Aulas Teóricas de Álgebra Linear

Instituto Superior Técnico - 1^o Semestre 2009/2010

MEAmbi - MEBiol

Matrizes

Definição 1. Uma **matriz** A, do tipo $m \times n$ (m por n), é uma tabela de mn números dispostos em m linhas e n colunas:

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

A **linha** i de A é:

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix},$$

para $i = 1, ..., m$. A **coluna** j de A é:

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

para $j = 1, ..., n$. Usa-se também a notação $A = (a_{ij})_{m \times n}$ na qual a_{ij} é a **entrada** (i, j) da matriz A.

Se $m = n$, diz-se que A é uma **matriz quadrada** do tipo $n \times n$ e as entradas a_{11}, a_{22}, ..., a_{nn} formam a chamada **diagonal principal** de A.

Exemplo 1. As matrizes

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & -2 & 0 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 0 & 7 \end{bmatrix} \text{ e } D = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

são dos seguintes tipos: A é 2×2, B é 2×4, C é 1×3, A é 4×1. Tem-se, por exemplo, $a_{21} = -2$, $b_{13} = 3$, $c_{12} = 0$ e $d_{41} = 1$.

Observação 1. Uma matriz (real) A do tipo $m \times n$ é uma aplicação:

$$\begin{array}{rcl} A : \{1, ..., m\} \times \{1, ..., n\} & \longrightarrow & \mathbb{R} \\ (i, j) & \longrightarrow & a_{ij} \end{array}$$

Notação 1. O conjunto de todas as matrizes reais (complexas) do tipo $m \times n$ é denotado por $\mathcal{M}_{m \times n}(\mathbb{R})$ ($\mathcal{M}_{m \times n}(\mathbb{C})$).

Definição 2. Duas matrizes são iguais se forem do mesmo tipo e se as entradas correspondentes forem iguais, isto é, $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{p\times q}$ são **iguais** se $m = p$, $n = q$ e $a_{ij} = b_{ij}$, para $i = 1, ..., m$ e $j = 1, ..., n$.

Definição 3. A **soma** de duas matrizes do mesmo tipo

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

é a matriz

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

Exemplo 2. Sejam

$$A = \begin{bmatrix} 1 & 4 & -1 \\ -3 & 2 & 6 \end{bmatrix}, B = \begin{bmatrix} 0 & -3 & 2 \\ 4 & -1 & -5 \end{bmatrix}, C = \begin{bmatrix} -1 \\ -1/2 \\ 2 \end{bmatrix} \text{ e } D = \begin{bmatrix} -2 & \sqrt{3} \end{bmatrix}.$$

Tem-se $A + B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ e não é possível somar C com D.

Definição 4. O **produto de um escalar** (número real ou complexo) α **por uma matriz** $A = (a_{ij})_{m\times n}$ é a matriz:

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

Notação 2. A matriz $(-1)A$ será denotada por $-A$.

Exemplo 3. Seja $A = \begin{bmatrix} 1 & 4 & -1 \\ -3 & 2 & 6 \end{bmatrix}$. Tem-se, por exemplo,

$$-2A = \begin{bmatrix} -2 & -8 & 2 \\ 6 & -4 & -12 \end{bmatrix}.$$

Definição 5. O **produto** AB de duas matrizes A e B só pode ser efectuado se o número de colunas da 1^a matriz, A, fôr igual ao número de linhas da 2^a matriz, B. Nesse caso, o produto AB de $A = (a_{ij})_{m\times p}$ por $B = (b_{ij})_{p\times n}$ é definido por:

$$AB = \left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)_{m\times n},$$

isto é,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \vdots & & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{p} a_{1k}b_{k1} & \cdots & \sum_{k=1}^{p} a_{1k}b_{kn} \\ \cdots & \sum_{k=1}^{p} a_{ik}b_{kj} & \cdots \\ \sum_{k=1}^{p} a_{mk}b_{k1} & \cdots & \sum_{k=1}^{p} a_{mk}b_{kn} \end{bmatrix}$$

Exemplo 4. Sejam A, B, C e D as matrizes do exemplo 2. Não é possível efectuar, por exemplo, AB. No entanto, tem-se:

$$AC = \begin{bmatrix} -5 \\ 14 \end{bmatrix} \quad \text{e} \quad CD = \begin{bmatrix} 2 & -\sqrt{3} \\ 1 & -\sqrt{3}/2 \\ -4 & 2\sqrt{3} \end{bmatrix}.$$

Observação 2. O produto de matrizes não é comutativo. Por exemplo, para

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \text{ tem-se } AB = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ e } BA = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Logo $AB \neq BA$.

Definição 6. A **transposta** de uma matriz $A = (a_{ij})_{m\times n}$ é a matriz

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

que se obtem trocando as linhas com as colunas de A.

Exemplo 5. Sejam A e C as matrizes do exemplo 2. Tem-se

$$A^T = \begin{bmatrix} 1 & -3 \\ 4 & 2 \\ -1 & 6 \end{bmatrix} \quad \text{e} \quad C^T = \begin{bmatrix} -1 & -\frac{1}{2} & 2 \end{bmatrix}.$$

Teorema 1. Sejam A, B, C e D matrizes de tipos apropriados, α e β escalares. São válidas as seguintes propriedades para as operações matriciais.

(a) (Comutatividade da soma) $A + B = B + A$.

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

(c) (Elemento neutro da soma) Existe uma única matriz $\mathbf{0}$ do tipo $m\times n$ tal que $A+\mathbf{0} = A$, para toda a matriz A do tipo $m \times n$. À matriz $\mathbf{0}$, cujas entradas são todas iguais a zero, chama-se **matriz nula**.

(d) (Simétrico) Para cada matriz A existe uma única matriz B tal que $A + B = \mathbf{0}$. Esta matriz B denota-se por $-A$.

(e) (Associatividade do produto por escalares) $\alpha(\beta A) = (\alpha\beta) A$.

(f) (Distributividade) $(\alpha + \beta) A = \alpha A + \beta A$.

(g) (Distributividade) $\alpha(A + B) = \alpha A + \alpha B$.

(h) (Associatividade do produto de matrizes) $A(BC) = (AB) C$.

(i) (Distributividade) $A(B + C) = AB + AC \quad \text{e} \quad (B + C) D = BD + CD$.

(j) $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

(k) $\left(A^T\right)^T = A$.

(l) $(A+B)^T = A^T + B^T$.

(m) $(\alpha A)^T = \alpha A^T$.

(n) $(AB)^T = B^T A^T$.

(o) $(A_1 A_2 ... A_n)^T = A_n^T ... A_2^T A_1^T$, com A_1, A_2, ..., A_n matrizes de tipos apropriados.

(p) À matriz, do tipo $n \times n$, cujas entradas da diagonal principal sejam iguais a 1 e as restantes sejam iguais a 0:

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

chama-se **matriz identidade** (de ordem n) e é tal que

$$AI = A \quad \text{e} \quad IB = B,$$

para todas as matrizes $A = (a_{ij})_{m \times n}$ e $B = (b_{ij})_{n \times m}$.

Definição 7. **(i)** A **diferença** entre duas matrizes A e B do mesmo tipo é definida por

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

ou seja, é a soma de A com o simétrico de B.

(ii) Sejam A uma matriz do tipo $n \times n$ e $p \in \mathbb{N}$. A **potência** p de A é definida por

$$A^p = \underbrace{A...A}_{p \text{ vezes}} \ \text{ e } \ \text{para } p = 0 \ \text{ define-se } \ A^0 = I.$$

(iii) À matriz do tipo $n \times n$

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

cujas entradas fora da diagonal principal são nulas, chama-se **matriz diagonal**.

Observação 3. Tem-se: $1A = A$, $0A = \mathbf{0}$, $A + A = 2A$, $\underbrace{A + \ldots + A}_{n \text{ vezes}} = nA$.

Definição 8. Seja $A = (a_{ij})_{n \times n}$ uma matriz do tipo $n \times n$. Diz-se que A é **simétrica** se $A = A^T$, isto é, se $a_{ij} = a_{ji}$, para $i, j = 1, ..., n$. Diz-se que A é **anti-simétrica** se $A = -A^T$, isto é, se $a_{ij} = -a_{ji}$, para $i, j = 1, ..., n$.

Sistemas de Equações Lineares

Definição 9. Uma **equação linear** com n incógnitas $x_1, x_2, ..., x_n$ é uma equação da forma

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b,$$

em que $a_1, a_2, ..., a_n$ e b são constantes (reais ou complexas).

Definição 10. Um **sistema de** m **equações lineares** com n incógnitas é um conjunto de equações da forma

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

em que a_{ij} e b_k são constantes (reais ou complexas), para $i, k = 1, ..., m$ e $j = 1, ..., n$.

Observação 4. Usando o produto de matrizes definido na secção anterior, o sistema linear acima pode ser escrito como uma equação matricial

$$AX = B,$$

em que

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

A matriz A é a **matriz dos coeficientes do sistema**, X é a matriz coluna das incógnitas e B é a matriz coluna dos termos independentes. Uma solução do sistema linear $(*)$ é uma matriz

$$S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

tal que as equações do sistema são satisfeitas quando substituímos

$$x_1 = s_1, x_2 = s_2, ..., x_n = s_n.$$

Ao conjunto de todas as soluções do sistema chama-se **conjunto solução** ou **solução geral** do sistema.

Exemplo 6. O sistema linear de duas equações e duas incógnitas $\left\{\begin{array}{l} x + 2y = 1 \\ 2x + y = 0 \end{array}\right.$ pode ser escrito do seguinte modo:

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

A solução (geral) do sistema acima é $x = -1/3$ e $y = 2/3$ (verifique!), isto é, $X = \begin{bmatrix} -1/3 \\ 2/3 \end{bmatrix}$.

Observação 5. De modo a facilitar a resolução de um sistema linear, este pode ser sempre substituído por outro que tenha o mesmo conjunto solução. Esse outro é obtido depois de aplicar sucessivamente operações sobre as equações do sistema inicial que não alterem a solução do mesmo. As operações são:

- Trocar a posição de duas equações do sistema;
- Multiplicar uma equação por um escalar diferente de zero;
- Somar a uma equação um múltiplo escalar de outra equação.

Estas são as chamadas operações elementares. Quando aplicamos operações elementares às equações de um sistema linear, só os coeficientes e os termos independentes do sistema são alterados. Assim, podemos aplicar as operações à matriz

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

à qual se dá o nome de **matriz aumentada do sistema**.

Definição 11. As **operações elementares** que podem ser aplicadas às linhas de uma matriz são as seguintes:

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

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

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

Teorema 2. Se dois sistemas lineares $AX = B$ e $CX = D$ são tais que a matriz aumentada $[C \mid D]$ é obtida de $[A \mid B]$ através de uma operação elementar, então os dois sistemas têm o mesmo conjunto solução, isto é, são **equivalentes**.

Definição 12. Uma **matriz** $A = (a_{ij})_{m\times n}$ diz-se **em escada de linhas** se:

(i) Todas as linhas nulas (formadas inteiramente por zeros) estão por baixo das linhas não nulas;

(ii) Por baixo (e na mesma coluna) do primeiro elemento não nulo de cada linha e por baixo dos elementos nulos anteriores da mesma linha, todas as entradas são nulas. Esse primeiro elemento não nulo de cada linha tem o nome de **pivot**.

Definição 13. O método de resolver sistemas lineares que consiste em aplicar operações elementares às linhas da matriz aumentada do respectivo sistema de modo a que essa matriz fique em escada de linhas, chama-se **método de eliminação de Gauss**.

Exemplo 7. O sistema linear de variáveis reais x, y e z

$$\begin{cases} x+z=3 \\ \\ x+2y+2z=6 \\ \\ 3y+3z=6 \end{cases} \qquad \text{é equivalente a} \quad \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 0 & 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}.$$

Consideremos então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 1 & 2 & 2 & 6 \\ 0 & 3 & 3 & 6 \end{array}\right] \underset{-L_1+L_2\to L_2}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 2 & 1 & 3 \\ 0 & 3 & 3 & 6 \end{array}\right] \underset{-\frac{3}{2}L_2+L_3\to L_3}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & \frac{3}{2} & \frac{3}{2} \end{array}\right].$$

Logo,

$$\begin{cases} x+z=3 \\ \\ 2y+z=3 \\ \\ \frac{3}{2}z=\frac{3}{2} \end{cases} \Leftrightarrow \begin{cases} x=2 \\ \\ y=1 \\ \\ z=1. \end{cases}$$

Neste exemplo o sistema tem **solução única** e diz-se **possível e determinado**.

Exemplo 8. O sistema linear de variáveis reais x, y, z e w

$$\begin{cases} 3z-9w=6 \\ \\ 5x+15y-10z+40w=-45 \\ \\ x+3y-z+5w=-7 \end{cases} \quad \text{é equivalente a} \quad \begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 6 \\ -45 \\ -7 \end{bmatrix}.$$

Consideremos então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] \underset{\substack{L_1\leftrightarrow L_3 \\ \frac{1}{5}L_2\to L_2}}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{-L_1+L_2\to L_2}{\longrightarrow}$$

$$\underset{-L_1+L_2\to L_2}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{3L_2+L_3\to L_3}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Logo,

$$\begin{cases} x+3y-z+5w=-7 \\ \\ -z+3w=-2 \end{cases} \Leftrightarrow \begin{cases} x=-3y-2w-5 \\ \\ z=3w+2. \end{cases}$$

As incógnitas y e w são livres e as incógnitas x e z são não livres. A solução geral do sistema é:

$$X = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -3s-2t-5 \\ s \\ 3t+2 \\ t \end{bmatrix},$$

para quaisquer $s, t \in \mathbb{R}$, isto é, o conjunto solução é dado por:

$$S = \{(-3s - 2t - 5, s, 3t + 2, t) : s, t \in \mathbb{R}\}.$$

Neste exemplo o sistema tem **infinitas soluções** e diz-se **possível e indeterminado**.

Exemplo 9. Seja $a \in \mathbb{R}$. O sistema linear de variáveis reais x, y e z

$$\begin{cases} x + 2y + z = 3 \\ x + y - z = 2 \\ x + y + (a^2 - 5)\, z = a \end{cases} \qquad \text{é equivalente a} \quad \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & a^2 - 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ a \end{bmatrix}.$$

Consideremos então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 1 & 1 & -1 & 2 \\ 1 & 1 & a^2 - 5 & a \end{array}\right] \underset{\substack{-L_1 + L_2 \to L_2 \\ -L_1 + L_3 \to L_3}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -1 & -2 & -1 \\ 0 & -1 & a^2 - 6 & a - 3 \end{array}\right] \underset{-L_2 + L_3 \to L_3}{\longrightarrow}.$$

$$\underset{-L_2 + L_3 \to L_3}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -1 & -2 & -1 \\ 0 & 0 & a^2 - 4 & a - 2 \end{array}\right].$$

Se $a = 2$, então o sistema é possível e indeterminado:

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

a incógnita z é livre, as incógnitas x e y são não livres e a solução geral do sistema é

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3t + 1 \\ -2t + 1 \\ t \end{bmatrix},$$

para qualquer $t \in \mathbb{R}$, isto é, o conjunto solução é dado por:

$$S = \{(3t + 1, -2t + 1, t) : t \in \mathbb{R}\}.$$

Assim, se $a = 2$, o sistema tem **infinitas soluções** e diz-se **possível e indeterminado**.

Se $a = -2$, o sistema **não tem solução** e diz-se **impossível**.

Se $a \neq -2$ e $a \neq 2$, o sistema tem a **solução única**:

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (a + 5)/(a + 2) \\ a/(a + 2) \\ 1/(a + 2) \end{bmatrix}$$

e diz-se **possível e determinado**.

Definição 14. Seja A uma matriz em escada de linhas. Ao n^o de colunas de A que não contêm pivots chama-se **nulidade** de A e escreve-se $\operatorname{nul} A$. Ao n^o de pivots de A, isto é, ao n^o de linhas não nulas de A, dá-se o nome de **característica** de A e escreve-se $\operatorname{car} A$. Se A fôr a matriz em escada de linhas obtida de C através de operações elementares então diz-se que a **característica** de C é $\operatorname{car} A$, tendo-se $\operatorname{car} C = \operatorname{car} A$.

Exemplo 10. As seguintes matrizes estão em escada de linhas:

$$A_1 = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & -5 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & -1 & 1/2 & 0 & 0 \\ 0 & 0 & -3 & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Pivot de A_1: 4. Pivots de A_2: $1, -5$. Pivots de A_3: $2, -3, -5$. Tem-se: $\operatorname{car} A_1 = 1$, $\operatorname{car} A_2 = 2$ e $\operatorname{car} A_3 = 3$. Além disso: $\operatorname{nul} A_1 = 1$, $\operatorname{nul} A_2 = 2$ e $\operatorname{nul} A_3 = 2$.

Observação 6. Seja $[A \mid B]$ a matriz aumentada associada a um sistema linear com n incógnitas.

(i) Se $\operatorname{car} A = \operatorname{car} [A \mid B] = n$ então o sistema é **possível e determinado** (tem uma única solução).

(ii) Se $\operatorname{car} A = \operatorname{car} [A \mid B] < n$ então o sistema é **possível e indeterminado** (tem um n^o infinito de soluções).

(iii) Se $\operatorname{car} A < \operatorname{car} [A \mid B]$ então o sistema é **impossível** (não tem solução).

(iv) As **incógnitas livres** (podem tomar valores arbitrários) do sistema são aquelas que correspondem às colunas, que não contenham pivots, da matriz em escada de linhas obtida de A através de operações elementares.

(v) As **incógnitas não livres** do sistema são aquelas que correspondem às colunas, que contenham pivots, da matriz em escada de linhas obtida de A através de operações elementares.

(vi) $\operatorname{car} A =$ n^o de linhas não nulas da matriz em escada de linhas obtida de A $=$ n^o de pivots $=$ n^o de incógnitas não livres.

$\operatorname{nul} A =$ n^o de incógnitas livres.

(vii) Seja A uma matriz do tipo $m \times n$. Então:

$$0 \leq \operatorname{car} A \leq \min \{m, n\} \qquad \text{e} \qquad \operatorname{car} A + \text{nulA} = n.$$

Teorema 3. Sejam A uma matriz do tipo $m \times n$ e B uma matriz do tipo $m \times 1$. Se o sistema linear $AX = B$ tem duas soluções distintas X_0 e X_1 $(X_0 \neq X_1)$, então terá infinitas soluções.

Dem. Basta verificar que $X_\lambda = (1 - \lambda) X_0 + \lambda X_1$ é solução do sistema $AX = B$, para qualquer $\lambda \in \mathbb{R}$.

Definição 15. Um sistema linear da forma $\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0 \end{cases}$ tem o nome de **sistema linear homogéneo**. Este sistema poder ser escrito na forma $AX = \mathbf{0}$.

Observação 7. **(i)** Todo o sistema linear homogéneo $AX = \mathbf{0}$ admite pelo menos a **solução trivial**:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Assim, todo o sistema linear homogéneo tem solução. Além disso, ou tem apenas a solução trivial ou tem infinitas soluções.

(ii) Como iremos ver num próximo capítulo, à solução geral do sistema linear homogéneo $AX = \mathbf{0}$ dá-se o nome de **núcleo** de A e escreve-se $\mathcal{N}(A)$.

Teorema 4. Se $A = (a_{ij})_{m\times n}$ é tal que $m < n$, então o sistema linear homogéneo $AX = \mathbf{0}$ tem infinitas soluções.

Dem. Como o sistema tem menos equações do que incógnitas ($m < n$), o n^o de linhas não nulas r da matriz em escada de linhas obtida da matriz aumentada do sistema também é tal que $r < n$. Assim, há r pivots e $n-r$ incógnitas livres as quais podem assumir qualquer valor. Logo, o sistema linear homogéneo $AX = \mathbf{0}$ tem infinitas soluções.

Teorema 5. Sejam $A = (a_{ij})_{m\times n}$ e α, β escalares.

(i) Se Y e W são soluções do sistema $AX = \mathbf{0}$, então $Y + W$ também o é.

(ii) Se Y é solução do sistema $AX = \mathbf{0}$, então αY também o é.

(iii) Se Y e W são soluções do sistema $AX = \mathbf{0}$, então $\alpha Y + \beta W$ também o é.

(iv) Sejam Y e W soluções do sistema $AX = B$. Se $\alpha Y + \beta W$ (para quaisquer escalares α, β) também é solução de $AX = B$, então $B = \mathbf{0}$. (Sugestão: basta fazer $\alpha = \beta = 0$.)

Teorema 6. Seja A uma matriz do tipo $m \times n$ e $B \neq \mathbf{0}$ uma matriz do tipo $m \times 1$. Qualquer solução X do sistema $AX = B$ escreve-se na forma $X = X_0 + Y$ onde X_0 é uma solução particular do sistema $AX = B$ e Y é uma solução do sistema homogéneo $AX = \mathbf{0}$. Assim:

$$\left\{ \begin{array}{c} \text{solução geral de} \\ AX = B \end{array} \right\} = \left\{ \begin{array}{c} \text{solução particular de} \\ AX = B \end{array} \right\} + \left\{ \begin{array}{c} \text{solução geral de} \\ AX = \mathbf{0} \end{array} \right\}.$$

Dem. Sendo X_0 uma solução particular do sistema $AX = B$, basta escrever

$$X = X_0 + (X - X_0)$$

e mostrar que $X - X_0$ é solução do sistema homogéneo $AX = \mathbf{0}$.

Matrizes Elementares e Matriz Inversa

Definição 16. Uma **matriz elementar** do tipo $n \times n$ é uma matriz obtida da matriz identidade I (do tipo $n \times n$) através de uma única operação elementar.

(i) A matriz P_{ij}, chamada **matriz de permutação**, é a matriz elementar obtida por troca da linha i com a linha j da matriz I. Tem-se:

$$P_{ij} = \begin{bmatrix} 1 & 0 & \cdots & & & & & & & \cdots & 0 \\ 0 & \ddots & \ddots & & & & & & & & \vdots \\ \vdots & \ddots & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & \ddots & \vdots \\ \vdots & & & & & & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & & & & & \cdots & 0 & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{array} .$$

(ii) A matriz $E_i(\alpha)$ é a matriz elementar obtida da matriz I através do produto do escalar $\alpha \neq 0$ pela linha i da matriz I. Tem-se:

$$E_i(\alpha) = \begin{bmatrix} 1 & 0 & \cdots & & & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & \cdots & 0 & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{array} .$$

(iii) A matriz $E_{ij}(\alpha)$ é a matriz elementar obtida da matriz I por soma da linha j com um múltiplo escalar α da linha i. Por exemplo para $i < j$ tem-se:

$$E_{ij}(\alpha) = \begin{bmatrix} 1 & 0 & \cdots & & & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & & & & \\ & & & \ddots & & & \\ & & \alpha & & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & \cdots & 0 & 1 \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{array} .$$

Observação 8. As matrizes elementares $E_{ij}(\alpha)$, com $i < j$, são **matrizes triangulares inferiores**, pois todas as entradas por cima das respectivas diagonais principais são nulas.

Exemplo 11. Sejam α, β escalares com $\alpha \neq 0$. As matrizes elementares do tipo 2×2 são:

$$P_{12} = P_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_1(\alpha) = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}, \quad E_2(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix},$$

$$E_{12}(\beta) = \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} \qquad \text{e} \qquad E_{21}(\beta) = \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}.$$

Teorema 7. Sejam E uma matriz elementar do tipo $m \times m$ e A uma matriz qualquer do tipo $m \times n$. Então, EA é a matriz obtida de A através da mesma operação elementar que originou E. Isto é, aplicar uma operação elementar a uma matriz corresponde a multiplicar essa matriz à esquerda por uma matriz elementar.

Exemplo 12. Consideremos a matriz aumentada do exemplo 8:

$$\left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right].$$

A operação elementar:

$$\left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right].$$

A operação elementar:

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{\frac{1}{5}L_2 \to L_2}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right].$$

A operação elementar:

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{-L_1 + L_2 \to L_2}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right].$$

Finalmente, a operação elementar:

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{3L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Tem-se então:

$$E_{23}(3)\, E_{12}(-1)\, E_2\left(\frac{1}{5}\right) P_{13} \left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Definição 17. Uma matriz A (do tipo $n \times n$) diz-se invertível se existir uma matriz B (do tipo $n \times n$) tal que

$$AB = BA = I.$$

À matriz B chama-se **matriz inversa** de A e denota-se por A^{-1}.

Observação 9. Obviamente que resulta da definição de matriz inversa o seguinte facto: sendo A^{-1} a matriz inversa de A, então A^{-1} é invertível e a sua inversa é a própria matriz A, isto é, $(A^{-1})^{-1} = A$.

Exemplo 13. As matrizes $A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} -1/2 & 1/6 \\ 0 & 1/3 \end{bmatrix}$ são a inversa uma da outra.

Teorema 8. A inversa de uma matriz é única.

Dem. Sejam B e C as inversas de A. Então, $B = BI = B(AC) = (BA)C = IC = C$.

Teorema 9. **(i)** Se $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ são duas matrizes invertíveis, então AB é invertível e

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

(ii) Seja $m \in \mathbb{N}$. Se $A = (a_{ij})_{n\times n}$ é invertível, então A^m é invertível e $(A^m)^{-1} = (A^{-1})^m$ e escreve-se $A^{-m} = (A^m)^{-1}$.

(iii) Se $A=(a_{ij})_{n\times n}$ é invertível, então A^T é invertível e $\left(A^T\right)^{-1}=\left(A^{-1}\right)^T$.

(iv) Seja $A=(a_{ij})_{n\times n}$. Se existir $l\in\mathbb{N}$ tal que $A^l=\mathbf{0}$ então A não é invertível.

Definição 18. Uma matriz $A=(a_{ij})_{n\times n}$ diz-se **não singular** se após o método de eliminação de Gauss esta fôr transformada numa **matriz triangular superior** (matriz cujas entradas por baixo da diagonal principal são todas nulas) cujas entradas da diagonal principal sejam todas não nulas. Uma matriz $A=(a_{ij})_{n\times n}$ diz-se **singular** se após o método de eliminação de Gauss existir (pelo menos) uma linha nula na matriz obtida de A.

Teorema 10. Seja $A=(a_{ij})_{n\times n}$. A é invertível $\Leftrightarrow$ A é não singular $\Leftrightarrow$ $\operatorname{car} A=n$.

Teorema 11. Toda a matriz elementar é invertível e a respectiva inversa é também uma matriz elementar. Tem-se:

(i) $(P_{ij})^{-1}=P_{ij}$.

(ii) $(E_i(\alpha))^{-1}=E_i(1/\alpha)$, para $\alpha\neq 0$.

(iii) $(E_{ij}(\alpha))^{-1}=E_{ij}(-\alpha)$.

Observação 10. Uma matriz A é invertível se e só se fôr igual ao produto de matrizes elementares.

Teorema 12. Seja A uma matriz do tipo $n\times n$.

(i) O sistema associado a $AX=B$ tem solução única se e só se A fôr invertível. Neste caso a solução é $X=A^{-1}B$.

(ii) O sistema homogéneo $AX=\mathbf{0}$ tem solução não trivial se e só se A fôr singular (não invertível).

Teorema 13. Sejam A e B duas matrizes do tipo $n\times n$. Se AB é invertível, então A e B são invertíveis.

Dem. Considere o sistema $(AB)\,X=\mathbf{0}$. Se B não fosse invertível, então pelo teorema 12 existiria $X\neq\mathbf{0}$ tal que $BX=\mathbf{0}$. Logo, $X\neq\mathbf{0}$ seria solução não trivial de $ABX=\mathbf{0}$, o que contraria o teorema 12 uma vez que por hipótese AB é invertível. Assim, B é invertível. Finalmente, A é invertível por ser o produto de duas matrizes invertíveis: $A=(AB)\,B^{-1}$.

Observação 11. (**Como inverter matrizes do tipo** $n\times n$). Seja A uma matriz do tipo $n\times n$ e consideremos a equação $AX=B$. **Se** A **fôr invertível** temos

$$AX=B\Leftrightarrow X=A^{-1}B,\qquad\text{isto é,}\qquad AX=IB\Leftrightarrow IX=A^{-1}B.$$

Assim, para determinar a inversa de A, iremos transformar a matriz aumentada $[A\mid I]$ na matriz $[I\mid A^{-1}]$, por meio de operações elementares aplicadas às linhas de $[A\mid I]$. Este método tem o nome de **método de eliminação de Gauss-Jordan** e consistirá na continuação do método de eliminação de Gauss agora aplicado a [matriz triangular superior $\mid *$], efectuando-se as eliminações de baixo para cima de modo a obter-se $[I\mid A^{-1}]$.

