* AB como a matriz $P = (p_{ij})_{m\times p}$, cujas entradas são:

$$p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}\,.$$

Por exemplo:

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 0 & -1 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 2.0 + 3.3 & 2.(-1) + 3.2 & 2.3 + 3.1 \\ 3.0 + 4.3 & 3.(-1) + 4.2 & 3.3 + 4.1 \\ 4.0 + 5.3 & 4.(-1) + 5.2 & 4.3 + 5.1 \end{bmatrix} = \begin{bmatrix} 9 & 4 & 9 \\ 12 & 5 & 13 \\ 15 & 6 & 17 \end{bmatrix}.$$

A multiplicação de matrizes é associativa, bem como distributiva (à direita e à esquerda) com respeito à adição. Mais especificamente, dadas matrizes $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{n\times p}$, $C = (c_{ij})_{p\times q}$, $D = (d_{ij})_{n\times p}$ e $E = (e_{ij})_{m\times n}$, são válidas as seguintes igualdades:

- $A(BC) = (AB)C$;
- $A(B + D) = AB + AD$;
- $(A + E)B = AB + EB$.

Provemos, pois, a primeira delas. Com esse propósito, escrevamos:

- $AB = (a'_{ij})_{m\times p}$ e $BC = (b'_{ij})_{n\times q}$;
- $A(BC) = (m_{ij})_{m\times q}$ e $(AB)C = (n_{ij})_{m\times q}$.

Assim, temos que:

$$\begin{aligned} m_{ij} &= \sum_{k=1}^{n} a_{ik}b'_{kj} = \sum_{k=1}^{n}\left(a_{ik}\sum_{l=1}^{p} b_{kl}c_{lj}\right) = \sum_{k=1}^{p}\left(\sum_{l=1}^{n} a_{il}b_{lk}\right)c_{kj} \\ &= \sum_{k=1}^{p} a'_{ik}c_{kj} = n_{ij} \, , \end{aligned}$$

donde se conclui que $A(BC) = (AB)C$.

Alertamos para o fato de que, quando duas matrizes A e B são quadradas e de mesma ordem, os dois produtos AB e BA estão bem definidos. No entanto, de modo geral, eles não são iguais, isto é, *o produto de matrizes quadradas de mesma ordem não é comutativo.* Quando, excepcionalmente, ocorre a igualdade $AB = BA$, dizemos que A e B *comutam.*

A associatividade do produto de matrizes, entretanto, nos permite definir, sem ambiguidade, a n-ésima potência de uma matriz A através da igualdade:

$$A^n = A.A\ldots A \ \ (n \text{ fatores}).$$

Vejamos, agora, alguns tipos especiais de matrizes quadradas.

Exemplo 1.1 (MATRIZES TRIANGULARES E DIAGONAIS)**.** Uma matriz quadrada $A = (a_{ij})_{n\times n}$ é dita *triangular*, quando ocorre uma das possibilidades:

$$a_{ij} = 0\ \forall i > j \quad \text{ou} \quad a_{ij} = 0\ \forall i < j.$$

No primeiro caso, ela é dita *triangular superior* e, no segundo, *triangular inferior.* Uma matriz que é triangular superior e inferior é dita *diagonal.* Note que duas matrizes diagonais de mesma ordem sempre comutam.

Exemplo 1.2 (MATRIZ IDENTIDADE)**.** A matriz diagonal $n \times n$ cujas entradas não nulas são todas iguais a 1 é chamada de *matriz identidade* de ordem n, a qual denota-se por I_n ou, simplesmente, por I. Tal matriz constitui o elemento neutro da multiplicação de matrizes quadradas, pois, para toda matriz quadrada A de ordem n, vale a igualdade:

$$AI = IA = A.$$

Exemplo 1.3 (MATRIZES SIMÉTRICAS)**.** Diz-se que uma matriz quadrada de ordem n, $A = (a_{ij})$, é *simétrica*, quando $a_{ij} = a_{ji}\ \forall 1 \le i, j \le n$. Evidentemente, toda matriz diagonal é simétrica. Um dos resultados mais importantes da Álgebra Linear, o célebre Teorema Espectral, estabelece uma espécie de recíproca dessa propriedade, uma vez que, segundo esse teorema, toda matriz simétrica é *diagonalizável.* Voltaremos a essa questão ao final deste capítulo, onde tornaremos isso mais claro.

Dada uma matriz $A = (a_{ij})_{m\times n}$, a *transposta* de A, por definição, é a matriz $A^* := (a^*_{ij})_{n\times m}$, em que $a^*_{ij} = a_{ji}$. Dito de forma simples, a transposta de uma matriz $m \times n$ é a matriz $n \times m$ obtida desta permutando-se suas linhas pelas colunas correspondentes.

Por exemplo, a transposta da matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ é a matriz

$$A^* = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Segue-se diretamente da definição de transposição que, para quaisquer matrizes A, $B \in \mathcal{M}(m,n)$ e $\lambda \in \mathbb{R}$, tem-se

$$(A+\lambda B)^* = A^* + \lambda B^* \quad \text{e} \quad (A^*)^* = A. \tag{1.1}$$

Observe também que uma matriz quadrada é simétrica se, e somente se, é igual à sua transposta.

1.2 Resolução de Sistemas Lineares

Dados dois números naturais m, n, mn números reais a_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$, e $b_1, \cdots, b_m \in \mathbb{R}$, chama-se o conjunto de equações

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \right. \tag{1.2}$$

de *sistema linear* nas variáveis $x_1, \cdots, x_n$. Uma *solução* do sistema (1.2) é uma n-upla de números reais, $(\lambda_1, \cdots, \lambda_n)$, que satisfaz cada uma das suas m equações, isto é,

$$\left\{ \begin{array}{ccccccccc} a_{11}\lambda_1 & + & a_{12}\lambda_2 & + & \cdots & + & a_{1n}\lambda_n & = & b_1 \\ a_{21}\lambda_1 & + & a_{22}\lambda_2 & + & \cdots & + & a_{2n}\lambda_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}\lambda_1 & + & a_{m2}\lambda_2 & + & \cdots & + & a_{mn}\lambda_n & = & b_m \end{array} \right. .$$

O *conjunto solução* de um sistema linear é o conjunto formado por todas as suas soluções. Note que, considerando-se a matriz $A = (a_{ij})_{m \times n}$, bem como

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

tem-se que o sistema (1.2) equivale à seguinte equação matricial

$$AX = B. \tag{1.3}$$

Nesse caso, dizemos que A é a *matriz do sistema* (1.2). Quando $B = \mathbf{0}$, o sistema (1.2) é dito *homogêneo.* Observe que todo sistema homogêneo admite a solução $X = \mathbf{0}$, dita *trivial.*

Neste momento, convém mencionar que, na igualdade (1.3), reside uma das razões pelas quais o produto de matrizes se define de uma maneira, digamos, "não ortodoxa". Mais precisamente, o produto de matrizes, tal como introduzido, permite reduzir um sistema linear, com inúmeras equações, a uma única equação matricial. Como consequência, o estudo dos sistemas lineares simplifica-se consideravelmente, pois finda por resumir-se ao estudo das matrizes.

Dado $n \in \mathbb{R}$, relembremos que o espaço $\mathbb{R}^n$ define-se como o produto cartesiano de n cópias de $\mathbb{R}$. Cada elemento de $\mathbb{R}^n$ é, portanto, uma n-upla de números reais $(x_1, \ldots, x_n)$, a qual é dita um *vetor* de $\mathbb{R}^n$. Para cada matriz $A = (a_{ij})_{m\times n}$, associamos à sua i-ésima linha e à sua j-ésima coluna, respectivamente, os seguintes vetores

$$\ell_i = (a_{i1}, a_{i2}, \cdots, a_{in}) \in \mathbb{R}^n \quad \text{e} \quad v_j = (a_{1j}, a_{2j}, \cdots, a_{mj}) \in \mathbb{R}^m.$$

Cada ℓ_i é dito um *vetor linha* de A, e cada v_j um *vetor coluna* de A. Assim, representamos a matriz A das seguintes formas:

$$A = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_m \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

Por exemplo, $(1, 3, -1)$, $(0, 1, 2) \in \mathbb{R}^3$ são os vetores linha da matriz

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

enquanto $(1, 0)$, $(3, 1)$, $(-1, 2) \in \mathbb{R}^2$ são seus vetores coluna.

Introduziremos, agora, a relação de equivalência mais natural entre sistemas lineares.

Definição 1.1 (SISTEMAS LINEARES EQUIVALENTES)**.** Dois sistemas lineares são ditos *equivalentes*, quando têm o mesmo conjunto solução.

No que se segue, introduziremos uma técnica de resolução de sistemas lineares que consiste em transformar um sistema dado em um sistema que lhe é equivalente e de solução imediata. Para tanto, dado um sistema linear $AX = B$, consideraremos a sua *matriz aumentada* $(A|B)$, que se obtém posicionando-se a *matriz coluna* B à direita da matriz A, isto é,

$$(A|B) = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix}.$$

Determinaremos, então, o conjunto solução de um sistema linear, efetuando uma série de operações, ditas *elementares*, sobre as linhas de sua matriz aumentada. São elas:

($\mathbf{E_1}$) Troca de posição entre duas linhas ℓ_i e ℓ_j ;

($\mathbf{E_2}$) Multiplicação de uma linha ℓ_i por um escalar $\lambda \neq 0$;

($\mathbf{E_3}$) Substituição de uma linha ℓ_j por $\ell_j + \lambda\ell_i$, sendo λ um real não nulo.

Denotaremos as operações $\mathbf{E_1}$, $\mathbf{E_2}$ e $\mathbf{E_3}$, respectivamente, por

$$\ell_i \leftrightarrow \ell_j \qquad\qquad \ell_i \to \lambda\ell_i \qquad\qquad \ell_j \to \ell_j + \lambda\ell_i\,.$$

Definição 1.2 (EQUIVALÊNCIA POR LINHAS)**.** Diz-se que uma matriz B é *linha equivalente* a uma matriz A, quando B é obtida de A efetuando-se nesta uma sequência de operações elementares.

É imediato que qualquer uma das operações elementares sobre uma matriz pode ser invertida e essa inversão também é uma operação elementar.

Especificamente, com a notação acima, as inversas de $\mathbf{E_1}$, $\mathbf{E_2}$ e $\mathbf{E_3}$ são, respectivamente,

$$\ell_j \leftrightarrow \ell_i \qquad\qquad \ell_i \to \frac{1}{\lambda}\ell_i \qquad\qquad \ell_j \to \ell_j - \lambda\ell_i\,.$$

Logo, se B é linha equivalente à matriz A, então A é linha equivalente à matriz B. Melhor dizendo, a relação de linha-equivalência é simétrica. É fácil ver também que a relação de linha-equivalência é reflexiva e transitiva (verifique!), isto é, *a relação de linha-equivalêcia é uma relação de equivalência entre matrizes.*

Proposição 1.1. *Dois sistemas lineares $AX = B$ e $A'X = B'$ são equivalentes, se suas matrizes aumentadas $(A|B)$ e $(A'|B')$ são linha equivalentes.*

Demonstração. Cada linha da matriz aumentada $(A|B)$ corresponde a uma equação do sistema $AX = B$. Daí, vê-se imediatamente que se $(A'|B')$ é obtida de $(A|B)$ efetuando-se nesta uma das duas primeiras operações elementares, então os sistemas correspondentes são equivalentes. Quanto à terceira operação elementar, basta ver que $(\lambda_1, \cdots, \lambda_n)$ é solução de

$$\left\{\begin{array}{ccccccc} a_{i1}x_1 & + & \cdots & + & a_{in}x_n & = & b_i \\ \\ a_{j1}x_1 & + & \cdots & + & a_{jn}x_n & = & b_j \end{array}\right. \tag{1.4}$$

se, e somente se, é solução de

$$\left\{\begin{array}{ccccccc} a_{i1}x_1 & + & \cdots & + & a_{in}x_n & = & b_i \\ \\ (a_{j1}+\lambda a_{i1})x_1 & + & \cdots & + & (a_{jn}+\lambda a_{in})x_n & = & b_j + \lambda b_i\,, \end{array}\right. \tag{1.5}$$

o que se verifica por substituição direta. ■

Exemplo 1.4. Consideremos o sistema

$$\left\{\begin{array}{ccccc} x & + & 2y & = & 1 \\ 3x & - & y & = & 2 \end{array}\right.,$$

cuja matriz aumentada é

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

Efetuando-se nessa matriz a operação $\ell_2 \to \ell_2 - 3\ell_1$, obtém-se

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & -7 & -1 \end{bmatrix}.$$

Segue-se, então, da Proposição 1.1, que o sitema dado é equivalente a

$$\left\{ \begin{array}{rcrcr} x & + & 2y & = & 1 \\ & - & 7y & = & -1 \end{array} \right. ,$$

cuja única solução, claramente, é $(5/7, 1/7)$.

Note que, nesse exemplo, a operação realizada sobre a matriz aumentada do sistema resultou na eliminação da variável x da segunda equação, transformando-a numa equação linear de uma única variável, donde se obteve facilmente o conjunto solução. Esse processo de resolução de sistemas lineares nos leva naturalmente ao conceito de *matriz escalonada*, que introduzimos a seguir.

Definição 1.3 (MATRIZ ESCALONADA)**.** Diz-se que uma matriz $A = (a_{ij})_{m\times n}$ é *escalonada*, quando cumpre as seguintes condições:

i) O primeiro elemento não nulo de uma linha está à esquerda do primeiro elemento não-nulo da linha subsequente;

ii) As linhas nulas, caso existam, estão abaixo das demais.

As matrizes,

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 2 & 3 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 2 & 3 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 5 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

por exemplo, são escalonadas, enquanto

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

não é uma matriz escalonada.

Não é difícil ver que, efetuando-se adequadamente operações elementares sobre as linhas de uma matriz, podemos transformá-la numa matriz escalonada (a este processo chamamos *escalonamento* da matriz), isto é, *toda matriz é linha equivalente a alguma matriz escalonada.*

Consideremos, por exemplo, a matriz

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

Para escaloná-la, operamos sobre suas linhas da seguinte maneira:

$$\begin{bmatrix} 1 & -2 & 3 & 1 \\ 2 & 1 & -1 & 2 \\ 4 & -3 & 5 & 4 \end{bmatrix} \overset{(1)\text{ e }(2)}{\longrightarrow} \begin{bmatrix} 1 & -2 & 3 & 1 \\ 0 & 5 & -7 & 0 \\ 0 & 5 & -7 & 0 \end{bmatrix} \overset{(3)\text{ e }(4)}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1/5 & 1 \\ 0 & 5 & -7 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

em que $(1), (2), (3), (4)$ denotam as seguintes operações

$$(1)\ \ell_2 \to \ell_2 - 2\ell_1 \quad (2)\ \ell_3 \to \ell_3 - 4\ell_1 \quad (3)\ \ell_1 \to \ell_1 + \frac{2}{5}\ell_1 \quad (4)\ \ell_3 \to \ell_3 - \ell_2\,.$$

Estamos, agora, em condições de aplicar o *Método de Gauss* para resolução de sistemas lineares. Ele consiste em escalonar a matriz aumentada $(A|B)$ de um sistema linear $AX = B$, transformando-a numa matriz $(A'|B')$. Como resultado, obtém-se o sistema $A'X = B'$ que, pela Proposição 1.1, é equivalente ao sistema $AX = B$, e que é de fácil resolução, pois $(A'|B')$ é escalonada.

Ilustraremos esse método considerando, inicialmente, o sistema linear

$$\left\{ \begin{array}{rcrcrcl} x & - & 2y & + & 3z & = & 1 \\ 2x & + & y & - & z & = & 2 \\ 4x & - & 3y & + & 5z & = & 4 \end{array} \right. .$$

Observando-se que sua matriz aumentada é a matriz A do exemplo acima, temos que o sistema dado é equivalente a

$$\left\{\begin{array}{rcrcrcl} x & & & + & \frac{1}{5}z & = & 1 \\ & & 5y & - & 7z & = & 0 \\ & & & & 0 & = & 0 \end{array}\right. ,$$

cujo conjunto solução é

$$S = \left\{\left(\frac{5-z}{5}, \frac{7z}{5}, z\right);\ z \in \mathbb{R}\right\} \subset \mathbb{R}^3.$$

Consideremos agora o sistema

$$(*)\quad \left\{\begin{array}{rcrcrcrcl} 2x & - & y & & & + & 4t & = & 9 \\ x & + & y & - & z & + & 2t & = & 7 \\ -x & + & 2y & + & z & - & t & = & 3 \\ & & 4y & - & z & + & 3t & = & 13 \end{array}\right. .$$

Procedendo-se como indicado acima (verifique, em cada passagem, quais foram as operações elementares realizadas), temos

$$\begin{bmatrix} 2 & -1 & 0 & 4 & 9 \\ 1 & 1 & -1 & 2 & 7 \\ -1 & 2 & 1 & -1 & 3 \\ 0 & 4 & -1 & 3 & 13 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & -1 & 2 & 7 \\ 2 & -1 & 0 & 4 & 9 \\ -1 & 2 & 1 & -1 & 3 \\ 0 & 4 & -1 & 3 & 13 \end{bmatrix} \longrightarrow$$

$$\begin{bmatrix} 1 & 1 & -1 & 2 & 7 \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 3 & 0 & 1 & 10 \\ 0 & 4 & -1 & 3 & 13 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{3} & 2 & \frac{16}{3} \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & \frac{5}{3} & 3 & \frac{19}{3} \end{bmatrix} \longrightarrow$$

$$\begin{bmatrix} 1 & 0 & -\frac{1}{3} & 2 & \frac{16}{3} \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & 0 & \frac{13}{3} & \frac{13}{3} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{3} & 2 & \frac{16}{3} \\ 0 & -3 & 2 & 0 & -5 \\ 0 & 0 & 2 & 1 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Dessa forma, o sistema $(*)$ é equivalente a

$$\begin{cases} x & & - & \frac{1}{3}z & + & 2t & = & \frac{16}{3} \\ & - & 3y & + & 2z & & & = & -5 \\ & & & & 2z & + & t & = & 5 \\ & & & & & & t & = & 1 \end{cases}$$

cujo conjunto solução é $S = \{(4,3,2,1)\}$.

Consideremos, finalmente, o sistema

$$\begin{cases} x + 3y = 5 \\ 2x + 6y = 8 \end{cases}.$$

Escalonando-se sua matriz aumentada, concluí-se facilmente que este é equivalente a

$$\begin{cases} x + 3y = 5 \\ \qquad\ 0 = -2 \end{cases},$$

cujo conjunto solução, evidentemente, é vazio.

No Capítulo 3, voltaremos à discussão sobre sistemas lineares. Lá, estabeleceremos condições sobre as matrizes A e B para que um sistema linear $AX = B$ admita uma ou infinitas soluções.

1.3 Inversão e Semelhança de Matrizes

Nesta seção, introduziremos dois conceitos fundamentais da teoria das matrizes, os quais, dentre outras características, determinam uma relação de equivalência na classe das matrizes quadradas de mesma ordem. Vejamos, então, o primeiro deles.

Definição 1.4 (MATRIZES INVERTÍVEIS)**.** Diz-se que uma matriz quadrada A é *invertível*, quando existe uma matriz quadrada B, de mesma ordem que A, tal que $AB = BA = I$. A matriz B é dita a *inversa* de A, e escreve-se $B = A^{-1}$.

Observemos que, se A é uma matriz invertível, então todo sistema linear $AX = B$ admite uma única solução, a saber, $X = A^{-1}B$. Assim, matrizes invertíveis definem sistemas lineares determinados, isto é, que têm solução única.

Definição 1.5 (SEMELHANÇA DE MATRIZES)**.** Dizemos que duas matrizes quadradas $n \times n$, A e B, são *semelhantes*, e escrevemos $A \sim B$, quando existe uma matriz $n \times n$ invertível, M, a qual cumpre a igualdade $AM = MB$.

É imediato que a semelhança entre matrizes é uma relação reflexiva e simétrica. Ademais, se A é semelhante a B, e B é semelhante a C, existem matrizes invertíveis, M, N, tais que $AM = MB$ e $BN = NC$. Logo, $A(MN) = (MN)C$. Porém, sendo M e N invertíveis, temos que MN é invertível (vide Exercício 5), donde se conclui que A é semelhante a C e, portanto, que a semelhança de matrizes é também transitiva. Assim, denotando-se por $\mathcal{M}(n)$ o conjunto das matrizes reais $n \times n$, vale o resultado seguinte.

Proposição 1.2. *A semelhança $\sim$ é uma relação de equivalência em $\mathcal{M}(n)$.*

Duas funções importantes, definidas em $\mathcal{M}(n)$, são o *traço* e o *determinante*, as quais são invariantes em classes de matrizes semelhantes, isto é, quando duas matrizes são semelhantes, seus traços e determinantes coincidem. Estudaremos os determinantes num capítulo posterior (vide, entretanto, Exercício 17). O traço, por outro lado, é um conceito mais simples e pode ser considerado agora.

Definição 1.6 (TRAÇO DE UMA MATRIZ)**.** Chama-se *traço* de uma matriz quadrada, $A = (a_{ij})_{n \times n}$, a soma das entradas de sua diagonal, isto é,

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

Vejamos, então, que matrizes semelhantes têm o mesmo traço. Para tanto, tomemos matrizes $n \times n$ arbitrárias, A e B, e observemos que

$$\text{traço}\,(AB) = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}b_{ki}\right) = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} b_{ki}a_{ik}\right) = \text{traço}\,(BA).$$

Suponhamos, agora, que A e B sejam semelhantes. Assim, existe uma matriz invertível, M, tal que $A = MBM^{-1}$. Logo, $\text{traço}\,(A) = \text{traço}\,(MBM^{-1}) = \text{traço}\,(MM^{-1}B) = \text{traço}\,(B)$. Vale, dessa forma, a implicação:

$$A \sim B \;\Rightarrow\; \text{traço}\,(A) = \text{traço}\,(B).$$

Diz-se que uma matriz quadrada A é *diagonalizável*, quando é semelhante a uma matriz diagonal. Conforme mencionamos anteriormente, o Teorema Espectral afirma que toda matriz simétrica é diagonalizável. A importância desse teorema, o qual estabeleceremos posteriormente, deve-se ao fato de que as matrizes diagonais, sob diversos aspectos, são as mais simples. Isso pode ser constatado, por exemplo, quando multiplicamos matrizes diagonais, ou quando calculamos seus determinantes.

1.4 Exercícios

Seção 1.1

1. Determine (caso exista) a matriz A, tal que $AB = C$, em que

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

2. Mostre que se A é uma matriz triangular superior, então A^2 é também uma matriz triangular superior.

3. Sejam A e B matrizes $m \times n$ e $n \times p$, respectivamente. A implicação

$$AB = 0 \;\Rightarrow\; A = 0 \;\text{ ou }\; B = 0$$

é verdadeira ou falsa? No caso negativo, exiba um contraexemplo.

4. Encontre uma matriz $A = (a_{ij})_{2\times 2}$, não nula, tal que $A^2 = \mathbf{0}$.

5. Sejam A e B matrizes $m \times n$ e $n \times p$, respectivamente. Prove que

$$(AB)^* = B^* A^*.$$

6. Uma matriz quadrada A é dita *antissimétrica*, quando $A^* = -A$. Prove que toda matriz quadrada se expressa como a soma de uma matriz simétrica com uma matriz antissimétrica.

7. Dado $v = (x_1, \ldots, x_n)$, defina $[v] = (x_{i1})_{n\times 1}$ por $x_{i1} = x_i$, isto é,

$$[v] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Sejam A uma matriz $m \times n$, e B uma matriz $n \times p$, tais que

$$B = \begin{bmatrix} v_1 & v_2 & \cdots & v_p \end{bmatrix} \quad \text{e} \quad AB = \begin{bmatrix} w_1 & w_2 & \cdots & w_p \end{bmatrix}.$$

Mostre que vale a igualdade:

$$[w_j] = A[v_j] \ \ \forall j = 1, \ldots p.$$

Dito de forma simples, *as colunas de AB são as respectivas colunas de B multiplicadas, à esquerda, pela matriz A.*

Seção 1.2

8. Determine o conjunto solução dos seguintes sistemas lineares:

a) $\left\{ \begin{array}{rcrcrcr} x & - & 2y & - & 3z & = & 0 \\ 3x & + & y & - & z & = & -1 \end{array} \right.$

b) $\left\{ \begin{array}{rcrcrcr} x & + & y & + & z & = & 2 \\ 2x & - & y & + & 3z & = & 9 \\ x & + & 2y & - & z & = & -3 \end{array} \right.$

9. Mostre que é vazio o conjunto solução do sistema linear:

$$\begin{cases} x_1 & - & 2x_2 & + & x_3 & + & 2x_4 & = & 1 \\ x_1 & + & x_2 & - & x_3 & + & x_4 & = & 2 \\ x_1 & + & 7x_2 & - & 5x_3 & - & x_4 & = & 3 \end{cases}.$$

10. Seja A a matriz 4×4:

$$\begin{bmatrix} 3 & -6 & 2 & -1 \\ -2 & 4 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix}.$$

Para quais matrizes coluna B, o sistema $AX = B$ tem solução?

Seção 1.3

11. Prove que, se A e B são matrizes $n \times n$ e $\lambda \in \mathbb{R}$, então são verdadeiras as seguintes igualdades:

    a) $\text{traço}\,(A + B) = \text{traço}\,(A) + \text{traço}\,(B)$;

    b) $\text{traço}\,(\lambda A) = \lambda\,\text{traço}\,(A)$.

12. Seja C uma matriz 2×2. Mostre que uma condição necessária e suficiente para que existam matrizes A e B, 2×2, as quais cumprem

$$AB - BA = C,$$

é a de que C tenha traço nulo.

13. Prove que uma matriz quadrada A é invertível se, e somente se, sua transposta A^* é invertível, e que, no caso afirmativo, $(A^*)^{-1} = (A^{-1})^*$.

14. Mostre que, se A e B são matrizes $n \times n$, ambas invertíveis, então AB é invertível e vale a igualdade $(AB)^{-1} = B^{-1}A^{-1}$.

15. Seja A uma matriz quadrada, a qual cumpre $A^2 - A + I = 0$. Mostre que A é invertível.

16. Prove que uma matriz quadrada A não é invertível, em qualquer das seguintes ocorrências:

    a) As entradas de uma de suas linhas são todas nulas;

    b) As entradas de uma de suas colunas são todas nulas;

    c) Dois de seus vetores coluna são múltiplos, um do outro;

    d) Dois de seus vetores linha são múltiplos, um do outro.

    Conclua que o produto de uma matriz $n \times 1$ por uma matriz $1 \times n$ é uma matriz $n \times n$, a qual não é invertível.

17. Suponha que $f : \mathcal{M}(n) \to \mathbb{R}$ seja uma função, tal que $f(I) = 1$ e

$$f(AB) = f(A)f(B) \;\; \forall A, B \in \mathcal{M}(n).$$

    Prove que:

    a) $A \sim B \Rightarrow f(A) = f(B)$;

    b) A invertível $\;\Rightarrow\;$ $f(A) \neq 0$ e $f(A^{-1}) = 1/f(A)$.

2

Espaços Vetoriais

Neste capítulo, introduzimos os espaços vetoriais, os quais constituem os principais objetos de nosso estudo. Um tal espaço, simplesmente, é uma estrutura formada por um conjunto e duas operações que possuem certas propriedades, ditas fundamentais. Os espaços euclidianos, os quais consideraremos frequentemente, admitem uma estrutura natural de espaço vetorial que os torna, nesse contexto, os exemplos padrão.

Estabeleceremos os conceitos básicos de subespaço vetorial, dependêcia linear, base e dimensão, e os aplicaremos para obter um critério de inversão de matrizes. Concluiremos, então, introduzindo o conceito de soma de subespaços vetoriais.

2.1 Espaços e Subespaços Vetoriais

Um *espaço vetorial (real)* é um conjunto V munido de duas operações:

$$\begin{array}{ccc} V\times V & \to & V \\ (v,w) & \mapsto & v+w \end{array} \quad \text{e} \quad \begin{array}{ccc} \mathbb{R}\times V & \to & V \\ (\lambda,v) & \mapsto & \lambda v \end{array},$$

chamadas, respectivamente, de *adição* e *multiplicação por escalar*, as quais têm as seguintes propriedades:

P1: $u+v=v+u \ \ \forall u,v\in V.$

P2: $(u+v)+w=u+(v+w) \ \ \forall u,v,w\in V.$

P3: Existe $\mathbf{0} \in V$, tal que $v + \mathbf{0} = v \ \forall v \in V$.

P4: Para todo $v \in V$, existe $-v \in V$, tal que $v + (-v) = \mathbf{0}$.

P5: $\lambda(\mu v) = (\lambda\mu)v$.

P6: $1.v = v \ \forall v \in V$.

P7: $\lambda(u + v) = \lambda u + \lambda v \ \forall \lambda \in \mathbb{R},\ u, v \in V$.

P8: $(\lambda + \mu)v = \lambda v + \mu v \ \forall \lambda, \mu \in \mathbb{R},\ v \in V$.

Nesse contexto, os elementos de V são chamados de *vetores*, enquanto os números reais são chamados de *escalares*. Um vetor $u = v + w$ é dito a *soma* de v e w. Assim, as propriedades (**P1**) e (**P2**) expressam, respectivamente, a *comutatividade* e *associatividade* da adição de vetores. A propriedade (**P3**), por sua vez, nos diz que essa adição possui um *elemento neutro* $\mathbf{0} \in V$, o qual é chamado de *vetor nulo* ou, simplesmente, de *zero* de V (vide Exercício 1). A propriedade (**P4**) estabelece que todo vetor de V possui um *oposto* com respeito à adição. As propriedades (**P5**) e (**P6**) correspondem à associatividade e existência de elemento neutro da multiplicação por escalar. Finalmente, as propriedades (**P7**) e (**P8**) expressam a distributividade da adição e da multiplicação por escalar, uma com respeito à outra.

Decorre diretamente das propriedades fundamentais da adição e multiplicação de números reais, que os conjuntos descritos nos quatro exemplos a seguir, juntamente com as correspondentes operações, são espaços vetoriais.

Exemplo 2.1 (*Os espaços euclidianos*)**.** $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n);\ x_i \in \mathbb{R}\}$. Dados os vetores $v = (x_1, \ldots, x_n)$, $w = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ e o escalar $\lambda \in \mathbb{R}$, definem-se $v + w$ e λv, respectivamente, por

$$v + w = (x_1 + y_1, \ldots, x_n + y_n) \quad \text{e} \quad \lambda v = (\lambda x_1, \ldots, \lambda x_n).$$

Nesse espaço, o vetor nulo é $\mathbf{0} = (0, \ldots, 0)$, e o oposto do vetor $v = (x_1, \ldots, x_n)$ é vetor $-v = (-x_1, \ldots, -x_n)$.

Exemplo 2.2 (*Os espaços de matrizes*)**.** $\mathcal{M}(m,n) = \{\text{matrizes reais } m \times n\}$ com as operações usuais da adição de matrizes e multiplicação de número real por matriz. O vetor nulo desse espaço é a matriz nula de $\mathcal{M}(m,n)$, isto é, aquela cujas entradas são todas nulas. O oposto da matriz $A = (a_{ij})$ é a matriz $-A = (-a_{ij})$.

Exemplo 2.3 (*O espaço dos polinômios*)**.** O conjunto $\mathcal{P}[t,\mathbb{R}]$ formado por todos os polinômios de variável t e coeficientes em $\mathbb{R}$. As operações de adição de polinômios e multiplicação de um número real por um polinômio fazem de $\mathcal{P}[t,\mathbb{R}]$ um espaço vetorial cujo vetor nulo é o polinômio nulo. Um vetor $v \in \mathcal{P}[t,\mathbb{R}]$ se escreve como $v = a_0 + a_1t + \cdots + a_{n-1}t^{n-1} + a_nt^n$, $a_i \in \mathbb{R}$, e o seu oposto é o vetor $-v = -a_0 - a_1t - \cdots - a_{n-1}t^{n-1} - a_nt^n$.

Exemplo 2.4 (*Os espaços de funções*)**.** O conjunto $\mathcal{F}(A,\mathbb{R})$ formado por todas as funções $f : A \to \mathbb{R}$, em que A é um conjunto arbitrário. Dados $f, g \in \mathcal{F}(A,\mathbb{R})$, definem-se $f+g$ e λf, $\lambda \in \mathbb{R}$, por

$$\begin{array}{rccc} f+g: & A & \to & \mathbb{R} \\ & x & \mapsto & f(x)+g(x) \end{array} \quad \text{e} \quad \begin{array}{rccc} \lambda f: & A & \to & \mathbb{R} \\ & x & \mapsto & \lambda f(x) \end{array}.$$

A função identicamente nula de A em $\mathbb{R}$ é o vetor nulo de $\mathcal{F}(A,\mathbb{R})$ e, dado $f \in \mathcal{F}(A,\mathbb{R})$, seu oposto é a função $-f$ que, a cada $x \in A$, associa o real $-f(x)$.

Proposição 2.1. *Dado um espaço vetorial V, valem, para quaisquer $\lambda \in \mathbb{R}$ e $v \in V$, as seguintes igualdades:*

i) $0v = \mathbf{0}$.

ii) $\lambda\mathbf{0} = \mathbf{0}$.

iii) $-v = (-1)v$.

Demonstração. De fato, $0v = (0+0)v = 0v + 0v$, isto é, $0v = 0v + 0v$. Adicionando-se $-0v$ a ambos os membros dessa última igualdade, obtém-se $0v = \mathbf{0}$, o que prova (i). De modo inteiramente análogo, prova-se (ii).

Da igualdade (i), temos que $\mathbf{0} = 0v = (-1+1)v$, donde $(-1)v + v = \mathbf{0}$. Adicionando-se, então, $-v$ a ambos os membros, obtém-se $-v = (-1)v$, o que prova a igualdade (iii) e conclui a demonstração. ■

Note que a propriedade (iii) na proposição acima implica na unicidade do vetor oposto e nos sugere definir a *diferença*

$$u - v = u + (-v).$$

Definição 2.1 (SUBESPAÇO VETORIAL)**.** Um subconjunto $W \subset V$ de um espaço vetorial V é chamado de *subespaço vetorial* (ou simplesmente *subespaço*) de V, se cumpre as condições seguintes:

i) $\mathbf{0} \in W$;

ii) $u + v \in W$ sempre que $u, v \in W$;

iii) $\lambda u \in W \ \ \forall \lambda \in \mathbb{R},\ u \in W$.

