Álgebra

Mestrado Integrado em Engenharia Informática e Computação

Ana Maria Mendonça
Maria Cristina Ribeiro

FEUP

2007/08

Conteúdo

Lista de Figuras

Capítulo 1

Matrizes

1.1 Conceitos gerais

A Álgebra Linear é muitas vezes referida como o primeiro curso em Matemática onde são introduzidos alguns conceitos matemáticos muito abstractos; um desses conceitos fundamentais é o conceito de **matriz**.

Definição 1.1 - *Matriz de tipo m×n sobre um corpo*

Uma matriz $A = [A_{ij}], (i = 1, 2, ..., m; j = 1, 2, ..., n)$ *de tipo* $m \times n$ *sobre um corpo* Ω *é um quadro com m linhas e n colunas cujos elementos* A_{ij} *são escalares de* Ω.

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

A - matriz; A_{ij} - elemento ou entrada da matriz.

Exemplo 1.1 A matriz $A = \begin{bmatrix} 1.5 & 7 & \sqrt{2} \\ 0 & 1 & 0.3 \end{bmatrix}$ é uma matriz 2×3 sobre o corpo **R** dos números reais; é uma **matriz real**.$\diamondsuit$

Exemplo 1.2 A matriz $B = \begin{bmatrix} 2+j & 7-j \\ -j & 2+2j \\ 0 & 1 \end{bmatrix}$ é uma matriz 3×2 sobre o corpo **C** dos números complexos; é uma **matriz complexa**.$\diamondsuit$

Os elementos de uma matriz não necessitam obrigatoriamente de ser valores numéricos. Por exemplo, se considerarmos o conjunto das funções $\mathbf{F}_{[0,1]}$, funções reais de variável real de domínio $[0, 1]$, é possível definir matrizes cujos elementos pertencem a $\mathbf{F}_{[0,1]}$; estas matrizes são matrizes de funções ou matrizes funcionais. Assim, a matriz $\begin{bmatrix} x^2 & x+1 \\ sen\,(x) & \sqrt{x} \end{bmatrix}$ é uma matriz 2×2 sobre $\mathbf{F}_{[0,1]}$. Contudo, não se trata de uma matriz sobre um corpo, mas sim sobre um anel comutativo com elemento unidade.

Nomenclatura e notações

As matrizes serão doravante denotadas por letras maiúsculas $(A, B, ...)$ e os respectivos elementos pelo identificador da matriz afectado de dois índices: o índice de linha e o índice de coluna. Logo, A_{ij} designa o elemento da linha i e coluna j da matriz A.

A linha i da matriz A representa-se por

$$A_{(i)} = \begin{bmatrix} A_{i1} & A_{i2} & \cdots & A_{in} \end{bmatrix}$$

e a coluna j representa-se por

$$A^{(j)} = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix}$$

Assim, temos:

$$A = [A_{ij}] = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} = \begin{bmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(m)} \end{bmatrix} = \begin{bmatrix} A^{(1)} & A^{(2)} & \cdots & A^{(n)} \end{bmatrix}$$

O tamanho de uma matriz é determinado pelo respectivo número de linhas e número de colunas. Para designar o tamanho de uma matriz são vulgarmente utilizados os termos **dimensão** ou **ordem da matriz**. Duas matrizes têm dimensão igual quando possuem o mesmo número de linhas e o mesmo número de colunas.

Exemplo 1.3 As matrizes $\begin{bmatrix} 1 & 2 & 4 \\ -3 & 5 & 6 \end{bmatrix}$ e $\begin{bmatrix} 2 & -1 & 5 \\ 6 & 0 & 1 \end{bmatrix}$ têm a mesma dimensão, mas as matrizes $\begin{bmatrix} 1 & 2 & 4 \\ -3 & 5 & 6 \end{bmatrix}$ e $\begin{bmatrix} 2 & 6 \\ -1 & 0 \\ 5 & 1 \end{bmatrix}$ são de dimensão diferente.$\diamondsuit$

Em muitas aplicações, o conceito de **matriz quadrada** é muito importante. Uma matriz quadrada de ordem n é uma matriz de tipo $n \times n$, isto é, com número idêntico de linhas e colunas.

Exemplo 1.4 A matriz $A = \begin{bmatrix} 1.5 & 3 & 2 \\ 0 & 5 & 7 \\ 2.5 & 1 & 2.4 \end{bmatrix}$ é uma matriz quadrada de ordem 3 sobre o corpo dos números reais $\mathbf{R}$.$\diamondsuit$

Quando uma matriz não é quadrada, ou seja, quando tem um número de linhas diferente do número de colunas, diz-se **rectangular**.

Uma matriz de tipo $1 \times n$ é uma **matriz-linha**; uma matriz de tipo $m \times 1$ é uma **matriz-coluna**.

Em matrizes do mesmo tipo, **elementos homólogos** são os que têm índices iguais.

Em matrizes quadradas, os elementos que têm os dois índices iguais são designados **elementos principais** da matriz. Os elementos principais de uma matriz quadrada constituem a respectiva **diagonal principal**. Chama-se **traço de uma matriz quadrada de ordem n** à soma dos seus elementos principais, e denota-se vulgarmente por $\mathbf{Tr}(A)$ = traço da matriz A.

1.2 Igualdade de matrizes

Definição 1.2 *Duas matrizes $A = [A_{ij}]$ e $B = [B_{ij}]$ são iguais se e só se todos os elementos homólogos forem iguais.*

$$A_{ij} = B_{ij} \ (i = 1...m; \ j = 1...n)$$

Duas matrizes com dimensões diferentes nunca são iguais.

1.3 Adição de matrizes

Designemos por $M_{m\times n}$ o conjunto das matrizes de ordem $m \times n$.

Definição 1.3 *A soma de duas matrizes* $A = [A_{ij}]$ *e* $B = [B_{ij}] \in M_{m\times n}$ *é uma matriz* $C = [C_{ij}]$ *cujos elementos são iguais à soma dos elementos homólogos de* A *e* B*.*

$$C_{ij} = A_{ij} + B_{ij} \ (i = 1...m;\ j = 1...n)$$

Exemplo 1.5 Se A e B forem as matrizes

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 4 \\ -1 & 5 \end{bmatrix}$$

então, a sua soma é a matriz C

$$C = A + B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 4 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 6 \\ 2 & 9 \end{bmatrix} . \diamondsuit$$

Propriedades da adição de matrizes

1. $A, B \in M_{m\times n} \Rightarrow A + B \in M_{m\times n}$
2. $(A + B) + C = A + (B + C)$
3. $A + B = B + A$
4. $A + O = O + A = A$
5. $A + (-A) = (-A) + A = O$
6. $A = B$ e $C = D \Rightarrow A + C = B + D$

Notas sobre a adição de matrizes:

a) O elemento neutro da adição é a matriz nula (de ordem $m \times n$);

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

b) $-A = [-A_{ij}]$ é a matriz simétrica de $A = [A_{ij}]$;
c) $A + (-B) = A - B$ (subtracção de matrizes);
d) $\mathbf{Tr}(A + B) = \mathbf{Tr}(A) + \mathbf{Tr}(B)$.

1.4 Multiplicação por escalares

Definição 1.4 - *Multiplicação de uma matriz por um escalar*
O produto de uma matriz $A = [A_{ij}] \in M_{m\times n}$ por um escalar $\lambda \in \Omega$ é uma matriz cujos elementos são iguais ao produto do escalar por cada elemento da matriz A.

$$\lambda.A = [\lambda.A_{ij}] \qquad (i = 1...m; j = 1...n)$$

Exemplo 1.6 Se $\lambda = 2$ e $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ então $\lambda.A = 2.\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}.\diamondsuit$

Propriedades da multiplicação por escalares

1. $A \in M_{m\times n}, \lambda \in \Omega \Rightarrow \lambda.A \in M_{m\times n}$
2. $(\lambda.\mu).A = \lambda.(\mu.A)$
3. $1.A = A$
4. $\lambda.(A + B) = \lambda.A + \lambda.B$
5. $(\lambda + \mu).A = \lambda.A + \mu.A$
6. $A = B \Rightarrow \lambda.A = \lambda.B$
7. $\mathbf{Tr}(\lambda.A) = \lambda.\mathbf{Tr}(A)$;

Como será posteriormente referido, o conjunto $M_{m\times n}$, das matrizes de ordem $m \times n$ sobre o corpo Ω, com as operações antes definidas, adição e multiplicação por um escalar, tem estrutura de espaço vectorial sobre o corpo Ω.

1.5 Multiplicação de matrizes

Definição 1.5 - *Multiplicação de matrizes*

Considerem-se duas matrizes A e B sobre o mesmo corpo tais que o número de colunas de A é igual ao número de linhas de B, isto é, a matriz A é de ordem m×n e a matriz B é de ordem n×p.

O produto das matrizes A e B é uma matriz P=A.B de ordem m×p onde

$$P_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} \quad (i = 1...m; \quad j = 1...p)$$

A lei de formação da matriz P a partir das duas matrizes A e B é conhecida por **multiplicação de linhas por colunas**, e o elemento genérico da matriz P, P_{ij}, é a soma dos produtos que se obtêm multiplicando os elementos A_{ik} (da linha i da matriz A) pelos elementos B_{kj} correspondentes (da coluna j da matriz B).

$$\begin{bmatrix} & & \vdots & & \\ A_{i1} & \cdots & A_{ik} & \cdots & A_{in} \\ & & \vdots & & \end{bmatrix} . \begin{bmatrix} & B_{1j} & \\ & \vdots & \\ \cdots & B_{kj} & \cdots \\ & \vdots & \\ & B_{nj} & \end{bmatrix} = \begin{bmatrix} & \vdots & \\ ... & P_{ij} & ... \\ & \vdots & \end{bmatrix}$$

Exemplo 1.7

$$\underbrace{\underbrace{A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}}_{2\times2} \; \underbrace{B = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 5 & 0 \end{bmatrix}}_{2\times3}}_{matriz \quad produto \quad \rightarrow 2\times3}$$

$$P = A.B = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} . \begin{bmatrix} 2 & 1 & 1 \\ 3 & 5 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 1\times2+3\times3 & 1\times1+3\times5 & 1\times1+3\times0 \\ 2\times2+4\times3 & 2\times1+4\times5 & 2\times1+4\times0 \end{bmatrix} = \begin{bmatrix} 11 & 16 & 1 \\ 16 & 22 & 2 \end{bmatrix} . \diamondsuit$$

Propriedades da multiplicação de matrizes

Seja M_n o conjunto das matrizes quadradas de ordem n sobre um corpo Ω. A multiplicação de matrizes verifica as seguintes propriedades:

1. $A, B \in M_n$ então $A.B \in M_n$
2. $(A.B).C = A.(B.C)$
3. Existe uma matriz $I \in M_n$ tal que qualquer que seja $A \in M_n$: $A.I = I.A = A$
4. $(A + B).C = A.C + B.C$ (distributividade à direita em relação à adição de matrizes)
5. $A.(B + C) = A.B + A.C$ (distributividade à esquerda em relação à adição de matrizes)
6. $A, B, C, D \in M_n$, se $A = B$ e $C = D$ então $A.C = B.D$ (compatibilidade)

Notas sobre a multiplicação de matrizes:

a) O produto de matrizes rectangulares, A e B, só é determinado quando o número de colunas de A é igual ao número de linhas de B. Pode reter-se a mnemónica

$$(m \times n).(n \times p) = m \times p$$

b) As propriedades de associatividade, distributividade à direita e à esquerda e compatibilidade com a igualdade de matrizes são verdadeiras para matrizes rectangulares desde que os produtos sejam determinados.

c) De um modo geral, **a multiplicação de matrizes não é comutativa**.

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

Quando $A.B = B.A$, diz-se que A e B são **matrizes permutáveis**.

d) O elemento neutro da multiplicação de matrizes quadradas de ordem n chama-se **matriz identidade** (de ordem n) e representa-se por

$$I_n = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}}_{\text{n colunas}} \left.\vphantom{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}\right\} \begin{array}{c}\text{n}\\ \text{linhas}\end{array}$$

A matriz identidade é sempre permutável com qualquer outra matriz; o mesmo acontece com qualquer **matriz escalar**, isto é, uma matriz cujos elementos da diagonal principal são todos iguais e os restantes elementos são nulos.

$$\begin{bmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a \end{bmatrix}$$

matriz escalar

e) Chama-se **matriz diagonal** a uma matriz quadrada com todos os elementos nulos excepto os da diagonal principal.

Exemplo 1.8 A matriz $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{bmatrix}$ é uma matriz diagonal e a matriz $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ é uma matriz escalar (também é diagonal).

1.6 Transposição de matrizes

Definição 1.6 - *Transposição de uma matriz sobre um corpo Ω.*

*A transposição é uma operação que a cada matriz A de ordem $m \times n$ sobre o corpo Ω faz corresponder uma outra matriz obtida pela troca ordenada de linhas e colunas. A nova matriz chama-se **matriz transposta da matriz A** e representa-se por A^t. A matriz A^t é de tipo $n \times m$.*

Propriedades da transposição de matrizes

1. $(A^{\mathrm{t}})^{\mathrm{t}} = A$

2. $(A + B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$

3. $(A.B)^{\mathrm{t}} = B^{\mathrm{t}}.A^{\mathrm{t}}$

Definição 1.7 - *Matriz simétrica*
A matriz A, quadrada de ordem n sobre Ω, é ***simétrica*** *se e só se é igual à sua transposta.*

$$A = A^t$$

Nota: não se deve confundir matriz simétrica ($A = A^{\mathrm{t}}$) com a simétrica da matriz A (a matriz $-A$).

Uma matriz é simétrica se e só se os elementos opostos, ou seja, A_{ij} e A_{ji}, são iguais para todos os valores de i e j.

Exemplo 1.9 A matriz real $A = \begin{bmatrix} 2 & 0 & 7 \\ 0 & 3 & 5 \\ 7 & 5 & 1 \end{bmatrix}$ é uma matriz simétrica.$\diamondsuit$

Definição 1.8 - *Matriz hemi-simétrica*
A matriz A, quadrada de ordem n sobre Ω, é ***hemi-simétrica*** *se e só se é igual à simétrica da sua transposta.*

$$A = -A^t$$

Exemplo 1.10 A matriz real $A = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}$ é uma matriz hemi-simétrica.$\diamondsuit$

Notas:
a) Qualquer que seja A, quadrada de ordem n, $A - A^{\mathrm{t}}$ é hemi-simétrica.
b) Qualquer matriz A, quadrada de ordem n, é igual à soma de uma matriz simétrica com uma matriz hemi-simétrica.

1.7 Conjugação de matrizes

Definição 1.9 - *Conjugação de uma matriz sobre o corpo **C** dos números complexos*

*A conjugação é uma operação que a cada matriz $A = [A_{ij}]$ de ordem $m \times n$ sobre o corpo **C** dos números complexos faz corresponder uma outra matriz cujo elemento genérico é o complexo conjugado de A_{ij}. Esta matriz chama-se **matriz conjugada de A** e representa-se por $\overline{A} = \left[\overline{A_{ij}}\right]$. A matriz $\overline{A}$ é de ordem $m \times n$.*

Propriedades da conjugação de matrizes

1. $\overline{\overline{A}} = A$
2. $\overline{A+B} = \overline{A} + \overline{B}$
3. $\overline{A.B} = \overline{A}.\overline{B}$

Exemplo 1.11 Se $A = \begin{bmatrix} 2+3j & 1-2j \\ j & 2 \end{bmatrix}$ então $\overline{A} = \begin{bmatrix} 2-3j & 1+2j \\ -j & 2 \end{bmatrix}$.$\diamondsuit$

1.8 Transconjugação de matrizes

Definição 1.10 - *Transconjugação de uma matriz sobre o corpo **C** dos números complexos*

*A transconjugação é uma operação que a cada matriz A de ordem $m \times n$ sobre o corpo **C** dos números complexos faz corresponder uma outra matriz que se obtém transpondo e conjugando a matriz A. Esta nova matriz chama-se **matriz transconjugada de A** ou **matriz adjunta de A** e representa-se por*

$$A^* = \overline{(A^t)} = \left(\overline{A}\right)^t$$

A matriz A^* é de ordem $n \times m$.

Propriedades da transconjugação de matrizes

1. $(A^*)^* = A$
2. $(A+B)^* = A^* + B^*$
3. $(A.B)^* = B^*.A^*$

Exemplo 1.12 Se $A = \begin{bmatrix} 2+j & 3 & j \\ 1 & -2j & 5 \end{bmatrix}$ então $A^* = \begin{bmatrix} 2-j & 1 \\ 3 & 2j \\ -j & 5 \end{bmatrix}$.$\diamondsuit$

Definição 1.11 - *Matriz hermiteana*

A matriz A, quadrada de ordem n sobre ***C****, é* ***hermiteana*** *se e só se é igual à sua transconjugada.*

$$A = A^*$$

É fácil concluir que uma matriz é hermiteana se e só se os elementos opostos são complexos conjugados e os elementos diagonais são reais.

Definição 1.12 - *Matriz hemi-hermiteana*

A matriz A, quadrada de ordem n sobre ***C****, é* ***hemi-hermiteana*** *se e só se é igual à simétrica da sua transconjugada.*

$$A = -A^*$$

Notas:

a) Uma matriz é hemi-hermiteana se e só se os elementos opostos são simétrico-conjugados e os elementos diagonais são nulos ou imaginários.

b) Qualquer que seja A, quadrada de ordem n, a matriz $A - A^*$ é hemi-hermiteana.

c) Qualquer que seja A, quadrada de ordem n, é igual à soma de uma matriz hermiteana com uma matriz hemi-hermiteana.

1.9 Inversão de matrizes

Definição 1.13 - *Inversa esquerda e inversa direita de uma matriz*

Seja A uma matriz de ordem $m \times n$ sobre um corpo Ω.

Chama-se inversa esquerda de A a qualquer matriz M de ordem $n \times m$ tal que $M.A = I_n$.

Chama-se inversa direita de A a qualquer matriz N de ordem $n \times m$ tal que $A.N = I_m$.

Verifica-se que se A tem inversa esquerda e inversa direita então as duas matrizes são iguais e A é uma matriz quadrada.

Definição 1.14 - *Inversa de uma matriz*

Seja A uma matriz sobre um corpo Ω. Chama-se ***matriz inversa de A*** *a uma matriz que se representa por A^{-1} tal que*

$$A^{-1}.A = A.A^{-1} = I$$

É condição necessária para A ter inversa que A seja uma matriz quadrada. Existindo inversa esta é única, mas nem todos as matrizes quadradas têm inversa.

Exemplo 1.13 Cálculo da inversa de uma matriz através da definição

Seja $A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$. A matriz inversa de A será também uma matriz quadrada de ordem 2. Vamos admitir que $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Pela definição de matriz inversa podemos escrever

$$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} . \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1.a + 0.c & 1.b + 0.d \\ 2.a + 3.c & 2.b + 3.d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Resolvendo o sistema de equações que se obtém da equação matricial anterior, é possível calcular os elementos a, b, c e d da matriz inversa de A.

$$\begin{cases} a & = & 1 \\ b & = & 0 \\ 2a + 3c & = & 0 \\ 2b + 3d & = & 1 \end{cases}$$

Assim $\begin{cases} a & = & 1 \\ b & = & 0 \\ c & = & -\frac{2}{3} \\ d & = & \frac{1}{3} \end{cases}$, e a matriz $A^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}. \diamondsuit$

Posteriormente serão estudados outros processos de cálculo da matriz inversa.

Propriedades da inversão de matrizes

1. Se A e B são matrizes quadradas de ordem n invertíveis, então $(A.B)^{-1} = B^{-1}.A^{-1}$.
2. Se A é quadrada de ordem n e invertível então $(A^{-1})^{-1} = A$.
3. $(A^t)^{-1} = (A^{-1})^t$.
4. $I^{-1} = I$.
5. É simples verificar que $A.B = O$ não implica $A = O$ ou $B = O$. Contudo, se A for invertível, então $A.B = O$ se e só se $B = O$. Com efeito, $A^{-1}.(A.B) = O \Rightarrow B = O$.
6. Se A é invertível então $(A^K)^{-1} = (A^{-1})^K$.

1.10 Operações elementares e matrizes elementares

Seja $M_{m\times n}$ o conjunto das matrizes $m \times n$ sobre um corpo. Recordemos que $A \in M_{m\times n}$ pode ser considerada como uma matriz com m linhas

$$A = \begin{bmatrix} A_{(1)} \\ A_{(2)} \\ ... \\ A_{(m)} \end{bmatrix}$$

ou uma matriz com n colunas.

$$A = \begin{bmatrix} A^{(1)} & A^{(2)} & ... & A^{(n)} \end{bmatrix}$$

Sobre estas matrizes podem ser realizadas operações elementares de três tipos, quer sobre as linhas, quer sobre as colunas.

Definição 1.15 - *Operações elementares*

Seja $A \in M_{m\times n}$. São designadas operações elementares sobre as linhas (colunas) de A:

(a) troca de duas linhas (colunas) de A;

(b) multiplicação de qualquer linha (coluna) de A por uma constante não nula;

(c) adição a uma linha (coluna) de outra linha (coluna) multiplicada por uma constante.

As operações elementares são do tipo 1, tipo 2 ou tipo 3 consoante são obtidas por (a), (b) ou (c).

Exemplo 1.14 Consideremos a matriz A

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

A troca da linha $A_{(2)}$ com a linha $A_{(1)}$ é uma operação elementar do tipo 1 sobre as linhas da matriz. A matriz resultante é

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

A multiplicação de $A^{(2)}$ por 3 é um exemplo de uma operação elementar do tipo 2 sobre as colunas. A matriz resultante é

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

Finalmente, a adição de $A_{(1)}$ com o resultado da multiplicação de $A_{(3)}$ por 4 é um exemplo de uma operação elementar de tipo 3 sobre as linhas. A matriz resultante é

$$D = \begin{bmatrix} 17 & 2 & 7 & 12 \\ 2 & 1 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{bmatrix}. \diamondsuit$$

Definição 1.16 - *Matriz elementar*

Uma matriz elementar $n \times n$ é uma matriz obtida pela execução de uma operação elementar sobre I_n.

A matriz elementar é de tipo 1, 2 ou 3 consoante a operação elementar que lhe deu origem é de tipo 1, 2 ou 3, respectivamente.

Exemplo 1.15 A troca das duas primeiras linhas de I_3 produz a matriz elementar

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

É importante notar que E também pode ser obtida pela troca das duas primeiras colunas de I_3.

De facto, qualquer matriz elementar pode ser conseguida pelo menos de duas maneiras distintas: pela realização de uma operação elementar sobre as linhas, ou pela execução de uma operação elementar sobre as colunas.

Exemplo 1.16 A matriz

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

também é uma matriz elementar pois pode ser obtida de I_3 somando à terceira coluna a primeira multiplicada por -2 (operação elementar sobre as colunas); porém, a mesma matriz pode resultar da soma da primeira linha com a terceira linha multiplicada por -2 (operação elementar sobre as linhas).$\diamondsuit$

Teorema 1.1 - *Seja $A \in M_{m\times n}$, e admitamos que B é obtida de A pela execução de uma operação elementar sobre as linhas (colunas). Então, existe uma matriz elementar E, quadrada de ordem m (n), tal que*

$$B = E.A \ (B = A.E).$$

Com efeito, a matriz elementar E é conseguida pela execução da operação elementar correspondente sobre as linhas (colunas) de $I_m(I_n)$.

Em contrapartida, se E for uma matriz elementar m×m (n×n), então $E.A(A.E)$ é uma matriz que pode ser obtida pela execução directa de uma operação elementar sobre as linhas (colunas) de A.

Exemplo 1.17 Seja B=$\begin{bmatrix} 2 & 1 & -1 & 3 \\ 1 & 2 & 3 & 4 \\ 4 & 0 & 1 & 2 \end{bmatrix}$ a matriz do exemplo anterior. Esta matriz foi conseguida a partir de A pela troca das suas duas primeiras linhas. Realizando uma operação idêntica sobre I_3 obtém-se a matriz elementar E

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

Efectuando o produto das matrizes $E.A$, verifica-se que $B = E.A$.

$$B = E.A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 & 3 \\ 1 & 2 & 3 & 4 \\ 4 & 0 & 1 & 2 \end{bmatrix}$$

Nos exemplos anteriormente apresentados, C resultou de A por multiplicação da segunda coluna por 3. Executando esta operação em I_4 obtém-se a seguinte matriz elementar:

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

Calculando o produto $A.E$, conclui-se que a matriz produto é a matriz C antes apresentada.

$$C = A.E = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 3 & 4 \\ 2 & 3 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{bmatrix} . \diamondsuit$$

É muito importante para as matérias que serão estudadas nas secções seguinte a garantia de existência da inversa de uma matriz elementar. Este facto é estabelecido no seguinte teorema:

Teorema 1.2 - *As matrizes elementares são invertíveis, e a inversa de uma matriz elementar é ainda uma matriz elementar do mesmo tipo.*

1.11 Estudo da característica de uma matriz

Consideremos novamente o conjunto das matrizes de ordem $m \times n$ sobre o corpo Ω, $M_{m\times n}$.

Definição 1.17 - *Combinação linear de linhas de uma matriz*

Seja $A \in M_{m\times n}$. Chama-se combinação linear das m linhas da matriz A à matriz linha definida pela expressão

$$X = \lambda_1 A_{(1)} + \lambda_2 A_{(2)} + ... + \lambda_m A_{(m)}$$

onde $\lambda_i \in \Omega$ $(i = 1...m)$.

Do mesmo modo, designa-se por combinação linear das n colunas da matriz A a matriz coluna que se obtém por

$$X = \lambda_1 A^{(1)} + \lambda_2 A^{(2)} + ... + \lambda_n A^{(n)},\ \lambda_i \in \Omega\ (i = 1...n)$$

Definição 1.18 - *Dependência linear das linhas de uma matriz*

Diz-se que as m linhas de $A \in M_{m \times n}$ são linearmente dependentes se e só se existem em Ω escalares $\lambda_1, \lambda_2, ...\lambda_m$ não todos nulos tais que

$$\lambda_1 A_{(1)} + \lambda_2 A_{(2)} + ... + \lambda_m A_{(m)} = O$$

Definição 1.19 - *Independência linear das linhas de uma matriz*

As linhas de uma matriz $A \in M_{m \times n}$ são linearmente independentes se se verifica

$$\lambda_1 A_{(1)} + \lambda_2 A_{(2)} + ... + \lambda_m A_{(m)} = O$$

só com todos os escalares λ_i $(i = 1...m)$ nulos.

Definições análogas às que acabámos de apresentar podem ser associadas às colunas de uma matriz.

Definição 1.20 - *Característica de uma matriz*

A característica de uma matriz é o número máximo de linhas (colunas) da matriz que são linearmente independentes.

Um dos processos de cálculo da característica de uma matriz de ordem $m \times n$ é o método designado por **condensação da matriz**. Para utilizar esse método é necessário começar por apresentar algumas definições e teoremas.

Definição 1.21 *Matriz triangular*

Chama-se matriz triangular a uma matriz quadrada em que são nulos todos os elementos para um dos lados da diagonal principal.

Deve distinguir-se entre:

- matriz triangular inferior, quando acima da diagonal só existem zeros;

- matriz triangular superior, quando abaixo da diagonal todos os elemento são nulos.

Exemplo 1.18 A matriz $\begin{bmatrix} 6 & 0 & 0 \\ 8 & 4 & 0 \\ 0 & 3 & -9 \end{bmatrix}$ é triangular inferior, enquanto a matriz $\begin{bmatrix} 6 & -2 & 8 \\ 0 & 5 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ é triangular superior.$\diamondsuit$

Teorema 1.3 - *Se T é uma matriz triangular, com elementos diagonais não nulos, então as suas linhas são linearmente independentes.*

Este teorema também é válido para as colunas.