Exemplo 14. **(i)** Seja $A = \begin{bmatrix} 0 & -1 & 1 \\ -1 & \frac{5}{4} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & 0 \end{bmatrix}$. Tem-se $[A \mid I] = \left[\begin{array}{ccc|ccc} 0 & -1 & 1 & 1 & 0 & 0 \\ -1 & \frac{5}{4} & -\frac{1}{2} & 0 & 1 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 & 0 & 1 \end{array}\right] \underset{L_1 \leftrightarrow L_2}{\longrightarrow}$

$$\underset{L_1 \leftrightarrow L_2}{\longrightarrow} \left[\begin{array}{ccc|ccc} -1 & \frac{5}{4} & -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 & 0 & 1 \end{array}\right] \underset{L_1+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{ccc|ccc} -1 & \frac{5}{4} & -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \\ 0 & \frac{3}{4} & -\frac{1}{2} & 0 & 1 & 1 \end{array}\right] \underset{\frac{3}{4}L_2+L_3 \to L_3}{\longrightarrow}$$

$$\underset{\frac{3}{4}L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{ccc|ccc} -1 & \frac{5}{4} & -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} & 1 & 1 \end{array}\right] \underset{\substack{2L_3+L_1 \to L_1 \\ -4L_3+L_2 \to L_2}}{\longrightarrow} \left[\begin{array}{ccc|ccc} -1 & \frac{5}{4} & 0 & \frac{3}{2} & 3 & 2 \\ 0 & -1 & 0 & -2 & -4 & -4 \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} & 1 & 1 \end{array}\right] \underset{\frac{5}{4}L_2+L_1 \to L_1}{\longrightarrow}$$

$$\underset{\frac{5}{4}L_2+L_1 \to L_1}{\longrightarrow} \left[\begin{array}{ccc|ccc} -1 & 0 & 0 & -1 & -2 & -3 \\ 0 & -1 & 0 & -2 & -4 & -4 \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} & 1 & 1 \end{array}\right] \underset{\substack{-L_2 \to L_2 \\ -L_1 \to L_1 \\ 4L_3 \to L_3}}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & 3 \\ 0 & 1 & 0 & 2 & 4 & 4 \\ 0 & 0 & 1 & 3 & 4 & 4 \end{array}\right].$$

Logo, $\begin{bmatrix} 0 & -1 & 1 \\ -1 & \frac{5}{4} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 4 & 4 \end{bmatrix}$.

Verifique(!) que: $\begin{bmatrix} 0 & -1 & 1 \\ -1 & \frac{5}{4} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 4 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(ii) Seja $A = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$. Tem-se

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 9 & 8 & 7 & 1 & 0 & 0 \\ 6 & 5 & 4 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right] \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \left[\begin{array}{ccc|ccc} 3 & 2 & 1 & 0 & 0 & 1 \\ 6 & 5 & 4 & 0 & 1 & 0 \\ 9 & 8 & 7 & 1 & 0 & 0 \end{array}\right] \underset{\substack{-2L_1+L_2 \to L_2 \\ -3L_1+L_3 \to L_3}}{\longrightarrow}$$

$$\underset{\substack{-2L_1+L_2 \to L_2 \\ -3L_1+L_3 \to L_3}}{\longrightarrow} \left[\begin{array}{ccc|ccc} 3 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 2 & 4 & 1 & 0 & -3 \end{array}\right] \underset{-2L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{ccc|ccc} 3 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right]$$

Logo, A é singular e como tal não é invertível.

(iii) Sejam $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} -4 & 0 \\ 0 & 8 \end{bmatrix}$, $C = \begin{bmatrix} 0 & \frac{1}{8} \\ -\frac{1}{4} & 0 \end{bmatrix}$. Determine-se X tal que $A\left(2X^T\right)^{-1} B^{-1} = C$.

Tem-se

$$A\left(2X^T\right)^{-1} B^{-1} = C \Leftrightarrow \left(2X^T\right)^{-1} = A^{-1}CB \Leftrightarrow 2X^T = \left(A^{-1}CB\right)^{-1} \Leftrightarrow$$

$$\Leftrightarrow X^T = \frac{1}{2}B^{-1}C^{-1}A \Leftrightarrow X = \frac{1}{2}A^T\left(C^T\right)^{-1}\left(B^T\right)^{-1} \Leftrightarrow$$

$$\Leftrightarrow X = \frac{1}{2}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 0 & -\frac{1}{4} \\ \frac{1}{8} & 0 \end{bmatrix}^{-1}\begin{bmatrix} -4 & 0 \\ 0 & 8 \end{bmatrix}^{-1} \Leftrightarrow X = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ 2 & 1 \end{bmatrix}.$$

Determinante

Definição 19. Dados os números naturais $1, 2, ..., n$ chama-se **permutação** desses n números a qualquer lista em que os mesmos sejam apresentados por ordem arbitrária.

Definição 20. Seja $(i_1 i_2 ... i_n)$ uma permutação dos números naturais $1, 2, ..., n$. Diz-se que um par $(i_j i_k)$ é uma **inversão** quando $(j-k)(i_j - i_k) < 0$ (isto é, quando i_j e i_k aparecerem na permutação por ordem decrescente).

Definição 21. Uma permutação $(i_1 i_2 ... i_n)$ diz-se **par** (**ímpar**) quando o n^o máximo de inversões incluídas fôr par (ímpar).

Exemplo 15. A permutação (21453) é ímpar pois contem as inversões $(21), (43)$ e (53).

Definição 22. Seja A uma matriz do tipo $n \times n$. Chama-se **determinante** de A, e escreve-se $|A|$ ou $\det A$, o número que se obtem do seguinte modo:

(i) Formam-se todos os produtos possíveis de n factores em que intervenha um elemento de cada linha e, simultaneamente, um elemento de cada coluna de A.

(ii) Afecta-se cada produto do sinal $+$ ou do sinal $-$ conforme as permutações (dos números naturais $1, 2, ..., n$) que figuram nos índices de linha e de coluna tenham a mesma paridade ou não.

(iii) Somam-se as parcelas obtidas.

Em resumo: fixando, por exemplo, a permutação $(i_1 i_2 ... i_n)$ de $1, 2, ..., n$

$$|A| = \sum_{\substack{(j_1 j_2 ... j_n) \\ \text{permutação de } 1,2,...,n}} (-1)^\sigma a_{i_1 j_1} a_{i_2 j_2} ... a_{i_n j_n},$$

em que

$$\sigma = \begin{cases} 0 & \text{se } (i_1 i_2 ... i_n) \text{ e } (j_1 j_2 ... j_n) \text{ têm a mesma paridade} \\ 1 & \text{se } (i_1 i_2 ... i_n) \text{ e } (j_1 j_2 ... j_n) \text{ têm paridade diferente.} \end{cases}$$

Observação 12. Podemos ainda escrever de modo equivalente:

(i)

$$|A| = \sum_{\substack{(j_1 j_2 ... j_n) \\ \text{permutação de } 1,2,...,n}} (-1)^\sigma a_{1 j_1} a_{2 j_2} ... a_{n j_n},$$

em que

$$\sigma = \begin{cases} 0 & \text{se } (j_1 j_2 ... j_n) \text{ é par} \\ 1 & \text{se } (j_1 j_2 ... j_n) \text{ é ímpar.} \end{cases}$$

(ii)

$$|A| = \sum_{\substack{(i_1 i_2 ... i_n) \\ \text{permutação de } 1,2,...,n}} (-1)^\sigma a_{i_1 1} a_{i_2 2} ... a_{i_n n},$$

em que

$$\sigma = \begin{cases} 0 & \text{se } (i_1 i_2 ... i_n) \text{ é par} \\ \\ 1 & \text{se } (i_1 i_2 ... i_n) \text{ é ímpar.} \end{cases}$$

Teorema 14. (i) Sendo $A = (a_{ij})_{n \times n}$, chama-se **traço** de A a: $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$. Se A é do tipo 2×2, então:

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \underset{\substack{\text{Só no caso} \\ 2\times 2}}{=} \frac{1}{2}\left((\operatorname{tr} A)^2 - \operatorname{tr}\left(A^2\right)\right).$$

(ii) Seja A uma matriz do tipo 3×3. Então

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

Observação 13. Se A é uma matriz do tipo $n \times n$ então $|A|$ tem $n!$ parcelas.

Exemplo 16. (i) $\begin{vmatrix} 1 & -1 \\ 2 & -2 \end{vmatrix} = 1(-2) - (-1)2 = 0.$

(ii) $\begin{vmatrix} 1 & 2 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & -3 \end{vmatrix} = 1(-1)(-3) + 3 + 8 - 1(-1)2 - 6(-3) - 2 = 32.$

Definição 23. Seja $A = (a_{ij})$ uma matriz do tipo $n \times n$, com $n > 1$. Seja A_{ij} a matriz do tipo $(n-1) \times (n-1)$ que se obtem de A suprimindo a linha i e a coluna j de A. Chama-se a A_{ij} o **menor-ij** da matriz A.

Teorema 15. (Fórmula de Laplace.) Seja A uma matriz do tipo $n \times n$, com $n > 1$. Tem-se

$$\det A = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij},$$

com $i \in \{1, \ldots, n\}$ fixo.

Observação 14. Seja A uma matriz do tipo $n \times n$, com $n > 1$. Tem-se

$$\det A = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij},$$

com $j \in \{1, \ldots, n\}$ fixo.

Exemplo 17.

$$\begin{vmatrix} 1 & 0 & -2 & 3 \\ 2 & 1 & -1 & 4 \\ 0 & -1 & 0 & -2 \\ 1 & 0 & -2 & -3 \end{vmatrix} = (-1)(-1)^{3+2} \begin{vmatrix} 1 & -2 & 3 \\ 2 & -1 & 4 \\ 1 & -2 & -3 \end{vmatrix} + (-2)(-1)^{3+4} \begin{vmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \\ 1 & 0 & -2 \end{vmatrix} =$$

$$= (-1)(-3) + (-2)4 + 2(-2)3 - (-1)3 - (-2)2(-3) - 4(-2) + 2\left[(-2) - (-2)\right] = -18$$

Teorema 16. Sejam A e B matrizes do tipo $n \times n$. Seja λ um escalar.

(i) Se A fôr uma matriz diagonal, triangular superior ou triangular inferior então o determinante de A é igual ao produto dos elementos da diagonal principal de A.

(ii) Se A tiver uma linha (ou coluna) nula então $\det A = 0$.

(iii) $\det\left(A^T\right) = \det A$.

(iv) Se B fôr obtida de A trocando duas linhas (ou colunas) de A então $\det B = -\det A$.

(v) Se B fôr obtida de A multiplicando uma linha (ou coluna) de A por um escalar λ então $\det B = \lambda \det A$.

(vi) Se B fôr obtida de A somando a uma linha (ou coluna) de A um múltiplo escalar λ de uma outra linha (ou coluna) de A então $\det B = \det A$.

(vii) Se duas linhas (ou colunas) de A forem iguais então $\det A = 0$.

(viii) $\det A \neq 0$ se e só se A é invertível.

(ix) $\det(\lambda A) = \lambda^n \det A$.

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

(xi) Seja $l \in \mathbb{N}$. Sejam $A_1, A_2, \ldots, A_l$ matrizes do tipo $n \times n$. Tem-se: $\det(A_1 A_2 \ldots A_l) = \det A_1 \det A_2 \ldots \det A_l$.

(xii) Se A fôr invertível $\det(A^{-1}) = \dfrac{1}{\det A}$.

(xiii) $\det(AB) = 0 \Rightarrow \det A = 0 \;\text{ ou }\; \det B = 0$.

(xiv) $\det(AB) = \det(BA)$.

Definição 24. Seja $A = (a_{ij})$ uma matriz do tipo $n \times n$, com $n > 1$. Seja $a'_{ij} = (-1)^{i+j} \det A_{ij}$ onde A_{ij} é o menor-ij da matriz A. Chama-se a a'_{ij} o **cofactor-**ij da matriz A e à matriz $\operatorname{cof} A = (a'_{ij})$ do tipo $n \times n$, com $n > 1$, a matriz dos cofactores de A.

Teorema 17. Para qualquer matriz A do tipo $n \times n$, com $n > 1$, tem-se

$$A\left(\operatorname{cof} A\right)^T = \left(\operatorname{cof} A\right)^T A = (\det A)\, I.$$

Se $\det A \neq 0$ então A é invertível e

$$A^{-1} = \frac{1}{\det A} (\operatorname{cof} A)^T = \left(\underbrace{\frac{1}{\det A} (-1)^{j+i} \det A_{ji}}_{\text{entrada } (i,j) \text{ de } A^{-1}} \right)_{n \times n}.$$

Exemplo 18. **(i)** Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tal que $\det A \neq 0$. Então A é invertível e

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

(Veja por exemplo o ex^o 4 (i),(ii),(iv),(viii) da ficha 2.) Note que $ad - bc = \det A$.

(ii) Podemos usar o teorema 17 para calcular não só a inversa de uma matriz (não singular) mas também entradas concretas dessa inversa. Seja

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

A entrada $(2, 3)$ da matriz A^{-1} é dada por

$$(A^{-1})_{23} = \frac{1}{\det A} \left((\operatorname{cof} A)^T \right)_{23} = \frac{1}{\det A} \left((-1)^{3+2} \det A_{32} \right) = \frac{1}{-3} \left(-\det \left(\begin{bmatrix} 1 & 0 \\ 4 & 6 \end{bmatrix} \right) \right) = 2.$$

Teorema 18. **(Regra de Cramer.)** Seja A uma matriz do tipo $n \times n$ tal que A é não singular. Então a única solução do sistema de equações lineares $AX = B$ é dada por

$$X = A^{-1}B = \frac{1}{\det A} (\operatorname{cof} A)^T B.$$

Isto é, sendo $X = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ e $B = \begin{bmatrix} b_1 & \ldots & b_n \end{bmatrix}^T$ tem-se, para $j = 1, \ldots, n$,

$$x_j = \frac{1}{\det A} \sum_{i=1}^{n} a'_{ji} b_i = \frac{\det \left((B_j)^T \right)}{\det A} = \frac{\det B_j}{\det A},$$

onde B_j é a matriz obtida de A substituindo a coluna j de A pela matriz coluna B dos termos independentes.

Exemplo 19. O sistema de equações lineares $\begin{cases} 2x + y = 8 \\ -x + 2y + 4z = 7 \\ -x + z = 1 \end{cases}$ pode ser resolvido usando a regra de Cramer:

$$x = \frac{\begin{vmatrix} 8 & 1 & 0 \\ 7 & 2 & 4 \\ 1 & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 0 \\ -1 & 2 & 4 \\ -1 & 0 & 1 \end{vmatrix}} = 13, \quad y = \frac{\begin{vmatrix} 2 & 8 & 0 \\ -1 & 7 & 4 \\ -1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 0 \\ -1 & 2 & 4 \\ -1 & 0 & 1 \end{vmatrix}} = -18 \quad \text{e} \quad z = \frac{\begin{vmatrix} 2 & 1 & 8 \\ -1 & 2 & 7 \\ -1 & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 1 & 0 \\ -1 & 2 & 4 \\ -1 & 0 & 1 \end{vmatrix}} = 14.$$

Espaços Lineares

Definição 25. Um conjunto não vazio V é um **espaço linear** (real) se existirem duas operações associadas a V, uma soma de elementos de V e um produto de escalares (números reais) por elementos de V, com as seguintes propriedades:

(a) (Fecho da soma). Para quaisquer $u, v \in V$ tem-se

$$u + v \in V.$$

(b) (Fecho do produto por escalares). Para quaisquer $\alpha \in \mathbb{R}$ e $u \in V$ tem-se

$$\alpha u \in V.$$

(c) (Comutatividade da soma). Para quaisquer $u, v \in V$,

$$u + v = v + u.$$

(d) (Associatividade da soma). Para quaisquer $u, v, w \in V$,

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

(e) (Elemento neutro da soma). Existe um elemento de V designado por $\mathbf{0}$ tal que, para qualquer $u \in V$,

$$u + \mathbf{0} = u.$$

(f) (Simétrico). Para cada (qualquer) $u \in V$ existe $v \in V$ tal que

$$u + v = \mathbf{0}.$$

A v chama-se o **simétrico** de u e denota-se por $-u$.

(g) (Associatividade do produto por escalares). Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u \in V$,

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

(h) (Distributividade em relação à soma de vectores). Para quaisquer $\alpha \in \mathbb{R}$ e $u, v \in V$,

$$\alpha (u + v) = \alpha u + \alpha v.$$

(i) (Distributividade em relação à soma de escalares). Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u \in V$,

$$(\alpha + \beta) u = \alpha u + \beta u.$$

(j) Para qualquer $u \in V$,

$$1u = u.$$

Observação 15. Aos elementos de V chamaremos vectores.

Exemplo 20. Exemplos de espaços lineares. Seja $\alpha \in \mathbb{R}$.

(i) $\mathbb{R}^n = \{(x_1, x_2, ..., x_n) : x_1, x_2, ..., x_n \in \mathbb{R}\}$, com as operações usuais:

$$(u_1, u_2, ..., u_n) + (v_1, v_2, ..., v_n) = (u_1 + v_1, u_2 + v_2, ..., u_n + v_n),$$

$$\alpha(u_1, u_2, ..., u_n) = (\alpha u_1, \alpha u_2, ..., \alpha u_n).$$

(ii) $\mathcal{M}_{m\times n}(\mathbb{R})$ (conjunto de todas as matrizes reais do tipo $m \times n$), com as operações (usuais): $A + B$ e αA.

(iii) O conjunto de todas as funções reais de variável real definidas num conjunto $S \subset \mathbb{R}$, com as operações usuais:

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

$$(\alpha f)(x) = \alpha f(x).$$

(iv) O conjunto $\mathcal{P} = \{a_0 + a_1 t + ... + a_s t^s : a_0, a_1, ..., a_s \in \mathbb{R} \text{ e } s \in \mathbb{N}_0\}$ de todos os polinómios reais de variável real, com as operações usuais.

(v) Seja $n \in \mathbb{N}$ fixo. O conjunto $\mathcal{P}_n = \{a_0 + a_1 t + ... + a_n t^n : a_0, a_1, ..., a_n \in \mathbb{R}\}$ de todos os polinómios reais de variável real e de grau menor ou igual a n, com as operações usuais.

$$(a_0 + a_1 t + \cdots + a_n t^n) + (b_0 + b_1 t + \cdots + b_n t^n) = a_0 + b_0 + (a_1 + b_1)\, t + \cdots + (a_n + b_n)\, t^n$$

$$\alpha\,(a_0 + a_1 t + \cdots + a_n t^n) = \alpha a_0 + (\alpha a_1)\, t + \cdots + (\alpha a_n)\, t^n.$$

Observação 16. Existem espaços lineares com operações não usuais:

(i) O conjunto dos números reais $\mathbb{R}$, com a soma definida por

$$u \boxplus v = u + v + 1,$$

e o produto por escalares definido por

$$\alpha \cdot u = \alpha u + \alpha - 1,$$

é um espaço linear. (Neste caso o elemento neutro é -1.)

(ii) O conjunto dos números reais maiores do que zero, com a soma definida por

$$u \boxplus v = uv,$$

e o produto por escalares definido por

$$\alpha \cdot u = u^\alpha,$$

é um espaço linear. (Neste caso o elemento neutro é 1.)

Observação 17. Alterações nos conjuntos considerados anteriormente podem resultar em conjuntos que não são espaços lineares.

(i) O conjunto $\{(x, y) \in \mathbb{R}^2 : x \geq 0 \text{ e } y \geq 0\}$, com as operações usuais, não é um espaço linear. Por exemplo, os simétricos não estão no conjunto.

(ii) O conjunto $V = \{a_0 + a_1t + ... + a_nt^n : a_0, a_1, ..., a_n \in \mathbb{R} \text{ e } a_n \neq 0\}$ de todos os polinómios reais de grau igual a n, com as operações usuais, não é um espaço linear. Por exemplo, para $n > 1$:

$$t^n, -t^n + t \in V, \quad \text{mas} \quad t^n + (-t^n + t) = t \notin V.$$

(iii) O conjunto $U = \{f : \mathbb{R} \longrightarrow \mathbb{R} \text{ tais que } f(1) = 2\}$, com as operações usuais, não é um espaço linear. Por exemplo, se $f_1, f_2 \in U$,

$$(f_1 + f_2)(1) = f_1(1) + f_2(1) = 2 + 2 = 4 \neq 2.$$

Logo, $f_1 + f_2 \notin U$.

Definição 26. Seja V um espaço linear. Diz-se que S é um **subespaço** de V se S é um subconjunto de V e se S, com as operações de V, fôr um espaço linear.

Observação 18. No entanto, para mostrar que um certo conjunto $S \subset V$ é um subespaço do espaço linear V, não será necessário verificar as 10 propriedades da definição 25, como se pode ver no seguinte teorema.

Teorema 19. Um subconjunto não vazio S de um espaço linear V é um subespaço de V se e só se:

(i) Para quaisquer $u, v \in S$ tem-se $u + v \in S$.

(ii) Para quaisquer $\alpha \in \mathbb{R}$ e $u \in S$ tem-se $\alpha u \in S$.

Exemplo 21. Exemplos de subespaços:

(i) Os únicos subespaços do espaço linear $\mathbb{R}$, com as operações usuais, são $\{0\}$ e $\mathbb{R}$.

(ii) Os subespaços do espaço linear $\mathbb{R}^3$, com as operações usuais, são: $\{(0, 0, 0)\}$, $\mathbb{R}^3$, todas as rectas que passam pela origem e todos os planos que passam pela origem.

(iii) O conjunto de todas as matrizes (reais) triangulares superiores (do tipo $n \times n$) é um subespaço do espaço linear $\mathcal{M}_{n \times n}(\mathbb{R})$, com as operações usuais.

(iv) O conjunto de todas as funções reais definidas e contínuas em $I \subset \mathbb{R}$ (I é um intervalo) é um subespaço do espaço linear de todas as funções $f : I \to \mathbb{R}$, com as operações usuais.

(v) Seja A uma matriz (real) do tipo $m \times n$. O conjunto

$$\mathcal{C}(A) = \{b \in \mathbb{R}^m : Au = b \text{ tem pelo menos uma solução } u\}$$

é um subespaço do espaço linear $\mathbb{R}^m$, com as operações usuais, ao qual se dá o nome de **espaço das colunas** de A.

(vi) Seja A uma matriz (real) do tipo $m \times n$. O conjunto

$$\mathcal{N}(A) = \{u \in \mathbb{R}^n : Au = \mathbf{0}\}$$

é um subespaço do espaço linear $\mathbb{R}^n$, com as operações usuais, ao qual se dá o nome de **espaço nulo ou núcleo** de A.

Observação 19. **(i)** Se A é invertível então $\mathcal{N}(A) = \{\mathbf{0}\}$.

(ii) Se $\mathcal{N}(A) = \{\mathbf{0}\}$ então A é invertível.

Definição 27. Seja S um subconjunto não vazio de um espaço linear V. Diz-se que um vector u é **combinação linear** finita dos elementos de S, se existir um n^o finito de elementos de S, $u_1, ..., u_k$, e de escalares $\lambda_1, ..., \lambda_k$ tais que

$$u = \lambda_1 u_1 + ... + \lambda_k u_k = \sum_{i=1}^{k} \lambda_i u_i.$$

Ao cojunto de todas as combinações lineares finitas de elementos de S chama-se **expansão linear** de S e designa-se por $L(S)$. Se S é o conjunto vazio $\varnothing$, escreve-se $L(\varnothing) = \{0\}$.

Teorema 20. Seja S um subconjunto não vazio de um espaço linear V. A expansão linear $L(S)$ de S é o menor subespaço de V que contém S. Deste modo, a $L(S)$ também se chama o **subespaço gerado** por S, e diz-se que S **gera** $L(S)$.

Observação 20. **(i)** Seja S e T dois subconjuntos não vazios de um espaço linear V, com $S \subset T$. Se $L(S) = V$ então $L(T) = V$.

(ii) Todo o subespaço do espaço linear $\mathbb{R}^n$ pode ser escrito como o núcleo de uma matriz.

Exemplo 22. **(i)** O espaço linear $\mathbb{R}^2$ é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{(1,0),(0,1)\}, \quad \{(1,2),(-1,11)\} \quad \text{e} \quad \{(23,8),(6,14)\}.$$

(ii) O subespaço $\{(x,y) \in \mathbb{R}^2 : y = 2x\}$ do espaço linear $\mathbb{R}^2$ é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{(1,2)\}, \quad \{(-2,-4)\} \quad \text{e} \quad \{(77,154)\}.$$

(iii) O espaço linear $\mathcal{P}_n$ de todos os polinómios reais de variável real e de grau menor ou igual a n, é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{1, t, t^2, ..., t^n\}, \quad \{1, 1+t, (1+t)^2, ..., (1+t)^n\} \quad \text{e} \quad \{1, \frac{t}{1!}, \frac{t^2}{2!}, ..., \frac{t^n}{n!}\}.$$

(iv) O espaço linear $\mathcal{P}$ de todos os polinómios reais de variável real, é gerado pelo conjunto infinito de vectores:

$$\{1, t, t^2, ...\}.$$

(v) Seja U o espaço linear de todas as funções reais com primeira derivada contínua em $\mathbb{R}$ (isto é, pertencentes a $C^1(\mathbb{R})$) e tais que $f'(x) = af(x)$ (em $\mathbb{R}$) com $a \in \mathbb{R}$. Então U é gerado pela função $g(x) = e^{ax}$, tendo-se $U = L(\{g\})$.

(vi) Seja A uma matriz (real) do tipo $m \times n$. O espaço das colunas de A,

$$\mathcal{C}(A) = \{b \in \mathbb{R}^m : Au = b \text{ tem pelo menos uma solução } u\},$$

é o subespaço (do espaço linear $\mathbb{R}^m$) gerado pelas colunas de A, uma vez que:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = u_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + u_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + ... + u_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

(vii) Seja A uma matriz (real) do tipo $m \times n$. Ao subespaço linear de $\mathbb{R}^n$ gerado pelas linhas de A dá-se o nome de **espaço das linhas** de A e designa-se por $\mathcal{L}(A)$.

(viii) Sejam

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -3 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 2 \\ 2 & -4 \\ -2 & 4 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}.$$

Tem-se

$$\mathcal{C}(A) = \{(0,0)\}, \qquad \mathcal{N}(A) = \mathbb{R}^3, \qquad \mathcal{L}(A) = \{(0,0,0)\}.$$

$$\mathcal{C}(B) = L(\{(1,0,0), (1,7,0)\}), \quad \mathcal{N}(B) = L(\{(3,1,0)\}), \quad \mathcal{L}(B) = L(\{(1,-3,1), (0,0,7)\}).$$

$$\mathcal{C}(C) = L(\{(-1,2,-2)\}), \qquad \mathcal{N}(C) = L(\{(2,1)\}) \qquad \text{e} \qquad \mathcal{L}(C) = L(\{(-1,2)\}).$$

$$\mathcal{C}(D) = L(\{(2,0), (0,-1)\}), \ \mathcal{N}(D) = \{(0,0)\} \ \text{ e } \ \mathcal{L}(D) = L(\{(2,0), (0,-1)\}).$$

(ix) Seja $U = \{A \in \mathcal{M}_{3\times 2}(\mathbb{R}) : a_{12} = a_{21} = a_{32} = 0 \text{ e } a_{11} + 2a_{31} = 0\}$. Tem-se, para $A \in U$,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} -2a_{31} & 0 \\ 0 & a_{22} \\ a_{31} & 0 \end{bmatrix} = a_{31} \begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

com $a_{31}, a_{22} \in \mathbb{R}$. Logo,

$$U = L\left(\left\{ \begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}\right).$$

(x) Seja $U = \{p(t) = a_0 + a_1 t + a_2 t^2 \in \mathcal{P}_2 : p(1) = p(0)\}$. Tem-se, para $p(t) \in U$,

$$p(1) = p(0) \Leftrightarrow a_0 + a_1 + a_2 = a_0 \Leftrightarrow a_1 + a_2 = 0 \Leftrightarrow a_1 = -a_2.$$

Logo, $p(t) = a_0 - a_2 t + a_2 t^2 = a_0 1 + a_2 \left(-t + t^2\right)$, com $a_0, a_2 \in \mathbb{R}$. Assim,

$$U = L\left(\left\{1, -t + t^2\right\}\right).$$

Teorema 21. Se U e V são subespaços do espaço linear W, então:

(i) O conjunto $U \cap V$ é um subespaço linear de W.

(ii) O conjunto $U + V = \{u + v : u \in U \quad \text{e} \quad v \in V\}$ é um subespaço de W. É o menor subespaço de W que contém $U \cup V$. O conjunto $U \cup V$ em geral não é um subespaço. Escreve-se $U + V = L(U \cup V)$.