É imediato que V e $\{\mathbf{0}\}$ são subespaços de V. Eles são ditos *subespaços triviais*. O subespaço $\{\mathbf{0}\}$ é também dito *nulo*. Note que *todo subespaço vetorial é um espaço vetorial.*

Exemplo 2.5. Verifiquemos que $W = \{(x, y) \in \mathbb{R}^2;\ x = y\}$ é um subespaço vetorial de $\mathbb{R}^2$. Com efeito, claramente, $\mathbf{0} = (0, 0) \in W$. Ademais, dados $u = (x_1, y_1)$, $v = (x_2, y_2) \in W$, temos que $x_1 = y_1$ e $x_2 = y_2$. Logo, $x_1 + x_2 = y_1 + y_2$ e, para todo $\lambda \in \mathbb{R}$, $\lambda x_1 = \lambda y_1$, donde $u + v = (x_1 + x_2, y_1 + y_2) \in W$ e $\lambda u = (\lambda x_1, \lambda y_2) \in W$. Segue-se que W é um subespaço vetorial de $\mathbb{R}^2$.

No exemplo acima, podemos observar que, geometricamente, W é uma reta r do plano $\mathbb{R}^2$ (aquela de equação cartesiana $y = x$) que passa pela origem do mesmo. Tomando-se uma equação paramétrica de r, vemos que $W = \{v \in \mathbb{R}^2;\ v = t(1, 1),\ t \in \mathbb{R}\}$, isto é, W é o conjunto de todos os múltiplos[(i)] de $(1, 1)$. Dessa forma, podemos dizer que *o vetor* $(1, 1)$ *determina o subespaço vetorial* W.

[(i)] Um vetor w é um *múltiplo* de outro vetor v, se existe $\lambda \in \mathbb{R}$, tal que $w = \lambda v$.

Exemplo 2.6. Seja W o seguinte subconjunto do espaço $\mathbb{R}^3$:

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

É imediato que o vetor nulo $0 = (0, 0, 0)$ pertence a W. Além disso, dados $u = (x_1, y_1, z_1)$ e $v = (x_2, y_2, z_2)$ em W, temos que $2x_1 + y_1 - z_1 = 0$ e $2x_2 + y_2 - z_2 = 0$. Adicionando-se estas igualdades membro a membro, obtém-se $2(x_1 + x_2) + (y_1 + y_2) - (z_1 + z_2) = 0$. Além disso, multiplicando-se a primeira delas por $\lambda \in \mathbb{R}$, tem-se $2(\lambda x_1) + (\lambda y_1) - (\lambda z_1) = 0$. Dessa forma, $u + v = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ e $\lambda u = (\lambda x_1, \lambda x_2, \lambda x_3)$ são vetores de W. Logo, W é um subespaço vetorial de $\mathbb{R}^3$.

É sabido que o conjunto W do exemplo anterior é um plano de $\mathbb{R}^3$ que contém a origem $(0, 0, 0)$. Tomando-se, por exemplo, os vetores $v_1 = (0, 1, 1)$ e $v_2 = (1, 0, 2)$, de W, obtém-se uma representação paramétrica para o mesmo, qual seja,

$$W = \{v \in \mathbb{R}^3;\ v = \lambda_1 v_1 + \lambda_2 v_2\,,\ \lambda_1, \lambda_2 \in \mathbb{R}\}.$$

Assim, podemos afirmar que *os vetores* $(0, 1, 1)$ *e* $(1, 0, 2)$ *determinam o subespaço* W.

Os exemplos acima motivam introduzir, no contexto dos espaços vetoriais, o conceito de *subespaço gerado* por um certo subconjunto de vetores. Com essa terminologia, temos, no Exemplo 2.5, que W é o subespaço gerado pelo vetor $(1, 1)$, enquanto no Exemplo 2.6, W é o subespaço gerado pelos vetores $(0, 1, 1)$ e $(1, 0, 2)$. No que se segue, tornaremos essa ideia mais precisa.

Definição 2.2 (COMBINAÇÃO LINEAR)**.** Diz-se que um vetor v, de um espaço vetorial V, é uma *combinação linear* dos vetores $v_1\,, \dots, v_n \in V$, quando existem escalares $\lambda_1\,, \dots, \lambda_n \in \mathbb{R}$, tais que

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n\,.$$

O vetor $(1, 2) \in \mathbb{R}^2$, por exemplo, é uma combinação linear dos vetores $v_1 = (1, 0)$ e $v_2 = (0, 1)$ de $\mathbb{R}^2$, pois $(1, 2) = 1(1, 0) + 2(0, 1)$. Note que todo vetor $v = (x, y) \in \mathbb{R}^2$ tem essa propriedade, já que

$$(x, y) = x(1, 0) + y(0, 1)\ \ \forall (x, y) \in \mathbb{R}^2.$$

Exemplo 2.7. Verifiquemos que, em $\mathbb{R}^2$, o vetor $(2,-3)$ é combinação linear dos vetores $(1,-1)$ e $(2,1)$. Para isso, devemos encontrar escalares λ_1, λ_2, tais que $(2,-3) = \lambda_1(1,-1) + \lambda_2(2,1) = (\lambda_1 + 2\lambda_2, -\lambda_1 + \lambda_2)$. Sendo assim, devemos ter $\lambda_1 + 2\lambda_2 = 2$ e $-\lambda_1 + \lambda_2 = -3$, isto é, $\lambda_1 = 8/3$ e $\lambda_2 = -1/3$. Logo, vale a seguinte igualdade:

$$(2,-3) = \frac{8}{3}(1,-1) - \frac{1}{3}(2,1).$$

Exemplo 2.8. Consideremos em $\mathcal{M}(2,2)$ os vetores

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Dado um vetor $v = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}(2,2)$, tem-se

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

donde se conclui que todo vetor $v \in \mathcal{M}(2,2)$ é combinação linear dos vetores v_1, v_2, v_3, v_4.

Exemplo 2.9. Segue-se diretamente da definição de polinômio que todo vetor $v \in \mathcal{P}[t,\mathbb{R}]$, de grau n, é uma combinação linear dos vetores $1, t, \ldots, t^n$.

Observação 2.1. O conceito de combinação linear pode ser usado para reinterpretar o produto de uma matriz $A = (a_{ij})_{m\times n}$ por uma *matriz coluna* $X = [v]$, em que $v = (x_1, \ldots, x_n)$ é um vetor de $\mathbb{R}^n$ (vide Exercício 7 – Cap. 1). Com efeito, indicando-se por $v_1, \ldots, v_n \in \mathbb{R}^n$ os vetores coluna de A, tem-se, pela definição de produto de matrizes, que

$$AX = A[v] = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = x_1[v_1] + \cdots + x_n[v_n].$$

Logo, *o produto de uma matriz* $m \times n$, A, *por uma matriz coluna* $n \times 1$, X, *é uma combinação linear dos vetores coluna de* A, *cujos coeficientes são as correspondentes entradas de* X.

Proposição 2.2 (COMBINAÇÕES LINEARES E SUBESPAÇOS). *Sejam V um espaço vetorial e $v_1, \ldots, v_n$ vetores de V. Considere o conjunto W formado por todas as combinações lineares desses vetores, isto é,*

$$W = \{v \in V;\, v = \lambda_1 v_1 + \cdots + \lambda_n v_n\,,\ \lambda_i \in \mathbb{R}\}.$$

Então, W é um subespaço vetorial de V.

Demonstração. Temos que $\mathbf{0} = 0v_1 + \cdots + 0v_n$, isto é, $\mathbf{0} \in W$. Tomemos $u = \mu_1 v_1 + \cdots + \mu_n v_n$ e $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ em W, e observemos que $u + v = (\mu_1 + \lambda_1)v_1 + \cdots + (\mu_n + \lambda_n)v_n \in W$. Ademais, para todo $\lambda \in \mathbb{R}$, $\lambda v = (\lambda\lambda_1)v_1 + \cdots + (\lambda\lambda_n)v_n \in W$. Logo, W é um subespaço de V. ■

Definição 2.3 (SUBESPAÇO GERADO). O subespaço W da Proposição 2.2 é chamado de *subespaço gerado* pelos vetores $v_1, \ldots, v_n$, o qual denotaremos por $\mathrm{Ger}\{v_1, \ldots, v_n\}$. Assim,

$$\mathrm{Ger}\{v_1, \ldots, v_n\} = \{v \in V;\ v = \lambda_1 v_1 + \cdots + \lambda_n v_n\,,\ \lambda_i \in \mathbb{R}\}\,.$$

Pelo Exemplo 2.6, bem como pelas observações feitas após o mesmo, temos que $\mathrm{Ger}\{(0,1,1),(1,0,2)\} = \{(x,y,z) \in \mathbb{R}^3;\, 2x + y - z = 0\} \subset \mathbb{R}^3$. Segue-se também de considerações anteriores que $\mathrm{Ger}\{(1,0),(0,1)\} = \mathbb{R}^2$. Note ainda que os vetores v_1, v_2, v_3, v_4 do Exemplo 2.8 geram $\mathcal{M}(2,2)$, isto é, $\mathrm{Ger}\{v_1, v_2, v_3, v_4\} = \mathcal{M}(2,2)$.

Essencialmente, a única operação entre conjuntos que é fechada na classe dos subespaços vetoriais é a interseção, conforme a proposição seguinte.

Proposição 2.3 (INTERSEÇÃO DE SUBESPAÇOS). *Sejam U e W subespaços de um espaço vetorial V. Então, $U \cap W$ é não vazio e constitui um subespaço de V.*

Demonstração. Temos, pela definição de subespaço vetorial, que $\mathbf{0} \in U \cap W$. Logo, $U \cap W \neq \emptyset$. Sejam $u, w \in U \cap W$ e $\lambda \in \mathbb{R}$. Em particular, $u, w \in U$, o que nos dá $u + w \in U$ e $\lambda u \in U$, já que U é um subespaço de V. Analogamente, $u + w \in W$ e $\lambda u \in W$. Dessa forma, $u + w \in U \cap W$ e $\lambda u \in U \cap W$, donde se conclui que $U \cap W$ é um subespaço vetorial de V. ■

Proposição 2.4. *Seja* $\{v_1, \dots, v_n\}$ *um subconjunto finito de um espaço vetorial* V*. Se* $w \in \mathrm{Ger}\{v_1, \dots, v_n\}$*, então* $\mathrm{Ger}\{v_1, \dots, v_n, w\} = \mathrm{Ger}\{v_1, , \dots, v_n\}$.

Demonstração. Dado $v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in \mathrm{Ger}\{v_1, \dots, v_n\}$, podemos escrever $v = \lambda_1 v_1 + \cdots + \lambda_n v_n + 0w$, donde $v \in \mathrm{Ger}\{v_1, \dots, v_n, w\}$. Logo, $\mathrm{Ger}\{v_1, , \dots, v_n, w\} \supset \mathrm{Ger}\{v_1, , \dots, v_n\}$.

Agora, uma vez que $w \in \mathrm{Ger}\{v_1, \dots, v_n\}$, existem escalares $\lambda_1, \lambda_2, \dots, \lambda_n$, tais que $w = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Dado, então, $v \in \mathrm{Ger}\{v_1, \dots, v_n, w\}$, sejam $\mu_1, \dots, \mu_{n+1}$ escalares tais que $v = \mu_1 v_1 + \cdots + \mu_n v_n + \mu_{n+1} w$. Assim, temos

$$\begin{aligned} v &= \mu_1 v_1 + \cdots + \mu_n v_n + \mu_{n+1}(\lambda_1 v_1 + \cdots + \lambda_n v_n) \\ &= (\mu_1 + \mu_{n+1}\lambda_1) v_1 + \cdots + (\mu_n + \mu_{n+1}\lambda_n) v_n , \end{aligned}$$

donde se infere que $v \in \mathrm{Ger}\{v_1, \dots, v_n\}$. Logo, vale a inclusão

$$\mathrm{Ger}\{v_1, \dots, v_n, w\} \subset \mathrm{Ger}\{v_1, \dots, v_n\},$$

e, portanto, $\mathrm{Ger}\{v_1, , \dots, v_n, w\} = \mathrm{Ger}\{v_1, , \dots, v_n\}$, como queríamos demonstrar. ∎

2.2 Dependência Linear

No que diz respeito a subespaços gerados, a Proposição 2.4 da seção anterior nos leva a concluir que, num conjunto $\mathcal{A}$ de geradores de um tal subespaço, pode haver vetores "obsoletos", isto é, o conjunto obtido de $\mathcal{A}$ por exclusão desses vetores gera o mesmo subespaço que $\mathcal{A}$. Nosso objetivo, agora, será o de determinar condições que, ao serem verificadas, nos permitirão decidir se há ou não vetores obsoletos num conjunto de geradores de um subespaço vetorial. Isso nos leva à noção de *dependência linear* de vetores.

Consideremos, inicialmente, o espaço $\mathbb{R}^2$.

Lema 2.1. *Sejam* $v = (a, b)$ *e* $w = (c, d)$ *vetores não nulos de* $\mathbb{R}^2$*. Então,* v *é múltiplo de* w *se, e somente se,* $ad - bc = 0$.

Demonstração. Com efeito, se existe $\lambda \in \mathbb{R}$ satisfazendo $v = \lambda w$, então $\lambda \neq 0$, pois $v, w \neq 0$. Além disso, devemos ter $a = \lambda c$ e $b = \lambda d$. Logo, $ad - bc = (\lambda c)d - (\lambda d)c = 0$. Reciprocamente, suponhamos que $ad - bc = 0$ se cumpra. Como v e w são não nulos, temos que $a \neq 0$ ou $b \neq 0$ e $c \neq 0$ ou $d \neq 0$. Suponhamos, então, sem perda de generalidade, que tenhamos $d, b \neq 0$ e façamos $\lambda = b/d$. Nesse caso, segue-se da hipótese que $a = \lambda c$, donde $v = (a, b) = (\lambda c, \lambda d) = \lambda(c, d) = \lambda w$, isto é, v é múltiplo de w. ■

Proposição 2.5. *Sejam $v = (a, b)$ e $w = (c, d)$ vetores não nulos de $\mathbb{R}^2$. Então,* $\mathrm{Ger}\{v, w\} = \mathbb{R}^2$ *se, e somente se, v não é múltiplo de w.*

Demonstração. Suponhamos que $\mathrm{Ger}\{v, w\} = \mathbb{R}^2$. Então, para todo vetor $u = (x, y) \in \mathbb{R}^2$, existem escalares μ, λ satisfazendo $u = \mu v + \lambda w$. Daí, tem-se $x = \mu a + \lambda c$ e $y = \mu b + \lambda d$. Multiplicando-se a primeira destas equações por $-b$, a segunda por a e adicionando-as, obtemos $-bx + ay = \lambda(ad - bc)$. Façamos agora $u = (-b, a)$. Teremos então $a^2 + b^2 = \lambda(ad - bc)$. No entanto, $v = (a, b) \neq 0$, isto é, $a^2 + b^2 > 0$. Segue-se que $ad - bc \neq 0$ e, pelo Lema 2.1, que v não é múltiplo de w.

Reciprocamente, suponhamos que $v = (a, b)$ não seja múltiplo do vetor $w = (c, d)$, isto é, que $ad - bc \neq 0$. Dado então $u = (x, y) \in \mathbb{R}^2$, tomando-se

$$\lambda = \frac{ay - bx}{ad - bc} \quad \text{e} \quad \mu = \frac{dx - cy}{ad - bc},$$

verifica-se facilmente que $u = \mu v + \lambda w$, isto é, $u \in \mathrm{Ger}\{v, w\}$. Uma vez que u é arbitrário, segue-se que $\mathrm{Ger}\{v, w\} = \mathbb{R}^2$. ■

Segue-se das Proposições 2.4 e 2.5 que, se $v, w \in \mathbb{R}^2$ são não-nulos e um é múltiplo do outro, então $\mathrm{Ger}\{v, w\} = \mathrm{Ger}\{v\} = \mathrm{Ger}\{w\} \subsetneq \mathbb{R}^2$. Nesse caso, v e w satisfazem a relação $v - \lambda w = 0$ para algum escalar λ. Por outro lado, se $\mathrm{Ger}\{v, w\} = \mathbb{R}^2$, a única relação entre os vetores v e w do tipo

$$\mu v + \lambda w = \mathbf{0} \tag{2.1}$$

é aquela em que $\mu = \lambda = 0$. De fato, se tivéssemos, por exemplo, $\mu \neq 0$, teríamos, por (2.1), que $v = -\frac{\lambda}{\mu} w$, o que contradiria o resultado da Proposição 2.5.

Essas considerações motivam a definição que se segue.

Definição 2.4 (DEPENDÊNCIA LINEAR). Seja $\mathcal{A} = \{v_1, \ldots, v_n\}$ um conjunto de n vetores de um espaço vetorial V. Uma combinação linear desses vetores,

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = v,$$

é dita *nula*, se o vetor resultante v é o vetor nulo. Uma combinação linear nula $\lambda_1 v_1 + \cdots + \lambda_n v_n = \mathbf{0}$ é dita *trivial*, se todos os escalares λ_i são nulos, caso contrário, ela é dita *não trivial*. Diz-se, então, que o conjunto $\mathcal{A}$ é *linearmente dependente* (LD), quando existe uma combinação linear nula e não trivial de seus elementos, caso contrário, ele é dito *linearmente independente* (LI).

Sejam v e w vetores de um espaço vetorial V. Se v é múltiplo de w, então $\{v, w\}$ é LD pois, nesse caso, existe um escalar λ, tal que $v = \lambda w$, isto é, $v - \lambda w = \mathbf{0}$. Já o conjunto $\{(1,0),(0,1)\} \subset \mathbb{R}^2$ é LI. Com efeito, considerando-se uma combinação linear nula de seus elementos, $\lambda(1,0) + \mu(0,1) = (0,0)$, teremos $(\lambda, \mu) = (0,0)$, implicando $\lambda = \mu = 0$. Assim, a única combinação linear nula possível de $(1,0)$ e $(0,1)$ é a trivial. Logo, $\{(1,0),(0,1)\}$ é LI.

Exemplo 2.10. Consideremos $\mathcal{A} = \{(1,2,-1),(1,0,1),(0,-1,1)\} \subset \mathbb{R}^3$ e verifiquemos se este conjunto é LI ou LD. Para tanto, tomemos uma combinação linear nula de seus elementos

$$x(1,2,-1) + y(1,0,1) + z(0,-1,1) = (0,0,0)\,.$$

Daí, resulta o seguinte sistema linear

$$\left\{\begin{array}{rcrcrcl} x & + & y & & & = & 0 \\ 2x & & & - & z & = & 0 \\ -x & + & y & + & z & = & 0 \end{array}\right. .$$

Resolvendo-o, verifica-se que qualquer terno $(x, -x, 2x)$ é uma solução do mesmo. Tomando-se, então, $x \neq 0$, obtém-se uma combinação linear nula e não trivial dos vetores de $\mathcal{A}$, qual seja,

$$x(1,2,-1) - x(1,0,1) + 2x(0,-1,1) = \mathbf{0}.$$

Logo, $\mathcal{A}$ é linearmente dependente.

Procedendo-se como no exemplo acima, obtém-se facilmente o resultado seguinte (vide, também, Exercício 11).

Proposição 2.6. *Dado um subconjunto finito* $\mathcal{C} = \{w_1, \ldots, w_n\} \subset \mathbb{R}^m$, *seja* $A = [w_1 \cdots w_n]$ *a matriz* $m \times n$ *cujos vetores coluna são os elementos de* $\mathcal{C}$. *Então,* $\mathcal{C}$ *é* LI *se, e somente se, o sistema linear homogêneo* $AX = \mathbf{0}$ *admite apenas a solução trivial.*

Exemplo 2.11. Sejam f, $g : \mathbb{R} \to \mathbb{R}$ as função cosseno e seno, isto é, $f(x) = \cos x$ e $g(x) = \operatorname{sen} x$. O conjunto $\mathcal{A} = \{f, g\}$ é LI no espaço de funções $\mathcal{F}(\mathbb{R}, \mathbb{R})$. De fato, dada uma combinação linar nula de f e g, $\lambda f + \mu g = 0$, devemos ter, *para todo* $x \in \mathbb{R}$,

$$\lambda \cos x \,+\, \mu \operatorname{sen} x \,=\, 0\,.$$

Fazendo-se $x = 0$ e, em seguida, $x = \pi/2$, obtém-se $\lambda = \mu = 0$. Logo, $\mathcal{A}$ é um conjunto LI de vetores de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Proposição 2.7. *Sejam* $\mathcal{A} \subset \mathcal{B} \subset \mathcal{C}$ *subconjuntos finitos de um espaço vetorial* V. *Então, são verdadeiras as seguintes afirmações:*

i) $\mathcal{A}$ *é* LI, *se* $\mathcal{B}$ *é* LI;

ii) $\mathcal{C}$ *é* LD, *se* $\mathcal{B}$ *é* LD;

isto é, todo subconjunto de um conjunto LI *é* LI, *e todo conjunto que contém um conjunto* LD *é* LD.

Demonstração. Façamos $\mathcal{A} = \{v_1, \ldots, v_k\}$, $\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_m\}$ e $\mathcal{C} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_m, v_{m+1}, \ldots, v_n\}$, em que $k \le m \le n$. Suponhamos que $\mathcal{B}$ seja LI e tomemos uma combinação linear nula dos vetores de $\mathcal{A}$, $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$. Se algum dos λ_i's fosse diferente de zero, teríamos uma combinação linear nula não trivial dos elementos de $\mathcal{B}$, a saber,

$$\lambda_1 v_1 + \cdots + \lambda_k v_k + 0 v_{k+1} + \cdots + 0 v_m = \mathbf{0},$$

contrariando o fato de $\mathcal{B}$ ser LI. Segue-se que cada λ_i é nulo e, portanto, que $\mathcal{A}$ é LI, o que prova (i). Quanto ao item (ii), suponhamos que $\mathcal{B}$ seja

LD. Então, existe uma combinação linear nula e não trivial dos vetores de $\mathcal{B}$, $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$. Dessa forma,

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + 0v_{m+1} + \cdots + 0v_n = \mathbf{0}$$

é uma combinação linear nula e não trivial dos vetores de $\mathcal{C}$, donde se infere que $\mathcal{C}$ é LD. ■

Proposição 2.8. *Um subconjunto $\mathcal{A} \subset V$ de um espaço vetorial V é LD se, e somente se, um de seus vetores se expressa como combinação linear dos demais.*

Demonstração. Suponhamos que $\mathcal{A} = \{v_1, \ldots, v_n\}$ seja LD e tomemos uma combinação linear nula e não trivial dos vetores de $\mathcal{A}$,

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = \mathbf{0}. \tag{2.2}$$

Então, para algum $i \in \{1, 2, \ldots, n\}$, tem-se $\lambda_i \neq 0$. Daí e da igualdade (2.2), obtém-se

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} \cdots - \frac{\lambda_n}{\lambda_i} v_n\,,$$

isto é, v_i é combinação linear de $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.

Agora, se para algum $i \in \{1, 2, \ldots, n\}$, tem-se

$$v_i = \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n\,,$$

obtém-se daí a combinação linear nula e não trivial

$$\lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} - v_i + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n = \mathbf{0},$$

donde $\mathcal{A}$ é LD. ■

O conceito de dependência linear pode ser facilmente estendido a subconjuntos infinitos de um espaço vetorial V. Basta dizer que um tal conjunto é LI se qualquer um de seus subconjuntos finitos é LI no sentido da Definição 2.4. Caso contrário, diz-se que este conjunto é LD. Deixamos a cargo do leitor a tarefa de verificar que o conjunto

$$\mathcal{A} = \{1, t, t^2, \ldots, t^n, t^{n+1}, \ldots\}$$

é LI no espaço $\mathcal{P}[t, \mathbb{R}]$.

2.3 Bases – Dimensão

Nesta seção, consideraremos subconjuntos de espaços vetoriais, os quais chamamos de *base*, que possuem duas características: são linearmente independentes e geram o espaço vetorial que os contém. Constataremos, então, que duas bases finitas quaisquer de um espaço vetorial V têm, necessariamente, o mesmo número de elementos, o qual chamamos *dimensão* de V.

Muitas propriedades de um espaço vetorial são verificadas através de suas bases. Isto se dá, especialmente, quando do estudo de aplicações entre espaços vetoriais, como veremos nos capítulos subsequentes.

Definição 2.5 (BASE – DIMENSÃO FINITA)**.** Um subconjunto $\mathcal{B} \subset V$ de um espaço vetorial V é dito uma *base* do mesmo, se

i) $\mathcal{B}$ é LI;

ii) $\operatorname{Ger}\mathcal{B} = V$.

Diz-se, então, que V tem *dimensão finita*, quando contém uma base finita. Caso contrário, diz-se que V tem *dimensão infinita*.

Para estabelecer o principal resultado desta seção (Teorema 2.1), usaremos a proposição abaixo, que refina o resultado da Proposição 2.8.

Proposição 2.9. *Um conjunto de vetores não nulos de um espaço vetorial* V, $\mathcal{A} = \{v_1, \dots, v_n\}$, *é* LD *se, e somente se, para algum* $k \in \{2, \dots, n\}$, *o vetor* v_k *é uma combinação linear de* $v_1, \dots, v_{k-1}$.

Demonstração. Suponhamos que $\mathcal{A}$ seja LD. Defina $k \in \{2, \dots, n\}$ como o menor inteiro para o qual o conjunto $\{v_1, v_2, \dots, v_k\}$ seja LD (note que, se v_1 fosse o vetor nulo, um tal k não existiria). Dessa forma, existe uma combinação linear nula e não trivial de $v_1, \dots, v_k$,

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = \mathbf{0}.$$

Além disso, devemos ter $\lambda_k \neq 0,$ caso contrário $\lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} = \mathbf{0}$ seria uma combinação linear nula e não trivial de $v_1, \dots, v_{k-1}$, o que, juntamente com a Proposição 2.8, implicaria que estes vetores são LD, contrariando, dessa forma, a definição de k. Segue-se que $\lambda_k \neq 0$ e, portanto, que

$$v_k = -\frac{\lambda_1}{\lambda_k}v_1 - \cdots - \frac{\lambda_{k-1}}{\lambda_k}v_{k-1}\,,$$

isto é, v_k é combinação linear de $v_1, \ldots, v_{k-1}$.

Reciprocamente, suponhamos que, para algum $k \in \{2, \ldots, n\}$, tenha-se $v_k \in \operatorname{Ger}\{v_1, \ldots, v_{k-1}\}$. Nesse caso, pela Proposição 2.8, $\{v_1, \ldots, v_{k-1}, v_k\}$ é LD. Segue-se, então, da Proposição 2.7-(ii) que $\mathcal{A} = \{v_1, \ldots, v_k, \ldots, v_n\}$ é LD. ■

Corolário 2.1. *Sejam V um espaço vetorial e $\mathcal{A} = \{v_1, \ldots, v_n\} \subset V$ um subconjunto* LI *de V. Nessas condições, se $v \in V - \operatorname{Ger}\mathcal{A}$, então $\{v_1, \ldots v_n, v\}$ é* LI.

Demonstração. Pela Proposição 2.9, nenhum dos vetores v_i é combinação linear dos anteriores, uma vez que $\mathcal{A}$ é LI. Além disso, por hipótese, v não é combinação linear dos vetores de $\mathcal{A}$. Logo, pela Proposição 2.9, $\{v_1, \ldots v_n, v\}$ é linearmente independente. ■

O teorema a seguir nos permitirá introduzir um dos conceitos fundamentais da Álgebra Linear, o de *dimensão*.

Teorema 2.1. *Seja V um espaço vetorial de dimensão finita. Então, a quantidade de elementos de um conjunto* LI*, de V, nunca excede a de um conjunto de geradores desse espaço. Mais precisamente, se $\mathcal{A} = \{v_1, \ldots, v_n\} \subset V$ é* LI*, e $\mathcal{B} = \{w_1, \ldots, w_m\}$ gera V, então, $m \geq n$.*

Demonstração. Consideremos o conjunto $\mathcal{B}' = \{v_n, w_1, \ldots, w_m\}$. Uma vez que $\mathcal{B}$ gera V, temos, em particular, que $v_n \in \operatorname{Ger}\{w_1, \ldots, w_m\}$. Logo, pela Proposição 2.8, $\mathcal{B}'$ é LD. Então, pela Proposição 2.9, um dos vetores w_i, de $\mathcal{B}'$, é uma combinação linear dos anteriores a ele em $\mathcal{B}'$. Segue-se, então, da Proposição 2.4 e do fato de $\mathcal{B}'$ gerar V (já que $\mathcal{B}$ gera V), que

$$\mathcal{B}_1 = \{v_n, w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_m\}$$

gera V. Consideremos agora o conjunto

$$\mathcal{B}_1' = \{v_{n-1}, v_n, w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_m\}.$$

Como no caso de $\mathcal{B}'$, $\mathcal{B}_1'$ é LD e gera V. Procendo-se, então, de modo análogo, obtemos

$$\mathcal{B}_2 = \{v_{n-1}, v_n, w_1, \ldots, w_{i-1}, w_{i+1}, \ldots w_{j-1}, w_{j+1}, \ldots, w_m\},$$

tal que $\operatorname{Ger}\mathcal{B}_2 = V$.

Repetimos esse processo (de se retirar um elemento de $\mathcal{B}$ a cada elemento de $\mathcal{A}$ acrescentado) até que não haja mais v_i's a acrescentar (o que ocorrerá se $m \geq n$) ou w_i's a retirar (o que ocorrerá se $m \leq n$). Dessa forma, se tivéssemos $m < n$, terminaríamos com um subconjunto próprio de $\mathcal{A}$, $\{v_k, \ldots, v_n\}$, $k = (n-m)+1$, que geraria V. Em particular, cada um dos vetores $v_1, \ldots, v_{k-1}$ seria uma combinação linear de $v_1, \ldots, v_k$, contrariando o fato de $\mathcal{A}$ ser LI. Segue-se desta contradição que devemos ter, necessariamente, $m \geq n$, como queríamos demonstrar. ∎

Corolário 2.2. *Duas bases quaisquer de um espaço vetorial de dimensão finita têm a mesma quantidade de vetores.*

Demonstração. Com efeito, suponha que $\mathcal{A}$ e $\mathcal{B}$ sejam bases distintas de um tal espaço vetorial, com m e n elementos, respectivamente. Uma vez que, ambos, $\mathcal{A}$ e $\mathcal{B}$, são LI e geram V, pelo Teorema 2.1, devemos ter $m \geq n$ e $n \geq m$, donde $m = n$. ∎

Definição 2.6 (DIMENSÃO)**.** Seja V um espaço vetorial de dimensão finita. A *dimensão* de V, que será denotada por $\dim V$, define-se como o número de elementos de uma qualquer de suas bases. O espaço vetorial nulo, por convenção, tem dimensão 0.

Os vetores

$$e_1 = (1, 0, \ldots, 0)\,,\; e_2 = (0, 1, 0, \ldots, 0)\,,\; \ldots\,,\; e_n = (0, 0, \ldots 1)\,,$$

claramente, formam uma base de $\mathbb{R}^n$, dita a *base canônica* desse espaço. Segue-se que $\dim \mathbb{R}^n = n$. É fácil ver também que os vetores v_1, v_2, v_3, v_4, definidos no Exemplo 2.8, constituem uma base de $\mathcal{M}(2,2)$, donde se infere que $\dim \mathcal{M}(2,2) = 4$. De modo geral, tem-se $\dim \mathcal{M}(m,n) = mn$.

Dada uma base $\mathcal{B} = \{v_1, \dots, v_n\}$ de um espaço vetorial V de dimensão n, para todo $v \in V$, os escalares $\lambda_1, \dots, \lambda_n$ que cumprem a igualdade

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

são únicos. Eles são ditos as *coordenadas* de v com relação à base $\mathcal{B}$. Para constatarmos isso, suponhamos que

$$v = \mu_1 v_1 + \cdots + \mu_n v_n.$$

Nesse caso, devemos ter

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n ,$$

donde $(\lambda_1 - \mu_1)v_1 + \cdots + (\lambda_n - \mu_n)v_n = \mathbf{0}$. Uma vez que $\mathcal{B}$ é LI, segue-se que $\lambda_1 - \mu_1 = \cdots = \lambda_n - \mu_n = 0$. Logo, $\lambda_i = \mu_i \, \forall i \in \{1, \dots, n\}$.

Exemplo 2.12. O conjunto $\mathcal{B} = \{(1,1), (-1,1)\}$ é uma base de $\mathbb{R}^2$ (verifique). Dado um vetor $(x,y) \in \mathbb{R}^2$, determinemos suas coordenadas com respeito a essa base. Assim, buscamos escalares $\lambda, \mu \in \mathbb{R}$, os quais cumpram $(x,y) = \lambda(1,1) + \mu(-1,1)$, isto é, λ e μ são as variáveis do sistema:

$$\left\{ \begin{array}{ccccc} \lambda & - & \mu & = & x \\ \lambda & + & \mu & = & y \end{array} \right. ,$$

cuja solução é $\lambda = (x+y)/2$ e $\mu = (x-y)/2$. Dessa forma, temos

$$(x,y) = \frac{x+y}{2}(1,1) + \frac{x-y}{2}(-1,1).$$

Exemplo 2.13. No espaço $\mathbb{R}^3$, as retas que contém a origem $\mathbf{0} = (0,0,0)$ são os subespaços de dimensão 1, enquanto os planos que contém $\mathbf{0}$ são seus subespaços de dimensão 2. Esses são todos os subespaços não triviais de $\mathbb{R}^3$.

Exemplo 2.14. Seja $W \subset \mathcal{M}(2,2)$ o conjunto das matrizes quadradas de ordem 2 que têm traço nulo. Então, todo elemento de W se escreve como:

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix}, \; a, b, c \in \mathbb{R}.$$

Além disso, vale a seguinte igualdade:

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Logo, W é o subespaço de $\mathcal{M}(2,2)$ gerado pelo conjunto

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}.$$

Pode-se verificar facilmente que $\mathcal{B}$ é também LI. Dessa forma, $\mathcal{B}$ é uma base de W, donde se infere que $\dim W = 3$.