Por outro lado, se T é uma matriz triangular com algum elemento diagonal nulo, então as linhas (colunas) são linearmente dependentes.

Teorema 1.4 - *As operações elementares sobre linhas ou colunas de uma matriz não alteram a sua característica.*

Teorema 1.5 - *Seja $A \in M_{m\times n}$ uma matriz com característica r. Então $r \leq m$, $r \leq n$, e, através de um número finito de operações elementares sobre as linhas e colunas, A pode ser transformada numa matriz D tal que:*

$$\begin{array}{l} (a) D_{ij} = 0 \text{ para } i \neq j \\ (b) D_{ii} = 1 \text{ para } i \leq r \\ (c) D_{ii} = 0 \text{ para } i > r \end{array}$$

Antes de apresentarmos a prova do teorema que acabamos de enunciar, vamos exemplificar o processo de condensação da matriz para determinação da sua característica.

Exemplo 1.19 Condensação de uma matriz para determinação da sua característica Consideremos a matriz

$$A = \begin{bmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix}, A \in M_{4\times 5}$$

Através de uma série de operações elementares vamos apresentar um exemplo de transformação de uma matriz A numa matriz D tal como é definida no teorema anterior; os números apresentados por cimas das setas indicam o número de operações elementares efectuadas.

$$\begin{bmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix} \xrightarrow{1} \begin{bmatrix} 4 & 4 & 4 & 8 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix} \xrightarrow{1} \begin{bmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix} \xrightarrow{2}$$

$$\begin{bmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{bmatrix} \xrightarrow{3} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{bmatrix} \xrightarrow{1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{bmatrix} \xrightarrow{2}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & 4 & 0 & 8 \\ 0 & 0 & 2 & 0 & 4 \end{bmatrix} \xrightarrow{3} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 8 \\ 0 & 0 & 2 & 0 & 4 \end{bmatrix} \xrightarrow{1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 & 4 \end{bmatrix} \xrightarrow{1}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = D,\ r(\mathrm{A}) = r(\mathrm{D}) = 3.$$

Prova: Se A é a matriz nula, então r(A)=r=0 ($D = A$).

Vamos admitir que $A \neq O$ e que r=r(A); então r(A)<0. A prova do teorema será efectuada por indução matemática em m, o número de linhas da matriz A.

Suponhamos $m = 1$. Através de, no máximo, uma operação elementar de tipo 1 sobre as colunas e de, no máximo, uma operação do tipo 2 sobre as colunas, A pode ser transformada numa matriz com o valor 1 na primeira posição. Com um número máximo de $n - 1$ operações do tipo 3 sobre as colunas é possível chegar à seguinte matriz

$$D = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

A matriz D contém apenas uma coluna linearmente independente; logo, r(D)=r(A)=1. Assim, o teorema está provado para $m = 1$.

Seguidamente, vamos admitir que o teorema continua a ser válido para qualquer matriz com $m-1$ linhas ($m > 1$). Vamos então provar que o teorema é válido para qualquer matriz com m linhas.

Admitamos que A é uma matriz qualquer de ordem $m \times n$. Se $n = 1$, o teorema pode ser demonstrado seguindo um processo semelhante ao que acabamos de expor para $m = 1$.

Consideremos agora $n > 1$. Dado que $A \neq O, A_{ij} \neq 0$ para algum par de índices (i, j). Por intermédio de uma operação sobre as linhas e de uma operação sobre as colunas (ambas de tipo 1) é possível deslocar a entrada não nula para a posição de índices (1,1). Através de, no máximo, uma operação adicional de tipo 2, garante-se um valor unitário na posição de índices (1,1). Com um número máximo de $m-1$ operações de tipo 3 sobre as linhas, e de $n-1$ operações tipo 3 sobre as colunas, podem ser eliminadas todas as entradas não nulas na primeira linha e na primeira coluna.

Por conseguinte, com um número finito de operações elementares, a matriz A pode ser transformada na matriz

$$B = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{bmatrix}$$

onde B' é uma matriz de dimensão $(m-1) \times (n-1)$. Como se pode observar, r(B')=r(B)-1. Dado que r(A)=r(B)=r, r(B')= $r-1$. Pela hipótese do método indutivo, $r-1 \leq m-1$ e $r-1 \leq n-1$; em consequência $r \leq m$ e $r \leq n$.

Também pela hipótese do método indutivo, B' pode ser transformada através de um número finito de operações elementares sobre as linhas e colunas numa matriz D' de ordem $(m-1) \times (n-1)$ tal que

(a) $D'_{ij} = 0$ para $i \neq j$
(b) $D'_{ii} = 1$ para $i \leq r-1$
(c) $D'_{ii} = 0$ para $i \geq r$

Assim, é possível obter D a partir de B por transformação de B'em D'.

Consequentemente, uma vez que A pode ser transformada em B, e B pode ser transformada em D apenas com um número finito de operações elementares em cada etapa, A pode ser transformada em D com um número finito de operações elementares. Finalmente, dado que $D^{'}$ contém valores 1 nas suas primeiras $r-1$ posições ao longo da diagonal principal e zeros nas restantes posições, D contém valores unitários apenas nas r primeiras posições diagonais. Logo, o teorema 1.5 está provado. ♣

Corolário 1.5.1 *Seja A uma matriz $m \times n$ com característica r. Então existem matrizes invertíveis B e C de dimensões $(m \times m)$ e $(n \times n)$, respectivamente, tais que $D = BAC$, sendo D uma matriz satisfazendo*

(a) $D_{ij} = 0$ para $i \neq j$
(b) $D_{ii} = 1$ para $i \leq r$
(c) $D_{ii} = 0$ para $i > r$

Prova: Pelo teorema 1.5, A pode ser transformada numa matriz D através de um número finito de operações elementares. Reportando-nos ao teorema 1.1, relembramos que a realização de uma operação elementar sobre uma matriz é equivalente à multiplicação dessa mesma matriz por uma matriz elementar. Então existem matrizes elementares $E_1, E_2, ..., E_p$ e $G_1, G_2, ..., G_q$ tais que

$$D = E_p E_{p-1} ... E_1 A G_1 G_2 ... G_q$$

O teorema 1.2 garante que cada uma das matrizes E_j e G_j é invertível. Designemos por B o produto $E_p E_{p-1} ... E_1$ e por C o produto $G_1 G_2 ... G_q$. Então B e C são invertíveis e $D = BAC$. ♣

Corolário 1.5.2 *Seja A uma matriz $m \times n$. Então r(A^t)=r(A).*

Corolário 1.5.3 *Qualquer matriz invertível é o produto de matrizes elementares.*

Exemplo 1.20 Seja $A = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 1 & 1 & -1 & 1 \end{bmatrix}$. Deve notar-se que as duas linhas de A são linearmente independentes pois não são múltiplas uma da outra; logo, r(A)=2.$\diamondsuit$

Exemplo 1.21 Seja $A = \begin{bmatrix} 1 & 3 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 3 & 0 & 0 \end{bmatrix}$. Por aplicação de uma operação elementar tipo 3 envolvendo as linhas 1 e 2 da matriz, resulta

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

Uma vez que na última matriz a terceira linha é múltipla da segunda, apenas as duas primeiras linhas são linearmente independentes. Portanto, r(A)=2.

Um método alternativo, mais simples neste exemplo particular, seria constatar que as colunas 1, 3 e 4 da matriz A são iguais, logo linearmente dependentes, pelo que r(A)=2.$\diamondsuit$

Exemplo 1.22 Seja A=$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 \end{bmatrix}$. Usando diversas operações elementares sobre as linhas e colunas é possível obter a sequência de matrizes:

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

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

É evidente que a última matriz da sequência tem 3 linhas linearmente independentes e consequentemente r(A)=3.$\diamondsuit$

1.12 Cálculo da inversa de uma matriz

Já foi antes apresentado um processo de inversão de matrizes que resultava directamente da definição de matriz inversa. Vai agora ser exposto um outro método de cálculo baseado na aplicação de operações elementares. Este método também é conhecido por **método de determinação da matriz inversa por condensação**.

Teorema 1.6 - *Uma matriz A, quadrada de ordem n, tem inversa A^{-1} se e só se a característica de A é igual a n.*

Definição 1.22 - *Sejam A e B matrizes ($m \times n$) e ($m \times p$), respectivamente. Por* ***matriz aumentada*** *entende-se a matriz $m \times (n+p)$*

$$\left[\begin{array}{c|c} A & B \end{array}\right] = \left[\begin{array}{ccc|ccc} A^{(1)} & \cdots & A^{(n)} & B^{(1)} & \cdots & B^{(p)} \end{array}\right]$$

onde $A^{(i)}$ é a coluna i da matriz A e $B^{(j)}$ é a coluna j da matriz B.

Seja A uma matriz invertível, e consideremos a matriz aumentada $n \times 2n$, $C = \left[\begin{array}{c|c} A & I_n \end{array}\right]$.

$$A^{-1}C = \left[\begin{array}{c|c} A^{-1}A & A^{-1}I_n \end{array}\right] = \left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$$

Pelo corolário 4.0.1, A^{-1} é o produto de matrizes elementares. Seja

$$A^{-1} = E_p E_{p-1} \ldots E_1.$$

Então podemos escrever

$$A^{-1}C = E_p E_{p-1} \ldots E_1 \left[\begin{array}{c|c} A & I_n \end{array}\right] = \left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right].$$

Uma vez que multiplicar à esquerda por uma matriz elementar equivale à execução da mesma operação elementar sobre as linhas da matriz, chegamos ao seguinte resultado:

- Se A é uma matriz quadrada de ordem n invertível, então é possível transformar a matriz $\left[\begin{array}{c|c} A & I_n \end{array}\right]$ na matriz $\left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$ através de um número finito de operações elementares sobre as linhas.

O resultado anterior pode ainda ser estendido com a seguinte conclusão:

- Se A é uma matriz quadrada de ordem n invertível e se a matriz $\left[\begin{array}{c|c} A & I_n \end{array}\right]$ for transformada numa matriz da forma $\left[\begin{array}{c|c} I_n & B \end{array}\right]$ através de um número finito de operações elementares sobre as linhas, então $B = A^{-1}$.

Prova: Sejam $E_1, E_2, \ldots, E_p$ as matrizes elementares associadas às operações elementares sobre as linhas. Então

$$E_p E_{p-1} \ldots E_1 \left[\begin{array}{c|c} A & I_n \end{array}\right] = \left[\begin{array}{c|c} I_n & B \end{array}\right].$$

Se $M = E_p E_{p-1} ... E_1$ a equação anterior pode ser escrita na forma

$$M \left[\begin{array}{c|c} A & I_n \end{array} \right] = \left[\begin{array}{c|c} MA & M \end{array} \right] = \left[\begin{array}{c|c} I_n & B \end{array} \right].$$

Assim, por um lado $MA = I_n$, ou seja, $M = A^{-1}$ e, em simultâneo, $M = B$. Logo terá que se verificar que $M = B = A^{-1}$. ♣

Exemplo 1.23 Cálculo da matriz inversa da matriz A=$\begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 0 & 2 & 4 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 1 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|ccc} 2 & 4 & 2 & 0 & 1 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array} \right] \longrightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1/2 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1/2 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 0 & -3 & -2 & 0 & -3/2 & 1 \end{array} \right] \longrightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1/2 & 0 \\ 0 & 1 & 2 & 1/2 & 0 & 0 \\ 0 & -3 & -2 & 0 & -3/2 & 1 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1/2 & 0 \\ 0 & 1 & 2 & 1/2 & 0 & 0 \\ 0 & 0 & 4 & 3/2 & -3/2 & 1 \end{array} \right] \longrightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & 1/2 & 0 \\ 0 & 1 & 2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 3/8 & -3/8 & 1/4 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -3/8 & 7/8 & -1/4 \\ 0 & 1 & 0 & -1/4 & 3/4 & -1/2 \\ 0 & 0 & 1 & 3/8 & -3/8 & 1/4 \end{array} \right] \longrightarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -3/8 & 7/8 & -1/4 \\ 0 & 1 & 0 & -1/4 & 3/4 & -1/2 \\ 0 & 0 & 1 & 3/8 & -3/8 & 1/4 \end{array} \right] \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/8 & -5/8 & 3/4 \\ 0 & 1 & 0 & -1/4 & 3/4 & -1/2 \\ 0 & 0 & 1 & 3/8 & -3/8 & 1/4 \end{array} \right]$$

Portanto, A^{-1}=$\begin{bmatrix} 1/8 & -5/8 & 3/4 \\ -1/4 & 3/4 & -1/2 \\ 3/8 & -3/8 & 1/4 \end{bmatrix}$.♢

Capítulo 2

Determinantes

2.1 Definição de determinante

Em Álgebra elementar aprende-se que o determinante de uma matriz de 2x2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

é calculado pela fórmula

$$\det(A) = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Podemos assim observar que o determinante da matriz A, vulgarmente designado simplesmente por **determinante de** A, é um escalar pertencente ao conjunto de definição dos elementos da matriz.

Definição 2.1 - *Determinante de uma matriz quadrada*

O determinante de uma matriz A de ordem $n \times n$ sobre um corpo Ω é um escalar pertencente a Ω, denotado por $\det(A)$ ou $|A|$, que pode ser calculado da forma que a seguir se indica:

1. *se A tem dimensão 1×1, então $\det(A) = A_{11}$;*
2. *se A tem dimensão 2×2, então $\det(A) = A_{11}A_{22} - A_{12}A_{21}$;*
3. *se A tem dimensão $n \times n$, $n > 2$, então o determinante de A pode exprimir-se como a soma dos produtos de cada um dos elementos de uma fila de A (linha ou coluna) pelo determinante da matriz (n-1)×(n-1) obtida eliminando*

em A a linha e a coluna contendo o elemento em questão; antes da soma ser efectuada, as parcelas produto são afectadas de um sinal positivo ou negativo. A expressão de cálculo do determinante é:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(\widetilde{A_{ij}}) \tag{2.1}$$

se o determinante for calculado usando os elementos da linha i da matriz, ou

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(\widetilde{A_{ij}}) \tag{2.2}$$

se o determinante for calculado usando os elementos da coluna j da matriz.

As expressões que acabámos de apresentar para o cálculo do determinante de uma matriz A são também referidas como **desenvolvimento Laplaciano** ao longo da linha (expressão 2.1) ou ao longo da coluna (expressão 2.2).

A matriz $\widetilde{A_{ij}}$ é uma submatriz de A obtida por eliminação da linha i e coluna j. Nas fórmulas anteriores, o escalar $(-1)^{i+j} \det\widetilde{(A_{ij})}$ é chamado **cofactor** ou **complemento algébrico** do elemento A_{ij}.

Em linguagem corrente, dizemos que o determinante de A é calculado pela soma dos produtos que se obtêm multiplicando cada um dos elementos de uma fila (linha ou coluna) da matriz A pelo respectivo cofactor.

Exemplo 2.1 Cálculo do determinante da matriz A

$$A = \begin{bmatrix} 2 & 1 & 1 & 5 \\ 1 & 1 & -4 & -1 \\ 2 & 0 & -3 & 1 \\ 3 & 6 & 1 & 2 \end{bmatrix} \qquad \det(A) = \begin{vmatrix} 2 & 1 & 1 & 5 \\ 1 & 1 & -4 & -1 \\ 2 & 0 & -3 & 1 \\ 3 & 6 & 1 & 2 \end{vmatrix}$$

Calculando o determinante da matriz A através do desenvolvimento ao longo da quarta linha da matriz

$$\det(A) = (-1)^{4+1}.3. \begin{vmatrix} 1 & 1 & 5 \\ 1 & -4 & -1 \\ 0 & -3 & 1 \end{vmatrix} + (-1)^{4+2}.6. \begin{vmatrix} 2 & 1 & 5 \\ 1 & -4 & -1 \\ 2 & -3 & 1 \end{vmatrix}$$

$$+(-1)^{4+3}.1. \begin{vmatrix} 2 & 1 & 5 \\ 1 & 1 & -1 \\ 2 & 0 & 1 \end{vmatrix} + (-1)^{4+4}.2. \begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & -4 \\ 2 & 0 & -3 \end{vmatrix}$$

Os determinantes de terceira ordem assim obtidos podem agora ser calculados usando desenvolvimentos Laplacianos ao longo das suas linhas ou colunas.

Desses desenvolvimentos resultam determinantes de segunda ordem cujos valores podem ser calculados directamente a partir das fórmulas apresentadas na definição 2.1.

Por exemplo, calculando o primeiro determinante de terceira ordem através do seu desenvolvimento ao longo da primeira coluna

$$\begin{vmatrix} 1 & 1 & 5 \\ 1 & -4 & -1 \\ 0 & -3 & 1 \end{vmatrix} = (-1)^{1+1}.1.\begin{vmatrix} -4 & -1 \\ -3 & 1 \end{vmatrix} + (-1)^{2+1}.1.\begin{vmatrix} 1 & 5 \\ -3 & 1 \end{vmatrix}$$

$$+(-1)^{1+1}.0.\begin{vmatrix} 1 & 5 \\ -4 & -1 \end{vmatrix} = -1 - 16 = -17.\diamondsuit$$

No entanto, para o cálculo de determinantes de ordem 3 existe uma regra prática, designada **Regra de Sarrus**. Com efeito, o determinante de uma matriz A de ordem 3 é calculado pela expressão

$$\det(A) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - \ c_1b_2a_3 - c_2b_3a_1 - c_3b_1a_2.$$

Os esquemas indicados na figura a seguir facilitam a memorização da regra de Sarrus para cálculo de determinantes de ordem 3.

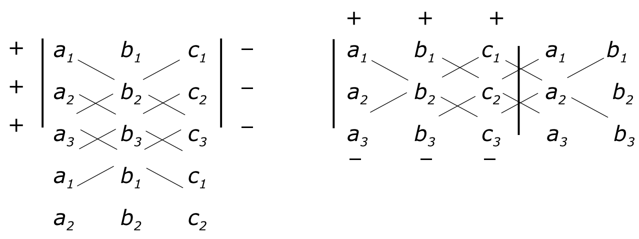

Figura 2.1: Regra de Sarrus

Notas sobre a definição de determinante

Há várias formas alternativas de introduzir o conceito de determinante. Neste curso foi escolhida a definição baseada nos desenvolvimentos Laplacianos ao longo de linhas ou de colunas. Esta abordagem foi seleccionada porque fornece de imediato um processo de cálculo simples.

Contudo, há uma definição alternativa fundamentada no conceito de permutações pares e ímpares de n objectos. As regras de cálculo de determinantes de segunda e terceira ordem derivam directamente da definição baseada em permutações.

Antes de introduzirmos a definição alternativa de determinante, torna-se necessário apresentar alguns conceitos em que essa definição se fundamenta.

Definição 2.2 -*Termo de uma matriz quadrada de ordem n*

Seja A uma matriz quadrada de ordem n sobre um corpo. Chama-se termo da matriz A a qualquer produto de n elementos da matriz

$$A_{i1j1}A_{i2j2}...A_{injn}$$

com um e só um factor em cada linha e com um e só um factor em cada coluna.

Dado que a ordem por que se apresentam os factores num termo é irrelevante, dois termos só são distintos quando, por reordenação, não se podem fazer coincidir. Por exemplo,

$$A_{12}A_{23}A_{31} \text{ e } A_{31}A_{12}A_{23}$$

são o mesmo termo, mas

$$A_{12}A_{23}A_{31} \text{ e } A_{23}A_{32}A_{11}$$

são termos distintos.

Daqui resulta que um dado termo pode sempre ser apresentado com os índices das linhas (ou das colunas) por ordem crescente. Assim, o número de termos distintos de uma matriz quadrada de ordem n é exactamente n!, ou seja, há tantos termos distintos quantas as permutações possíveis dos índices das colunas (ou das linhas).

Exemplo 2.2 O termo $A_{12}A_{23}A_{31}$ apresenta-se com os índices de linha ordenados por ordem crescente, enquanto o termo $A_{31}A_{12}A_{23}$ tem os seus índices de coluna ordenados. Em contrapartida, no termo $A_{23}A_{11}A_{32}$ nem os índices de linha nem os índices de coluna estão ordenados.$\diamondsuit$

Definição 2.3 - *Sinal de um termo*

Chama-se sinal de um termo a uma potência de (-1), (-1)$^\alpha$, cujo expoente α é igual à soma do número de inversões dos índices das linhas com o número de inversões dos índices das colunas.

Definição 2.4 - *Determinante de uma matriz quadrada sobre um corpo*

Determinante de uma matriz A, quadrada de ordem n sobre um corpo Ω, é um escalar de Ω igual à soma dos termos distintos de A afectados do respectivo sinal.

Exemplo 2.3 Vamos usar a definição alternativa que acabámos de apresentar para calcular o determinante da matriz de quarta ordem

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

$$\begin{aligned} \det A = {} & (-1)^0 A_{11}A_{22}A_{33}A_{44} + (-1)^1 A_{11}A_{22}A_{34}A_{43} + (-1)^1 A_{11}A_{23}A_{32}A_{44} + \\ & (-1)^2 A_{11}A_{23}A_{34}A_{42} + (-1)^2 A_{11}A_{24}A_{32}A_{43} + (-1)^3 A_{11}A_{24}A_{33}A_{42} + \\ & (-1)^1 A_{12}A_{21}A_{33}A_{44} + (-1)^2 A_{12}A_{21}A_{34}A_{43} + (-1)^3 A_{12}A_{23}A_{34}A_{41} + \\ & (-1)^2 A_{12}A_{23}A_{31}A_{44} + (-1)^3 A_{12}A_{24}A_{31}A_{43} + (-1)^4 A_{12}A_{24}A_{33}A_{41} + \\ & (-1)^2 A_{13}A_{21}A_{32}A_{44} + (-1)^3 A_{13}A_{21}A_{34}A_{42} + (-1)^3 A_{13}A_{22}A_{31}A_{44} + \\ & (-1)^4 A_{13}A_{22}A_{34}A_{41} + (-1)^4 A_{13}A_{24}A_{31}A_{42} + (-1)^5 A_{13}A_{24}A_{32}A_{41} + \\ & (-1)^3 A_{14}A_{21}A_{32}A_{43} + (-1)^4 A_{14}A_{21}A_{33}A_{42} + (-1)^4 A_{14}A_{22}A_{31}A_{43} + \\ & (-1)^5 A_{14}A_{22}A_{33}A_{41} + (-1)^5 A_{14}A_{23}A_{31}A_{42} + (-1)^6 A_{14}A_{23}A_{32}A_{41} \end{aligned}$$

Considerando agora a matriz que já antes utilizámos como exemplo

$$A = \begin{bmatrix} 2 & 1 & 1 & 5 \\ 1 & 1 & -4 & -1 \\ 2 & 0 & -3 & 1 \\ 3 & 6 & 1 & 2 \end{bmatrix}$$

$$\det(A) = (-1)^0.2.1.(-3).2 + (-1)^1.2.1.1.1 + (-1)^1.2.(-4).0.2 + (-1)^2.2.(-4).1.6$$

$$+(-1)^2.2.(-1).0.1 + (-1)^3.2.(-1)(-3)(6) + (-1)^1.1.1.(-3).2 + (-1)^2.1.1.1.1$$

$$+(-1)^3.1.(-4).1.3 + (-1)^2.1.(-4).2.2 + (-1)^3.1.(-1).2.1 + (-1)^4.1.(-1).(-3).3$$

$$+(-1)^2.1.1.0.2 + (-1)^3.1.1.1.6 + (-1)^3.1.1.2.2 + (-1)^4.1.1.1.3$$

$$+(-1)^4.1.(-1).2.6 + (-1)^5.1.(-1).0.3 + (-1)^3.5.1.0.1 + (-1)^4.5.1.(-3).6$$

$$+(-1)^4.5.1.2.1 + (-1)^5.5.1.(-3).3 + (-1)^5.5.(-4).2.6 + (-1)^6.5.(-4).0.3$$

$$\det(A) = -12 - 2 - 0 - 48 + 0 - 36 + 6 + 1 + 12 - 16 + 2 + 9 + 0$$

$$-6 - 4 + 3 - 12 + 0 - 0 - 90 + 10 + 45 + 240 + 0 = 102\diamondsuit$$

2.2 Propriedades dos determinantes

D1 - Se B é uma matriz obtida pela troca de duas filas paralelas (linhas ou colunas) de uma matriz A, então
$$\det(B) = -\det(A)$$

D2 - Se B é uma matriz que resulta de A por multiplicação de cada um dos elementos de uma fila (linha ou coluna) por um escalar λ, então

$$\det(B) = \lambda \det(A)$$

D3 - Se B é uma matriz obtida adicionando à linha (coluna) j um múltiplo da linha (coluna) i da matriz A, com i$\neq$j, então

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

Exemplo 2.4 Para ilustrar o uso destas três propriedades no cálculo de determinantes, vamos voltar a calcular o determinante da matriz A já considerada antes.

$$A = \begin{bmatrix} 2 & 1 & 1 & 5 \\ 1 & 1 & -4 & -1 \\ 2 & 0 & -3 & 1 \\ 3 & 6 & 1 & 2 \end{bmatrix}$$

Através da aplicação sucessiva da propriedade D3, vamos introduzir zeros na segunda coluna de A; depois o valor do determinante será muito facilmente calculado por expansão ao longo dessa coluna.

$$\det(A) = \begin{vmatrix} 2 & 1 & 1 & 5 \\ 1 & 1 & -4 & -1 \\ 2 & 0 & -3 & 1 \\ 3 & 6 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 1 & 5 \\ -1 & 0 & -5 & -6 \\ 2 & 0 & -3 & 1 \\ -9 & 0 & -5 & -28 \end{vmatrix} = 1(-1)^{1+2} \begin{vmatrix} -1 & -5 & -6 \\ 2 & -3 & 1 \\ -9 & -5 & -28 \end{vmatrix}$$

O determinante de terceira ordem que resultou do desenvolvimento Laplaciano ao longo da coluna pode ser calculado por processo análogo. Se, de novo por aplicação da propriedade D3, fizermos aparecer zeros na coluna 1 obtemos

$$\det(A) = 1(-1)^3 \begin{vmatrix} -1 & -5 & -6 \\ 0 & -13 & -11 \\ 0 & 40 & 26 \end{vmatrix} = (-1)\left[(-1)(-1)^{1+1} \begin{vmatrix} -13 & -11 \\ 40 & 26 \end{vmatrix}\right]$$
$$= 102. \diamondsuit$$

D4 - $\det(I) = 1$.

D5 - Se duas filas paralelas (linhas ou colunas) da matriz forem iguais, então o determinante da matriz é nulo.

D6 - Se uma fila (linha ou coluna) da matriz for nula, então o seu determinante é nulo.

D7 - As filas paralelas (linhas ou colunas) de uma matriz são linearmente independentes se e só se o determinante da matriz for não nulo.

D8 - O determinante de uma matriz triangular é o produto dos elementos da diagonal principal.

D9 - $\det(A) = \det(A^t)$.

D10 - Quaisquer que sejam $A, B \in M_n$, $\det(AB) = \det(A).\det(B)$.

2.3 Cálculo da matriz inversa usando determinantes

Vamos agora estudar outro processo de cálculo da matriz inversa de uma matriz A, quadrada de ordem n, usando determinantes.

Teorema 2.1 - *Uma matriz A, quadrada de ordem n, tem inversa A^{-1} se e só se o seu determinante é diferente de zero.*

Prova: Para demonstrar este teorema basta notar que A tem inversa A^{-1} sse r(A)=n. Porém, r(A)=n se e só se $\det(A) \neq 0$. Para justificar esta última afirmação, é suficiente pensar no processo de condensação da matriz para determinar a sua característica, pois a aplicação de operações elementares tem como consequência, na pior da hipóteses, a multiplicação do valor de $\det(A)$ por uma constante não nula. ♣

Vamos agora relacionar o determinante de uma matriz A com o determinante da sua matriz inversa A^{-1}, e procurar um processo de cálculo desta última matriz.