Exemplo 23. **(i)** Em $\mathbb{R}^3$, considere os subespaços:

$$U = \{(x, y, z) \in \mathbb{R}^3 : x + y - 2z = 0\} \quad \text{e} \quad V = L\left(\{(1, 1, -1), (1, 2, 1)\}\right).$$

Seja $v \in V$, então

$$v = \alpha(1, 1, -1) + \beta(1, 2, 1) = (\alpha + \beta, \alpha + 2\beta, -\alpha + \beta),$$

com $\alpha, \beta \in \mathbb{R}$. Para que v esteja também em U é preciso que:

$$(\alpha + \beta) + (\alpha + 2\beta) - 2\left(-\alpha + \beta\right) = 0.$$

A última equação é equivalente a $4\alpha + \beta = 0 \Leftrightarrow \beta = -4\alpha$. Logo,

$$U \cap V = \{(-3\alpha, -7\alpha, -5\alpha) : \alpha \in \mathbb{R}\} = \{\alpha(-3, -7, -5) : \alpha \in \mathbb{R}\} = L\left(\{(3, 7, 5)\}\right).$$

(ii) Em $\mathbb{R}^3$, considere os subespaços:

$$U = L\left(\{(1, -1, 1), (1, 2, 2)\}\right) \quad \text{e} \quad V = L\left(\{(2, 1, 1), (-1, 1, 3)\}\right).$$

Seja $v \in U$, então

$$v = \alpha(1,-1,1) + \beta(1,2,2) = (\alpha+\beta, -\alpha+2\beta, \alpha+2\beta),$$

com $\alpha, \beta \in \mathbb{R}$. Para que v esteja também em V é preciso que:

$$(\alpha+\beta, -\alpha+2\beta, \alpha+2\beta) = \lambda(2,1,1) + \mu(-1,1,3) = (2\lambda-\mu, \lambda+\mu, \lambda+3\mu)\,,$$

com $\lambda, \mu \in \mathbb{R}$. Deste modo,

$$\begin{cases} \alpha+\beta = 2\lambda-\mu \\ -\alpha+2\beta = \lambda+\mu \\ \alpha+2\beta = \lambda+3\mu. \end{cases}$$

Considerando a matriz aumentada tem-se

$$\begin{bmatrix} 1 & 1 & | & 2\lambda-\mu \\ -1 & 2 & | & \lambda+\mu \\ 1 & 2 & | & \lambda+3\mu \end{bmatrix} \underset{\substack{L_1+L_2\to L_2 \\ -L_1+L_3\to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 1 & | & 2\lambda-\mu \\ 0 & 3 & | & 3\lambda \\ 0 & 1 & | & -\lambda+4\mu \end{bmatrix} \underset{-\frac{1}{3}L_2+L_3\to L_3}{\longrightarrow} \begin{bmatrix} 1 & 1 & | & 2\lambda-\mu \\ 0 & 3 & | & 3\lambda \\ 0 & 0 & | & -2\lambda+4\mu \end{bmatrix}.$$

Logo,

$$\begin{cases} \alpha+\beta = 2\lambda-\mu \\ \beta = \lambda \\ 0 = -2\lambda+4\mu. \end{cases} \Leftrightarrow \begin{cases} \alpha = \mu \\ \beta = 2\mu \\ \lambda = 2\mu. \end{cases}$$

Assim,

$$\alpha(1,-1,1) + \beta(1,2,2) = \mu(1,-1,1) + 2\mu(1,2,2) = (3\mu, 3\mu, 5\mu) = \mu(3,3,5).$$

Logo,

$$U \cap V = \{(3\mu, 3\mu, 5\mu) : \mu \in \mathbb{R}\} = \{\mu(3,3,5) : \mu \in \mathbb{R}\} = L\left(\{(3,3,5)\}\right).$$

Observação 21. Neste exemplo (ii), os subespaços U e V poderiam ter sido apresentados inicialmente na forma:

$$U = \{(x,y,z) \in \mathbb{R}^3 : 4x+y-3z=0\} \quad \text{e} \quad V = \{(x,y,z) \in \mathbb{R}^3 : 2x-7y+3z=0\},$$

uma vez que

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & 2 \\ x & y & z \end{bmatrix} \underset{\substack{-L_1+L_2\to L_2 \\ -xL_1+L_3\to L_3}}{\longrightarrow} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & 1 \\ 0 & x+y & z-x \end{bmatrix} \underset{-\frac{1}{3}(x+y)L_2+L_3\to L_3}{\longrightarrow} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & z-\frac{4}{3}x-\frac{1}{3}y \end{bmatrix}$$

e logo $z-\frac{4}{3}x-\frac{1}{3}y = 0 \Leftrightarrow 4x+y-3z=0$. Por outro lado,

$$U = \{(x,y,z) \in \mathbb{R}^3 : 4x+y-3z=0\} = L\left(\{(1,-4,0),(0,3,1)\}\right) = L\left(\{(1,-1,1),(1,2,2)\}\right)$$

pois sendo $y = -4x+3z$,

$$U = \{(x,-4x+3z,z) : x,z \in \mathbb{R}\} = \{x(1,-4,0)+z(0,3,1) : x,z \in \mathbb{R}\} = L\left(\{(1,-4,0),(0,3,1)\}\right).$$

Como

$$(1,-4,0)) = 2(1,-1,1)-(1,2,2) \quad \text{e} \quad (0,3,1) = -(1,-1,1)+(1,2,2)$$

$$(1,-1,1)=(1,-4,0)+(0,3,1) \quad \text{e} \quad (1,2,2)=(1,-4,0)+2(0,3,1)$$

ou seja

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

em que $\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$, tem-se $U = L\left(\{(1,-4,0),(0,3,1)\}\right) = L\left(\{(1,-1,1),(1,2,2)\}\right)$.

Analogamente se mostra que

$V = \{(x,y,z) \in \mathbb{R}^3 : 2x-7y+3z=0\} = L\left(\{(7,2,0),(-3,0,2)\}\right) = L\left(\{(2,1,1),(-1,1,3)\}\right).$

(iii) Seja U o subespaço de $\mathcal{M}_{n\times n}(\mathbb{R})$ das matrizes triangulares superiores e seja V o subespaço de $\mathcal{M}_{n\times n}(\mathbb{R})$ das matrizes triangulares inferiores. Então

$$U+V=\mathcal{M}_{n\times n}(\mathbb{R}) \quad \text{e} \quad U\cap V = \text{ subespaço das matrizes diagonais.}$$

(iv) Sejam $U=L(\{(1,0)\})$ e $V=L(\{(0,1)\})$ subespaços de $\mathbb{R}^2$. O conjunto

$$U\cup V=\{(x,y)\in\mathbb{R}^2 : x=0 \vee y=0\}$$

não é um espaço linear pois $\underbrace{(1,0)}_{\in U}+\underbrace{(0,1)}_{\in V}=(1,1)\notin U\cup V$. No entanto, tem-se $U+V=\mathbb{R}^2$.

Teorema 22. Se U e V subespaços do espaço linear W, então $U\cup V$ é subespaço de W se e só se $U\subset V$ ou $V\subset U$.

Teorema 23. Sejam W_1 e W_2 subespaços de um espaço linear V tais que $W_1\cap W_2=\{\mathbf{0}\}$. Se $V=W_1+W_2$ então todo o vector $v\in V$ pode ser escrito de modo único na forma

$$v=w_1+w_2$$

com $w_1\in W_1$ e $w_2\in W_2$. Neste caso escreve-se $V=W_1\oplus W_2$ e diz-se que V é a **soma directa** dos espaços W_1 e W_2.

Teorema 24. Seja $A\in\mathcal{M}_{m\times n}(\mathbb{R})$. Tem-se $\mathcal{C}(A)=\mathcal{L}(A^T)$ e $\mathcal{L}(A)\cap\mathcal{N}(A)=\{\mathbf{0}\}$.

Observação 22. Seja $A\in\mathcal{M}_{m\times n}(\mathbb{R})$. Se A' fôr a matriz em escada que se obtem de A por aplicação do método de eliminação de Gauss, tem-se

$$\mathcal{C}(A)\neq\mathcal{C}(A').$$

Teorema 25. Seja $A\in\mathcal{M}_{m\times n}(\mathbb{R})$. O espaço das linhas $\mathcal{L}(A)$ e o núcleo $\mathcal{N}(A)$ mantêm-se invariantes por aplicação do método de eliminação de Gauss. Isto é, sendo A' a matriz em escada que se obtem de A por aplicação desse método, tem-se

$$\mathcal{L}(A)=\mathcal{L}(A') \quad \text{e} \quad \mathcal{N}(A)=\mathcal{N}(A').$$

Independência linear

Definição 28. Seja V um espaço linear. Seja

$$S = \{v_1, v_2, ..., v_k\} \subset V.$$

Diz-se que o conjunto S é **linearmente dependente** se e só se algum dos vectores de S se escrever como combinação linear dos restantes, isto é, se e só se existir algum $i \in \{1, 2, ..., k\}$ e escalares $\lambda_1, \lambda_2, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_k \in \mathbb{R}$ tais que

$$v_i = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + ... + \lambda_k v_k.$$

Definição 29. Seja V um espaço linear. Seja

$$S = \{v_1, v_2, ..., v_k\} \subset V.$$

Diz-se que o conjunto S é **linearmente independente** se e só se nenhum dos vectores de S se puder escrever como combinação linear dos restantes, isto é, se e só a única solução do sistema homogéneo

$$\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k = \mathbf{0}$$

fôr a solução trivial, ou seja, $\lambda_1 = \lambda_2 = ... = \lambda_k = 0$. No caso em que $V = \mathbb{R}^n$, sendo A a matriz cujas colunas são os vectores de $S \subset V$, diz-se que S é **linearmente independente** se e só se $\mathcal{N}(A) = \{\mathbf{0}\}$.

Teorema 26. Seja A' uma matriz em escada de linhas.

(i) As colunas de A' que contêm pivots são linearmente independentes.

(ii) As linhas não nulas de A' são linearmente independentes.

(iii) O n^o de linhas independentes e o n^o de colunas independentes (de A') são ambos iguais à característica de A'.

Observação 23. **(i)** Assim, atendendo ao teorema anterior, a independência linear de $S = \{v_1, v_2, ..., v_k\} \subset V$ (espaço linear) pode ser decidida aplicando o método de eliminação à matriz A cujas colunas são os vectores de S, de modo a colocá-la em escada de linhas. Sendo A' essa matriz em escada, tem-se pelo teorema 25

$$\mathcal{N}(A) = \mathcal{N}(A') \quad (*).$$

Uma vez que as colunas de A' que contêm pivots são linearmente independentes então, devido a (*), as colunas de A nas posições correspondentes também serão linearmente independentes.

(ii) Em $\mathbb{R}$, quaisquer dois vectores são linearmente dependentes.

(iii) Em $\mathbb{R}^2$, dois vectores são linearmente independentes se não forem colineares.

(iv) Em $\mathbb{R}^3$, três vectores são linearmente independentes se não forem coplanares.

(v) Qualquer conjunto que contenha o vector nulo (elemento neutro) é linearmente dependente. Em particular, o conjunto $\{\mathbf{0}\}$, formado apenas pelo vector nulo, é linearmente dependente.

(vi) O conjunto vazio $\varnothing$ é linearmente independente.

Teorema 27. Sejam S_1 e S_2 dois subconjuntos finitos de um espaço linear, tais que $S_1 \subset S_2$.

(i) Se S_1 é linearmente dependente então S_2 também é linearmente dependente.

(ii) Se S_2 é linearmente independente então S_1 também é linearmente independente.

Observação 24. Sejam S_1 e S_2 dois subconjuntos finitos de um espaço linear, tais que $S_1 \subset S_2$.

(i) Se S_2 fôr linearmente dependente então S_1 tanto pode ser linearmente dependente como linearmente independente.

(ii) Se S_1 fôr linearmente independente então S_2 tanto pode ser linearmente dependente como linearmente independente.

Exemplo 24. Seja $S = \{(1,0,2),(2,0,4),(0,1,2)\}$. Tem-se

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 2 & 4 & 2 \end{bmatrix} \underset{-2L_1+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \underset{-2L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = A'.$$

Logo, como apenas existem dois pivots e portanto uma variável livre, as três colunas de A são linearmente dependentes, isto é, o conjunto S é linearmente dependente. O subconjunto de S:

$$\{(1,0,2),(2,0,4)\}$$

também é linearmente dependente. No entanto, uma vez que a 1^a e 3^a colunas de A são independentes pois correspondem às colunas da matriz em escada A' que contêm os pivots, o subconjunto de S:

$$\{(1,0,2),(0,1,2)\}$$

é linearmente independente.

Bases e Dimensão de um Espaço Linear

Definição 30. Chama-se **base** de um espaço linear V a qualquer subconjunto $\mathcal{B}$ de V que verifique as duas condições:

(i) $\mathcal{B}$ gera V, isto é, $L(\mathcal{B}) = V$.

(ii) $\mathcal{B}$ é linearmente independente.

Teorema 28. Qualquer espaço linear $V \neq \{\mathbf{0}\}$ tem pelo menos uma base.

Observação 25. Qualquer espaço linear $V \neq \{\mathbf{0}\}$ tem um n^o infinito de bases. Por exemplo, se $\mathcal{B} = \{u_1, ..., u_k\}$ fôr uma base de V então para cada $\alpha \neq 0$ o conjunto $\{\alpha u_1, ..., \alpha u_k\}$ é também uma base de V.

Teorema 29. Todas as bases de um espaço linear $V \neq \{\mathbf{0}\}$ têm o mesmo n^o de vectores.

Definição 31. Chama-se **dimensão** de um espaço linear $V \neq \{\mathbf{0}\}$ ao n^o de vectores de uma base qualquer de V, e escreve-se $\dim V$. Se $V = \{\mathbf{0}\}$ então $\dim V = 0$ uma vez que o conjunto vazio $\varnothing$ é base de $\{\mathbf{0}\}$. Um espaço linear terá dimensão finita se uma sua base tiver um n^o finito de vectores.

Observação 26. A dimensão de um espaço linear, isto é, o n^o de elementos de uma sua base é igual ao n^o mínimo de vectores possam constituir um conjunto gerador desse espaço e é também igual ao n^o máximo de vectores que possam constituir um conjunto linearmente independente nesse espaço.

Exemplo 25. **(i)** O conjunto $\{1\}$ é uma base de $\mathbb{R}$, chamada base canónica ou natural de $\mathbb{R}$. Logo,

$$\dim \mathbb{R} = 1.$$

(ii) O conjunto $\{(1,0),(0,1)\}$ é uma base de $\mathbb{R}^2$, chamada base canónica ou natural de $\mathbb{R}^2$. Logo,

$$\dim \mathbb{R}^2 = 2.$$

(iii) O conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$ é uma base de $\mathbb{R}^3$, chamada base canónica ou natural de $\mathbb{R}^3$. Logo,

$$\dim \mathbb{R}^3 = 3.$$

(iv) O conjunto

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

é uma base de $\mathcal{M}_{2\times 3}(\mathbb{R})$, chamada base canónica ou natural de $\mathcal{M}_{2\times 3}(\mathbb{R})$. Logo,

$$\dim \mathcal{M}_{2\times 3}(\mathbb{R}) = 6.$$

(v) Tem-se

$$\dim \mathbb{R}^n = n \quad \text{e} \quad \dim \mathcal{M}_{m\times n}(\mathbb{R}) = mn.$$

(vi) O conjunto $\{1, t, t^2, ..., t^n\}$ é uma base de $\mathcal{P}_n$ (espaço linear de todos os polinómios reais de variável real e de grau menor ou igual a n), chamada base canónica ou natural de $\mathcal{P}_n$. Logo,

$$\dim \mathcal{P}_n = n + 1.$$

(vii) O conjunto $\{1, t, t^2, ...\}$ é uma base de $\mathcal{P}$ (espaço linear de todos os polinómios reais de variável real), chamada base canónica ou natural de $\mathcal{P}$. Logo,

$$\dim \mathcal{P} = \infty.$$

Observação 27. Seja A uma matriz do tipo $m \times n$. Tem-se

$$\operatorname{nul} A = \dim \mathcal{N}(A) \qquad \text{e} \qquad \operatorname{car} A = \dim \mathcal{L}(A).$$

Teorema 30. Seja A uma matriz do tipo $m \times n$. Tem-se

$$\dim \mathcal{C}(A) = \dim \mathcal{L}(A) = \operatorname{car} A.$$

Dem. Suponhamos que $\operatorname{car} A = k$. Sendo A' a matriz $m \times n$ em escada (reduzida) de linhas, então A' tem exactamente k linhas não nulas. Sejam $R_1, R_2, \ldots, R_k$ essas linhas. Como $\mathcal{L}(A) = \mathcal{L}(A')$, então as linhas $L_1, L_2, \ldots, L_m$ de A podem ser expressas como combinações lineares das linhas $R_1, R_2, \ldots, R_k$, ou seja, existem escalares c_{ij}, com $i = 1, \ldots, m$ e $j = 1, \ldots, k$ tais que

$$\begin{gathered} L_1 = c_{11}R_1 + c_{12}R_2 + \cdots + c_{1k}R_k \\ L_2 = c_{21}R_1 + c_{22}R_2 + \cdots + c_{2k}R_k \\ \cdots \\ L_m = c_{m1}R_1 + c_{m2}R_2 + \cdots + c_{mk}R_k \end{gathered}$$

Para $i = 1, \ldots, m$, sejam a_{ij} e r_{ij} as componentes j das linhas L_i e R_i respectivamente. Assim, tem-se

$$\begin{gathered} a_{1j} = c_{11}r_{1j} + c_{12}r_{2j} + \cdots + c_{1k}r_{kj} \\ a_{2j} = c_{21}r_{1j} + c_{22}r_{2j} + \cdots + c_{2k}r_{kj} \\ \cdots \\ a_{mj} = c_{m1}r_{1j} + c_{m2}r_{2j} + \cdots + c_{mk}r_{kj} \end{gathered}$$

ou seja, matricialmente,

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = r_{1j} \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + r_{2j} \begin{bmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{bmatrix} + \cdots + r_{kj} \begin{bmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{bmatrix}.$$

Como $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$ é a coluna j de A, a última igualdade mostra que os vectores

$$\begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix}, \begin{bmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{bmatrix}$$

geram $\mathcal{C}(A)$. Logo, tem-se

$$\dim \mathcal{C}(A) \leq k = \dim \mathcal{L}(A).$$

Deste modo, substituindo A por A^T tem-se também

$$\underbrace{\dim \mathcal{C}\left(A^T\right)}_{=\dim \mathcal{L}(A)} \leq \underbrace{\dim \mathcal{L}\left(A^T\right)}_{=\dim \mathcal{C}(A)}.$$

Ou seja, tem-se

$$\dim \mathcal{C}(A) \leq \dim \mathcal{L}(A) \qquad \text{e} \qquad \dim \mathcal{L}(A) \leq \dim \mathcal{C}(A).$$

Isto é,

$$\dim \mathcal{C}(A) = \dim \mathcal{L}(A).$$

Teorema 31. Sejam W_1 e W_2 dois subespaços de dimensão finita de um espaço linear V. Então,

$$\dim (W_1 + W_2) = \dim W_1 + \dim W_2 - \dim (W_1 \cap W_2).$$

Teorema 32. Sejam V um espaço linear de dimensão finita e W um subespaço de V.

(i) Seja $S = \{u_1, ..., u_k\} \subset V$. Se S é linearmente independente então S será um subconjunto de uma base de V e ter-se-á $\dim V \geq k$.

(ii) Se $\dim V = n$, então quaisquer m vectores de V, com $m > n$, são linearmente dependentes.

(iii) Se $\dim V = n$, então nenhum conjunto com m vectores de V, em que $m < n$, pode gerar V.

(iv) O subespaço W tem dimensão finita e $\dim W \leq \dim V$.

(v) Se $\dim W = \dim V$, então $W = V$.

(vi) Se $\dim V = n$, então quaisquer n vectores de V linearmente independentes constituem uma base de V.

(vii) Se $\dim V = n$, então quaisquer n vectores geradores de V constituem uma base de V.

Exemplo 26. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Recorde que:

$$\operatorname{car} A + \operatorname{nul} A = n.$$

Como $\mathcal{L}(A)$ e $\mathcal{N}(A)$ são subespaços de $\mathbb{R}^n$ então

$$\mathcal{L}(A) + \mathcal{N}(A) = L\left(\mathcal{L}(A) \cup \mathcal{N}(A)\right)$$

é também um subepaço de $\mathbb{R}^n$. Por outro lado, atendendo a que

$$\mathcal{L}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}$$

(teorema 24), tem-se

$$\dim(\mathcal{L}(A) \cap \mathcal{N}(A)) = 0.$$

Assim,

$$\begin{aligned}
\dim(\mathcal{L}(A) + \mathcal{N}(A)) &= \dim \mathcal{L}(A) + \dim \mathcal{N}(A) - \dim(\mathcal{L}(A) \cap \mathcal{N}(A)) = \\
&= \operatorname{car} A + \operatorname{nul} A - 0 = \\
&= n.
\end{aligned}$$

Logo, pelo teorema 32 (v), tem-se

$$\mathbb{R}^n = \mathcal{L}(A) \oplus \mathcal{N}(A).$$

Exemplo 27. **(i)** Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}$:

$$\{0\} \quad \text{e} \quad \mathbb{R}.$$

(ii) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}^2$:

$$\{(0,0)\}\,, \quad \text{todas as rectas que contêm a origem} \quad \text{e} \quad \mathbb{R}^2.$$

(iii) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}^3$:

$$\{(0,0,0)\}\,, \text{ todas as rectas que contêm a origem,}$$

$$\text{todos os planos que contêm a origem e } \mathbb{R}^3.$$

Observação 28. O método de eliminação de Gauss permite determinar a dimensão e uma base quer para o espaço das linhas $\mathcal{L}(A)$ quer para o espaço das colunas $\mathcal{C}(A)$ de uma matriz A. Seja A' a matriz em escada que se obtem de A por aplicação do método de eliminação de Gauss. Então,

(i) Uma base para $\mathcal{L}(A)$ será formada pelas linhas não nulas de A'.

(ii) Uma base para $\mathcal{C}(A)$ será formada pelas colunas de A que correspondem às posições das colunas de A' que contêm os pivots.

Exemplo 28. Seja

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

Tem-se

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 3 & 3 \\ -6 & -3 & 1 & 1 \end{bmatrix} \underset{\substack{-2L_1+L_2\to L_2 \\ 3L_1+L_3\to L_3}}{\longrightarrow} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 4 \end{bmatrix} \underset{-4L_2+L_3\to L_3}{\longrightarrow} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A'.$$

Logo, $\{(2,1,1,1),(0,0,1,1)\}$ é uma base de $\mathcal{L}(A)$ e $\{(2,4,-6),(1,3,1)\}$ é uma base de $\mathcal{C}(A)$. Assim,

$$\dim \mathcal{L}(A) = 2 = \dim \mathcal{C}(A)$$

e

$$\mathcal{L}(A) = L\left(\{(2,1,1,1),(0,0,1,1)\}\right), \;\; \mathcal{C}(A) = L\left(\{(2,4,-6),(1,3,1)\}\right).$$

Por outro lado,

$$\begin{aligned} \mathcal{N}(A') &= \left\{ (x,y,z,w) \in \mathbb{R}^4 : A' \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\} = \\ &= \{(x,-2x,-w,w) : x,w \in \mathbb{R}\} = \\ &= L\left(\{(1,-2,0,0),(0,0,-1,1)\}\right). \end{aligned}$$

Como o conjunto $\{(1,-2,0,0),(0,0,-1,1)\}$ é linearmente independente e gera $\mathcal{N}(A')$ então é uma base de $\mathcal{N}(A')$. Finalmente, uma vez que $\mathcal{N}(A) = \mathcal{N}(A')$, o conjunto

$$\{(1,-2,0,0),(0,0,-1,1)\}$$

é uma base de $\mathcal{N}(A)$ e portanto $\dim \mathcal{N}(A) = 2$, com

$$\mathcal{N}(A) = L\left(\{(1,-2,0,0),(0,0,-1,1)\}\right).$$

Exemplo 29. Seja $S = \{1,2,-1),(2,1,1),(-1,-2,1),(0,1,0)\} \subset \mathbb{R}^3$. Determinemos uma base para $L(S)$.

Considerando a matriz cujas colunas são os vectores de S, tem-se

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & -2 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix} \underset{\substack{-2L_1+L_2\to L_2 \\ L_1+L_3\to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 3 & 0 & 0 \end{bmatrix} \underset{L_2+L_3\to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Logo, $S' = \{1,2,-1),(2,1,1),(0,1,0)\}$ é uma base de $L(S)$. Como $\dim \mathbb{R}^3 = 3$, então tem-se mesmo: $L(S) = \mathbb{R}^3$ e S' é uma base de $\mathbb{R}^3$.

Resolução alternativa: Considerando a matriz cujas colunas são os vectores de S, tem-se

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underset{\substack{-2L_1+L_2\to L_2 \\ L_1+L_3\to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \underset{L_3\leftrightarrow L_4}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underset{\frac{1}{3}L_2+L_3\to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Logo, $S' = \{1, 2, -1), (0, -3, 3), (0, 0, 1)\}$ é uma base de $L(S)$. Como $\dim \mathbb{R}^3 = 3$, então tem-se mesmo: $L(S) = \mathbb{R}^3$ e S' é uma base de $\mathbb{R}^3$.

Exemplo 30. Seja $S_{a,b} = \{1, 0, 1), (0, 1, a), (1, 1, b), (1, 1, 1)\} \subset \mathbb{R}^3$. Determinemos os valores dos parâmetros a e b para os quais $S_{a,b}$ não gere $\mathbb{R}^3$.

Considerando a matriz cujas colunas são os vectores de S, tem-se

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & a & b & 1 \end{bmatrix} \underset{-L_1+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & a & b-1 & 0 \end{bmatrix} \underset{-aL_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & b-a-1 & -a \end{bmatrix}.$$

Logo, $S_{a,b}$ não gera $\mathbb{R}^3$ se e só se $b - a - 1 = 0$ e $-a = 0$, isto é, se e só se $a = 0$ e $b = 1$.

Teorema 33. **(i)** Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. As colunas de A geram $\mathbb{R}^m$ se e só se $\operatorname{car} A = m$.

(ii) Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. As colunas de A são linearmente independentes se e só se $\operatorname{car} A = n$.

(iii) Seja $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. A matriz A é invertível se e só se as colunas de A (ou as linhas de A) formarem uma base de $\mathbb{R}^n$. No caso de A ser invertível tem-se

$$\mathcal{C}(A) = \mathcal{L}(A) = \mathbb{R}^n.$$

Observação 29. Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e considere o sistema de equações lineares $Au = b$.

(i) O sistema $Au = b$ é **impossível** (não tem solução) se e só se $b \notin \mathcal{C}(A)$, isto é, se e só se $\operatorname{car} A < \operatorname{car} [A \mid b]$.

(ii) O sistema $Au = b$ é **possível e indeterminado** (tem um n^o infinito de soluções) se e só se $b \in \mathcal{C}(A)$ e as colunas de A forem linearmente dependentes, isto é, se e só se $\operatorname{car} A = \operatorname{car} [A \mid b] < n$, isto é, se e só se $\operatorname{car} A = \operatorname{car} [A \mid b]$ e $\operatorname{nul} A \neq 0$.

(iii) O sistema $Au = b$ é **possível e determinado** (tem uma única solução) se e só se $b \in \mathcal{C}(A)$ e as colunas de A forem linearmente independentes, isto é, se e só se $\operatorname{car} A = \operatorname{car} [A \mid b] = n$, isto é, se e só se $\operatorname{car} A = \operatorname{car} [A \mid b]$ e $\operatorname{nul} A = 0$.

Observação 30. Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e considere o sistema de equações lineares $Au = b$.

(i) Existência de solução: Se $m \leq n$ então o sistema $Au = b$ tem pelo menos uma solução u para cada $b \in \mathbb{R}^m$ se e só se $\operatorname{car} A = m$.

(ii) Unicidade de solução: Se $m \geq n$ então o sistema $Au = b$ tem no máximo uma solução u para cada $b \in \mathbb{R}^m$ se e só se $\operatorname{car} A = n$, isto é, se e só se $\operatorname{nul} A = 0$.

(iii) Existência e unicidade de solução: Se $m = n$ então o sistema $Au = b$ tem solução única u para cada $b \in \mathbb{R}^m$ se e só se A fôr invertível.

Teorema 34. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. As seguintes afirmações são equivalentes.

(i) A é não singular.

(ii) A é igual ao produto de matrizes elementares.

(iii) A é invertível.

(iv) $Au = \mathbf{0}$ tem apenas a solução trivial $u = \mathbf{0}$.

(v) $Au = b$ tem solução única u para cada $b \in \mathbb{R}^n$.

(vi) $\mathcal{N}(A) = \{\mathbf{0}\}$.

(vii) $\operatorname{nul} A = 0$.

(viii) $\operatorname{car} A = n$.

(ix) As colunas de A geram $\mathbb{R}^n$.

(x) As colunas de A são independentes.

(xi) As colunas de A formam uma base de $\mathbb{R}^n$.

(xii) As linhas de A geram $\mathbb{R}^n$.