De forma análoga, mostra-se que o conjunto formado pelas matrizes quadradas de ordem n que têm traço nulo é um subespaço de $\mathcal{M}(n,n)$, cuja dimensão é $n-1$.

Proposição 2.10. *Seja V um espaço vetorial de dimensão $n > 0$. Então, são verdadeiras as seguintes afirmações:*

i) *Todo subconjunto finito de V, o qual possui mais de n vetores, é, necessariamente,* LD*;*

ii) *Todo subconjunto* LI *de V que possui n vetores é uma base de V;*

iii) *Todo subespaço W de V é de dimensão finita, e cumpre* $\dim W \leq \dim V$, *em que a igualdade ocorre se, e somente se, $V = W$;*

iv) *Para todo subconjunto* LI, $\{v_1, \ldots, v_k\} \subset V$, $1 \leq k < n$, *existem vetores $v_{k+1}, \ldots, v_n$, tais que $\{v_1, \ldots, v_k, \ldots, v_n\}$ constitui uma base de V (isto é, todo subconjunto* LI, *de V, pode ser "completado" de modo a formar uma base).*

Demonstração. (i) Seja $\mathcal{A}$ um subconjunto finito de V com mais de n vetores. Se $\mathcal{A}$ fosse LI, teríamos uma contradição com o Teorema 2.1, pois toda base de V tem n elementos e gera V. Logo, $\mathcal{A}$ é LD.

(ii) Consideremos um conjunto LI, $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$. Se $\mathcal{B}$ não gerasse V, existiria um vetor $v \in V$, tal que $v \notin \operatorname{Ger}\mathcal{B}$. Pelo Corolário 2.1, o conjunto $\{v_1, \ldots v_n, v\}$ seria LI, contradizendo (i). Dessa forma, $\mathcal{B}$ gera V e, portanto, é uma base desse espaço.

(iii) Se o subespaço W não fosse de dimensão finita, ou tivesse dimensão finita e maior que $\dim V$, existiria um conjunto LI, $\{w_1, \ldots w_m\} \subset W$, tal que $m > n$, o que contradiz (i). Assim, W tem dimensão finita e $\dim W \leq \dim V$. Ocorrendo a igualdade, tem-se, por (ii), que toda base de W é também uma base de V, implicando que $V = W$. A recíproca é imediata.

(iv) O conjunto $\{v_1, \ldots, v_k\} \subset V$, sendo LI, não pode gerar V, senão teríamos uma base de V com menos de n vetores. Logo, existe $v_{k+1} \in V - \operatorname{Ger}\{v_1, \ldots, v_k\}$, donde $\{v_1, \ldots, v_k, v_{k+1}\}$ é LI. Procedendo-se indutivamente, obtém-se um conjunto LI, $\{v_1, \ldots, v_k, \ldots, v_n\}$, o qual, por (ii), constitui uma base de V. ∎

Exemplo 2.15. Consideremos $\mathcal{B} = \{(1, -1, 1), (0, 1, 2), (1, 0, -2)\} \subset \mathbb{R}^3$ e denotemos por A a matriz cujas colunas são os vetores de $\mathcal{B}$. Resolvendo-se o sistema linear homogêneo associado, $AX = \mathbf{0}$, verifica-se que o mesmo admite apenas a solução trivial, donde se infere que $\mathcal{B}$ é LI. Uma vez que a dimensão de $\mathbb{R}^3$ é 3, segue-se da Proposição 2.10-(ii) que $\mathcal{B}$ é uma base de $\mathbb{R}^3$.

Adotemos a seguinte notação: Um vetor v de um espaço vetorial V, cujas coordenadas com respeito a uma base $\mathcal{B} = \{v_1, \ldots v_n\}$ de V são $x_1, \ldots, x_n$, será indicado por $(x_1, , \ldots, x_n)_{\mathcal{B}}$, isto é, valem, por definição, as igualdades:

$$v = (x_1, , \ldots, x_n)_{\mathcal{B}} = x_1 v_1 + \cdots + x_n v_n \,.$$

Como vimos fazendo, omitiremos a indicação de $\mathcal{B}$ quando se tratar da base canônica de $\mathbb{R}^n$.

Exemplo 2.16. Constatamos no Exemplo 2.12 que um vetor v de $\mathbb{R}^2$, cujas coordenadas com respeito à base canônica são x e y, tem $\frac{x+y}{2}$ e $\frac{x-y}{2}$ como coordenadas com respeito à base $\mathcal{B} = \{(1,1),(-1,1)\}$. Nesse caso, escrevemos

$$v = (x,y) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right)_{\mathcal{B}}.$$

Nos será conveniente, também, associarmos a v uma matriz coluna $[v]_{\mathcal{B}}$, cujas entradas são suas coordenadas com respeito à base $\mathcal{B}$, isto é,

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

No caso em que $V = \mathbb{R}^n$ e $\mathcal{B}$ é a sua base canônica, denotamos $[v]_{\mathcal{B}}$ simplesmente por $[v]$ (vide Exercício 7 – Cap. 1).

Segue-se diretamente da unicidade das coordenadas de um vetor com respeito a uma base, bem como das propriedades da adição e multiplicação por escalar em V, que a correspondência

$$v \in V \leftrightarrow [v]_{\mathcal{B}} \in \mathcal{M}(n,1)$$

é biunívoca, e cumpre as seguintes igualdades:

- $[u+v]_{\mathcal{B}} = [u]_{\mathcal{B}} + [v]_{\mathcal{B}}$;
- $[\lambda u]_{\mathcal{B}} = \lambda [u]_{\mathcal{B}}$.

2.4 O Posto de uma Matriz

Nesta seção, temos como objetivo desenvolver um método que nos permita extrair uma base de um conjunto de geradores de um subespaço W de um espaço de dimensão finita, V. Esse processo nos conduzirá naturalmente ao conceito de *posto* de uma matriz. Uma vez que usaremos coordenadas, nos limitaremos ao caso $V = \mathbb{R}^n$.

Dada uma matriz $A = (a_{ij})_{m\times n}$, o subespaço $\mathscr{L}_A \subset \mathbb{R}^n$, gerado pelos vetores linha de A, é dito o *espaço linha* de A, enquanto o subespaço $\mathscr{C}_A \subset \mathbb{R}^m$, gerado pelos vetores coluna de A, é chamado de *espaço coluna* de A.

Consideremos, por exemplo, a matriz

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

Seus espaços linha e coluna são $\mathscr{L}_A = \operatorname{Ger}\{(1,2,-1),(2,-1,1)\} \subset \mathbb{R}^3$ e $\mathscr{C}_A = \operatorname{Ger}\{(1,2),(2,-1),(-1,1)\} \subset \mathbb{R}^2$. Note que,

$$\dim \mathscr{L}_A = \dim \mathscr{C}_A = 2.$$

No caso da matriz

$$B = \begin{bmatrix} 2 & -4 \\ 1 & -2 \\ 3 & -6 \end{bmatrix},$$

temos que $\mathscr{L}_B = \operatorname{Ger}\{(2,-4),(1,-2),(3,-6)\} = \operatorname{Ger}\{(1,-2)\} \subset \mathbb{R}^2$, enquanto $\mathscr{C}_B = \operatorname{Ger}\{(2,1,3),(-4,-2,-6)\} = \operatorname{Ger}\{(2,1,3)\} \subset \mathbb{R}^3$. Em particular, $\dim \mathscr{L}_B = \dim \mathscr{C}_B = 1$.

Os exemplos acima sugerem o resultado seguinte, cuja peculiaridade reside no fato de os espaços linha e coluna de uma matriz $m \times n$, com $m \neq n$, serem subespaços de espaços euclidianos distintos.

Proposição 2.11. *O espaço linha e o espaço coluna de uma matriz arbitrária A têm mesma dimensão, a qual chama-se* posto *de A.*

Demonstração. Consideremos uma matriz $A = (a_{ij})_{m\times n}$, e designemos os seus vetores linha e coluna, respectivamente, por $u_1, \ldots, u_m$ e $v_1, \ldots, v_n$, isto é,

$$A = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}.$$

Tomemos uma base $\mathcal{B} = \{w_1, \cdots, w_k\}$, de $\mathscr{L}_A$, e escrevamos

$$w_p = (\lambda_{p1}, \ldots, \lambda_{pn}) = \sum_{j=1}^{n} \lambda_{pj} e_j, \quad p = 1, \ldots k, \tag{2.3}$$

em que $\{e_1\,, \ldots, e_n\}$ é a base canônica de $\mathbb{R}^n$. Designemos por $\mu_{ip}\,, p = 1, \ldots, k$, as coordenadas de u_i com respeito a $\mathcal{B}$. Assim, temos que,

$$u_i = \sum_{p=1}^{k} \mu_{ip} w_p = \sum_{p=1}^{k} \mu_{ip} \left(\sum_{j=1}^{n} \lambda_{pj} e_j \right) = \sum_{j=1}^{n} \left(\sum_{p=1}^{k} \mu_{ip} \lambda_{pj} \right) e_j\,. \tag{2.4}$$

Uma vez que

$$u_i = (a_{i1}\,, \ldots, a_{in}) = \sum_{j=1}^{n} a_{ij} e_j\,,$$

segue-se da unicidade das coordenadas de u_i com respeito à base canônica, que

$$a_{ij} = \sum_{p=1}^{k} \mu_{ip} \lambda_{pj} \quad \forall i \in \{1, \ldots, m\},\; j \in \{1, \ldots, n\}. \tag{2.5}$$

Definindo-se, então, para cada $p \in \{1, \cdots, k\}$, o vetor

$$w'_p = \sum_{i=1}^{p} \mu_{ip} e_i,$$

a igualdade (2.5) nos dá

$$v_j = \sum_{p=1}^{k} \lambda_{pj} w'_p\,.$$

Logo, $\mathscr{C}_A \subset \operatorname{Ger}\{w'_1, \ldots w'_k\}$, donde

$$\dim \mathscr{C}_A \leq k = \dim \mathscr{L}_A\,. \tag{2.6}$$

Finalmente, aplicando-se (2.6) à transposta A^* da matriz A, obtém-se

$$\dim \mathscr{C}_A = \dim \mathscr{L}_{A^*} \geq \dim \mathscr{C}_{A^*} = \dim \mathscr{L}_A\,,$$

o que implica $\dim \mathscr{L}_A = \dim \mathscr{C}_A$, como desejado. ■

Proposição 2.12. *Os vetores linha não nulos de uma matriz escalonada são linearmente independentes. Em particular, eles formam uma base do espaço linha dessa matriz.*

Demonstração. De fato, sejam $u_1, \ldots, u_r \subset \mathbb{R}^n$ os vetores linha não nulos de uma matriz escalonada $A = (a_{ij})_{m\times n}$. Nesse caso, se a_{1k}, $1 \leq k \leq n$, é a primeira coordenada não nula de u_1, então a k-ésima coordenada dos demais vetores, $u_2, \ldots, u_r$, é igual a zero. Assim, u_1 não é combinação linear de $u_2, \ldots, u_r$. Analogamente, u_2 não é combinação linear de $u_3, \ldots, u_r$ e assim por diante, isto é, considerando-se os vetores linha de A na ordem

$$u_r\,, u_{r-1}\,, \ldots, u_1\,,$$

tem-se que nenhum deles se escreve como combinação linear dos anteriores. O resultado segue-se, então, da Proposição 2.9. ■

Proposição 2.13. *Sejam $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$ matrizes linha equivalentes. Então, $\mathscr{L}_A = \mathscr{L}_B$. Em particular, o posto de A é igual ao de B.*

Demonstração. Seja $L = \{u_1, \ldots, u_m\} \subset \mathbb{R}^n$ o conjunto formado pelas linhas não nulas de A. Evidentemente, a permutação de linhas, ou a multiplicação de uma linha por um escalar não nulo, não altera o espaço $\mathscr{L}_A = \operatorname{Ger} L$. Consideremos, então, o conjunto $L' = \{u_1, \ldots, u_{j-1}, u_j + \lambda u_i, u_{j+1}, \ldots, u_m\}$, obtido de L substituindo-se o seu j-ésimo vetor u_j por $u_j + \lambda u_i$, em que λ é um escalar não nulo, e $1 \leq i \neq j \leq m$. Devemos, dessa forma, mostrar que $\operatorname{Ger} L' = \operatorname{Ger} L$. Para tanto, observemos que, dado $v \in \operatorname{Ger} L$, existem escalares $\mu_1, \ldots, \mu_m$, tais que

$$\begin{aligned}
v &= \mu_1 u_1 + \cdots + \mu_i u_i + \cdots \mu_j u_j + \cdots + \mu_m u_m \\
&= \mu_1 u_1 + \cdots + (\mu_i + \lambda\mu_j - \lambda\mu_j)u_i + \cdots + \mu_j u_j + \cdots + \mu_m u_m \\
&= \mu_1 u_1 + \cdots + (\mu_i - \lambda\mu_j)u_i + \cdots + \mu_j(u_j + \lambda u_i) + \cdots + \mu_m u_m\,,
\end{aligned}$$

donde $v \in \operatorname{Ger} L'$, isto é, $\operatorname{Ger} L \subset \operatorname{Ger} L'$.

Fazendo-se, agora, $u'_j = u_j + \lambda u_i$, tem-se que L e L' se escrevem como:

- $L' = \{u_1, \ldots, u_{j-1}, u'_j, u_{j+1}, \ldots, u_m\}$;
- $L = \{u_1, \ldots, u_i, \ldots, u'_j - \lambda u_i, \ldots, u_m\}$.

Logo, os argumentos acima se aplicam igualmente, implicando que $\operatorname{Ger} L' \subset \operatorname{Ger} L$ e, portanto, que $\operatorname{Ger} L' = \operatorname{Ger} L$, como queríamos demonstrar. ■

Exemplo 2.17. Determinemos, através dos resultados acima, uma base para o subespaço W, de $\mathbb{R}^4$, gerado pelo seguinte conjunto:

$$L = \{(1,0,2,-1),(3,-2,1,1),(4,-2,3,0),(2,-2,-1,2)\}.$$

Para tanto, basta considerarmos a matriz A cujas linhas são os vetores de L, e verificar que a mesma é linha equivalente à matriz escalonada

$$E = \begin{bmatrix} 1 & 0 & 2 & -1 \\ 0 & -2 & -5 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Logo, $\mathcal{B} = \{(1,0,2,-1),(0,-2,-5,4)\}$ é uma base de $\mathscr{L}_E = \mathscr{L}_A = W$. Em particular, $\dim W = 2$.

Proposição 2.14. *São equivalentes as seguintes afirmações acerca de uma matriz quadrada $A \in \mathcal{M}(n,n)$:*

i) *A é invertível;*

ii) *Para toda matriz $B \in \mathcal{M}(n,1)$, o sistema linear $AX = B$ admite uma única solução;*

iii) *A tem posto n.*

Demonstração. (i) $\Rightarrow$ (ii): Dada $B \in \mathcal{M}(n,1)$, sendo A invertível, temos que $X = A^{-1}B$ é, claramente, a única solução de $AX = B$.

(ii) $\Rightarrow$ (i): Seja $N = [w_1 \cdots w_n] \in \mathcal{M}(n)$ uma matriz arbitrária. Por hipótese, para cada $i = 1, \ldots, n$, existe um único $v_i \in \mathbb{R}^n$, tal que $A[v_i] = [w_i]$. Logo, $M = [v_1 \cdots v_n]$ é a única matriz de $\mathcal{M}(n)$ que satisfaz a igualdade $AM = N$. Dito de outra forma, a equação matricial

$$AY = N, \quad N \in \mathcal{M}(n), \tag{2.7}$$

admite sempre uma única solução $Y \in \mathcal{M}(n)$. Em particular, se $N = A$, então $Y = I$. Agora, fazendo-se $N = I$, temos que existe $M \in \mathcal{M}(n,n)$, tal que

$AM = I$. Escrevendo-se, então, $I' = MA$, e multiplicando-se à esquerda por A, obtém-se $AI' = A$, implicando que $I' = I$. Logo, $AM = MA = I$, donde A é invertível e $M = A^{-1}$.

(ii) $\Rightarrow$ (iii): Por hipótese, o sistema $AX = \mathbf{0}$ admite apenas a solução trivial. Logo, pela Proposição 2.6, os n vetores coluna de A são LI, donde se infere que o posto de A é n.

(iii) $\Rightarrow$ (ii): Escrevendo-se $A = [v_1 \cdots v_n]$, tem-se que $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$, já que, por hipótese, A tem posto n. Dada $B \in \mathcal{M}(n,1)$, seja $w \in \mathbb{R}^n$, tal que $B = [w]$. Fazendo-se, então, $C = [w]_{\mathcal{B}}$, tem-se, pelas considerações da Observação 2.1, que C é a única solução do sistema $AX = B$. ■

2.5 Somas de Subespaços

Sejam V um espaço vetorial e $U, W \subset V$ subespaços de V. Definimos a *soma* de U e W, $U + W$, como o subconjunto de V formado por todos as somas do tipo $u + w$, em que $u \in U$ e $w \in W$, isto é,

$$U + W = \{v \in V;\, v = u + w,\ u \in U,\ w \in W\}.$$

No caso em que $U \cap W = \{\mathbf{0}\}$, dizemos que essa soma é *direta* e a denotamos por $U \oplus W$.

Proposição 2.15. *Sejam U e W subespaços de um espaço vetorial V. Então, $U + W$ é um subespaço de V.*

Demonstração. Temos que $\mathbf{0} \in U \cap W$. Logo, $\mathbf{0} = \mathbf{0} + \mathbf{0} \in U + W$. Tomemos $v, v' \in U + W$. Então, existem $u, u' \in U$ e $w, w' \in W$, tais que $v = u + w$ e $v' = u' + w'$. Uma vez que U e W são subespaços de V, temos que $u + u' \in U$, e $w + w' \in W$. Além disso, dado $\lambda \in \mathbb{R}$, tem-se $\lambda u \in U$ e $\lambda w \in W$. Logo, $v + v' = (u + u') + (w + w') \in U + W$ e $\lambda v = \lambda u + \lambda w \in U + W$, donde se conclui que $U + W$ é um subespaço vetorial de V. ■

Exemplo 2.18. Sejam $U = \mathrm{Ger}\,\{(1,0)\}$ e $W = \mathrm{Ger}\,\{(1,1)\}$ os subespaços de $\mathbb{R}^2$ gerados, respectivamente, por $(1,0)$ e $(1,1)$. Claramente, $\{(1,0),(1,1)\}$ é uma base de $\mathbb{R}^2$. Logo, cada $v \in \mathbb{R}^2$ se escreve como $v = \lambda(1,0) + \mu(1,1)$. Fazendo-se $\lambda(1,0) = u \in U$ e $\mu(1,1) = w \in W$, temos que $v \in U + W$. Além disso, $U \cap W = \{\mathbf{0}\}$. Logo, $\mathbb{R}^2 = U \oplus W$.

Exemplo 2.19. Considere, em $\mathbb{R}^3$, os vetores $v_1 = (1,0,0)$, $v_2 = (0,1,0)$ e $v_3 = (1,1,-1)$, juntamente com os subespaços $U = \mathrm{Ger}\{v_1, v_2\}$ e $W = \mathrm{Ger}\{v_1, v_3\}$. Pode-se verificar facilmente que $\mathbb{R}^3 = U + W$ e que $U \cap W = \mathrm{Ger}\{v_1\} \neq \{\mathbf{0}\}$. Em particular, essa soma não é direta.

Proposição 2.16. *Sejam U e W subespaços de um espaço vetorial V. Então, se $U \cap W = \{\mathbf{0}\}$, são verdadeiras as seguintes afirmarções:*

i) *Para todo $v \in U \oplus W$, existem* únicos *vetores $u \in U$ e $w \in W$, tais que $v = u + w$;*

ii) $\dim(U \oplus W) = \dim U + \dim W$, *se V tiver dimensão finita.*

Demonstração. Dado $v \in U \oplus W$, suponhamos que

$$v = u + w \quad \text{e} \quad v = u' + w', \;\; u, u' \in U, \;\; w, w' \in W.$$

Daí, temos que $u + w = u' + w'$, isto é, $u - u' = w' - w$. Uma vez que U e W são subespaços de V, devemos ter $u - u' \in U$ e $w' - w \in W$. Porém $U \cap W = \{\mathbf{0}\}$. Logo, $u - u' = w' - w = \mathbf{0}$, o que nos dá $u = u'$ e $w = w'$, provando, assim, a asserção (i).

Quanto a (ii), tomemos bases $\mathcal{A} = \{u_1, \cdots, u_m\}$ e $\mathcal{B} = \{w_1, \cdots, w_n\}$, de U e W, respectivamente. Provemos, então, que $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ é uma base de $U \oplus W$, donde se seguirá o resultado.

É imediato que $\mathrm{Ger}\, \mathcal{C} \subset U \oplus W$. Para a inclusão contrária, consideremos $u \in U$ e $w \in W$, e os escalares $\lambda_1, \cdots, \lambda_m$ e $\mu_1, \ldots, \mu_n$, tais que

$$u = \sum_{i=1}^{m} \lambda_i u_i \quad \text{e} \quad w = \sum_{j=1}^{n} \mu_j w_j.$$

Assim, temos que

$$u + w = \sum_{i,j=1}^{m,n} (\lambda_i u_i + \mu_j w_j) \in \mathrm{Ger}\, \mathcal{C},$$

ou seja, $U \oplus W \subset \operatorname{Ger} \mathcal{C}$. Dessa forma, $U \oplus W = \operatorname{Ger} \mathcal{C}$. Finalmente, tomando-se uma combinação linear nula dos vetores de $\mathcal{C}$,

$$\lambda_1 u_1 + \cdots + \lambda_m u_m + \mu_1 w_1 + \cdots + \mu_n w_n = \mathbf{0}\,,$$

devemos ter, pelo resultado do item (i),

$$\lambda_1 u_1 + \cdots + \lambda_m u_m = \mathbf{0} \quad \text{e} \quad \mu_1 w_1 + \cdots + \mu_n w_n = \mathbf{0}\,,$$

pois $\lambda_1 u_1 + \cdots + \lambda_m u_m \in U$ e $\mu_1 w_1 + \cdots + \mu_n w_n \in W$. Porém $\{u_1, \cdots u_m\}$ e $\{w_1, \cdots w_n\}$ são bases. Em particular, estes conjuntos são LI. Logo,

$$\lambda_1 = \cdots = \lambda_m = \mu_1 = \cdots = \mu_m = 0.$$

Dessa forma, $\mathcal{C}$ é LI e, portanto, constitui uma base de $U \oplus W$. ■

2.6 Exercícios

Seção 2.1

1. Mostre a unicidade do vetor nulo de um espaço vetorial arbitrário.

2. Verifique, em cada um dos itens abaixo, se o conjunto $W \subset V$ é um subespaço do espaço vetorial V.

   a) $W = \{(x, y) \in \mathbb{R}^2\,;\, x = -y\},\ V = \mathbb{R}^2$;

   b) $W = \{(x, y) \in \mathbb{R}^2\,;\, y = x^2\},\ V = \mathbb{R}^2$;

   c) $W = \{(x, y, z) \in \mathbb{R}^3\,;\, x + y + z \geq 0\},\ V = \mathbb{R}^3$;

   d) $W = \{(x, y, z) \in \mathbb{R}^3\,;\, z = 0\},\ V = \mathbb{R}^3$;

   e) $W = \{A \in \mathcal{M}(n)\,;\, A \text{ é diagonal}\},\ V = \mathcal{M}(n)$;

   f) $W = \{X \in \mathcal{M}(n, 1)\,;\, AX = \mathbf{0},\ A \in \mathcal{M}(m, n)\},\ V = \mathcal{M}(n, 1)$;

   g) $W = \{p \in \mathcal{P}[\mathbb{R}, t]\,;\, p = \mathbf{0} \text{ ou } \operatorname{grau}(p) \leq n,\ n \in \mathbb{N}\},\ V = \mathcal{P}[\mathbb{R}, t]$;

   h) $W = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R})\,;\, f(x) = f(-x)\},\ V = \mathcal{F}(\mathbb{R}, \mathbb{R})$.

i) $W = \{f \in \mathcal{F}(\mathbb{R},\mathbb{R})\,;\, f(x) = -f(-x)\}$, $V = \mathcal{F}(\mathbb{R},\mathbb{R})$.

3. Sejam U e W subespaços de um espaço vetorial V. Mostre que $U \cup W$ é um subespaço de V se, e somente se, $U \subset W$ ou $W \subset U$.

4. Determine se, em $\mathbb{R}^4$, o vetor $(4,-2,5,-1)$ está no subespaço gerado pelos vetores $(1,-1,0,1), (0,1,1,-1)$ e $(1,-1,1,0)$.

5. Seja $\mathcal{A}$ um subconjunto arbitrário de um espaço vetorial V, não necessariamente finito. Mostre que o conjunto de todas as combinações lineares de vetores de $\mathcal{A}$ é um subespaço de V, o qual denotamos, como no caso finito, por $\operatorname{Ger}\mathcal{A}$. Mostre que $\operatorname{Ger}\mathcal{A}$ coincide com a interseção de todos os subespaços de V que contêm $\mathcal{A}$.

6. Dado $\mathcal{C} \subset \mathbb{R}$, defina $N(\mathcal{C}) = \{f \in \mathcal{F}(\mathbb{R},\mathbb{R})\,;\, f(x) = 0\,\forall x \in \mathcal{C}\}$. Prove, então, as seguintes afirmações:

   a) $N(\mathcal{C})$ é um subespaço de $\mathcal{F}(\mathbb{R},\mathbb{R})$ $\forall \mathcal{C} \subset \mathbb{R}$;

   b) $\mathcal{C}_1 \subset \mathcal{C}_2 \;\Rightarrow\; N(\mathcal{C}_2) \subset N(\mathcal{C}_1)$;

   c) $N(\mathcal{C}_1 \cup \mathcal{C}_2) = N(\mathcal{C}_1) \cap N(\mathcal{C}_2)$;

   d) $N(\mathcal{C}_1 \cap \mathcal{C}_2) = N(\mathcal{C}_1) + N(\mathcal{C}_2)$;

   e) $N(\mathcal{C}) = \{\mathbf{0}\} \Leftrightarrow \mathcal{C} = \mathbb{R}$;

   f) $N(\mathcal{C}_1) \oplus N(\mathcal{C}_2) = \mathcal{F}(\mathbb{R},\mathbb{R}) \Leftrightarrow \mathcal{C}_2 = \mathbb{R} - \mathcal{C}_1$.

7. Dada uma matriz $A \in \mathcal{M}(m,n)$, prove que o conjunto solução do sistema linear homogêneo $AX = \mathbf{0}$ é um subespaço vetorial de $\mathbb{R}^n$.

8. Encontre um sistema linear homogêneo cujo conjunto solução seja o subespaço de $\mathbb{R}^4$ gerado pelos vetores $(-1,0,1,2), (3,4,-2,5)$ e $(1,4,0,9)$.

Seção 2.2

9. Determine três vetores linearmente dependentes em $\mathbb{R}^3$, tais que dois quaisquer deles sejam linearmente independentes.

10. Para que valores de λ os vetores $(\lambda, 1, 0)$, $(1, \lambda, 1)$ e $(0, 1, \lambda)$ são linearmente dependentes em $\mathbb{R}^3$?

11. Demonstre a Proposição 2.6 com base na Observação 2.1.

12. Sejam u, v, w vetores LI de um espaço vetorial V. Determine se o conjunto $\{u+v, u+w, v+w\}$ é LI ou LD.

Seção 2.3

13. Em cada item abaixo, encontre uma base para o subespaço $W \subset \mathbb{R}^n$.

    a) $W = \operatorname{Ger}\{(1,-4,3),(3,-14,17),(0,-2,5)\} \subset \mathbb{R}^3$;

    b) $W = \operatorname{Ger}\{(1,2,3,4),(5,6,7,8),(9,10,11,12)\} \subset \mathbb{R}^4$;

14. Mostre que cada um dos seguintes conjuntos é um subespaço de $\mathcal{M}(n)$ e, em seguida, determine sua dimensão.

    a) $\{A \in \mathcal{M}(n)\,;\ A$ é triangular superior$\}$;

    b) $\{A \in \mathcal{M}(n)\,;\ A$ é triangular inferior$\}$;

    c) $\{A \in \mathcal{M}(n)\,;\ A$ é simétrica$\}$;

    d) $\{A \in \mathcal{M}(n)\,;\ A$ é antissimétrica$\}$;

    e) $\{A \in \mathcal{M}(n)\,;\ A$ tem traço nulo$\}$.

15. Exiba bases de $\mathcal{M}(2)$, bem como de $\mathcal{M}(3)$, cujas matrizes sejam todas invertíveis.

Seção 2.4

16. Prove que $\mathscr{L}_A = \mathscr{C}_A$, em que A é a matriz

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

17. Seja $A = (a_{ij})$ uma matriz quadrada de ordem $n \geq 3$. Suponha que existam reais não nulos λ, μ, tais que

$$a_{ij} = \lambda i + \mu j \;\; \forall 1 \leq i, j \leq n.$$

Prove que A não é invertível e conclua que vale o mesmo para a matriz

$$\begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix}.$$

18. Seja A uma matriz 5×5, tal que seus vetores linha u_i, $i = 1, \ldots 5$, satisfazem a seguinte relação:

$$u_1 + u_2 - 2u_4 + 3u_5 = \mathbf{0}.$$

Encontre uma matriz C, de posto 3, tal que CA tenha linhas u_1, u_4 e $\mathbf{0}$.

19. Dada uma matriz $A = (a_{ij})_{2\times 2}$, prove que todo sistema linear $AX = B$ admite uma única solução se, e somente se, $a_{11}a_{22} - a_{12}a_{21} \neq 0$.

Seção 2.5

20. Prove que, se U e W são subespaços de um espaço vetorial V de dimensão finita, então $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

21. Seja $\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ uma base de um espaço vetorial V. Defina, $U = \mathrm{Ger}\,(v_1, \ldots, v_k)$ e $W = \mathrm{Ger}\,(v_{k+1}, \ldots, v_n)$, e mostre que $V = U \oplus W$.

22. Sejam W_p e W_i os subespaços de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ definidos nos itens (h) e (i) do Exercício 2. Prove que $\mathcal{F}(\mathbb{R}, \mathbb{R}) = W_p \oplus W_i$.

3

Transformações Lineares

No capítulo anterior, vimos que os espaços vetoriais, juntamente com suas operações, constituem uma determinada estrutura. Neste, vamos estudar as funções entre esses espaços, as quais, num certo sentido, preservam essa estrutura, ditas *transformações lineares*.

Através das transformações lineares, estabelece-se uma relação de equivalência entre espaços vetoriais, dita *isomorfismo*. Constataremos, então, que dois espaços vetoriais de dimensão finita são equivalentes (isto é, *isomorfos*) se, e somente se, têm mesma dimensão. Em particular, todo espaço vetorial de dimensão finita é isomorfo a algum espaço euclidiano $\mathbb{R}^n$.

A propriedade fundamental de uma transformação linear é a de ser determinada por seus valores numa base, pois, a partir dela, verifica-se a existência de uma íntima relação entre as transformações lineares e as matrizes. Mais especificamente, para cada transformação linear entre espaços de dimensão finita com bases predeterminadas, existe uma única matriz a ela associada. Como consequência, o estudo das transformações lineares reduz-se, em muitos aspectos, ao das matrizes. Esse fato, na verdade, é uma das características mais marcantes da Álgebra Linear, conforme verificaremos neste e nos capítulos subsequentes.

3.1 Linearidade de Aplicações

As funções entre espaços vetoriais (reais) são designadas como tal, quando o seu contradomínio é o conjunto dos números reais. Caso contrário, elas são chamadas de *transformações* ou *aplicações*. As transformações lineares entre espaços vetoriais são aquelas que preservam as suas estruturas, conforme a definição que se segue.

Definição 3.1 (TRANSFORMAÇÃO LINEAR)**.** Sejam V e W espaços vetoriais. Uma aplicação $T : V \to W$ é dita uma *transformção linear*, se ela preserva as operações de adição e produto por escalar, isto é, se, para quaisquer $u, v \in V$ e $\lambda \in \mathbb{R}$, tem-se:

- $T(u + v) = T(u) + T(v)$;
- $T(\lambda u) = \lambda T(u)$.

Uma tal transformação em que $W = \mathbb{R}$ é dita um *funcional linear* de V, e transformações lineares de um espaço vetorial nele mesmo são chamadas também de *operadores lineares*.

Segue-se da definição que *toda transformação linear* $T : V \to W$ *leva o vetor nulo de* V *no vetor nulo de* W. De fato, $T(\mathbf{0}) = T(0.\mathbf{0}) = 0.T(\mathbf{0}) = \mathbf{0}$.

Dados espaços vetoriais V e W, verifica-se facilmente que são lineares a *aplicação identidade* de V e a *aplicação nula* de V em W, definidas por

$$\begin{array}{rccc} I: & V & \to & V \\ & v & \mapsto & v \end{array} \quad \text{e} \quad \begin{array}{rccc} T: & V & \to & W \\ & v & \mapsto & \mathbf{0} \end{array}.$$

Temos também que a *homotetia* de fator $a \in \mathbb{R}$, definida por

$$\begin{array}{rccc} T: & V & \to & V \\ & v & \mapsto & av \end{array},$$

é um operador linear. Com efeito, dados $u, v \in V$ e $\lambda \in \mathbb{R}$, tem-se

- $T(u + v) = a(u + v) = au + av = T(u) + T(v)$;

- $T(\lambda u) = a(\lambda u) = \lambda(au) = \lambda T(u)$.

Observemos que, se $T : V \to W$ for uma transformação linear, então

$$T(u + \lambda v) = T(u) + \lambda T(v) \;\; \forall u, v \in V,\, \lambda \in \mathbb{R}. \tag{3.1}$$

Reciprocamente, a validez de (3.1) implica que T é linear. Com efeito, tomando-se $\lambda = 1$, tem-se $T(u + v) = T(u) + T(v)$ e, tomando-se $u = \mathbf{0}$, tem-se $T(\lambda v) = \lambda T(v)$. Assim, podemos adotar a condição (3.1) como critério de linearidade de aplicações.