Admitamos que A é invertível; então $AA^{-1} = I$. Atendendo à propriedade D10

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

Logo,

$$\det(A^{-1}) = \frac{1}{\det(A)} = [\det(A)]^{-1}.$$

Definição 2.5 - *Matriz dos cofactores*

Seja A uma matriz quadrada de ordem n. Chama-se matriz dos cofactores de A, $C(A)$ à matriz que se obtém substituindo cada elemento A_{ij} de A pelo respectivo cofactor. Esta matriz também é conhecida por matriz dos complementos algébricos.

$$C(A) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}$$

Definição 2.6 - *Matriz adjunta clássica*

A matriz adjunta clássica de uma matriz A, $Adj(A)$, é a transposta da matriz dos cofactores.

$$Adj\,(A) = [C\,(A)]^t = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

Nota: Quando não houver possibilidade de confusão com a matriz conjugada transposta (matriz adjunta) a matriz adjunta clássica será simplesmente designada por matriz adjunta.

A matriz inversa de uma matriz A invertível pode ser calculada pela fórmula

$$A^{-1} = \frac{1}{\det(A)}\,[C\,(A)]^t = \frac{1}{\det(A)}\,[Adj\,(A)]\,.$$

Para provar esta afirmação vamos calcular o produto $P = A\,[Adj\,(A)]$.

$$P = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}C_{11} + A_{12}C_{12} + ...A_{1n}C_{1n} & \cdots & A_{11}C_{n1} + A_{12}C_{n2} + ...A_{1n}C_{nn} \\ A_{21}C_{11} + A_{22}C_{12} + ...A_{2n}C_{1n} & \cdots & A_{21}C_{n1} + A_{22}C_{n2} + ...A_{2n}C_{nn} \\ \vdots & \ddots & \vdots \\ A_{n1}C_{11} + A_{n2}C_{12} + ...A_{nn}C_{1n} & \cdots & A_{n1}C_{n1} + A_{n2}C_{n2} + ...A_{nn}C_{nn} \end{bmatrix}$$

Assim, podemos escrever

$$P_{ij} = \sum_{j=1}^{n} A_{ij}C_{kj} \left\{ \begin{array}{c} |A| \text{ para } i = k \\ 0 \text{ para } i \neq k \end{array} \right.$$

Logo,

$$P = A\,[Adj\,(A)] = \begin{bmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |A| \end{bmatrix} = |A|\,I_n$$

$$A.\left[\frac{1}{|A|}\left[Adj\left(A\right)\right]\right]=I_n$$

ou

$$A^{-1}=\tfrac{1}{|A|}\left[Adj\left(A\right)\right].$$

Em conclusão, a expressão que acabamos de apresentar constitui um método alternativo para a inversão de matrizes. ♣

Exemplo 2.5 Cálculo da matriz inversa de $A=\begin{bmatrix}1 & 2 & -1\\ -1 & 1 & 2\\ 2 & -1 & 1\end{bmatrix}$

Comecemos por determinar o valor de $|A|$.

$$|A|=\begin{vmatrix}1 & 2 & -1\\ -1 & 1 & 2\\ 2 & -1 & 1\end{vmatrix}=$$

$$=1.1.1+(-1).(-1).(-1)+2.2.2-(-1).1.2-2.(-1).1-1.2.(-1)=14$$

Os cofactores dos elementos de A são:

$$C_{11}=(-1)^{1+1}\begin{vmatrix}1 & 2\\ -1 & 1\end{vmatrix}=3 \quad C_{12}=(-1)^{1+2}\begin{vmatrix}-1 & 2\\ 2 & 1\end{vmatrix}=5$$

$$C_{13}=(-1)^{1+3}\begin{vmatrix}-1 & 1\\ 2 & -1\end{vmatrix}=-1$$

$$C_{21}=(-1)^{2+1}\begin{vmatrix}2 & -1\\ -1 & 1\end{vmatrix}=-1 \quad C_{22}=(-1)^{2+2}\begin{vmatrix}1 & -1\\ 2 & 1\end{vmatrix}=3$$

$$C_{23}=(-1)^{2+3}\begin{vmatrix}1 & 2\\ 2 & -1\end{vmatrix}=5$$

$$C_{31}=(-1)^{3+1}\begin{vmatrix}2 & -1\\ 1 & 2\end{vmatrix}=5 \quad C_{32}=(-1)^{3+2}\begin{vmatrix}1 & -1\\ -1 & 2\end{vmatrix}=-1$$

$$C_{33}=(-1)^{3+3}\begin{vmatrix}1 & 2\\ -1 & 1\end{vmatrix}=3$$

A matriz adjunta (clássica) de A é

$$Adj(A) = \begin{bmatrix} 3 & -1 & 5 \\ 5 & 3 & -1 \\ -1 & 5 & 3 \end{bmatrix}$$

Uma vez que

$$A^{-1} = \tfrac{1}{|A|} \left[Adj\,(A) \right]$$

então a matriz inversa de A é

$$A^{-1} = \frac{1}{14} \begin{bmatrix} 3 & -1 & 5 \\ 5 & 3 & -1 \\ -1 & 5 & 3 \end{bmatrix} = \begin{bmatrix} \frac{3}{14} & \frac{-1}{14} & \frac{5}{14} \\ \frac{5}{14} & \frac{3}{14} & \frac{-1}{14} \\ \frac{-1}{14} & \frac{5}{14} & \frac{3}{14} \end{bmatrix} . \diamondsuit$$

Capítulo 3

Sistemas de Equações Lineares

3.1 Conceitos gerais

A resolução de sistemas de equações é provavelmente uma das aplicações mais importantes da Álgebra Linear. O conhecido método de eliminação para a determinação das soluções de um sistema de equações lineares foi já estudado em cursos anteriores e envolve a eliminação de variáveis de modo a obter um sistema mais simples. A técnica utilizada para eliminar variáveis recorre a três tipos de operações:

1. troca de duas equações do sistema;

2. multiplicação de qualquer equação do sistema por uma constante não nula;

3. soma de um múltiplo de uma equação a outra equação.

Neste curso vamos ver como um sistema de equações lineares se pode exprimir por uma única equação matricial, e, de seguida, estudar dois métodos para determinar as soluções do sistema. O primeiro método recorre ao conceito de determinante, estudado no capítulo anterior; o segundo método é uma sistematização do método de eliminação, e baseia-se na utilização de operações elementares sobre as linhas das matrizes que representam o sistema.

Comecemos, porém, por estabelecer algumas definições importantes.

Definição 3.1 - *Sistema de m equações lineares em n incógnitas sobre um corpo*
O sistema de equações

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

onde a_{ij} *e* b_j $(1 \leq i \leq m, 1 \leq j \leq n)$ *são elementos de um corpo* Ω *e* $x_1, x_2, ..., x_n$ *são* n *variáveis tomando valores em* Ω, *é chamado* **sistema de m equações lineares em n incógnitas sobre o corpo** Ω.

Os escalares a_{ij} são designados por **coeficientes das incógnitas** e os escalares b_i são os **termos conhecidos** ou **termos independentes**.

A matriz $m \times n$

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

é chamada **matriz dos coeficientes do sistema (S)** ou, simplesmente, **matriz do sistema (S)**.

Se estabelecermos

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

então o sistema (S) pode ser reescrito como uma equação matricial única

$$AX = B.$$

Uma solução do sistema (S) é um conjunto de n escalares s_i

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

que verifique $As = B$.

A colecção de todos as soluções do sistema (S) designa-se **conjunto de soluções do sistema**.

A matriz aumentada $\left[\begin{array}{c|c} A & B \end{array}\right]$, que se obtém de A ampliando-a com a matriz coluna dos termos conhecidos, chama-se **matriz completa do sistema**.

Classificação dos sistemas quanto às soluções

Se o sistema (S) tiver pelo menos uma solução o sistema é **possível**; diz-se ainda que as m equações são **compatíveis**.

Se p das equações do sistema ($p < m$) têm uma solução em comum, então qualquer das outras equações que seja satisfeita por essa mesma solução diz-se compatível com as outras p equações.

Quando existe uma única solução para o sistema, este é **determinado**; pelo contrário, se tem mais do que uma solução, o sistema é **indeterminado**.

Um **sistema impossível** é aquele que não tem qualquer solução; as suas equações são **incompatíveis**.

Em resumo

$$\text{Sistema}\begin{cases} \text{possível (tem soluções)} \begin{cases} \text{determinado} & \text{(uma única solução)} \\ \\ \text{indeterminado} & \text{(várias soluções)} \end{cases} \\ \\ \text{impossível (não tem solução)} \end{cases}$$

Quando os termos conhecidos (escalares b_i) são todos nulos, as equações e o sistema dizem-se **homogéneos**. Um sistema homogéneo é sempre possível uma vez que admite sempre a solução nula $x_j = 0$ ($j = 1...n$).

3.2 Resolução de sistemas de equações lineares

Resolver um sistema é determinar todas as suas soluções, ou concluir que o sistema é impossível.

Dois sistemas dizem-se **equivalentes** se e só se têm as mesmas soluções. A equivalência entre sistemas de equações permite encontrar as soluções de um sistema resolvendo qualquer outro sistema equivalente ao sistema dado. Alguns métodos de resolução de sistemas consistem precisamente em passar, em etapas sucessivas, do sistema inicialmente proposto para outro mais simples de resolver; todos os sistemas intermédios são equivalentes entre si. As soluções do sistema final devem poder ser determinadas de forma quase imediata.

3.3 Resolução de sistemas de equações usando determinantes

Vamos considerar um sistema de m equações lineares a n incógnitas (S), e estudar a sua resolução usando o conceito de determinante. Podem distinguir-se três casos diferentes:

1. a característica da matriz do sistema é igual ao número de equações e ao número de incógnitas;

2. a característica da matriz do sistema é igual ao número de equações, mas o número de incógnitas é superior;

3. a característica da matriz do sistema é inferior ao número de equações.

1º Caso r(A)=m=n

Considere-se um sistema de equações com igual número de equações e de incógnitas ($m = n$) e de modo a que a matriz quadrada dos coeficientes, A, tenha característica igual à ordem da matriz (ou $\det(A) \neq 0$). Um sistema nas condições mencionadas é vulgarmente chamado **sistema de Cramer**.

A resolução do sistema de Cramer, $AX = B$, pode ser efectuada facilmente calculando a matriz inversa de A, A^{-1}, e multiplicando-a por B.

$$X = A^{-1}B$$

Vamos, contudo, apresentar um teorema que fundamenta uma regra prática de resolução de sistema nas condições antes mencionadas; esta regra é conhecida por **regra de Cramer**.

Teorema 3.1 (Cramer) - *o valor de cada incógnita x_j é obtido pelo quociente de dois determinantes: o determinante do denominador é o determinante $|A|$ da matriz dos coeficientes das incógnitas; o determinante do numerador é o determinante que resulta de $|A|$ substituindo a coluna dos coeficientes da incógnita x_j pela coluna dos termos independentes.*

Prova: Como já se viu $X = A^{-1}B$, onde $A^{-1} = \frac{1}{|A|}\left[Adj(A)\right]$. Então

$$X = \frac{1}{|A|}\left[Adj(A)\right]B.$$

Logo, podemos escrever

$$x_j = \frac{C_{1j}b_1 + C_{2j}b_2 + \cdots + C_{nj}b_n}{|A|}$$

ou ainda

$$x_j = \frac{\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}}$$

pois fazendo o desenvolvimento Laplaciano do determinante do numerador ao longo da coluna j, concluímos

$$\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} = b_1C_{1j} + b_2C_{2j} + \cdots + b_nC_{nj}.$$

O teorema de Cramer fica assim demonstrado. ♣

Exemplo 3.1 Considere-se o sistema sobre o corpo dos números reais $\mathbf{R}$

$$\left\{ \begin{array}{cccc} x_1 & +\,2x_2 & & = & 5 \\ 2x_1 & +\,x_2 & +\,x_3 & = & 1 \\ x_1 & & +\,x_3 & = & 7 \end{array} \right.$$

O determinante da matriz dos coeficientes

$$\det(A) = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix}$$

pode ser calculado, por exemplo, por desenvolvimento ao longo da terceira linha resultando

$$\det(A) = (-1)^4.1. \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} + (-1)^6.1. \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = (2) + (1-4) = -1.$$

Dado que det(A)$\neq$ 0, o sistema é de Cramer. Logo,

$$x_1 = \frac{\begin{vmatrix} 5 & 2 & 0 \\ 1 & 1 & 1 \\ 7 & 0 & 1 \end{vmatrix}}{-1} = -17 \qquad x_2 = \frac{\begin{vmatrix} 1 & 5 & 0 \\ 2 & 1 & 1 \\ 1 & 7 & 1 \end{vmatrix}}{-1} = 11 \qquad x_3 = \frac{\begin{vmatrix} 1 & 2 & 5 \\ 2 & 1 & 1 \\ 1 & 0 & 7 \end{vmatrix}}{-1} = 24.\diamondsuit$$

2º Caso r(A)= m $<$ n

Neste caso, o sistema tem um número de incógnitas superior ao número de equações. Porém, a característica da matriz dos coeficientes é igual ao número de equações. Isto significa que as m linhas da matriz são linearmente independentes, o que é equivalente a dizer que as m equações do sistema são linearmente independentes.

Um sistema com estas características pode ser apresentado da seguinte maneira

$$(S) \qquad \left\{ \begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m + a_{1(m+1)}x_{(m+1)} + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mm}x_m + a_{m(m+1)}x_{(m+1)} + \cdots + a_{mn}x_n = b_m \end{array} \right.$$

ou, em notação matricial,

$$AX = B$$

onde A é uma matriz $m \times n, m < n$, e r(A)=m.

Como a característica de A é m, a matriz tem m colunas linearmente independentes. Vamos pois admitir que são as primeiras m colunas (no caso das m primeiras colunas não serem linearmente independentes, é sempre possível alterar a sua ordem trocando também a ordem das respectivas incógnitas).

Então existe uma submatriz de A, A_1 quadrada de ordem m, com determinante diferente de zero.

Seja $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$, onde A_1 é de ordem $m \times m$ e A_2 é de ordem $m \times (n-m)$, e $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, com

$$X_1 = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \text{ e } X_2 = \begin{bmatrix} x_{m+1} \\ \vdots \\ xn \end{bmatrix}.$$

As incógnitas incluídas em X_1 designam-se **incógnitas principais**; as incógnitas incluídas em X_2 chamam-se **incógnitas não principais, ou livres.**

O sistema $AX = B$ pode agora ser reescrito na forma

$$A_1X_1 + A_2X_2 = B.$$

Passando para o segundo membro as incógnitas não principais, obtém-se um sistema (S')

$$(S') \qquad A_1X_1 = B - A_2X_2.$$

Se considerarmos que as incógnitas não principais integram os termos do segundo membro, o sistema (S') obtido pelo processo que descrevemos é de Cramer e pode ser resolvido usando a regra de Cramer, antes exposta.

Sempre que se fixam valores para as incógnitas não principais, determina–se uma solução particular para o sistema; mas, atendendo a que se podem atribuir valores arbitrários a estas incógnitas, o sistema admite um conjunto infinito de soluções. Diz-se que o sistema é indeterminado, com grau de indeterminação d, onde $d = n - m$ ($m =$r(A)).

A solução geral do sistema (S') pode ser apresentada na forma a seguir indicada.

$$\begin{cases} X_1 &= A_1^{-1}B - A_1^{-1}A_2X_2 \\ X_2 &= I_{n-m}X_2 \end{cases}$$

onde I_{n-m} é a matriz identidade de ordem $n - m$.

Exemplo 3.2 Considere-se o sistema sobre o corpo dos números reais $\mathbf{R}$

$$\begin{cases} 2x_1 & + x_2 & + x_3 & & +2x_5 & = & 1 \\ x_1 & & + x_3 & + x_4 & & = & 1 \\ & + x_2 & - x_3 & + x4 & - x_5 & = & 1 \end{cases}$$

Comecemos por determinar a característica da matriz dos coeficientes.

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

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

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

Embora o processo de condensação da matriz não esteja ainda completo, a última matriz determinada permite-nos concluir desde já que a característica da matriz é 3 (por troca das colunas 3 e 4 foi possível colocar um valor não nulo na posição de índice 3,3 da matriz).

O método de condensação da matriz permite constatar que as 3 primeiras colunas da matriz A original não são linearmente independentes; porém, a submatriz formada pelas colunas 1, 2 e 4 de A têm determinante não nulo.

Assim,

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ -1 & -1 \end{bmatrix},$$

$$\textit{incógnitas principais}\; \underbrace{X_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix}}\; \textit{incógnitas não principais}\; \underbrace{X_2 = \begin{bmatrix} x_3 \\ x_5 \end{bmatrix}}.$$

Passando para o segundo membro as incógnitas não principais, chega-se ao seguinte sistema (de Cramer)

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_5 \end{bmatrix}$$

Depois de conhecermos o valor do determinante da nova matriz dos coeficientes, é possível aplicar a regra de Cramer para chegar ao conjunto de soluções do sistema. Contudo, podemos concluir desde já que o sistema é indeterminado com grau de indeterminação $d = 2$ ($d = n - r = 5 - 3$).

$$\begin{vmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = 2.0.1 + 1.1.0 + 0.1.1 - 0.0.0 - 1.1.2 - 1.1.1 = -3$$

$$x_1 = \frac{\begin{vmatrix} 1 - x_3 - 2x_5 & 1 & 0 \\ 1 - x_3 & 0 & 1 \\ 1 + x_3 + x_5 & 1 & 1 \end{vmatrix}}{-3} = \frac{1}{3} - x_3 - x_5$$

$$x_2 = \frac{\begin{vmatrix} 2 & 1 - x_3 - 2x_5 & 0 \\ 1 & 1 - x_3 & 1 \\ 0 & 1 + x_3 + x_5 & 1 \end{vmatrix}}{-3} = \frac{1}{3} + x_3$$

$$x_4 = \frac{\begin{vmatrix} 2 & 1 & 1 - x_3 - 2x_5 \\ 1 & 0 & 1 - x_3 \\ 0 & 1 & 1 + x_3 + x_5 \end{vmatrix}}{-3} = \frac{2}{3} + x_5$$

A solução geral do sistema é pois

$$\begin{cases} x_1 = \frac{1}{3} - x_3 - x_5 \\ x_2 = \frac{1}{3} + x_3 \\ x_4 = \frac{2}{3} + x_5 \\ x_3 = x_3 \\ x_5 = x_5 \end{cases} . \diamondsuit$$

3^{o} Caso r(A)=r $<$ m

Considere-se um sistema (S) com m equações em n incógnitas. A matriz dos coeficientes, A, é uma matriz de ordem $m \times n$. Vamos também admitir que a característica de A é inferior ao número de equações.

Neste caso, procure-se uma submatriz de A, quadrada e de ordem igual à característica da matriz A, r=r(A), com determinante não nulo, e represente-se por Δ_r o determinante desta submatriz; o determinante Δ_r é vulgarmente chamado **determinante principal do sistema (S)**.

Uma escolha de Δ_r dá lugar a uma classificação das equações e das incógnitas em principais e não principais. As equações associadas às r linhas de Δ_r são as **equações principais**; as outras equações dizem-se **não principais**. Do mesmo modo, as incógnitas correspondentes às r colunas de Δ_r são **incógnitas principais,** enquanto as restantes se designam por **incógnitas não principais (ou livres)**. No caso particular de $r(A) = n$, tem-se que o n^o de incógnitas não principais é zero.

Sem perda de generalidade, pode-se considerar que as r equações principais correspondem às r primeiras linhas de A e as incógnitas principais às r primeiras colunas de A. Então

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

onde

$A_1 = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{bmatrix}$ é a submatriz dos coeficientes das incógnitas principais nas equações principais;

$A_2 = \begin{bmatrix} a_{1,r+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{r,r+1} & \cdots & a_{rn} \end{bmatrix}$ é a submatriz dos coeficientes das incógnitas não principais nas equações principais;

$A_3 = \begin{bmatrix} a_{r+1,1} & \cdots & a_{r+1,r} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mr} \end{bmatrix}$ é a submatriz dos coeficientes das incógnitas principais nas equações não principais;

$A_4 = \begin{bmatrix} a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & \vdots & \vdots \\ a_{m,r+1} & \cdots & a_{mn} \end{bmatrix}$ é a submatriz dos coeficientes das incógnitas não principais nas equações não principais.

Do mesmo modo,

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \text{com } X_1 = \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} \text{ e } X_2 = \begin{bmatrix} x_{r+1} \\ \vdots \\ x_n \end{bmatrix}$$

e

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \text{com } B_1 = \begin{bmatrix} b_1 \\ \vdots \\ b_r \end{bmatrix} \text{ e } B_2 = \begin{bmatrix} b_{r+1} \\ \vdots \\ b_m \end{bmatrix}.$$

O sistema $AX = B$ pode-se escrever na forma equivalente

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

ou

$$\begin{cases} A_1X_1 + A_2X_2 = B_1 \\ A_3X_1 + A_4X_2 = B2 \end{cases}$$

Vamos começar por considerar apenas as r equações principais, ou seja, a equação matricial

$$A_1X_1 + A_2X_2 = B_1.$$

Se em cada equação principal passarmos para o segundo membro as parcelas correspondentes às incógnitas não principais obtemos

$$(S') \qquad A_1X_1 = B_1 - A_2X_2.$$

O sistema (S') é um sistema de r equações nas mesmas condições do caso anterior, isto é, um sistema possível. No caso geral (S') será um sistema indeterminado com grau de indeterminação $d = n - r$; porém, se $d = 0$ $(A_2 = O, r = n)$ o sistema é determinado.

Uma vez que se pretende resolver o sistema inicial (S) e não apenas o sistema (S'), resta verificar se as soluções de (S') são também soluções de cada uma das equações não principais. Diz-se que se vai *verificar a compatibilidade de cada uma das equações não principais.*

A análise da compatibilidade de cada equação não principal consiste no cálculo do respectivo **determinante característico**.

Definição 3.2 - *Determinante característico de uma equação não principal*

O determinante característico de uma equação não principal é o determinante da matriz que se obtém juntando à matriz dos coeficientes das incógnitas principais nas equações principais, A_1, uma linha constituída pelos coeficientes das incógnitas principais dessa equação, e uma coluna formada pelos termos independentes das equações principais e o termo independente da equação não principal.

A cada uma das equações não principais do sistema está associado o respectivo determinante característico da forma:

$$\Delta_s = \begin{vmatrix} & & & b_1 \\ & A_1 & & \vdots \\ & & & b_r \\ a_{s1} & \cdots & a_{sr} & b_s \end{vmatrix}. \qquad (s = r+1, ..., m)$$

Teorema 3.2 (Rouché) - *É condição necessária e suficiente para que um sistema de equações seja possível que todos os determinantes característicos, caso existam, sejam nulos.*

Exemplo 3.3 Considere-se o sistema sobre o corpo dos números reais **R**

$$(S) \qquad \begin{cases} 2x_1 + x_2 = 5 \\ x_1 + x_3 = 4 \\ x_1 + x_2 - x_3 = 1 \\ x_2 + x_3 = 3 \\ 2x_1 - x_3 = 2 \end{cases}$$

com

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} e\ B = \begin{bmatrix} 5 \\ 4 \\ 1 \\ 3 \\ 2 \end{bmatrix}.$$

Determinação da característica de A

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

$$\rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow r(A) = 3$$

Uma vez que o número de equações m é 5, verifica-se que $r < m$ $(3 < 5)$.

A sequência de operações elementares aplicada para determinar a característica da matriz permite-nos chegar à conclusão de que o determinante da submatriz constituída pelas linhas 1, 2 e 4 de A é não nulo. Logo, escolhendo este determinante para determinante principal e reordenando as equações podemos escrever

$$\Delta_r = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -3 \neq 0$$

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \end{bmatrix}$$

$$X = X_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Como se pode observar, não existem incógnitas não principais.

O passo seguinte consiste na verificação da compatibilidade das equações não principais, pois caso um dos determinantes característicos seja não nulo podemos concluir de imediato que o sistema é impossível.

Determinante característico da terceira equação

$$\Delta_3 = \begin{vmatrix} 2 & 1 & 0 & 5 \\ 0 & 0 & 1 & 4 \\ 0 & 1 & 1 & 3 \\ 1 & 1 & -1 & 1 \end{vmatrix} = 0$$

Determinante característico da quinta equação

$$\Delta_5 = \begin{vmatrix} 2 & 1 & 0 & 5 \\ 0 & 0 & 1 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & 0 & -1 & 2 \end{vmatrix} = 0$$

Como $\Delta_3 = \Delta_5 = 0$ o sistema é possível, e, atendendo a que todas as incógnitas são principais, é também determinado. Finalmente, aplicando a regra de Cramer para encontrar as soluções do sistema (S') constituído pelas equações principais chegamos aos resultados que a seguir se indicam.

$$x_1 = \frac{\begin{vmatrix} 5 & 1 & 0 \\ 4 & 0 & 1 \\ 3 & 1 & 1 \end{vmatrix}}{-3} = 2; \qquad x_2 = \frac{\begin{vmatrix} 2 & 5 & 0 \\ 0 & 4 & 1 \\ 0 & 3 & 1 \end{vmatrix}}{-3} = 1; \qquad x_3 = \frac{\begin{vmatrix} 2 & 1 & 5 \\ 0 & 0 & 4 \\ 0 & 1 & 3 \end{vmatrix}}{-3} = 2.\diamondsuit$$

3.4 Resolução de sistemas de equações lineares usando operações elementares

Nesta secção vamos estudar outro método de resolução de sistemas de equações lineares usando operações elementares sobre as linhas da matriz completa do sistema. A essência desta técnica reside na transformação do sistema inicial num sistema equivalente, isto é, com o mesmo conjunto de soluções, mas mais simples de resolver.

O teorema a seguir apresentado e o seu corolário indicam um processo simples para a obtenção de sistemas equivalentes.

Teorema 3.3 - *Seja* $(S) : AX = B$ *um sistema de* m *equações em* n *incógnitas, e seja* C *uma matriz invertível de ordem* m. *Então, o sistema* $(S') : (CA)X = CB$ *é equivalente a* (S).

Prova: Sejam K e K' respectivamente os conjuntos de soluções de (S) e (S').

Se $w \in K$ então $Aw = B$. Porém $CAw = CB$, logo $w \in K'$. Portanto, $K \subseteq K'$.

Por outro lado, se $w \in K'$ então $CAw = CB$.

Então, $Aw = C^{-1}(CAw) = C^{-1}(CB) = B$; logo, $w \in K$ e $K' \subseteq K$.

Uma vez que $K \subseteq K'$ e $K' \subseteq K$, então $K = K'$. $\clubsuit$

Corolário 3.3.1 - *Seja* $AX = B$ *um sistema de* m *equações em* n *incógnitas. Se* $[A'|B']$ *é obtida de* $[A|B]$ *através de um número finito de operações elementares sobre as linhas, então o sistema* $A'X = B'$ *é equivalente ao sistema original.*

Prova: Vamos admitir que $[A'|B']$ é obtida de $[A|B]$ através de operações elementares sobre as linhas. Estas podem ser realizadas pela multiplicação por matrizes elementares de ordem $m \times m$, $E_1, ..., E_p$. Seja $C = E_p...E_1$. Então

$$[A'|B'] = C\,[A|B] = [CA|CB]$$

Dado que $E_i (i = 1...p)$ são matrizes invertíveis, C é também uma matriz invertível. Então $A^{'} = CA$ e $B^{'} = CB$.

Logo, pelo teorema 3.3 o sistema $A'X = B'$ é equivalente ao sistema $AX = B$. ♣

O procedimento que vamos usar para resolver sistemas de equações é a seguir descrito.