(xiii) As linhas de A são independentes.

(xiv) As linhas de A formam uma base de $\mathbb{R}^n$.

(xv) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é sobrejectiva. (Num próximo capítulo.)

(xvi) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é injectiva. (Num próximo capítulo.)

(xvii) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é bijectiva. (Num próximo capítulo.)

(xviii) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é invertível. (Num próximo capítulo.)

(xix) $\det A \neq 0$.

(xx) 0 não é valor próprio de A. (Num próximo capítulo.)

(xxi) $A^T A$ é invertível.

(xxii) $(\mathcal{N}(A))^\perp = \mathbb{R}^n$. (Num próximo capítulo.)

(xxiii) $(\mathcal{L}(A))^\perp = \{\mathbf{0}\}$. (Num próximo capítulo.)

Coordenadas de um vector numa base e matriz de mudança de base

Definição 32. Seja $\mathcal{B} = \{v_1, v_2, ..., v_k\}$ uma base ordenada de um espaço linear V e seja u um vector de V. Chamam-se **coordenadas** do vector u na base ordenada $\mathcal{B}$ aos escalares $\lambda_1, \lambda_2, ..., \lambda_k$ da combinação linear:

$$u = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_k v_k.$$

Teorema 35. Seja V um espaço linear.

(i) Um conjunto $\mathcal{B}$ de vectores não nulos de V é uma base de V se e só se todo o vector de V puder ser escrito de modo único como combinação linear dos vectores de $\mathcal{B}$.

(ii) Se $\dim V = n$, então dados $u, w \in V$ e $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma base ordenada de V, tem-se $u = w$ se e só se as coordenadas de u e de w na base $\mathcal{B}$ forem iguais.

Teorema 36. Seja V um espaço linear de dimensão n. Sejam $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_n\}$ e $\mathcal{B}_2 = \{w_1, w_2, \ldots, w_n\}$ duas bases ordenadas de V. Seja $S_{\mathcal{B}_1 \to \mathcal{B}_2}$ a matriz cujas colunas são as coordenadas dos vectores de $\mathcal{B}_1$ em relação à base $\mathcal{B}_2$. Isto é,

$$S_{\mathcal{B}_1 \to \mathcal{B}_2} = (s_{ij})_{n \times n} \quad \text{com} \quad v_j = \sum_{i=1}^{n} s_{ij} w_i \quad \text{para todo o } j = 1, ..., n.$$

A matriz $S_{\mathcal{B}_1 \to \mathcal{B}_2}$ é não singular e chama-se **matriz de mudança de base** (da base $\mathcal{B}_1$ para $\mathcal{B}_2$). Assim, se tivermos

$$u = \sum_{i=1}^{n} \lambda_i v_i,$$

isto é, se $(\lambda_1, ..., \lambda_n)$ forem as coordenadas do vector u na base $\mathcal{B}_1$ então as coordenadas $(\mu_1, ..., \mu_n)$ de u na base $\mathcal{B}_2$ são dadas por

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = S_{\mathcal{B}_1 \to \mathcal{B}_2} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Dem. Tem-se

$$u = \sum_{i=1}^{n} \mu_i w_i = \sum_{j=1}^{n} \lambda_j v_j = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} s_{ij} w_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} s_{ij} \lambda_j \right) w_i.$$

Atendendo ao teorema 35 (i), as coordenadas de um vector u numa base são únicas. Logo, para todo o $i = 1, ..., n$,

$$\mu_i = \left(\sum_{j=1}^{n} s_{ij} \lambda_j \right). \quad \text{Isto é,} \quad \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = S_{\mathcal{B}_1 \to \mathcal{B}_2} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Observação 31. Tem-se
$$S_{\mathcal{B}_2 \to \mathcal{B}_1} = \left(S_{\mathcal{B}_1 \to \mathcal{B}_2}\right)^{-1}.$$

Exemplo 31. Consideremos
$$\mathcal{B}_c = \{(1,0), (0,1)\}$$
a base canónica de $\mathbb{R}^2$. Seja
$$\mathcal{B} = \{(1,2), (3,4)\}$$
uma outra base ordenada de $\mathbb{R}^2$. Sejam $(5,6)$ as coordenadas de um vector u na base canónica $\mathcal{B}_c$ e determinemos as coordenadas de u na base $\mathcal{B}$ usando a matriz de mudança de base $S_{\mathcal{B}_c \to \mathcal{B}}$. Tem-se
$$S_{\mathcal{B}_c \to \mathcal{B}} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix},$$
uma vez que
$$(1,0) = -2(1,2) + 1(3,4) \quad \text{e} \quad (0,1) = \frac{3}{2}(1,2) - \frac{1}{2}(3,4). \quad (*)$$
Logo, as coordenadas de u na base $\mathcal{B}$ são dadas por
$$S_{\mathcal{B}_c \to \mathcal{B}} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$
Logo, -1 e 2 são as coordenadas de $(5,6)$ na base ordenada $\mathcal{B}$, isto é
$$(5,6) = -1(1,2) + 2(3,4).$$

Observação 32. Colocando os vectores em coluna, note que as duas igualdades em (*) podem ser escritas na forma:
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}$$
sendo esta última igualdade equivalente a
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}}_{= \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}^{-1}} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \Leftrightarrow \begin{cases} (1,2) = 1(1,0) + 2(0,1) \\ \\ (3,4) = 3(1,0) + 4(0,1) \end{cases}$$
querendo isto dizer que as coordenadas dos vectores $(1,2)$ e $(3,4)$ relativamente à base canónica (ordenada) $\{(1,0), (0,1)\}$ são respectivamente $(1,2)$ e $(3,4)$.

Transformações Lineares

Definição 33. Sejam U e V espaços lineares. Diz-se que

$$T : U \to V$$

é uma **transformação linear** se e só se verificar as duas condições:

(i) $T(u + v) = T(u) + T(v)$, para todos os $u, v \in U$.

(ii) $T(\lambda u) = \lambda T(u)$, para todos os $u \in U$ e escalares λ.

Observação 33. Sejam U e V espaços lineares. Sejam $\mathbf{0}$ o vector nulo de U e $\mathbf{0}'$ o vector nulo de V.

(i) Se $T : U \to V$ fôr uma transformação linear então $T(U)$ é um subespaço de V e além disso tem-se $T(\mathbf{0}) = \mathbf{0}'$ $(T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) \Leftrightarrow T(\mathbf{0}) = \mathbf{0}')$. Logo, se T não verificar $T(\mathbf{0}) = \mathbf{0}'$ então T não será uma transformação linear.

(ii) $T : U \to V$ é uma transformação linear se e só se

$$T(\lambda u + \mu v) = \lambda T(u) + \mu T(v),$$

para todos os $u, v \in U$ e escalares λ, μ.

(iii) Seja $T : U \to V$ uma transformação linear, com $U = L(\{v_1, v_2, \ldots, v_n\})$. Seja $u \in U$. Logo, existem escalares $\lambda_1, \lambda_2, ..., \lambda_n$ tais que

$$u = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n.$$

Tem-se então

$$T(u) = \lambda_1 T(v_1) + \lambda_2 T(v_2) + ... + \lambda_n T(v_n).$$

Exemplo 32. Consideremos a base canónica $\{(1, 0), (0, 1)\}$ de $\mathbb{R}^2$. Seja

$$T : \mathbb{R}^2 \to \mathbb{R}$$

uma transformação linear tal que $T(1, 0) = 1$ e $T(0, 1) = 1$.

Para qualquer $(x, y) \in \mathbb{R}^2$ tem-se

$$(x, y) = x(1, 0) + y(0, 1).$$

Então,

$$T(x, y) = T(x(1, 0) + y(0, 1)) = xT(1, 0) + yT(0, 1) = x + y.$$

Logo, $T : \mathbb{R}^2 \to \mathbb{R}$ é a transformação linear definida explicitamente por

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

Teorema 37. Sejam U e V espaços lineares e seja $\{v_1, v_2, \ldots, v_n\}$ uma base de U. Sejam $T_1, T_2 : U \to V$ duas transformações lineares.

$$\text{Se } \; T_1(v_i) = T_2(v_i) \quad \text{para todo o } i = 1, \ldots, n, \text{ então } \quad T_1(u) = T_2(u),$$

para todo o $u \in U$, isto é, $T_1 = T_2$.

Exemplo 33. **(i)** Sejam U e V espaços lineares e seja $\mathbf{0}$ o vector nulo de V. Seja $O : U \to V$ definida por

$$O(u) = \mathbf{0},$$

para todo o $u \in U$. O é uma transformação linear e chama-se **transformação nula**.

(ii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Seja

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

definida por

$$T(u) = Au,$$

para todo o $u \in \mathbb{R}^n$. T é uma transformação linear.

(iii) Sejam V um espaço linear e k um escalar (fixo). Seja $T_k : V \to V$ definida por

$$T_k(v) = kv,$$

para todo o $v \in V$.

T_k é uma transformação linear. Diz-se que T_k é uma **homotetia**.

Se $0 < k < 1$ diz-se que T_k é uma **contracção**.

Se $k > 1$ diz-se que T_k é uma **dilatação**.

Se $k = 1$ então chama-se a T_1 a **transformação identidade** e denota-se por I. Tem-se

$$I(u) = u,$$

para todo o $u \in U$.

(iv) $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (1 - y, 2x)$ **não** é uma transformação linear.

(v) $T : \mathbb{R}^2 \to \mathbb{R}$ definida por $T(x, y) = xy$ **não** é uma transformação linear.

(vi) Seja $T : \mathcal{P}_2 \to \mathcal{P}_3$ definida por

$$T\left(p\left(t\right)\right) = tp\left(t\right).$$

T é uma transformação linear.

(vii) Seja $T : \mathcal{P}_3 \to \mathcal{P}_1$ definida por

$$T\left(p\right) = p''.$$

T é uma transformação linear.

(viii) Seja $T : C^1\left(\mathbb{R}\right) \to C\left(\mathbb{R}\right)$ definida por

$$T\left(f\right) = f',$$

onde $C^1(\mathbb{R})$ é o espaço linear de todas as funções reais com primeira derivada contínua em $\mathbb{R}$ e $C(\mathbb{R})$ é o espaço linear de todas as funções reais contínuas em $\mathbb{R}$. T é uma transformação linear.

(ix) Seja $a \in \mathbb{R}$ (fixo). Seja $T : C^1(\mathbb{R}) \to \mathbb{R}$ definida por

$$T(f) = f'(a).$$

T é uma transformação linear.

(x) Seja $n \in \mathbb{N}$. Seja $T : C^n(\mathbb{R}) \to C(\mathbb{R})$ definida por

$$T(f) = f^{(n)},$$

onde $f^{(n)}$ é a derivada de ordem n de f, $C^n(\mathbb{R})$ é o espaço linear de todas as funções reais com derivada de ordem n contínua em $\mathbb{R}$ e $C(\mathbb{R})$ é o espaço linear de todas as funções reais contínuas em $\mathbb{R}$. T é uma transformação linear.

(xi) Seja $T : C(\mathbb{R}) \to C^1(\mathbb{R})$ definida por

$$T(f) = \int_0^x f(t)\, dt.$$

T é uma transformação linear.

(xii) Seja $T : C([a,b]) \to \mathbb{R}$ definida por

$$T(f) = \int_a^b f(t)\, dt.$$

T é uma transformação linear.

(xiii) Seja $T : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathcal{M}_{n\times n}(\mathbb{R})$ definida por

$$T(X) = X^T.$$

T é uma transformação linear.

(xiv) Seja $T : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathcal{M}_{n\times n}(\mathbb{R})$ definida por

$$T(X) = AX,$$

com $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ fixa. T é uma transformação linear.

(xv) Seja

$$\operatorname{tr} : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$$

definida por

$$\operatorname{tr}(A) = a_{11} + a_{22} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii},$$

para todo o $A = (a_{ij})_{n\times n} \in \mathcal{M}_{n\times n}(\mathbb{R})$. tr (traço) é uma transformação linear.

Observação 34. Para matrizes quadradas $A = (a_{ij})_{n\times n}$ define-se o **traço** de A, $\operatorname{tr}(A)$, como sendo a soma de todas as entradas da diagonal principal de A, isto é, $\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}$. Sejam $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ duas matrizes do tipo $n \times n$ e α um escalar. Tem-se:

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

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

(iii)
$$\operatorname{tr}(A^T) = \operatorname{tr}(A),$$

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

Definição 34. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Seja $\mathbf{0}$ o vector nulo de V.

(i) Chama-se **contradomínio** ou imagem de T ao conjunto

$$T(U) = \{T(u) : u \in U\},$$

que também se denota por $\mathcal{I}(T)$.

Note-se que se existir $\{u_1, \ldots, u_k\} \subset U$ tal que $U = L(\{u_1, \ldots, u_k\})$ então

$$\mathcal{I}(T) = L(\{T(u_1), \ldots, T(u_k)\}).$$

(ii) Chama-se **núcleo** ou espaço nulo de T ao conjunto

$$\mathcal{N}(T) = \{u \in U : T(u) = \mathbf{0}\}.$$

Teorema 38. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Então, os conjuntos $\mathcal{N}(T)$ e $\mathcal{I}(T)$ são subespaços de U e V respectivamente.

Exemplo 34. **(i)** Sejam U e V espaços lineares. Sejam $\mathbf{0}$ e $\mathbf{0}'$ os vectores nulos de U e V respectivamente.

Considere a transformação nula $O : U \to V$ definida por

$$O(u) = \mathbf{0}',$$

para todo o $u \in U$. Tem-se

$$\mathcal{N}(O) = U \quad \text{e} \quad \mathcal{I}(O) = \{\mathbf{0}'\}.$$

(ii) Considere a transformação identidade $I : U \to U$ definida por

$$I(u) = u,$$

para todo o $u \in U$. Tem-se

$$\mathcal{N}(I) = \{\mathbf{0}\} \quad \text{e} \quad \mathcal{I}(I) = U.$$

(iii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Seja

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

definida por

$$T(u) = Au,$$

para todo o $u \in \mathbb{R}^n$. Tem-se

$$\mathcal{N}(T) = \mathcal{N}(A) \quad \text{e} \quad \mathcal{I}(T) = \mathcal{C}(A).$$

(iv) Seja $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ definida por $T(f) = f'$. Tem-se

$$\mathcal{N}(T) = \{f : \mathbb{R} \to \mathbb{R} \text{ tal que } f \text{ é constante em } \mathbb{R}\} \qquad \text{e} \qquad \mathcal{I}(T) = C(\mathbb{R}).$$

(v) Seja $T : C^2(\mathbb{R}) \to C(\mathbb{R})$ definida por $T(f(t)) = f''(t) + \omega^2 f(t)$, com $\omega \in \mathbb{R}\backslash\{0\}$. Tem-se (pág. 72 de [2])

$$\mathcal{N}(T) = L(\{\cos(\omega t), \operatorname{sen}(\omega t)\}),$$

onde $\{\cos(\omega t), \operatorname{sen}(\omega t)\}$ é uma base de $\mathcal{N}(T)$. Observe-se que $\mathcal{N}(T)$ é precisamente a solução geral da equação diferencial linear homogénea $f''(t) + \omega^2 f(t) = 0$.

(vi) Seja $T : C^2(\mathbb{R}) \to C(\mathbb{R})$ definida por $T(f(t)) = f''(t) - \omega^2 f(t)$, com $\omega \in \mathbb{R}\backslash\{0\}$. Tem-se (pág. 74 de [2])

$$\mathcal{N}(T) = L\left(\left\{e^{-\omega t}, e^{\omega t}\right\}\right),$$

onde $\{e^{-\omega t}, e^{\omega t}\}$ é uma base de $\mathcal{N}(T)$. Note-se que $\mathcal{N}(T)$ é precisamente a solução geral da equação diferencial linear homogénea $f''(t) - \omega^2 f(t) = 0$.

Definição 35. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear.

(i) Chama-se **característica** de T à dimensão de $\mathcal{I}(T)$, isto é, $\operatorname{car} T = \dim \mathcal{I}(T)$.

(ii) Chama-se **nulidade** de T à dimensão de $\mathcal{N}(T)$, isto é, $\operatorname{nul} T = \dim \mathcal{N}(T)$.

Teorema 39. Sejam U um espaço linear de dimensão finita e T uma transformação linear definida em U. Então, o subespaço $\mathcal{I}(T)$ tem dimensão finita e

$$\dim \mathcal{N}(T) + \dim \mathcal{I}(T) = \dim U.$$

Definição 36. Sejam U e V espaços lineares e $S, T : U \to V$ transformações lineares. Seja λ um escalar. Sejam $S + T, \lambda T : U \to V$ definidas por

$$(S + T)(u) = S(u) + T(u) \qquad \text{e} \qquad (\lambda T)(u) = \lambda T(u),$$

para todo o $u \in U$.

$S + T$ e λT são transformações lineares.

Definição 37. Sejam U e V espaços lineares. Chama-se a $\mathfrak{L}(U,V)$ o conjunto de **todas** as transformações lineares de U em V.

Teorema 40. Sejam U e V espaços lineares. O conjunto $\mathfrak{L}(U,V)$, com as operações da definição 36, é um espaço linear.

Exemplo 35. Seja $\mathcal{B} = \{T_1, T_2, T_3, T_4\}$ com $T_1, T_2, T_3, T_4 \in \mathfrak{L}(\mathbb{R}^2, \mathbb{R}^2)$ definidas por

$$T_1(x,y) = (x,0), \ \ T_2(x,y) = (y,0), \ \ T_3(x,y) = (0,x) \ \text{ e } \ T_4(x,y) = (0,y),$$

para todo o $(x,y) \in \mathbb{R}^2$. O conjunto $\mathcal{B}$ é uma base de $\mathfrak{L}(\mathbb{R}^2, \mathbb{R}^2)$. Logo,

$$\dim \mathfrak{L}(\mathbb{R}^2, \mathbb{R}^2) = 4.$$

Definição 38. Sejam U, V e W espaços lineares e, $T : U \to V$ e $S : V \to W$ transformações lineares. Seja $S \circ T$ (ou ST): $U \to W$ definida por

$$(S \circ T)(u) = S(T(u)),$$

para todo o $u \in U$. $S \circ T$ é uma transformação linear. Chama-se a $S \circ T$ (ou ST) a **composição de** S com T.

Observação 35. Em geral, tem-se $S \circ T \neq T \circ S$.

Teorema 41. **(i)** Sejam $T : U \to V$, $S : V \to W$ e $R : W \to X$. Então, tem-se $R \circ (S \circ T) = (R \circ S) \circ T$.

(ii) Sejam $R, S : U \to V$ e $T : V \to W$. Seja $\lambda \in \mathbb{R}$. Então, tem-se

$$T \circ (R + S) = T \circ R + T \circ S \ \text{ e } \ T \circ (\lambda R) = \lambda (T \circ R).$$

Se o contradomínio de Q estiver contido em U então

$$(R + S) \circ Q = R \circ Q + S \circ Q \ \text{ e } \ (\lambda R) \circ Q = \lambda (R \circ Q).$$

Definição 39. Define-se $T^0 = I$ e $T^k = T \circ T^{k-1}$, para todo o $k = 1, 2, ...$

Observação 36. Tem-se $T^{m+n} = T^m \circ T^n$ para todos os $m, n \in \mathbb{N}$.

Definição 40. **(i)** $T : U \to V$ diz-se **injectiva** se e só se

$$T(u) = T(w) \ \ \Rightarrow \ \ u = w,$$

para todos os $u, w \in U$, isto é, se e só se

$$u \neq w \ \ \Rightarrow \ \ T(u) \neq T(w),$$

para todos os $u, w \in U$.

(ii) $T : U \to V$ diz-se **sobrejectiva** se e só se $T(U) = V$.

(iii) $T : U \to V$ diz-se **bijectiva** se e só se fôr injectiva e sobrejectiva.

Definição 41. Sejam U e V espaços lineares. Diz-se que U e V são isomorfos se e só se existir um **isomorfismo** entre U e V, isto é, se e só se existir uma transformação linear bijectiva $T : U \to V$. Sendo U e V isomorfos escreve-se

$$U \cong V.$$

Teorema 42. Sejam U e V dois espaços lineares de dimensões finitas. Então, U e V são isomorfos se e só se $\dim U = \dim V$.

Observação 37. **(i)** Qualquer espaço linear real de dimensão n é isomorfo a $\mathbb{R}^n$.

(ii) Sejam U e V dois espaços lineares de dimensões finitas. A transformação linear $T : U \to V$ é sobrejectiva se e só se T transformar um qualquer conjunto gerador de U num conjunto gerador de V.

(iii) Sejam U e V dois espaços lineares de dimensões finitas. Se a transformação linear $T : U \to V$ fôr sobrejectiva então $\dim V \leq \dim U$.

(iv) Sejam U e V dois espaços lineares de dimensões finitas. Se a transformação linear $T : U \to V$ fôr injectiva então $\dim U \leq \dim V$.

Exemplo 36. **(i)** A transformação linear $T : \mathbb{R}^n \to \mathcal{M}_{n\times 1}(\mathbb{R})$ definida por

$$T(a_1, \ldots, a_n) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},$$

é um isomorfismo. Logo $\mathbb{R}^n \cong \mathcal{M}_{n\times 1}(\mathbb{R})$.

(ii) A transformação linear $T : \mathcal{M}_{m\times n}(\mathbb{R}) \to \mathbb{R}^{mn}$ definida por

$$T\left(\begin{bmatrix} a_{11} & & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & & a_{mn} \end{bmatrix}\right) = (a_{11}, \ldots, a_{m1}, \ldots, a_{1n}, \ldots, a_{mn}),$$

é um isomorfismo. Logo $\mathcal{M}_{m\times n}(\mathbb{R}) \cong \mathbb{R}^{mn}$.

(iii) A transformação linear $T : \mathbb{R}^{n+1} \to \mathcal{P}_n$ definida por

$$T(a_0, a_1, \ldots, a_n) = a_0 + a_1 t + \cdots + a_n t^n,$$

é um isomorfismo. Logo $\mathbb{R}^{n+1} \cong \mathcal{P}_n$.

(iv) Seja A uma matriz $m \times n$. Os espaços $\mathcal{C}(A)$ e $\mathcal{L}(A)$ são isomorfos pois têm a mesma dimensão (car A).

$$\mathcal{C}(A) \cong \mathcal{L}(A).$$

Teorema 43. Sejam U e V espaços lineares de dimensões finitas tais que

$$\dim U = \dim V.$$

Seja $T : U \to V$ uma transformação linear. Então, T é injectiva se e só se T é sobrejectiva.

Definição 42. Diz-se que $T : U \to V$ é invertível se existir $S : T(U) \to U$ tal que

$$S \circ T = I_U \ \text{ e } \ T \circ S = I_{T(U)},$$

onde I_U e $I_{T(U)}$ são as funções identidade em U e $T(U)$ respectivamente. Chama-se a S a inversa de T e escreve-se

$$S = T^{-1}.$$

Teorema 44. Sejam U e V espaços lineares de dimensões finitas. Seja

$$T : U \to V$$

uma transformação linear. Seja $\mathbf{0}$ o vector nulo de U. As seguintes afirmações são equivalentes.

(i) T é injectiva.

(ii) T é invertível e a inversa $T^{-1} : T(U) \to U$ é linear.

(iii) $\mathcal{N}(T) = \{\mathbf{0}\}$.

(iv) $\dim U = \dim T(U)$.

(v) T transforma vectores linearmente independentes de U em vectores linearmente independentes de V.

(vi) T transforma bases de U em bases de $T(U)$.

Teorema 45. Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Seja $b \in V$. Então:

(i) Existência de solução: a equação linear $T(u) = b$ tem sempre solução (para qualquer b) se e só se T fôr sobrejectiva ($T(U) = V$);

(ii) Unicidade de solução: a equação linear $T(u) = b$ a ter solução, ela é única se e só se T fôr injectiva;

(iii) Existência e unicidade de solução: a equação linear $T(u) = b$ tem solução única u se e só se T fôr bijectiva.

Teorema 46. Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Seja $b \in V$. A solução geral da equação linear $T(u) = b$ obtém-se somando a uma solução particular dessa equação, a solução geral da equação linear homogénea $T(u) = \mathbf{0}$.

Teorema 47. (Representação matricial de uma transformação linear). Sejam U e V espaços lineares de dimensões finitas tais que $\dim U = n$ e $\dim V = m$. Sejam $\mathcal{B}_1 = \{u_1, u_2, \ldots, u_n\}$ e $\mathcal{B}_2 = \{v_1, v_2, \ldots, v_m\}$ duas bases ordenadas de U e V respectivamente. Seja $T : U \to V$ uma transformação linear. Considere-se a matriz $A = (a_{ij})_{m \times n}$ cuja coluna j, para cada $j = 1, ..., n$, é formada pelas coordenadas de $T(u_j)$ na base $\mathcal{B}_2$. Isto é,

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i.$$

Chama-se a esta matriz A a **representação matricial** de T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$ e escreve-se

$$A = M(T; \mathcal{B}_1; \mathcal{B}_2).$$

Além disso, sendo $\alpha_1, \alpha_2, ..., \alpha_n$ as coordenadas de um vector $u \in U$ na base ordenada $\mathcal{B}_1$ então as coordenadas $\beta_1, \beta_2, ..., \beta_m$ de $T(u) \in V$ na base ordenada $\mathcal{B}_2$ são dadas por

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} = M(T; \mathcal{B}_1; \mathcal{B}_2) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Observação 38. (i) Nas condições do teorema anterior, tem-se

$$u \in \mathcal{N}(T) \Leftrightarrow (\alpha_1, \alpha_2, ..., \alpha_n) \in \mathcal{N}(A)$$

uma vez que

$$T(u) = T\left(\sum_{j=1}^{n} \alpha_j u_j\right) \underset{T \text{ é linear}}{=} \sum_{j=1}^{n} \alpha_j T(u_j) = \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{m} a_{ij} v_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \alpha_j\right) v_i$$

e sendo $\{v_1, v_2, \ldots, v_m\}$ uma base de V tem-se

$$u \in \mathcal{N}(T) \Leftrightarrow T(u) = \mathbf{0} \Leftrightarrow \left(\sum_{j=1}^{n} a_{ij} \alpha_j = 0, \text{ para } i = 1, ..., m\right) \Leftrightarrow (\alpha_1, \alpha_2, ..., \alpha_n) \in \mathcal{N}(A).$$

Além disso:

$$\mathcal{I}(T) = L\left(\{a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m, \ldots, a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m\}\right) =$$
$$= L\left(\{T(u_1), \ldots, T(u_n)\}\right).$$

(ii) Seja V um espaço linear de dimensão finita, com $\dim V = n$. Sejam $\mathcal{B}_1 = \{u_1, u_2, \ldots, u_n\}$ e $\mathcal{B}_2 = \{v_1, v_2, \ldots, v_n\}$ duas bases ordenadas de V. A representação matricial da transformação identidade $I : V \to V$ em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$ é igual à matriz de mudança da base $\mathcal{B}_1$ para $\mathcal{B}_2$. Isto é,

$$M(I; \mathcal{B}_1; \mathcal{B}_2) = S_{\mathcal{B}_1 \to \mathcal{B}_2}.$$

Teorema 48. Sejam U e V espaços lineares tais que $\dim U = n$ e $\dim V = m$. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ bases (ordenadas) de U e V respectivamente. Seja

$$A = M(T; \mathcal{B}_1; \mathcal{B}_2) \in \mathcal{M}_{m\times n}(\mathbb{R})$$

a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$. Tem-se então:

(i) $\dim \mathcal{N}(T) = \operatorname{nul} A$;

(ii) $\dim \mathcal{I}(T) = \operatorname{car} A$;

(iii) T é injectiva se e só se $\operatorname{nul} A = 0$, isto é, se e só se $\operatorname{car} A = n$;

(iv) T é sobrejectiva se e só se $\operatorname{car} A = m$.