Exemplo 3.1 (PROJEÇÕES)**.** Suponha que U e W sejam subespaços de V, tais que $V = U \oplus W$. Nesse caso, segue-se da Proposiçao 2.16 que cada $v \in V$ se escreve de modo único como $v = u + w$, em que $u \in U$ e $w \in W$. Assim, ficam bem definidas as *projeções* de V sobre U e W, as quais são definidas, respectivamente, por

$$\begin{array}{rccc} P_U : & V & \to & U \\ & u+w & \mapsto & u \end{array} \quad \text{e} \quad \begin{array}{rccc} P_W : & V & \to & W \\ & u+w & \mapsto & w \end{array}.$$

Dados $v = u + w$ e $v' = u' + w'$ em V, e $\lambda \in \mathbb{R}$, temos que

$$P_U(v + \lambda v') = P_U((u + \lambda u') + (w + \lambda w')) = u + \lambda u' = P_U(v) + \lambda P_U(v'),$$

donde se conclui que P_U é linear. De modo análogo, verifica-se que a projeção P_W , igualmente, é linear.

Exemplo 3.2. A função $f : \mathbb{R}^2 \to \mathbb{R}$, definida por $f(x, y) = 2x + y$, é um funcional linear, pois, dados $u = (x_1, y_1)$, $v = (x_2, y_2) \in \mathbb{R}^2$, e $\lambda \in \mathbb{R}$, tem-se

$$\begin{aligned} f(u + \lambda v) &= f(x_1 + \lambda x_2, y_1 + \lambda y_2) \\ &= 2(x_1 + \lambda x_2) + (y_1 + \lambda y_2) \\ &= (2x_1 + y_1) + \lambda(2x_2 + y_2) \\ &= f(u) + \lambda f(v)\,. \end{aligned}$$

Exemplo 3.3. Consideremos agora a aplicação $T : \mathbb{R}^2 \to \mathbb{R}^3$, em que

$$T(x, y) = (x - y, x + y, y).$$

Dados $u = (x_1, y_1)$, $v = (x_2, y_2) \in \mathbb{R}^2$, tem-se

$$\begin{aligned} T(u + \lambda v) &= T(x_1 + \lambda x_2, y_1 + \lambda y_2) \\ &= ((x_1 + \lambda x_2) - (y_1 + \lambda y_2), (x_1 + \lambda x_2) + (y_1 + \lambda y_2), y_1 + \lambda y_2) \\ &= (x_1 - y_1, x_1 + y_1, y_1) + \lambda(x_2 - y_2, x_2 + y_2, y_2) \\ &= T(u) + \lambda T(v), \end{aligned}$$

donde se infere que T é uma transformação linear.

Exemplo 3.4. A aplicação $T : \mathbb{R}^2 \to \mathbb{R}^2$, em que $T(x, y) = (x^2, xy)$, *não* é linear. Com efeito, dados $u = (x, y) \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$, tem-se

$$T(\lambda u) = T(\lambda x, \lambda y) = (\lambda^2 x^2, \lambda^2 xy) = \lambda^2 T(u).$$

Assim, se λ for diferente de zero e de 1, $T(\lambda u) \neq \lambda T(u)$, donde T não é linear.

Exemplo 3.5 (LINEARIDADE DO TRAÇO)**.** Conforme verificamos anteriormente, para quaisquer matrizes quadradas A, B e $\lambda \in \mathbb{R}$, valem as seguintes igualdades:

$$\text{traço}(A + B) = \text{traço}(A) + \text{traço}(B) \ \text{ e } \ \text{traço}(\lambda A) = \lambda\, \text{traço}(A).$$

Segue-se, então, desse fato que

$$\begin{array}{rccc} f: & \mathcal{M}(n) & \to & \mathbb{R} \\ & A & \mapsto & \text{traço}(A) \end{array}$$

é um funcional linear de $\mathcal{M}(n)$.

Verifiquemos agora que uma transformação linear fica determinada por seus valores numa base. Mais especificamente, vale o resultado seguinte.

Proposição 3.1. *Sejam V e W espaços vetoriais de dimensão finita, e $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V. Então, dados n vetores $w_1, \dots, w_n$ em W, existe uma única transformação linear $T : V \to W$, tal que*

$$T(v_i) = w_i \ \ \forall i = 1, \dots, n. \tag{3.2}$$

Demonstração. Provemos, inicialmente, a existência de T. Dado $u \in V$, sejam $x_1, \ldots, x_n$ suas coordenadas com respeito à base $\mathcal{B}$, isto é,

$$u = (x_1, \ldots, x_n)_{\mathcal{B}} = x_1 v_1 + \cdots + x_n v_n .$$

Defina, então, $T : V \to W$ por

$$T(u) = T(x_1, \ldots, x_n)_{\mathcal{B}} = \sum_{i=1}^{n} x_i w_i .$$

É imediato que $T(v_i) = w_i \, \forall i = 1, \ldots, n$. Além disso, para quaisquer $v = (y_1, \ldots, y_n)_{\mathcal{B}} \in V$ e $\lambda \in \mathbb{R}$, tem-se

$$\begin{aligned} T(u + \lambda v) &= T(x_1 + \lambda y_1, \ldots, x_n + \lambda y_n)_{\mathcal{B}} \\ &= \sum_{i=1}^{n} (x_i + \lambda y_i) w_i = \sum_{i=1}^{n} x_i w_i + \lambda \sum_{i=1}^{n} \lambda y_i w_i \\ &= T(u) + \lambda T(v), \end{aligned}$$

donde T é linear.

Suponhamos, agora, que exista uma transformação linear $T' : V \to W$, tal que $T'(v_i) = w_i \, \forall i = 1, \ldots n$. Nesse caso, para todo $u = (x_1, \ldots, x_n)_{\mathcal{B}}$ in V, segue-se da linearidade de T' que

$$T'(u) = T'\left(\sum_{i=1}^{n} x_i v_i \right) = \sum_{i=1}^{n} x_i T'(v_i) = \sum_{i=1}^{n} x_i w_i = T(u),$$

o que prova a unicidade de T com respeito à igualdade (3.2) e conclui, dessa forma, a demonstração. ∎

Exemplo 3.6. Consideremos a base canônica de $\mathbb{R}^3$, $\mathcal{B} = \{e_1, e_2, e_3\}$, e determinemos a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$, tal que

$$T(e_1) = (-1, 2), \quad T(e_2) = (1, 1) \quad \text{e} \quad T(e_3) = (0, 1).$$

Temos que,

$$\begin{aligned} T(x, y, z) &= T(xe_1 + ye_2 + ze_3) = xT(e_1) + yT(e_2) + zT(e_3) \\ &= x(-1, 2) + y(1, 1) + z(0, 1), \end{aligned}$$

isto é,

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

3.2 Transformações Lineares e Matrizes

O teorema seguinte estabelece a relação fundamental entre transformações lineares e matrizes, a qual mencionamos à introdução deste capítulo. Do mesmo, infere-se que, fixando-se bases em espaços vetoriais de dimensão finita, V e W, existe uma correspondência biunívoca entre as transformações lineares de V em W e as matrizes reais $m \times n$, em que $n = \dim V$ e $m = \dim W$.

Teorema 3.1 (TRANSFORMAÇÕES LINEARES E MATRIZES). *Sejam V e W espaços vetoriais de dimensões n e m, respectivamente, e $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$ bases desses espaços. Então, dada $A \in \mathcal{M}(m,n)$, existe uma única transformação linear*

$$\begin{array}{rccc} T_A : & V & \to & W \\ & v & \mapsto & T_A(v) \end{array},$$

a qual cumpre a igualdade

$$[T_A(v)]_{\mathcal{B}'} = A[v]_{\mathcal{B}} \ \ \forall v \in V. \tag{3.3}$$

Reciprocamente, para toda transformação linear $T : V \to W$, existe uma única matriz $A \in \mathcal{M}(m,n)$, tal que $T = T_A$.

Demonstração. Uma vez que as relações

$$v \in V \leftrightarrow [v]_{\mathcal{B}} \in \mathcal{M}(n,1) \quad \text{e} \quad w \in W \leftrightarrow [w]_{\mathcal{B}'} \in \mathcal{M}(m,1)$$

são biunívocas, temos que a igualdade (3.3), na verdade, define a aplicação T_A , isto é, para cada $v \in V$, existe um único vetor $T_A(v) \in W,$ para o qual a igualdade (3.3) se cumpre.

Resta-nos, pois, provar que T_A é linear. Para tanto, tomemos $u, v \in V$ e $\lambda \in \mathbb{R}$, e observermos que

$$\begin{aligned} [T_A(u+\lambda v)]_{\mathcal{B}'} &= A[u+\lambda v]_{\mathcal{B}} = A([u]_{\mathcal{B}} + [\lambda v]_{\mathcal{B}}) = A[u]_{\mathcal{B}} + \lambda A[v]_{\mathcal{B}} \\ &= [T_A(u)]_{\mathcal{B}'} + \lambda [T_A(v)]_{\mathcal{B}'} = [T_A(u) + \lambda T_A(v)]_{\mathcal{B}'} \,. \end{aligned}$$

Dessa forma,

$$T_A(u+\lambda v) = T_A(u) + \lambda T_A(v) \quad \forall u, v \in V,\ \lambda \in \mathbb{R},$$

donde T_A é linear.

A fim de verificar a unicidade de T_A com respeito à igualdade (3.3), tomemos uma transformação linear $T : V \to W$, a qual cumpre

$$[T(v)]_{\mathcal{B}'} = A[v]_{\mathcal{B}} \ \forall v \in V.$$

Nesse caso, para todo $v \in V$, temos que $[T(v)]_{\mathcal{B}'} = [T_A(v)]_{\mathcal{B}'}$ e, portanto, $T(v) = T_A(v)\,\forall v \in V$, isto é, $T = T_A$.

Reciprocamente, suponhamos que $T : V \to W$ seja uma transformação linear, escrevamos $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$, $\mathcal{B}' = \{w_1, \dots, w_m\} \subset W$, e consideremos a matriz $A = (a_{ij}) \in \mathcal{M}(m,n)$, cuja j-ésima coluna é formada pelas coordenadas do vetor Tv_j com respeito à base $\mathcal{B}'$. Mais precisamente, as entradas a_{ij} são definidas pela igualdade

$$Tv_j = \sum_{i=1}^{m} a_{ij}w_i = a_{1j}w_1 + \cdots a_{mj}w_m\,, \ \ 1 \leq j \leq n. \tag{3.4}$$

Assim, pelas considerações da Observação 2.1 e pela linearidade de T, para um dado $v = x_1v_1 + \cdots + x_nv_n \in V$, tem-se

$$A[v]_{\mathcal{B}} = x_1[T(v_1)]_{\mathcal{B}'} + \cdots + x_n[T(v_n)]_{\mathcal{B}'} = [T(x_1v_1 + \cdots + x_nv_n)]_{\mathcal{B}'} = [T(v)]_{\mathcal{B}'}\,.$$

Daí, e da unicidade de T_A com respeito à igualdade (3.3), segue-se que $T = T_A$.

Finalmente, suponhamos que exista $B \in \mathcal{M}(m,n)$, tal que $T = T_B$. Então, $T_A = T_B$, donde, para todo $v \in V$, tem-se $A[v]_{\mathcal{B}} = B[v]_{\mathcal{B}}$. Fazendo-se $v = v_j$, segue-se dessa última igualdade que as colunas correspondentes de A e B coincidem e, portanto, que $A = B$. Isso conclui a demonstração. ■

Nas condições do teorema acima, dizemos que A é a *matriz de $T = T_A$ com respeito às bases $\mathcal{B}$ e $\mathcal{B}'$*, e a indicamos por $[T]_{\mathcal{B}}^{\mathcal{B}'}$. Sendo assim, a igualdade (3.3) assume a seguinte forma

$$[T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}}\,.$$

No caso particular em que $V = W$ e $\mathcal{B} = \mathcal{B}'$, escrevemos $[T]_{\mathcal{B}}$, ao invés de $[T]_{\mathcal{B}}^{\mathcal{B}}$. Além disso, omitiremos a indicação das bases $\mathcal{B}$ e $\mathcal{B}'$, isto é, escreveremos simplesmente $[T]$, quando V e W forem espaços euclidianos, e $\mathcal{B}$ e $\mathcal{B}'$ suas respectivas bases canônicas.

Observação 3.1. Convém mencionar que, com a notação acima, em coordenadas com respeito à base $\mathcal{B}'$, $T(v_1), \ldots, T(v_n)$ *são justamente os vetores coluna de* $[T]_{\mathcal{B}}^{\mathcal{B}'}$.

Exemplo 3.7. Tomemos, com a notação do Teorema 3.1, $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, e suponhamos que $\mathcal{B} \subset \mathbb{R}^2$ e $\mathcal{B}' \subset \mathbb{R}^3$ sejam suas respectivas bases canônicas. Nessas condições, a transformação $T : \mathbb{R}^2 \to \mathbb{R}^3$, em que

$$[T] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix},$$

cumpre, para $v = (x, y) \in \mathbb{R}^2$,

$$[T(v)] = [T][v] = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x - y \\ x + y \\ y \end{bmatrix},$$

donde T é definida por

$$T(x, y) = (x - y, x + y, y).$$

Note que T é justamente a transformação linear do Exemplo 3.3.

Exemplo 3.8. Dada a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$,

$$T(x, y, z) = (x + \pi y + ez, \sqrt{2}x - \log 3y + \sqrt{3}z),$$

obtenhamos sua matriz $[T] = (a_{ij})_{2\times 3}$ com respeito às bases canônicas de $\mathbb{R}^3$ e $\mathbb{R}^2$. Para tanto, observemos que, pela igualdade (3.4), basta obtermos as imagens dos vetores da base canônica de $\mathbb{R}^3$, pois os mesmos determinarão as colunas de $[T]$. Assim, uma vez que

$$T(1, 0, 0) = (1, \sqrt{2}), \quad T(0, 1, 0) = (\pi, -\log 3), \quad T(0, 0, 1) = (e, \sqrt{3}),$$

temos que $[T] = \begin{bmatrix} 1 & \pi & e \\ \sqrt{2} & -\log 3 & \sqrt{3} \end{bmatrix}$.

Tomando-se a notação introduzida no fim da Seção 2.3, infere-se do Teorema 3.1 que toda transformação linear $T : V \to W$ se exprime de maneira única como

$$T(x_1\,, \ldots, x_n)_{\mathcal{B}} = (a_{11}x_1 + \cdots + a_{1n}x_n,\ \ldots\ , a_{m1}x_1 + \cdots + a_{mn}x_n)_{\mathcal{B}'},$$

em que $[T]_{\mathcal{B}}^{\mathcal{B}'} = (a_{ij})_{m\times n}$.

Exemplo 3.9. Tomemos as bases $\mathcal{B} = \{(1,0,1),(0,1,-1),(1,2,-3)\} \subset \mathbb{R}^3$ e $\mathcal{B}' = \{(1,0),(1,-1)\} \subset \mathbb{R}^2$, e consideremos a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$, cuja matriz com respeito a essas bases é:

$$[T]_{\mathcal{B}}^{\mathcal{B}'} = \begin{bmatrix} 2 & -1 & 2 \\ 4 & 1 & -2 \end{bmatrix}.$$

Assim, em coordenadas com respeito a $\mathcal{B}$ e $\mathcal{B}'$, T se escreve como:

$$T(x,y,z)_{\mathcal{B}} = (2x - y + 2z, 4x + y - 2z)_{\mathcal{B}'}.$$

A fim de expressar T em coordenadas com respeito às bases canônicas de $\mathbb{R}^3$ e $\mathbb{R}^2$, podemos proceder como no Exemplo 2.12 e obter as seguintes relações:

- $(x,y,z) = \left(\frac{x+y+z}{2}, -x + 2y + z, \frac{x-y-z}{2}\right)_{\mathcal{B}}$.
- $(x,y)_{\mathcal{B}'} = (x+y, -y)$.

Logo, valem as igualdades:

$$\begin{aligned} T(x,y,z) &= T\left(\frac{x+y+z}{2}, -x+2y+z, \frac{x-y-z}{2}\right)_{\mathcal{B}} \\ &= (3x - 2y - z, 5y + 4z)_{\mathcal{B}'} \\ &= (3x + 3y + 3z, -5y - 4z). \end{aligned}$$

Em particular,

$$[T] = \begin{bmatrix} 3 & 3 & 3 \\ 0 & -5 & 4 \end{bmatrix}.$$

Exemplo 3.10. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ o operador linear, tal que

$$[T] = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix},$$

isto é, $T(x,y) = (x+2y, 4x+3y)$. Consideremos a base $\mathcal{B} = \{(1,-1),(1,2)\} \subset \mathbb{R}^2$ e determinemos a matriz de T com respeito a $\mathcal{B}$. Para tanto, como no exemplo anterior, buscamos inicialmente a relação entre as coordenadas com respeito à base canônica e as coordenadas com respeito à base $\mathcal{B}$, obtendo

$$(x,y)_{\mathcal{B}} = (x+y, -x+2y) \quad \text{e} \quad (x,y) = \left(\frac{2x-y}{3}, \frac{x+y}{3}\right)_{\mathcal{B}}. \tag{3.5}$$

Logo,

$$T(x,y)_{\mathcal{B}} = T(x+y, -x+2y) = (-x+5y, x+10y) = (-x, 5y)_{\mathcal{B}},$$

donde se infere que

$$[T]_{\mathcal{B}} = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}.$$

Note que traço $[T]$ = traço $[T]_{\mathcal{B}}$ (vide último parágrafo da próxima seção).

3.3 Mudança de Coordenadas

Nos dois últimos exemplos da seção anterior, constatamos que, em coordenadas, uma transformação linear $T : V \to W$ pode se expressar de formas distintas, a dependender das bases consideradas em V e W. No caso do operador do Exemplo 3.10, vê-se ainda que a mudança da base canônica de $\mathbb{R}^2$ para a base $\mathcal{B}$ tornou a expressão de T mais simples, pois sua matriz com respeito a $\mathcal{B}$ resultou ser diagonal.

Essas considerações nos levam a buscar um método efetivo de se obter as relações entre dois sistemas de coordenadas, quando os mesmos são relativos a bases distintas. Com esse propósito, tomemos bases $\mathcal{B}$, $\mathcal{B}'$ de um espaço vetorial V, e denotemos por I a aplicação identidade de V. Nessas condições, a matriz $[I]_{\mathcal{B}}^{\mathcal{B}'}$ é dita a *matriz de mudança* da base $\mathcal{B}$ para a base $\mathcal{B}'$. Note que essa nomenclatura deve-se ao fato de que, para todo $v \in V$,

$$[v]_{\mathcal{B}'} = [Iv]_{\mathcal{B}'} = [I]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}},$$

isto é, a multiplicação de $[I]_{\mathcal{B}}^{\mathcal{B}'}$ por $[v]_{\mathcal{B}}$ produz o efeito de mudar as coordenadas de v da base $\mathcal{B}$ para a base $\mathcal{B}'$.

Exemplo 3.11. Considere as bases $\mathcal{B} = \{(1,0),(0,1)\}$ e $\mathcal{B}' = \{(1,-1),(1,2)\}$ de $\mathbb{R}^2$. Pondo-se

$$(1,0) = x_1(1,-1) + y_1(1,2) \quad \text{e} \quad (0,1) = x_2(1,-1) + y_2(1,2),$$

obtém-se facilmente $x_1 = 2/3$, $y_1 = 1/3$ e $x_2 = -y_2 = -1/3$. Logo,

$$[I]_{\mathcal{B}}^{\mathcal{B}'} = \begin{bmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \end{bmatrix}.$$

Assim, dado $v = (x,y) \in \mathbb{R}^2$,

$$[v]_{\mathcal{B}'} = [I]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}} = \begin{bmatrix} 2/3 & -1/3 \\ 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2x-y \\ x+y \end{bmatrix},$$

isto é (compare com (3.5)),

$$v = (x,y) = \left(\frac{2x-y}{3}, \frac{x+y}{3}\right)_{\mathcal{B}'}.$$

Proposição 3.2 (SEMELHANÇA ENTRE AS MATRIZES DE UM OPERADOR)**.** *Sejam $\mathcal{B}$ e $\mathcal{B}'$ bases de um espaço vetorial V de dimensão finita. Então, para todo operador linear $T : V \to V$, fazendo-se $M = [I]_{\mathcal{B}}^{\mathcal{B}'}$, tem-se*

$$[T]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}M,$$

isto é, $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{B}'}$ são matrizes semelhantes.

Demonstração. Dado $v \in V$, temos que $M[v]_{\mathcal{B}} = [v]_{\mathcal{B}'}$ e, portanto, $M^{-1}[v]_{\mathcal{B}'} = [v]_{\mathcal{B}}$. Além disso, temos que $[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}$ e $[T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}'}[v]_{\mathcal{B}'}$. Logo,

$$(M^{-1}[T]_{\mathcal{B}'}M)[v]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}[v]_{\mathcal{B}'} = M^{-1}[T(v)]_{\mathcal{B}'} = [T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}},$$

donde se infere que $[T]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}M$, como desejado. ∎

Como sabemos, matrizes semelhantes têm traços iguais. Assim, a proposição acima nos permite definir o *traço* de um operador linear $T : V \to V$ como o traço da matriz $[T]_{\mathcal{B}}$, em que $\mathcal{B}$ é uma base arbitrária de V.

3.4 Núcleo e Imagem

Dada uma transformação linear $T : V \to W$, associamos a ela dois subespaços, um de V, outro de W, os quais designamos, respectivamente, por *núcleo* e *imagem* de T. Mostraremos que o comportamento de uma transformação linear, no que diz respeito a injetividade ou sobrejetividade, é relativo às propriedades desses subespaços.

Definição 3.2 (NÚCLEO – IMAGEM)**.** Seja $T : V \to W$ uma transformação linear entre dois espaços vetoriais V e W. O *núcleo* de T, denotado por $\operatorname{Nuc} T$, é o conjunto formado por todos os vetores de V cuja imagem, por T, é o vetor nulo de W. A *imagem* de T, denotada por $\operatorname{Im} T$, é o conjunto dos vetores de W que são imagem, por T, de (pelo menos) um vetor de V. Em símbolos,

$$\operatorname{Nuc} T = \{v \in V\,; T(v) = \mathbf{0}\} \quad \text{e} \quad \operatorname{Im} T = \{w \in W\,; w = T(v),\ v \in V\}.$$

Dado um conjunto arbitrário $\Omega \subset V$, denota-se por $T(\Omega)$ o conjunto formado por todos os elementos de W que são imagens de elementos de Ω, isto é, $T(\Omega) = \{w \in W; w = T(v)\,, v \in \Omega\}$. Em particular, $\operatorname{Im} T = T(V)$.

Proposição 3.3. *Dada uma transformação linear $T : V \to W$, os conjuntos $\operatorname{Nuc} T$ e $\operatorname{Im} T$ são subespaços de V e W, respectivamente.*

Demonstração. Sendo T uma aplicação linear, tem-se $T(\mathbf{0}) = \mathbf{0}$. Logo, $\operatorname{Nuc} T$ e $\operatorname{Im} T$ contêm os vetores nulos dos respectivos espaços em que estão contidos. Além disso, dados $u, v \in \operatorname{Nuc} T$ e $\lambda \in \mathbb{R}$, tem-se $T(u+\lambda v) = T(u)+\lambda T(v) = \mathbf{0}$, donde $u + \lambda v \in \operatorname{Nuc} T$. Logo, $\operatorname{Nuc} T$ é um subespaço de V.

Consideremos agora $z, w \in \operatorname{Im} T$ e $\lambda \in \mathbb{R}$. Existem, então, $u, v \in V$ satisfazendo $T(u) = z$ e $T(v) = w$. Assim, $T(u+\lambda v) = T(u)+\lambda T(v) = z+\lambda w$, implicando que $z+\lambda w \in \operatorname{Im} T$ e, portanto, que $\operatorname{Im} T$ é um subespaço de W. ■

Definição 3.3 (POSTO – NULIDADE)**.** Dada uma transformação linear T, chama-se de *posto* de T a dimensão de sua imagem, e de *nulidade* de T a dimensão de seu núcleo.

A proposição seguinte é de grande valia na determinação da imagem de uma transformação linear.

Proposição 3.4. *Sejam V um espaço vetorial e $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de V. Então, se $T : V \to W$ é uma transformação linear, tem-se*

$$\operatorname{Im} T = \operatorname{Ger}\{T(v_1), \dots, T(v_n)\}.$$

Demonstração. Seja $w = T(v) \in \operatorname{Im} T \subset W$, em que $v = (x_1, \dots, x_n)_{\mathcal{B}} \in V$. Sendo assim, tem-se

$$w = T(v) = x_1 T(v_1) + \cdots + x_n T(v_n) \in \operatorname{Ger}\{T(v_1), \dots, T(v_n)\},$$

ou seja, $\operatorname{Im} T \subset \operatorname{Ger}\{T(v_1), \dots, T(v_n)\}$.

Temos que $\operatorname{Im} T$ é um subespaço de W e que $\{T(v_1), \dots, T(v_n)\} \subset \operatorname{Im} T$. Logo, $\operatorname{Ger}\{T(v_1), \dots, T(v_n)\} \subset \operatorname{Im} T$ e, portanto,

$$\operatorname{Im} T = \operatorname{Ger}\{T(v_1), \dots, T(v_n)\},$$

como desejávamos provar. ■

Exemplo 3.12. Determinemos o núcleo e a imagem da transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$, dada por $T(x, y, z) = (x + y + 2z, x - y - 3z)$. Para tanto, observemos que a igualdade $T(x, y, z) = \mathbf{0}$ equivale ao sistema linear

$$\left\{ \begin{array}{ccccccc} x & + & y & + & 2z & = & 0 \\ x & - & y & - & 3z & = & 0 \end{array} \right. .$$

Resolvendo-o, concluímos que

$$\operatorname{Nuc} T = \{(x, -5x, 2x) \in \mathbb{R}^3;\, x \in \mathbb{R}\} = \operatorname{Ger}\{(1, -5, 2)\} \subset \mathbb{R}^3.$$

Já a imagem de T, pela Proposicao 3.4, é o espaço gerado pelos vetores $T(1, 0, 0) = (1, 1)$, $T(0, 1, 0) = (1, -1)$ e $T(0, 0, 1) = (2, -3)$, donde se conclui que $\operatorname{Im} T = \mathbb{R}^2$. Em particular, T é sobrejetiva.

Exemplo 3.13. Tomemos agora a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^3$, definida por $T(x, y) = (x - 4y, 3x + y, x - 7y)$. Procedendo-se de forma análoga à do exemplo anterior, conclui-se que $\operatorname{Nuc} T = \{\mathbf{0}\}$. Temos ainda que $T(1, 0) = (1, 3, 1)$ e $T(0, 1) = (-4, 1, -7)$. Logo, pela Proposição 3.4, tem-se $\operatorname{Im} T = \operatorname{Ger}\{(1, 3, 1), (-4, 1, -7)\}$. Em particular, o posto de T é 2, pois $(1, 3, 1)$ e $(-4, 1, -7)$ são, claramente, LI.

Exemplo 3.14. Seja $f : \mathcal{M}(n, n) \to \mathbb{R}$, $n > 1$, o funcional linear traço, introduzido no Exemplo 3.5. O núcleo de f é, dessa forma, composto pelas matrizes de $\mathcal{M}(n, n)$ que têm traço nulo. Portanto, $\dim(\operatorname{Nuc} f) = n - 1$ (vide Exemplo 2.14). Temos, também, que f é sobrejetivo, pois, dado $a \in \mathbb{R}$, $a \neq 0$, a matriz $A \in \mathcal{M}(n, n)$ cujas entradas a_{ij} são todas nulas, exceto por $a_{11} = a$, é tal que $f(A) = a$.

Exemplo 3.15. Seja $\mathcal{P}_n$ o espaço dos polinômios de variável t e grau menor que, ou igual a, n. Consideremos a transformação linear

$$\begin{array}{rccc} T: & \mathcal{P}_2 & \to & \mathcal{P}_1 \\ & at^2 + bt + c & \mapsto & 2at + b \end{array},$$

que associa a cada polinômio $p(t) \in \mathcal{P}_2$ sua derivada $p'(t) \in \mathcal{P}_1$ (verifique que T é, de fato, linear). Assim, o núcleo de T é formado pelos polinômios $p(t)$ cuja derivada é o polinômio nulo, isto é, aqueles para os quais $a = b = 0$. Logo, $\operatorname{Nuc} T$ é o conjunto dos polinômios que se escrevem como $p(t) = c$. Em particular, $\dim(\operatorname{Nuc} T) = 1$.

Observe-se também que T é sobrejetiva, pois, dado $q = q(t) = \lambda t + \mu \in \mathcal{P}_1$, o polinômio $p = p(t) = \frac{\lambda}{2}t^2 + \mu t \in \mathcal{P}_2$ satisfaz $T(p) = q$.

Proposição 3.5 (INJETIVIDADE DE TRANSFORMAÇÕES LINEARES). *São equivalentes as seguintes afirmações a respeito de uma transformação linear $T : V \to W$:*

i) *T é injetiva;*

ii) $\operatorname{Nuc} T = \{\mathbf{0}\}$*;*

iii) *T leva vetores* LI *de V em vetores* LI *de W.*

No caso em que V tem dimensão finita, cada uma das afirmações acima é equivalente a:

iv) *Existe uma base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V, tal que $\{T(v_1), \dots, T(v_n)\}$ é* LI *em W.*

Demonstração. (i) $\Rightarrow$ (ii). Seja $v \in \operatorname{Nuc} T$. Então $T(v) = \mathbf{0} = T(\mathbf{0})$. Uma vez que estamos supondo que T é injetiva, devemos ter $v = \mathbf{0}$, isto é, $\operatorname{Nuc} T = \{\mathbf{0}\}$.

(ii) $\Rightarrow$ (iii). Tomemos vetores $v_1, \cdots, v_n \in V$ e suponhamos que eles sejam LI. Devemos provar que $T(v_1), \dots, T(v_n)$ são LI em W. Para tanto, tomemos uma combinação linear nula desses vetores, $\lambda_1 T(v_1) + \cdots + \lambda_n T(v_n) = 0$. Daí e da linearidade de T, tem-se $T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0$, donde $v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in \operatorname{Nuc} T$. Pela hipótese, devemos ter $v = 0$, o que nos dá $\lambda_1 = \cdots = \lambda_n = 0$, já que os vetores $v_1, \dots, v_n$ são LI. Logo, os vetores $T(v_1), \dots, T(v_n)$ são LI.

(iii) $\Rightarrow$ (i). Sejam $v_1,\ v_2 \in V$, tais que $T(v_1) = T(v_2)$. Nesse caso, $T(v_1 - v_2) = \mathbf{0}$. Se tivéssemos $v = v_1 - v_2 \neq \mathbf{0}$, teríamos que a imagem do conjunto LI, $\{v\} \subset V$, seria o conjunto LD, $\{\mathbf{0}\} \subset W$, contrariando a hipótese. Segue-se que $v_1 = v_2$ e, portanto, que T é injetiva.

Suponhamos agora que V tenha dimensão finita. Nesse caso, a implicação (iii) $\Rightarrow$ (iv) é imediata. Vejamos, então, que (iv)$\Rightarrow$ (ii). Para tanto, tomemos $v = (x_1, \dots, x_n)_{\mathcal{B}} \in \operatorname{Nuc} T \subset V$. Então,

$$\mathbf{0} = T(v) = T(x_1 v_1 + \cdots + x_n v_n) = x_1 T(v_1) + \cdots + x_n T(v_n).$$

Daí e da independência linear de $T(v_1), \dots, T(v_n)$, segue-se que $x_1 = \cdots = x_n = 0$, o que nos dá $v = \mathbf{0}$. Isso prova que $\operatorname{Nuc} T = \{\mathbf{0}\}$ e conclui, dessa forma, a demonstração. ■

Infere-se da Proposição 3.5 que a transformação linear do Exemplo 3.13 é injetiva, enquanto as dos Exemplos 3.12, 3.14 e 3.15 não o são. Ainda com respeito a esses exemplos, observa-se que, em todos eles, a dimensão do domínio da transformação linear considerada é a soma de sua nulidade com seu posto. Provemos, agora, que esse fato é geral.

Teorema 3.2 (DO NÚCLEO E DA IMAGEM)**.** *Seja $T : V \to W$ uma transformação linear. Então, se V tem dimensão finita, vale a igualdade:*

$$\dim V = \dim(\operatorname{Nuc} T) + \dim(\operatorname{Im} T). \tag{3.6}$$

Demonstração. Suponhamos, inicialmente, que $\operatorname{Nuc} T = \{0\}$, e tomemos uma base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de V. Pelas Proposições 3.4 e 3.5, temos que o conjunto $\mathcal{B}' = \{T(v_1), \ldots, T(v_n)\}$ gera $\operatorname{Im} T$ e é LI. Logo, $\mathcal{B}'$ é uma base de $\operatorname{Im} T$, donde a igualdade (3.6) se cumpre.