Partindo da equação matricial $AX = B$ executamos os seguintes passos:

1. Construção da matriz aumentada $[A|B]$;

2. Por aplicação de operações elementares às linhas desta matriz aumentada, transformamos $[A|B]$ numa nova matriz $[F|K]$; esta nova matriz apresenta características especiais dado que é uma **matriz em formato de linhas escalonadas**. O novo sistema é equivalente ao original e pode ser muito facilmente resolvido.

Definição 3.3 - *Matriz em formato de linhas escalonadas*

Diz-se que uma matriz se apresenta em formato de linhas escalonadas quando são satisfeitas as seguintes condições:

(a) qualquer linha contendo uma entrada não nula precede as linhas cujas entradas são todas nulas;

(b) a primeira entrada não nula em cada linha é a única entrada não nula na respectiva coluna;

(c) a primeira entrada não nula em cada linha é 1 e ocorre numa coluna à direita do primeiro valor 1 em qualquer linha anterior.

É possível mostrar que uma matriz só tem uma maneira de ser apresentada em formato de linhas escalonadas; contudo, podem ser usadas diversas sequências de operações elementares para representar a matriz neste formato especial.

O procedimento usado para transformar a matriz completa de um sistema de equações numa matriz em formato de linhas escalonadas é designado por **eliminação Gaussiana**, e nele podem ser distinguidos dois passos:

1. no primeiro passo, a matriz completa do sistema é transformada numa matriz triangular superior na qual a primeira entrada não nula de cada linha tem o valor 1, e ocorre numa coluna à direita da primeira entrada não nula de cada uma das linhas precedentes;

2. no segundo passo, a matriz triangular é transformada numa matriz com as linhas escalonadas.

Para ilustrar este método de resolução de sistemas de equações, vamos resolver o sistema a seguir indicado.

Exemplo 3.4

$$\begin{cases} 2x_1 + 3x_2 + x_3 + 4x_4 - 9x_5 = 17 \\ x_1 + x_2 + x_3 + x_4 - 3x_5 = 6 \\ x_1 + x_2 + x_3 + 2x_4 - 5x_5 = 8 \\ 2x_1 + 2x_2 + 2x_3 + 3x_4 - 8x_5 = 14 \end{cases}$$

A matriz completa do sistema é

$$\begin{bmatrix} 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 1 & -3 & 6 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{bmatrix}$$

A aplicação de eliminação Gaussiana a esta matriz gera a seguinte sequência de matrizes

1º Passo:

$$\begin{bmatrix} 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 1 & -3 & 6 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & -3 & 6 \\ 2 & 3 & 1 & 4 & -9 & 17 \\ 1 & 1 & 1 & 2 & -5 & 8 \\ 2 & 2 & 2 & 3 & -8 & 14 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 & -2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow$$

2º Passo:

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

O sistema de equações correspondente a esta última matriz é

$$\begin{cases} x_1 + 2x_3 - 2x_5 = 3 \\ x_2 - \ x_3 + x_5 = 1 \\ x_4 - 2x_5 = 2 \end{cases}$$

A última linha é ignorada uma vez que é constituída apenas por zeros.

A resolução de um sistema de equações usando este método permite dividir as incógnitas x_i em dois conjuntos. O primeiro conjunto, das *variáveis principais*, contém as variáveis colocadas mais à esquerda nas equações, que no exemplo anterior são x_1, x_2 e x_4. O segundo conjunto é constituído pelas restantes variáveis, neste caso x_3 e x_5; são as *variáveis não principais.*

Passando as variáveis não principais para o segundo membro, chegamos finalmente à solução geral do sistema

$$\begin{cases} x_1 = 3 - 2x_3 + 2x_5 \\ x_2 = 1 + \ x_3 - x_5 \\ x_4 = 2 + 2x_5 \\ x_3 = x_3 \\ x_5 = x_5 \end{cases}$$

A cada uma das variáveis não principais é ainda possível atribuir um valor paramétrico $t_1, t_2 ...$, obtendo-se

$$\begin{cases} x_1 = 3 - 2t_1 + 2t_2 \\ x_2 = 1 + \ t_1 - t_2 \\ x_4 = 2 + 2t_2 \qquad\qquad t_1, t_2 \in R \\ x_3 = t_1 \\ x_5 = t_2 \end{cases}$$

Uma solução arbitrária do sistema (S) pode apresentar-se na forma

$$s = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 3-2t_1+2t_2 \\ 1+\ t_1-t_2 \\ t_1 \\ 2+2t_2 \\ t_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, t_1, t_2 \in \mathbf{R}.\diamondsuit$$

O procedimento que acabámos de apresentar também permite detectar facilmente se um sistema de equações é impossível. Com efeito, quando o sistema não tem soluções, como resultado da redução da matriz completa para o formato de linhas escalonadas surge pelo menos uma linha em que todas as entradas nulas excepto a entrada da última coluna (coluna dos termos independentes).

Os factos que acabamos de expôr são estabelecidos no seguinte teorema.

Teorema 3.4 - *Seja $AX = B$ um sistema de equações lineares. O sistema tem pelo menos uma solução se e só se r(A)=r(A|B).*

Exemplo 3.5 Resolução do sistema de equações

$$\begin{cases} \qquad\quad 2x_2 + 4\,x_3 = 2 \\ 2x_1 + 4x_2 + 2x_3 = 3 \\ 3x_1 + 3x_2 + x_3 = 1 \\ 3x_1 + x_2 - 3x_3 = 1 \end{cases}$$

Matriz completa $[A|B]$

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

1º Passo da eliminação Gaussiana

$$\begin{bmatrix} 0 & 2 & 4 & 2 \\ 2 & 4 & 2 & 3 \\ 3 & 3 & 1 & 1 \\ 3 & 1 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 3/2 \\ 0 & 2 & 4 & 2 \\ 3 & 3 & 1 & 1 \\ 3 & 1 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 3/2 \\ 0 & 2 & 4 & 2 \\ 0 & -3 & -2 & -7/2 \\ 0 & -5 & -4 & -7/2 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 2 & 1 & 3/2 \\ 0 & 1 & 2 & 1 \\ 0 & -3 & -2 & -7/2 \\ 0 & -5 & -4 & -7/2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 3/2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 4 & -1/2 \\ 0 & 0 & 6 & 3/2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 3/2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -1/8 \\ 0 & 0 & 6 & 3/2 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 2 & 1 & 3/2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -1/8 \\ 0 & 0 & 0 & 6/8 \end{bmatrix} \leftarrow \textit{sistema impossível}$$

O sistema equivalente correspondente à última matriz que resultou do primeiro passo de eliminação gaussiana é impossível pois contém a equação

$$0.x_1 + 0.x_2 + 0.x_3 = 6/8$$

que, obviamente, não tem solução.$\diamondsuit$

3.5 Resolução de sistemas homogéneos

Definição 3.4 - *Sistema homogéneo*
Um sistema $AX = B$ de m equações em n incógnitas é homogéneo se $B = O$.

Qualquer sistema homogéneo tem pelo menos uma solução, a solução nula

$$s = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Esta solução também é chamada **solução trivial do sistema homogéneo**.

É, no entanto, importante conhecer em que condições um sistema homogéneo tem soluções não nulas.

Teorema 3.5 - *O sistema homogéneo $AX = O$ tem soluções não nulas se e só se r(A)= $r < n$, isto é, se e só se a característica da matriz dos coeficientes for inferior ao número de incógnitas.*

Prova: Se $r = n$ o sistema é possível e determinado, portanto só tem uma solução que é, obrigatoriamente, a solução trivial.

Quando $r < n$ o sistema é indeterminado, logo tem um número infinito de soluções. $\clubsuit$

Definição 3.5 - *Conjunto fundamental de soluções de um sistema homogéneo*

Um conjunto $X_{1,}X_2, ..., X_k$ *de soluções linearmente independentes do sistema* $AX = O$ *é um conjunto fundamental de soluções se qualquer solução do sistema é uma combinação linear das soluções* $X_{1,}X_2, ..., X_k$.

O número de soluções em qualquer conjunto fundamental é igual ao grau de indeterminação do sistema.

Relação entre as soluções de um sistema e as soluções do sistema homogéneo associado

É fácil verificar que a diferença de duas soluções do sistema não homogéneo $AX = B$ é uma solução do sistema homogéneo $AX = O$. Com efeito, se X_1 e X_2 forem soluções de $AX = B$, então

$$AX_1 = B,\ AX_2 = B \Rightarrow A(X_1 - X_2) = O.$$

Por outro lado, se X é uma solução qualquer do sistema homogéneo e se X_0 é uma solução particular do sistema não homogéneo, $X + X_0$ é uma solução do sistema não homogéneo

$$AX_0 = B, AX = O \Rightarrow A(X_0 + X) = B.$$

Assim, a solução geral do sistema $AX = B$ pode ser obtida somando uma solução particular deste sistema com a solução geral do sistema homogéneo associado.

Exemplo 3.6 Considere-se o sistema de equações sobre $\mathbf{R}$

$$\begin{cases} x_1 + \ \ x_2 + \ \ x_3 + \ \ \ x_4 = 0 \\ x_1 + 3x_2 + 2x_3 + 4\,x_4 = 0 \\ 2x_1 \qquad\ \ + \ \ x_3 - \ \ \ x_4 = 0 \end{cases}$$

No caso de sistemas homogéneos é desnecessário trabalhar com a matriz completa do sistema uma vez que a coluna dos termos independentes é constituída apenas por valores nulos que não são alterados pelas operações elementares sobre as linhas. Portanto basta aplicar eliminação gaussiana à matriz dos coeficientes.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 2 & 4 \\ 2 & 0 & 1 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 3 \\ 0 & -2 & -1 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1/2 & 3/2 \\ 0 & -2 & -1 & -3 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1/2 & -1/2 \\ 0 & 1 & 1/2 & 3/2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{cases} x_1 = -\frac{1}{2}x_3 + \frac{1}{2}x_4 \\ x_2 = -\frac{1}{2}x_3 - \frac{3}{2}x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases}$$

ou

$$\begin{cases} x_1 = -\frac{1}{2}t_1 + \frac{1}{2}t_2 \\ x_2 = -\frac{1}{2}t_1 - \frac{3}{2}t_2 \\ x_3 = t_1 \\ x_4 = t_2 \end{cases},$$

O sistema tem grau de indeterminação $d = 4 - 2 = 2$ e a sua solução geral é

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t_1 \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 1/2 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}$$

Se se escolher $\begin{cases} t_1 = 1 \\ t_2 = 0 \end{cases}$ e $\begin{cases} t_1 = 0 \\ t_2 = 1 \end{cases}$ obtemos um conjunto fundamental de soluções, pois

$$\begin{cases} x_1 = -\frac{1}{2} \\ x_2 = -\frac{1}{2} \\ x_3 = 1 \\ x_4 = 0 \end{cases} \text{ e } \begin{cases} x_1 = \frac{1}{2} \\ x_2 = -\frac{3}{2} \\ x_3 = 0 \\ x_4 = 1 \end{cases}$$

são duas soluções linearmente independentes.

Se se arbitrar $\begin{cases} t_1 = 1 \\ t_2 = 1 \end{cases}$ e $\begin{cases} t_1 = 3 \\ t_2 = 1 \end{cases}$ obtemos outro conjunto fundamental de soluções formado por

$$\begin{cases} x_1 = -\frac{1}{2} \\ x_2 = -\frac{1}{2} \\ x_3 = 1 \\ x_4 = 0 \end{cases} \text{ e } \begin{cases} x_1 = \frac{1}{2} \\ x_2 = -\frac{3}{2} \\ x_3 = 0 \\ x_4 = 1 \end{cases} \diamondsuit.$$

Consideremos agora o sistema $AX = B$ associado ao sistema homogéneo indicado no exemplo anterior, com $B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

Matriz completa do sistema $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 0 & 1 & -1 & 2 \end{bmatrix}$

Resolvendo o sistema obtemos

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 0 & 1 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 3 & 1 \\ 0 & -2 & -1 & -3 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1/2 & 3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 0 & 1/2 & -1/2 & 1 \\ 0 & 1 & 1/2 & 3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{cases} x_1 = 1 - \frac{1}{2}x_3 + \frac{1}{2}x_4 \\ x_2 = -\frac{1}{2}x_3 - \frac{3}{2}x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases}$$

O resultado final é

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 1/2 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}$$

ou seja, a solução geral do sistema não homogéneo é a soma de uma solução particular

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

com a solução geral do sistema homogéneo

$$t_1 \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 1/2 \\ -3/2 \\ 0 \\ 1 \end{bmatrix} . \diamondsuit$$

Capítulo 4

Espaços vectoriais

4.1 Introdução

É comum no tratamento de muitas grandezas físicas a necessidade de considerar tanto um valor de amplitude como uma direcção. Como exemplos podemos tomar grandezas como força, velocidade e aceleração. Este tipo de entidades que associam amplitude e direcção são designadas *vectores* e representam-se habitualmente como setas nas quais o comprimento da seta denota a amplitude do vector e a direcção da seta a direcção do vector. Em geral só a amplitude e direcção de um vector são relevantes, e consideram-se iguais vectores de idêntica amplitude e direcção independentemente da sua posição.

Intuitivamente, é fácil concluir que quando dois vectores actuam simultaneamente num ponto a amplitude do vector resultante (o vector soma dos dois vectores) não tem de ser a soma das amplitudes dos vectores originais. Tomando um exemplo, se um nadador nada contra a corrente a uma taxa de 2 km/h contra uma corrente de 1km/h ele não progride a uma velocidade de 3 km/h; uma vez que o movimento do nadador e da corrente se opõem, ele apenas se deslocará 1 km/h para montante. No entanto, se ele estiver a nadar a favor da corrente progredirá à velocidade de 3 km/h.

Os vectores são adicionados usando a *regra do paralelogramo* ilustrada na figura a seguir.

Sendo os lados opostos do paralelogramo paralelos e de comprimentos iguais, a extremidade Q da seta que representa $x + y$ pode também ser obtida aplicando x em P e y na extremidade de x. Da mesma forma se pode obter a extremidade de $x + y$ aplicando primeiro y em P e x na extremidade de y.

Além da adição, outra operação básica sobre vectores é a sua ampliação ou

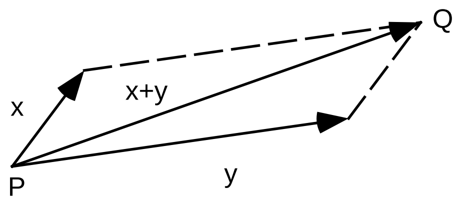

Figura 4.1: Regra do paralelogramo

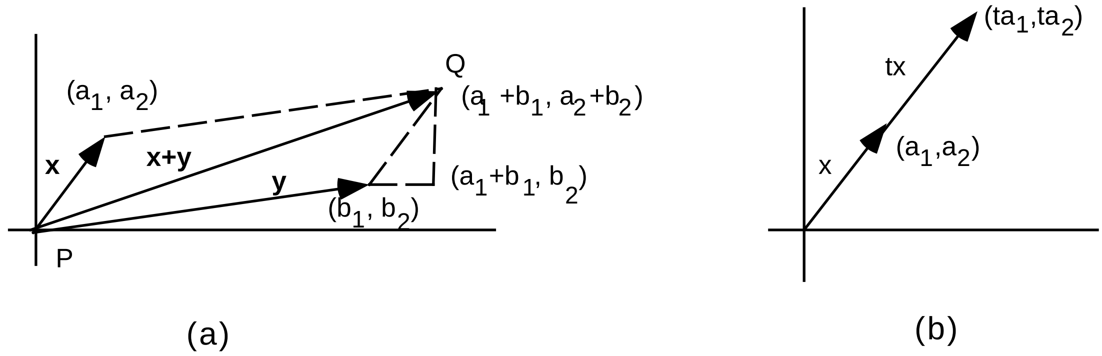

Figura 4.2: Adição e multiplicação escalar de vectores.

contracção sem mudança de direcção. Esta operação designa-se *multiplicação escalar* e consiste na multiplicação do vector por um número real. Sendo o vector x representado por uma seta, então para qualquer número real $t \geq 0$ o vector tx será representado por uma seta com a mesma direcção da que representa x, mas tendo comprimento t vezes o comprimento da seta que representa x. Sendo $t < 0$, o vector tx será representado por uma seta de direcção oposta à de x e de comprimento $|t|$ vezes o da seta que representa x. Dois vectores não nulos x e y dizem-se paralelos se $y = tx$ para algum número real não nulo t.

Para descrever a multiplicação escalar algebricamente, representemos novamente x na origem de um sistema de coordenadas. Tendo o extremo de x coordenadas (a_1, a_2), então as coordenadas do extremo de tx serão (ta_1, ta_2), como se representa na figura 4.2.

As descrições algébricas da adição de vectores e multiplicação escalar resultam nas seguintes propriedades para vectores arbitrários x, y e z, e números reais arbitrários a e b:

1. $x + y = y + x$.

2. $(x + y) + z = x + (y + z)$.

3. Existe um vector designado 0 tal que $x + 0 = x$ para cada vector x.
4. Para cada vector x existe um vector y tal que $x + y = 0$.
5. $1x = x$.
6. $(ab)x = a(bx)$.
7. $a(x + y) = ax + ay$.
8. $(a + b)x = ax + bx$.

É fácil verificar que as 8 propriedades enunciadas acima, bem como as interpretações geométricas da adição de vectores e multiplicação escalar, são ainda válidas para vectores no espaço tridimensional. Vamos usá-los para escrever equações de rectas e planos no espaço.

Consideremos primeiro a equação de uma recta no espaço que passa por 2 pontos distintos P e Q. Seja O a origem de um sistema de coordenadas e u e v os vectores que começam em O e terminam respectivamente em P e Q. Se w designar o vector começando em P e terminando em Q, então pela adição temos que $u + w = v$, e portanto $w = v - u$, em que $-u$ designa o vector $(-1)u$ (ver figura 4.3 em que $OPQR$ é um paralelogramo). Como um vector que seja múltiplo escalar de w é paralelo a w, mas possivelmente de comprimento diferente do deste, qualquer ponto na recta que une P e Q pode ser obtido como extremo de um vector começando em P e da forma tw para algum real t. Inversamente, o extremo de todo o vector da forma tw que começa em P fica na recta ligando P e Q. Então a equação da recta passando por P e Q é $x = u + tw = u + t(v - u)$, em que t é um n úmero real e x um ponto arbitrário sobre a recta. Notar ainda que o extremo R do vector $v - u$ na figura 4.3 tem coordenadas iguais à diferença das coordenadas de Q e P.

Exemplo 4.1 Vamos determinar a equação da recta que passa pelos pontos P e Q de coordenadas $(-2, 0, 1)$ e $(4, 5, 3)$, respectivamente. O extremo R do vector aplicado na origem e com a direcção do vector começando em P e terminando em Q tem coordenadas $(4, 5, 3) - (-2, 0, 1) = (6, 5, 2)$. A equação da recta é então

$$x = (-2, 0, 1) + t(6, 5, 2). \diamondsuit$$

Sejam agora P, Q e R três pontos não colineares no espaço. Estes pontos determinam um plano único, cuja equação pode ser encontrada usando as observações anteriores acerca de vectores. Sejam u e v os vectores começando em P e terminando

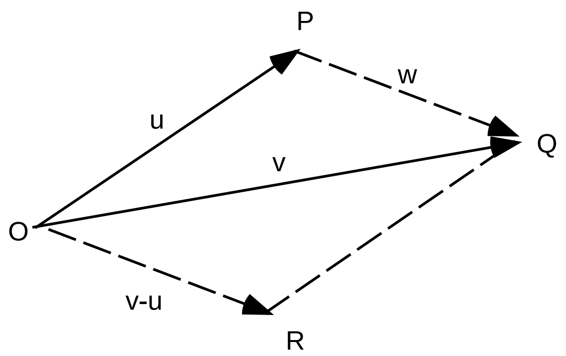

Figura 4.3: Equação da recta que passa por 2 pontos

em Q e R, respectivamente. Notar que qualquer ponto no plano que contém P, Q e R é o extremo S de um vector x que começa em P e tem a forma $t_1u + t_2v$ para algum par de valores reais t_1 e t_2. O extremo de t_1u será o ponto de intersecção da recta que passa por P e Q com a que passa por S e é paralela à recta que passa por P e R (ver figura 4.4). De forma semelhante se pode localizar o extremo de t_2v. Podemos ver também que para quaisquer números reais t_1 e t_2, $t_1u + t_2v$ é um vector no plano que contém P, Q e R. Donde se conclui que a equação do plano que contém P, Q e R é

$$x = P + t_1u + t_2v$$

em que t_1 e t_2 são números reais arbitrários e x um ponto arbitrário no plano.

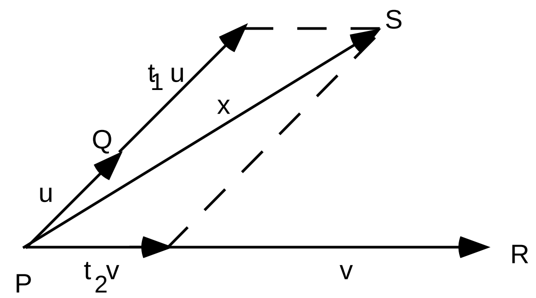

Figura 4.4: Equação do plano que passa por 3 pontos

Exemplo 4.2 Sejam P, Q e R os pontos de coordenadas $(1, 0, 2)$, $(-3, -2, 4)$ e $(1, 8, -5)$, respectivamente. O extremo do vector colocado na origem e com o comprimento e direcção do vector começando em P e terminando em Q é

$$(-3, -2, 4) - (1, 0, 2) = (-4, -2, 2).$$

Do mesmo modo, o extremo do vector colocado na origem e com o comprimento e direcção do vector começando em P e terminando em R é

$$(1, 8, -5) - (1, 0, 2) = (0, 8, -7).$$

A equação do plano que passa nos 3 pontos é então:

$$x = (1, 0, 2) + t_1(-4, -2, 2) + t_2((0, 8, -7). \ \diamondsuit$$

Qualquer estrutura matemática que goze das 8 propriedades enunciadas anteriormente é designada "espaço vectorial". Vamos de seguida definir formalmente um espaço vectorial e considerar exemplos diversos.

4.2 Espaços vectoriais

É interessante verificar que para entidades tão diversas como as forças que actuam sobre um corpo em movimento e os polinómios de coeficientes reais se podem definir as operações de adição e multiplicação escalar, para as quais se verificam as propriedades enunciadas atrás. Parece portanto natural abstrair essas propriedades, e daí decorre a definição de espaço vectorial.

Definição 4.1 - *Espaço vectorial*
Um espaço vectorial (ou espaço linear) V sobre um corpo F *é um conjunto sobre o qual estão definidas duas operações (*adição *e* multiplicação escalar*) tais que para cada par de elementos x e y de V existe um e um só elemento $x + y$ em V, e para cada elemento a de F e cada elemento x de V existe um e um só elemento ax em V, para o qual se verifica:*

(EV 1) *Para todo o x, y em V, $x + y = y + x$ (comutatividade da adição).*

(EV 2) *Para todo o x, y, z em V, $(x + y) + z = x + (y + z)$ (associatividade da adição).*

(EV 3) *Existe um elemento 0 em V tal que $x + 0 = x$ para todo o x de V.*

(EV 4) *Para cada elemento x de V existe um elemento y de V tal que $x + y = 0$.*

(EV 5) *Para cada elemento x de V, $1x = x$.*

(EV 6) *Para cada par a, b de elementos de F e cada elemento x de V, $(ab)x = a(bx)$.*

(EV 7) *Para cada a de F e cada par x, y de V, $a(x+y) = ax + ay$.*

(EV 8) *Para cada par a, b de elementos de F e cada elemento x de V,* $(a+b)x = ax + bx$.

Os elementos $x + y$ e ax são designados soma *de x e y e* produto *de a e x respectivamente.*

Os elementos do corpo F são designados *escalares* e os elementos do espaço vectorial V *vectores*. Notar que deixamos de restringir o uso da designação **vector** às entidades físicas usadas na secção anterior; o termo será usado para qualquer elemento de um espaço vectorial.

Vamos considerar exemplos de espaços vectoriais aos quais nos referiremos na sequência do texto. Para cada espaço vectorial teremos de especificar o corpo de escalares respectivo, e a definição das operações de soma e multiplicação escalar.

Exemplo 4.3 O conjunto de todos os n-tuplos de elementos de um corpo F forma um espaço vectorial, que denotaremos F^n, sob as operações de adição coordenada a coordenada e multiplicação escalar a seguir definidas. Sendo $x = (a_1, \ldots, a_n) \in F^n, y = (b_1, \ldots, b_n) \in F^n$ e $c \in F$, então

$$x + y = (a_1 + b_1, \ldots, a_n + b_n) \text{ e } cx = (ca_1, \ldots, ca_n)$$

Os elementos de F^n são frequentemente escritos na forma de *vectores coluna*:

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

em vez da forma *vector linha* $[a_1, \ldots, a_n]$. Como os 1-tuplos de elementos de F podem ser considerados como elementos de F eles próprios, escreve-se F e não F^1 o espaço vectorial dos 1-tuplos de F. $\diamondsuit$

Exemplo 4.4 O conjunto de todas as matrizes $m \times n$ com elementos num corpo F constitui um espaço vectorial, denotado $M_{m\times n}(F)$, sob as operações de adição e multiplicação por um escalar definidas para matrizes. $\diamondsuit$

Exemplo 4.5 Seja S um conjunto não vazio e F um corpo, e seja $\mathcal{F}(S, F)$ o conjunto de todas as funções de S em F. Define-se que dois elementos f e g de $\mathcal{F}(S, F)$ são iguais quando $f(s) = g(s)$ para cada $s \in S$. O conjunto

$\mathcal{F}(S, F)$ é um espaço vectorial sob as operações de adição e multiplicação escalar definidas para $f, g \in \mathcal{F}(S, F)$ e $c \in F$ como

$$(f+g)(s) = f(s) + g(s) \text{ e } (cf)(s) = c[f(s)]$$

para cada $s \in S$. Notar que estas são as definições habituais das operações de adição e multiplicação escalar usadas na álgebra e na análise matemática. $\diamondsuit$

Exemplo 4.6 O conjunto de todos os polinómios com coeficientes de um corpo F é um espaço vectorial, que denotaremos $P(F)$, sob as operações seguintes. Sendo

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

e

$$g(x) = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0$$

de $P(F)$ e $c \in F$,

$$(f+g)(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_0 + b_0)$$

e

$$cf(x) = ca_n x^n + ca_{n-1} x^{n-1} + \cdots + ca_0. \ \diamondsuit$$

Exemplo 4.7 Sendo F um corpo, define-se *sequência* em F como uma função σ dos inteiros positivos em F. A sequência σ tal que $\sigma(n) = a_n$ será denotada $\{a_n\}$. Seja V o conjunto de todas as sequências $\{a_n\}$em F que têm somente um número finito de termos não nulos a_n. Se $\{a_n\}$e $\{b_n\}$ são sequências de V e $t \in F$, então $\{a_n\} + \{b_n\}$ é a sequência $\{c_n\}$ de V tal que $c_n = a_n + b_n$, $(n = 1, 2, \ldots)$ e $t\{a_n\}$ é a sequência $\{d_n\}$ de V tal que $d_n = ta_n$, $(n = 1, 2, \ldots)$. Com estas definições de soma e multiplicação escalar, V é um espaço vectorial. $\diamondsuit$

Os exemplos seguintes são de conjuntos onde se definem uma adição e uma multiplicação escalar, mas que não são espaços vectoriais.