Teorema 49. Sejam $\mathcal{B}_c^n = \{e_1, e_2, \ldots, e_n\}$ e $\mathcal{B}_c^m = \{e_1', e_2', \ldots, e_m'\}$ as bases canónicas (ordenadas) de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Considere-se a matriz $A = (a_{ij})_{m\times n} = M(T; \mathcal{B}_c^n; \mathcal{B}_c^m) \in \mathcal{M}_{m\times n}(\mathbb{R})$ cuja coluna j, para cada $j = 1, ..., n$, é formada pelas coordenadas de $T(e_j)$ na base $\mathcal{B}_c^m$. Isto é,

$$T(e_j) = \sum_{i=1}^{m} a_{ij} e_i' = a_{1j} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + ... + a_{mj} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Então, tem-se, para todo o $u \in \mathbb{R}^n$,

$$T(u) = Au.$$

Dem. Seja $u \in \mathbb{R}^n$. Então, existem $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$ tais que

$$u = \lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n = \sum_{j=1}^{n} \lambda_j e_j.$$

Uma vez que, para todo o $j = 1, ..., n$, $T(e_j) = \sum_{i=1}^{m} a_{ij} e_i'$, tem-se

$$T(u) = T\left(\sum_{j=1}^{n} \lambda_j e_j\right) \underset{\text{T é linear}}{=} \sum_{j=1}^{n} \lambda_j T(e_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} e_i' = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\lambda_j\right) e_i' =$$

$$= \left(\sum_{j=1}^{n} a_{1j}\lambda_j, ..., \sum_{j=1}^{n} a_{mj}\lambda_j\right) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = Au.$$

Observação 39. No caso em que $U = \mathbb{R}^n$, $V = \mathbb{R}^m$ e $\mathcal{B}_1 = \mathcal{B}_c^n$, $\mathcal{B}_2 = \mathcal{B}_c^m$, como caso particular da alínea (i) da observação 38, tem-se:

$$\mathcal{N}(T) = \mathcal{N}(A) \quad \text{ e } \quad \mathcal{I}(T) = \mathcal{C}(A),$$

uma vez que neste caso as coordenadas de um vector numa base coincidem com o próprio vector.

Exemplo 37. **(i)** Seja $T : \mathbb{R}^4 \to \mathbb{R}^3$ definida por

$$T(x, y, z, w) = (3x + y - 2z, 0, x + 4z).$$

T é uma transformação linear e a matriz $M(T; \mathcal{B}_c^4; \mathcal{B}_c^3)$ que representa T em relação às bases canónicas (ordenadas) $\mathcal{B}_c^4$ e $\mathcal{B}_c^3$ de $\mathbb{R}^4$ e $\mathbb{R}^3$ respectivamente, é dada por

$$A = M(T; \mathcal{B}_c^4; \mathcal{B}_c^3) = \begin{bmatrix} 3 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 4 & 0 \end{bmatrix},$$

uma vez que $T(1, 0, 0, 0) = (3, 0, 1)$, $T(0, 1, 0, 0) = (1, 0, 0)$, $T(0, 0, 1, 0) = (-2, 0, 4)$ e $T(0, 0, 0, 1) = (0, 0, 0)$.

Tem-se então:

$$T(x, y, z, w) = M(T; \mathcal{B}_c^4; \mathcal{B}_c^3) \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = (3x + y - 2z, 0, x + 4z).$$

Além disso, tem-se

$$\begin{aligned} \mathcal{N}(T) &= \mathcal{N}(A) = \left\{(x, y, z, w) \in \mathbb{R}^4 : y = 14z \ \text{ e } \ x = -4z\right\} = \\ &= \{(-4z, 14z, z, w) : z, w \in \mathbb{R}\} = L(\{(-4, 14, 1, 0), (0, 0, 0, 1)\}) \end{aligned}$$

e

$$\mathcal{I}(T) = \mathcal{C}(A) = L(\{(3, 0, 1), (1, 0, 0)\}).$$

Uma base de $\mathcal{I}(T) : \{(3, 0, 1), (1, 0, 0)\}$. Uma base de $\mathcal{N}(T) : \{(-4, 14, 1, 0), (0, 0, 0, 1)\}$.

(ii) Sejam $\mathcal{B}_1 = \{1, t, t^2\}$ e $\mathcal{B}_2 = \{1, t, t^2, t^3\}$ as bases canónicas (ordenadas) de $\mathcal{P}_2$ e $\mathcal{P}_3$ respectivamente. Seja $D : \mathcal{P}_2 \to \mathcal{P}_3$ tal que $D(1) = 0$, $D(t) = 1$ e $D(t^2) = 2t$. D é uma transformação linear e a matriz $M(D; \mathcal{B}_1; \mathcal{B}_2)$ que representa D em relação às bases canónicas $\mathcal{B}_1$ e $\mathcal{B}_2$, é dada por

$$M(D; \mathcal{B}_1; \mathcal{B}_2) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Além disso tem-se

$$M(D; \mathcal{B}_1; \mathcal{B}_2) \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \\ 0 \end{bmatrix},$$

isto é, $D(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$, com $a_0, a_1, a_2 \in \mathbb{R}$.

Além disso, como

$$\mathcal{N}(D) = \left\{a_0 + a_1 t + a_2 t^2 : D\left(a_0 + a_1 t + a_2 t^2\right) = \mathbf{0}\right\} = \left\{a_0 + a_1 t + a_2 t^2 : a_1 = a_2 = 0 \text{ e } a_0 \in \mathbb{R}\right\},$$

tem-se

$$\mathcal{N}(D) = \{a_0 : a_0 \in \mathbb{R}\} = L(\{1\}) \qquad \text{e} \qquad \mathcal{I}(D) = L(\{1, 2t\}).$$

Uma base de $\mathcal{I}(D)$: $\{1, 2t\}$. Uma base de $\mathcal{N}(D)$: $\{1\}$.

(iii) Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ a transformação linear cuja matriz que a representa em relação às bases $\mathcal{B}_1 = \{(1,1,1), (0,1,1), (0,0,1)\}$ e $\mathcal{B}_2 = \{(1,1), (1,-1)\}$ de $\mathbb{R}^3$ e $\mathbb{R}^2$ respectivamente, é dada por

$$A = M(T; \mathcal{B}_1; \mathcal{B}_2) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}.$$

Seja $u \in \mathbb{R}^3$ e sejam $(\alpha_1, \alpha_2, \alpha_3)$ as coordenadas de u em relação à base $\mathcal{B}_1$. Tem-se

$$u \in \mathcal{N}(T) \Leftrightarrow (\alpha_1, \alpha_2, \alpha_3) \in \mathcal{N}(A)$$

e como

$$\mathcal{N}(A) = \mathcal{N}\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}\right) = \{(-2y - 3z, y, z) : y, z \in \mathbb{R}\} = L(\{(-2,1,0), (-3,0,1)\}),$$

$$\begin{aligned} \mathcal{N}(T) &= \{(-2y - 3z)(1,1,1) + y(0,1,1) + z(0,0,1) : y, z \in \mathbb{R}\} = \\ &= \{(-2y - 3z, -y - 3z, -y - 2z) : y, z \in \mathbb{R}\} = L(\{(-2,-1,-1), (-3,-3,-2)\}) \end{aligned}$$

Quanto ao contradomínio:

$$\begin{aligned} \mathcal{I}(T) &= L(\{1(1,1) + 2(1,-1), 2(1,1) + 4(1,-1), 3(1,1) + 6(1,-1)\}) = \\ &= L(\{(3,-1), (6,-2), (9,-3)\}) = L(\{(3,-1)\}). \end{aligned}$$

Uma base de $\mathcal{I}(T)$: $\{(3,-1)\}$. Uma base de $\mathcal{N}(T)$: $\{(-2,-1,-1), (-3,-3,-2)\}$.

Note-se que:

$$\dim \mathcal{N}(T) = \operatorname{nul} A \qquad \text{e} \qquad \dim \mathcal{I}(T) = \operatorname{car} A.$$

Teorema 50. Sejam U, V e W espaços lineares de dimensões finitas. Sejam $\mathcal{B}_1, \mathcal{B}_2$ e $\mathcal{B}_3$ bases de U, V e W respectivamente. Sejam $T \in \mathfrak{L}(U,V)$ e $S \in \mathfrak{L}(V,W)$. Então, tem-se

$$M(S \circ T; \mathcal{B}_1; \mathcal{B}_3) = M(S; \mathcal{B}_2; \mathcal{B}_3) M(T; \mathcal{B}_1; \mathcal{B}_2).$$

Teorema 51. Sejam U e V dois espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases ordenadas de U e V respectivamente. Seja $A = M(T; \mathcal{B}_1; \mathcal{B}_2)$ a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$.

Se $V = T(U)$ então T é invertível se e só se A fôr uma matriz quadrada não singular. Tem-se então

$$A^{-1} = M(T^{-1}; \mathcal{B}_2; \mathcal{B}_1),$$

isto é, A^{-1} será a matriz que representa T^{-1} em relação às bases $\mathcal{B}_2$ e $\mathcal{B}_1$.

Teorema 52. Sejam U e V espaços lineares de dimensões finitas respectivamente n e m. Isto é,

$$\dim U = n \quad \text{e} \quad \dim V = m.$$

Então, os espaços lineares $\mathfrak{L}(U,V)$ e $\mathcal{M}_{m\times n}(\mathbb{R})$ são isomorfos e escreve-se

$$\mathfrak{L}(U,V) \cong \mathcal{M}_{m\times n}(\mathbb{R}).$$

Dem. Fixando bases $\mathcal{B}_1$ e $\mathcal{B}_2$ para U e V respectivamente,

$$\begin{array}{ccc} \mathfrak{L}(U,V) & \rightarrow & \mathcal{M}_{m\times n}(\mathbb{R}) \\ T & \rightarrow & M(T;\mathcal{B}_1;\mathcal{B}_2) \end{array}$$

é uma transformação linear bijectiva.

Observação 40. No teorema anterior tem-se $\dim \mathfrak{L}(U,V) = mn$.

Teorema 53. Seja V um espaço linear de dimensão finita. Seja $T : V \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases ordenadas de V. Seja $M(T;\mathcal{B}_1;\mathcal{B}_1)$ a matriz que representa T em relação à base $\mathcal{B}_1$.

Então, a matriz $M(T;\mathcal{B}_2;\mathcal{B}_2)$ que representa T em relação à base $\mathcal{B}_2$, é dada por

$$M(T;\mathcal{B}_2;\mathcal{B}_2) = S_{\mathcal{B}_1\to\mathcal{B}_2} M(T;\mathcal{B}_1;\mathcal{B}_1)\left(S_{\mathcal{B}_1\to\mathcal{B}_2}\right)^{-1},$$

onde $S_{\mathcal{B}_1\to\mathcal{B}_2}$ é a matriz de mudança da base $\mathcal{B}_1$ para $\mathcal{B}_2$.

Isto é, o diagrama seguinte é comutativo.

$$\begin{array}{ccc} (V,\mathcal{B}_1) & \overset{M(T;\mathcal{B}_1;\mathcal{B}_1)}{\underset{T}{\longrightarrow}} & (V,\mathcal{B}_1) \\ S_{\mathcal{B}_1\to\mathcal{B}_2} \downarrow I & & I \downarrow S_{\mathcal{B}_1\to\mathcal{B}_2} \\ (V,\mathcal{B}_2) & \overset{T}{\underset{M(T;\mathcal{B}_2;\mathcal{B}_2)}{\longrightarrow}} & (V,\mathcal{B}_2) \end{array}$$

Teorema 54. Caso geral. Sejam U e V dois espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_1'$ duas bases ordenadas de U. Sejam $\mathcal{B}_2$ e $\mathcal{B}_2'$ duas bases ordenadas de V. Seja $M(T;\mathcal{B}_1;\mathcal{B}_2)$ a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$.

Então, a matriz $M(T;\mathcal{B}_1';\mathcal{B}_2')$ que representa T em relação às bases $\mathcal{B}_1'$ e $\mathcal{B}_2'$, é dada por

$$M(T;\mathcal{B}_1';\mathcal{B}_2') = S_{\mathcal{B}_2\to\mathcal{B}_2'} M(T;\mathcal{B}_1;\mathcal{B}_2)\left(S_{\mathcal{B}_1\to\mathcal{B}_1'}\right)^{-1},$$

onde $S_{\mathcal{B}_2\to\mathcal{B}_2'}$ e $S_{\mathcal{B}_1\to\mathcal{B}_1'}$ são as matrizes de mudança das bases $\mathcal{B}_2$ para $\mathcal{B}_2'$ e de $\mathcal{B}_1$ para $\mathcal{B}_1'$ respectivamente.

Isto é, o diagrama seguinte é comutativo.

$$\begin{array}{ccc} (U,\mathcal{B}_1) & \overset{M(T;\mathcal{B}_1;\mathcal{B}_2)}{\underset{T}{\longrightarrow}} & (V,\mathcal{B}_2) \\ S_{\mathcal{B}_1\to\mathcal{B}_1'} \downarrow I & & I \downarrow S_{\mathcal{B}_2\to\mathcal{B}_2'} \\ (U,\mathcal{B}_1') & \overset{T}{\underset{M(T;\mathcal{B}_1';\mathcal{B}_2')}{\longrightarrow}} & (V,\mathcal{B}_2') \end{array}$$

Exemplo 38. Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x,y) = (y, x, y-x)$. T é uma transformação linear. A matriz $M(T;\mathcal{B}_c^2;\mathcal{B}_c^3)$ que representa T em relação à base canónica (ordenada) $\mathcal{B}_c^2$ de $\mathbb{R}^2$ e à base canónica (ordenada) $\mathcal{B}_c^3$ de $\mathbb{R}^3$, é dada por

$$M(T;\mathcal{B}_c^2;\mathcal{B}_c^3) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

Sejam $\mathcal{B}_1 = \{(1,1), (-1,1)\}$ uma base ordenada de $\mathbb{R}^2$ e $\mathcal{B}_2 = \{(0,0,1), (0,1,1), (1,1,1)\}$ uma base ordenada de $\mathbb{R}^3$

A matriz $M(T;\mathcal{B}_1;\mathcal{B}_2)$ que representa T em relação à base ordenada $\mathcal{B}_1$ de $\mathbb{R}^2$ e à base ordenada $\mathcal{B}_2$ de $\mathbb{R}^3$, é dada por

$$M(T;\mathcal{B}_1;\mathcal{B}_2) = \begin{bmatrix} -1 & 3 \\ 0 & -2 \\ 1 & 1 \end{bmatrix},$$

uma vez que

$$\begin{aligned} T(1,1) &= (1,1,0) = -(0,0,1) + 0(0,1,1) + 1\,(1,1,1) \\ T(-1,1) &= (1,-1,2) = 3(0,0,1) - 2(0,1,1) + 1\,(1,1,1)\,. \end{aligned}$$

Vamos agora verificar que se tem

$$M(T;\mathcal{B}_1;\mathcal{B}_2) = S_{\mathcal{B}_c^3 \to \mathcal{B}_2} M(T;\mathcal{B}_c^2;\mathcal{B}_c^3) \left(S_{\mathcal{B}_c^2 \to \mathcal{B}_1}\right)^{-1}.$$

Uma vez que

$$\begin{aligned} (1,0,0) &= 0(0,0,1) - 1(0,1,1) + 1\,(1,1,1)\,, \quad (0,1,0) = -(0,0,1) + 1(0,1,1) + 0\,(1,1,1)\,, \\ (0,0,1) &= 1(0,0,1) + 0(0,1,1) + 0\,(1,1,1) \end{aligned}$$

tem-se então

$$S_{\mathcal{B}_c^3 \to \mathcal{B}_2} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Logo,

$$S_{\mathcal{B}_c^3 \to \mathcal{B}_2} M(T;\mathcal{B}_c^2;\mathcal{B}_c^3) \left(S_{\mathcal{B}_c^2 \to \mathcal{B}_1}\right)^{-1} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} S_{\mathcal{B}_1 \to \mathcal{B}_c^2} =$$

$$= \begin{bmatrix} -2 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 0 & -2 \\ 1 & 1 \end{bmatrix} = M(T;\mathcal{B}_1;\mathcal{B}_2).$$

Por exemplo, para $(2,1) \in \mathbb{R}^2$, tem-se:

$$\begin{array}{ccc} \text{coordenadas de } (2,1) & \overset{M(T;\mathcal{B}_c^2;\mathcal{B}_c^3)}{\underset{T}{\longrightarrow}} & \text{coordenadas de } T(2,1) \\ \text{na base } \mathcal{B}_c^2 & & \text{na base } \mathcal{B}_c^3 \\ & & \\ S_{S_{\mathcal{B}_c^2 \to \mathcal{B}_1}} \downarrow I & & I \downarrow S_{\mathcal{B}_c^3 \to \mathcal{B}_2} \\ & & \\ \text{coordenadas de } (2,1) & \overset{T}{\underset{M(T;\mathcal{B}_1;\mathcal{B}_2)}{\longrightarrow}} & \text{coordenadas de } T(2,1) \\ \text{na base } \mathcal{B}_1 & & \text{na base } \mathcal{B}_2. \end{array}$$

Valores próprios e vectores próprios. Diagonalização.

Definição 43. Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Diz-se que um escalar λ é um **valor próprio** de T se existir um vector não nulo $u \in U$ tal que

$$T(u) = \lambda u.$$

Aos vectores não nulos u que satisfaçam a equação anterior chamam-se **vectores próprios** associados ao valor próprio λ. Dado um valor próprio λ de T, o conjunto

$$E_\lambda = \{u \in U : T(u) = \lambda u\} = \mathcal{N}(T - \lambda I)$$

é um subespaço linear de U. Chama-se a E_λ o **subespaço próprio** associado ao valor próprio λ. À dimensão de E_λ chama-se **multiplicidade geométrica** de λ e denota-se por $m_g(\lambda)$, isto é,

$$\dim \mathcal{N}(T - \lambda I) = m_g(\lambda).$$

Exemplo 39. (a) Seja V um espaço linear e $I : V \to V$ a transformação identidade. Então todos os vectores de V, exceptuando o vector nulo, são vectores próprios de T associados ao valor próprio 1.

(b) Seja V o espaço linear das funções reais indefinidamente diferenciáveis em R e $T : V \to V$ a (transfomação) função derivada. Como, por exemplo

$$T\left(e^{2x}\right) = 2e^{2x}$$

então e^{2x} é vector próprio de T associado ao valor próprio 2.

(c) Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida por $T(x, y) = (-x, y)$. Como

$$T(x,y) = \lambda(x,y) \Leftrightarrow (-x,y) = \lambda(x,y) \Leftrightarrow \begin{cases} -x = \lambda x \\ y = \lambda y \end{cases}$$

se $x \neq 0$ e $y \neq 0$ então $\lambda = -1$ e $\lambda = 1$, pelo que não existem vectores próprios neste caso;

se $x = 0$ e $y = 0$ então (x, y) não é vector próprio;

se $x \neq 0$ e $y = 0$ então $\lambda = -1$ é valor próprio e os vectores próprios são $(x, 0)$;

se $x = 0$ e $y \neq 0$ então $\lambda = 1$ é valor próprio e os vectores próprios são $(0, y)$;

(d) Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida por $T(x, y) = (y, -x)$. Como

$$T(x,y) = \lambda(x,y) \Leftrightarrow (y,-x) = \lambda(x,y) \Leftrightarrow \begin{cases} y = \lambda x \\ -x = \lambda y \end{cases} \Leftrightarrow \begin{cases} y = \lambda x \\ -x = \lambda^2 x \end{cases} \Leftrightarrow \begin{cases} y = \lambda x \\ x\left(\lambda^2 + 1\right) = 0 \end{cases}$$

não existem vectores (x, y) com $x \neq 0$ e $y \neq 0$ e $\lambda \in \mathbb{R}$ que satisfaçam o sistema anterior. Logo T não tem vectores próprios.

Observação 41. **(i)** Sejam V um espaço linear e $\mathbf{0}$ o vector nulo de V. Seja $T: V \to V$ uma transformação linear. Um escalar λ é um valor próprio de T se e só se $\mathcal{N}(T-\lambda I) \neq \{\mathbf{0}\}$.

(ii) Se o espaço linear V tiver dimensão finita n e se $A = M(T;\mathcal{B};\mathcal{B})$ fôr a matriz $n \times n$ que representa T em relação a uma base ordenada $\mathcal{B}$ de V, então um escalar λ é um valor próprio de T se e só se esse escalar λ fôr solução da equação

$$\det(A - \lambda I) = 0,$$

uma vez que se tem, para $v \in V$,

$$(T - \lambda I)\, v = \mathbf{0} \Leftrightarrow (A - \lambda I) \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \mathbf{0}.$$

onde $\alpha_1, ..., \alpha_n$ são as coordenadas de v na base ordenada $\mathcal{B}$, daí que

$$\lambda \text{ é um valor próprio de } T \;\Leftrightarrow\; \mathcal{N}(T - \lambda I) \neq \{\mathbf{0}\} \Leftrightarrow \mathcal{N}(A - \lambda I) \neq \{\mathbf{0}\} \Leftrightarrow \det(A - \lambda I) = 0$$

isto é

$$\lambda \text{ é um valor próprio de } T \;\Leftrightarrow\; \det(A - \lambda I) = 0$$

Além disso, tem-se

$$v \text{ é um vector próprio de } T \;\Leftrightarrow\; v \in \mathcal{N}(T - \lambda I) \setminus \{\mathbf{0}\} \Leftrightarrow (\alpha_1, ..., \alpha_n) \in \mathcal{N}(A - \lambda I) \setminus \{\mathbf{0}\}$$

isto é

$$v \text{ é um vector próprio de } T \;\Leftrightarrow\; (\alpha_1, ..., \alpha_n) \in \mathcal{N}(A - \lambda I) \setminus \{\mathbf{0}\}$$

e

$$m_g(\lambda) = \dim \mathcal{N}(T - \lambda I) = \dim \mathcal{N}(A - \lambda I).$$

(iii) No caso em que $V = \mathbb{R}^n$ e $A = M(T;\mathcal{B}_c^n;\mathcal{B}_c^n)$, como (neste caso) $v = (\alpha_1, ..., \alpha_n)$, tem-se

$$\mathcal{N}(T - \lambda I) = \mathcal{N}(A - \lambda I).$$

Definição 44. Seja A uma matriz $n \times n$. Chama-se ao polinómio

$$p(\lambda) = \det(A - \lambda I)$$

o **polinómio característico** da matriz A. Este polinómio tem grau n, o coeficiente do termo de grau n é $(-1)^n$, o coeficiente do termo de grau $n-1$ é $(-1)^{n-1} \operatorname{tr} A$ e o termo constante é $p(0) = \det A$.

Definição 45. Seja A uma matriz $n \times n$. Chama-se **valor próprio** de A a qualquer escalar λ tal que $A - \lambda I$ seja singular, isto é, tal que $\det(A - \lambda I) = 0$. Ao conjunto de todos os valores próprios de A chama-se **espectro** de A. À multiplicidade de λ como raíz do polinómio $\det(A-\lambda I)$ chama-se **multiplicidade algébrica** de λ e denota-se por $m_a(\lambda)$. Chama-se **vector próprio** de A, associado ao valor próprio λ de A, a qualquer vector não nulo v que verifique

$$(A-\lambda I)v = \mathbf{0},$$

isto é, a qualquer vector

$$v \in \mathcal{N}(A-\lambda I) \setminus \{\mathbf{0}\}.$$

Observação 42. (i) Seja A uma matriz $n \times n$. O escalar 0 é valor próprio de A se e só se A fôr singular. Isto é, a matriz A é invertível se e só se 0 não fôr valor próprio de A.

(ii) Seja A uma matriz $n \times n$. Então o polinómio característico de A pode ser escrito na forma:

$$p(\lambda) = \det(A-\lambda I) = (\lambda_1 - \lambda)^{m_1}(\lambda_2 - \lambda)^{m_2} \cdots (\lambda_k - \lambda)^{m_k},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_k$ são os valores próprios distintos de A e $m_1, m_2, \ldots, m_k$ são as **multiplicidades algébricas** desses valores próprios respectivamente, com $m_1+m_2+\cdots+m_k = n$.

(iii) Seja A uma matriz $n \times n$. Tem-se $m_g(\lambda) \leq m_a(\lambda)$, para qualquer valor próprio λ de A.

(iv) Seja A uma matriz $n \times n$, com os valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_n$ (repetidos de acordo com a respectiva multiplicidade algébrica). Então, atendendo à alínea anterior e à Definição 45 tem-se

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n \qquad \text{e} \qquad \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Definição 46. Sejam A e B matrizes $n \times n$. As matrizes A e B dizem-se **semelhantes** se existir uma matriz S invertível tal que

$$B = SAS^{-1}.$$

Teorema 55. Duas matrizes são semelhantes se e só se existirem bases em relação às quais essas matrizes representem a mesma transformação linear.

Teorema 56. Sejam A e B matrizes $n \times n$. Se A e B forem semelhantes **então** A e B têm o(a) mesmo(a)**:**

(i) determinante; **(ii)** característica; **(iii)** nulidade; **(iv)** traço;

(v) polinómio característico, e portanto têm os mesmos valores próprios com as mesmas multiplicidades algébricas e geométricas.

Dem. (Matrizes semelhantes têm o mesmo polinómio característico.)

$$\begin{aligned}
\det(B-\lambda I) &= \det(SAS^{-1} - \lambda I) = \det(SAS^{-1} - \lambda SS^{-1}) = \\
&= \det(S(A-\lambda I)S^{-1}) = \det S \det(A-\lambda I) \det\left(S^{-1}\right) = \\
&= \det S \det(A-\lambda I) \frac{1}{\det S} = \det(A-\lambda I).
\end{aligned}$$

Teorema 57. **(i)** Seja V um espaço linear. Seja $T : V \to V$ uma transformação linear. Se T tiver valores próprios distintos $\lambda_1, ..., \lambda_k$ e se $u_1, ..., u_k$ forem os vectores próprios associados a cada um destes valores próprios, então os vectores $u_1, ..., u_k$ são linearmente independentes.

(ii) Seja A uma matriz $n \times n$. Se A tiver valores próprios distintos $\lambda_1, ..., \lambda_k$ e se $u_1, ..., u_k$ forem os vectores próprios associados a cada um destes valores próprios, então os vectores $u_1, ..., u_k$ são linearmente independentes.

Dem. (ii) Seja $r = \dim L(\{u_1, ..., u_k\})$. Suponhamos que $r < k$. Suponhamos ainda, a menos de uma reordenação, que o conjunto $\{u_1, ..., u_r\}$ é linearmente independente.

Como o conjunto $\{u_1, ..., u_r, u_{r+1}\}$ é linearmente dependente, então existem escalares não todos nulos $c_1, ..., c_r, c_{r+1}$ tais que

$$c_1 u_1 + \cdots + c_r u_r + c_{r+1} u_{r+1} = \mathbf{0}. \quad (*)$$

Note que tem que se ter $c_{r+1} \neq 0$ caso contrário o conjunto $\{u_1, ..., u_r\}$ é linearmente dependente. Logo

$$c_{r+1} u_{r+1} \neq \mathbf{0}$$

e assim $c_1, ..., c_r$ não podem ser todos nulos.

Por outro lado, atendendo a (*) tem-se

$$A(c_1 u_1 + \cdots + c_r u_r + c_{r+1} u_{r+1}) = \mathbf{0} \Leftrightarrow c_1 A u_1 + \cdots + c_r A u_r + c_{r+1} A u_{r+1} = \mathbf{0} \Leftrightarrow$$

$$\Leftrightarrow c_1 \lambda_1 u_1 + \cdots + c_r \lambda_r u_r + c_{r+1} \lambda_{r+1} u_{r+1} = \mathbf{0} \quad (**).$$

Logo, multiplicando $-\lambda_{r+1}$ por (*) e somando a (**) tem-se:

$$-\lambda_{r+1}(c_1 u_1 + \cdots + c_r u_r + c_{r+1} u_{r+1}) + c_1 \lambda_1 u_1 + \cdots + c_r \lambda_r u_r + c_{r+1} \lambda_{r+1} u_{r+1} = \mathbf{0} \Leftrightarrow$$

$$\Leftrightarrow c_1(\lambda_1 - \lambda_{r+1}) u_1 + \cdots + c_r(\lambda_r - \lambda_{r+1}) u_r = \mathbf{0}$$

Assim, sendo os escalares $c_1, ..., c_r$ não todos nulos e sendo os escalares $\lambda_1, ..., \lambda_k$ todos distintos, então o conjunto $\{u_1, ..., u_r\}$ seria linearmente dependente, contrariando a hipótese de o mesmo ser linearmente independente. Logo, tem que se ter $r = k$.

Definição 47. **(i)** Seja A uma matriz $n \times n$. Se existir uma matriz P invertível tal que

$$D = PAP^{-1},$$

com D **matriz diagonal**, então diz-se que A é uma **matriz diagonalizável** e que P é a **matriz diagonalizante**. No caso de A ser uma matriz diagonal, a matriz diagonalizante é a matriz identidade.

(ii) Seja V um espaço linear tal que $\dim V = n$. Seja $T : V \to V$ uma transformação linear. Diz-se que T **é diagonalizável** se existir uma base $\mathcal{B}$ de V em relação à qual a matriz $M(T; \mathcal{B}; \mathcal{B})$ que representa T nessa base seja uma matriz diagonal.

Teorema 58. Seja $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. A matriz A é diagonalizável se e só se existir uma base $\mathcal{B}_{vp}$ de $\mathbb{R}^n$ apenas constituída por vectores próprios de A. Neste caso, as entradas da diagonal principal da matriz diagonal D serão os valores próprios de A apresentados pela ordem dos vectores próprios correspondentes na base $\mathcal{B}_{vp}$. Além disso, a matriz P^{-1} será a matriz cujas colunas serão os vectores próprios de A, da base $\mathcal{B}_{vp}$ de $\mathbb{R}^n$ dispostos pela mesma ordem, tendo-se

$$D = PAP^{-1}.$$

O mesmo se aplica a $\mathbb{C}^n$.