Suponhamos agora que $\dim(\operatorname{Nuc} T) = k > 0$. Pela Proposição 2.10-(iv), existe uma base $\mathcal{B} = \{v_1, \cdots, v_k, v_{k+1}, \ldots, v_n\}$ de V, tal que seus primeiros k vetores, $v_1, \ldots, v_k$, formam uma base de $\operatorname{Nuc} T$. Além disso, pondo-se $V_0 = \operatorname{Ger}\{v_{k+1}, \ldots, v_n\}$, tem-se $V_0 \cap \operatorname{Nuc} T = \{\mathbf{0}\}$, pois $V = \operatorname{Nuc} T \oplus V_0$ (vide Exercício 21 – Cap. 2). Em particular, pela Proposição 3.5, $T|_{V_0} : V_0 \to W$ é injetiva. Porém, pela Proposição 3.4, o conjunto

$$\{T(v_1), \ldots, T(v_k), \ldots, T(v_{k+1}), \ldots, T(v_n)\}$$

gera $\operatorname{Im} T$. Uma vez que $T(v_1) = \cdots = T(v_k) = \mathbf{0}$, segue-se que

$$\mathcal{B}' = \{T(v_{k+1}), \cdots, T(v_n)\},$$

igualmente, gera $\operatorname{Im} T$. No entanto, $\mathcal{B}'$ é a imagem, pela aplicação injetiva $T|_{V_0}$, do conjunto LI $\{v_{k+1}, \ldots, v_n\}$. Ainda pela Proposição 3.5, $\mathcal{B}'$ é LI, donde $\mathcal{B}'$ é uma base de $\operatorname{Im} T$. Dessa forma,

$$\dim V = n = k + (n - k) = \dim(\operatorname{Nuc} T) + \dim(\operatorname{Im} T),$$

como queríamos demonstrar. ■

Corolário 3.1. *Sejam V um espaço vetorial de dimensão finita e $T : V \to W$ uma transformação linear. São válidas as seguintes asserções:*

i) *Se T é bijetiva, então W tem dimensão finita e* $\dim V = \dim W$;

ii) *Se* $\dim V = \dim W$, *então, uma condição necessária e suficiente para que T seja bijetiva, é a de que seja injetiva ou sobrejetiva.*

Demonstração. Suponha que T seja bijetiva. Nesse caso, tem-se $W = \operatorname{Im} T$ e $\operatorname{Nuc} T = \{\mathbf{0}\}$. Em particular, W tem dimensão finita. Além disso, por (3.6), $\dim V = \dim(\operatorname{Nuc} T) + \dim(\operatorname{Im} T) = \dim W,$ o que prova (i).

Suponhamos agora que $\dim V = \dim W$. Se T for bijetiva, por definição, será injetiva e sobrejetiva. Reciprocamente, se T for injetiva, teremos $\dim(\operatorname{Nuc} T) = 0$ e, então, por (3.6), $\dim W = \dim V = \dim(\operatorname{Im} T)$, donde T será sobrejetiva e, portanto, bijetiva.

Finalmente, supondo-se que T seja sobrejetiva, segue-se de (3.6) que

$$\dim(\operatorname{Nuc} T) = \dim V - \dim(\operatorname{Im} T) = \dim V - \dim W = 0,$$

isto é, T é injetiva. Logo, bijetiva. Isso prova (ii) e conclui a demonstração. ■

Segue-se do Corolário 3.1-(ii) que o operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$, tal que

$$[T] = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix},$$

é bijetivo. Para verificarmos isso, basta observarmos que os vetores coluna de $[T]$, os quais geram $\operatorname{Im} T$, são LI. Logo, eles constituem uma base de $\mathbb{R}^3$, donde se infere que $\operatorname{Im} T = \mathbb{R}^3$. Assim, T é sobrejetivo. Logo, pelo Corolário 3.1-(ii), T é bijetivo.

Exemplo 3.16 (BIJETIVIDADE DA TRANSPOSIÇÃO DE MATRIZES)**.** Seja

$$T : \mathcal{M}(m, n) \to \mathcal{M}(n, m)$$

a trasposição de matrizes, isto é, $T(X) = X^*$, $X \in \mathcal{M}(m, n)$. Temos que T é linear (vide igualdade (1.1)). Além disso, $T(X) = \mathbf{0}$ se, e só se, $X^* = \mathbf{0}$ ou, equivalentemente, $X = \mathbf{0}$. Logo, $\operatorname{Nuc} T = \{\mathbf{0}\}$, donde T é injetiva. Uma vez que $\dim \mathcal{M}(m, n) = \dim \mathcal{M}(n, m) = mn$, segue-se do Corolário 3.1-(ii) que T é bijetiva.

3.5 Operações com Transformações Lineares

Introduziremos agora as operações de adição e produto de transformações lineares, e constataremos que, quando os espaços envolvidos são de dimensão finita, essas operações determinam uma álgebra para transformações lineares, a qual é idêntica àquela das matrizes. Com esse propósito, consideremos espaços vetoriais V e W, e denotemos por $\mathscr{L}(V,W)$ o conjunto formado por todas as transformações lineares de V em W, isto é,

$$\mathscr{L}(V,W) = \{T : V \to W\,;\ T \text{ é linear}\}.$$

Dadas T_1, $T_2 \in \mathscr{L}(V,W)$ e $\lambda \in \mathbb{R}$, definimos a *adição* $T_1 + T_2 : V \to W$ e o *produto por escalar* $\lambda T_1 : V \to W$ por

$$(T_1+T_2)(v) = T_1(v) + T_2(v) \quad \text{e} \quad (\lambda T_1)(v) = \lambda T_1(v).$$

É imediato que, para quaisquer $T_1, T_2 \in \mathscr{L}(V,W)$ e $\lambda \in \mathbb{R}$, tem-se $T_1+T_2 \in \mathscr{L}(V,W)$ e $\lambda T_1 \in \mathscr{L}(V,W)$. É também de fácil verificação que essas operações têm as propriedades fundamentais (comutatividade, associatividade, etc.) que conferem à $\mathscr{L}(V,W)$ uma estrutura de espaço vetorial. Ademais, tomando-se bases $\mathcal{B} \subset V$ e $\mathcal{B}' \subset W$, para quaisquer $T_1, T_2 \in \mathscr{L}(V,W)$, $v \in V$ e $\lambda \in \mathbb{R}$, tem-se

$$\begin{aligned}
[(T_1+\lambda T_2)(v)]_{\mathcal{B}'} &= [T_1(v) + \lambda T_2(v)]_{\mathcal{B}'} = [T_1(v)]_{\mathcal{B}'} + \lambda [T_2(v)]_{\mathcal{B}'} \\
&= [T_1]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}} + \lambda [T_2]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}} = ([T_1]_{\mathcal{B}}^{\mathcal{B}'} + \lambda [T_2]_{\mathcal{B}}^{\mathcal{B}'})[v]_{\mathcal{B}}\,.
\end{aligned}$$

Segue-se, então, dessa última igualdade, juntamente com a unicidade da matriz de uma transformação linear com respeito a bases predeterminadas, o resultado seguinte.

Proposição 3.6 (A MATRIZ DA SOMA É A SOMA DAS MATRIZES)**.** *Sejam* T_1, $T_2 \in \mathscr{L}(V,W)$, *em que* V *e* W *são espaços vetoriais de dimensões* n *e* m, *respectivamente. Nessas condições, para quaisquer bases* $\mathcal{B} \subset V$, $\mathcal{B}' \subset W$, *e todo* $\lambda \in \mathbb{R}$, *tem-se*

$$[T_1+T_2]_{\mathcal{B}}^{\mathcal{B}'} = [T_1]_{\mathcal{B}}^{\mathcal{B}'} + \lambda [T_2]_{\mathcal{B}}^{\mathcal{B}'}\,.$$

Em particular, a aplicação

$$\begin{array}{ccc} \mathscr{L}(V,W) & \to & \mathcal{M}(m,n) \\ T & \mapsto & [T]_{\mathcal{B}}^{\mathcal{B}'} \end{array}$$

é uma transformação linear.

Vejamos agora o comportamento das transformações lineares com respeito à composição de aplicações. Para tanto, tomemos espaços vetoriais U, V, W, e aplicações lineares $T_1 : U \to V$, $T_2 : V \to W$. Verifiquemos, então, que a composta $T_2 \circ T_1 : U \to W$ é linear. Com efeito, dados $u, v \in U$ e $\lambda \in \mathbb{R}$, tem-se

$$\begin{aligned} (T_2 \circ T_1)(u + \lambda v) &= T_2(T_1(u + \lambda v)) = T_2(T_1(u) + \lambda T_1(v)) \\ &= T_2(T_1(u)) + \lambda T_2(T_1(v)) = (T_2 \circ T_1)(u) + \lambda (T_2 \circ T_1)(v), \end{aligned}$$

donde $T_2 \circ T_1$ é linear.

Diz-se, então, que $T = T_2 \circ T_1 : U \to W$ é o *produto* de T_2 e T_1 e escreve-se $T = T_2 T_1$. Dado um operador linear $T : V \to V$, o produto $TT \dots T$, com k fatores, denota-se também por T^k. Essa nomenclatura justifica-se pelo resultado seguinte que, grosso modo, nos diz que a matriz do produto de transformações lineares é o produto das matrizes correspondentes.

Proposição 3.7 (A MATRIZ DO PRODUTO É O PRODUTO DAS MATRIZES)**.** *Sejam U, V e W espaços vetoriais de dimensão finita, e $T_1 : U \to V$, $T_2 : V \to W$ aplicações lineares. Então, para quaisquer bases $\mathcal{A} \subset U$, $\mathcal{B} \subset V$ e $\mathcal{C} \subset W$, vale a seguinte igualdade:*

$$[T_2 T_1]_{\mathcal{A}}^{\mathcal{C}} = [T_2]_{\mathcal{B}}^{\mathcal{C}} [T_1]_{\mathcal{A}}^{\mathcal{B}}. \tag{3.7}$$

Demonstração. Dado $u \in U$, considere os vetores

$$v = T_1(u) \in V \quad \text{e} \quad w = T_2(v) = (T_2 T_1)(u) \in W.$$

Nessas condições, temos

$$[(T_2 T_1)u]_{\mathcal{C}} = [T_2(v)]_{\mathcal{C}} = [T_2]_{\mathcal{B}}^{\mathcal{C}} [v]_{\mathcal{B}} = [T_2]_{\mathcal{B}}^{\mathcal{C}} [T_1(u)]_{\mathcal{B}} = ([T_2]_{\mathcal{B}}^{\mathcal{C}} [T_1]_{\mathcal{A}}^{\mathcal{B}}) [u]_{\mathcal{A}} \,,$$

o que implica na igualdade (3.7) ■

Uma vez que transformações lineares entre espaços de dimensão finita são univocamente associadas a matrizes, segue-se das Proposições 3.6 e 3.7 que a álgebra das transformações lineares tem as mesmas propriedades da álgebra das matrizes. Mais precisamente, se T_1, T_2, T_3 são transformações lineares entre espaços de dimensão finita, valem, sempre que os produtos e adições mencionados estejam definidos, as seguintes igualdades:

i) $T_1\mathbf{0} = \mathbf{0}T_1 = \mathbf{0}$;

ii) $T_1 I = I T_1 = T_1$, em que I é a identidade;

iii) $T_1(T_2 + T_3) = T_1T_2 + T_1T_3$;

iv) $(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$;

v) $(T_1T_2)T_3 = T_1(T_2T_3)$;

vi) $(\lambda T_1)T_2 = T_1(\lambda T_2) = \lambda(T_1T_2)\ \forall \lambda \in \mathbb{R}$.

Convém mencionar que, se $T_1, T_2 : V \to V$ são operadores distintos, tem-se, em geral, $T_1T_2 \neq T_2T_1$, isto é, assim como o produto de matrizes, *o produto de operadores lineares não é comutativo.* Quando ocorre $T_1T_2 = T_2T_1$, dizemos que T_1 e T_2 *comutam.* Segue-se, portanto, da Proposição 3.7, que, para qualquer base $\mathcal{B}$ de V, T_1 e T_2 comutam se, e somente se, as matrizes $[T_1]_{\mathcal{B}}$ e $[T_2]_{\mathcal{B}}$ comutam.

Exemplo 3.17. Sejam $T_1, T_2, T_3 : \mathbb{R}^2 \to \mathbb{R}^2$ os operadores lineares dados por

$$T_1(x,y) = (-y, x), \quad T_2(x,y) = (-x, y) \quad \text{e} \quad T_3(x,y) = (2x, y).$$

Temos que suas respectivas matrizes com respeito à base canônica de $\mathbb{R}^2$ são

$$[T_1] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad [T_2] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{e} \quad [T_3] = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Daí, verifica-se facilmente que $[T_1]$ não comuta com $[T_2]$ ou $[T_3]$, e que $[T_2]$ e $[T_3]$ comutam. Logo, valem as mesmas afirmações para as transformações lineares correspondentes.

3.6 Isomorfismos

Estudaremos agora as transformações lineares bijetivas, chamadas de *isomorfismos*. Os isomorfismos, dentre outras características, são importantes por determinarem uma relação de equivalência na classe dos espaços vetoriais, conforme constaremos.

Proposição 3.8 (LINEARIDADE DA INVERSA)**.** *Seja $T : V \to W$ uma aplicação linear bijetiva. Então, a inversa de T, $T^{-1} : W \to V$, é linear.*

Demonstração. Dados $w_1, w_2 \in W$ e $\lambda \in \mathbb{R}$, sejam $v_1 = T^{-1}(w_1)$ e $v_2 = T^{-1}(w_2)$. Uma vez que T é linear, tem-se $T(v_1 + \lambda v_2) = T(v_1) + \lambda T(v_2) = w_1 + \lambda w_2$. Logo, $T^{-1}(w_1 + \lambda w_2) = v_1 + \lambda v_2 = T^{-1}(w_1) + \lambda T^{-1}(w_2)$, donde T^{-1} é linear. ■

Uma aplicação linear bijetiva $T : V \to W$, como mencionamos acima, é dita um *isomorfismo*. Nesse caso, pela Proposição 3.8, $T^{-1} : W \to V$ é também um isomorfismo. Nessas condições, podemos dizer que os espaços V e W são *isomorfos*, e indicar essa relação por $V \sim W$. Assim, $\sim$ é simétrica. Além disso, observando-se que a aplicação identidade de um espaço vetorial é um isomorfismo, infere-se que $\sim$ é reflexiva. Por fim, uma vez que uma composição de bijeções é também uma bijeção, conclui-se que $\sim$ é também transitiva.

Em suma, vale o resultado seguinte.

Proposição 3.9. *A relação de isomorfismo é uma equivalência na classe dos espaços vetoriais.*

Exemplo 3.18. A aplicação

$$\begin{array}{rccl} T: & \mathcal{P}_2 & \to & \mathbb{R}^3 \\ & at^2 + bt + c & \mapsto & (a,b,c) \end{array}$$

é linear e bijetiva (verifique!), donde se conclui que o espaço dos polinômios de grau menor que, ou igual a, 2 é isomorfo a $\mathbb{R}^3$. Analogamente, o espaço $\mathcal{P}_n$ é isomorfo a $\mathbb{R}^n$.

Mais geralmente, vale o seguinte resultado.

Proposição 3.10. *Dois espaços vetoriais de dimensão finita são isomorfos se, e somente se, têm a mesma dimensão.*

Demonstração. Sejam V e W espaços vetoriais de dimensão finita. Se $V \sim W$, segue-se do Corolário 3.1-(i) que $\dim V = \dim W$.

Suponhamos, então, que $\dim V = \dim W = n$, e tomemos bases $\mathcal{B} = \{v_1, \ldots v_n\}$ e $\mathcal{B}' = \{w_1, \cdots w_n\}$, de V e W, respectivamente. Pela Proposição 3.1, existe uma única transformação linear $T : V \to W$ que satisfaz $T(v_i) = w_i \,\forall i \in \{1, 2, \cdots n\}$. Esse fato, juntamente com os resultados da Proposição 3.5 e do Corolário 3.1-(ii), nos leva à conclusão de que T é bijetiva e, portanto, um isomorfismo. ■

Corolário 3.2. *Todo espaço vetorial de dimensão n é isomorfo ao espaço $\mathbb{R}^n$.*

Proposição 3.11. *Sejam V e W espaços vetoriais de mesma dimensão finita. Então, dadas bases $\mathcal{B} \subset V$ e $\mathcal{B}' \subset W$, tem-se que uma transformação linear $T : V \to W$ é um isomorfismo se, e somente se, a matriz $[T]_{\mathcal{B}}^{\mathcal{B}'}$ é invertível. No caso afirmativo, vale a igualdade:*

$$([T]_{\mathcal{B}}^{\mathcal{B}'})^{-1} = [T^{-1}]_{\mathcal{B}'}^{\mathcal{B}}\,.$$

Demonstração. Suponhamos que $T : V \to W$ seja um isomorfismo. Nesse caso, designando-se por $I = W \to W$ a aplicação identidade de W, temos que $TT^{-1} = I$. Logo, pela Proposição 3.7,

$$[I]_{\mathcal{B}'} = [TT^{-1}]_{\mathcal{B}'} = [T]_{\mathcal{B}}^{\mathcal{B}'}[T^{-1}]_{\mathcal{B}'}^{\mathcal{B}}\,.$$

Uma vez que $[I]_{\mathcal{B}'}$ é matriz identidade de ordem $n = \dim V = \dim W$, segue-se que $[T]_{\mathcal{B}}^{\mathcal{B}'}$ é invertível e que $(\,[T]_{\mathcal{B}}^{\mathcal{B}'})^{-1} = [T^{-1}]_{\mathcal{B}'}^{\mathcal{B}}$.

Reciprocamente, supondo-se que $[T]_{\mathcal{B}}^{\mathcal{B}'}$ seja invertível, temos que o núcleo de T é trivial, donde T é injetiva e, portanto, bijetiva, já que V e W têm mesma dimensão. Com efeito, se $T(v) = 0$, então, $[\mathbf{0}]_{\mathcal{B}'} = [T(v)]_{\mathcal{B}'} = [T]_{\mathcal{B}}^{\mathcal{B}'}[v]_{\mathcal{B}}$. Logo, $[v]_{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}'})^{-1}[\mathbf{0}]_{\mathcal{B}'} = [\mathbf{0}]_{\mathcal{B}'}$, donde $v = 0$ e, portanto, $\operatorname{Nuc}(T) = \{\mathbf{0}\}$. ■

Corolário 3.3. *Sejam $T \in \mathscr{L}(V)$ um operador linear num espaço V de dimensão finita, e $\mathcal{B}$ uma base arbitrária de V. Então, T é invertível se, e somente se, $[T]_{\mathcal{B}}$ é invertível. No caso afirmativo, tem-se $[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}$.*

Decorre diretamente do Teorema 3.1 e da Proposição 3.6 o resultado seguinte.

Teorema 3.3. *Sejam V e W espaços vetoriais de dimensões n e m, respectivamente. Dadas bases $\mathcal{B} \subset V$ e $\mathcal{B}' \subset W$, a aplicação*

$$\begin{array}{ccc} \mathscr{L}(V,W) & \to & \mathcal{M}(m,n) \\ T & \mapsto & [T]_{\mathcal{B}}^{\mathcal{B}'} \end{array}$$

que associa a cada transformação linear $T \in \mathscr{L}(V,W)$ a sua matriz com respeito a $\mathcal{B}$ e $\mathcal{B}'$, é um isomorfismo. Em particular, $\dim \mathscr{L}(V,W) = mn$.

Se V e W são como no teorema acima, segue-se dele e do Corolário 3.2 que

$$\mathscr{L}(V,W) \sim \mathcal{M}(m,n) \sim \mathbb{R}^{mn}.$$

3.7 Aplicação aos Sistemas Lineares

Nesta seção que encerra o capítulo, faremos uma breve aplicação dos resultados obtidos ao estudo dos sistemas lineares. Em particular, estabeleceremos condições sobre as matrizes A e B de um sistema linear $AX = B$ para que o mesmo admita solução. Antes, consideremos o seguinte resultado.

Lema 3.1. *Suponha que o sistema linear $AX = B$ admita uma solução particular X_p. Então, $\overline{X}$ é solução de $AX = B$ se, e somente se, $\overline{X} = X_h + X_p$, em que X_h é uma solução do sistema homogêneo $AX = \mathbf{0}$. Em particular, X_p é uma solução única de $AX = B$ se, e somente se, $AX = \mathbf{0}$ admite apenas a solução trivial.*

Demonstração. Suponhamos que X_h seja uma solução de $AX = \mathbf{0}$. Então, para $\overline{X} = X_h + X_p$, teremos

$$A\overline{X} = A(X_h + X_p) = AX_h + AX_p = AX_p = B,$$

donde $\overline{X}$ será uma solução de $AX = B$.

Reciprocamente, suponha que $\overline{X}$ seja uma solução de $AX = B$. Nesse caso, fazendo-se $X_h = \overline{X} - X_p$, verifica-se igualmente que X_h é uma solução do sistema homogêneo $AX = \mathbf{0}$. ■

Conforme constatamos anteriormente, para toda matriz $A \in \mathcal{M}(m,n)$, o conjunto solução do sistema linear homogêneo $AX = 0$ é um subespaço vetorial de $\mathbb{R}^n$, o qual, naturalmente, corresponde ao núcleo da transformação linear associada, $T_A : \mathbb{R}^n \to \mathbb{R}^m$. Segue-se, portanto, do Lema 3.1, que o conjunto solução do sistema linear $AX = B$ é a *variedade afim*:

$$\operatorname{Nuc} T_A + v_p = \{v + v_p \in \mathbb{R}^n \,;\, v \in \operatorname{Nuc} T_A\},$$

em que $v_p \in \mathbb{R}^n$ satisfaz $A[v_p] = B$ (Fig. 3.1).

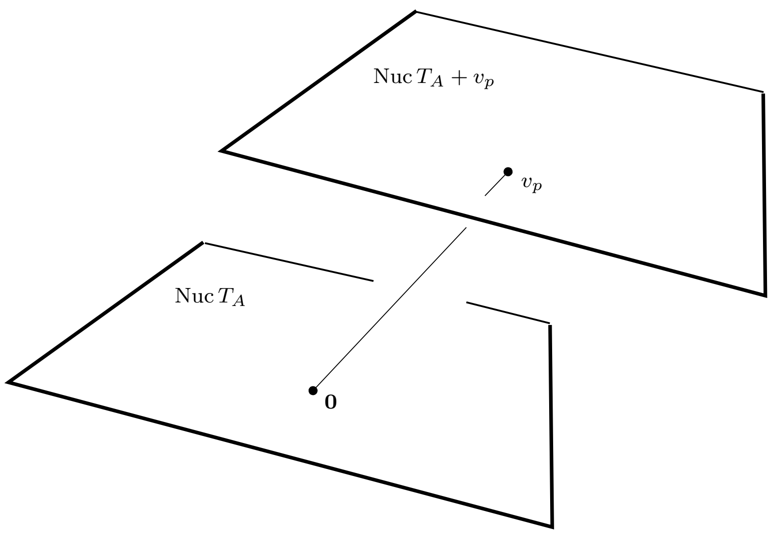


Figura 3.1: O conjunto solução de $AX = B$ como um subespaço afim de $\mathbb{R}^n$.

Teorema 3.4 (SOLUBILIDADE DE SISTEMAS LINEARES)**.** *Dadas matrizes* $A \in \mathcal{M}(m,n)$ *e* $B \in \mathcal{M}(m,1)$, *o sistema linear* $AX = B$ *admite solução se, e somente se,*

$$\operatorname{posto} A = \operatorname{posto}\,[A\,|B]. \tag{3.8}$$

No caso afirmativo, a solução é única se, e somente se, $\operatorname{posto} A = n$.

Demonstração. Sejam $w_1, \ldots, w_n \in \mathbb{R}^m$ os vetores coluna de A. Fazendo-se $B = [w]$, $w \in \mathbb{R}^m$, pelo Teorema 3.1, existe uma solução de $AX = B$ se, e somente se, $w \in \operatorname{Im} T_A$, o que ocorre se, e somente se, $w \in \operatorname{Ger}\{w_1, \dots, w_n\}$ (vide Proposição 3.4 e Observação 3.1). Uma vez que $[A\,|B] = [w_1 \cdots w_n\ w]$, temos que essa última condição é claramente equivalente à igualdade (3.8).

Pelo Lema 3.1, se existe uma solução particular X_p de $AX = B$, então essa solução é única se, e somente se, o sistema homogêneo associado admite apenas a solução trivial. Esse fato, porém, é equivalente à independência linear dos vetores coluna de A (vide Proposição 2.6), isto é, à igualdade posto $A = n$. ■

3.8 Exercícios

Seção 3.1

1. Verifique, em cada item, se a aplicação T é ou não uma transformação linear.

   a) $T : \mathbb{R}^2 \to \mathbb{R}^3,\ T(x, y) = (3x - 2y, x - y, -2x + 2y)$;

   b) $T : \mathbb{R}^4 \to \mathbb{R}^2,\ T(x, y, z, w) = (xy, zw)$;

   c) $T : \mathbb{R}^3 \to \mathbb{R}^3,\ T(x, y, z) = (x + y - 2z, x - 2y + z, x - z)$;

   d) $T : \mathbb{R}^2 \to \mathbb{R}^2,\ T(x, y) = (x + y - 2, 4x - y/2, x + 9y)$.

2. Sejam $T : V \to W$ uma transformação linear e $W_0 \subset W$ um subespaço de W. Prove que $T^{-1}(W_0) := \{v \in V\,;\, T(v) \in W_0\}$ é um subespaço de V.

3. Seja $T : V \to W$ uma transformação linear. Prove que, se o conjunto $\{T(v_1), \cdots, T(v_k)\}$ é LI em W, então $\{v_1, \cdots, v_k\}$ é LI em V.

4. Sejam V_0, $V_1 \subset V$ subespaços de V, e $T : V \to W$ uma transformação linear. Mostre que $T(V_0 + V_1) = T(V_0) + T(V_1)$.

5. Sejam $T : V \to V$ um operador linear e $\lambda \in \mathbb{R}$. Prove que:

   a) $V_\lambda = \{v \in V\,;\, T(v) = \lambda v\} \subset V$ é um subespaço de V;

   b) Se $v_1 \in V_{\lambda_1}$ e $v_2 \in V_{\lambda_2}$ são não nulos e $\lambda_1 \neq \lambda_2$, então v_1 e v_2 são LI.

Seção 3.2

6. Seja A uma matriz quadrada de ordem n que possui linhas ou colunas nulas. Prove que a transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$, tal que $A = [T]$, não é injetiva ou sobrejetiva.

7. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ um operador linear, tal que sua matriz $[T]$ é simétrica. Prove que existem reais λ_1 , λ_2 , e uma base de $\mathbb{R}^2$, $\mathcal{B} = \{v_1, v_2\}$, tais que $Tv_1 = \lambda_1 v_1$ e $Tv_2 = \lambda_2 v_2$.

8. Sejam A e B matrizes $n \times n$, e V um espaço vetorial de dimensão n. Prove que, se A e B são semelhantes, então existem bases $\mathcal{B}$, $\mathcal{B}'$ de V, e um operador linear $T : V \to V$, tais que $[T]_{\mathcal{B}} = A$ e $[T]_{\mathcal{B}'} = B$.

9. Seja $f : \mathbb{R}^n \to \mathbb{R}$ um funcional linear não identicamente nulo. Prove que f é sobrejetivo e que existe uma base $\mathcal{B}$ de $\mathbb{R}^n$, tal que
$$f(x_1, \cdots, x_n)_{\mathcal{B}} = ax_1\ ,\ \ a \in \mathbb{R}.$$
Exiba, então, a matriz de f com respeito à base $\mathcal{B}$.

10. Dada $A \in \mathcal{M}(2,2)$, seja $T : \mathcal{M}(2,2) \to \mathcal{M}(2,2)$ a aplicação definida por $T(X) = AX,\ X \in \mathcal{M}(2,2)$. Prove que T é linear e determine a matriz de T com respeito à base canônica de $\mathcal{M}(2,2)$.

Seção 3.3

11. Considere, em $\mathbb{R}^3$, as bases
$$\mathcal{B} = \{(1,1,0), (-1,1,1), (0,1,2)\} \ \text{ e } \ \mathcal{B}' = \{(2,1,1), (0,0,1), (1,-1,1)\},$$
e determine as matrizes de mudança de base, $[I]_{\mathcal{B}}^{\mathcal{B}'}$ e $[I]_{\mathcal{B}'}^{\mathcal{B}}$. Em seguida, determine as coordenadas de $v = (1,2,3)_{\mathcal{B}}$ com respeito a $\mathcal{B}'$, bem como as coordenadas de $w = (1,2,3)_{\mathcal{B}'}$ com respeito a $\mathcal{B}$.

12. Sejam $\mathcal{B}$ e $\mathcal{B}'$ como no exercício anterior, e $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear definido por $T(x,y,z) = (x - y + z, 5x + 2y - z, x + y + 3z)$. Determine as matrizes $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{B}'}$ e compare os seus respectivos traços.

Seção 3.4

13. Prove que existe uma transformação linear $T : \mathbb{R}^4 \to \mathbb{R}^4$ que satisfaz:

    a) $\operatorname{Nuc} T = \operatorname{Ger}\{(1,0,1,0),(-1,0,0,1)\}$;

    b) $\operatorname{Im} T = \operatorname{Ger}\{(1,-1,0,2),(0,1,-1,0)\}$.

14. Determine o núcleo e a imagem de cada transformação do Exercício 1 que é linear. Em seguida, investigue o comportamento destas transformações lineares no que diz respeito à injetividade e sobrejetividade.

15. Determine o núcleo e a imagem da transformação linear $T : \mathcal{P}_2 \to \mathcal{P}_3$ definida pela igualdade:

$$T(at^2 + bt + c) = \frac{a}{3}t^3 + \frac{b}{2}t^2 + ct.$$

16. Sejam $T_1\,, T_2 \in \mathscr{L}(V,V)$ operadores lineares de um espaço vetorial V de dimensão finita. Prove que:

    a) Existe $T \in \mathscr{L}(V,V)$ satisfazendo $T_1 = T_2T \;\Leftrightarrow\; \operatorname{Im} T_1 \subset \operatorname{Im} T_2\,$;

    b) Existe $T \in \mathscr{L}(V,V)$ satisfazendo $T_1 = TT_2 \;\Leftrightarrow\; \operatorname{Nuc} T_2 \subset \operatorname{Nuc} T_1\,$.

17. Sejam V e W espaços vetoriais de dimensão finita, e $W_0 \subset W$ um subespaço de W. Dada uma transformação linear $T : V \to W$, prove que $\dim(T^{-1}(W_0)) \geq \dim V - \dim W + \dim W_0\,$.

18. Sejam $T_1, T_2 : V \to V$ operadores lineares num espaço vetorial V de dimensão finita. Prove que:

    a) $\operatorname{Nuc} T_2 \subset \operatorname{Nuc}(T_1T_2)\,$;

    b) $T_2(\operatorname{Nuc}(T_1T_2)) \subset \operatorname{Nuc}(T_1)$;

    e conclua que $\dim(\operatorname{Nuc}(T_1T_2)) \leq \dim(\operatorname{Nuc} T_1) + \dim(\operatorname{Nuc} T_2)$.

Seção 3.5

19. Considere transformações lineares $T_1, T_2 \in \mathscr{L}(U, V)$ e $T_3 \in \mathscr{L}(V, W)$. Mostre, então, que:

    a) $\operatorname{posto}(T_1 + T_2) \leq \operatorname{posto}(T_1) \operatorname{posto}(T_2)$;

    b) $\operatorname{posto}(T_3 T_1) \leq \min\{\operatorname{posto}(T_1), \operatorname{posto}(T_3)\}$.

20. Seja $T : V \to V$ um operador linear. Suponha que exista $v \in V$ satisfazendo $T^k(v) = \mathbf{0}$ e $T^{k-1}(v) \neq \mathbf{0}$ para algum $k \in \mathbb{N}$. Prove que o conjunto $\{v, T(v), T^2(v), \dots, T^{k-1}(v)\}$ é LI.

21. Seja $T : V \to V$ uma transformação linear injetiva. Prove que, se $T_1, T_2 \in \mathscr{L}(V, V)$ satisfazem $TT_1 = TT_2$, então $T_1 = T_2$.

22. Seja $T : V \to V$ um operador linear. Prove que $T^2 = \mathbf{0}$ se, e somente se, $\operatorname{Im} T \subset \operatorname{Nuc} T$.

Seção 3.6

23. Considere um isomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$, e suponha que a, b, c, $d \in \mathbb{R}^2$ sejam vértices de uma paralelogramo. Mostre que vale o mesmo para $T(a)$, $T(b)$, $T(c)$ e $T(d)$.

24. Diz-se que um operador $T : V \to V$ é *nilpotente*, quando existe $k \in \mathbb{N}$, tal que $T^k = \mathbf{0}$. Mostre que, se $T : V \to V$ é nilpotente, então $T - I : V \to V$ é um isomorfismo, em que I denota o operador identidade de V.

25. Aplique o Corolário 3.3 para mostrar que, dadas matrizes A e B em $\mathcal{M}(n)$, se $AB = I$, então $BA = I$, isto é, A e B são invertíveis, sendo uma a inversa da outra.

26. Dadas matrizes A, $B \in \mathcal{M}(n, n)$, seja $T : \mathcal{M}(n, n) \to \mathcal{M}(n, n)$ a transformação linear definida por $T(X) = AX - XB$. Mostre que, se A e B são semelhantes, então T não é um isomorfismo.

27. Sejam V e W espaços vetoriais, tais que $\dim V = n$ e $\dim W = m$. Dadas bases $\mathcal{B}_1, \mathcal{B}_2 \subset V$, e $\mathcal{B}'_1, \mathcal{B}'_2 \subset W$, e uma transformação linear $T : V \to W$,

mostre que existem matrizes invertíveis, $M \in \mathcal{M}(m,m)$ e $N \in \mathcal{M}(n,n)$, tais que $M[T]_{\mathcal{B}_1}^{\mathcal{B}'_1} = [T]_{\mathcal{B}_2}^{\mathcal{B}'_2} N$.

4

Determinantes

Neste capítulo, estudaremos os determinantes. O determinante de uma matriz quadrada A é um número real, a ela associado, cujas características refletem certas propriedades de A. Estabelece-se, por exemplo, que A será invertível se, e somente se, seu determinante for não nulo.

Conforme constataremos nos capítulos posteriores, o determinante desempenha também um papel fundamental na teoria de operadores lineares em espaços de dimensão finita, bem como na conceituação de volume em espaços euclidianos. Como consequência, determinantes têm uma grande incidência em diversas teorias da matemática, especialmente em Análise e Geometria.

Assim, para cada $n \in \mathbb{N}$, definiremos uma *função determinante* em $\mathcal{M}(n)$. Verificaremos, então, suas propriedades fundamentais e, através das mesmas, obteremos alguns resultados importantes, dentre os quais, o supracitado critério de inversão de matrizes.