Exemplo 4.8 Seja $S = \{(a_1, a_2) : a_1, a_2 \in R\}$. Para $(a_1, a_2), (b_1, b_2) \in S$ e $c \in R$, defina-se

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 - b_2) \text{ e } c(a_1, a_2) = (ca_1, ca_2).$$

Não se verificam as propriedades (EV1), (EV2) e (EV8), logo S com estas operações não é um espaço vectorial. $\diamondsuit$

Exemplo 4.9 Seja S como no exemplo anterior. Para $(a_1, a_2), (b_1, b_2) \in S$ e $c \in R$ defina-se

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, 0) \text{ e } c(a_1, a_2) = (ca_1, 0).$$

S não é um espaço vectorial com estas operações uma vez que não se verificam (EV3), (EV4) e (EV5). $\diamondsuit$

Passamos agora a enunciar algumas consequências elementares da definição de espaço vectorial.

Proposição 4.1 (Lei do cancelamento na adição de vectores.) *Sendo x, y e z elementos de um espaço vectorial V tais que $x + z = y + z$, então $x = y$.*

Prova: Existe um elemento v em V tal que $z + v = 0$ (EV4). Então

$$\begin{aligned} x &= x + 0 = x + (z + v) = (x + z) + v = \\ &= (y + z) + v = y + (z + v) = y + 0 = y \end{aligned}$$

por (EV2) e (EV3).$\clubsuit$

Corolário 4.0.1 *O vector 0 descrito em (EV3) é único.*

Corolário 4.0.2 *O vector y descrito em (EV4) é único.*

O vector 0 em (EV3) é chamado *vector zero* de V, e o vector y de (EV4) (o vector único tal que $x + y = 0$) é chamado o *simétrico* de x e denotado $-x$.

O resultado seguinte contém propriedades elementares da multiplicação escalar.

Proposição 4.2 *Em todo o espaço vectorial V as seguintes afirmações são verdadeiras:*

(a) $0x = O$ *para cada* $x \in V$.

(b) $(-a)x = -(ax) = a(-x)$ *para cada* $a \in F$ *e cada* $x \in V$.

(c) $a0 = 0$ *para cada* $a \in F$.

Prova: (a) Por (EV8), (EV1) e (EV3) tem-se que

$$0x + 0x = (0 + 0)x = 0x = O + 0x.$$

Então $0x = O$ pela Proposição 4.1.(b) O elemento $-(ax)$ é o elemento único de V tal que $ax + [-(ax)] = 0$. Então se $ax + (-a)x = 0$, o Corolário 4.0.2 implica que $(-a)x = -(ax)$. Mas, por (EV8),

$$ax + (-a)x = [a + (-a)]x = 0x = 0$$

por (a). Então $(-a)x = -(ax)$.(c) Prova semelhante à de (a).$\clubsuit$

4.3 Subespaços

Sendo o espaço vectorial uma estrutura algébrica tem interesse identificar subconjuntos cuja estrutura é a mesma do conjunto inicial.

Definição 4.2 - *Subespaço vectorial*
Um subconjunto W de um espaço vectorial V sobre um corpo F é designado subespaço *de V se W é um espaço vectorial sobre F com as operações de adição e multiplicação escalar definidas para V.*

Em qualquer espaço vectorial V verifica-se que V e $\{0\}$ são subespaços. O último é designado *subespaço zero* de V. Não é necessário verificar todas as condições da definição de um espaço vectorial para mostrar que um subconjunto W de um espaço vectorial V é um subespaço. Algumas das condições são trivialmente verificadas pelo facto de os elementos de W serem elementos de V.

Teorema 4.1 *Seja V um espaço vectorial e W um subconjunto de V. Então W é um subespaço de V se e só se as 3 condições seguintes se verificarem para as operações definidas em V:*

(a) $0 \in W$.

(b) $x + y \in W$ *para todo* $x \in W$ *e* $y \in W$.

(c) $ax \in W$ *para todo* $a \in F$ *e* $x \in W$.

As condições expostas neste teorema fornecem um método simples para determinar se um subconjunto de um espaço vectorial é um subespaço.

Exemplo 4.10 Seja o conjunto W das matrizes simétricas ($M^t = M$), subconjunto de$M_{n\times n}(F)$. W é um subespaço de $M_{n\times n}(F)$ porque:

(a) A matriz zero é igual à sua transposta logo pertence a W.

(b) Se $A \in W$ e $B \in W$, então $A = A^t$ e $B = B^t$. Como $(A + B)^t = A^t + B^t$ então $(A + B)^t = A + B$ e portanto $A + B \in W$.

(c) Se $A \in W$ então $A = A^t$. Logo para qualquer $a \in F$, $(aA)^t = aA^t = aA$ donde $aA \in W$. $\diamondsuit$

Exemplo 4.11 O conjunto de todas as matrizes diagonais em $M_{n\times n}(F)$ é um subespaço de $M_{n\times n}(F)$. $\diamondsuit$

Exemplo 4.12 Seja n um inteiro não negativo e $P_n(F)$ o conjunto de todos os polinómios em $P(F)$ com grau menor ou igual a n. $P_n(F)$ é um subespaço de $P(F)$. $\diamondsuit$

Exemplo 4.13 O conjunto $C(R)$ que contém todas as funções reais de variável real contínuas é um subespaço de $F(R, R)$, o espaço de funções definido na secção anterior. $\diamondsuit$

Exemplo 4.14 O conjunto das matrizes $n \times n$ com traço igual a zero é um subespaço de $M_{n\times n}(F)$. $\diamondsuit$

Exemplo 4.15 O conjunto das matrizes $M_{m\times n}(F)$ com entradas não negativas não é um subespaço de $M_{m\times n}(F)$ pois não verifica a condição (c) do Teorema 4.1. $\diamondsuit$

Podemos combinar subespaços de um espaço vectorial para obter novos subespaços.

Teorema 4.2 *Qualquer intersecção de subespaços de um espaço vectorial V é ainda um subespaço de V.*

Prova: Seja C uma colecção de subespaços de V, e W a intersecção de todos os subespaços em C. Uma vez que todos os subespaços contêm o vector zero, $0 \in W$. Seja $a \in F$ e $x, y \in W$; x e y são portanto elementos de cada subespaço em C. Então $x + y$ e ax são elementos de cada subespaço em C (porque a soma de vectores de um subespaço e o produto de um vector de um subespaço por um escalar pertencem ainda ao subespaço). Então $x + y \in W$ e $ax \in W$, e W é um subespaço de V pelo Teorema 4.1. $\clubsuit$

Será interessante verificar se um resultado semelhante se verifica para a união de subespaços. Vê-se que a união de subespaços verifica as condições (a) e (c) do Teorema 4.1 mas que a condição (b) não se verifica sempre. De facto, a união de 2 subespaços é um subespaço se e só se um dos subespaços é subconjunto do outro. Há no entanto outras formas de combinar subespaços de forma a obter subespaços maiores.

Definição 4.3 - *Soma de subespaços*

Sendo S_1 e S_2 subconjuntos não vazios de um espaço vectorial V, então a soma *de S_1 e S_2, denotada $S_1 + S_2$ é o conjunto $\{x + y : x \in S_1 \text{ e } y \in S_2\}$.*

Definição 4.4 *Um espaço vectorial V diz-se a* soma directa *de W_1 e W_2, denotado $V = W_1 \oplus W_2$ se W_1 e W_2 são subespaços de V tais que $W_1 \cap W_2 = \{0\}$ e $W_1 + W_2 = V$.*

4.4 Combinações Lineares e Sistemas de Equações Lineares

Vimos anteriormente que a equação de um plano passando por três pontos não colineares P, Q e R no espaço é $x = P + t_1 u + t_2 v$, onde u e v denotam vectores que começam em P e terminam em Q e R respectivamente e t_1 e t_2 denotam números reais arbitrários. Um caso particular importante ocorre quando P é a origem das coordenadas. A equação do plano simplifica-se neste caso para $x = t_1 u + t_2 v$, e o conjunto de pontos neste plano é um subespaço de R^3. As expressões da forma $t_1 u + t_2 v$ têm grande importância na teoria dos espaços vectoriais. A definição seguinte é uma generalização destas expressões.

Definição 4.5 - *Soma directa de subespaços*

Seja V um espaço vectorial e S um subconjunto não vazio de V. Um vector x de V diz-se uma combinação linear de elementos de S *se existe um número finito de elementos $y_1, \ldots, y_n$ de S e escalares $a_1, \ldots, a_n$ de F tais que $x = a_1 y_1 + \cdots + a_n y_n$. Neste caso é costume dizer-se que x é uma* combinação linear *de $y_1, \ldots, y_n$.*

Notar que em qualquer espaço vectorial V, $0x = 0$ para cada $x \in V$. O vector zero é portanto combinação linear de qualquer subconjunto não vazio de V.

Na sequência deste curso encontraremos muitas situações onde é necessário determinar se um vector pode ou não ser expresso como combinação linear de outros vectores e, no caso afirmativo, como. Este problema reduz-se ao de resolver um sistema de equações lineares. O exemplo seguinte ilustra este problema.

Exemplo 4.16 Determinemos se o vector (2, 6, 8) pode ser expresso como combinação de

$$y_1 = (1, 2, 1), y_2 = (-2, -4, -2), y_3 = (0, 2, 3), y_4 = (2, 0, -3) \text{ e } y_5 = (-3, 8, 16).$$

Temos portanto de determinar se existem escalares a_1, a_2, a_3, a_4, a_5 tais que

$$\begin{aligned}
(2, 6, 8) &= a_1 y_1 + a_2 y_2 + a_3 y_3 + a_4 y_4 + a_5 y_5 \\
&= a_1(1, 2, 1) + a_2(-2, -4, -2) + a_3(0, 2, 3) + a_4(2, 0, -3) + a_5(-3, 8, 16) \\
&= (a_1 - 2a_2 + 2a_4 - 3a_5, 2a_1 - 4a_2 + 2a_3 + 8a_5, a_1 - 2a_2 + 3a_3 - 3a_4 + 16a_5)
\end{aligned}$$

Portanto $(2, 6, 8)$ pode ser expresso como combinação linear de y_1, y_2, y_3, y_4, y_5 se e só se existir um tuplo de 5 escalares $(a_1, a_2, a_3, a_4, a_5)$ que satisfaça o sistema de equações lineares

$$\begin{cases} a_1 - 2a_2 + 2a_4 - 3a_5 = 2 \\ 2a_1 - 4a_2 + 2a_3 + 8a_5 = 6 \\ a_1 - 2a_2 + 3a_3 - 3a_4 + 16a_5 = 8 \end{cases}$$

obtido igualando coordenadas correspondentes na equação anterior. O sistema pode ser resolvido por eliminação gaussiana, do que resulta

$$\begin{cases} a_1 = 2a_2 - a_5 - 4 \\ a_3 = -3a_5 + 7 \\ a_4 = 2a_5 + 3. \end{cases}$$

Para qualquer escolha dos escalares a_2 e a_5, um vector da forma

$$(a_1, a_2, a_3, a_4, a_5) = (2a_2 - a_5 - 4, a_2, -3a_5 + 7, 2a_5 + 3, a_5)$$

é solução do sistema. Em particular, o vector (-4, 0, 7, 3, 0) obtido fazendo $a_2 = 0$ e $a_5 = 0$ é uma solução. Temos portanto que

$$(2, 6, 8) = -4y_1 + 0y_2 + 7y_3 + 3y_4 + 0y_5,$$

e concluímos que $(2, 6, 8)$ é uma combinação linear de y_1, y_2, y_3, y_4, y_5.$\diamondsuit$

O conjunto de combinações lineares de elementos de um subconjunto não vazio de um espaço vectorial é mais um exemplo de subespaço.

Teorema 4.3 *Sendo S um subconjunto não vazio de um espaço vectorial V, o conjunto W de todas as combinações lineares de elementos de S é um subespaço de V. W é ainda o mais pequeno subespaço de V que contém S, isto é, W é um subconjunto de qualquer subespaço de V que contenha S.*

Definição 4.6 - *Span(S)*

O subespaço W descrito no Teorema 4.3 é designado subespaço gerado pelos elementos de S *e denotado* span(S). *Por conveniência define-se span($\emptyset$) = $\{0\}$.*

Definição 4.7 - *Conjunto gerador*

Um subconjunto S de um espaço vectorial V gera V *se span(S)=V. Diz-se também, nesta situação, que os elementos de S geram V.*

Exemplo 4.17 Os vectores (1,1,0), (1,0,1) e (0,1,1) geram R^3 porque um elemento arbitrário (a_1, a_2, a_3) de R^3 é uma combinação linear dos 3 vectores dados. Os escalares r, s e t para os quais

$$r(1,1,0) + s(1,0,1) + t(0,1,1) = (a_1, a_2, a_3)$$

são

$$r = \frac{1}{2}(a_1 + a_2 - a_3),\ s = \frac{1}{2}(a_1 - a_2 + a_3)\ e\ t = \frac{1}{2}(-a_1 + a_2 + a_3).\ \diamondsuit$$

Exemplo 4.18 Os polinómios $x^2 + 3x - 2$, $2x^2 + 5x - 3$ e $-x^2 - 4x + 4$ geram $P_2(R)$ pois cada um deles pertence a $P_2(R)$ e cada polinómio $ax^2 + bx + c$ de $P_2(R)$ é combinação linear de estes 3. Mais concretamente

$$ax^2 + bx + c = (-8a + 5b + 3c)(x^2 + 3x - 2) +$$

$$(4a - 2b - c)(2x^2 + 5x - 3) + (-a + b + c)(-x^2 - 4x + 4).\ \diamondsuit$$

Exemplo 4.19 As matrizes

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ e } \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

geram $M_{2\times2}(R)$ uma vez que um elemento arbitrário de $M_{2\times2}(R)$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

pode ser expresso como uma combinação linear das 4 matrizes dadas da seguinte forma:

$$\begin{aligned}
& (\frac{1}{3}a_{11} + \frac{1}{3}a_{12} + \frac{1}{3}a_{21} - \frac{2}{3}a_{22}) \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \\
+ \quad & (\frac{1}{3}a_{11} + \frac{1}{3}a_{12} - \frac{2}{3}a_{21} + \frac{1}{3}a_{22}) \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\
+ \quad & (\frac{1}{3}a_{11} - \frac{2}{3}a_{12} + \frac{1}{3}a_{21} + \frac{1}{3}a_{22}) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\
+ \quad & (-\frac{2}{3}a_{11} + \frac{1}{3}a_{12} + \frac{1}{3}a_{21} + \frac{1}{3}a_{22}) \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \\
= \quad & \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}
\end{aligned}$$

$\diamondsuit$

4.5 Dependência e Independência Linear

No início da secção anterior viu-se que a equação de um plano passando por 3 pontos não colineares no espaço, um dos quais é a origem, é da forma $x = t_1u + t_2v$, em que $u, v \in R^3$ e t_1 e t_2 são escalares. Portanto, um vector x de R^3 é uma combinação linear de $u, v \in R^3$ se e só se x se situa no plano definido por u e v. Em R^3 o espaço gerado por 2 vectores não paralelos tem assim uma interpretação geométrica simples. O mesmo se pode dizer do espaço gerado por um só vector não zero de R^3.

Na equação $x = t_1u + t_2v$, x depende de u e v na medida em que x é combinação linear de u e v. Um conjunto no qual pelo menos um dos vectores é combinação linear dos outros é dito linearmente dependente.

Definição 4.8 - *Subconjunto linearemente dependente*

Um subconjunto S de um espaço vectorial V diz-se linearmente dependente *se existe um número finito de vectores distintos $x_1, x_2, \ldots, x_n$ de S e escalares $a_1, a_2, \ldots, a_n$, não todos zero, tais que*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0.$$

Neste caso diz-se também que os elementos de S são linearmente dependentes.

Vê-se facilmente que em qualquer espaço vectorial um subconjunto S que contenha o vector zero é forçosamente linearmente dependente. Uma vez que $0 = 1 \,.\, 0$, o vector zero é combinação linear de elementos de S na qual algum coeficiente é não nulo.

Exemplo 4.20 Em R^4 o conjunto

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

é linearmente dependente porque

$$4(1, 3, -4, 2) - 3(2, 2, -4, 0) + 2(1, -3, 2, -4) + 0(-1, 0, 1, 0) = (0, 0, 0, 0). \ \diamondsuit$$

Definição 4.9 - *Subconjunto linearemente independente*

Um subconjunto S de um espaço vectorial que não é linearmente dependente diz-se linearmente independente. *Os elementos de um tal conjunto dizem-se também linearmente independentes.*

Os seguintes factos acerca de conjuntos linearmente independentes são verdadeiros em qualquer espaço vectorial.

1. O conjunto vazio é linearmente independente, pois os conjuntos linearmente dependentes têm de ser não vazios.

2. Um conjunto constituído por um só vector não nulo é linearmente independente. Se o conjunto $\{x\}$ for linearmente dependente então $ax = 0$ para algum escalar não zero a. Donde

$$x = a^{-1}(ax) = a^{-1}0 = 0.$$

3. Para quaisquer vectores $x_1, x_2, \ldots, x_n$, temos $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ se $a_1 = 0, a_2 = 0, \ldots, a_n = 0$. Chama-se a isto a representação trivial de 0 como combinação linear de $x_1, x_2, \ldots, x_n$. Um conjunto é linearmente independente se e só se as únicas representações de 0 como combinações lineares dos seus elementos forem as triviais.

Teorema 4.4 *Seja V um espaço vectorial e $S_1 \subseteq S_2 \subseteq V$. Se S_1 é linearmente dependente, então S_2 é também linearmente dependente.*

Corolário 4.4.1 *Seja V um espaço vectorial e $S_1 \subseteq S_2 \subseteq V$. Se S_2 é linearmente independente, então S_1 é linearmente independente.*

4.6 Bases e dimensões

Um subconjunto S de um espaço vectorial V que é linearmente independente e gera V goza de uma propriedade muito útil - cada elemento de V pode ser expresso de uma e uma só maneira como combinação linear de elementos de S. Este resultado transforma os conjuntos geradores linearmente independentes em blocos básicos dos espaços vectoriais.

Definição 4.10 - *Base para um espaço vectorial*

Uma base *β para um espaço vectorial V é um subconjunto linearmente independente de V que gera V. (Sendo β uma base de V, diz-se também que os elementos de β formam uma base de V.)*

Exemplo 4.21 Lembrando que $\text{span}(\emptyset) = \{0\}$ e que $\emptyset$ é linearmente independente, vemos que $\emptyset$ é uma base para o espaço vectorial $\{0\}$. $\diamondsuit$

Exemplo 4.22 Em F^n, seja $e_1 = (1, 0, 0, \ldots, 0, 0)$, $e_2 = (0, 1, 0, \ldots, 0, 0)$, $\ldots$, $e_n = (0, 0, 0, \ldots, 0, 1)$. É fácil ver que o conjunto $\{e_1, e_2, \ldots, e_n\}$ é uma base de F^n e é chamada a *base canónica* de F^n. $\diamondsuit$

Exemplo 4.23 Em $M_{m\times n}(F)$ seja M^{ij} a matriz cuja única entrada não nula é um 1 na linha i e coluna j. Então $\{M^{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ é uma base para $M_{m\times n}(F)$. $\diamondsuit$

Exemplo 4.24 Em $P_n(F)$ o conjunto $\{1, x, x^2, \ldots, x^n\}$ é uma base. É designada a *base canónica* de $P_n(F)$. $\diamondsuit$

Exemplo 4.25 Em $P(F)$ o conjunto $\{1, x, x^2, \ldots\}$ é uma base. $\diamondsuit$

Notar no último exemplo que uma base não precisa de ser finita. Nem todos os espaços vectoriais têm base finita.

O teorema seguinte exprime uma das mais importantes propriedades de uma base.

Teorema 4.5 *Seja V um espaço vectorial e $\beta = \{x_1, \ldots, x_n\}$ um subconjunto de V. Então β é uma base para V se e só se cada vector y de V puder ser expresso de forma única como combinação linear de vectores de β, ou seja, puder ser expresso na forma*

$$y = a_1x_1 + \cdots + a_nx_n$$

para escalares únicos $a_1, \ldots a_n$.

O Teorema 4.5 mostra que cada vector v num espaço vectorial V de base $\beta = \{x_1, \ldots, x_n\}$ pode ser expresso de forma única como

$$v = a_1x_1 + \cdots + a_nx_n$$

para escalares $a_1, \ldots a_n$ apropriados. Portanto, v determina um n-tuplo de escalares único $(a_1, \ldots a_n)$ e, inversamente, cada n-tuplo de escalares determina um único vector v usando as entradas do n-tuplo como coeficientes de uma combinação linear de vectores de β. Isto sugere que V se assemelha ao espaço vectorial F^n, em que n é o número de vectores numa base de V.

Enunciaremos agora alguns resultados relativos a espaços vectoriais e bases respectivas.

- Seja S um subconjunto linearmente independente de um espaço vectorial V, e x um elemento de V que não pertence a S. Então $S \cup \{x\}$ é linearmente dependente se e só se $x \in span(S)$.

- Se um espaço vectorial V é gerado por um conjunto finito S_0, então um subconjunto de S_0 é uma base para V, e portanto V tem uma base finita.

Exemplo 4.26 Pode verificar-se que os vectores $(2,-3,5),(8,-12,20),(1,0,-2)$, $(0,2,-1)$ e $(7,2,0)$ geram R^3. Seleccionaremos uma base para R^3 de entre estes elementos. Para começar seleccionamos qualquer elemento não zero do conjunto gerador, por exemplo $(2,-3,5)$, como um dos elementos da base. Uma vez que $4(2,-3,5)=(8,-12,20)$, o conjunto

$$\{(2,-3,5),(8,-12,20)\}$$

é linearmente dependente. Portanto não incluiremos $(8,-12,20)$ na nossa base. Uma vez que $(1,0,-2)$ não é múltiplo de $(2,-3,5)$, o conjunto

$$\{(2,-3,5),(1,0,-2)\}$$

é linearmente independente. Logo incluiremos $(1,0,-2)$ na base.

Continuando para o próximo elemento do conjunto gerador, incluiremos ou não $(0,2,-1)$ na base conforme o conjunto

$$\{(2,-3,5),(1,0,-2),(0,2,-1)\}$$

seja linearmente independente ou linearmente dependente. É fácil verificar que o conjunto é linearmente independente. Portanto incluiremos $(0,2,-1)$ na base.

O último elemento do conjunto gerador, $(7,2,0)$, é incluído ou não na base conforme o conjunto

$$\{(2,-3,5),(1,0,-2),(0,2,-1),(7,2,0)\}$$

seja linearmente independente ou linearmente dependente. Como

$$2(2,-3,5)+3(1,0,-2)+4(0,2,-1)-(7,2,0)=(0,0,0)$$

excluiremos $(7,2,0)$ da base. O conjunto

$$\{(2,-3,5),(1,0,-2),(0,2,-1)\}$$

constitui assim uma base para R^3. $\diamondsuit$

Os resultados mais importantes deste capítulo são enunciados no teorema seguinte e seus corolários.

Teorema 4.6 (Teorema da substituição) *Seja V um espaço vectorial tendo uma base β com exactamente n elementos. Seja $S = \{y_1, y_2, \ldots y_m\}$ um subconjunto linearmente independente de V contendo exactamente m elementos, com $m \leq n$. Então existe um subconjunto S_1 de β contendo exactamente $n-m$ elementos tal que $S \cup S_1$ gera V.*

Para ilustrar este teorema, notar que $S = \{x^2 - 4, x + 6\}$ é um subconjunto linearmente independente de $P_2(F)$. Como $\beta = \{1, x, x^2\}$ é uma base de $P_2(F)$, tem de haver um subconjunto S_1 de β com $3 - 2 = 1$ elementos tal que $S \cup S_1$ gera $P_2(F)$. Neste exemplo qualquer subconjunto de β tendo 1 elemento serve para S_1. Vemos assim que o conjunto S_1 no Teorema 4.6 não tem de ser único.

Corolário 4.6.1 *Seja V um espaço vectorial tendo uma base β com exactamente n elementos. Então qualquer subconjunto linearmente independente de V com exactamente n elementos é uma base de V.*

Corolário 4.6.2 *Seja V um espaço vectorial tendo uma base β com exactamente n elementos. Então qualquer subconjunto de V com mais de n elementos é linearmente dependente. Em consequência, qualquer subconjunto linearmente independente de V tem no máximo n elementos.*

Exemplo 4.27 Seja $S = \{x^2+7, 8x^2-2x, 4x-3, 7x+2\}$. Embora se possa provar directamente que S é um subconjunto linearmente dependente de $P_2(F)$, esta conclusão é imediata pelo Corolário 4.6.2 porque $\{1, x, x^2\}$ é uma base para $P_2(F)$ tendo menos elementos que S. $\diamondsuit$

Corolário 4.6.3 *Seja V um espaço vectorial tendo uma base β com exactamente n elementos. Então qualquer base de V tem exactamente n elementos.*

As definições seguintes dizem respeito ao número de elementos na base de um espaço.

Definição 4.11 - *Dimensão de um espaço vectorial*

Um espaço vectorial V diz-se de dimensão finita *se tem uma base com um número finito de elementos. O número único de elementos em cada base para V chama-se dimensão de V e denota-se* dim(V). *Se um espaço vectorial não é de dimensão finita diz-se de* dimensão infinita.

Os exemplos seguintes focam a dimensão de espaços referidos anteriormente.

Exemplo 4.28 O espaço vectorial $\{0\}$ tem dimensão 0. $\diamondsuit$

Exemplo 4.29 O espaço vectorial F^n tem dimensão n. $\diamondsuit$

Exemplo 4.30 O espaço vectorial $M_{m\times n}(F)$ tem dimensão mn. $\diamondsuit$

Exemplo 4.31 O espaço vectorial $P_n(F)$ tem dimensão $n+1$. $\diamondsuit$

Exemplo 4.32 O espaço vectorial $P(F)$ é de dimensão infinita. $\diamondsuit$

Nos exemplos seguintes faz-se notar que a dimensão de um espaço vectorial depende do seu corpo de escalares.

Exemplo 4.33 Sobre o corpo dos números complexos, o espaço vectorial dos números complexos tem dimensão 1. (Uma base é $\{1\}$.) $\diamondsuit$

Exemplo 4.34 Sobre o corpo dos números reais, o espaço vectorial dos números complexos tem dimensão 2. (Uma base é $\{1, j\}$.) $\diamondsuit$

Corolário 4.6.4 *Sejam V um espaço vectorial de dimensão n e S um subconjunto de V que gera V e tem no máximo n elementos. Então S é uma base para V e portanto tem exactamente n elementos.*

Corolário 4.6.5 *Seja β uma base para um espaço vectorial de dimensão finita V e S umsubconjunto linearmente independente de V. Existe um subconjunto S_1 de β tal que $S \cup S_1$ é uma base para V. Portanto, cada subconjunto linearmente independente de V pode ser estendido para uma base de V.*

O Teorema 4.6 e os seus corolários condensam o essencial das relações entre subconjuntos linearmente independentes, bases e dimensões. Vamos resumir os resultados principais.

Uma base para um espaço vectorial V é um subconjunto linearmente independente que gera V. Se V tem uma base finita, então qualquer base para V tem o mesmo número de vectores. Este número é chamado a **dimensão de** V, e V é dito de dimensão finita.

Portanto, se a dimensão de V é n, toda a base para V tem exactamente n vectores. Mais ainda, cada subconjunto linearmente independente de V não tem mais de n vectores e pode ser estendido para uma base de V incluindo vectores escolhidos apropriadamente. Cada conjunto gerador de V tem pelo menos n vectores e pode ser reduzido a uma base de V excluindo vectores escolhidos convenientemente. O diagrama na Figura 4.5 mostra estas relações.