Teorema 59. Seja A uma matriz $n \times n$. Então as afirmações seguintes são equivalentes:

(i) A é diagonalizável.

(ii) A tem n vectores próprios linearmente independentes.

(iii) A soma das multiplicidades geométricas dos valores próprios de A é n.

(iv) A multiplicidade geométrica de cada valor próprio de A é igual à multiplicidade algébrica desse valor próprio.

Observação 43. **(i)** Seja V um espaço linear tal que $\dim V = n$. Seja $A = M(T, \mathcal{B}, \mathcal{B})$ a matriz $n \times n$ que representa a transformação linear $T : V \to V$ em relação à base ordenada $\mathcal{B}$. No caso de haver uma base $\mathcal{B}_{vp}$ (ordenada) de V apenas constituída por vectores próprios de T, então tem-se

$$M(T, \mathcal{B}_{vp}, \mathcal{B}_{vp}) = PAP^{-1},$$

onde $P^{-1} = S_{\mathcal{B}_{vp} \to \mathcal{B}}$, sendo deste modo $M(T, \mathcal{B}_{vp}, \mathcal{B}_{vp})$ a matriz diagonal cujas entradas da diagonal principal são os valores próprios de A apresentados pela ordem dos vectores próprios correspondentes na base $\mathcal{B}_{vp}$. Assim, T é representada relativamente a uma base por uma matriz diagonal, isto é, T **é diagonalizável**.

No caso de se ter $V = \mathbb{R}^n$ e $\mathcal{B} = \mathcal{B}_c^n$ (base canónica) as colunas da matriz $P^{-1} = S_{\mathcal{B}_{vp} \to \mathcal{B}_c^n}$ são os vectores próprios de A da base $\mathcal{B}_{vp}$ dispostos pela mesma ordem.

(ii) No caso de se ter $D = PAP^{-1}$, com P invertível e D matriz diagonal, tem-se, para $k \in \mathbb{N}$,

$$D^k = PA^kP^{-1}, \qquad \text{ou seja}, \qquad A^k = P^{-1}D^kP.$$

Exemplo 40. Nos exemplos que se seguem as matrizes A consideradas poderão ser vistas como matrizes que representam transformações lineares T relativamente à base canónica (ou outras) de $\mathbb{R}^3$, tendo-se nesse casos, para todo o $u \in \mathbb{R}^3$,

$$T(u) = Au.$$

Deste modo, os valores próprios e vectores próprios de T serão respectivamente os valores próprios e vectores próprios de A.

(i) Uma matriz com valores próprios distintos.

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

O polinómio característico é dado por

$$
\begin{aligned}
\det(A-\lambda I) &= \begin{vmatrix} 1-\lambda & 5 & -1 \\ 0 & -2-\lambda & 1 \\ -4 & 0 & 3-\lambda \end{vmatrix} = \\
&= (1-\lambda)(-2-\lambda)(3-\lambda) - 20 + 4(2+\lambda) = \\
&= (1-\lambda)(-2-\lambda)(3-\lambda) + 4\lambda - 12 = \\
&= (3-\lambda)[(\lambda-1)(\lambda+2) - 4] = \\
&= (3-\lambda)\left(\lambda^2+\lambda-6\right) = \\
&= (3-\lambda)(\lambda-2)(\lambda+3).
\end{aligned}
$$

Os valores próprios de A são os valores de λ para os quais $\det(A-\lambda I)=0$. Logo, os valores próprios de A são

$$\lambda_1 = 3, \quad \lambda_2 = 2 \quad \text{e} \quad \lambda_3 = -3.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $u \in \mathbb{R}^3$ para os quais

$$(A-\lambda I)\,u = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A-\lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$. Tem-se

$$\mathcal{N}(A-\lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} -2 & 5 & -1 \\ 0 & -5 & 1 \\ -4 & 0 & 0 \end{bmatrix}\right) = L(\{(0,1,5)\}).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}(A-\lambda_1 I) = L(\{(0,1,5)\}).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$ são

$$u = (0,s,5s), \ \text{ com } \ s \in \mathbb{R}\backslash\{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$. Tem-se

$$\mathcal{N}(A-\lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} -1 & 5 & -1 \\ 0 & -4 & 1 \\ -4 & 0 & 1 \end{bmatrix}\right) = L(\{(1,1,4)\}).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}(A-\lambda_2 I) = L(\{(1,1,4)\}).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$ são

$$u = (s,s,4s), \ \text{ com } \ s \in \mathbb{R}\backslash\{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_3 = -3$. Tem-se

$$\mathcal{N}(A-\lambda_3 I) = \mathcal{N}\left(\begin{bmatrix} 4 & 5 & -1 \\ 0 & 1 & 1 \\ -4 & 0 & 6 \end{bmatrix}\right) = L(\{(3,-2,2)\}).$$

Logo, o subespaço próprio E_{λ_3} é dado por

$$E_{\lambda_3} = \mathcal{N}\left(A - \lambda_3 I\right) = L\left(\{(3, -2, 2)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_3 = -3$ são

$$u = (3s, -2s, 2s), \quad \text{com} \quad s \in \mathbb{R} \backslash \{0\}.$$

Atendendo a que os valores próprios de A são distintos, pelo teorema 57, os vectores próprios de A associados a esses valores próprios são linearmente independentes. Como $\dim \mathbb{R}^3 = 3$, então 3 vectores em $\mathbb{R}^3$ linearmente independentes formarão desde logo uma base de $\mathbb{R}^3$. Logo, o conjunto

$$\mathcal{B} = \{(0, 1, 5), (1, 1, 4), (3, -2, 2)\}$$

é uma base de $\mathbb{R}^3$. Deste modo, temos uma base de $\mathbb{R}^3$ formada só por vectores próprios de A. Logo, a matriz A é diagonalizável, isto é, existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} é diagonal, tendo-se

$$D = PAP^{-1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \quad \text{com} \quad P^{-1} = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 1 & -2 \\ 5 & 4 & 2 \end{bmatrix}.$$

Note que cada coluna de P^{-1} é formada pelo vector próprio associado ao valor próprio respectivo e na posição respectiva. Além disso, tem-se

$$\begin{array}{ccc} (\mathbb{R}^3, \mathcal{B}_c^3) & \underset{T}{\overset{M(T;\mathcal{B}_c^3;\mathcal{B}_c^3)}{\longrightarrow}} & (\mathbb{R}^3, \mathcal{B}_c^3) \\ S_{\mathcal{B}_c^3 \to \mathcal{B}} \downarrow I & & I \downarrow S_{\mathcal{B}_c^3 \to \mathcal{B}} \\ (\mathbb{R}^3, \mathcal{B}) & \underset{M(T;\mathcal{B};\mathcal{B})}{\overset{T}{\longrightarrow}} & (\mathbb{R}^3, \mathcal{B}) \end{array}$$

com

$$S_{\mathcal{B}_c^3 \to \mathcal{B}} = P, \quad M(T; \mathcal{B}; \mathcal{B}) = D \quad \text{e} \quad M(T; \mathcal{B}_c^3; \mathcal{B}_c^3) = A.$$

(ii) Uma matriz com valores próprios repetidos mas diagonalizável.

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

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 3 & 3 & 4-\lambda \end{vmatrix} = \\ &= (2-\lambda)(3-\lambda)(4-\lambda) + 6 + 6 - 3(3-\lambda) - 6(2-\lambda) - 2(4-\lambda) = \\ &= -\lambda^3 + 9\lambda^2 - 15\lambda + 7 = \\ &= -(\lambda - 1)(\lambda - 1)(\lambda - 7). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A-\lambda I)=0$. Logo, os valores próprios de A são

$$\lambda_1 = 1 \quad \text{e} \quad \lambda_2 = 7.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $u \in \mathbb{R}^3$ para os quais

$$(A-\lambda I)\, u = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}\,(A-\lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 1$. Tem-se

$$\mathcal{N}\,(A-\lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}\right) = L\left(\{(-1,1,0),(-1,0,1)\}\right).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}\,(A-\lambda_1 I) = L\left(\{(-1,1,0),(-1,0,1)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 1$ são

$$u = (-s-t, s, t)\,, \text{ com } s \text{ e } t \text{ não simultâneamente nulos.}$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 7$. Tem-se

$$\mathcal{N}\,(A-\lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} -5 & 1 & 1 \\ 2 & -4 & 2 \\ 3 & 3 & -3 \end{bmatrix}\right) = L\left(\{(1,2,3)\}\right).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}\,(A-\lambda_2 I) = L\left(\{(1,2,3)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 7$ são

$$u = (s, 2s, 3s)\,, \text{ com } s \in \mathbb{R}\backslash\{0\}\,.$$

Atendendo a que

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} = 3,$$

podemos ter a seguinte base de $\mathbb{R}^3$ formada só por vectores próprios de A

$$\mathcal{B} = \{(-1,1,0),(-1,0,1),(1,2,3)\}\,.$$

Logo, a matriz A é diagonalizável, isto é, existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} é diagonal, tendo-se

$$D = PAP^{-1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \quad \text{com} \quad P^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}.$$

Note que cada coluna de P^{-1} é formada pelo vector próprio associado ao valor próprio respectivo e na posição respectiva. Além disso, tem-se

$$
\begin{array}{ccc}
(\mathbb{R}^3, \mathcal{B}_c^3) & \underset{T}{\overset{M(T;\mathcal{B}_c^3;\mathcal{B}_c^3)}{\longrightarrow}} & (\mathbb{R}^3, \mathcal{B}_c^3) \\
S_{\mathcal{B}_c^3 \to \mathcal{B}} \downarrow I & & I \downarrow S_{\mathcal{B}_c^3 \to \mathcal{B}} \\
(\mathbb{R}^3, \mathcal{B}) & \underset{M(T;\mathcal{B};\mathcal{B})}{\overset{T}{\longrightarrow}} & (\mathbb{R}^3, \mathcal{B})
\end{array}
$$

com

$$S_{\mathcal{B}_c^3 \to \mathcal{B}} = P, \quad M(T;\mathcal{B};\mathcal{B}) = D \quad \text{e} \quad M(T;\mathcal{B}_c^3;\mathcal{B}_c^3) = A.$$

(iii) Uma matriz com valores próprios repetidos e não diagonalizável.

$$A = \begin{bmatrix} 7 & 5 & -1 \\ 0 & -2 & 1 \\ 20 & 0 & 3 \end{bmatrix}$$

O polinómio característico é dado por

$$
\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} 7-\lambda & 5 & -1 \\ 0 & -2-\lambda & 1 \\ 20 & 0 & 3-\lambda \end{vmatrix} = \\
&= (7-\lambda)(-2-\lambda)(3-\lambda) + 100 - 20(2+\lambda) = \\
&= (3-\lambda)\left[(7-\lambda)(-2-\lambda) + 20\right] = \\
&= (3-\lambda)\left(\lambda^2 - 5\lambda + 6\right) = \\
&= (3-\lambda)(\lambda-3)(\lambda-2).
\end{aligned}
$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 3 \quad \text{e} \quad \lambda_2 = 2.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $u \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, u = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A - \lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$. Tem-se

$$\mathcal{N}(A - \lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} 4 & 5 & -1 \\ 0 & -5 & 1 \\ 20 & 0 & 0 \end{bmatrix}\right) = L\left(\{(0,1,5)\}\right).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}(A - \lambda_1 I) = L\left(\{(0,1,5)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$ são

$$u = (0, s, 5s), \quad \text{com} \quad s \in \mathbb{R} \backslash \{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$. Tem-se

$$\mathcal{N}\left(A - \lambda_2 I\right) = \mathcal{N}\left(\begin{bmatrix} 5 & 5 & -1 \\ 0 & -4 & 1 \\ 20 & 0 & 1 \end{bmatrix}\right) = L\left(\{(1, -5, -20)\}\right).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}\left(A - \lambda_2 I\right) = L\left(\{(1, -5, -20)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$ são

$$u = (s, -5s, -20s), \text{ com } s \in \mathbb{R} \backslash \{0\}.$$

Atendendo a que

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} = 2 < 3,$$

não é possível ter uma base de $\mathbb{R}^3$ formada só por vectores próprios de A. Logo, a matriz A não é diagonalizável, isto é, não existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} seja diagonal.

(iv) Uma matriz com apenas um valor próprio real.

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

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = \\ &= \lambda^2 (1 - \lambda) + (1 - \lambda) = \\ &= (1 - \lambda)\left(\lambda^2 + 1\right). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 1, \quad \lambda_2 = i \quad \text{e} \quad \lambda_3 = -i.$$

Logo, a matriz A não é diagonalizável numa matriz de entradas reais, isto é, não existe uma matriz invertível P diagonalizante tal que a matriz PAP^{-1} seja diagonal com entradas reais. No entanto e atendendo a que os três valores próprios são distintos, a matriz A é diagonalizável numa matriz de entradas complexas:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$$

Exemplo 41. A sucessão de Fibonacci (Leonardo de Pisa, 1202). Seja $(u_n)_{n \in \mathbb{N}}$ tal que

$$u_1 = 1, \quad u_2 = 1 \quad \text{e} \quad u_{n+2} = u_n + u_{n+1}, \ n \in \mathbb{N}.$$

Considerando a igualdade $u_{n+1}=u_{n+1}$, podemos escrever o sistema

$$\left\{\begin{array}{l} u_{n+1}=u_{n+1} \\ u_{n+2}=u_n+u_{n+1} \end{array}\right.$$

isto é

$$\begin{bmatrix} u_{n+1} \\ u_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_n \\ u_{n+1} \end{bmatrix}$$

para todo o $n \in \mathbb{N}$. Aplicando sucessivamente a igualdade anterior tem-se

$$\begin{bmatrix} u_{n+1} \\ u_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_n \\ u_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_{n-1} \\ u_n \end{bmatrix} =$$

$$= \cdots = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Calculemos agora os valores próprios de $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$:

$$\det \begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = 0 \Leftrightarrow (-\lambda)(1-\lambda)-1=0 \Leftrightarrow \lambda^2-\lambda-1=0 \Leftrightarrow$$

$$\Leftrightarrow \lambda = \frac{1+\sqrt{5}}{2} \quad \text{ou} \quad \lambda = \frac{1-\sqrt{5}}{2}.$$

Valores próprios: $\lambda_1 = \frac{1+\sqrt{5}}{2}$ e $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

Atendendo a que

$$\mathcal{N}\begin{bmatrix} -\lambda_1 & 1 \\ 1 & 1-\lambda_1 \end{bmatrix} = \mathcal{N}\begin{bmatrix} 0 & 1+\lambda_1-\lambda_1^2 \\ 1 & 1-\lambda_1 \end{bmatrix} = \mathcal{N}\begin{bmatrix} 0 & 0 \\ 1 & \frac{1-\sqrt{5}}{2} \end{bmatrix} = L\left(\left\{\left(\frac{-1+\sqrt{5}}{2},1\right)\right\}\right)$$

$\left(\frac{-1+\sqrt{5}}{2},1\right)$ é um vector próprio associado ao valor próprio $\frac{1+\sqrt{5}}{2}$, sendo todos os vectores próprios associados ao valor próprio $\frac{1+\sqrt{5}}{2}$ dados por $L\left(\left\{\left(\frac{-1+\sqrt{5}}{2},1\right)\right\}\right) \setminus \{(0,0)\}$.

Atendendo a que

$$\mathcal{N}\begin{bmatrix} -\lambda_2 & 1 \\ 1 & 1-\lambda_2 \end{bmatrix} = \mathcal{N}\begin{bmatrix} 0 & 1+\lambda_2-\lambda_2^2 \\ 1 & 1-\lambda_2 \end{bmatrix} = \mathcal{N}\begin{bmatrix} 0 & 0 \\ 1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} = L\left(\left\{\left(-\frac{1+\sqrt{5}}{2},1\right)\right\}\right)$$

$\left(-\frac{1+\sqrt{5}}{2},1\right)$ é um vector próprio associado ao valor próprio $\frac{1-\sqrt{5}}{2}$, sendo todos os vectores próprios associados ao valor próprio $\frac{1-\sqrt{5}}{2}$ dados por $L\left(\left\{\left(-\frac{1+\sqrt{5}}{2},1\right)\right\}\right) \setminus \{(0,0)\}$.

Como existe uma base de $\mathbb{R}^2$ formada por só por vectores próprios (os dois valores próprios são distintos logo os vectores próprios correspondentes são linearmente independentes) então a matriz $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ é diagonalizável. Assim, fazendo

$$P^{-1} = \begin{bmatrix} \frac{-1+\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$$

tem-se

$$P=\begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \end{bmatrix}$$

e

$$D=\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}=P\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}P^{-1}\Leftrightarrow$$

$$\Leftrightarrow\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}=P^{-1}\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}P.$$

Logo

$$\begin{bmatrix} u_{n+1} \\ u_{n+2} \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n\begin{bmatrix} 1 \\ 1 \end{bmatrix}=\left(P^{-1}\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}P\right)^n\begin{bmatrix} 1 \\ 1 \end{bmatrix}=$$

$$=P^{-1}\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^n P\begin{bmatrix} 1 \\ 1 \end{bmatrix}=P^{-1}\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix}P\begin{bmatrix} 1 \\ 1 \end{bmatrix}=$$

$$=\begin{bmatrix} \frac{-1+\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix}\begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}=$$

$$=\begin{bmatrix} \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n+\frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n \\ \frac{5-3\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n+\frac{3\sqrt{5}+5}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n \end{bmatrix}.$$

Isto é,

$$u_{n+1}=\frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n+\frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

para todo o $n\in\mathbb{N}$, com $u_1=1$.

Verifique que (por exemplo) $u_2=1$, $u_3=2$, $u_4=3$.

Exemplo 42. (Um processo de difusão.) Considere duas células adjacentes separadas por uma membrana permeável e suponha que um fluído passa da 1^a célula para a 2^a a uma taxa (em mililitros por minuto) numericamente igual a 4 vezes o volume (em mililitros) do fluído da 1^a célula. Em seguida, passa da 2^a célula para a 1^a a uma taxa (em mililitros por minuto) numericamente igual a 5 vezes o volume (em mililitros) do fluído da 2^a célula.

Sejam $v_1(t)$ e $v_2(t)$ respectivamente o volume da 1^a célula e o volume da 2^a célula no instante t. Suponha que inicialmente a primeira célula tem 10 mililitros de fluído e que a segunda tem 8 mililitros de fluído, isto é $v_1(0)=10$ e $v_2(0)=8$.

Determinemos o volume de fluído de cada célula no instante t.

Tem-se

$$\begin{cases} v_1'(t)=-4v_1(t) \\ v_2'(t)=4v_1(t)-5v_2(t) \end{cases}$$

isto é

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix}=\begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix}\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}. \quad (*)$$

-4 e -5 são os valores próprios da matriz $\begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix}$, sendo os vectores próprios associados $(1,4)$ e $(0,1)$ respectivamente.

Como existe uma base de $\mathbb{R}^2$ formada por só por vectores próprios (os dois valores próprios são distintos logo os vectores próprios correspondentes são linearmente independentes) então a matriz $\begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix}$ é diagonalizável. Assim, fazendo

$$P^{-1} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \quad \text{tem-se} \quad P = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} = P \begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix} P^{-1} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix} = P^{-1} \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} P.$$

Assim, considerando a mudança de variável

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = P^{-1} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

o sistema (*) é equivalente a

$$P^{-1} \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix} P^{-1} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow P^{-1} \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = \left(P^{-1} \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} P \right) P^{-1} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow P^{-1} \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = P^{-1} \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \Leftrightarrow \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} u_1'(t) = -4u_1(t) \\ \\ u_2'(t) = -5u_2(t) \end{cases} \underset{\substack{\text{Se } u_1 \neq 0 \\ \text{e } u_2 \neq 0}}{\Leftrightarrow} \begin{cases} \dfrac{u_1'(t)}{u_1(t)} = -4 \\ \\ \dfrac{u_2'(t)}{u_2(t)} = -5 \end{cases} \Leftrightarrow \begin{cases} \log|u_1(t)| = -4t + k_1 \\ \\ \log|u_2(t)| = -5t + k_2 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} u_1(t) = e^{-4t+k_1} = c_1 e^{-4t} \\ \\ u_2(t) = e^{-5t+k_2} = c_2 e^{-5t} \end{cases}$$

com $c_1 = e^{k_1}, c_2 = e^{k_2} \in \mathbb{R}$. Logo

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = P^{-1} \begin{bmatrix} c_1 e^{-4t} \\ c_2 e^{-5t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} c_1 e^{-4t} \\ c_2 e^{-5t} \end{bmatrix} = \begin{bmatrix} c_1 e^{-4t} \\ 4c_1 e^{-4t} + c_2 e^{-5t} \end{bmatrix}.$$

Como $\begin{cases} v_1(0) = 10 \\ v_2(0) = 8 \end{cases}$ então $c_1 = 10$ e $c_2 = -32$ e assim a solução geral do sistema de equações diferenciais lineares

$$\begin{cases} v_1'(t) = -4v_1(t) \\ v_2'(t) = 4v_1(t) - 5v_2(t) \end{cases}$$

com os valores iniciais $\begin{cases} v_1(0) = 10 \\ v_2(0) = 8 \end{cases}$ é dada por

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} 10e^{-4t} \\ 40e^{-4t} - 32e^{-5t} \end{bmatrix}.$$

Produtos internos e Ortogonalização

Definição 48. Sejam V um espaço linear real e $\mathbf{0}$ o vector nulo de V. Chama-se **produto interno** em V a uma aplicação

$$\langle , \rangle : V \times V \to \mathbb{R}$$

$$(u, v) \;\; \to \langle u, v \rangle$$

que verifique as três condições seguintes.

(i) Simetria: para todos os $u, v \in V$

$$\langle u, v \rangle = \langle v, u \rangle .$$

(ii) Linearidade: para todo o $v \in V$ (fixo) a aplicação

$$V \to \mathbb{R}$$

$$u \to \langle u, v \rangle$$

é linear.

(iii) Positividade: para todo o $u \in V$ tal que $u \neq \mathbf{0}$,

$$\langle u, u \rangle > 0.$$

Tendo-se $\langle u, u \rangle = 0$ se e só se $u = \mathbf{0}$.

Observação 44. (a) Um produto interno num espaço linear real é uma forma **bilinear**, **simétrica** e **definida positiva**.

(b) Num espaço linear V sobre $\mathbb{C}$ (espaço linear complexo), um produto interno é uma aplicação que a cada par de vectores $(u, v) \in V \times V$ associa o número complexo $\langle u, v \rangle$ e que verifica as seguintes condições:

(i) Para todos os $u, v \in V$

$$\langle u, v \rangle = \overline{\langle v, u \rangle}.$$

(ii) Para todo o $v \in V$ (fixo) tem-se

$$\langle \alpha u + \beta w, v \rangle = \overline{\alpha} \langle u, v \rangle + \overline{\beta} \langle w, v \rangle$$

para todos os $u, w \in V$ e $\alpha, \beta \in \mathbb{C}$, (onde por exemplo $\overline{\alpha} = a - bi$ se $\alpha = a + bi$) e a aplicação, para todo o $u \in V$ (fixo)

$$V \to \mathbb{C}$$

$$v \to \langle u, v \rangle$$

é linear.

(iii) Para todo o $u \in V$ tal que $u \neq \mathbf{0}$,

$$\langle u, u \rangle > 0.$$

Tendo-se $\langle u, u \rangle = 0$ se e só se $u = \mathbf{0}$.

(c) A um espaço linear real de dimensão finita com um produto interno chama-se **espaço euclidiano**. A um espaço linear complexo de dimensão finita com um produto interno chama-se **espaço unitário**.

Observação 45. **(i)** Seja V um espaço linear real. Seja $\mathcal{B} = \{w_1, w_2, ..., w_n\}$ uma base de V. Sejam $u, v \in V$. Sejam

$$\alpha_1, \alpha_2, ..., \alpha_n \quad \text{e} \quad \beta_1, \beta_2, ..., \beta_n$$

as coordenadas de u e de v na base $\mathcal{B}$ respectivamente, isto é,

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n = \sum_{i=1}^{n} \alpha_i w_i$$

e

$$v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n = \sum_{i=1}^{n} \beta_i w_i.$$

Logo,

$$\langle u, v \rangle = \left\langle \sum_{i=1}^{n} \alpha_i w_i, \sum_{i=1}^{n} \beta_i w_i \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \langle w_i, w_j \rangle =$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix} \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \dots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \dots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \dots & \langle w_n, w_n \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Assim, a aplicação $\langle , \rangle : V \times V \to \mathbb{R}$ que a cada $(u, v) \in V \times V$ faz corresponder $\langle u, v \rangle$, é um produto interno em V se e só se a matriz

$$G = \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \dots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \dots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \dots & \langle w_n, w_n \rangle \end{bmatrix}$$

fôr simétrica ($G = G^T$) e definida positiva ($u^T G u > 0$, para todo o $u \neq \mathbf{0}$). Note-se que a linearidade é consequência das propriedades referentes às operações matriciais envolvidas.

(ii) À matriz G anterior dá-se o nome de matriz da métrica do produto interno.

(iii) Num próximo capítulo (teorema 78 (i)), como consequência da diagonalização ortogonal (teorema 75 e observação 59), sendo G simétrica ($G = G^T$), será estabelecida a equivalência:

$$(u^T G u > 0, \quad \text{para todo o} \quad u \neq \mathbf{0}) \Leftrightarrow (\text{todos os valores próprios de } G \text{ são positivos}).$$

(iv) Observe-se ainda que no caso de se ter um espaço unitário pode-se encontrar uma matriz G cujos valores próprios sejam todos positivos e tal que $G = \overline{G}^T$, (onde $\overline{G}$ é a matriz que se obtem de G passando todas as entradas desta ao complexo conjugado), tendo-se

$$\langle u, v \rangle = \begin{bmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \dots & \overline{\alpha_n} \end{bmatrix} G \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Uma matriz A que satisfaça a condição $A = \overline{A}^T$ diz-se **hermiteana**.

Teorema 60. Seja V um espaço linear real com $\dim V = n$. Seja $\{w_1, w_2, ..., w_n\}$ uma base de V. Então, uma aplicação

$$\langle , \rangle : V \times V \to \mathbb{R}$$

é um produto interno (em V) se e só se

$$\langle u, v \rangle = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix} G \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix},$$

com

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n$$

e

$$v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n$$

e G é uma matriz simétrica cujos valores próprios são todos positivos. Se a aplicação $\langle , \rangle$ fôr um produto interno tem-se

$$G = \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \dots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \dots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \dots & \langle w_n, w_n \rangle \end{bmatrix}.$$

Exemplo 43. **(i)** Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$. Esta aplicação é um produto interno em $\mathbb{R}^2$ a que se dá o nome de produto interno usual em $\mathbb{R}^2$, uma vez que

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} G \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com

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

A matriz G é simétrica e o único valor próprio de G é $1 > 0$.

(ii) Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = -2\alpha_1\beta_1 + 3\alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$. Esta aplicação não é um produto interno em $\mathbb{R}^2$, uma vez que

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = -2\alpha_1\beta_1 + 3\alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} G \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com

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

A matriz G é simétrica, no entanto, os valores próprios de G: -2 e 3 não são ambos positivos.

(iii) O produto interno usual em $\mathbb{R}^n$ é dado por:

$$\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

$$(u, v) \quad \to \langle u, v \rangle = u^T v,$$

onde $u^T = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}$ e $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

(iv) O produto interno usual em $\mathbb{C}^n$ é dado por:

$$\langle , \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$$

$$(u, v) \quad \to \langle u, v \rangle = u^H v,$$

onde $u^H = \overline{u}^T = \begin{bmatrix} \overline{u_1} & \overline{u_2} & \dots & \overline{u_n} \end{bmatrix}$ e $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

(v) Um produto interno em $\mathcal{M}_{m\times n}(\mathbb{R})$.

$$\langle , \rangle : M_{m\times n}(\mathbb{R}) \times M_{m\times n}(\mathbb{R}) \to \mathbb{R}$$

$$(A, B) \quad \to \langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} = \operatorname{tr}\left(A^T B\right).$$

(vi) Um produto interno em $C([a, b])$.

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

$$(f,g) \quad \to \langle f,g\rangle = \int_a^b f(x)\,g(x)\,dx.$$

Exemplo 44. $\mathbb{R}^2$ **com um produto interno não usual.** Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = 2\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$.

É fácil ver que esta aplicação é simétrica e linear em relação a (α_1, α_2) (fixando (β_1, β_2)). Vejamos por exemplo que a condição

$$\langle (\alpha_1, \alpha_2), (\alpha_1, \alpha_2) \rangle > 0, \quad \text{para todo o } (\alpha_1, \alpha_2) \neq (0,0),$$

é satisfeita.