4.1 Determinantes em $\mathcal{M}(2)$

Consideremos uma matriz quadrada de segunda ordem, $A = (a_{ij})_{2\times 2}$, e observemos que, pela Proposição 2.5, um sistema linear do tipo

$$AX = B$$

admite uma única solução se, e somente se, as entradas a_{ij} de A satisfazem

$$a_{11}a_{22} - a_{12}a_{21} \neq 0. \tag{4.1}$$

Em particular, a condição (4.1) é equivalente à invertibilidade de A e, portanto, à invertibilidade do operador linear $T \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^2)$, tal que $[T] = A$.

Esses fatos sugerem o estudo da função $\mathfrak{D} : \mathcal{M}(2) \to \mathbb{R}$, definida por

$$\mathfrak{D}(A) = a_{11}a_{22} - a_{12}a_{21}\,, \quad A = (a_{ij})_{2\times 2}\,. \tag{4.2}$$

Para tanto, identificaremos $\mathcal{M}(2)$ com $\mathbb{R}^2 \times \mathbb{R}^2$ através da correspondência

$$A = (a_{ij}) \in \mathcal{M}(2) \quad \cong \quad (v_1, v_2) \in \mathbb{R}^2 \times \mathbb{R}^2,$$

em que $v_1 = (a_{11}, a_{21})$ e $v_2 = (a_{12}, a_{22})$ são os vetores coluna de A.

Verifica-se, então, que a função $\mathfrak{D}$ tem as seguintes propriedades, as quais são válidas para quaisquer $v_1\,, v_2\,, v_3 \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$:

i) $\mathfrak{D}(v_1 + \lambda v_2, v_3) = \mathfrak{D}(v_1, v_3) + \lambda\mathfrak{D}(v_2, v_3)$;

ii) $\mathfrak{D}(v_1, v_2 + \lambda v_3) = \mathfrak{D}(v_1, v_2) + \lambda\mathfrak{D}(v_1, v_3)$;

iii) $\mathfrak{D}(v_1, v_1) = 0$;

iv) $\mathfrak{D}(e_1, e_2) = 1$.

Assim, $\mathfrak{D}$ é *bilinear* (propriedades (i) e (ii)), *alternada* (propriedade (iii)) e vale 1 na matriz identidade de $\mathcal{M}(2)$ (propriedade (iv)). Além disso, como consequencia das três primeiras propriedades, tem-se que $\mathfrak{D}$ é *antissimétrica*, isto é, para quaisquer $v_1\,, v_2 \in \mathbb{R}^2$, vale a igualdade

v) $\mathfrak{D}(v_1, v_2) = -\mathfrak{D}(v_2, v_1)$.

De fato, segue-se de (i), (ii) e (iii) que:

$$\begin{aligned} 0 &= \mathfrak{D}(v_1 + v_2, v_1 + v_2) \\ &= \mathfrak{D}(v_1, v_1) + \mathfrak{D}(v_1, v_2) + \mathfrak{D}(v_2, v_1) + \mathfrak{D}(v_2, v_2) \\ &= \mathfrak{D}(v_1, v_2) + \mathfrak{D}(v_2, v_1). \end{aligned}$$

Suponhamos agora que $\mathfrak{D}' : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ seja uma função que cumpra as propriedades (i)–(iv) acima e, consequentemente, a propriedade (v). Nessas condições, dados $v_1 = (a_{11}, a_{21})$ e $v_2 = (a_{12}, a_{22})$ em $\mathbb{R}^2$, tem-se

$$\begin{aligned}
\mathfrak{D}'(v_1, v_2) &= \mathfrak{D}'(\, a_{11}e_1 + a_{21}e_2\,,\, a_{12}e_1 + a_{22}e_2\,) \\
&= a_{11}a_{12}\mathfrak{D}'(e_1, e_1) + a_{11}a_{22}\mathfrak{D}'(e_1, e_2) \\
&+ a_{21}a_{12}\mathfrak{D}'(e_2, e_1) + a_{21}a_{22}\mathfrak{D}'(e_2, e_2) \\
&= a_{11}a_{22} - a_{12}a_{21} \\
&= \mathfrak{D}(v_1, v_2).
\end{aligned}$$

Segue-se dessas considerações que a única função definida em $\mathbb{R}^2 \times \mathbb{R}^2$ que é bilinear, alternada e que vale 1 em (e_1, e_2) é a a função $\mathfrak{D}$ definida em (4.2), a qual denomina-se *determinante*.

Dada uma matriz $A = (a_{ij})_{2\times 2}$ em $\mathcal{M}(2)$, denota-se o determinante de A por $\det A$ e escreve-se

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

Temos, então, o seguinte resultado:

Proposição 4.1. *São equivalentes as seguintes afirmações acerca de uma matriz $A \in \mathcal{M}(2)$:*

- *Para toda matriz coluna $B \in \mathcal{M}(2, 1)$, o sistema linear $AX = B$ admite uma única solução;*
- *A é invertível;*
- $\det A \neq 0$.

Segue-se da Proposição 4.1 que a matriz $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$, contrariamente à matriz $B = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$, é invertível, pois $\det A = 1 \neq 0$ e $\det B = 0$.

4.2 Determinantes em $\mathcal{M}(n)$

A fim de estender para $\mathcal{M}(n)$ a noção de determinante introduzida na seção anterior, identifiquemos esse espaço com $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ (n cópias) através da seguinte correspondência:

$$A = (a_{ij})_{n\times n} \in \mathcal{M}(n) \;\; \cong \;\; (v_1\,, \dots, v_n) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n,$$

em que, para cada $j \in \{1, \dots, n\}$, $v_j = (a_{1j}, \dots, a_{nj})$ é o j-ésimo vetor coluna da matriz A.

Definição 4.1 (n-FORMA)**.** Uma função $\mathfrak{f} : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ é dita *n-linear* ou, equivalentemente, uma *n-forma* em $\mathbb{R}^n$, se, para todo índice $j \in \{1, \dots, n\}$, vale a igualdade

$$\mathfrak{f}(v_1\,, \dots, \lambda v_j + w_j\,, \dots, v_n) = \lambda\mathfrak{f}(v_1\,, \dots, v_j\,, \dots, v_n) + \mathfrak{f}(v_1\,, \dots, w_j\,, \dots, v_n)$$

quaisquer que sejam $\lambda \in \mathbb{R}$ e $v_1\,, \dots, v_j\,, w_j\,, \dots, v_n \in \mathbb{R}^n$.

Exemplo 4.1. A função

$$\mathfrak{f}(v_1\,, \dots, v_n) = a_{11} \dots a_{nn}\,, \;\; v_j = (a_{1j}\,, \dots, a_{nj}) \in \mathbb{R}^n,$$

é uma n-forma em $\mathbb{R}^n$. Com efeito, dados $j \in \{1, \dots, n\}$ e $\lambda \in \mathbb{R}$, tem-se, para quaisquer $v = (x_1, \dots, x_n)$, $w = (y_1\,, \dots, y_n) \in \mathbb{R}^n$, que

$$\begin{aligned}\mathfrak{f}(v_1, \dots, v + \lambda w, \dots, v_n) &= a_{11} \dots (x_j + \lambda y_j) \dots a_{nn} \\ &= a_{11} \dots x_j \dots a_{nn} + \lambda a_{11} \dots y_j \dots a_{nn}\,,\end{aligned}$$

donde se conclui que

$$\mathfrak{f}(v_1, \dots, v + \lambda w, \dots, v_n) = \mathfrak{f}(v_1, \dots, v, \dots, v_n) + \lambda\mathfrak{f}(v_1, \dots, w, \dots, v_n).$$

De modo análogo, prova-se que, dados $i_1, \dots, i_n \in \{1, \dots, n\}$, dois a dois distintos, a função $\mathfrak{f} : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$, definida por

$$\mathfrak{f}(v_1\,, \dots, v_n) = a_{i_1 1} a_{i_2 2} \dots a_{i_n n}\,, \;\; v_j = (a_{1j}\,, \dots, a_{nj})\,,$$

é uma n-forma em $\mathbb{R}^n$.

Definição 4.2 (n-FORMA ALTERNADA – ANTISSIMÉTRICA)**.** Diz-se que uma n-forma $\mathfrak{f}$ em $\mathbb{R}^n$ é *alternada* se, para quaisquer n vetores $v_1, \dots, v_n \in \mathbb{R}^n$, tem-se

$$\mathfrak{f}(v_1, \dots, v_n) = 0,$$

sempre que $v_i = v_j$ para algum par de índices $i \neq j \in \{1, \dots, n\}$. Diz-se que $\mathfrak{f}$ é *antissimétrica*, se

$$\mathfrak{f}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -\mathfrak{f}(v_1, \dots, v_j, \dots, v_i, \dots, v_n),$$

quaisquer que sejam $i, j \in \{1, \dots, n\}$.

Dadas duas n-formas em $\mathbb{R}^n$, $\mathfrak{f}$ e $\mathfrak{g}$, verifica-se facilmente que, para todo $\lambda \in \mathbb{R}$, a função $\mathfrak{f} + \lambda\mathfrak{g}$ é uma n-forma em $\mathbb{R}^n$. Além disso, $\mathfrak{f} + \lambda\mathfrak{g}$ é alternada, se $\mathfrak{f}$ e $\mathfrak{g}$ o forem. Daí, segue-se que

$$\mathcal{F}(\mathbb{R}^n) = \{n\text{-formas em}\,\mathbb{R}^n\} \;\text{ e }\; \mathcal{A}(\mathbb{R}^n) = \{n\text{-formas alternadas em}\,\mathbb{R}^n\}$$

são espaços vetoriais, sendo que $\mathcal{A}(\mathbb{R}^n)$, evidentemente, é um subespaço de $\mathcal{F}(\mathbb{R}^n)$.

Proposição 4.2. *Sejam $\mathfrak{f}$ uma n-forma alternada em $\mathbb{R}^n$ e $\{v_1, \dots, v_n\} \subset \mathbb{R}^n$ um conjunto* LD. *Então,* $\mathfrak{f}(v_1, \dots, v_n) = 0$.

Demonstração. Com efeito, pela n-linearidade de $\mathfrak{f}$, se um dos vetores v_i for nulo, teremos $\mathfrak{f}(v_1, \dots, v_n) = 0$. Caso contrário, pela Proposição 2.9, existem $k \in \{2, \dots, n\}$ e $\lambda_1, \dots, \lambda_{k-1} \in \mathbb{R}$, tais que $v_k = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1}$. Logo,

$$\mathfrak{f}(v_1, \dots, v_k, \dots, v_n) = \mathfrak{f}(v_1, \dots, v_{k-1}, \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1}, v_{k+1}, \dots, v_n) = 0\,,$$

uma vez que $\mathfrak{f}$ é n-linear e alternada. ■

Proposição 4.3. *Uma n-forma em $\mathbb{R}^n$ é alternada se, e somente se, é antissimétrica.*

Demonstração. Seja $\mathfrak{f}$ uma n-forma alternada em $\mathbb{R}^n$. Nesse caso, dados $i, j \in \{1, \ldots, n\}$ e $v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n \in \mathbb{R}^n$, tem-se

$$\begin{array}{rcl} 0 & = & \mathfrak{f}(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) \\ & = & \mathfrak{f}(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + \mathfrak{f}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) \\ & + & \mathfrak{f}(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + \mathfrak{f}(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) \\ & = & \mathfrak{f}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + \mathfrak{f}(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n), \end{array}$$

donde $\mathfrak{f}$ antissimétrica.

Reciprocamente, suponhamos que $\mathfrak{f}$ seja uma n-forma antissimétrica em $\mathbb{R}^n$. Então, para quaisquer $v_1, \ldots, v_n, v \in \mathbb{R}^n$, vale a igualdade

$$\mathfrak{f}(v_1, \ldots, v, \ldots, v, \ldots, v_n) = -\mathfrak{f}(v_1, \ldots, v, \ldots, v, \ldots, v_n).$$

Logo, $\mathfrak{f}(v_1, \ldots, v, \ldots, v, \ldots, v_n) = 0$, isto é, $\mathfrak{f}$ é alternada. ■

4.2.1 Existência do Determinante

Definição 4.3 (FUNÇÃO DETERMINANTE)**.** Uma n-forma alternada $\mathfrak{D}$ em $\mathbb{R}^n$, a qual cumpre $\mathfrak{D}(e_1, \ldots, e_n) = 1$, é dita uma *função determinante.*

Claramente, a função identidade é a única função determinante em $\mathbb{R}$. Conforme vimos na seção anterior, existe também uma única função determinante em $\mathbb{R}^2 \times \mathbb{R}^2 \cong \mathcal{M}(2)$.

Vejamos o caso $n = 3$, isto é, suponhamos que $\mathfrak{D} : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ seja uma função determinante. Escrevendo-se $v_j = \sum_{i=1}^{3} a_{ij} e_i$ e valendo-se das propriedades de $\mathfrak{D}$, tem-se

$$\begin{array}{rcl} \mathfrak{D}(v_1, v_2, v_3) & = & \mathfrak{D}\left(\sum_{i_1=1}^{3} a_{i_1 1} e_{i_1}, \sum_{i_2=1}^{3} a_{i_2 2} e_{i_2}, \sum_{i_3=1}^{3} a_{i_3 3} e_{i_3}\right) \\ & = & a_{11}a_{22}a_{33}\mathfrak{D}(e_1, e_2, e_3) + a_{11}a_{32}a_{23}\mathfrak{D}(e_1, e_3, e_2) \\ & + & a_{21}a_{12}a_{33}\mathfrak{D}(e_2, e_1, e_3) + a_{21}a_{32}a_{13}\mathfrak{D}(e_2, e_3, e_1) \\ & + & a_{31}a_{12}a_{23}\mathfrak{D}(e_3, e_1, e_2) + a_{31}a_{22}a_{13}\mathfrak{D}(e_3, e_2, e_1). \end{array}$$

Porém, por definição, $\mathfrak{D}(e_1, e_2, e_3) = 1$ e, pela Proposição 4.3, $\mathfrak{D}$ é antissimétrica. Logo, pondo-se $A \cong (v_1\,, v_2\,, v_3) \in \mathcal{M}(3)$, tem-se

$$\mathfrak{D}(A) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}\,.$$

Observemos agora que essa última igualdade pode ser escrita como

$$\begin{aligned}\mathfrak{D}(A) &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) \\ &= a_{11}\det A_{11} - a_{21}\det A_{21} + a_{31}\det A_{31}\,, \end{aligned} \tag{4.3}$$

em que $A_{1j}\,,\ j \in \{1, 2, 3\}$, é a matriz 2×2 obtida de A por eliminação de suas primeira linha e j-ésima coluna.

A igualdade (4.3) sugere fortemente a construção de funções determinante por indução, conforme o faremos na demonstração do Teorema 4.1. Antes, porém, consideremos o lema seguinte.

Lema 4.1. *Seja $\mathfrak{f}$ uma n-forma em $\mathbb{R}^n$ com a seguinte propriedade: Para todo $j \in \{2, \dots, n\}$ e quaisquer $v_1, \dots, v_{j-1}, v_j, \dots, v_n$, tem-se*

$$\mathfrak{f}(v_1, \dots, v_{j-1}, v_j, \dots, v_n) = 0$$

sempre que $v_{j-1} = v_j$. Então, $\mathfrak{f}$ é alternada.

Demonstração. O argumento usado na primeira parte da demonstração da Proposição 4.3 aplicado à n-forma $\mathfrak{f}$ prova que

$$\mathfrak{f}(v_1, \dots, v_{j-1}, v_j, \dots, v_n) = -\mathfrak{f}(v_1, \dots, v_j, v_{j-1}, \dots, v_n),$$

quaisquer que sejam $j \in \{1, \dots, n\}$ e $v_1, \dots, v_{j-1}, v_j, \dots, v_n \in \mathbb{R}^n$. Além disso, toda n-upla $(v_1, \dots, v_i, \dots, v_j, \dots, v_n)$ de vetores de $\mathbb{R}^n$ pode ser transformada na n-upla $(v_1, \dots, v_i, v_j, \dots, v_n)$ através de permutações sucessivas de vetores adjacentes, isto é, do tipo v_{k-1}, v_k. Segue-se, então, da igualdade acima, que

$$\mathfrak{f}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = (-1)^p\, \mathfrak{f}(v_1, \dots, v_i, v_j, \dots, v_n),$$

em que p é o número de permutações realizadas. Em particular, se $v_i = v_j$, tem-se $\mathfrak{f}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = 0$, donde se conclui que $\mathfrak{f}$ é alternada. ■

Teorema 4.1 (EXISTÊNCIA DO DETERMINANTE EM $\mathcal{M}(n)$). *Para todo $n \in \mathbb{N}$, existe uma função determinante em $\mathcal{M}(n)$.*

Demonstração. O resultado é imediato para $n = 1$. Suponhamos, então, que, dado $n \geq 2$, exista uma função determinante $\mathfrak{D}$ em $\mathcal{M}(n-1)$, e definamos a função $\mathfrak{D}_1$ em $\mathcal{M}(n)$ por

$$\mathfrak{D}_1(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \mathfrak{D}(A_{1j}) \,, \tag{4.4}$$

isto é

$$\mathfrak{D}_1(A) = a_{11}\mathfrak{D}(A_{11}) - a_{12}\mathfrak{D}(A_{12}) + \cdots + (-1)^{1+n} a_{1n}\mathfrak{D}(A_{1n}) \,,$$

em que A_{1j}, $j \in \{1, \ldots, n\}$, é a matriz $(n-1)\times(n-1)$ obtida de A eliminando-se sua primeira linha e sua j-ésima coluna.

Mostremos, pois, que $\mathfrak{D}_1$ é uma função determinante em $\mathcal{M}(n)$, o que, juntamente com o Princípio da Indução, nos levará à conclusão do teorema.

Com esse propósito, consideremos a aplicação linear $T : \mathbb{R}^n \to \mathbb{R}^{n-1}$, definida por $T(x_1, \ldots, x_n) = (x_2, \ldots, x_n)$, e observemos que, para cada $j \in \{1, \ldots, n\}$, vale

$$A \cong (v_1, \ldots, v_n) \;\Rightarrow\; A_{1j} \cong (T(v_1), \ldots, T(v_{j-1}), T(v_{j+1}), \ldots, T(v_n)).$$

Em particular, a aplicação

$$A \cong (v_1, \ldots, v_n) \in \mathcal{M}(n) \mapsto A_{1j} \in \mathcal{M}(n-1)$$

é linear com respeito a cada uma das variáveis $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$. Consequentemente, uma vez que $\mathfrak{D}$ é $(n-1)$-linear, vale o mesmo para a função

$$A \cong (v_1, \ldots, v_n) \in \mathcal{M}(n) \;\mapsto\; \mathfrak{f}_j(v_1, \ldots, v_n) = a_{1j}\mathfrak{D}(A_{1j}). \tag{4.5}$$

Além disso, tomando-se $\lambda \in \mathbb{R}$ e $v_j' = (a_{1j}', \ldots, a_{nj}')$ em $\mathbb{R}^n$, tem-se

$$\begin{aligned} \mathfrak{f}_j(v_1, \ldots, v_j + \lambda v_j', \ldots, v_n) &= (a_{1j} + \lambda a_{1j}')\mathfrak{D}(A_{1j}) \\ &= \mathfrak{f}_j(v_1, \ldots, v_j, \ldots, v_n) + \lambda \mathfrak{f}_j(v_1, \ldots, v_j', \ldots, v_n), \end{aligned}$$

ou seja, $\mathfrak{f}_j$ é linear, também, com respeito a v_j.

Segue-se que, para todo $j = 1, \dots, n$, a função $\mathfrak{f}_j$ é n-linear. Portanto, $\mathfrak{D}$, sendo uma combinação linear dessas funções, é n-linear.

Tomemos agora $A \cong (v_1, \dots, v_n) \in \mathcal{M}(n)$ e suponhamos que, para algum $k \in \{2, \dots, n\}$, tenha-se $v_{k-1} = v_k$. Nesse caso, para todo $j \in \{1, \dots, n\}$ diferente de $k-1$ e de k, a matriz A_{1j} tem duas colunas iguais. Além disso, as matrizes $A_{1(k-1)}$ e A_{1k} são iguais. Logo,

$$\mathfrak{D}_1(v_1, \dots, v_n) = (-1)^k a_{1(k-1)} \mathfrak{D}(A_{1(k-1)}) + (-1)^{k+1} a_{1k} \mathfrak{D}(A_{1k}) = 0.$$

Dessa igualdade e do Lema 4.1, conclui-se que $\mathfrak{D}$ é alternada.

Por fim, verifica-se por substituição direta que $\mathfrak{D}_1(e_1, \dots, e_n) = 1$, donde se infere que $\mathfrak{D}_1$ é uma função determinante em $\mathcal{M}(n)$, como queríamos demonstrar. ■

Através de um argumento análogo ao da demonstração acima, prova-se que, para todo $i \in \{1, \dots, n\}$, a função

$$D_i(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \mathfrak{D}(A_{ij}), \quad A \in \mathcal{M}(n), \tag{4.6}$$

é uma função determinante em $\mathcal{M}(n)$, dita o *desenvolvimento de Laplace* de A com repeito à sua i-ésima linha.

Consideremos, por exemplo, a matriz

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

Temos que:

- $\mathfrak{D}_1(A) = 1. \begin{vmatrix} 2 & -3 \\ 0 & 1 \end{vmatrix} + 1. \begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix} + 0. \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} = 9;$

- $\mathfrak{D}_2(A) = -\begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} + 2. \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} + 3. \begin{vmatrix} 1 & -1 \\ 2 & 0 \end{vmatrix} = 9;$

- $\mathfrak{D}_3(A) = 2.\begin{vmatrix} -1 & 0 \\ 2 & -3 \end{vmatrix} - 0.\begin{vmatrix} 1 & 0 \\ 1 & -3 \end{vmatrix} + 1.\begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 9;$

o que sugere a unicidade da função determinante.

4.2.2 Unicidade do Determinante

No que se segue, constataremos que as n expressões (4.6) definem a mesma função determinante, o que decorrerá do fato de que, na verdade, para todo $n \in \mathbb{N}$, existe uma única função determinante em $\mathcal{M}(n)$. A demonstração desse fato nos levará a fazer algumas considerações sobre permutações.

Definição 4.4 (PERMUTAÇÃO – TRANSPOSIÇÃO)**.** Dado $n \in \mathbb{N}$, uma *permutação* (*de grau* n) é uma função bijetiva $\sigma : N_n \to N_n$, em que $N_n = \{1, \ldots, n\}$. Uma permutação σ de grau n é dita uma *transposição*, se existem $i, j \in N_n$, tais que:

- $i \neq j$;
- $\sigma(i) = j$ e $\sigma(j) = i$;
- $\sigma(k) = k \, \forall k \in N_n - \{i, j\}$.

Adotaremos a seguinte notação: Dada uma permutação $\sigma : N_n \to N_n$, escreveremos $\sigma(1) = \sigma_1, \ldots, \sigma(n) = \sigma_n$. Quando σ e τ forem, ambas, permutações de grau n, denotaremos a composta $\sigma \circ \tau$ como o produto $\sigma\tau$. Evidentemente, o produto de permutações de grau n é também uma permutação de grau n.

Dito de forma simples, uma permutação de grau n nada mais é que uma reordenação da sequência $1, \ldots, n$. A partir desse ponto de vista, é fácil ver que toda permutação se expressa como um produto de transposiçõoes, isto é, qualquer permutação preestabelecida pode ser obtida, reordenando-se os elementos de N_n de dois em dois. Assim, para toda permutação $\sigma : N_n \to N_n$, existem transposições $\tau_1, \ldots, \tau_k$, tais que

$$\sigma = \tau_1 \ldots \tau_k \,. \tag{4.7}$$

À igualdade (4.7), chamaremos uma *representação* de σ.

É claro que, em geral, uma permutação admite mais de uma representação. No entanto, o número k de transposições de uma representação qualquer de uma permutação σ é sempre par ou sempre ímpar, ou seja, a paridade de k depende de σ, não da representação em si. Curiosamente, podemos demonstrar facilmente esse fato, de caráter puramente combinatório, usando a teoria de determinantes. Com efeito, seja $\mathfrak{D}$ uma função determinante em $\mathcal{M}(n)$. Uma vez que $\mathfrak{D}$ é antissimétrica, se σ for uma permutação de grau n representada como em (4.7), teremos que

$$\mathfrak{D}(e_{\sigma_1}, \dots, e_{\sigma_n}) = (-1)^k \mathfrak{D}(e_1, \cdots, e_n) = (-1)^k.$$

Dessa forma, k será par ou ímpar, conforme tenha-se $\mathfrak{D}(e_{\sigma_1}, \dots, e_{\sigma_n}) = 1$ ou $\mathfrak{D}(e_{\sigma_1}, \cdots, e_{\sigma_n}) = -1$, respectivamente.

Definição 4.5 (PARIDADE DE PERMUTAÇÕES)**.** Diz-se que uma permutação é *par*, quando é representada por um número par de transposições. Caso contrário, ela é dita *ímpar*. Define-se o *sinal* de uma permutação σ, $\operatorname{sgn}\sigma$, como 1, para permutações pares, e -1, para permutações ímpares.

Tomemos agora uma n-forma alternada $\mathfrak{f}$ em $\mathbb{R}^n$. Pondo-se, para cada índice $j \in \{1, \cdots, n\}$,

$$v_j = (a_{1j}, \cdots, a_{nj}) = \sum_{i=1}^{n} a_{ij} e_i \in \mathbb{R}^n,$$

e procedendo-se como o fizemos no caso particular $n = 3$, concluí-se, igualmente, que $\mathfrak{f}(v_1, \dots, v_n)$ é uma soma de $n!$ parcelas do tipo

$$a_{\sigma_1 1} a_{\sigma_2 2} \dots a_{\sigma_n n}\, \mathfrak{f}(e_{\sigma_1}, e_{\sigma_2}, \dots, e_{\sigma_n}),$$

em que cada σ é uma permutação de grau n. Uma vez que $\mathfrak{f}$ é antissimétrica, essa expressão nada mais é que

$$(\operatorname{sgn}\sigma) a_{\sigma_1 1} a_{\sigma_2 2} \dots a_{\sigma_n n}\, \mathfrak{f}(e_1, e_2, \dots, e_n),$$

isto é, dada $A \in \mathcal{M}(n)$, tem-se

$$\mathfrak{f}(A) = \sum_{\sigma \in \Lambda} (\operatorname{sgn} \sigma) a_{\sigma_1 1} a_{\sigma_2 2} \dots a_{\sigma_n n}\, \mathfrak{f}(I), \tag{4.8}$$

em que $\Lambda = \{\text{permutações de grau } n\}$.

Segue-se da igualdade (4.8) que existe uma única função determinante em $\mathcal{M}(n)$, a qual denotamos por $\det$, isto é,

$$\det A = \sum_{\sigma \in \Lambda} (\operatorname{sgn} \sigma) a_{\sigma_1 1} a_{\sigma_2 2} \cdots a_{\sigma_n n}\,, \quad A = (a_{ij}) \in \mathcal{M}(n). \tag{4.9}$$

Em suma, tem-se o resultado seguinte.

Teorema 4.2 (UNICIDADE DO DETERMINANTE)**.** *Para todo* $n \in \mathbb{N}$, *existe uma única função determinante em* $\mathcal{M}(n)$, *a qual denota-se por* $\det$ *e define-se através da igualdade* (4.9). *Além disso, se* $\mathfrak{f}$ *for uma* n*-forma alternada, para toda matriz* $A \in \mathcal{M}(n)$, *valerá a seguinte igualdade:*

$$f(A) = \det A f(I).$$

Em particular, $\dim \mathcal{A}(\mathbb{R}^n) = 1$.

Dada uma matriz $A = (a_{ij}) \in \mathcal{M}(n)$, observemos que

$$a_{\sigma_1 1} a_{\sigma_2 2} \dots a_{\sigma_n n} = a_{1\sigma_1^{-1}} a_{2\sigma_2^{-1}} \dots a_{n\sigma_n^{-1}} \;\; \forall \sigma \in \Lambda.$$

Além disso, se σ se expressa como em (4.7), então, $\sigma^{-1} = \tau_k^{-1} \dots \tau_1^{-1}$, donde $\operatorname{sgn} \sigma = \operatorname{sgn} \sigma^{-1}$. Dessas considerações e da igualdade (4.9), segue-se que

$$\det A = \det A^* \;\; \forall A \in \mathcal{M}(n).$$

Dada uma matriz $A = (a_{ij}) \in \mathcal{M}(n)$, escreveremos

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

4.3 Propriedade Fundamental – Aplicações

Estabeleceremos nessa seção a propriedade fundamental dos determinantes, qual seja, que o determinante do produto de duas matrizes (quadradas e de mesma ordem) é igual ao produto de seus determinantes. Aplicaremos essa propriedade para mostrar que uma matriz quadrada é invertível se, e somente se, seu determinante é não nulo, o que no permitirá definir o determinante de um operador linear de um espaço de dimensão finita. Obteremos, também, uma fórmula de inversão para matrizes, bem como um método para calcular o posto de uma matriz arbitrária (não necessariamente quadrada). Concluiremos, então, apresentando um resultado sobre polinômios, o qual envolve o conceito de determinante, conhecido como *critério de Sylvester.*

Teorema 4.3. *Dadas matrizes* A, $B \in \mathcal{M}(n)$, *tem-se*

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

Demonstração. Fixemos $A \in \mathcal{M}(n)$ e consideremos a função $\mathfrak{f} : \mathcal{M} \to \mathbb{R}$, dada por $\mathfrak{f}(B) = \det(AB)$. Observemos, então, que se $B \cong (v_1, \ldots, v_n)$, tem-se $AB \cong (Av_1, \ldots, Av_n)$, isto é, $\mathfrak{f}(v_1, \ldots, v_n) = \det(Av_1, \ldots, Av_n)$. Daí, conclui-se facilmente que $\mathfrak{f}$ é n-linear e alternada. Logo, pelo Teorema 4.2, tem-se $\mathfrak{f}(B) = \det(B)\mathfrak{f}(I)$, donde $\det(AB) = \det(B)\det(A)$, como desejado. ■

Sejam $A \in \mathcal{M}(n)$ uma matriz invertível e A^{-1} sua inversa. Segue-se da igualdade $AA^{-1} = I$ e do Teorema 4.3, que $\det A$ e $\det A^{-1}$ são não nulos e cumprem $\det A = 1/\det A^{-1}$. Por outro lado, se $A \cong (v_1, \ldots, v_n)$ não for invertível, conforme vimos no capítulo anterior, o posto de A será menor que n. Em particular, os vetores coluna de A serão LD. Logo, pela Proposição 4.2, teremos $\det A = 0$.

Vale, portanto, o seguinte resultado.

Proposição 4.4. *Uma condição necessária e suficiente para que uma matriz* $A \in \mathcal{M}(n)$ *seja invertível, é a de que seu determinante seja não nulo. No caso afirmativo, valerá a igualdade*

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

Exemplo 4.2. Conforme constatamos anteriormente, a matriz

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

tem determinante igual a $9 \neq 0$. Logo, A é invertível e $\det A^{-1} = 1/9$.

Consideremos um espaço vetorial V de dimensão finita, e bases $\mathcal{B}\,, \mathcal{B}' \subset V$. Temos, pela Proposição 3.2, que as respectivas matrizes de um operador linear $T : V \to V$ com respeito a $\mathcal{B}$ e $\mathcal{B}'$, $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{B}'}$, são semelhantes, isto é, existe uma matriz invertível M (a matriz de mudança de $\mathcal{B}$ para $\mathcal{B}'$), tal que

$$[T]_{\mathcal{B}} = M^{-1}[T]_{\mathcal{B}'}M. \tag{4.10}$$

Segue-se, portanto, de (4.10), do Teorema 4.3 e da Proposição 4.4 que

$$\det[T]_{\mathcal{B}} = \det[T]_{\mathcal{B}'}\,,$$

igualdade esta que justifica a definição a seguir.

Definição 4.6 (DETERMINANTES DE OPERADORES)**.** O *determinante* de um operador linear $T : V \to V$ num espaço vetorial V de dimensão finita, denotado por $\det T$, é definido por

$$\det T = \det\,[T]_{\mathcal{B}}\,,$$

em que $\mathcal{B}$ é uma base arbitrária de V.

Observe que, pela Proposição 4.4, *um operador* $T : V \to V$ $(\dim V < \infty)$ *é invertível se, e somente se*, $\det T \neq 0$.

4.3.1 Inversão de Matrizes

Dada uma matriz $A \in \mathcal{M}(n)$, seja $C = (c_{ij}) \in \mathcal{M}(n)$ a matriz tal que

$$c_{ij} = (-1)^{i+j} \det A_{ij}\,. \tag{4.11}$$

Cada entrada c_{ij} é dita um *cofator* de A, e C é dita, então, a *matriz dos cofatores* de A. Define-se a *adjunta clássica* de A como a matriz $\operatorname{adj} A = C^*$, isto é,

$$\operatorname{adj} A = (d_{ij})_{n\times n}\,, \;\; d_{ij} = (-1)^{i+j}\det A_{ji}\,.$$

Verifiquemos que

$$A(\operatorname{adj} A) = (\det A)I \quad \forall A \in \mathcal{M}(n). \tag{4.12}$$

De fato, escrevendo-se $B = A(\operatorname{adj} A)$, tem-se, para todo $i \in \{1, \ldots, n\}$, que

$$b_{ii} = \sum_{k=1}^{n} a_{ik}d_{ki} = \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\det A_{ik} = \det A. \tag{4.13}$$

Agora, dados $1 \leq i \neq j \leq n$, seja $A' = (a'_{ij}) \in \mathcal{M}(n)$ a matriz obtida de A substituindo-se sua j-ésima linha por sua i-ésima linha. Assim, A' terá duas linhas iguais e, portanto, $\det A' = 0$. Note também que $A'_{jk} = A_{jk}\,\forall k \in \{1, \ldots, n\}$. Dessa forma,

$$b_{ij} = \sum_{k=1}^{n} a_{ik}d_{kj} = \sum_{k=1}^{n}(-1)^{j+k}a_{ik}\det A_{jk} = \sum_{k=1}^{n}(-1)^{j+k}a'_{jk}\det A'_{jk} = \det A' = 0.$$

Combinando-se essa última igualdade com (4.13), obtém-se (4.12). Em particular, se A for invertível, valerá a igualdade

$$A^{-1} = \frac{1}{\det A}(\operatorname{adj} A). \tag{4.14}$$

Retomando-se a matriz

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

do Exemplo 4.2, a qual verificamos ser invertível, pode-se obter facilmente sua inversa por meio da fórmula (4.14). Com efeito, calculando-se os cofatores (c_{ij}) de A através de (4.11) e lembrando-se que $\det A = 9$, obtém-se

$$A^{-1} = \frac{1}{9}\begin{bmatrix} 2 & 1 & 3 \\ -7 & 1 & 3 \\ -4 & -2 & 3 \end{bmatrix}.$$

4.3.2 Determinação do Posto de uma Matriz

Conideremos a matriz

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 5 & 1 & 0 \\ 0 & 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

cujo posto, claramente, é igual a 3. Observemos que, exluindo-se sua última linha e suas duas últimas colunas, obtém-se a seguinte *submatriz* 3×3 de A:

$$A' = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix},$$

cujo determinante é igual a $1 \neq 0$ e que, portanto, é invertível. Em particular, posto $(A') = 3$. Note que a escolha da linha e colunas excluídas foi determinada pelo fato de o espaço linha de A ser gerado pelos três primeiros vetores linha, e seu espaço coluna ser gerado pelos três primeiros vetores coluna.