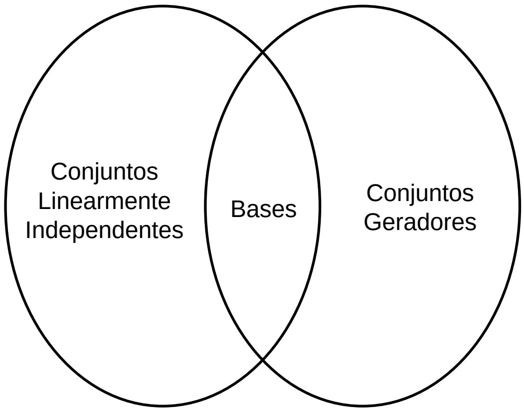

Figura 4.5: Relação entre bases, conjuntos linearmente independentes e conjuntos geradores.

A dimensão dos Subespaços

Vamos agora relacionar a dimensão de um subespaço com a dimensão do espaço vectorial que o contém.

Teorema 4.7 *Seja W um subespaço de um espaço vectorial de dimensão finita V. Então W tem dimensão finita e dim$(W) \leq$ dim(V). Para além disto, se dim(W) =dim(V) então $W = V$.*

Exemplo 4.35 Seja

$$\{(a_1, a_2, a_3, a_4, a_5) \in F^5 : a_1 + a_3 + a_5 = 0, a_2 = a_4\}.$$

Pode ver-se facilmente que W é um subespaço de F^5 com base

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

Portanto $\dim(W) = 3$. $\diamondsuit$

Exemplo 4.36 O conjunto das matrizes diagonais $n \times n$ é um subespaço W de $M_{n\times n}(F)$. Uma base para W é

$$\{M^{11}, M^{22}, \ldots, M^{nn}\},$$

em que M^{ij} é a matriz onde a única entrada não nula é um 1 na linha i coluna j. Então $\dim(W) = n$. $\diamondsuit$

Exemplo 4.37 Vimos antes que o conjunto das matrizes simétricas $n \times n$ é um subespaço W de $M_{n\times n}(F)$. Uma base para W é

$$\{A^{ij} : 1 \leq i \leq j \leq n\},$$

em que A^{ij} é a matriz $n \times n$ tendo um 1 na linha i coluna j, um 1 na coluna i linha j, e 0 em todos os restantes elementos. Vemos que

$$dim(W) = n + (n-1) + \cdots + 1 = \frac{1}{2}n(n-1). \diamondsuit$$

Corolário 4.7.1 *Sendo W um subespaço de um espaço vectorial V de dimensão finita, então W tem uma base finita, e qualquer base de W é um subconjunto de uma base de V.*

Exemplo 4.38 O conjunto dos polinómios da forma

$$a_{18}x^{18} + a_{16}x^{16} + \cdots + a_2x^2 + a_0,$$

onde $a_{18}, a_{16}, \ldots, a_2, a_0 \in F$, é um subespaço W de$P_{18}(F)$. Uma base para W é

$$\{1, x^2, \ldots, x^{16}, x^{18}\},$$

que é um subconjunto da base canónica para $P_{18}(F)$ $\diamondsuit$

Podemos usar o Teorema 4.7 para determinar os subespaços de R^2 e R^3. Como R^2 tem dimensão 2, os subespaços de R^2 só podem ter dimensão 0, 1 ou 2. Os únicos subespaços de dimensão 0 e 2 são $\{0\}$e R^2, respectivamente. Qualquer subespaço de R^2 de dimensão 1 é constituído por todos os múltiplos escalares de algum vector não nulo de R^2.

Identificando um ponto de R^2 da forma habitual com um ponto do plano Euclideano, é possível descrever geometricamente os subespaços de R^2: um subespaço de R^2 de dimensão 0 é a origem do plano Euclideano; um subespaço de R^2 de dimensão 1 identifica-se com uma recta passando pela origem; um subespaço de R^2 de dimensão 2 é todo o plano Euclideano.

De forma semelhante, os subespaços de R^3 terão de ter dimensões 0, 1, 2 ou 3. Interpretando geometricamente, vemos que um subespaço de dimensão 0 será a origem do espaço Euclideano tridimensional, um subespaço de dimensão 1 uma recta passando pela origem, um subespaço de dimensão 2 um plano passando pela origem e um subespaço de dimensão 3 o próprio espaço tridimensional.

Capítulo 5

Transformações lineares

Usando a teoria dos espaços vectoriais, vamos considerar funções definidas sobre estes espaços, e em particular as que preservam a estrutura de espaço vectorial. As funções nestas condições são designadas **transformações lineares** e são comuns em áreas como a análise matemática e a geometria.

Neste capítulo vamos definir e dar exemplos de transformações lineares e estudar a relação entre transformações lineares e matrizes. No que se segue assume-se que os espaços vectoriais são definidos sobre um corpo F.

5.1 Transformações lineares

Definição 5.1 - *Transformação linear*

Sejam V e W espaços vectoriais. Uma função $T : V \to W$ é chamada transformação linear de V em W *se para todo o $x, y \in V$ e $c \in F$ se verificam as seguintes condições:*

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

(b) $T(cx) = cT(x)$.

Para uma transformação $T : V \to W$ verificam-se os seguintes factos:

1. Se T é linear então $T(0) = 0$.
2. T é linear se e só se $T(ax + y) = aT(x) + T(y)$ para todo o $x, y \in V$ e $a \in F$.
3. T é linear se e só se para $x_1, \ldots x_n \in V$ e $a_1, \ldots a_n \in F$ se verifica

$$T(\sum_{i=1}^{n} a_i x_i) = \sum_{i=1}^{n} a_i T(x_i).$$

A propriedade (2) acima é habitualmente usada para demonstrar que uma dada transformação é linear.

Exemplo 5.1 Seja $V = P_n(R)$ e $W = P_{n-1}(R)$. Vamos definir $T : V \to W$ como $T(f) = f'$, onde f' denota a derivada de f. Para mostrar que T é linear, sejam g e h vectores de $P_n(R)$ e $a \in R$. Teremos

$$T(ag + h) = (ag + h)' = ag' + h' = aT(g) + T(h).$$

Portanto, pela propriedade enunciada acima, T é linear. $\diamondsuit$

Exemplo 5.2 Seja $V = C(R)$ o espaço vectorial das funções reais de variável real contínuas e $a, b \in R, a < b$. Defina-se $T : V \to R$ como

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

para todo o $f \in V$. Então T é uma transformação linear pelas propriedades elementares do integral. $\diamondsuit$

Dois exemplos muito importantes de transformações lineares de que faremos uso frequente são a **transformação identidade** e a **transformação zero**.

Sobre um espaço vectorial V define-se a *transformação identidade* $I_V : V \to V$ como

$$I_V(x) = x \text{ para todo o } x \in V.$$

Para espaços vectoriais V e W define-se a *transformação zero* $T_0 : V \to W$ como

$$T_0(x) = 0 \text{ para todo o } x \in V.$$

É fácil verificar que estas transformações são lineares. I_V será abreviada para I quando for claro qual o espaço V em causa.

Exemplo 5.3 Seja $T : R^2 \to R^2$

$$T(a_1, a_2) = (2a_1 + a_2, a_1).$$

Para mostrar que T é linear tomamos $c \in F$ e $x, y \in R^2$ em que $x = (b_1, b_2)$ e $y = (d_1, d_2)$. Então

$$cx + y = (cb_1 + d_1, cb_2 + d_2),$$

e teremos

$$T(cx + y) = (2(cb_1 + d_1) + cb_2 + d_2, cb_1 + d_1).$$

Por outro lado

$$\begin{aligned} cT(x) + T(y) &= c(2b_1 + b_2, b_1) + (2d_1 + d_2, d_1) \\ &= (2cb_1 + cb_2 + 2d_1 + d_2, cb_1 + d_1) \\ &= (2(cb_1 + d_1) + cb_2 + d_2, cb_1 + d_1). \end{aligned}$$

Portanto, T é linear. $\diamondsuit$

Exemplo 5.4 Seja $T : M_{m\times n}(F) \to M_{n\times m}(F)$

$$T(A) = A^t$$

em que A^t designa a matriz transposta de A. Então T é linear pelas propriedades da transposição de matrizes. $\diamondsuit$

A álgebra linear tem larga aplicação na geometria. Isto deve-se ao facto de as transformações geométricas mais importantes serem transformações lineares. Vamos considerar nos exemplos seguintes as transformações correspondentes à rotação, reflexão e projecção. A prova da linearidade destas transformações é deixada como exercício.

Exemplo 5.5 Para $0 \leq \theta < 2\pi$ define-se $T_\theta : R^2 \to R^2$ como

$$T_\theta(a_1, a_2) = (a_1 cos\theta - a_2 sen\theta, a_1 sen\theta + a_2 cos\theta).$$

T_θ é designada *rotação de* θ (ver a figura 5.1). $\diamondsuit$

Exemplo 5.6 Seja $T : R^2 \to R^2$ definida como $T(a_1, a_2) = (a_1, -a_2)$. T é designada *reflexão em torno do eixo dos xx* (ver a figura 5.1). $\diamondsuit$

Exemplo 5.7 Seja $T : R^2 \to R^2$ definida como $T(a_1, a_2) = (a_1, 0)$. T é designada *projecção sobre o eixo dos xx* (ver a Figura 5.1). $\diamondsuit$

Vamos considerar dois conjuntos muito importantes associados a cada transformação linear: o "núcleo" e a "imagem". A determinação destes conjuntos permite analisar as propriedades de uma transformação linear.

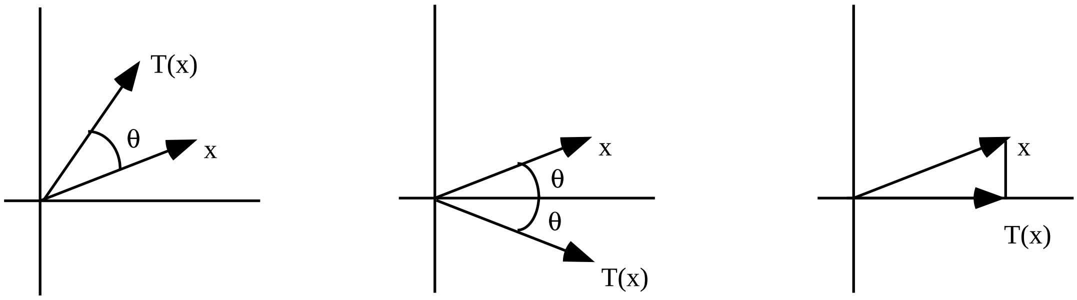

Figura 5.1: Transformações geométricas.

Núcleo e imagem

Sejam V e W espaços vectoriais e $T : V \to W$ uma transformação linear.

Definição 5.2 - *Núcleo de uma transformação linear*
Define-se núcleo *(ou* espaço nulo*) de T, $N(T)$, como o conjunto de todos os vectores $x \in V$ tais que $T(x) = 0$; ou seja,*

$$N(T) = \{x \in V : T(x) = 0\}.$$

Definição 5.3 - *Imagem de uma transformação linear*
Define-se imagem *(ou* contradomínio*) de T, $R(T)$, como o subconjunto de W que contém todas as imagens, por T, dos elementos de V; ou seja,*

$$R(T) = \{T(x) : x \in V\}.$$

Exemplo 5.8 Sejam V e W espaços vectoriais e $I : V \to V$ e $T_0 : V \to W$ as transformações identidade e zero, respectivamente, tal como definidas atrás. Então $N(I) = \{0\}$, $R(I) = V$, $N(T_0) = V$ e $R(T_0) = \{0\}$. $\diamondsuit$

Exemplo 5.9 Seja $T : R^3 \to R^2$

$$T(a_1, a_2, a_3) = (a_1 - a_2, 2a_3).$$

Pode verificar-se que $N(T) = \{(a, a, 0) : a \in R\}$ e $R(T) = R^2$. $\diamondsuit$

Nos exemplos anteriores pode observar-se que o núcleo e imagem de cada uma das transformações são subespaços. O próximo resultado mostra que é assim em geral.

Teorema 5.1 *Sejam V e W espaços vectoriais e $T : V \to W$ uma transformação linear. Então $N(T)$ e $R(T)$ são subespaços de V e W, respectivamente.*

O próximo teorema fornece um método para determinar um conjunto gerador para a imagem de uma transformação linear.

Teorema 5.2 *Sejam V e W espaços vectoriais e $T : V \to W$ uma transformação linear. Se V tiver uma base $\beta = \{x_1, \ldots x_n\}$ então*

$$R(T) = span(\{T(x_1), \ldots T(x_n)\}.$$

Exemplo 5.10 Define-se uma transformação linear $T : P_2(R) \to M_{2\times 2}(R)$:

$$T(f) = \begin{bmatrix} f(1) - f(2) & 0 \\ 0 & f(0) \end{bmatrix}$$

Como sabemos que $\beta = \{1, x, x^2\}$ é uma base para $P_2(R)$ teremos

$$\begin{aligned} R(T) &= span(T(\beta)) = span(\{T(1), T(x), T(x^2)\}) \\ &= span\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right) \\ &= span\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right). \end{aligned}$$

Encontrámos assim uma base para $R(T)$, e portanto $dim(R(T)) = 2$. $\diamondsuit$

A dimensão do subespaço $R(T)$ é designada **característica** da transformação linear.

Considerando o efeito de uma transformação linear, vemos intuitivamente que quanto maior for a dimensão do núcleo, tanto menor será a dimensão da imagem. Ou seja, quantos mais vectores forem transformados em 0, menor será o conjunto das imagens. Isto é expresso formalmente no próximo teorema.

Teorema 5.3 (Teorema da Dimensão) *Sejam V e W espaços vectoriais e $T : V \to W$ uma transformação linear. Se V for de dimensão finita então*

$$dim(N(T)) + característica(T) = dim(V).$$

Exemplo 5.11 Aplicando o teorema da dimensão à transformação no exemplo anterior concluimos que $\dim(N(T)) + 2 = 3$, logo $\dim(N(T)) = 1$.

São aplicáveis às transformações lineares as noções de injectividade e sobrejectividade usadas para as funções. Vamos ver que estas características se relacionam com a dimensão do núcleo da transformação.

Teorema 5.4 *Sejam V e W espaços vectoriais e $T : V \to W$ uma transformação linear. Então T é injectiva se e só se $N(T) = \{0\}$.*

Usando o teorema anterior verificamos que a transformação do exemplo anterior não é injectiva.

Exemplo 5.12 Seja $T : P_2(R) \to P_3(R)$ definido como

$$T(f)(x) = 2f'(x) + \int_0^x 3f(t)dt.$$

Então

$$R(T) = span(\{T(1), T(x), T(x^2)\}) = span(\{3x, 2 + \frac{3}{2}x^2, 4x + x^3\}).$$

Logo a característica de T é 3. Como $\dim(P_3(R)) = 4$, T não é sobrejectiva. Pelo teorema da dimensão podemos concluir que $\dim(N(T))$=0. E portanto, pelo Teorema 5.4, T é injectiva. $\diamondsuit$

Vejamos agora o caso em que uma transformação linear é simultaneamente injectiva e sobrejectiva.

Teorema 5.5 *Sejam V e W espaços vectoriais de igual dimensão e $T : V \to W$ uma transformação linear. A transformação T é injectiva se e só se T é sobrejectiva.*

Exemplo 5.13 $T : F^2 \to F^2$ é definida como $T(a_1, a_2) = (a_2 + a_1, a_1)$. Vê-se facilmente que $N(T) = \{0\}$, portanto T é injectiva. Logo, pelo teorema anterior, T é sobrejectiva. $\diamondsuit$

Sendo T uma transformação linear e injectiva, então um subconjunto S é linearmente independente se e só se $T(S)$ é linearmente independente. O exemplo seguinte ilustra este resultado.

Exemplo 5.14 $T : P_2(R) \to R^3$ é definida como $T(a_0 + a_1x + a_2x^2) = (a_0, a_1, a_2)$. É óbvio que T é injectiva. Seja $S = \{2 - x + 3x^2, x + x^2, 1 - 2x^2\}$. Então S é linearmente independente em $P_2(R)$ se e só se

$$T(S) = \{(2, -1, 3), (0, 1, 1), (1, 0, -2)\}$$

for linearmente independente em R^3. $\diamondsuit$

Uma das propriedades mais importantes das transformações lineares é que são completamente determinadas pela sua acção sobre uma base. Este resultado, expresso no próximo teorema e seu corolário, será usado frequentemente.

Teorema 5.6 *Sejam V e W espaços vectoriais e suponhamos que V tem dimensão finita e base $\{x_1, \ldots, x_n\}$. Para quaisquer vectores $y_1, \ldots, y_n$ de W existe exactamente uma transformação linear $T : V \to W$ tal que $T(x_i) = y_i$ para $i = 1, \ldots, n$.*

Corolário 5.6.1 *Sejam V e W espaços vectoriais e suponhamos que V tem uma base finita $\{x_1, \ldots, x_n\}$. Se $U, T : V \to W$ são lineares e $U(x_i) = T(x_i)$ para $i = 1, \ldots, n$, então $U = T$.*

Exemplo 5.15 Seja $T : R^2 \to R^2$ definida como $T(a_1, a_2) = (2a_2 - a_1, 3a_1)$ e $U : R^2 \to R^2$ linear. Sabendo que $U(1, 2) = (3, 3)$ e $U(1, 1) = (1, 3)$, então $U = T$. Isto resulta do corolário anterior e do facto de $\{(1, 2), (1, 1)\}$ ser uma base para R^2. $\diamondsuit$

5.2 Representação Matricial de uma Transformação Linear

O aspecto mais interessante das transformações lineares é admitirem uma representação matricial. Esta qualidade uniformiza o seu estudo e estabelece uma relação um para um entre matrizes e transformações lineares.

A primeira noção necessária para abordar as transformações lineares do ponto de vista matricial é a de *base ordenada.*

Definição 5.4 - *Base ordenada*

Seja V um espaço vectorial de dimensão finita. Uma base ordenada *para V é uma base na qual se impõe uma ordem determinada. Uma base ordenada é assim uma sequência de elementos linearmente independentes de V que geram V.*

Exemplo 5.16 Seja $\beta = \{x_1, x_2, x_3\}$ uma base ordenada de V.

Então $\gamma = \{x_2, x_1, x_3\}$ é também uma base ordenada, mas $\beta \neq \gamma$ dado que as bases são ordenadas. $\diamondsuit$

No espaço vectorial F^n chama-se ao conjunto $\{e_1, \ldots, e_n\}$ a base canónica ordenada de F^n. Do mesmo modo no espaço vectorial $P_n(F)$ chama-se ao conjunto $\{1, x, \ldots, x^n\}$ a base canónica ordenada de $P_n(F)$.

Dispondo do conceito de base ordenada podemos identificar vectores abstractos num espaço vectorial n-dimensional com n-tuplos. Esta identificação é feita usando **vectores de coordenadas**.

Definição 5.5 - *Vector de coordenadas*

Seja $\beta = \{x_1, \ldots, x_n\}$ *uma base ordenada para um espaço vectorial de dimensão finita* V. *Para* $x \in V$ *define-se o* vector das coordenadas de x relativamente a β, $[x]_\beta$, *como*

$$[x]_\beta = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},$$

em que

$$x = \sum_{i=1}^{n} a_i x_i.$$

Notar que $[x_i]_\beta = e_i$ na definição acima. A correspondência $x \to [x]_\beta$ é ela própria uma transformação linear de V em F^n.

Exemplo 5.17 Seja $V = P_2(R)$ e $=\{1, x, x^2\}$ a base canónica ordenada para V. Sendo $f(x) = 4 + 6x - 7x^2$ teremos

$$[f]_\beta = \begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}. \diamondsuit$$

Podemos agora avançar para a definição de representação matricial de uma transformação linear.

Definição 5.6 - *Representação matricial de uma transformação linear*

Sejam V *e* W *espaços vectoriais de dimensão finita com bases ordenadas* $\beta = \{x_1, \ldots, x_n\}$ *e* $\gamma = \{y_1, \ldots, y_m\}$, *respectivamente. Seja* $T : V \to W$ *linear. Então existem escalares únicos* $a_{ij} \in F$ *(*$i = 1, \ldots, m$ *e* $j = 1, \ldots, n$*) tais que*

$$T(x_j) = \sum_{i=1}^{m} a_{ij} y_i \qquad \textit{para} \ \ 1 \leq j \leq n.$$

A matriz $m \times n$ A *definida como* $A_{ij} = a_{ij}$ *é designada a* representação matricial de T nas bases ordenadas β e γ *e escreve-se* $A = [T]_\beta^\gamma$. *Sendo* $V = W$ *e* $\beta = \gamma$ *escreve-se* $A = [T]_\beta$.

Notar que a coluna índice j de A é simplesmente $[T(x_j)]_\gamma$. A obtenção da representação matricial de uma transformação linear é ilustrada nos próximos exemplos.

Exemplo 5.18 Definindo

$$T : P_3(R) \to P_2(R) \qquad T(f) = f'.$$

Sejam β e γ as bases canónicas ordenadas para $P_3(R)$ e $P_2(R)$ respectivamente. Então teremos

$$\begin{array}{rcl} T(1) & = & 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ T(x) & = & 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ T(x^2) & = & 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \\ T(x^3) & = & 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 \end{array}$$

Portanto

$$[T]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}. \ \diamondsuit$$

Exemplo 5.19 Definindo

$$T : R^2 \to R^3 \qquad T(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2).$$

Sejam β e γ as bases canónicas ordenadas para R^2 e R^3 respectivamente. Teremos

$$T(1,0) = (1,0,2) = 1e_1 + 0e_2 + 2e_3$$

e

$$T(0,1) = (3,0,-4) = 3e_1 + 0e_2 - 4e_3.$$

Portanto

$$[T]_\beta^\gamma = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{bmatrix}.$$

Se fizermos $\gamma' = \{e_3, e_2, e_1\}$ então

$$[T]_\beta^{\gamma'} = \begin{bmatrix} 2 & -4 \\ 0 & 0 \\ 1 & 3 \end{bmatrix}. \ \diamondsuit$$

Adição de transformações lineares

Definição 5.7 *Sejam $T, U : V \to W$ funções arbitrárias, sendo V e W espaços vectoriais, e seja $a \in F$. Define-se $T + U : V \to W$ como*

$$(T + U)(x) = T(x) + U(x) \text{ para todo o } x \in V$$

e

$$(aT)(x) = aT(x) \text{ para todo o } x \in V.$$

Esta definição é a usual para a adição e multiplicação escalar de funções. O resultado interessante é que as somas e múltiplos escalares de transformações lineares são ainda lineares.

Teorema 5.7 *Sejam V e W espaços vectoriais e $T, U : V \to W$ lineares. Então para todo o $a \in F$*

(a) $aT + U$ é linear.

(b) Usando as operações de adição e multiplicação escalar tal como definidas acima, a colecção de todas as transformações lineares de V em W é um espaço vectorial sobre F. Este espaço vectorial é denotado $Ł(V, W)$.

Por meio da definição de representação matricial de uma transformação linear, pode fazer-se a correspondência entre $Ł(V, W)$ e o espaço vectorial $M_{m \times n}(F)$, em que n e m são as dimensões de V e W, respectivamente. Esta correspondência é estabelecida no próximo teorema.

Teorema 5.8 *Sejam V e W espaços vectoriais de dimensão finita com bases ordenadas β e γ respectivamente, e $T, U : V \to W$ transformações lineares. Então*

(a) $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$.

(b) $[aT]_\beta^\gamma = a[T]_\beta^\gamma$ para todo o $a \in F$.

Exemplo 5.20 Definindo

$$T : R^2 \to R^3 \qquad T(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2)$$

e

$$U : R^2 \to R^3 \qquad U(a_1, a_2) = (a_1 - a_2, 2a_1, 3a_1 + 2a_2).$$

Sejam β e γ as bases canónicas ordenadas para R^2 e R^3 respectivamente. Então

$$[T]_\beta^\gamma = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{bmatrix}$$

como vimos no exemplo anterior e

$$[U]_\beta^\gamma = \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 2 \end{bmatrix}.$$

Se determinarmos $T+U$ usando a definição das transformações teremos:

$$(T+U)(a_1, a_2) = (2a_1 + 2a_2, 2a_1, 5a_1 - 2a_2).$$

Portanto

$$[T+U]_\beta^\gamma = \begin{bmatrix} 2 & 2 \\ 2 & 0 \\ 5 & -2 \end{bmatrix},$$

o que corresponde exactamente a $[T]_\beta^\gamma + [U]_\beta^\gamma$, ilustrando o teorema anterior.
$\diamondsuit$

Composição de transformações lineares

Definição 5.8 - *Composição de transformações lineares*
Sejam V, W e Z espaços vectoriais e $T : V \to W$ e $U : W \to Z$ transformações lineares. Define-se $UT : V \to Z$ como

$$(UT)(x) = U(T(x)) \text{ para todo o } x \in V.$$

A composição de transformações lineares é também é uma transformação linear.

Teorema 5.9 *Sejam V, W e Z espaços vectoriais e $T : V \to W$ e $U : W \to Z$ lineares. Então $UT : V \to Z$ é linear.*

Listam-se a seguir propriedades da composição de transformações.

Teorema 5.10 *Seja V um espaço vectorial e $T, U_1, U_2 \in Ł(V)$. Então:*

(a) $T(U_1 + U_2) = TU_1 + TU_2$.

(b) $T(U_1U_2) = (TU_1)U_2$.

(c) $TI = IT = T$.

(d) $a(U_1U_2) = aU_1(U_2) = U_1(aU_2)$ *para todo o* $a \in F$.

O próximo resultado relaciona a composição de transformações lineares com o produto das matrizes respectivas.

Teorema 5.11 *Sejam V, W e Z espaços vectoriais de dimensão finita com bases ordenadas α, β e γ respectivamente. Sejam $T : V \to W$ e $U : W \to Z$ transformações lineares. Então*

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta.$$

Usando a representação matricial de uma transformação linear e a multiplicação de matrizes podemos obter o transformado de qualquer vector. Este resultado justifica o interesse da representação matricial de transformações lineares.

Teorema 5.12 *Sejam V e W espaços vectoriais de dimensão finita com bases ordenadas β e γ respectivamente, e $T : V \to W$ linear. Então para cada $x \in V$ temos*

$$[T(x)]_\gamma = [T]_\beta^\gamma \cdot [x]_\beta.$$

Exemplo 5.21 Seja $T : P_3(R) \to P_2(R)$ tal que $T(f) = f'$, e β e γ as bases canónicas ordenadas para $P_3(R)$ e $P_2(R)$ respectivamente. Vimos já que a matriz da transformação nestas bases será

$$A = [T]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Pode ilustrar-se o teorema anterior verificando que $[T(p)]_\gamma = [T]_\beta^\gamma \cdot [p]_\beta$, em que $p \in P_3(R)$ é o polinómio $p(x) = 2 - 4x + x^2 + 3x^3$. Sendo $q = T(p)$ teremos $q(x) = p'(x) = -4 + 2x + 9x^2$. Então

$$[T(p)]_\gamma = [q]_\gamma = \begin{bmatrix} -4 \\ 2 \\ 9 \end{bmatrix}.$$

Por outro lado

$$[T]_\beta^\gamma \cdot [p]_\beta = A \cdot [p]_\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ 9 \end{bmatrix}. \diamondsuit$$

Vamos agora definir a transformação "multiplicação à esquerda", designada L_A onde A é uma matriz $m \times n$. Esta transformação permite estabelecer uma correspondência directa entre transformações lineares e matrizes, e transferir propriedades acerca de matrizes para transformações e vice-versa.

Definição 5.9 - *Transformação de "multiplicação à esquerda"*
Seja A uma matriz $m \times n$ com elementos de um corpo F. Denota-se L_A a transformação $L_A : F^n \to F^m$ definida por $L_A(x) = Ax$ (o produto matricial de A por x) para cada vector coluna $x \in F^n$. L_A designa-se transformação de multiplicação à esquerda*.*

Exemplo 5.22 Seja

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

Sendo $x = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ teremos

$$L_A(x) = Ax = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}. \diamondsuit$$

O operador L_A é obviamente linear. Enumeram-se a seguir algumas das suas propriedades, que podem ser utilizadas para derivar muitas das propriedades conhecidas das operações com matrizes.