Atendendo a que

$$\langle (\alpha_1, \alpha_2), (\alpha_1, \alpha_2) \rangle = 2\alpha_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2 = \alpha_1^2 + (\alpha_1 + \alpha_2)^2,$$

tem-se

$$\langle (\alpha_1, \alpha_2), (\alpha_1, \alpha_2) \rangle = 0 \Leftrightarrow (\alpha_1 = 0 \text{ e } \alpha_1 + \alpha_2 = 0) \Leftrightarrow$$
$$\Leftrightarrow (\alpha_1 = 0 \text{ e } \alpha_2 = 0) \Leftrightarrow (\alpha_1, \alpha_2) = (0,0).$$

Em alternativa, podemos escrever

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = 2\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} G \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com

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

A matriz G é simétrica e os valores próprios de G: $\frac{3+\sqrt{5}}{2}$ e $\frac{3-\sqrt{5}}{2}$ são ambos positivos.

Definição 49. Sejam V um espaço linear com um produto interno e $\mathbf{0}$ o vector nulo de V. Sejam $u, v \in V$.

(i) Chama-se **norma** de u a:

$$\|u\| = \sqrt{\langle u, u\rangle}.$$

(ii) Chama-se **projecção ortogonal** de v sobre $u \neq \mathbf{0}$ a:

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

(iii) Diz-se que u e v são **ortogonais** se $\langle u, v\rangle = 0$.

(iv) Chama-se **ângulo** entre dois vectores não nulos u e v tais que $\langle u, v\rangle \in \mathbb{R}$ a:

$$\theta = \arccos \frac{\langle u, v\rangle}{\|u\| \, \|v\|}.$$

Note que este ângulo está bem definido atendendo ao teorema 61.

Observação 46. **(i)** O ângulo θ entre dois vectores não nulos u e v é $\frac{\pi}{2}$ se e só se u e v são ortogonais.

(ii) Para cada $u \in V$ (fixo) com $u \neq \mathbf{0}$, a aplicação $\operatorname{proj}_u : V \to V$ que a cada $v \in V$ faz corresponder $\operatorname{proj}_u v$, é uma transformação linear.

Teorema 61. Desigualdade de Cauchy-Schwarz. Seja V um espaço linear com um produto interno. Então, para todos os $u, v \in V$,

$$|\langle u, v\rangle| \leq \|u\| \, \|v\|$$

Observação 47. **(i) Teorema de Pitágoras.** Sejam $u, v \in \mathbb{R}^2$. Tem-se u e v ortogonais se e só se

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

Dem.

$$\begin{aligned} \|u - v\|^2 &= \langle u - v, u - v\rangle = \langle u, u\rangle - \langle v, u\rangle - \langle u, v\rangle + \langle v, v\rangle = \\ &= \|u\|^2 - 2\langle u, v\rangle + \|v\|^2 = \|u\|^2 + \|v\|^2 \end{aligned}$$

se e só se

$$\langle u, v\rangle = 0,$$

isto é, se e só se u e v forem ortogonais.

(ii) Em $\mathbb{R}^2$ com o produto interno usual, a desigualdade de Cauchy-Schwarz é dada por

$$|\alpha_1\beta_1 + \alpha_2\beta_2| \leq \sqrt{\alpha_1^2 + \alpha_2^2}\sqrt{\beta_1^2 + \beta_2^2},$$

uma vez que

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = \alpha_1\beta_1 + \alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$.

(iii) Em $\mathbb{R}^n$ com o produto interno usual, a desigualdade de Cauchy-Schwarz é dada por

$$\left|\sum_{i=1}^{n} \alpha_i\beta_i\right| \leq \sqrt{\sum_{i=1}^{n} \alpha_i^2}\sqrt{\sum_{i=1}^{n} \beta_i^2},$$

uma vez que

$$\langle (\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n)\rangle = \alpha_1\beta_1 + ... + \alpha_n\beta_n,$$

com $(\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n) \in \mathbb{R}^n$.

Teorema 62. Sejam V um espaço linear com um produto interno e $\mathbf{0}$ o vector nulo de V. Sejam $u, v \in V$ e λ escalar. A norma é uma aplicação $\|\| : V \to \mathbb{R}$ que satisfaz as seguintes propriedades.

(i) Positividade: $\|u\| > 0$ se $u \neq \mathbf{0}$.

(ii) Homogeneidade: $\|\lambda u\| = |\lambda| \, \|u\|$

(iii) Desigualdade triangular: $\|u + v\| \leq \|u\| + \|v\|$

Observação 48. Pode definir-se **norma** num espaço linear V, sem estar associada a qualquer produto interno, como sendo uma aplicação de V em $\mathbb{R}$ que satisfaz as propriedades do teorema anterior. A um espaço linear com uma norma chama-se **espaço normado**.

Observação 49. Seja V um espaço linear com um produto interno. Sejam $u, v \in V$. Tem-se

$$\langle u, v \rangle = \frac{1}{2} \left(\|u + v\|^2 - \|u\|^2 - \|v\|^2 \right).$$

Observação 50. Seja V um espaço normado. Sejam $u, v \in V$. Então, a norma pode ser obtida de um produto interno na forma

$$\|u\| = \sqrt{\langle u, u \rangle}$$

se e só se

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

Esta última equação é conhecida por **lei do paralelogramo**.

Exemplo 45. Uma norma que não é obtida a partir de um produto interno. Seja $\|\| : \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por

$$\|(\alpha_1, \alpha_2)\| = |\alpha_1| + |\alpha_2| ,$$

com $(\alpha_1, \alpha_2) \in \mathbb{R}^2$. É fácil verificar que esta aplicação satisfaz as três condições do teorema 62. Logo, é uma norma. No entanto, é também fácil verificar que esta norma não satisfaz a lei do paralelogramo. Logo, esta norma não poderá ser obtida a partir de um produto interno.

Definição 50. (i) Sejam V um espaço euclidiano (ou unitário) e $S \subset V$. Diz-se que S é **ortogonal** se para todos os $u, v \in S$ com $u \neq v$, se tiver

$$\langle u, v \rangle = 0.$$

Diz-se que S é **ortonormado** se fôr ortogonal e se, para todo o $u \in S$, se tiver

$$\|u\| = 1.$$

(ii) Sejam V um espaço euclidiano (ou unitário) e U_1, U_2 dois subespaços de V. Diz-se que U_1 e U_2 são **ortogonais** e escreve-se $U_1 \perp U_2$ se para quaisquer $u \in U_1$ e $v \in U_2$ se tiver $\langle u, v\rangle = 0$.

Teorema 63. Sejam V um espaço linear com um produto interno e $S \subset V$. Seja $\mathbf{0}$ o vector nulo de V. Se S é ortogonal e $\mathbf{0} \notin S$ então S é linearmente independente. Em particular, se $n = \dim V$ então qualquer conjunto S ortogonal de n vectores não nulos é uma base de V.

Teorema 64. Seja V um espaço euclidiano (ou unitário) com $\dim V = n$. Seja $\mathcal{B} = \{u_1, ..., u_n\}$ uma base ortogonal de V. Então, as coordenadas de um vector $v \in V$ em relação à base $\mathcal{B}$ são dadas por:

$$\alpha_j = \frac{\langle v, u_j\rangle}{\langle u_j, u_j\rangle},$$

com $j = 1, ..., n$. Se $\mathcal{B}$ fôr ortonormada então as coordenadas de um vector $v \in V$ em relação à base $\mathcal{B}$ são dadas por:

$$\alpha_j = \langle v, u_j\rangle,$$

com $j = 1, ..., n$.

Teorema 65. Seja V um espaço euclidiano com $\dim V = n$. Seja $\mathcal{B} = \{w_1, ..., w_n\}$ uma base ortonormada de V. Então, para todos os $u, v \in V$, tem-se

$$\langle u, v\rangle = \sum_{i=1}^{n} \langle u, w_i\rangle \langle v, w_i\rangle \quad \text{(fórmula de Parseval)} \quad \text{e} \quad \|u\| = \sqrt{\sum_{i=1}^{n} \langle u, w_i\rangle^2}.$$

Observação 51. Seja V um espaço euclidiano com $\dim V = n$. Seja

$$\mathcal{B} = \{w_1, ..., w_n\}$$

uma base ortonormada de V. Sejam $u, v \in V$, com

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n, \quad v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n.$$

Então, atendendo ao teorema 64, a fórmula de Parseval é dada por:

$$\langle u, v\rangle = \sum_{i=1}^{n} \alpha_i \beta_i = \alpha_1\beta_1 + \alpha_2\beta_2 + ... + \alpha_n\beta_n \quad \text{e} \quad \text{tem-se} \quad \|u\| = \sqrt{\sum_{i=1}^{n} \alpha_i^2}.$$

Notação 3. Sejam V um espaço euclidiano (ou unitário) e $\mathbf{0}$ o vector nulo de V. Para qualquer $v \in V$, com $v \neq \mathbf{0}$, o vector $\dfrac{1}{\|v\|}v$ será denotado por $\dfrac{v}{\|v\|}$.

Teorema 66. Método de ortogonalização de Gram-Schmidt. Seja V um espaço euclidiano (ou unitário) não nulo. Então V tem bases ortonormadas. Mais concretamente, considere o conjunto linearmente independente:

$$\{v_1, v_2, ..., v_k\} \subset V$$

e sejam

$$\begin{aligned} u_1 &= v_1, \\ u_2 &= v_2 - \operatorname{proj}_{u_1} v_2, \\ &\;\;... \\ u_k &= v_k - \operatorname{proj}_{u_1} v_k - ... - \operatorname{proj}_{u_{k-1}} v_k \end{aligned}$$

então

(i) $L(\{u_1, u_2, ..., u_k\}) = L(\{v_1, v_2, ..., v_k\})$;

(ii) o conjunto $\{u_1, u_2, ..., u_k\}$ é uma base ortogonal de $L(\{v_1, v_2, ..., v_k\})$.

(iii) o conjunto $\left\{\frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}, ..., \frac{u_k}{\|u_k\|}\right\}$ é uma base ortonormada de $L(\{v_1, v_2, ..., v_k\})$.

Exemplo 46. Considere-se $\mathbb{R}^4$ com o produto interno usual. Seja

$$U = L(\{(1, 1, -1, -1), (1, 2, 3, 4), (2, 1, -6, -7), (1, 3, 7, 9)\}).$$

Determinemos a dimensão de U e uma base ortonormada para U. Tem-se

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ -1 & 3 & -6 & 7 \\ -1 & 4 & -7 & 9 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 4 & -4 & 8 \\ 0 & 5 & -5 & 10 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Logo, o conjunto $\{v_1, v_2\}$, com $v_1 = (1, 1, -1, -1)$ e $v_2 = (1, 2, 3, 4)$, é uma base de U e como tal $\dim U = 2$.

Sejam $u_1 = v_1$ e $u_2 = v_2 - \operatorname{proj}_{u_1} v_2$.

Logo, o conjunto $\{u_1, u_2\}$, com $u_1 = (1, 1, -1, -1)$ e

$$u_2 = (1, 2, 3, 4) - \frac{1 + 2 - 3 - 4}{4}(1, 1, -1, -1) = (2, 3, 2, 3),$$

é uma base ortogonal de U. Uma base ortonormada para U:

$$\left\{\frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}\right\} = \left\{\left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right), \left(\frac{\sqrt{26}}{13}, \frac{3\sqrt{26}}{26}, \frac{\sqrt{26}}{13}, \frac{3\sqrt{26}}{26}\right)\right\}$$

Teorema 67. Seja $\mathcal{B} = \{u_1, u_2, ..., u_n\}$ uma base de um espaço euclidiano (ou unitário). A base $\mathcal{B}$ é ortonormada se e só se a matriz da métrica G em relação a essa base fôr a matriz identidade. Em $\mathbb{R}^n$ o produto interno usual é aquele em relação ao qual a base canónica é ortonormada.

Teorema 68. Seja $\{v_1, v_2, ..., v_n\}$ uma base de $\mathbb{R}^n$. Então, existe um único produto interno em $\mathbb{R}^n$ para o qual esta base é ortonormada.

Exemplo 47. Considere em $\mathbb{R}^2$ a base $\mathcal{B} = \{v_1, v_2\}$, com $v_1 = (1,0)$ e $v_2 = (1,1)$. Vejamos que existe um e um só produto interno para o qual a base $\mathcal{B}$ é ortonormada.

Seja $\mathcal{B}_c^2 = \{(1,0),(0,1)\}$ a base canónica de $\mathbb{R}^2$. Tem-se

$$S_{B_c^2 \to \mathcal{B}} = \left(S_{\mathcal{B} \to B_c^2}\right)^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Sejam $u, v \in \mathbb{R}^2$. Tem-se

$$u = (\alpha_1, \alpha_2) \quad \text{e} \quad v = (\beta_1, \beta_2),$$

onde α_1, α_2 e β_1, β_2 são as coordenadas na base $\mathcal{B}_c^2$ de u e v respectivamente. Seja $P = S_{\mathcal{B}_c^2 \to \mathcal{B}}$. Logo, a aplicação $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2$ definida por

$$\langle u, v \rangle = (Pu)^T G \, (Pv)$$

com

$$G = \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

é um produto interno e é o único para o qual a base $\mathcal{B}$ é ortonormada. Tem-se então

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \left(\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \right)^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \right) =$$
$$= \alpha_1\beta_1 - \alpha_1\beta_2 - \alpha_2\beta_1 + 2\alpha_2\beta_2.$$

É fácil verificar que para este produto interno a base $\mathcal{B}$ é ortonormada:

$$\langle (1,0), (1,1) \rangle = 0 \;\; \text{e} \;\; \langle (1,0), (1,0) \rangle = \langle (1,1), (1,1) \rangle = 1.$$

Definição 51. Sejam V um espaço euclidiano (ou unitário) e U um subespaço de V. Diz-se que um elemento de V é **ortogonal a** U se fôr ortogonal a todos os elementos de U. Ao conjunto de todos os elementos ortogonais a U chama-se **complemento ortogonal** de U e designa-se por $U^\perp$.

Note-se que se U_1 e U_2 são dois subespaços de V tais que $U_1 \perp U_2$ (Def. 50) então tem-se $U_1 \subset U_2^\perp$ e $U_2 \subset U_1^\perp$.

Teorema 69. Qualquer que seja o subespaço U de um espaço euclidiano (ou unitário) V, também $U^\perp$ é um subespaço de V.

Exemplo 48. **(i)** Se $U \subset \mathbb{R}^3$ é um plano que passa pela origem, então $U^\perp$ é uma recta que passa pela origem e é perpendicular ao plano.

(ii) Se $U \subset \mathbb{R}^3$ é uma recta que passa pela origem, então $U^\perp$ é um plano que passa pela origem e é perpendicular à recta.

(iii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Então,

$$\mathcal{N}(A) = (\mathcal{L}(A))^\perp = \left(\mathcal{C}(A^T)\right)^\perp \qquad \text{e} \quad \mathcal{N}(A^T) = \left(\mathcal{L}(A^T)\right)^\perp = (\mathcal{C}(A))^\perp.$$

(iv) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é invertível. Então, $(\mathcal{N}(A))^\perp = \mathbb{R}^n$ e $(\mathcal{L}(A))^\perp = \{\mathbf{0}\}$.

Teorema 70. Se U é um subespaço de dimensão finita de um espaço euclidiano (ou unitário) V, então V é a soma directa de U e $U^{\perp}$, isto é, $V = U \oplus U^{\perp}$. Logo, cada elemento $v \in V$ pode ser escrito de modo único como soma de um elemento de U com um elemento de $U^{\perp}$:

$$v = v_U + v_{U^{\perp}}, \quad \text{com} \quad v_U \in U \quad \text{e} \quad v_{U^{\perp}} \in U^{\perp}.$$

À aplicação linear $P_U : V \to V$ definida por $P_S(v) = v_S$ e tal que $P_U = P_U \circ P_U$ chama-se **projecção ortogonal de** V **sobre** U e à aplicação linear $P_{U^{\perp}} : V \to V$ definida por $P_{U^{\perp}}(v) = v_{U^{\perp}}$ e tal que $P_{U^{\perp}} = P_{U^{\perp}} \circ P_{U^{\perp}}$ chama-se **projecção ortogonal de** V **sobre** $U^{\perp}$. Tem-se

$$I = P_U + P_{U^{\perp}}.$$

Se $\{w_1, w_2, ..., w_l\}$ é uma base ortonormada de U, então

$$P_U(v) = \sum_{i=1}^{l} \langle v, w_i \rangle \, w_i,$$

para todo o $v \in V$.

Se $\{u_1, u_2, ..., u_k\}$ é uma base ortonormada de $U^{\perp}$, então

$$P_{U^{\perp}}(v) = \sum_{j=1}^{k} \langle v, u_j \rangle \, u_j,$$

para todo o $v \in V$.

Neste caso, $\{w_1, w_2, ..., w_l, u_1, u_2, ..., u_k\}$ é uma base ortonormada de V.

As aplicações P_U e $P_{U^{\perp}}$ são transformações lineares de V em V que satisfazem as propriedades:

(i) $P_U(V) = U$

(ii) $P_{U^{\perp}}(V) = U^{\perp}$

(iii) $(P_U)^2 = P_U$

(iv) $(P_{U^{\perp}})^2 = P_{U^{\perp}}$

(v) $\langle P_U(u), v \rangle = \langle u, P_U(v) \rangle$, $\quad \langle P_{U^{\perp}}(u), v \rangle = \langle u, P_{U^{\perp}}(v) \rangle$, para todos os $u, v \in V$;

(vi) $\|u\|^2 = \|P_U(u)\|^2 + \|P_{U^{\perp}}(u)\|^2$, para todo o $u \in V$ (Teorema de Pitágoras);

Observação 52. Seja U é um subespaço de dimensão finita de um espaço euclidiano (ou unitário) V. Seja $v \in V$.

(i) $\dim U + \dim U^{\perp} = \dim V$

(ii) Se $\{v_1, v_2, ..., v_n\}$ é uma base de U então $v \in U^{\perp}$ se e só se

$$\langle v, v_1 \rangle = \langle v, v_2 \rangle = ... = \langle v, v_n \rangle = 0.$$

Teorema 71. Seja U é um subespaço de dimensão finita de um espaço euclidiano (ou unitário) V. Seja $v \in V$. Então, existe um **elemento de U mais próximo de v** do que qualquer dos outros pontos de U. **Este elemento é a projecção ortogonal $P_U(v)$ de v sobre U** e tem-se

$$\|v - P_U(v)\| \leq \|v - u\|,$$

para todo o $u \in U$, e a igualdade verifica-se se e só se $u = P_U(v)$. Além disso, tendo-se $\mathbf{0} \in U$, a **distância d de um ponto $b \in V$ a um subespaço U** é dada por:

$$d(b, U) = \|P_{U^\perp}(b - \mathbf{0})\| = \|P_{U^\perp}(b)\| = \|b - P_U(b)\|.$$

Definição 52. Seja V um espaço euclidiano (ou unitário). Seja U é um subespaço de V com $\dim U = k$. Seja $q \in V$. Chama-se ao conjunto

$$\{q\} + U$$

um k-plano. A **distância d de um ponto $p \in V$ a um k-plano $\mathcal{P} = \{q\} + U$** é dada por:

$$d(p, \mathcal{P}) = \|P_{U^\perp}(p - q)\|.$$

Observação 53. A distância entre dois k-planos paralelos $\mathcal{P}_1 = \{a\} + U$ e $\mathcal{P}_2 = \{b\} + U$ é dada por:

$$d(\mathcal{P}_1, \mathcal{P}_2) = \|P_{U^\perp}(a - b)\|.$$

Exemplo 49. Considere-se $\mathbb{R}^3$ com o produto interno usual.

(i) Seja $\mathcal{P}$ o plano (em $\mathbb{R}^3$) que passa pelos pontos: $(1,2,1), (1,0,-1)$ e $(1,1,1)$. Tem-se

$$\mathcal{P} = \{(1,2,1)\} + L(\{(0,-2,-2), (0,-1,0)\})$$

Equação vectorial de $\mathcal{P}$:

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

com $\alpha, \beta \in \mathbb{R}$.

Equações paramétricas de $\mathcal{P}$:

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

com $\alpha, \beta \in \mathbb{R}$.

Equação cartesiana de $\mathcal{P}$:
$$x = 1.$$

Podemos determinar uma **equação cartesiana de** $\mathcal{P}$ do seguinte modo. Atendendo a que
$$\mathcal{P} = \{(1,2,1)\} + L\left(\{(0,-2,-2),(0,-1,0)\}\right)$$
seja
$$U = L\left(\{(0,-2,-2),(0,-1,0)\}\right).$$
Logo,
$$U^{\perp} = \left\{(x,y,z) \in \mathbb{R}^3 : \langle (x,y,z),(0,-2,-2)\rangle = 0 \;\text{ e }\; \langle (x,y,z),(0,-1,0)\rangle = 0\right\} =$$
$$= \mathcal{N}\left(\begin{bmatrix} 0 & -2 & -2 \\ 0 & -1 & 0 \end{bmatrix}\right) = L\left(\{(1,0,0)\}\right)$$
e assim, a equação cartesiana do plano $\mathcal{P}$ que passa pelo ponto $(1,2,1)$ é dada por:
$$\left(\langle (x-1,y-2,z-1),(1,0,0)\rangle = 0\right) \Leftrightarrow$$
$$\Leftrightarrow \left(1\,(x-1) + 0\,(y-2) + 0\,(z-1) = 0\right),$$
ou seja por
$$x = 1.$$

(ii) Determinemos a **equação cartesiana** da recta que passa pelos pontos $(1,1,0)$ e $(1,2,1)$. Tem-se
$$r = \{(1,1,0)\} + L\left(\{(0,1,1)\}\right),$$
uma vez que $(0,1,1) = (1,2,1) - (1,1,0)$. Seja
$$U = L\left(\{(0,1,1)\}\right).$$
Logo,
$$U^{\perp} = \left\{(x,y,z) \in \mathbb{R}^3 : \langle (x,y,z),(0,1,1)\rangle = 0\right\} =$$
$$= \mathcal{N}\left(\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}\right) = L\left(\{(1,0,0),(0,1,-1)\}\right)$$
e assim, a equação cartesiana da recta r é dada por:
$$\left(\langle (x-1,y-1,z),(1,0,0)\rangle = 0 \;\text{ e }\; \langle (x-1,y-1,z),(0,1,-1)\rangle = 0\right) \Leftrightarrow$$
$$\Leftrightarrow \left(1\,(x-1) = 0 \;\text{ e }\; 1\,(y-1) - 1z = 0\right),$$
ou seja por
$$\begin{cases} x = 1 \\ \\ y - z = 1. \end{cases}$$

Produto Externo e Produto Misto

Definição 53. Sejam $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3) \in \mathbb{R}^3$. Então o **produto externo** (vectorial) de u por v, denotado por $u \times v$, é o vector de $\mathbb{R}^3$ definido por

$$u \times v = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1),$$

isto é,

$$\begin{aligned} u \times v &= \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right) = \\ &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} e_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} e_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} e_3 = \\ &= \begin{vmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}, \end{aligned}$$

onde $\{e_1, e_2, e_3\}$ é a base canónica de $\mathbb{R}^3$.

Observação 54. Sejam $u, v, w \in \mathbb{R}^3$ e $\alpha \in \mathbb{R}$. Então, tem-se:

(i) $e_1 \times e_2 = e_3$

(ii) $e_2 \times e_3 = e_1$

(iii) $e_3 \times e_1 = e_2$

(iv) $u \times v = -(v \times u)$

(v) $u \times (v + w) = u \times v + u \times w$

(vi) $(u + v) \times w = u \times w + v \times w$

(vii) $\alpha (u \times v) = (\alpha u) \times v = u \times (\alpha v)$

(viii) $u \times \mathbf{0} = \mathbf{0} \times u = \mathbf{0}$

(ix) $u \times u = \mathbf{0}$

(x) Se u e v forem paralelos então $u \times v = \mathbf{0}$

(xi) $u \times (v \times w) = \langle u, w \rangle v - \langle u, v \rangle w$

(xii) $(u \times v) \times w = \langle w, u \rangle v - \langle w, v \rangle u$

(xiii) $\|u \times v\|^2 + \langle u, v \rangle^2 = \|u\|^2 \|v\|^2$

(xiv) $u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0$

Teorema 72. Considere-se $\mathbb{R}^3$ com o produto interno usual. Sejam $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ e seja $\theta \in [0, \pi]$ o ângulo formado por u e v. Então tem-se:

(i) $\|u \times v\| = \|u\| \|v\| \operatorname{sen} \theta$.

(ii) A área do paralelogramo de lados adjacentes u e v é dada por:

$$A = \|u \times v\|.$$

Dem. **(i)** Como $\theta \in [0, \pi]$, tem-se $\operatorname{sen} \theta = \sqrt{1 - \cos^2 \theta}$ e deste modo,

$$\|u\| \|v\| \operatorname{sen} \theta =$$

$$\begin{aligned}
&= \|u\| \|v\| \sqrt{1 - \cos^2 \theta} = \|u\| \|v\| \sqrt{1 - \frac{\langle u, v\rangle^2}{\|u\|^2 \|v\|^2}} = \sqrt{\|u\|^2 \|v\|^2 - \langle u, v\rangle^2} = \\
&= \sqrt{\left(u_1^2 + u_2^2 + u_3^2\right)\left(v_1^2 + v_2^2 + v_3^2\right) - \left(u_1 v_1 + u_2 v_2 + u_3 v_3\right)^2} = \\
&= \sqrt{\left(u_2 v_3 - u_3 v_2\right)^2 + \left(u_3 v_1 - u_1 v_3\right)^2 + \left(u_1 v_2 - u_2 v_1\right)^2} = \\
&= \|\left(u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1\right)\| = \\
&= \|u \times v\|.
\end{aligned}$$

(ii) $A = (\text{base})\ (\text{altura}) = \|u\| \|v\| \operatorname{sen} \theta$.

Definição 54. Considere-se $\mathbb{R}^3$ com o produto interno usual. Sejam $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3), w = (w_1, w_2, w_3) \in \mathbb{R}^3$. À expressão

$$\langle u, v \times w\rangle$$

chama-se **produto misto** de u, v e w.

Observação 55. Considere-se $\mathbb{R}^3$ com o produto interno usual. Sejam $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3), w = (w_1, w_2, w_3) \in \mathbb{R}^3$. Então, tem-se:

(i)

$$\langle u, v \times w\rangle = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

(ii) Sendo θ o ângulo formado por u e $v \times w$, o volume do paralelepípedo com um vértice e arestas u, v, w com origem em $(0, 0, 0)$, é dado por

$$V = \underbrace{\|v \times w\|}_{\text{área da face determinada por } v \text{ e } w} \quad \underbrace{\|u\| |\cos \theta|}_{\text{altura}} = \langle u, v \times w\rangle = \left| \det \left(\begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \right) \right|$$

(iii) $\langle u, u \times v\rangle = 0$

(iv) $\langle v, u \times v\rangle = 0$

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

Matrizes Hermiteanas, Matrizes Simétricas e Matrizes Normais. Diagonalização Unitária e Diagonalização Ortogonal

Definição 55. Seja $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{C})$. Denota-se por A^H a matriz $\overline{A}^T$, isto é, a transposta da matriz conjugada $\overline{A} = (\overline{a_{ij}})$, onde $\overline{a_{ij}}$ é o complexo conjugado de a_{ij}. Ou seja, escreve-se

$$A^H = \overline{A}^T.$$

A matriz A diz-se **hermiteana** se $A^H = A$.

Observação 56. **(a)** Sejam $\alpha, \beta \in \mathbb{C}$, $A, C \in \mathcal{M}_{m\times n}(\mathbb{C})$ e $B \in \mathcal{M}_{n\times r}(\mathbb{C})$. Tem-se:
(i) $\left(A^H\right)^H = A$ **(ii)** $(\alpha A + \beta B)^H = \overline{\alpha} A^H + \overline{\beta} B^H$ **(iii)** $(AC)^H = C^H A^H$

(b) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica ($A^T = A$). Então A é hermiteana uma vez que sendo A real tem-se $A^H = A^T$.

Teorema 73. Todos os valores próprios de uma matriz hermiteana são reais. Além disso, os vectores próprios associados a valores próprios distintos são ortogonais.

Dem. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ tal que A é hermiteana. Seja λ um valor próprio de A e seja u um vector próprio associado. Seja $\alpha = u^H A u$. Então, tem-se

$$\overline{\alpha} = \alpha^H = \left(u^H A u\right)^H = u^H A^H \left(u^H\right)^H \underset{A \text{ é hermiteana}}{=} u^H A u = \alpha.$$

Ou seja, α é real. Por outro lado, como

$$\alpha = u^H A u = u^H \lambda u = \lambda \left\|u\right\|^2,$$

tem-se $\lambda = \dfrac{\alpha}{\left\|u\right\|^2} \in \mathbb{R}$.