De modo geral, se o posto de uma matriz não nula $A \in \mathcal{M}(m,n)$ for p, teremos, pela Proposição 2.11, que $\dim \mathscr{L}_A = \dim \mathscr{C}_A = p$. Nessas condições, excluindo-se, de A, $m-p$ vetores linha e $n-p$ vetores coluna, convenientemente escolhidos, determina-se uma submatriz $A' \in \mathcal{M}(p)$, cujo posto é p e que, portanto, tem determinante não nulo.

Por fim, observemos que se A' for uma submatriz de A e α' for um subconjunto LI do conjunto formado por seus vetores coluna, então o conjunto α, formado pelos vetores coluna correspondentes de A, será também LI. Logo, o posto de uma submatriz de A nunca é maior que o de A. Dessa forma, se o posto de uma matriz A for igual a p, nenhuma submatriz quadrada de A, de ordem maior que p, pode ter determinante não nulo.

Vale, portanto, o resultado seguinte.

Proposição 4.5. *O posto de uma matriz $A \in \mathcal{M}(m,n)$ é o maior inteiro p com a propriedade de que existe uma submatriz $A' \in \mathcal{M}(p)$, de A, tal que* $\det A' \neq 0$.

Exemplo 4.3. O posto da matriz

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

é 2, pois, designando-se por $A' \in \mathcal{M}(2)$ a submatriz de A obtida por exclusão de sua terceira linha, tem-se $\det A' = 1 \neq 0$.

4.3.3 Critério de Sylvester

Em conclusão, aplicaremos o Teorema 4.3 para estabelecer um elegante resultado sobre fatoração de polinômios, devido a James Sylvester (1814–1897), o qual envolve determinantes(i).

Dados dois polinômios $f,\, g \in \mathcal{P}[(t,\mathbb{R}]$, em que

$$f(t) = a_m t^m + \cdots + a_1 t + a_0 \quad \text{e} \quad g(t) = b_n t^n + \cdots + b_1 t + b_0\,, \;\; a_m\,, b_n \neq 0,$$

associamos aos mesmos uma matriz $(m+n)\times(m+n)$, dita a *matriz de Sylvester* do par (f,g), a qual se define como

$$S(f,g) = \left[\begin{array}{cccccc} a_m & & \cdots & & a_0 & \\ & \ddots & & & & \ddots \\ & & a_m & \cdots & & a_0 \\ b_n & & \cdots & b_0 & & \\ & \ddots & & & \ddots & \\ & & b_n & \cdots & & b_0 \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c} a \\ \ddots \\ a \end{array}}\right\} n \text{ linhas} \\ \left.\vphantom{\begin{array}{c} b \\ \ddots \\ b \end{array}}\right\} m \text{ linhas} \end{array} .$$

Na matriz $S(f,g)$, em cada linha, as entradas não especificadas são todas iguais a zero, enquanto as especificadas, como indicado, são os coeficientes de f (n primeiras linhas) ou de g (m últimas linhas). Note que, escrevendo-se $S(f,g) = (s_{ij})$, os coeficientes de f e g são dispostos ao longo das linhas de

(i) Vide: H. Scala, *An application of determinants.* American Mathemaical Monthly **78** (1971) 889–890.

$S(f,g)$ de modo que, se $s_{ij} = a_k$ (respec., $s_{ij} = b_k$), então $s_{(i+1)(j+1)} = a_k$ (respec., $s_{(i+1)(j+1)} = b_k$) , $1 \leq i, j < m+n$

Por fim, relembremos que todo polinômio $f \in \mathcal{P}[t,\mathbb{R}]$ admite uma fatoração única, a menos da ordem dos fatores, do tipo

$$f = f_1 \dots f_k\,,$$

em que cada fator f_i é um polinômio irredutível (isto é, que não admite fatoração em $\mathcal{P}[t,\mathbb{R}]$) de grau 1 ou 2. Nesse contexto, diz-se que f, $g \in \mathcal{P}[t,\mathbb{R}]$ têm um *fator comum*, se um dos fatores de f for igual a um dos fatores de g.

Teorema 4.4 (Critério de Sylvester)**.** *Dois polinômios f, $g \in \mathcal{P}[t,\mathbb{R}]$ têm um fator comum se, e somente se,* $\det S(f,g) = 0$.

Demonstração. Consideremos a seguinte matriz $(m+n)\times(m+n)$,

$$A = \begin{bmatrix} t^{n+m-1} & 0 & 0 & \cdots & 0 & 0 \\ t^{n+m-2} & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ t & 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

e observemos que $\det A = t^{n+m-1}$. Além disso, pondo-se $S = S(f,g)$, tem-se

$$SA = \begin{bmatrix} t^{n-1}f(t) & a_{m-1} & a_{m-2} & \cdots & 0 \\ t^{n-2}f(t) & a_m & a_{m-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ f(t) & \cdots & & \cdots & a_0 \\ t^{n-1}g(t) & b_{m-1} & b_{m-2} & \cdots & 0 \\ t^{n-2}g(t) & b_m & b_{m-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ g(t) & \cdots & & \cdots & b_0 \end{bmatrix}.$$

Logo, considerando-se o desenvolvimento de Laplace de SA relativo à sua primeira coluna, obtém-se

$$\det(SA) = p(t)f(t) + q(t)g(t),$$

em que $p\,,q \in \mathcal{P}[t,\mathbb{R}]$ são polinômios cujos graus não excedem $n-1$ e $m-1$, respectivamente. Porém, pelo Teorema 4.3, $\det(SA) = \det(S)\det A$, donde

$$t^{n+m-1}\det S = p(t)f(t) + q(t)g(t). \tag{4.15}$$

Assim, se $\det S = 0$, teremos $p(t)f(t) = -q(t)g(t)$ e, então, pf e $-qg$ terão os mesmos fatores. Em particular, todos os fatores de f dividirão q ou g. Porém, o grau de f é maior que o de q. Logo, algum fator de f não dividirá q e, portanto, dividirá g, isto é, f e g terão um fator comum.

Suponhamos agora que f e g tenham um fator comum r. Nesse caso, segue-se de (4.15) que $t^{n+m-1}\det S = \lambda(t)r(t)$, em que $\lambda \in \mathcal{P}[t,\mathbb{R}]$. Se λ for o polinômio nulo, então $\det S = 0$. Caso contrário, devemos ter $r(t) = ct^k$, em que $k = 1$ ou $k = 2$, e $c \neq 0$. Em particular, sendo r um fator de f e de g, tem-se $a_0 = b_0 = 0$, o que nos dá, igualmente, $\det S = 0$. ■

Aplicando-se o critério de Sylvester aos polinômios

$$f(t) = t^3 - 2t^2 + 3t - 2 \quad \text{e} \quad g(t) = 2t^2 - t - 1,$$

conclui-se que os mesmos têm um fator comum, pois sua matriz de Sylvester,

$$S(f,g) = \begin{bmatrix} 1 & -2 & 3 & -2 & 0 \\ 0 & 1 & -2 & 3 & -2 \\ 2 & -1 & -1 & 0 & 0 \\ 0 & 2 & -1 & -1 & 0 \\ 0 & 0 & 2 & -1 & -1 \end{bmatrix},$$

tem determinante nulo (verifique!).

4.4 Exercícios

Seções 4.1 e 4.2

1. Exiba matrizes $A, B \in \mathcal{M}(2)$, tais que

$$\det A = \det(A+I) \quad \text{e} \quad \det B = \det(B-I).$$

Prove que não existe uma matriz $A \in \mathcal{M}(2)$, tal que

$$\det A = \det(A + I) = \det(A - I).$$

2. Prove que o determinante de uma matriz triangular é igual ao produto das entradas de sua diagonal.

3. Seja A uma matriz $n \times n$ com mais de $n^2 - n$ entradas nulas. Mostre que $\det A = 0$.

4. Seja $A = (a_{ij}) \in \mathcal{M}(n)$ uma matriz quadrada cujas entradas a_{ij} são números inteiros. Suponha que a soma das entradas de cada linha de A seja igual a $k \in \mathbb{Z}$, isto é,

$$\sum_{j=1}^{n} a_{ij} = k \ \forall i \in \{1, \ldots, n\}.$$

Denote por $T \in \mathscr{L}(\mathbb{R}^n)$ o operador de $\mathbb{R}^n$, tal que $[T] = A$. Prove que, nessas condições:

- k divide $\det A$;
- Existe $v \in \mathbb{R}^n$, tal que $T(v) = kv$.

5. Seja $A = (a_{ij}) \in \mathcal{M}(3)$ uma matriz simétrica, tal que $a_{ii} = 1 \, \forall i \in \{1, 2, 3\}$ e $|a_{ij}| \leq 1 \, \forall i \neq j \in \{1, 2, 3\}$. Mostre que $\det A \leq 1$.

6. Prove que, para quaisquer reais a, b, c, vale a igualdade

$$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2).$$

Estabeleça, então, através do princípio da indução, que a *matriz de Vandermonde*,

$$A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \vdots & \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{bmatrix},$$

satisfaz

$$\det A = \prod_{i>j} (a_i - a_j).$$

7. Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear. Mostre que

$$\det(T(v_1), \cdots, T(v_n)) = \lambda \det(v_1, \cdots, v_n) \ \ \forall v_1, \cdots, v_n \in \mathbb{R}^n,$$

em que $\lambda = \det(T(e_1), \cdots, T(e_n))$.

Seção 4.3

8. Seja A uma matriz $n \times n$ *antissimétrica*, isto é, $A^* = -A$. Prove que $\det A = 0$, se n for ímpar.

9. Diz-se que uma matriz $A \in \mathcal{M}(n)$ é *ortogonal*, quando $AA^* = I$. Prove que o determinante de uma matriz ortogonal é sempre 1 ou -1.

10. Mostre que, quaisquer que sejam a, b, c, $\theta \in \mathbb{R}$, as matrizes

$$A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \ \text{ e } \ B = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

não são semelhantes, a menos que se tenha $A = B = I$.

11. Sejam A, $B \in \mathcal{M}(n)$ matrizes linha equivalentes. Mostre que $\det A = 0$ se, e somente se, $\det B = 0$. Se $\det A \neq 0$, sob que condições teremos $\det A = \det B$?

12. Prove que $\operatorname{adj} A^* = (\operatorname{adj} A)^* \ \forall A \in \mathcal{M}(n)$.

13. Sejam $A, B \in \mathcal{M}(n)$ matrizes invertíveis. Prove que:

    a) $\operatorname{adj}(AB) = (\operatorname{adj} B)(\operatorname{adj} A)$;

    b) $(\operatorname{adj} A)B = B(\operatorname{adj} A)$, se $AB = BA$.

14. (*Regra de Cramer*) Seja $A \cong (v_1, \ldots, v_n) \in \mathcal{M}(n)$ uma matriz invertível. Dado $b \in \mathbb{R}^n$, mostre que a (única) solução do sistema $AX = B$, $\mathcal{M}(n,1) \ni B \cong b$, é $X \cong (x_1, \ldots, x_n)$, tal que

$$x_j = \frac{\det A_j}{\det A}, \quad A_j \cong (v_1, \ldots, v_{j-1}, b, v_{j+1}, \ldots, v_n).$$

15. Dada uma matriz $A \in \mathcal{M}(n)$, defina $T : \mathcal{M}(n) \to \mathcal{M}(n)$ por

$$T(X) = AX - XA.$$

Prove que T é linear e que $\det T = 0$.

16. Seja $A \cong (v_1, \ldots, v_n) \in \mathcal{M}(n)$, $n > 1$, tal que, para quaisquer vetores $w_1, \ldots, w_n \in \mathbb{R}^n$, tenha-se

$$\sum_{j=1}^{n} \det(v_1, v_2, \ldots, v_{j-1}, w_j, v_{j+1}, \ldots, v_n) = 0.$$

Prove que o posto de A é menor que $n - 1$.

5

Produto Interno

Uma característica fundamental dos espaços vetoriais é a de admitir estruturas geométricas. Dito de outra forma, em todo espaço vetorial podem ser introduzidas as noções de *distância* e *ângulo* entre vetores, as quais generalizam aquelas da Geometria Euclidiana. Isso é feito através de uma função com propriedades especiais, denominada *produto interno*.

Assim, um produto interno é um acessório que, agregado a um espaço vetorial, enriquece sua estrutura, tornando-o um objeto matemático de grande interesse. Por exemplo, a geometria de uma superfície de $\mathbb{R}^3$, tal como uma esfera, um cilindro ou uma quádrica, é totalmente determinada pelo produto interno que se associa a cada um de seus planos tangentes, os quais são subespaços vetoriais de $\mathbb{R}^3$. Teorias mais avançadas, como Geometria Riemanniana, valem-se de ideias análogas. Em suma, parafraseando o célebre astrônomo Johannes Kepler, que afirmava: " Onde há matéria, há Geometria"; declaramos: Onde há Geometria, há produto interno.

No que se segue, discutiremos sobre o conceito de produto interno em espaços vetoriais, e de seu consequente, o de ortogonalidade. Como motivação, iniciaremos mostrando que a Geometria Euclidiana Plana é totalmente determinada por um produto interno, dito o produto escalar. Concluiremos, então, introduzindo a noção de volume (de paralelepípedos) em espaços euclidianos.

5.1 Geometria Plana e Produto Escalar

Consideremos, no plano $\mathbb{R}^2$, as noções de comprimento e ângulo da Geometria Euclidiana, as quais, como sabemos, relacionam-se através das lendárias igualdades trigonométricas em triângulos retângulos. Do ponto de vista da Álgebra Linear, o fato notável, o qual fundamenta a teoria que discutiremos nas próximas seções deste capítulo, é que essas noções e, portanto, toda a Geometria Euclidiana Plana, são determinadas por uma única função, dita o *produto escalar.* Vejamos, então, como isto se dá.

O *produto escalar* $\langle v, w\rangle$ entre os vetores $v = (x_1, x_2)$ e $w = (y_1, y_2)$ em $\mathbb{R}^2$ define-se através da expressão

$$\langle v, w\rangle = x_1y_1 + x_2y_2 \,.$$

Representando-se vetores em $\mathbb{R}^2$ como segmentos de reta orientados e com origem em $\mathbf{0} = (0, 0)$, segue-se diretamente do Teorema de Pitágoras que o comprimento de um vetor $v = (x_1, x_2) \in \mathbb{R}^2$, o qual chamamos *norma* de v e denotamos por $\|v\|$, expressa-se como

$$\|v\| = \sqrt{x_1^2 + x_2^2}\,.$$

Igualmente, se $w = (y_1, y_2)$, o comprimento do segmento de reta determinado pelas extremidades de v e w, dito a *distância* entre v e w e denotado por $\operatorname{dist}(v, w)$, satisfaz (vide Fig. 5.1):

$$\operatorname{dist}(v, w) = \|v - w\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}\,.$$

Observemos que, para qualquer vetor $v = (x_1, x_2) \in \mathbb{R}^2$, tem-se

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Em particular, vale a igualdade

$$\operatorname{dist}(v, w) = \sqrt{\langle v - w, v - w\rangle}.$$

Consequentemente, em $\mathbb{R}^2$, distância (ou comprimento) é um conceito que se exprime completamente através do produto escalar. Vejamos que o mesmo acontece com o conceito de ângulo. Mais precisamente, verifiquemos que o ângulo θ entre dois vetores LI, $v = (x_1, x_2)$ e $w = (y_1, y_2)$, satisfaz

$$\cos\theta = \frac{\langle v, w\rangle}{\|v\|\,\|w\|}\,\cdot \tag{5.1}$$

Em particular, v e w são ortogonais (isto é, perpendiculares entre si) se, e somente se, $\langle v, w\rangle = 0$.

Para obtermos (5.1), designemos por θ_v e θ_w os ângulos (orientados no sentido anti-horário) entre o eixo x_1, das abscissas, e os vetores v e w, respectivamente. Assim, admitindo-se $\theta_v > \theta_w$, tem-se que θ satisfaz $\theta = \theta_v - \theta_w$, donde $\cos\theta = \cos\theta_v \cos\theta_w + \operatorname{sen}\theta_v \operatorname{sen}\theta_w$. Logo (vide Figura 5.1),

$$\cos\theta = \frac{x_1}{\|v\|}\frac{y_1}{\|w\|} + \frac{x_2}{\|v\|}\frac{y_2}{\|w\|} = \frac{\langle v, w\rangle}{\|v\|\,\|w\|}\,\cdot$$

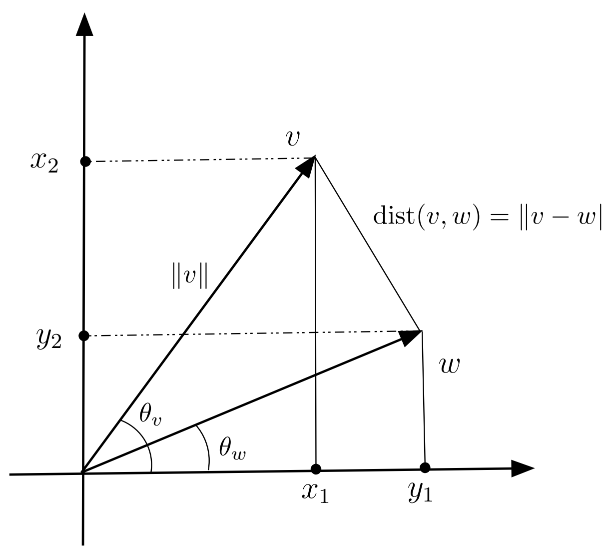


Figura 5.1

Segue-se das considerações acima que o produto escalar, de fato, determina a Geometria Eucliana Plana, no sentido de que os conceitos de comprimento e ângulo podem ser introduzidos a partir deste. Na próxima seção, veremos que o produto escalar é um caso particular de uma classe de funções definidas em espaços vetoriais, chamadas de produtos internos. Tais funções, conforme constataremos, generalizam o produto escalar e, analogamente, definem geometrias nos espaços vetoriais em que estão definidas.

5.2 Espaços Vetoriais com Produto Interno

Seja V um espaço vetorial. Uma função $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ é dita um *produto interno* em V, quando é *bilinear, simétrica* e *positiva definida.* Nessa ordem, essas propriedades significam que, para quaisquer $u, v, w \in V$ e $\lambda \in \mathbb{R}$, tem-se:

- $\langle u + \lambda v, w\rangle = \langle u, w\rangle + \lambda\langle v, w\rangle$ e $\langle u, v + \lambda w\rangle = \langle u, v\rangle + \lambda\langle u, w\rangle$;
- $\langle v, w\rangle = \langle w, v\rangle$;
- $\langle v, v\rangle \geq 0$, ocorrendo a igualdade se, e só se, $v = 0$.

Um espaço vetorial V munido de um produto interno $\langle\ ,\ \rangle$ será denotado pelo par $(V, \langle\ ,\ \rangle)$.

Exemplo 5.1 (PRODUTO INTERNO CANÔNICO DE $\mathbb{R}^n$)**.** Sejam $v = (x_1, \dots, x_n)$ e $w = (y_1, \dots, y_n)$ vetores de $\mathbb{R}^n$. Verifica-se facilmente que o *produto escalar*

$$\langle v, w\rangle = x_1y_1 + \cdots + x_ny_n$$

define em $\mathbb{R}^n$ um produto interno, dito *canônico.*

Advertimos que, sempre que considerado, o espaço $\mathbb{R}^n$ estará implicitamente munido de seu produto interno canônico, salvo menção em contrário.

Exemplo 5.2. Dadas matrizes $A, B \in \mathcal{M}(m, n)$, observemos que o produto A^*B é uma matriz quadrada de ordem n. Logo, a função

$$\langle A, B\rangle = \text{traço}\,(A^*B)$$

está bem definida. Além disso, dadas matrizes $A, B, C \in \mathcal{M}(m,n)$, e $\lambda \in \mathbb{R}$, segue-se das propriedades operatórias das matrizes que

$$(A+\lambda B)^*C = (A^* + \lambda B^*)C = A^*C + \lambda B^*C.$$

Logo (vide Exercício 11),

$$\begin{aligned} \langle A+\lambda B, C\rangle &= \text{traço}\,(A^*C + \lambda B^*C) \\ &= \text{traço}\,(A^*C) + \lambda\,\text{traço}\,(B^*C) = \langle A, C\rangle + \lambda\langle B, C\rangle. \end{aligned}$$

Analogamente, verifica-se que $\langle A, B+\lambda C\rangle = \langle A, B\rangle + \lambda\langle A, C\rangle$, donde se conclui que $\langle\ ,\ \rangle$ é bilinear. Ademais,

$$\langle A, B\rangle = \text{traço}\,(A^*B) = \text{traço}\,(A^*B)^* = \text{traço}\,(B^*A) = \langle B, A\rangle,$$

isto é, $\langle\ ,\ \rangle$ é também simétrica.

Por fim, verifiquemos que $\langle\ ,\ \rangle$ é positiva definida. Para tanto, escrevamos $A = (a_{ij})$ e observemos que, nesse caso,

$$\langle A, A\rangle = \text{traço}\,(A^*A) = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2\,,$$

donde se infere que $\langle A, A\rangle > 0$, a menos que a matriz A seja nula. Portanto, $\langle\ ,\ \rangle$ é positiva definida e, consequentemente, $\langle\ ,\ \rangle$ é um produto interno em $\mathcal{M}(m,n)$.

Tomemos, por exemplo, as seguintes matrizes em $\mathcal{M}(3,2)$:

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

Nesse caso, temos que

$$\langle A, B\rangle = \text{traço}\,(A^*B) = \text{traço}\begin{bmatrix} -5 & -3 \\ 4 & -4 \end{bmatrix} = -9.$$

Exemplo 5.3. Segue-se diretamente das propriedades fundamentais da integral de funções reais que

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx, \;\; f, g \in \mathscr{C}[a, b],$$

é um produto interno em $\mathscr{C}[a, b]$.

Vejamos agora que, assim como as transformações lineares, um produto interno num espaço vetorial de dimensão finita fica determinado por seus valores numa base. Em particular, qualquer tal espaço pode ser munido de um produto interno.

Proposição 5.1. *Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ uma base de um espaço vetorial V. Dados n^2 números reais a_{ij}, $i, j \in \{1, \dots, n\}$, tais que $a_{ii} > 0$ e $a_{ij} = a_{ji}$, existe um único produto interno $\langle\, ,\, \rangle$ em V, tal que*

$$\langle v_i , v_j\rangle = a_{ij} \;\; \forall i, j \in \{1, \dots, n\}. \tag{5.2}$$

Demonstração. Considere $v = (x_1, \dots, x_n)_{\mathcal{B}}$, $w = (y_1, \dots, y_n)_{\mathcal{B}} \in V$, e defina

$$\langle v, w\rangle = \sum_{i,j=1}^{n} a_{ij}x_iy_j\, . \tag{5.3}$$

Assim, dados $u = (x'_1, \dots, x'_n)_{\mathcal{B}} \in V$ e $\lambda \in \mathbb{R}$, tem-se

$$\langle u+\lambda v, w\rangle = \sum_{i,j=1}^{n} a_{ij}(x'_i+\lambda x_i)y_j = \sum_{i,j=1}^{n} a_{ij}x'_iy_j+\lambda \sum_{i,j=1}^{n} a_{ij}x_iy_j = \langle u, w\rangle+\lambda\langle v, w\rangle.$$

De forma inteiramente análoga, verifica-se que $\langle\, ,\, \rangle$ cumpre as outras propriedades que fazem dessa função um produto interno. Além disso, segue-se diretamente de (5.3) que $\langle v_i, v_j\rangle = a_{ij}\; \forall i, j \in \{1, \dots, n\}$.

Suponhamos agora que $\langle\, ,\, \rangle$ seja um produto interno arbitrário em V, o qual satisfaz (5.2). Dados $v, w \in V$ como acima, pela bilinearidade do produto interno, temos

$$\langle v, w\rangle = \left\langle \sum_{i=1}^{n} x_iv_i, \sum_{j=1}^{n} y_jv_j \right\rangle = \sum_{i,j=1}^{n} x_iy_j\langle v_i , v_j\rangle = \sum_{i,j=1}^{n} a_{ij}x_iy_j\, ,$$

isto é, o produto interno $\langle\, ,\,\rangle$ coincide com aquele definido em (5.3), provando, assim, sua unicidade. ■

Exemplo 5.4. Seja $\mathcal{B} = \{e_1\,, e_2\}$ a base canônica de $\mathbb{R}^2$. Dados $a_{ij} = i + j$, com $i, j \in \{1, 2\}$, o produto interno $\langle\, ,\,\rangle$ de $\mathbb{R}^2$ que satisfaz $\langle e_i, e_j\rangle = a_{ij}$ é dado por:

$$\langle v, w\rangle = \sum_{i,j=1}^{2} a_{ij} x_i y_j = 2x_1y_1 + 3x_1y_2 + 3x_2y_1 + 4x_2y_2\,,$$

em que $v = (x_1, x_2)$ e $w = (y_1, y_2)$.

Seja V um espaço vetorial munido de um produto interno $\langle\, ,\,\rangle$. A *norma* de um vetor $v \in V$ determinada por $\langle\, ,\,\rangle$ é o número real $\|v\|$ definido por

$$\|v\| = \sqrt{\langle v, v\rangle}. \tag{5.4}$$

A partir da norma, definimos a *distância* entre dois vetores $v, w \in V$ por

$$\operatorname{dist}(v, w) = \|v - w\|. \tag{5.5}$$

Observemos que, no espaço $\mathbb{R}^n$ (munido do produto interno canônico), a distância entre dois vetores $v = (x_1, \ldots, x_n)$ e $w = (y_1, \ldots, y_n)$ corresponde à distância euclidiana usual em $\mathbb{R}^n$, isto é,

$$\operatorname{dist}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

Assim, os vetores $v = (1, -1, 2)$ e $w = (2, 0, -1)$ de $\mathbb{R}^3$ têm normas $\|v\| = \sqrt{6}$ e $\|w\| = \sqrt{5}$, e a distância entre os mesmos é

$$\operatorname{dist}(v, w) = \|v - w\| = \|(-1, -1, 3)\| = \sqrt{11}.$$

Exemplo 5.5. Sejam $A, B \in \mathcal{M}(3, 2)$ as matrizes do Exemplo 5.2. Temos que

$$\|A\|^2 = \langle A, A\rangle = \text{traço}\,(A^*A) = \text{traço}\begin{bmatrix} 18 & 6 \\ 6 & 5 \end{bmatrix} = 23,$$

donde $\|A\| = \sqrt{23}$. Analogamente, obtém-se $\|B\| = \sqrt{10}$.

A relação fundamental entre o produto interno e a norma por ele definida expressa-se através da célebre Desigualdade de Cauchy-Schwarz, a qual estabeleceremos a seguir.

Teorema 5.1 (DESIGUALDADE DE CAUCHY-SCHWARZ)**.** *Seja V um espaço vetorial munido de um produto interno $\langle\, ,\, \rangle$. Então,*

$$|\langle v, w\rangle| \leq \|v\|\|w\| \;\; \forall v, w \in V, \tag{5.6}$$

ocorrendo a igualdade se, e somente se, v e w são múltiplos um do outro.

Demonstração. O resultado é imediato quando v ou w é o vetor nulo. Suponhamos, então, que v e w sejam ambos não nulos e definamos a função:

$$f(t) = \langle v - tw, v - tw\rangle, \;\; t \in \mathbb{R}.$$

Temos que $f(t) \geq 0 \, \forall t \in \mathbb{R}$, e que $f(t) = 0$ para algum $t \in \mathbb{R}$ se, e somente se, $v = tw$, pois o produto interno é positivo definido. Além disso, pela bilinearidade do produto interno,

$$f(t) = \|v\|^2 - 2\langle v, w\rangle t + \|w\|^2 t^2,$$

isto é, f é uma função quadrática não negativa de variável t. Em particular, o discriminante de f é não positivo. Logo,

$$4\langle v, w\rangle - 4\|v\|^2\|w\|^2 \leq 0,$$

o que implica (5.6). Temos também que o discriminante de f é nulo se, e somente se, f possui um único zero, donde se infere que ocorre a igualdade em (5.6) se, e somente se, $v = tw$ para algum $t \in \mathbb{R}$. ■

A desigualdade de Cauchy-Schwarz tem inúmeras consequências. A primeira que destacaremos é a prova do cumprimento da desigualdade triangular por normas advindas de produtos internos, conforme a proposição seguinte.

Proposição 5.2 (PROPRIEDADES FUNDAMENTAIS DA NORMA)**.** *A norma $\|\;\|$ advinda de um produto interno $\langle\, ,\, \rangle$ de um espaço vetorial V tem as seguintes propriedades, as quais são válidas para quaiquer $v, w \in V$ e $\lambda \in \mathbb{R}$:*

i) $\|v\| \geq 0$ *e* $\|v\| = 0 \Leftrightarrow v = 0$ *(positividade);*

ii) $\|\lambda v\| = |\lambda|\|v\|$ *(homogeneidade);*

iii) $\|v + w\| \leq \|v\| + \|w\|$ *(desigualdade triangular).*

Demonstração. As propriedades (i) e (ii) decorrem imediatamente da definição de norma e da positividade e bilinearidade do produto interno. Já a desigualdade triangular, ela decorre da Desigualdade de Cauchy-Schwarz. De fato, dados vetores $v, w \in V$, temos $\|v + w\|^2 = \langle v + w, v + w \rangle$. Logo,

$$\|v + w\|^2 = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2,$$

donde se obtém (iii). ■

A desigualdade de Cauchy-Schwarz também nos permite introduzir a noção de ângulo entre vetores em qualquer espaço vetorial V munido de um produto interno $\langle\ ,\ \rangle$. Mais precisamente, dados vetores não nulos $v, w \in V$, definimos o *ângulo* $\sphericalangle(v, w) \in [0, \pi]$ entre v e w por (compare com (5.1)):

$$\sphericalangle(v, w) := \arccos \frac{\langle v, w \rangle}{\|v\|\,\|w\|} \cdot \tag{5.7}$$

Note que, pela Desigualdade de Cauchy-Schwarz,

$$-1 \leq \frac{\langle v, w \rangle}{\|v\|\,\|w\|} \leq 1,$$

de modo que $\sphericalangle(v, w)$ está bem definido.

Calculemos o ângulo entre os vetores $v = (1, -1, 2)$ e $w = (2, 0, -1)$ de $\mathbb{R}^3$, considerados acima. Temos que,

$$\sphericalangle(v, w) = \arccos \frac{\langle v, w \rangle}{\|v\|\,\|w\|} = \arccos \frac{0}{\|\sqrt{6}\|\,\|\sqrt{5}\|} = 0,$$

isto é, $\sphericalangle(v, w) = \pi/2$.

Exemplo 5.6. Tomemos, uma vez mais, as matrizes A e B do Exemplo 5.2. Considerando-se também os resultados do Exemplo 5.5, temos que

$$\sphericalangle(A,B) = \arccos \frac{\langle A,B\rangle}{\|A\|\,\|B\|} = \arccos \frac{-9}{\sqrt{23}\sqrt{10}} \approx \frac{7\pi}{10}\,.$$

A desigualdade de Cauchy-Schwarz pode ser aplicada para se obter desigualdades numéricas não triviais, conforme ilustrado no exemplo a seguir.

Exemplo 5.7. Verifiquemos que, para quaisquer reais $a, b \geq 1$, tem-se

$$\sqrt{a^2-1} + \sqrt{b^2-1} \leq ab.$$

Com efeito, tomando-se em $\mathbb{R}^2$ o produto interno canônico e aplicando-se a desigualdade de Cauchy-Schwarz aos vetores $v = (\sqrt{a^2-1}, 1)$ e $w = (1, \sqrt{b^2-1})$, obtém-se

$$\sqrt{a^2-1} + \sqrt{b^2-1} = \langle v,w\rangle \leq \|v\|\|w\| = ab.$$

5.3 Ortogonalidade

Em Geometria Euclidiana Plana, o conceito de ortogonalidade está envolvido em vários resultados fundamentais dessa teoria, notadamente, no Teorema de Pitágoras. Considerando-se (5.7), este conceito se estende naturalmente ao contexto dos espaços vetoriais, conforme a definição seguinte.

Definição 5.1 (ORTOGONALIDADE)**.** Seja $(V, \langle\ ,\ \rangle)$ um espaço vetorial com produto interno $\langle\ ,\ \rangle$. Dizemos que $v, w \in V$ são *ortogonais*, quando $\langle v,w\rangle = 0$.

Dados vetores $v, w \in (V, \langle\ ,\ \rangle)$, tem-se

$$\|v+w\|^2 = \langle v+w, v+w\rangle = \|v\|^2 + 2\langle v,w\rangle + \|w\|^2.$$

Vale, portanto, o seguinte resultado.

Teorema 5.2. *Dois vetores $v, w \in (V, \langle\ ,\ \rangle)$ são ortogonais se, e somente se, cumprem a igualdade $\|v+w\|^2 = \|v\|^2 + \|w\|^2$.*

Note que a parte “somente se” do teorema acima constitui uma versão do Teorema de Pitágoras para espaços vetoriais, já que, em $\mathbb{R}^2$, quando v e w são ortogonais, v, w, e $v+w$ são os lados (ou vértices, se vistos como pontos) de um triângulo retângulo.