1. $[L_A]_\beta^\gamma = A$.
2. $L_A = L_B$ se e só se $A = B$.
3. $L_{A+B} = L_A + L_B$ e $L_{aA} = aL_A$ para todo o $a \in F$.
4. sendo $T : F^n \to F^m$ linear, então existe uma matriz $m \times n$ única C tal que $T = L_C$, sendo $C = [T]_\beta^\gamma$.

5. Sendo E uma matriz $n \times p$, então $L_{AE} = L_A L_E$.

6. Se $m = n$, então $L_{I_n} = I_{F^n}$.

5.3 Inversão de transformações, isomorfismos

Vamos agora definir a transformação inversa de uma transformação linear, verificando que é ainda linear. Em termos de representações matriciais, à transformação inversa vai corresponder a matriz inversa.

Definição 5.10 - *Transformação inversa*
Sejam V e W espaços vectoriais e $T : V \to W$ linear. T tem uma inversa *$U : W \to V$ se $TU = I_W$ e $UT = I_V$. A inversa é única, e escreve-se $U = T^{-1}$. Diz-se que T é* invertível *se T tem inversa.*

Para funções invertíveis T e U verifica-se

- $(TU)^{-1} = U^{-1}T^{-1}$;
- $(T^{-1})^{-1} = T$.

Uma função é invertível se e só se é bijectiva.

Exemplo 5.23 Define-se $T : P_1(R) \to R^2$ como $T(a + bx) = (a, a + b)$. Tem-se então que $T^{-1} : R^2 \to P_1(R)$ pode ser dada por $T^{-1}(c, d) = c + (d - c)x$. Note-se que T^{-1} é ainda linear. $\diamondsuit$

Teorema 5.13 *Sejam V e W espaços vectoriais e $T : V \to W$ linear e invertível. Então $T^{-1} : W \to V$ é linear.*

Deste teorema e do Teorema 5.5 resulta que sendo T uma transformação linear entre espaços de igual dimensão as condições de ser invertível, injectiva e sobrejectiva são todas equivalentes.

Podemos agora relacionar as representações matriciais de transformações inversas.

Teorema 5.14 *Sejam V e W espaços vectoriais de dimensão finita com bases ordenadas β e γ, respectivamente. Seja $T : V \to W$ linear. T é invertível se e só se $[T]_\beta^\gamma$ for invertível. Tem-se ainda que $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.*

Recorrendo à noção de invertibilidade podemos estabelecer formalmente a relação entre espaços vectoriais que são intuitivamente muito semelhantes apesar da forma dos seus elementos. Por exemplo, para os espaços $M_{2\times 2}(F)$ e F^4 vemos que, se associarmos a cada matriz

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

o 4-tuplo (a, b, c, d), as somas e multiplicações escalares se associam de forma semelhante; ou seja, em termos de estrutura de espaço vectorial, estes 2 espaços podem ser considerados **isomorfos**.

Definição 5.11 - *Isomorfismos*

Sejam V e W espaços vectoriais. Diz-se que V é isomorfo *de W se existe uma transformação linear $T : V \to W$ que seja invertível. Uma transformação linear nessas condições diz-se um* isomorfismo *de V em W.*

Exemplo 5.24 Seja $T : F^2 \to P_1(F)$ dada por $T(a_1, a_2) = a_1 + a_2x$. T é obviamente invertível. Logo F^2 é isomorfo de $P_1(F)$. $\diamondsuit$

Da definição de isomorfismo decorrem consequências interessantes para o estudo das transformações lineares. Para espaços V e W sobre o mesmo corpo F:

- V é isomorfo de W se e só se $\dim(V) = \dim(W)$;
- Se V é um espaço vectorial de dimensão n, V é isomorfo de F^n;
- Sendo β e γ bases ordenadas para V e W, a função $\Phi : L(V, W) \to M_{m\times n}(F)$ definida como $\Phi(T) = [T]_\beta^\gamma$ é um isomorfismo;
- Sendo n e m as dimensões de V e W, $L(V, W)$ tem dimensão mn.

5.4 Mudanças de Base

Quando referimos a uma base os elementos de um espaço vectorial, passamos a ter como representação de cada vector as suas coordenadas nessa base. A questão seguinte a levantar é: como se relacionam as coordenadas do mesmo vector em bases diferentes?

Teorema 5.15 *Sejam β e β' duas bases ordenadas para o espaço V, e a matriz $Q = [I_V]_{\beta'}^{\beta}$. Então:*

(a) Q é invertível.

(b) Para todo o $v \in V$, $[v]_\beta = Q[v]_{\beta'}$.

Prova: (a) Sendo I_V invertível, Q é invertível pelo Teorema 5.14.

(b) Para qualquer $v \in V$

$$[v]_\beta = [I_V(v)]_\beta = [I_V]_{\beta'}^{\beta}[v]_{\beta'} = Q[v]_{\beta'}.$$

♣

A matriz Q definida atrás é designada *matriz de mudança de coordenadas*. Atendendo à parte (b) do teorema acima, diremos que Q *muda coordenadas* β' *para coordenadas* β. Observe-se que sendo

$$\beta = \{x_1, x_2, \ldots, x_n\}$$

e

$$\beta' = \{x'_1, x'_2, \ldots, x'_n\},$$

então, pela definição de representação matricial de uma transformação,

$$I_V(x'_j) = \sum_{i=1}^{n} Q_{ij}x_i,$$

para $j = 1, 2, \ldots, n$; logo

$$x'_j = \sum_{i=1}^{n} Q_{ij}x_i,$$

ou seja a coluna de índice j em Q é $[x'_j]_\beta$.

Note-se que sendo Q a matriz de mudança de coordenadas β' para coordenadas β, então Q^{-1} é a matriz de mudança de coordenadas β para coordenadas β'.

Exemplo 5.25 Seja $V = R^2$, $=\{(1,1),(1,-1)\}$ e $\beta' = \{(2,4),(3,1)\}$. Sendo

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

a matriz que muda coordenadas β' para coordenadas β é

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

Teremos então, por exemplo,

$$[(2,4)]_\beta = Q[(2,4)]_{\beta'} = Q\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}. \diamondsuit$$

Suponhamos agora que $T : V \rightarrow V$ é uma transformação linear no espaço V e que β e β' são bases ordenadas para V. Então T pode ser representada pelas matrizes $[T]_\beta$ e $[T]_{\beta'}$. Põe-se a questão de relacionar estas duas matrizes.

Teorema 5.16 *Sejam $T : V \rightarrow V$ uma transformação linear no espaço V de dimensão finita, e β e β' bases ordenadas para V. Seja Q a matriz de mudança de coordenadas β' para coordenadas β. Então*

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q.$$

Prova: Seja I a transformação identidade em V. Então $T = IT = TI$ e teremos

$$\begin{aligned} Q[T]_{\beta'} &= [I]_{\beta'}^{\beta}[T]_{\beta'}^{\beta'} \\ &= [IT]_{\beta'}^{\beta} \\ &= [TI]_{\beta'}^{\beta} \\ &= [T]_{\beta}^{\beta}[I]_{\beta'}^{\beta} \\ &= [T]_\beta Q. \end{aligned}$$

Portanto, $[T]_{\beta'} = Q^{-1}[T]_\beta Q$. ♣

Exemplo 5.26 Seja $V = R^3$ e $T : V \rightarrow V$ definida como

$$T\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right) = \begin{bmatrix} 2a_1 + a_2 \\ a_1 + a_2 + 3a_3 \\ -a_2 \end{bmatrix}.$$

Seja β a base canónica ordenada para R^3 e

$$\beta' = \left\{ \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\},$$

uma nova base ordenada para R^3. Para ilustrar o Teorema 5.16, vejamos que

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

Seja Q a matriz que muda coordenadas β' para coordenadas β. Como β é a base canónica para R^3, as colunas de Q são simplesmente os elementos de β' escritos na mesma ordem. Então

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

Calculando Q^{-1} obtemos

$$Q^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Pelo Teorema 5.16 temos que $[T]_{\beta'} = Q^{-1}[T]_\beta Q$. Multiplicando obtemos

$$[T]_{\beta'} = \begin{bmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{bmatrix}.$$

Para mostrar que esta é a matriz correcta, podemos verificar que a imagem por T do elemento de índice j de β' é a combinação dos elementos de β' com os elementos da coluna de índice j como coeficientes. Por exemplo, para $j = 2$ teremos

$$T\left(\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right) = 2\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -1 \end{bmatrix}.$$

Notar que os coeficientes são os elementos da segunda coluna de $[T]_{\beta'}$. $\diamondsuit$

É muitas vezes conveniente aplicar o Teorema 5.16 ao contrário, como se mostra no exemplo seguinte.

Exemplo 5.27 Usamos já como exemplo a transformação linear correspondente à reflexão em torno do eixo dos xx. A regra $(x, y) \to (x, -y)$ para este caso é simples de obter. Vejamos agora como definir a regra menos óbvia para uma reflexão em torno da recta $y = 2x$ (ver a Figura 5.2).

Pretendemos uma expressão para $T(a, b)$ para todo o (a, b) em R^2. Sendo T uma transformação linear, é completamente determinada pelos seus valores numa base para R^2.

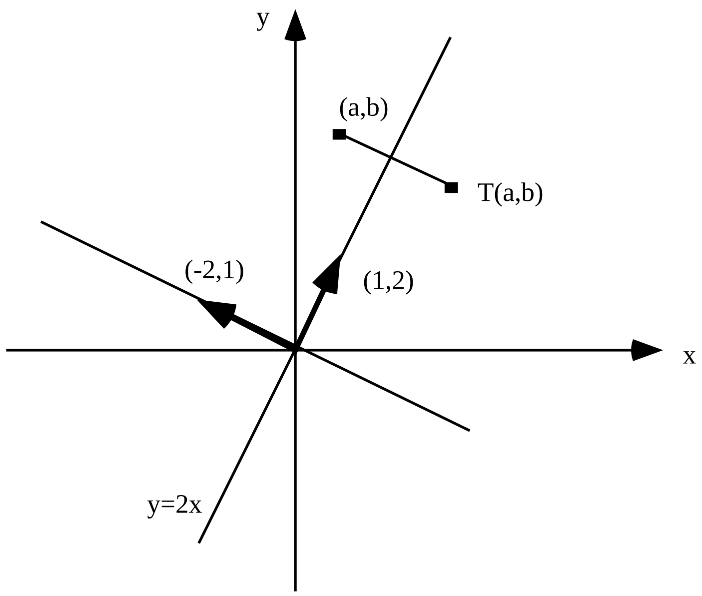

Figura 5.2: Reflexão em torno de $y = 2x$

É óbvio que $T(1,2) = (1,2)$ e que $T(-2,1) = -(-2,1) = (2,-1)$. Então, se fizermos

$$\beta' = \{(1,2), (-2,1)\},$$

β' é uma base ordenada para R^2 e

$$[T]_{\beta'} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Sendo β a base canónica para R^2, a matriz de mudança de coordenadas β' para β será

$$Q = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \quad \text{e} \quad Q^{-1}[T]_\beta Q = [T]_{\beta'}.$$

Podemos resolver esta igualdade para $[T]_\beta$ obtendo $[T]_\beta = Q[T]_{\beta'}Q^{-1}$. Uma vez que

$$Q^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix},$$

teremos

$$[T]_\beta = \frac{1}{5}\begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}.$$

Sendo β a base canónica de R^2, T será o operador de multiplicação à esquerda por $[T]_\beta$. Portanto, para qualquer (a,b) de R^2 teremos

$$T(a,b) = \frac{1}{5}(-3a+4b, 4a+3b).\ \diamondsuit$$

A relação entre as matrizes $[T]_{\beta'}$ e $[T]_\beta$ no Teorema 5.16, pela sua importância no estudo das transformações lineares, tem uma designação especial.

Definição 5.12 - *Matrizes semelhantes*

Sejam A e B elementos de $M_{m\times n}(F)$. Diz-se que B é semelhante *a A se existe uma matriz invertível $Q \in M_{m\times n}(F)$ tal que*

$$B = Q^{-1}AQ.$$

O enunciado do Teorema 5.16 pode ser reformulado recorrendo a esta definição: sendo $T : V \to V$ uma transformação linear num espaço vectorial de dimensão finita V, e β e β' bases de V, então $[T]_{\beta'}$ é semelhante a $[T]_\beta$.

Pode também generalizar-se este resultado para $T : V \to W$ sendo V e W distintos. Neste caso teremos mudança de coordenadas tanto em V como em W.

Teorema 5.17 *Seja $T : V \to W$ uma transformação linear do espaço V de dimensão finita no espaço W de dimensão finita. Sejam β e β' bases ordenadas para V e γ e γ' bases ordenadas para W. Então*

$$[T]_{\beta'}^{\gamma'} = P^{-1}[T]_\beta^\gamma Q$$

onde Q é a matriz de mudança de coordenadas β' para coordenadas β e P a matriz de mudança de coordenadas γ' para coordenadas γ.

Capítulo 6

Diagonalização

Vimos já que uma transformação linear pode ser representada por uma matriz, uma vez escolhidas bases para os espaços vectoriais domínio e imagem da transformação. Ficou também claro que a representação matricial depende da base, e que a mudança de base significa mudança da representação matricial. A questão que abordaremos agora é a de obter a "forma mais simples" para a representação matricial de uma transformação linear. Vendo a matriz de uma transformação como um operador que transforma coordenadas de vectores do domínio em coordenadas das imagens respectivas, podemos intuir que a "forma mais simples" deste operador corresponderá à matriz diagonal. De facto, numa transformação representada por uma matriz diagonal cada coordenada do vector imagem é determinada recorrendo a uma apenas das coordenadas do vector do domínio. Interessa-nos então, para uma transformação $T : V \to V$ procurar resposta para as duas perguntas seguintes:

- Será que existe uma base ordenada β de V tal que $[T]_\beta$ seja uma matriz diagonal?

- Existindo uma tal base, como determiná-la?

Para resolver a questão da diagonalização são necessários os conceitos de valor próprio e vector próprio que veremos de seguida.

6.1 Valores e Vectores Próprios

Iremos concentrar-nos aqui nas transformações lineares de um espaço vectorial nele próprio, que designaremos **operadores lineares**. Para um operador linear T num espaço vectorial V consideraremos as matrizes que o representam nas várias bases ordenadas para V. Em geral omite-se o qualificativo "ordenado" e fala-se simplesmente em "bases".

Vimos já que as matrizes que representam um operador $T : V \to V$ nas bases β e β' para o espaço V se relacionam por

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q$$

onde Q é a matriz de mudança de coordenadas β' para coordenadas β. Definimos ainda que matrizes nestas condições se dizem *semelhantes*. No próximo teorema exploramos um caso particular desta relação.

Teorema 6.1 *Seja $A \in M_{n\times n}(F)$ e $\gamma = \{x_1, x_2, \ldots, x_n\}$ uma base qualquer para F^n. Então $[L_A]_\gamma = Q^{-1}AQ$, em que Q é a matriz $n \times n$ cuja coluna de índice j é x_j para $j = 1, 2, \ldots, n$.*

Exemplo 6.1 Seja

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \in M_{2\times 2}(R) \text{ e } \gamma = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

Sendo

$$Q = \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix},$$

então

$$Q^{-1} = \begin{bmatrix} -2 & 1 \\ 3 & -1 \end{bmatrix}$$

e

$$[L_A]_\gamma = Q^{-1}AQ = \begin{bmatrix} -2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix}. \diamondsuit$$

As matrizes que representam a mesma transformação em bases diferentes são semelhantes. Veremos agora que o inverso também se verifica.

Teorema 6.2 *Seja T um operador linear num espaço vectorial V de dimensão n, e β uma base para V. Sendo B uma matriz $n \times n$ semelhante a $[T]_\beta$, então existe uma base β' para V tal que $B = [T]_{\beta'}$.*

O conceito de semelhança é útil para estudar o problema da diagonalização porque permite reformulá-lo no contexto das matrizes.

Definição 6.1 - *Operador diagonalizável*

Um operador linear T num espaço vectorial V de dimensão finita diz-se diagonalizável *se existe uma base β para V tal que $[T]_\beta$ é uma matriz diagonal.*

Uma matriz quadrada A diz-se diagonalizável *se é semelhante a uma matriz diagonal.*

O teorema seguinte relaciona estes dois conceitos.

Teorema 6.3 *Seja T um operador linear num espaço vectorial V de dimensão finita, e β uma base para V. Então T é diagonalizável se e só se $[T]_\beta$ é uma matriz diagonalizável.*

Corolário 6.3.1 *Uma matriz A é* diagonalizável *se e só se L_A é diagonalizável.*

A questão da diagonalizabilidade pode agora colocar-se em termos de matrizes:

- Será uma dada matriz A diagonalizável?
- Se A é diagonalizável, como determinar a matriz invertível Q tal que $Q^{-1}AQ$ é uma matriz diagonal?

Teorema 6.4 *Um operador linear T num espaço vectorial V de dimensão finita é diagonalizável se e só se existe uma base $\beta = \{x_1, \ldots, x_n\}$ para V e escalares $\lambda_1, \ldots, \lambda_n$ (não necessariamente distintos) tais que $T(x_j) = \lambda_j x_j$, para $1 \leq j \leq n$. Nestas condições tem-se*

$$[T]_\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Este resultado motiva as definicões seguintes.

Definição 6.2 - *Valor próprio e vector próprio*

Seja T um operador linear num espaço vectorial V. Um elemento não zero $x \in V$ diz-se vector próprio *de T se existe um escalar λ tal que $T(x) = \lambda x$. Ao escalar λ chama-se* valor próprio *correspondente ao vector próprio x.*

De forma semelhante, sendo A uma matriz $n \times n$ sobre um corpo F, um elemento não zero $x \in F^n$ diz-se vector próprio *de A se é um vector próprio de L_A; ou seja, $Ax = \lambda x$ para algum escalar λ. Ao escalar λ chama-se* valor próprio *correspondente ao vector próprio x.*

Podemos reformular o teorema anterior de acordo com esta terminologia:

Um operador linear T num espaço V é diagonalizável se e só se existe uma base β para V constituída por vectores próprios de T. Sendo T diagonalizável, $\beta = \{x_1, x_2, \ldots, x_n\}$ uma base de vectores próprios de T e $D = [T]_\beta$, então D é uma matriz diagonal e D_{ii} é o valor próprio correspondente a x_i $(i = 1, 2, \ldots, n)$.

Exemplo 6.2 Seja $C^\infty(R)$ o conjunto de todas as funções $F : R \to R$ tendo derivadas de todas as ordens. Vemos que $C^\infty(R)$ é um subespaço de $F(R, R)$, o espaço de todas as funções reais de variável real definido anteriormente. Define-se $T : C^\infty(R) \to C^\infty(R)$ por $T(y) = y'$, em que y' denota a derivada de y. Vê-se facilmente que T é um operador linear em $C^\infty(R)$. Vamos determinar os seus valores e vectores próprios.

Se λ é valor próprio de T, existe um vector próprio $y \in C^\infty(R)$ tal que $y' = T(y) = \lambda y$. Esta equação diferencial tem soluções da forma $y(t) = ce^{\lambda t}$ para alguma constante c. Portanto, todo o número real λ é um valor próprio de T, sendo os vectores próprios associados da forma $ce^{\lambda t}$ para $c \neq 0$. Sendo $\lambda = 0$ os vectores próprios são as funções constantes não zero. $\diamondsuit$

Exemplo 6.3

$$A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}, \quad x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

$$L_A(x_1) = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2x_1$$

x_1 é vector próprio de L_A, e portanto de A. $\lambda_1 = -2$ é o valor próprio correspondente a x_1.

$$L_A(x_2) = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 20 \end{bmatrix} = 5 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 5x_2$$

x_2 é vector próprio de L_A. $\lambda_2 = 5$ é o valor próprio correspondente a x_2.

$\beta = \{x_1, x_2\}$ é uma base de R^2 e pelo teorema anterior

$$[L_A]_\beta = \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}.$$

Sendo

$$Q = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}$$

teremos

$$Q^{-1}AQ = [L_A]_\beta = \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}. \diamondsuit$$

Este exemplo ilustra uma técnica de diagonalização para uma matriz A de ordem n: sendo β uma base constituída por vectores próprios de A e Q a matriz cujas colunas são os vectores próprios, então $Q^{-1}AQ$ é uma matriz diagonal. Para usar este procedimento é preciso um método para determinar os vectores próprios de uma matriz ou operador. Veremos que estes se determinam facilmente uma vez obtidos os valores próprios. Para este efeito introduz-se a noção de determinante de um operador linear.

Definição 6.3 - *Determinante de um operador linear*

Seja T um operador linear num espaço vectorial de dimensão finita V. Define-se o determinante *de T, denotado* det(T)*, como det($[T]_\beta$), para uma qualquer base β de V.*

Vê-se facilmente que a escolha da base β para $[T]_\beta$ não afecta o determinante. Para os determinantes de operadores lineares provam-se propriedades semelhantes às dos determinantes de matrizes. O resultado importante para o objectivo da diagonalização está expresso no teorema que se segue.

Teorema 6.5 *Seja T um operador linear num espaço vectorial de dimensão finita V sobre um corpo F. Um escalar $\lambda \in F$ é um valor próprio de T se e só se*

$$det(T - \lambda I) = 0.$$

Corolário 6.5.1 *Seja A uma matriz $n \times n$ sobre um corpo F. Um escalar $\lambda \in F$ é um valor próprio de A se e só se $det(A - \lambda I) = 0$.*

Exemplo 6.4 Seja

$$A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \in M_{2\times 2}(R).$$

$$det(A - \lambda I) = det\left(\begin{bmatrix} 1-\lambda & 1 \\ 4 & 1-\lambda \end{bmatrix}\right) = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1).$$

Os valores próprios de A são portanto 3 e -1. $\diamondsuit$

Exemplo 6.5 Seja $T : P_2(R) \to P_2(R)$ o operador linear definido por

$$T(f(x)) = f(x) + xf'(x) + f'(x)$$

e β a base canónica para $P_2(R)$. Então

$$[T]_\beta = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}.$$

$$\begin{aligned} det(T - \lambda I) = det([T]_\beta - \lambda I) &= det\left(\begin{bmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{bmatrix}\right) \\ &= (1-\lambda)(2-\lambda)(3-\lambda) \\ &= -(\lambda-1)(\lambda-2)(\lambda-3). \end{aligned}$$

Os valores próprios de T são portanto 1, 2 e 3. $\diamondsuit$

Nestes exemplos pode observar-se que sendo A uma matriz $n \times n$ então $det(A - \lambda I)$ é um polinómio em λ cujo coeficiente no termo de maior grau é $(-1)^n$. Os valores próprios de A são simplesmente os zeros deste polinómio.

Definição 6.4 - *Polinómio característico*

Sendo $A \in M_{n\times n}(F)$, o polinómio $det(A - tI_n)$ na variável t é designado polinómio característico *de A.*

Seja T um operador linear num espaço vectorial de dimensão finita V com base β. Define-se polinómio característico *de T como o polinómio característico de $A = [T]_\beta$, ou seja*

$$f(t) = det(A - tI_n).$$

Teorema 6.6 *O polinómio característico de $A \in M_{n\times n}(F)$ é um polinómio de grau n com coeficiente $(-1)^n$ no termo de maior grau.*

Corolário 6.6.1 *Seja A uma matriz $n \times n$ e $f(t)$ o seu polinómio característico. Então:*

(a) Um escalar λ é valor próprio de A se e só se λ é zero do polinómio $f(t)$, ou seja se $f(\lambda) = 0$.

(b) A tem no máximo n valores próprios distintos.

Corolário 6.6.2 *Seja T um operador linear num espaço vectorial V de dimensão n e $f(t)$ o seu polinómio característico. Então:*

(a) Um escalar λ é valor próprio de T se e só se λ é zero do polinómio $f(t)$, ou seja se $f(\lambda) = 0$.

(b) T tem no máximo n valores próprios distintos.

Os resultados anteriores fornecem o método para determinar todos os valores próprios de uma matriz ou operador linear. Veremos agora o procedimento para determinar os vectores próprios correspondentes a um determinado valor próprio.

Teorema 6.7 *Seja T um operador linear num espaço vectorial V e λ um valor próprio de T. Um vector $x \in V$ é um vector próprio de T correspondente a λ se e só se $x \neq 0$ e $x \in N(T - \lambda I)$.*

Podemos exemplificar a determinação de vectores próprios para os exemplos anteriores em que já obtivemos os valores próprios.

Exemplo 6.6 A matriz

$$A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$$

tem valores próprios $\lambda_1 = 3$ e $\lambda_2 = -1$. Comecemos pelos vectores para $\lambda_1 = 3$. Façamos

$$B = A - \lambda_1 I = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix}.$$

Então $x \in R^2$ é um vector próprio para o valor próprio λ_1 se e só se $x \neq 0$ e $x \in N(L_B)$, ou seja

$$\begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_1 + x_2 \\ 4x_1 - 2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

O conjunto de soluções deste sistema é

$$\left\{ k \begin{bmatrix} 1 \\ 2 \end{bmatrix} : k \in R \right\}$$

e portanto os vectores próprios correspondentes a $\lambda_1 = 3$ são da forma $k \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ para $k \neq 0$.

Considerando agora os vectores para $\lambda_2 = -1$ teremos

$$B = A - \lambda_2 I = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} - \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}.$$

Então $x \in R^2$ é um vector próprio de A para o valor próprio λ_2 se for solução do sistema

$$\begin{cases} 2x_1 & +x_2 = 0 \\ 4x_1 & +2x_2 = 0 \end{cases}.$$

Portanto

$$N(L_B) = \left\{ k \begin{bmatrix} 1 \\ -2 \end{bmatrix} : k \in R \right\}$$

e os vectores próprios correspondentes a $\lambda_2 = -1$ serão da forma $k \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ para $k \neq 0$. Notar que

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$$

é uma base para R^2 constituída por vectores próprios de A, e portanto A é diagonalizável. De facto, tomando

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

teremos

$$Q^{-1}AQ = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}. \diamondsuit$$

Exemplo 6.7 Consideremos de novo $T : P_2(R) \to P_2(R)$ tal que

$$T(f(x)) = f(x) + xf'(x) + f'(x)$$

para o qual, tomando a base canónica,

$$[T]_\beta = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}.$$

Concluimos já que os valores próprios de T são $\lambda_1 = 1$, $\lambda_2 = 2$ e $\lambda_3 = 3$. Para $\lambda_1 = 1$ teremos

$$B = A - \lambda_1 I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}.$$

É fácil ver que

$$N(L_B) = \left\{ k \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} : k \in R \right\}.$$

Os vectores próprios de T correspondentes a $\lambda_1 = 1$ são da forma $k \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ ou seja, atendendo à base escolhida para $P_2(R)$, são os polinómios constantes diferentes de 0.

Para $\lambda_2 = 2$ teremos

$$B = A - \lambda_2 I = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

de onde

$$N(L_B) = \left\{ k \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} : k \in R \right\}$$

e portanto os vectores próprios de T correspondentes a $\lambda_2 = 2$ têm coordenadas da forma $k \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ e são os polinómios da forma $k + kx$ para $k \neq 0$.