Sejam agora u_1 e u_2 vectores próprios associados respectivamente a valores próprios distintos λ_1 e λ_2. Então, tem-se

$$(Au_1)^H u_2 = u_1^H A^H u_2 \underset{A \text{ é hermiteana}}{=} u_1^H A u_2 = \lambda_2 u_1^H u_2$$

e

$$(Au_1)^H u_2 = \left(u_2^H A u_1\right)^H = \left(u_2^H \lambda_1 u_1\right)^H \underset{\lambda_1 \in \mathbb{R}}{=} \lambda_1 u_1^H u_2.$$

Logo, tem-se

$$\lambda_1 u_1^H u_2 = \lambda_2 u_1^H u_2 \Leftrightarrow (\lambda_1 - \lambda_2)\, u_1^H u_2 = 0.$$

E assim, como $\lambda_1 \neq \lambda_2$, então $\underbrace{u_1^H u_2}_{=\langle u_1, u_2\rangle} = 0$, ou seja, u_1 e u_2 são ortogonais.

Observação 57. Todos os valores próprios de uma matriz simétrica real são reais. Além disso, os vectores próprios associados a valores próprios distintos são ortogonais.

Definição 56. **(i)** Seja $U \in \mathcal{M}_{n\times n}(\mathbb{C})$. A matriz U diz-se **unitária** se se tiver $U^H U = I$, isto é, se $U^H = U^{-1}$, ou seja, se as colunas de U constituirem um conjunto ortonormado em $\mathbb{C}^n$.

(ii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz A diz-se **ortogonal** se se tiver $A^T A = I$, isto é, se $A^T = A^{-1}$, ou seja, se as colunas de A constituirem um conjunto ortonormado em $\mathbb{R}^n$.

Observação 58. **(i)** Seja $U \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que U é unitária. Como, sendo U real, se tem $U^H = U^T$, então U é ortogonal. Isto é, toda a matriz real unitária é ortogonal.

(ii) Seja A uma matriz hermiteana. Se os valores próprios de A forem distintos, então existe uma matriz unitária que diagonaliza A, isto é, existe U unitária tal que UAU^H é uma matriz diagonal, ou seja, A diz-se **unitariamente diagonalizável.**

(iii) A afirmação anterior (ii) continua válida mesmo se os valores próprios não forem distintos, como será provado no Teorema 78.

Teorema 74. Seja A uma matriz $n \times n$. Então, existe uma matriz unitária U tal que UAU^H é triangular superior.

Dem. A demonstração será efectuada por indução em n. O resultado é óbvio para $n = 1$. Suponhamos que a hipótese é válida para matrizes $k\times k$ e seja A uma matriz $(k+1)\times(k+1)$. Sejam λ_1 um valor próprio de A e w_1 um vector próprio associado de norma 1. Aplicando o método de ortogonalização de Gram-Schmidt, seja $\{w_1, \ldots, w_{k+1}\}$ uma base ortonormada para $\mathbb{C}^{k+1}$. Seja W^H a matriz cuja coluna i é igual ao vector w_i, para $i = 1, \ldots, k+1$. Então, por construção, a matriz W^H é unitária. Por outro lado, a primeira coluna de WAW^H é igual a WAw_1, tendo-se

$$WAw_1 = W\lambda_1 w_1 = \lambda_1 W w_1 = \lambda_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Logo, tem-se

$$WAW^H = \begin{bmatrix} \lambda_1 & | & * & \cdots & * \\ - & | & - & - & - \\ 0 & | & & & \\ \vdots & | & & M & \\ 0 & | & & & \end{bmatrix},$$

onde M é uma matriz $k \times k$.

Pela hipótese de indução, existe uma matriz $k \times k$ unitária V_1 tal que $V_1 M (V_1)^H = T_1$, onde T_1 é uma matriz triangular. Seja

$$V = \begin{bmatrix} 1 & | & 0 & \cdots & 0 \\ - & | & - & - & - \\ 0 & | & & & \\ \vdots & | & & V_1 & \\ 0 & | & & & \end{bmatrix}.$$

Então V é unitária e tem-se

$$(VW)\,A\,(VW)^H = VWAW^HV^H = \left[\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ - & - & - & - \\ 0 & & & \\ \vdots & & V_1M\,(V_1)^H & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ - & - & - & - \\ 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right] = T,$$

onde T é uma matriz triangular. Como a matriz VW é unitária, pondo $U = VW$, tem-se

$$UAU^H = T,$$

com T triangular e U unitária.

Teorema 75. Seja A uma matriz hermiteana. Então existe uma matriz unitária U que diagonaliza A, isto é, A é diagonalizável unitariamente. Ou seja, existe U unitária tal que tal que a matriz UAU^H é diagonal.

Dem. Pelo teorema anterior, existe uma matriz unitária U tal que a matriz UAU^H é triangular. Seja $T = UAU^H$. Tem-se então

$$T^H = \left(UAU^H\right)^H = \left(U^H\right)^H A^H U^H \underset{A \text{ é hermiteana}}{=} UAU^H = T.$$

Logo, como $T = T^H$ e T é triangular então T é diagonal.

Observação 59. Atendendo ao resultado anterior tem-se então o seguinte. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica, isto é, tal que $A = A^T$. Então existe uma matriz ortogonal P (matriz que verifica: $P^T = P^{-1}$) tal que PAP^T é diagonal, isto é, A é **ortogonalmente diagonalizável** relativamente a uma base ortonormada formada só por vectores próprios de A. A matriz P^T é a matriz cujas colunas são os vectores próprios de A que formam uma base ortonormada formada de $\mathbb{R}^n$, sendo PAP^T a matriz diagonal onde se coloca na entrada i da diagonal principal o valor próprio correspondente ao vector próprio da coluna i da matriz P^T.

Observação 60. **(i)** Existem matrizes não hermiteanas que são diagonalizáveis relativamente a bases ortonormadas formadas só por vectores próprios, como por exemplo as matrizes anti-hermiteanas ($A^H = -A$) e as matrizes anti-simétricas ($A^T = -A$).

(ii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Suponhamos que A é ortogonalmente diagonalizável relativamente a uma base ortonormada formada só por vectores próprios de A. Sejam D diagonal e P ortogonal tais que $A = P^TDP$. Então

$$A^T = \left(P^TDP\right)^T = P^TD^T\left(P^T\right)^T = P^TDP = A.$$

Logo A é simétrica. Tem-se então, atendendo tambem à observação 59, sendo A real:

$$A \text{ é ortogonalmente diagonalizável} \quad \Leftrightarrow \quad A \text{ é simétrica}$$

(iii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$. Suponhamos que A é **unitariamente diagonalizável** relativamente a uma base ortonormada formada só por vectores próprios de A. Sejam D diagonal e U unitária tais que $A = U^H DU$. Como em geral se tem $D^H \neq D$, então

$$A^H = \left(U^H DU\right)^H = U^H D^H U \neq U^H DU = A.$$

Logo A não tem que ser necessariamente hermiteana.

(iv) O próximo teorema diz quais são as matrizes **unitariamente diagonalizáveis.**

Definição 57. Uma matriz A diz-se **normal** se

$$AA^H = A^H A.$$

Teorema 76. Uma matriz A é normal se e só se fôr unitariamente diagonalizável relativamente a uma base ortonormada formada só por vectores próprios de A.

Dem. $(\Rightarrow)$ Suponhamos que A é normal. Pelo teorema 74, existe uma matriz unitária U e uma matriz triangular superior T tais que $T = UAU^H$. Vejamos que T é normal. Tem-se

$$\begin{aligned} T^H T &= \left(UAU^H\right)^H UAU^H = UA^H U^H UAU^H = UA^H AU^H \underset{A \text{ é normal}}{=} \\ &= UAA^H U^H = UAU^H UA^H U^H = TT^H. \end{aligned}$$

Logo T é normal. Seja $T = (t_{ij})$ do tipo $n \times n$. Comparando as entradas das diagonais principais de $T^H T$ e TT^H tem-se:

$$\begin{aligned} |t_{11}|^2 + |t_{12}|^2 + |t_{13}|^2 + \cdots + |t_{1n}|^2 &= |t_{11}|^2 \\ |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2 &= |t_{12}|^2 + |t_{22}|^2 \\ &\vdots \\ |t_{nn}|^2 &= |t_{1n}|^2 + |t_{2n}|^2 + |t_{3n}|^2 + \cdots + |t_{nn}|^2 \end{aligned}$$

e assim, $t_{ij} = 0$ sempre que $i \neq j$. Logo T é diagonal e portanto A é unitariamente diagonalizável.

$(\Leftarrow)$ Suponhamos agora que A é unitariamente diagonalizável. Queremos mostrar que A é normal. Sejam D diagonal e U unitária tais que $D = UAU^H$, ou seja, $A = U^H DU$. Tem-se

$$AA^H = U^H DU \left(U^H DU\right)^H = U^H DUU^H D^H U = U^H \left(DD^H\right) U$$

e

$$A^H A = \left(U^H DU\right)^H U^H DU = U^H D^H UU^H DU = U^H \left(D^H D\right) U.$$

Como

$$DD^H = D^H D = \begin{bmatrix} |\lambda_1|^2 & 0 & \cdots & 0 \\ 0 & |\lambda_2|^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & |\lambda_n|^2 \end{bmatrix},$$

então tem-se $AA^H = A^H A$ e assim A é normal.

Formas quadráticas

Definição 58. Uma **equação quadrática** em duas variáveis x_1 e x_2 é uma equação da forma

$$ax_1^2 + bx_2^2 + 2cx_1x_2 + dx_1 + ex_2 + f = 0$$

a qual pode ser escrita na forma

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & c \\ c & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + f = 0.$$

Sejam

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \text{e} \qquad A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}.$$

(A é uma matriz real simétrica). À função real a duas variáveis reais $Q : \mathbb{R}^2 \to \mathbb{R}$ definida por $Q(x) = x^T Ax$, com

$$x^T Ax = ax_1^2 + bx_2^2 + 2cx_1x_2$$

chama-se **forma quadrática** real a 2 variáveis reais associada à equação quadrática anterior.

Podem haver equações do 2^o grau e formas quadráticas com um n^o de variáveis superior a 2. Uma equação quadrática em n variáveis $x_1, x_2, \ldots, x_n$ é uma equação da forma

$$x^T Ax + Bx + \alpha = 0,$$

onde $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $A = (a_{ij})$ é uma matriz real simétrica do tipo $n \times n$, $B \in \mathcal{M}_{1\times n}(\mathbb{R})$ e α é um escalar. À função real a n variáveis reais $Q : \mathbb{R}^n \to \mathbb{R}$ definida por

$$Q(x) = x^T Ax = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}x_j \right) x_i$$

chama-se **forma quadrática** real a n variáveis reais associada à equação quadrática anterior.

Teorema 77. (Teorema dos eixos principais). Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica. Então existe uma mudança de variáveis ortogonal (teorema 75 e observação 59) que transforma a forma quadrática $x^T Ax$ na forma quadrática $y^T Dy$ sem termos cruzados. Isto é, se P diagonalizar A ortogonalmente ($D = PAP^T$), então a mudança de variáveis $x = P^T y$ transforma a forma quadrática $x^T Ax$ na forma quadrática $y^T Dy$:

$$x^T Ax = y^T PAP^T y = y^T Dy = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 =$$

$$= \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_n$ são os valores próprios de A associados respectivamente aos vectores próprios que constituem as colunas de P^T e que formam uma base ortonormada de $\mathbb{R}^n$.

Observação 61. **(i)** Chama-se **cónica** ou secção cónica à curva plana obtida por meio de um corte efectuado por um plano relativamente a uma superfície cónica. As secções cónicas que se obtêm quando o plano que efectua o corte não passa pelo vértice da superfície cónica, são elipses (os valores próprios têm o mesmo sinal) (podendo ter-se circunferências: quando o corte é efectuado perpendicularmente ao eixo de simetria do cone), parábolas (um dos dois valores próprios é zero) e hipérboles (os dois valores próprios têm sinais contrários).

(ii) Em $\mathbb{R}^3$ tem-se

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad A = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} g \\ h \\ i \end{bmatrix}$$

e

$$ax_1^2 + bx_2^2 + cx_3^2 + 2dx_1x_2 + 2ex_1x_3 + 2fx_2x_3 + gx_1 + hx_2 + ix_3 + \alpha = 0.$$

À superficie resultante da equação anterior chama-se **quádrica**. Existem quatro tipos de quádricas não degeneradas): elipsóides, hiperbolóides (de uma ou duas folhas), cones e parabolóides (elípticos ou hiperbólicos).

Exemplo 50. Considere-se a forma quadrática $Q : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$Q(x, y) = 3x^2 + 4xy + 3y^2.$$

Tem-se

$$Q(x, y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix},$$

com $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$. Os valores próprios de A são $\lambda_1 = 1$ e $\lambda_2 = 5$. Tem-se então a seguinte forma quadrática diagonal (isto é, sem termos cruzados)

$$Q(x', y') = \begin{bmatrix} x' & y' \end{bmatrix} D \begin{bmatrix} x' \\ y' \end{bmatrix} =$$

$$= \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

com

$$D = PAP^T, \qquad \begin{bmatrix} x' \\ y' \end{bmatrix} = P \begin{bmatrix} x \\ y \end{bmatrix},$$

e

$$P^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & \operatorname{sen}\frac{\pi}{4} \\ -\operatorname{sen}\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix},$$

em que P^T é a matriz diagonalizante obtida colocando na 1^a coluna um vector próprio de norma 1 associado ao valor próprio λ_1 e na 2^a coluna um vector próprio de norma 1 associado ao valor próprio λ_2, de tal modo que ambos os vectores próprios constituam uma base ortonormada de $\mathbb{R}^2$. Observe-se que a matriz P é ortogonal, isto é, tem-se $P^T = P^{-1}$.

Tem-se então

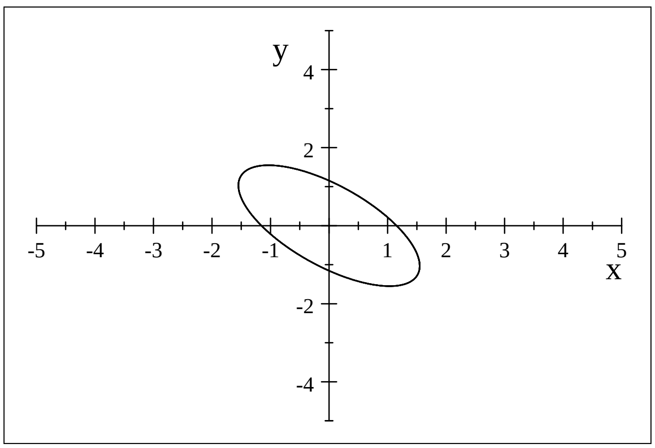
$$Q(x,y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} =$$

$$\begin{bmatrix} x & y \end{bmatrix} P^T D P \begin{bmatrix} x \\ y \end{bmatrix} =$$

$$= \left(P \begin{bmatrix} x \\ y \end{bmatrix} \right)^T D P \begin{bmatrix} x \\ y \end{bmatrix} =$$

$$= \begin{bmatrix} x' & y' \end{bmatrix} D \begin{bmatrix} x' \\ y' \end{bmatrix} = Q(x', y').$$

Por exemplo, relativamente à equação quadrática $3x^2 + 4xy + 3y^2 = 4$



tem-se a elipse:

$$(x')^2 + 5\,(y')^2 = 4 \Leftrightarrow \left(\frac{x'}{2}\right)^2 + \left(\frac{y'}{\frac{2\sqrt{5}}{5}}\right)^2 = 1.$$

Definição 59. Seja A uma matriz real simétrica do tipo $n \times n$. Diz-se que A e a forma quadrática $Q : \mathbb{R}^n \to \mathbb{R}$ dada por $Q(x) = x^T A x$ são:

(i) definidas positivas se $x^T A x > 0$, para todo o $x \neq \mathbf{0}$;

(ii) definidas negativas se $x^T A x < 0$, para todo o $x \neq \mathbf{0}$;

(iii) semidefinidas positivas se $x^T A x \geq 0$, para todo o x;

(iv) semidefinidas negativas se $x^T A x \leq 0$, para todo o x;

(v) indefinidas se existirem pontos onde $x^T A x$ seja positiva e pontos onde $x^T A x$ seja negativa.

Teorema 78. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica. Então,

(i) A é **definida positiva** se e só se todos os valores próprios de A forem positivos;

(ii) A é **definida negativa** se e só se todos os valores próprios de A forem negativos;

(iii) A é **semidefinida positiva** se e só se todos os valores próprios de A forem não negativos;

(iv) A é **semidefinida negativa** se e só se todos os valores próprios de A forem não positivos;

(v) A é **indefinida** se e só se A tiver pelo menos um valor próprio positivo e outro negativo.

Dem. (i) ($\Rightarrow$) Suponhamos que A é definida positiva, isto é,

$$x^T A x > 0,$$

para todo o $x \neq \mathbf{0}$. Seja λ um valor próprio de A. Então, para qualquer vector próprio x associado a λ tem-se

$$x^T A x = \lambda x^T x = \lambda \left\| x \right\|^2,$$

com $x \neq \mathbf{0}$. Logo

$$\lambda = \frac{x^T A x}{\left\| x \right\|^2} > 0.$$

($\Leftarrow$) Suponhamos que todos os valores próprios de A são positivos. Seja $\{x_1, \ldots, x_n\}$ um conjunto ortonormado de vectores próprios de A (existe esse conjunto pelo teorema 75 e observação 59). Logo $\{x_1, \ldots, x_n\}$ é uma base ortonormada de $\mathbb{R}^n$. Se $x \neq \mathbf{0}$, então tem-se

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n,$$

com $\alpha_1, ..., \alpha_n$ não todos nulos, pelo que

$$x^T A x = (\alpha_1 x_1 + \cdots + \alpha_n x_n)^T A (\alpha_1 x_1 + \cdots + \alpha_n x_n) =$$

$$= \left(\alpha_1 (x_1)^T + \cdots + \alpha_n (x_n)^T\right)(\alpha_1 A x_1 + \cdots + \alpha_n A x_n) =$$

$$= \left(\alpha_1 (x_1)^T + \cdots + \alpha_n (x_n)^T\right)(\alpha_1 \lambda_1 x_1 + \cdots + \alpha_n \lambda_n x_n) =$$

$$= \sum_{i=1}^{n} (\alpha_i)^2 \lambda_i > 0.$$

Logo A é definida positiva.

Mínimos Quadrados

Existem aplicações relativamente às quais os erros cometidos nas medições das entradas de A ou de b podem levar a que o sistema de equações lineares $Au = b$ não tenha solução, quando teoricamente deveria ter. Em tais casos é natural a procura da "melhor solução aproximada" para esse problema.

Definição 60. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^n$. Então, a $\widehat{u} \in \mathbb{R}^n$ chama-se melhor solução aproximada ou **solução de mínimos quadrados** de $Au = b$ se

$$\|b - A\widehat{u}\| \leq \|b - Au\|,$$

para qualquer $u \in \mathbb{R}^n$. Ao vector $b - A\widehat{u}$ chama-se vector erro de mínimos quadrados e ao escalar $\|b - A\widehat{u}\|$ chama-se erro de mínimos quadrados.

Observação 62. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^n$. Procuremos então um método para determinar as soluções de mínimos quadrados de $Au = b$. Atendendo a que $Au \in \mathcal{C}(A)$ para todo o $u \in \mathbb{R}^n$, então a distância $\|b - Au\|$ é mínima se

$$Au = P_{\mathcal{C}(A)}(b),$$

onde $P_{\mathcal{C}(A)}$ é a projecção ortogonal de $\mathbb{R}^n$ sobre $\mathcal{C}(A)$. Como $P_{\mathcal{C}(A)}(b) \in \mathcal{C}(A)$, a equação $Au = P_{\mathcal{C}(A)}(b)$ tem sempre solução e essas soluções são as soluções de mínimos quadrados de $Au = b$. Deste modo, qualquer sistema linear tem sempre pelo menos uma solução de mínimos quadrados.

Por outro lado, pode escrever-se a equação $Au = P_{\mathcal{C}(A)}(b)$ na forma

$$b - Au = b - P_{\mathcal{C}(A)}(b) = P_{\mathcal{N}(A^T)}(b)$$

tendo-se

$$A^T(b - Au) = A^T\left(b - P_{\mathcal{C}(A)}(b)\right) = A^T\left(P_{\mathcal{N}(A^T)}(b)\right) = \mathbf{0},$$

pois $(\mathcal{C}(A))^\perp = \mathcal{N}\left(A^T\right)$. Logo

$$A^TAu = A^Tb.$$

A esta equação chama-se equação normal associada a $Au = b$.

Teorema 79. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^n$.

(i) As soluções de mínimos quadrados do sistema linear

$$Au = b$$

são as soluções da equação normal

$$A^TAu = A^Tb.$$

(ii) Se $\text{car}\, A = n$ então a equação normal

$$A^T A u = A^T b$$

tem a solução única

$$\widehat{u} = \left(A^T A\right)^{-1} A^T b$$

e tem-se

$$P_{\mathcal{C}(A)}\left(b\right) = A\widehat{u} = A\left(A^T A\right)^{-1} A^T b,$$

isto é, $A\left(A^T A\right)^{-1} A^T$ é a matriz que representa a projecção ortogonal $P_{\mathcal{C}(A)}$.

Observação 63. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Vejamos que se tem

$$\text{car}\, A = \text{car}\left(A^T A\right).$$

Basta para isso, mostrar que

$$\mathcal{N}\left(A\right) = \mathcal{N}\left(A^T A\right).$$

Seja $u \in \mathcal{N}\left(A\right)$. Como $Au = \mathbf{0}$ então $A^T A u = A^T \mathbf{0} = \mathbf{0}$ e assim $u \in \mathcal{N}\left(A^T A\right)$.

Reciprocamente, seja $u \in \mathcal{N}\left(A^T A\right)$ e vejamos que $u \in \mathcal{N}\left(A\right)$. Tem-se $A^T A u = \mathbf{0}$, logo

$$Au \in \mathcal{N}\left(A^T\right) = \left(\mathcal{L}\left(A^T\right)\right)^{\perp} = \left(\mathcal{C}\left(A\right)\right)^{\perp}$$

e como tal

$$\langle Au, Au\rangle = 0,$$

ou seja $\|Au\|^2 = 0$ e então $Au = \mathbf{0}$, isto é, $u \in \mathcal{N}\left(A\right)$.

Observação 64. Vejamos agora o modo como se pode determinar uma curva (ou recta) específica que se possa "ajustar" a um conjunto de pontos determinados experimentalmente.

(i) A partir de dois ou mais pontos dados $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, pretende-se determinar uma recta $y = a_0 + a_1 x$ que seja a recta que "melhor aproxime" ou a recta de mínimos quadrados de melhor ajuste aos pontos dados (recta de regressão). Isto é, pretende-se determinar as soluções de mínimos quadrados de

$$\begin{cases} y_1 = a_0 + a_1 x_1 \\ y_2 = a_0 + a_1 x_2 \\ \quad\vdots \\ y_m = a_0 + a_1 x_m \end{cases}$$

ou seja de

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Sejam

$$A=\begin{bmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_m\end{bmatrix},\qquad u=\begin{bmatrix}a_0\\ a_1\end{bmatrix}\qquad \text{e}\qquad b=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix}.$$

Atendendo a que car $A=$ car $\left(A^TA\right)$, se houver pelo menos dois pontos distintos, tem-se car $A=2$ e nesse caso, a equação normal

$$A^TAu=A^Tb$$

tem como única solução de mínimos quadrados

$$\widehat{u}=\left(A^TA\right)^{-1}A^Tb.$$

Assim, a recta de mínimos quadrados $y=a+bx$ é a recta que torna mínimos os quadrados cuja soma

$$(y_1-(a_0+a_1x_1))^2+(y_2-(a_0+a_1x_2))^2+\cdots+(y_m-(a_0+a_1x_m))^2$$

é dada por

$$\|b-A\widehat{u}\|^2,$$

onde $\|b-A\widehat{u}\|$ é o erro de mínimos quadrados.

(ii) A partir de m pontos dados $(x_1,y_1),(x_2,y_2),\ldots,(x_m,y_m)$, pretende-se determinar um polinómio cujo gráfico esteja tão perto quanto possível desses m pontos dados. Isto é, com $m\in\mathbb{N}$ previamente fixo, pretende-se determinar as soluções de mínimos quadrados do sistema de m equações a $n+1$ incógnitas $(a_0,a_1,a_2,\ldots,a_n)$

$$\begin{cases} y_1=a_0+a_1x_1+a_2x_1^2+\cdots+a_nx_1^n\\ y_2=a_0+a_1x_2+a_2x_2^2+\cdots+a_nx_2^n\\ \vdots\\ y_m=a_0+a_1x_m+a_2x_m^2+\cdots+a_nx_m^n\end{cases}$$

ou seja de

$$\begin{bmatrix}1 & x_1 & x_1^2 & \cdots & x_1^n\\ 1 & x_2 & x_2^2 & \cdots & x_2^n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_m & x_m^2 & \cdots & x_m^n\end{bmatrix}\begin{bmatrix}a_0\\ a_1\\ \vdots\\ a_n\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix}.$$

Sejam

$$A=\begin{bmatrix}1 & x_1 & x_1^2 & \cdots & x_1^n\\ 1 & x_2 & x_2^2 & \cdots & x_2^n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_m & x_m^2 & \cdots & x_m^n\end{bmatrix},\qquad u=\begin{bmatrix}a_0\\ a_1\\ \vdots\\ a_n\end{bmatrix}\qquad \text{e}\qquad b=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_m\end{bmatrix}.$$

Note-se que se $n+1=m$ e se os pontos dados forem distintos, então existe um único polinómio de grau n (o chamado polinómio interpolador) que passa por todos esses m pontos.

Por outro lado, atendendo a que $\operatorname{car} A = \operatorname{car}\left(A^T A\right)$, se $n < m$ e pelo menos $n+1$ pontos forem distintos, tem-se $\operatorname{car} A = n+1$ e então a equação normal

$$A^T A u = A^T b$$

tem como única solução de mínimos quadrados

$$\widehat{u} = \left(A^T A\right)^{-1} A^T b.$$

Exemplo 51. Determinemos a recta de mínimos quadrados relativa aos pontos

$$(0,1), (1,3), (2,4) \text{ e } (3,4).$$

Sejam

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad \text{e} \qquad b = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix}.$$

Tem-se $\operatorname{car} A = 2$ e como tal a solução de mínimos quadrados é única e dada por:

$$\widehat{u} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \left(A^T A\right)^{-1} A^T b =$$

$$= \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix},$$

tendo-se

$$y = \frac{3}{2} + x.$$

O vector $b - A\widehat{u}$ é o vector erro de mínimos quadrados, sendo o erro de mínimos quadrados dado por:

$$\|b - A\widehat{u}\| =$$

$$= \sqrt{(y_1 - (a_0 + a_1 x_1))^2 + (y_2 - (a_0 + a_1 x_2))^2 + (y_3 - (a_0 + a_1 x_3))^2 + (y_4 - (a_0 + a_1 x_4))^2} =$$

$$= \sqrt{\left(1 - \left(\frac{3}{2} + 0\right)\right)^2 + \left(3 - \left(\frac{3}{2} + 1\right)\right)^2 + \left(4 - \left(\frac{3}{2} + 2\right)\right)^2 + \left(4 - \left(\frac{3}{2} + 3\right)\right)^2} =$$

$$= \sqrt{\left(1 - \left(\frac{3}{2} + 0\right)\right)^2 + \left(3 - \left(\frac{3}{2} + 1\right)\right)^2 + \left(4 - \left(\frac{3}{2} + 2\right)\right)^2 + \left(4 - \left(\frac{3}{2} + 3\right)\right)^2} =$$

$$= \sqrt{\frac{25}{16} + \frac{1}{4} + \frac{1}{16} + \frac{1}{4}} = \frac{\sqrt{34}}{4}.$$

Bibliografia

1. Seymour Lipschutz, Linear Algebra, Schaum´s Outline Series, fourth edition, McGraw-Hill, 2009.

2. Luis T. Magalhães, Álgebra Linear como Introdução à Matemática Aplicada, 9^a edição, Texto Editora, 2001.

3. Gilbert Strang, Introduction to Linear Algebra, 3rd edition, Wellesley Cambridge Pr. 2003.

4. Steven J. Leon, Linear Algebra with Applications, 6th edition, Prentice Hall, 2006.

5. Bernard Kolman, Introductory Linear Algebra with Applications, Prentice Hall, 1996.

6. Howard Anton and Robert C. Busby, Contemporary Linear Algebra, John Wiley & Sons, Inc., 2002.

7. Howard Anton and Chris Rorres, Elementary Linear Algebra with Applications, 9th edition, John Wiley & Sons, Inc., 2005.

8. António Monteiro e Gonçalo Pinto, Álgebra Linear e Geometria Analítica, McGraw-Hill, 1997.