Definição 5.2 (ORTOGONALIDADE DE CONJUNTOS)**.** Um subconjunto de $(V, \langle\ ,\ \rangle)$ é dito *ortogonal*, quando seus elementos são, dois a dois, ortogonais. Um subconjunto ortogonal cujos vetores são todos unitários (isto é, de norma 1) é dito *ortonormal*.

Observe que todo conjunto ortogonal $\mathcal{C} = \{v_1, \dots, v_k\} \subset (V, \langle\ ,\ \rangle)$ cujos vetores são não nulos é LI. A fim de verificarmos essa propriedade, consideremos uma combinação linear nula desses vetores, isto é,

$$\lambda_1 v_1 + \cdots + \lambda_i v_i + \cdots + \lambda_k v_k = \mathbf{0}.$$

Para um dado $i \in \{1, \dots, k\}$, tomando-se o produto interno por v_i em ambos os membros da igualdade acima, obtém-se $\lambda_i \langle v_i, v_i\rangle = 0$, o que nos dá $\lambda_i = 0$, provando, assim, a independência linear de $\mathcal{C}$.

No que se segue, através de um bem conhecido método denominado *ortogonalização de Gram-Schmidt*, mostraremos que todo espaço vetorial $(V, \langle\ ,\ \rangle)$ de dimensão finita admite uma base ortonormal. O fundamento desse método é o conceito de *projeção ortogonal*, que ora introduzimos.

Definição 5.3 (PROJEÇÃO ORTOGONAL)**.** Dados $v, w \in (V, \langle\ ,\ \rangle)$, $w \neq \mathbf{0}$, o vetor

$$P_w(v) = \frac{\langle v, w\rangle}{\langle w, w\rangle} w$$

denomina-se a *projeção ortogonal* de v sobre (o espaço gerado por) w.

O motivo da nomenclatura, na definição acima, reside no fato de o vetor $v - P_w(v)$ ser ortogonal a w (vide Fig. 5.2). Com efeito,

$$\langle v - P_w(v), w\rangle = \langle v, w\rangle - \langle P_w(v), w\rangle = \langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, w\rangle = 0.$$

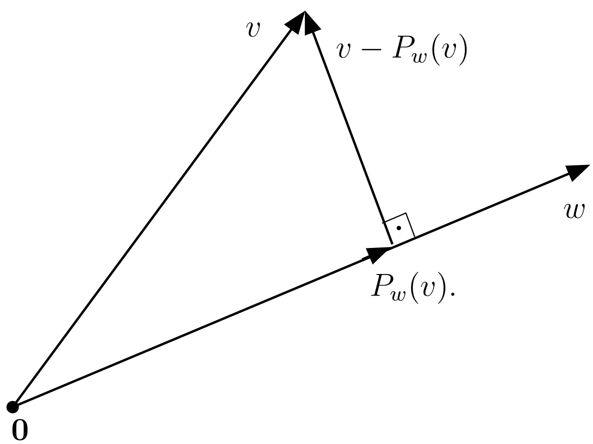


Figura 5.2: Projeção de v sobre Ger $\{w\}$.

Acrescentamos, ainda, que a sentença entre parênteses deve-se à igualdade

$$P_{\lambda w}(v) = P_w(v) \ \ \forall \lambda \neq 0,$$

a qual é de verificação imediata.

Observemos que, nas condições do parágrafo anterior, se v e w forem LI, então $\{v - P_v(w), w\}$ será uma base ortogonal do subespaço Ger $\{v, w\} \subset V$. Nesse caso, multiplicando-se cada um desses vetores pelos inversos de suas respectivas normas, obtém-se uma base ortonormal $\mathcal{B}$ de Ger $\{v, w\}$. Em suma, a partir da base $\{v, w\}$ de Ger $\{v, w\}$, obtivemos uma base ortogonal desse subespaço substituindo-se o vetor v por $v - P_w(v)$. Em seguida, "normalizamos" essa base ortogonal a fim de obter a base ortonormal $\mathcal{B}$.

Exemplo 5.8. Consideremos em $\mathbb{R}^3$ os vetores $v = (1, 0, -1)$ e $w = (1, 1, 0)$ e apliquemos ortogonalização de Gram-Schmidt para obter uma base ortonormal de Ger $\{v, w\}$. Temos que $\langle v, w\rangle / \langle w, w\rangle = 1/2$, donde

$$P_w(v) = \frac{\langle v, w\rangle}{\langle w, w\rangle} w = (1/2, 1/2, 0)$$

e, portanto, $v - P_w(v) = (1/2, -1/2, -1)$. Logo, $\{(1/2, -1/2, -1), (1, 1, 0)\}$ é uma base ortogonal de Ger $\{v, w\}$. Normalizando-a, obtemos a base ortonormal

$$\mathcal{B} = \{(\sqrt{6}/6, -\sqrt{6}/6, -\sqrt{2}/\sqrt{3}), (\sqrt{2}/2, \sqrt{2}/2, 0)\}.$$

Procedendo-se indutivamente, podemos aplicar ortogonalização de Gram-Schmidt a um conjunto arbitrário de vetores LI em $(V, \langle\ ,\ \rangle)$. Em particular, a uma base. Para constatarmos isso, basta observarmos que, se $\{v_1, \dots, v_k\}$ é uma base ortogonal de um subespaço W de V, e $v \in V - W$, então o vetor

$$w_k = \lambda_1 v_1 + \cdots + \lambda_k v_k, \quad \lambda_i = \frac{\langle v_i, v\rangle}{\langle v_i, v_i\rangle},$$

satisfaz a igualdade (verifique!)

$$\langle v - w_k, v_i\rangle = 0 \ \forall i \in 1, \dots, k.$$

Logo, $\{v_1, \dots, v_{k-1}, v - w_k\}$ é uma base ortogonal do subespaço de V, o qual é gerado por esses vetores.

Das considerações acima, segue-se o seguinte resultado.

Proposição 5.3 (ORTOGONALIZAÇÃO DE GRAM-SCHMIDT)**.** *Seja* $\{v_1, \dots, v_n\}$ *uma base de um espaço vetorial* $(V, \langle\ ,\ \rangle)$*. Então, o conjunto* $\{v_1', \dots, v_n'\} \subset V$ *definido pelas relações de recorrência:*

$$v_1' = v_1$$
$$v_{i+1}' = v_{i+1} - \sum_{k=1}^{i} \frac{\langle v_{i+1}, v_i\rangle}{\langle v_i, v_i\rangle} v_i, \ i \in \{1, \dots, n-1\},$$

constitui uma base ortogonal de V*. Em particular, o conjunto*

$$\mathcal{B} = \left\{ \frac{v_1'}{\|v_1'\|}, \dots, \frac{v_n'}{\|v_n'\|} \right\} \subset V$$

é uma base ortonormal de V*.*

Seja $\mathcal{B} = \{u_1, \dots, u_n\}$ uma base ortonormal de $(V, \langle\ ,\ \rangle)$. Dados vetores $v = (x_1, \dots, x_n)_{\mathcal{B}}$, $w = (y_1, \dots, y_n)_{\mathcal{B}} \in V$, temos que

$$\langle v, w\rangle = \langle x_1 u_1 + \cdots + x_n u_n, y_1 u_1 + \cdots + y_n u_n\rangle = x_1 y_1 + \cdots + x_n y_n,$$

isto é, em coordenadas com respeito a uma base ortonormal, o produto interno em $(V, \langle\ ,\ \rangle)$ expressa-se como o produto interno canônico de $\mathbb{R}^n$. Deste simples fato, segue-se que todo funcional linear $f \in (V, \mathbb{R})$ exprime-se como um produto interno por um vetor fixo, conforme o teorema seguinte.

Teorema de Riesz. Seja $(V, \langle\ ,\ \rangle)$ um espaço vetorial de dimensão finita munido de um produto interno $\langle\ ,\ \rangle$. Então, para todo funcional linear $f \in \mathscr{L}(V, \mathbb{R})$, existe um único vetor $a \in V$, tal que

$$f(v) = \langle a, v\rangle \, \forall v \in V.$$

Demonstração. Tome uma base ortornormal $\mathcal{B} = \{u_1\,, \dots, u_n\}$, de V, e defina o vetor $a = f(u_1)u_1 + \cdots + f(u_n)u_n$. Assim, dado $v = x_1u_1 + \cdots + x_nu_n \in V$, tem-se $f(v) = x_1f(u_1) + \cdots + x_nf(u_n) = \langle a, v\rangle$.

Suponhamos, agora, que exista $b \in V$, tal que $f(v) = \langle b, v\rangle \, \forall v \in V$. Nesse caso, para todo $v \in V$, teremos $\langle a, v\rangle = \langle b, v\rangle$, isto é,

$$\langle a - b, v\rangle = 0 \ \forall v \in V.$$

Fazendo-se, na igualdade acima, $v = a - b$, conclui-se imediatamente que $a = b$ e, portanto, que a é único. ■

A fim de estender o conceito de projeção ortogonal, tomemos um vetor $v \in (V, \langle\ ,\ \rangle)$ e um subespaço não trivial $W \subset V$. Verifiquemos, então, que existe um único vetor $w \in W$, tal que $v - w$ é ortogonal a todos os vetores de W.

Para tanto, valendo-nos da Proposição 5.3, fixemos uma base ortonormal $\{u_1, \dots, u_k\}$ de W. Definindo-se, então,

$$w = \sum_{i=1}^{k} \langle v, u_i\rangle u_i\,, \tag{5.8}$$

verifica-se facilmente que $\langle v - w, u_i\rangle = 0$ para todo $i = 1, \dots, k$. Logo, pela bilinearidade do produto interno, temos que $v-w$ é ortogonal a todos os vetores de W.

Suponhamos agora que exista um vetor $w' \in W$, tal que $v-w'$ seja ortogonal a todos os vetores de W. Nesse caso, escrevendo-se $w' = \lambda_1u_1 + \cdots + \lambda_ku_k$, para cada $i = 1, \dots, k$, tem-se

$$0 = \langle v - w', u_i\rangle = \langle v, u_i\rangle - \langle w', u_i\rangle = \langle v, u_i\rangle - \lambda_i\,,$$

donde $\lambda_i = \langle v, u_i \rangle$ e, portanto, $w' = w$.

O vetor w em (5.8) é dito a *projeção ortogonal* de v sobre W, e é denotado por $P_W(v)$. Dessa forma, fica definida a *aplicação projeção ortogonal* de V em W, $P_W : V \to W$, dada por

$$P_W(v) = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_k \rangle u_k\,, \;\; v \in V.$$

Note que, pela bilinearidade do produto interno, P_W é linear! Além disso, essa aplicação tem as seguintes propriedades, cuja verificação deixamos a cargo do leitor (vide Exercício 14).

- $P_W \circ P_W = P_W$ (*idempotência*);
- $\langle P_W(v), v' \rangle = \langle v, P_W(v') \rangle \; \forall v, v' \in V$ (*simetria*);
- $\langle P_W(v), v \rangle \geq 0 \; \forall v \in V$ (*positividade*);
- $\|P_W(v)\| \leq \|v\| \; \forall v \in V$ e $\|P_W(v)\| = \|v\| \Leftrightarrow v \in V$ (*semi-contratilidade*).

O conceito de ortogonalidade nos permite decompor um espaço vetorial $(V, \langle\,,\,\rangle)$ numa soma direta de subespaços, os quais são ortogonais entre si. Isto envolve o seguinte conceito.

Definição 5.4 (COMPLEMENTO ORTOGONAL)**.** Dado um subespaço vetorial W de $(V, \langle\,,\,\rangle)$, chama-se o conjunto

$$W^{\perp} = \{v \in V\,;\, \langle v, w \rangle = 0 \; \forall w \in W\}$$

de *complemento ortogonal* de W em V.

Proposição 5.4 (PROPRIEDADES DO COMPLEMENTO ORTOGONAL)**.** *O complemento ortogonal $W^{\perp}$ de um subespaço $W \subset (V, \langle\,,\,\rangle)$ é um subespaço vetorial de V, o qual tem as seguintes propriedades:*

i) $V = W \oplus W^{\perp}$;

ii) $W^{\perp\perp} = W$.

Demonstração. É imediato que $\mathbf{0} \in W^\perp$. Além disso, dados $v, v' \in W^\perp$, $\lambda \in \mathbb{R}$, e $w \in W$, tem-se $\langle v + \lambda v', w\rangle = \langle v, w\rangle + \lambda\langle v', w\rangle = 0$, donde se infere que $W^\perp$ é um subespaço vetorial de V.

A fim de provar as asserções (i) e (ii), tomemos uma base ortonormal de W, $\{u_1, \ldots, u_k\}$. Usando ortogonalização de Gram-Schmidt, essa base pode ser completada para uma base ortonormal de V, $\{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\}$. Assim, pondo-se $V' = \operatorname{Ger}\{u_{k+1}, \ldots, u_n\}$, tem-se $V = W \oplus V'$. Resta-nos, pois, mostrar que $V' = W^\perp$. A inclusão $V' \subset W^\perp$ é imediata. Ademais, se $v \in W^\perp$, tem-se $\langle v, u_i\rangle = 0\, \forall i = 1, \ldots, k$, donde $v \in V'$. Logo, $V' = W^\perp$, o que prova (i).

Quanto à afirmação (ii), observemos que, pelo estabelecido em (i),

$$W \oplus W^\perp = V = W^\perp \oplus W^{\perp\perp},$$

o que nos dá $\dim W^{\perp\perp} = \dim W$ (vide Proposição 2.16). Além disso, pelo provado em (i), dado $v \in W^{\perp\perp}$, existem únicos $w \in W$ e $w^\perp \in W^\perp$, tais que $v = w + w^\perp$. Porém, $0 = \langle v, w^\perp\rangle = \langle w^\perp, w^\perp\rangle$, donde $w^\perp = 0$, isto é, $v = w$. Segue-se que $W^{\perp\perp} \subset W$ e, portanto, que $W = W^{\perp\perp}$, uma vez que esses subespaços têm dimensões iguais. ■

5.4 Volume em Espaços Euclidianos

No que diz respeito a medição de objetos do espaço euclidiano $\mathbb{R}^3$, as "medidas" fundamentais são: comprimento, área e volume. Através da teoria que desenvolvemos neste capítulo e no anterior, podemos estender essas noções para o espaço $\mathbb{R}^n$, isto é, para todo $k \in \{1, \ldots, n\}$, podemos definir um "volume k-dimensional" de subconjuntos de $\mathbb{R}^n$.

A introdução de um conceito de volume para subconjuntos arbitrários de $\mathbb{R}^n$ é uma questão bastante delicada e não será tratada neste texto. Na verdade, esse problema é o objeto da uma vasta teoria, dita, *da medida.* No entanto, o primeiro passo na direção da construção do conceito geral de volume, o que constitui aqui nosso propósito, é o de estabelecê-lo para paralelepípedos de $\mathbb{R}^n$.

Assim como o fizemos com as noções de comprimento e ângulo, usaremos a Geometria Euclidiana Plana como guia. Mais precisamente, consideraremos o conceito de área de paralelogramo à luz da Álgebra Linear.

Com esse intuito, tomemos vetores linearmente independentes

$$v = (a, b) \text{ e } w = (c, d) \in \mathbb{R}^2,$$

denotemos por $\mathscr{P}(v, w)$ o paralelogramo de lados v e w, e provemos que

$$|\mathscr{P}(v, w)| = |\det(v, w)|, \tag{5.9}$$

em que $|\mathscr{P}(v, w)|$ denota a área de $\mathscr{P}(v, w)$, a ser entendida aqui como o produto do comprimento de um de seus lados, dito a base, pela altura correspondente. Tomando-se w como base de $\mathscr{P}(v, w)$, a altura será a norma da projeção de v sobre o vetor $w^\perp = (-d, c)$, pois $\langle w^\perp, w\rangle = 0$ (Fig. 5.4).

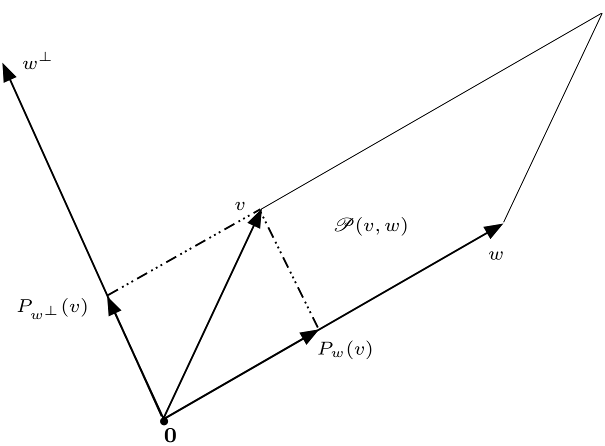


Figura 5.3

Logo, uma vez que $\|w\| = \|w^\perp\|$ e $\langle w^\perp, v\rangle = -ad + bc = -\det(v, w)$, tem-se

$$|\mathscr{P}(v, w)| = \|w\| \, \|P_{w^\perp}(v)\| = \|w\| \frac{|\langle v, w^\perp\rangle|}{\langle w^\perp, w^\perp\rangle} \|w^\perp\| = |\det(v, w)|.$$

Uma outra expressão da área de $\mathscr{P}(v, w)$ pode ser obtida a partir da igualdade $v = P_w(v) + P_{w^\perp}(v)$. De fato, pelo Teorema de Pitágoras, tem-se $\|v\|^2 = \|P_w(v)\|^2 + \|P_{w^\perp}(v)\|^2$. Desta forma,

$$\|v\|^2 = \frac{\langle v, w\rangle^2}{\langle w, w\rangle^2}\|w\|^2 + \frac{\langle v, w^\perp\rangle^2}{\langle w^\perp, w^\perp\rangle^2}\|w^\perp\|^2 = \frac{\langle v, w\rangle^2}{\langle w, w\rangle} + \frac{|\det(v,w)|^2}{\langle w, w\rangle}\,.$$

Daí e de (5.9), obtém-se

$$|\mathscr{P}(v,w)|^2 = \langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle^2 = \det\begin{bmatrix} \langle v, v\rangle & \langle v, w\rangle \\ \langle w, v\rangle & \langle w, w\rangle \end{bmatrix}.$$

A matriz na última igualdade acima é dita a matriz *de Gram* dos vetores v e w, e seu determinante é denotado por $g(v,w)$. Assim, temos

$$|\mathscr{P}(v,w)| = \sqrt{g(v,w)}. \tag{5.10}$$

Note que podemos expressar o comprimento de um vetor $v \in \mathbb{R}^2$ através da matriz de Gram 1×1, $[\langle v, v\rangle]$, uma vez que $\|v\| = \sqrt{\langle v, v\rangle} = \sqrt{g(v)}$. Convém mencionar também que podemos descrever $\mathscr{P}(v,w)$ como

$$\mathscr{P}(v,w) = \{\lambda v + \mu w\,;\ \lambda, \mu \in [0,1]\}.$$

Exemplo 5.9. Dados $v = (3,-1), w = (2,2) \in \mathbb{R}^2$, temos

$$|\mathscr{P}(v,w)| = |\det(v,w)| = 8.$$

Observemos também que

$$g(v,w) = \det\begin{bmatrix} \langle v, v\rangle & \langle v, w\rangle \\ \langle w, v\rangle & \langle w, w\rangle \end{bmatrix} = \det\begin{bmatrix} 10 & 4 \\ 4 & 8 \end{bmatrix} = 64 = |\mathscr{P}(v,w)|^2,$$

confirmando, assim, a igualdade (5.10).

Mais geralmente, o *paralelepípedo k-dimensional* de $\mathbb{R}^n$ determinado por k vetores linearmente independentes, $v_1, \dots, v_k$, é definido por:

$$\mathscr{P}(v_1, \dots, v_k) = \{\lambda_1 v_1 + \cdots + \lambda_k v_k\ ;\ \lambda_i \in [0,1]\ \forall i = 1, \dots, k\}.$$

Pondo-se $W_{k-1} = \text{Ger}\,\{v_1\,, \ldots, v_{k-1}\}$, o *volume k-dimensional* do paralelepípedo $\mathscr{P}(v_1\,, \ldots, v_k)$, a ser denotado por $|\mathscr{P}(v_1\,, \ldots, v_k)|$, é definido indutivamente da seguinte forma (Fig. 5.4):

$$|\mathscr{P}(v_1\,, \ldots, v_k)| \;=\; \begin{cases} \|v_1\|\,, & \text{se } k = 1. \\ |\mathscr{P}(v_1\,, \ldots, v_{k-1})|\,\|v_k - P_{W_{k-1}}(v_k)\|, & \text{se } k > 1. \end{cases}$$

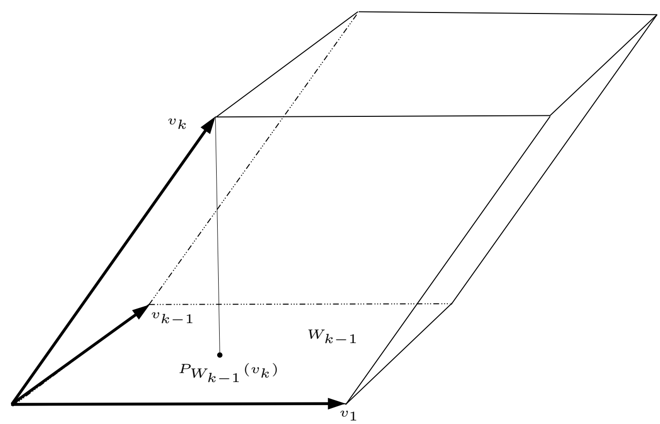


Figura 5.4: Paralelepípedo $\mathscr{P}(v_1\,, \ldots, v_k)$.

Consideremos, então, a *matriz de Gram* de $v_1\,, \ldots, v_k$,

$$G(v_1, \ldots, v_k) = (\langle v_i, v_j\rangle)_{k\times k}\ ,$$

e denotemos seu determinante por $g(v_1\,, \ldots, v_k)$.

Proposição 5.5. *Para todo paralelepípedo $\mathscr{P}(v_1, \ldots, v_k) \subset \mathbb{R}^n$, vale a igualdade*

$$|\mathscr{P}(v_1\,, \ldots, v_k)| \;=\; \sqrt{g(v_1\,, \ldots, v_k)}. \tag{5.11}$$

Demonstração. Temos que a igualdade (5.11) é trivial para $k = 1$. Suponhamos, por indução, que ela seja verdadeira para $k \in \mathbb{N}$. Pondo-se, então, $P_{W_k}(v_{k+1}) = \lambda_1 v_1 + \cdots + \lambda_k v_k$ e $w_{k+1} = v_{k+1} - P_{W_k}(v_{k+1})$, tem-se

$$v_{k+1} = \sum_{j=1}^{k} \lambda_j v_j + w_{k+1}\,.$$

Assim, denotando-se por $c_1, \dots, c_{k+1}$ os vetores coluna da matriz de Gram $G(v_1, \dots, v_{k+1})$, tem-se

$$
\begin{aligned}
c_{k+1} &= \sum_{i=1}^{k+1} \langle v_{k+1}, v_i \rangle e_i = \sum_{i=1}^{k+1} \left\langle \sum_{j=1}^{k} \lambda_j v_j + w_{k+1}, v_i \right\rangle e_i \\
&= \sum_{i=1}^{k+1} \left\langle \sum_{j=1}^{k} \lambda_j v_j, v_i \right\rangle e_i + \sum_{i=1}^{k+1} \langle w_{k+1}, v_i \rangle \, e_i \\
&= \sum_{j=1}^{k} \sum_{i=1}^{k+1} \langle \lambda_j v_j, v_i \rangle e_i + \langle w_{k+1}, v_{k+1} \rangle e_{k+1} \\
&= \sum_{j=1}^{k} \lambda_j c_j + \|w_{k+1}\|^2 e_{k+1}.
\end{aligned}
$$

Daí, observando-se que $\det(c_1, \dots, c_k, e_{k+1}) = g(v_1, \dots, v_k)$, e lembrando-se que det é uma forma multilinear alternada, obtém-se

$$
\begin{aligned}
g(v_1, \dots, v_{k+1}) &= \det(c_1, \dots, c_k, c_{k+1}) \\
&= \det\left(c_1, \dots, c_k, \sum_{j=1}^{k} \lambda_j c_j + \|w_{k+1}\|^2 e_{k+1} \right) \\
&= \|w_{k+1}\|^2 g(v_1, \dots, v_k) = \|w_{k+1}\|^2 |\mathscr{P}(v_1, \dots, v_k)|^2 \\
&= |\mathscr{P}(v_1, \dots, v_{k+1})|^2.
\end{aligned}
$$

O resultado, portanto, segue-se do Princípio da Indução. ■

Note que, na Proposição 5.5, provamos também que $g(v_1, \dots, v_k) > 0$ sempre que $v_1, \dots, v_k$ são LI.

Corolário 5.1. *Para todo paralelepípedo n-dimensional* $\mathscr{P}(v_1, \dots, v_n) \subset \mathbb{R}^n$, *vale a seguinte igualdade:*

$$
|\mathscr{P}(v_1, \dots, v_n)| = |\det(v_1, \dots, v_n)|. \tag{5.12}
$$

Demonstração. Basta observarmos que, se A é a matriz $n \times n$ cujos vetores linha são $v_1, \ldots, v_n$, então $G(v_1, \ldots, v_n) = AA^*$ (vide Exercício 12). Logo,

$$|\mathscr{P}(v_1, \ldots, v_n)|^2 = \det G(v_1, \ldots, v_n) = (\det A)^2,$$

donde se obtém (5.12). ■

Exemplo 5.10. Consideremos os vetores de $\mathbb{R}^3$, $v_1 = (1, 0, 1)$, $v_2 = (2, -1, 1)$ e $v_3 = (1, 1, -1)$. Um cálculo direto nos dá $g(v_1, v_2) = 3$, $g(v_1, v_3) = 6$ e $g(v_2, v_3) = 18$. Logo, as áreas dos correspondentes paralelogramos são:

$$|\mathscr{P}(v_1, v_2)| = \sqrt{3}, \;\; |\mathscr{P}(v_1, v_3)| = \sqrt{6} \;\text{ e }\; |\mathscr{P}(v_1, v_2)| = 3\sqrt{2}.$$

Quanto ao volume do paralelepípedo determinado por v_1, v_2 e v_3, temos

$$|\mathscr{P}(v_1, v_2, v_3)| = |\det(v_1, v_2, v_3)| = |3| = 3.$$

5.5 Exercícios

Seção 5.2

1. Sejam $\langle\, , \rangle_1$ e $\langle\, , \rangle_2$ produtos internos num espaço vetorial V e λ um número real positivo. Mostre que

$$\langle v, w\rangle = \langle v, w\rangle_1 + \lambda\langle v, w\rangle_2, \;\; v, w \in V,$$

é um produto interno em V.

2. Seja $T : V \to (W, \langle\, , \rangle)$ uma aplicação linear injetiva. Mostre que:

   i) A função $\langle u, v\rangle_0 = \langle T(u), T(v)\rangle$, $u, v \in V$, define um produto interno $\langle\, , \rangle_0$ em V, dito *induzido* por T.

   ii) O produto interno definido em $\mathcal{M}(m, n)$ no Exemplo 5.2 é aquele induzido pelo isomorfismo natural $T : \mathcal{M}(m, n) \to (\mathbb{R}^{mn}, \langle\, , \rangle)$, em que $\langle\, , \rangle$ é o produto interno canônico de $\mathbb{R}^{mn}$.

3. Mostre que a norma $\| \ \|$ de um espaço vetorial $(V, \langle \ , \ \rangle)$ tem as seguintes propriedades, válidas para quaisquer $v, w \in V$:

   i) $\|v+w\|^2+\|v-w\|^2 = 2(\|v\|^2+\|w\|^2)$ (*identidade do paralelogramo*);

   ii) $|\,\|v\| - \|w\|\,| \leq \|v - w\|$ (*segunda desigualdade triangular*).

4. Mostre que, para todo espaço vetorial $(V, \langle \ , \ \rangle)$, vale a seguinte *identidade de polarização:*

$$\langle v, w \rangle = \frac{1}{2}\left(\|v\|^2 + \|w\|^2 - \|v - w\|^2\right) \ \forall v, w \in V,$$

   isto é, todo produto interno é determinado pela norma a ele associada.

5. Mostre que a função distância dist definida em (5.5) tem as seguintes propriedades, válidas para quaisquer $u, v, w \in (V, \langle \ , \ \rangle)$:

   i) $\operatorname{dist}(v, w) \geq 0$ e $\operatorname{dist}(v, w) = 0 \ \Leftrightarrow v = w$ (*positividade*);

   ii) $\operatorname{dist}(v, w) = \operatorname{dist}(w, v)$ (*simetria*);

   iii) $\operatorname{dist}(u, w) \leq \operatorname{dist}(u, v) + \operatorname{dist}(v, w)$ (*desigualdade triangular*).

6. Dado $c \in \mathbb{R}$, considere a função $f(x) = (x + c)^2/(x^2 + 1)$, $x \in \mathbb{R}$. Use a desigualdade de Cauchy-Schwarz para provar que

$$f(x) \leq (1 + c)^2 \ \forall x \in \mathbb{R}.$$

7. Sejam u e v vetores LI num espaço vetorial $(V, \langle \ , \ \rangle)$. Determine o vetor w de Ger $\{u, v\}$, o qual é paralelo à bissetriz do ângulo $\sphericalangle(u, v)$, isto é, w é tal que $\sphericalangle(u, w) = \sphericalangle(w, v)$.

8. Considere o exercício anterior e determine o incentro (ponto de encontro das bissetrizes) do triângulo de $\mathbb{R}^2$ de vértices $v_1 = (0, 0)$, $v_2 = (1, 1)$ e $v_3 = (-2, 3)$.

Seção 5.3

9. Seja $\{v_1, \ldots, v_k\}$ uma base ortornormal de um subespaço $W \subset (V, \langle\ ,\ \rangle)$. Prove a *desigualdade de Bessel*:

$$\sum_{i=1}^{k} \langle v, v_i \rangle^2 \leq \|v\|^2 \ \ \forall v \in V.$$

10. Mostre que o operador linear $J : \mathbb{R}^2 \to \mathbb{R}^2$ definido por

$$J(x_1, x_2) = (-x_2, x_1)$$

tem as seguintes propriedades, as quais são válidas para quaisquer vetores $v, w \in \mathbb{R}^2$.

   i) $\langle J(v), v \rangle = 0$ (*J é a rotação positiva de ângulo $\pi/2$*);
   ii) $\langle J(v), J(w) \rangle = \langle v, w \rangle$ (*J preserva produto interno*);
   iii) $\|J(v)\| = \|v\|$ (*J preserva norma*).

11. Sejam W_1 e W_2 subespaços de $(V, \langle\ ,\ \rangle)$. Mostre que:

   i) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$;
   ii) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

12. Seja A uma matriz $m \times n$, $v_1, \ldots, v_m \in \mathbb{R}^n$ seus vetores linha, e W o espaço gerado por $v_1, \ldots, v_m$. Dado $v \in \mathbb{R}^n$, mostre que:

$$A[v] = \begin{bmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_m \rangle \end{bmatrix}.$$

Conclua que são válidas as seguintes igualdades:

   i) $AA^* = (\langle v_i, v_j \rangle)_{m \times m}$.
   ii) $\mathrm{Nuc}\,(T_A) = W^\perp$.

13. Seja $\mathcal{B} = \{u_1, \ldots, u_n\}$ uma base de $(V, \langle\ ,\ \rangle)$. Suponha que, para todo $v = (x_1, \ldots, x_n)_\mathcal{B} \in V$, tenha-se $\|v\|^2 = x_1^2 + \cdots + x_n^2$. Mostre que $\mathcal{B}$ é ortonormal.

14. Verifique as propriedades da projeção ortogonal $P_W : V \to W$ listadas na Seção 5.3.

15. Seja $\mathcal{B} = \{u_1, \ldots, u_n\}$ uma base ortonormal de $(V, \langle\ ,\ \rangle)$. Dada um operador linear $T : V \to V$, mostre que a matriz de T com respeito à base $\mathcal{B}$, $A = (a_{ij})_{n\times n}$, satisfaz:

$$a_{ij} = \langle Tu_i, u_j \rangle.$$

16. Seja $W \subset (V, \langle\ ,\ \rangle)$ um subespaço não trivial e $v \in V - W$. Mostre que:

$$\|v - P_W(v)\| < \|v - w\| \ \forall w \in W - P_W(v),$$

isto é, $P_W(v)$ é o ponto de W que está mais próximo de v.

17. Seja Π o plano de $\mathbb{R}^3$ de equação $x - 2y + z = 0$. Considere o exercício anterior e determine o ponto de Π que está mais próximo de $v = (1, 2, -1)$.

Seção 5.4

18. Seja $\{v_1, \ldots, v_n\}$ uma base de um espaço vetorial $(V, \langle\ ,\ \rangle)$. Dados n escalares $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, prove que existe um único vetor $v \in V$, tal que

$$\langle v, v_i \rangle = \lambda_i \ \ \forall i = 1, \ldots, n.$$

19. Seja $T : \mathbb{R}^n \to \mathbb{R}^n$ um operador linear. Mostre que, para todo paralelepípedo n-dimensional $\mathscr{P}(v_1, \ldots, v_n)$, tem-se

$$|T(\mathscr{P}(v_1, \ldots, v_n))| = |\det T|\,|\mathscr{P}(v_1, \ldots, v_n)|.$$

20. Considere uma base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de $\mathbb{R}^n$. Denote por W_k o subespaço gerado por $v_1, \ldots v_k$, e ponha $w_{k+1} = v_{k+1} - P_{W_k}(v_k)$, $1 \le k \le n-1$. Mostre que, nessas condições, vale a igualdade:

$$|\mathscr{P}(v_1, \ldots, v_n)| = \|v_1\|\,\|w_2\| \ldots \|w_n\|.$$

Conclua, então, a validez da *desigualdade de Hadamard*:

$$|\det(v_1, \ldots, v_n)| \le \|v_1\|\,\|v_2\| \ldots \|v_n\|,$$

em que a igualdade ocorre se, e somente se, a base $\mathcal{B}$ é ortogonal.