Finalmente para $\lambda_3 = 3$ virá

$$B = A - \lambda_3 I = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

e portanto

$$N(L_B) = \left\{ k \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} : k \in R \right\}.$$

Logo, as coordenadas dos vectores próprios de T correspondentes a $\lambda_3 = 3$ são da forma $k \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ e correspondem aos polinómios da forma $k + 2kx + kx^2$ para $k \neq 0$. $\diamondsuit$

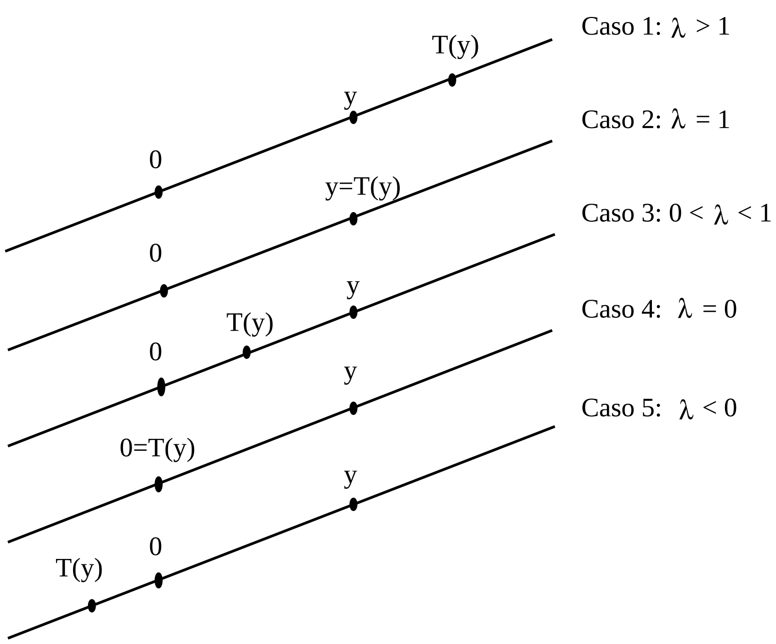

Figura 6.1: Efeito de T sobre o espaço de um seu vector próprio.

Os valores e vectores próprios de matrizes e transformações são susceptíveis de interpretação geométrica. Seja x um vector próprio de um operador T em V, e portanto $T(x) = \lambda x$ para algum escalar λ. Seja $W = span(\{x\})$ o subespaço unidimensional de V gerado por x. Sendo $y \in W$ teremos $y = cx$ para algum escalar c. Então

$$T(y) = T(cx) = cT(x) = c\lambda x = \lambda y \in W.$$

Portanto, T transforma W em si próprio. Sendo V um espaço sobre o corpo dos reais, W pode ser visto como uma recta passando pela origem de V. O operador T multiplica cada elemento de W pelo escalar λ. Dependendo do valor de λ, podemos identificar vários casos (ver Figura 6.1).

Caso 1 Se $\lambda > 1$ então T desloca elementos de W para pontos mais afastados de 0 por um factor de λ.

Caso 2 Se $\lambda = 1$ então T é a transformação identidade em W.

Caso 3 Se $0 < \lambda < 1$ então T desloca elementos de W para pontos mais próximos de 0 por um factor de λ.

Caso 4 Se $\lambda = 0$ então T é a transformação zero em W.

Caso 5 Se $\lambda < 0$ então T inverte a orientação de W, ou seja, desloca pontos de W de um lado de 0 para o outro.

Podemos ilustrar esta visão dos valores e vectores próprios com as transformações geométricas já usadas como exemplo.

Vimos que o operador $T : R^2 \to R^2$ definido por $T(x_1, x_2) = (x_1, -x_2)$ representava uma reflexão em relação ao eixo dos xx. Vemos que esta transformação transforma cada um dos eixos nele próprio. e_1 e e_2 serão vectores próprios da transformação, associados aos valores próprios 1 e -1. T funciona como identidade no eixo dos xx e inverte a orientação do eixo dos yy.

A transformação $U : R^2 \to R^2$ definido por $T(x_1, x_2) = (x_1, 0)$ representa a projecção sobre o eixo dos xx. Vemos geometricamente que U funciona como identidade no eixo dos xx e como transformação zero no eixo dos yy. Vemos que e_1 e e_2 são vectores próprios da transformação, associados aos valores próprios 1 e 0.

Consideremos ainda a transformação rotação de um ângulo θ, definida como $T_\theta : R^2 \to R^2$ tal que $T(x_1, x_2) = (x_1 cos\theta - x_2 sen\theta, x_1 sen\theta + x_2 cos\theta)$. Para $0 < \theta < \pi$ é óbvio que T_θ não transforma nenhum espaço unidimensional nele próprio. Daqui se conclui que T_θ não tem vectores próprios, e portanto também não tem valores próprios. Para confirmar esta conclusão podemos reparar que o polinómio característico de T_θ é

$$det(T_\theta - tI) = det\left(\begin{bmatrix} cos\theta - t & -sen\theta \\ sen\theta & cos\theta - t \end{bmatrix}\right) = t^2 - (2cos\theta)t + 1$$

que não tem zeros em R.

6.2 Diagonalizabilidade

Vimos já que nem todos os operadores lineares são diagonalizáveis. Dos resultados anteriores concluímos que é condição necessária e suficiente de diagonalizabilidade a existência de uma base de vectores próprios. O que nos falta ainda é uma forma simples de determinar se um operador linear é diagonalizável e, no caso de ser, um procedimento para determinar a base de vectores próprios.

Em exemplos anteriores construímos bases de vectores próprios escolhendo um vector próprio associado a cada valor próprio. Em geral este procedimento não produz uma base, mas qualquer conjunto obtido por este processo é linearmente independente.

Teorema 6.8 *Seja T um operador linear num espaço vectorial V e $\lambda_1, \lambda_2, \ldots \lambda_k$ valores próprios distintos de T. Sendo $x_1, x_2, \ldots, x_k$ vectores próprios distintos de T tais que λ_j corresponde a x_j ($1 \leq j \leq k$), então $\{x_1, x_2, \ldots, x_k\}$ é linearmente independente.*

Corolário 6.8.1 *Seja T um operador linear num espaço vectorial V de dimensão finita n. Se T tem n valores próprios distintos, então T é diagonalizável.*

Exemplo 6.8 Seja

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \in M_{2\times 2}(R).$$

O polinómio característico de A (e portanto de L_A) é

$$det(A - tI) = det\left(\begin{bmatrix} 1-t & 1 \\ 1 & 1-t \end{bmatrix}\right) = t(t-2).$$

Os valores próprios de A são portanto 0 e 2. Como L_A é um operador linear no espaço de dimensão 2, R^2, conclui-se que L_A (e portanto A) é diagonalizável. $\diamondsuit$

Definição 6.5 - *Polinómio factorizável*

Um polinómio $f(x)$ de $P(F)$ é factorizável em F se existem escalares a_0,a_1, ..., a_n (não necessariamente distintos) em F tais que

$$f(x) = a_0(x - a_1)(x - a_2)\cdots(x - a_n).$$

O polinómio $x^2 - 1 = (x+1)(x-1)$ é factorizável em R, mas $(x^2+1)(x-2)$ não o é pois $x^2 + 1$ não pode ser factorizado em factores lineares em R. No entanto, se considerarmos o corpo C, $(x^2+1)(x-2)$ é factorizável pois pode ser expresso como $(x+j)(x-j)(x-2)$. Sendo $f(t)$ o polinómio característico de um operador linear ou de uma matriz, o corpo a considerar será o corpo associado com o operador ou a matriz.

Teorema 6.9 *O polinómio característico de qualquer operador linear diagonalizável é factorizável.*

Prova: Supondo que T é diagonalizável, teremos na base β para a qual a sua representação é diagonal

$$D = [T]_\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

e sendo $f(t)$ o polinómio característico de T será

$$\begin{aligned} f(t) &= det(D-tI) = det\left(\begin{bmatrix} \lambda_1 - t & 0 & \cdots & 0 \\ 0 & \lambda_2 - t & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{bmatrix}\right) \\ &= (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t) = (-1)^n (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n). \end{aligned}$$

♣

Deste teorema decorre que se T é um operador linear diagonalizável num espaço de dimensão n e não tem n valores próprios distintos então o seu polinómio característico terá forçosamente zeros repetidos.

Definição 6.6 - *Multiplicidade algébrica de um valor próprio*
Seja λ um valor próprio de um operador linear ou matriz com polinómio característico $f(t)$. A multiplicidade algébrica *de λ é o maior inteiro positivo k para o qual $(t-\lambda)^k$ é factor de $f(t)$.*

Exemplo 6.9 Seja

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

Sendo $f(t)$ o polinómio característico de A, $f(t) = -(t-1)^2(t-2)$. Então $\lambda = 1$ é valor próprio de A com multiplicidade 2 e $\lambda = 2$ é valor próprio de A com multiplicidade 1. ♢

Sendo T um operador diagonalizável num espaço de dimensão finita V sabemos que existe uma base β para V constituída por vectores próprios de T. Sabemos ainda, do Teorema 6.4, que $[T]_\beta$ é uma matriz diagonal tal que os elementos na diagonal são valores próprios de T. Sendo o polinómio característico de T igual a $det([T]_\beta - tI)$, vemos que cada valor próprio de T deve aparecer na diagonal de $[T]_\beta$ um número de vezes igual à sua multiplicidade algébrica. Portanto β contém, para cada valor próprio, um número de vectores próprios (linearmente independentes) igual à multiplicidade desse valor próprio. Vemos portanto que o número de vectores próprios linearmente independentes correspondentes a um dado valor próprio é de grande interesse para determinar se o operador é diagonalizável. Lembrando que os vectores próprios de T para o valor próprio λ são os vectores não nulos do núcleo de $T - \lambda I$, interessa estudar este conjunto.

Definição 6.7 - *Espaço próprio*

Seja T um operador linear num espaço vectorial V e λ um seu valor próprio. Defina-se $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I_V)$. O conjunto E_λ é designado espaço próprio de T associado ao valor próprio λ. *De forma similar, o espaço próprio para uma matriz A será o espaço para o operador L_A.*

Teorema 6.10 *Seja T um operador linear num espaço vectorial de dimensão finita V. Se λ é um valor próprio de T com multiplicidade algébrica m então*

$$1 \leq dim(E_\lambda) \leq m.$$

A dimensão de E_λ designa-se também *multiplicidade geométrica de λ.*

Exemplo 6.10 Seja $T : P_2(R) \to P_2(R)$ definido por $T(f) = f'$, a derivada de f. A matriz de T nas base canónica de $P_2(R)$ é

$$[T]_\beta = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

O polinómio característico de T é

$$det([T]_\beta - tI) = det\left(\begin{bmatrix} -t & 1 & 0 \\ 0 & -t & 2 \\ 0 & 0 & -t \end{bmatrix}\right) = -t^3.$$

Portanto T tem apenas um valor próprio ($\lambda = 0$) de multiplicidade 3. Donde $E_\lambda = N(T - \lambda I) = N(T)$. E_λ é portanto o subespaço de $P_2(R)$ contendo os polinómios constantes. Neste caso $\{1\}$ é uma base para E_λ e $dim(E_\lambda) = 1$. Não existe portanto uma base para $P_2(R)$ constituída por vectores próprios de T, e T não é diagonalizável. $\diamondsuit$

Exemplo 6.11 Seja T um operador linear em R^3 definido por

$$T\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}\right) = \begin{bmatrix} 4a_1 & & & + & a_3 \\ 2a_1 & + & 3a_2 & + & 2a_3 \\ a_1 & & & + & 4a_3 \end{bmatrix}.$$

Determinemos o espaço próprio de T correspondente a cada um dos valores próprios. Sendo β a base canónica para R^3

$$[T]_\beta = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}.$$

O polinómio característico de T é

$$det([T]_\beta - tI) = det\left(\begin{bmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{bmatrix}\right) = -(t-5)(t-3)^2.$$

Os valores próprios de T são $\lambda_1 = 5$ e $\lambda_2 = 3$ com multiplicidades 1 e 2, respectivamente. Como

$$E_{\lambda_1} = N(T - \lambda_1 I) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in R^3 : \begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\},$$

E_{λ_1} é o espaço das soluções do sistema

$$\left\{ \begin{array}{rcrcrcl} -x_1 & & & + & x_3 & = & 0 \\ 2x_1 & - & 2x_2 & + & 2x_3 & = & 0 \\ x_1 & & & - & x_3 & = & 0 \end{array} \right. .$$

Resolvendo o sistema temos que $\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$ é uma base para E_{λ_1}. Logo $dim(E_{\lambda_1}) = 1$.

De forma semelhante temos que $E_{\lambda_2} = N(T - \lambda_2 I)$ é o espaço de soluções do sistema

$$\left\{ \begin{array}{rcrcl} x_1 & + & x_3 & = & 0 \\ 2x_1 & + & 2x_3 & = & 0 \\ x_1 & + & x_3 & = & 0 \end{array} \right. ,$$

donde $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ é uma base para E_{λ_2}, e portanto $dim(E_{\lambda_2}) = 2$.

Neste caso a multiplicidade de cada valor próprio iguala a dimensão do espaço próprio associado. Notar que

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

é uma base para R^3 composta apenas por vectores próprios de T. Em consequência, T é diagonalizável. $\diamondsuit$

Os exemplos sugerem que, para um operador linear cujo polinómio característico é factorizável, a diagonalizabilidade é equivalente à igualdade entre a multiplicidade geométrica e a multiplicidade algébrica para cada valor próprio.

Teorema 6.11 *Seja T um operador linear num espaço vectorial de dimensão finita V cujo polinómio característico é factorizável. Sejam $\lambda_1, \lambda_2, \ldots, \lambda_k$ os valores próprios distintos de T. Então*

(a) T é diagonalizável se e só se a multiplicidade algébrica de λ_i é igual a $dim(E_{\lambda_i})$ para todo o i.

(b) Se T é diagonalizável e S_i é uma base para E_{λ_i} para todo o i, então

$$\beta = S_1 \cup S_2 \cup \cdots \cup S_k$$

é uma base para V constituída por vectores próprios de T.

Os resultados quanto a diagonalizabilidade são resumidos nos pontos seguintes.

Teste de diagonalizabilidade

Seja T um operador linear num espaço vectorial de dimensão n. T é diagonalizável se e só se as seguintes condições se verificam:

1. O polinómio característico de T é factorizável.

2. A multiplicidade algébrica de λ é igual a $n - caract(T - \lambda I)$ para cada valor próprio λ de T.

Algoritmo de diagonalização

Seja T um operador linear diagonalizável num espaço vectorial de dimensão finita V e $\lambda_1, \lambda_2, \ldots, \lambda_k$ os valores próprios distintos de T. Para cada j, seja β_j uma base para $E_j = N(T - \lambda_j I)$ e $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. Então β é uma base para V, e $[T]_\beta$ é uma matriz diagonal.

Exemplo 6.12 Testemos a diagonalizabilidade da matriz

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \in M_{3\times 3}(R).$$

O polinómio característico de A é $det(A - tI) = -(t-4)(t-3)^2$. A tem portanto valores próprios $\lambda_1 = 4$ e $\lambda_2 = 3$ com multiplicidades 1 e 2, respectivamente. A condição 1 do teste de diagonalizabilidade é satisfeita e, uma vez que λ_1 tem multiplicidade 1, a condição 2 é satisfeita para λ_1. Resta testar a condição 2 para λ_2. Como

$$B = A - \lambda_2 I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

tem característica 2, $3 - caract(B) = 1$. Logo a condição 2 não é verificada para λ_2 e A não é diagonalizável. $\diamondsuit$

Exemplo 6.13 Seja $A = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}$.

Pretende-se mostrar que A é diagonalizável e encontrar uma matriz Q de ordem 2 tal que $Q^{-1}AQ$ seja uma matriz diagonal. Este resultado será usado para calcular A^n para qualquer inteiro positivo n.

A é diagonalizável se e só se L_A o for. O polinómio característico de A é $(t-1)(t-2)$. Portanto, A tem 2 valores próprios distintos e é diagonalizável. Para encontrar uma base β para R^2 tal que $[L_A]_\beta$ seja uma matriz diagonal, basta atender a que L_A tem valores próprios $\lambda_1 = 1$ e $\lambda_2 = 2$. É fácil encontrar bases para E_{λ_1} e E_{λ_2}. $\left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$ é uma base para E_{λ_1} e $\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ uma base para E_{λ_2}. Portanto, tomando a base

$$\beta = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

teremos

$$[L_A]_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Sabemos também que fazendo $Q = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$, teremos $Q^{-1}AQ = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. A pode portanto ser expressa como

$$A = Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^{-1}.$$

Portanto,

$$
\begin{aligned}
A^n &= \left(Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^{-1} \right)^n \\
&= Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^{-1} Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^{-1} \cdots Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^{-1} \\
&= Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^n Q^{-1} \\
&= Q \begin{bmatrix} 1^n & 0 \\ 0 & 2^n \end{bmatrix} Q^{-1} \\
&= \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2-2^n & 2-2^{n+1} \\ -1+2^n & -1+2^{n+1} \end{bmatrix} \diamondsuit
\end{aligned}
$$

6.3 Subespaços Invariantes

Vimos já que quando x é vector próprio de um operador linear T, o operador transforma o espaço gerado por $\{x\}$ em si próprio. Os subespaços que são transformados em si próprios são de grande interesse no estudo dos operadores lineares.

Definição 6.8 - *Subespaço invariante*

Seja T um operador linear num espaço vectorial V. Um subespaço W de V é dito invariante sob T *se $T(W) \subseteq W$, ou seja, se $T(x) \in W$ para todo o $x \in W$.*

Exemplo 6.14 Sendo T um operador linear num espaço vectorial V, os seguintes subespaços são invariantes sob T:

(a) $\{0\}$

(b) V

(c) R(T)

(d) N(T)

(e) E_l, para qualquer valor próprio λ de T.

$\diamondsuit$

Exemplo 6.15 Seja $T : R^3 \to R^3$ definida por

$$T(a, b, c) = (a + b, b + c, 0).$$

O plano $xy = \{(x, y, 0) : x, y \in R\}$ e o eixo dos $xx = \{(x, 0, 0) : x \in R\}$ são invariantes sob T. $\diamondsuit$

Consideremos de novo um operador linear T num espaço V, e um vector x de V não zero. O subespaço

$$W = span(\{x, T(x), T^2(x), \ldots\})$$

é designado *subespaço T-cíclico de V gerado por x*. É fácil mostrar que W é invariante sob T. W é de facto o mais pequeno subespaço de V invariante sob T que contém x.

Exemplo 6.16 Seja $T : R^3 \to R^3$ definido por

$$T(a, b, c) = (-b + c, a + c, 3c).$$

Determinemos o subespaço T-cíclico W gerado por $e_1 = (1, 0, 0)$. Uma vez que

$$T(e_1) = T(1, 0, 0) = (0, 1, 0) = e_2$$

e

$$T^2(e_1) = T(T(e_1)) = T(e_2) = (-1, 0, 0) = -e_1,$$

temos que

$$W_{e_1} = span(\{e_1, T(e_1), T^2(e_1), \ldots\}) = span(\{e_1, e_2\}) = \{(s, t, 0) : s, t \in R\}. \ \diamondsuit$$

Exemplo 6.17 Seja T o operador linear em $P(R)$ definido por $T(f) = f'$. Então o subespaço T-cíclico gerado por x^2 é $span(\{x^2, 2x, 2\}) = P_2(R)$. $\diamondsuit$

Em geral, a existência de um subespaço invariante permite definir um novo operador cujo domínio é o subespaço. Sendo T um operador linear em V e W um subespaço de V invariante sob T, então a restrição T_W de T a W é uma função de W em W, e é um operador linear em W. O operador T_W herda certas propriedades do operador T que lhe dá origem.

Teorema 6.12 *Seja T um operador linear num espaço V de dimensão finita, e W um subespaço de V invariante sob T. Então o polinómio característico de T_W é divisor do polinómio característico de T.*

Exemplo 6.18 Seja $T : R^4 \to R^4$ definida por

$$T(a, b, c, d) = (a + b + 2c - d, b + d, 2c - d, c + d),$$

e seja $W = \{(t, s, 0, 0) : t, s \in R\}$. Notar que W é um subespaço de R^4 invariante sob T, uma vez que

$$T(a, b, 0, 0) = (a + b, b, 0, 0) \in W.$$

Seja $\gamma = \{e_1, e_2\}$ uma base para W. Estenda-se γ para β, a base canónica de R^4. Então

$$B = [T_W]_\gamma = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ e } A = [T]_\beta = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Sendo $f(t)$ o polinómio característico de T e $g(t)$ o polinómio característico de T_W teremos

$$f(t) = det(A - tI_4) = det\left(\begin{bmatrix} 1-t & 1 & 2 & -1 \\ 0 & 1-t & 0 & 1 \\ 0 & 0 & 2-t & -1 \\ 0 & 0 & 1 & 1-t \end{bmatrix}\right) =$$

$$= det\left(\begin{bmatrix} 1-t & 1 \\ 0 & 1-t \end{bmatrix}\right) \cdot det\left(\begin{bmatrix} 2-t & -1 \\ 1 & 1-t \end{bmatrix}\right) =$$

$$= g(t) \cdot det\left(\begin{bmatrix} 2-t & -1 \\ 1 & 1-t \end{bmatrix}\right). \ \diamondsuit$$

Usando o teorema anterior podemos recorrer ao polinómio característico de T_W para obter informação sobre o polinómio característico de T. Os subespaços cíclicos são úteis porque o polinómio característico da restrição de T a um subespaço cíclico se obtém facilmente.

Teorema 6.13 *Seja T um operador linear num espaço V de dimensão finita, e W um subespaço T-cíclico de V gerado por x. Suponhamos que $dim(W) = k \geq 1$ (e portanto $x \neq 0$). Então:*

(a) $\{x, T(x), T^2(x), \ldots, T^{k-1}(x)\}$ é uma base de W.

(b) Sendo $T^k(x) = -a_0x - a_1T(x) - \cdots - a_{k-1}T^{k-1}(x)$ então o polinómio característico de T_W é $f(t) = (-1)^k(a_0 + a_1t + \cdots + a_{k-1}t^{k-1} + t^k)$.

Exemplo 6.19 Seja $T : R^3 \to R^3$ definida por $T(a, b, c) = (-b + c, a + c, 3c)$, a transformação usada num dos exemplos anteriores, e $W = span(\{e_1, e_2\})$, o subespaço T-cíclico gerado por e_1. Podemos obter o polinómio característico $f(t)$ de T_W de duas maneiras: usando o Teorema 6.13 ou por meio de determinantes.

(a) *Usando o Teorema 6.13.* Do exemplo anterior temos que $\{e_1, e_2\}$ gera W e que $T^2(e_1) = -e_1$. Então $f(t) = t^2 + 1$, pela parte (b) do Teorema 6.13.

(b) *Por meio de determinantes.* Seja $\beta = \{e_1, e_2\}$, que é uma base de W. Como $T(e_1) = e_2$ e $T(e_2) = -e_1$ teremos

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

e portanto

$$f(t) = det\left(\begin{bmatrix} -t & -1 \\ 1 & -t \end{bmatrix}\right) = t^2 + 1. \ \diamondsuit$$

6.4 O Teorema de Cayley-Hamilton

O teorema anterior pode ser usado para provar um dos resultados mais interessantes da teoria dos operadores lineares.

Teorema 6.14 (Cayley-Hamilton) *Seja T um operador linear num espaço vectorial V de dimensão finita, e seja $f(t)$ o polinómio característico de T. Então $f(T) = T_0$ (a transformação zero); isto é, T satisfaz o seu polinómio característico.*

Prova: Temos de mostrar que $f(T)(x) = 0$ para todo o $x \in V$. Sendo $x = 0$, então $f(T)(x) = 0$ pois $f(T)$ é uma transformação linear. Suponhamos então que $x \neq 0$, e seja W o subespaço T-cíclico de V gerado por x. Se $dim(W) = k$, então pelo Teorema 6.13 existem escalares $a_0, -a_1, \ldots, -a_{k-1}$ tais que

$$T^k(x) = -a_0 x - a_1 T(x) - \cdots - a_{k-1} T^{k-1}(x).$$

Ainda pelo Teorema 6.13 temos que

$$g(t) = (-1)^k (a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$$

é o polinómio característico de T_W. Combinando as duas equações resulta que

$$g(T)(x) = (-1)^k (a_0 I + a_1 T + \cdots + a_{k-1} T^{k-1} + T^k)(x) = 0.$$

Pelo Teorema 6.12, $g(t)$ divide $f(t)$; logo existe um polinómio $q(t)$ tal que $f(t) = q(t)g(t)$. Portanto

$$f(T)(x) = q(T)g(T)(x) = q(T)(g(T)(x)) = q(T)(0) = 0.$$

♣

Exemplo 6.20 Seja $T : R^2 \to R^2$ dado por $T(a, b) = (a + 2b, -2a + b)$, e $\beta = \{e_1, e_2\}$. Então

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

em que $A = [T]_\beta$. O polinómio característico de T é

$$f(t) = det(A - tI) = det\left(\begin{bmatrix} 1-t & 2 \\ -2 & 1-t \end{bmatrix}\right) = t^2 - 2t + 5.$$

Verifica-se facilmente que $T_0 = f(T) = T^2 - 2T + 5I$. Da mesma forma

$$f(A) = A^2 - 2A + 5I = \begin{bmatrix} -3 & 4 \\ -4 & -3 \end{bmatrix} + \begin{bmatrix} -2 & -4 \\ 4 & -2 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \ \diamondsuit$$

Este exemplo sugere também o resultado seguinte.

Corolário 6.14.1 (Teorema de Cayley-Hamilton para matrizes) *Seja A uma matriz $n \times n$ e $f(t)$ o seu polinómio característico. Então $f(A) = 0$, a matriz zero de ordem n.*

Capítulo 7

Geometria Analítica

O objectivo da Geometria Analítica é o estudo de problemas geométricos por via algébrica. Assim, cabe a esta disciplina o papel de "ponte" entre duas grandes divisões da Matemática tradicional: a Geometria, nascida da preocupação prática da medida topográfica, e a Álgebra, como ciência dos números. O problema central da geometria analítica é pois o estabelecimento da correspondência entre uma equação $\mathrm{F}(x, y) = 0$ (ou $\mathrm{F}(x, y, z) = 0$) e o conjunto, ou lugar geométrico, dos pontos do plano (do espaço) cujas coordenadas satisfazem a equação. As relações estreitas entre a geometria analítica e a álgebra vectorial levam-nos a começar por introduzir algumas noções básicas deste último tema.

7.1 Noções de Álgebra Vectorial

A abordagem à álgebra vectorial pode ser realizada por três vias diferentes:

1. geométrica;

2. analítica;

3. axiomática.

Neste capítulo vamos seguir a abordagem analítica. Assim, os vectores são descritos usando números, decorrendo as suas propriedades operatórias das dos números. Este método relaciona-se com a via geométrica pela fixação de um sistema de eixos coordenados, ou referencial. Uma abordagem axiomática foi já efectuada no capítulos anteriores quando do estudo dos espaços vectoriais.

Pontos e vectores no plano

Pontos no plano

Desde que se estabeleça no plano um sistema de eixos coordenados, cada ponto do plano pode ser caracterizado pelas suas coordenadas nesse referencial, tal como se mostra na figura 7.1. Há, portanto, uma correspondência biunívoca entre cada ponto P do plano e cada par ordenado (x, y) de números reais. O ponto O, de intersecção dos dois eixos coordenados, é designado origem do referencial, e as respectivas coordenadas são $(0, 0)$.

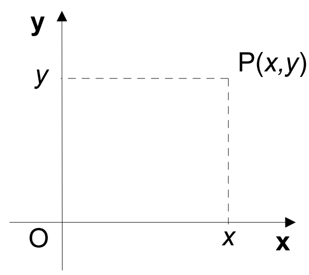

Figura 7.1: Ponto no plano.

Vectores no plano

Ao contrário das grandezas escalares, que podem ser caracterizadas apenas por um número, outras grandezas físicas, tais como força ou velocidade, exigem para a sua caracterização completa, além da amplitude, uma direcção e um sentido; estas entidades são designadas grandezas vectoriais, ou simplesmente vectores. Para especificar a velocidade de um ponto que se move no plano, é necessário determinar a amplitude da velocidade (relação entre espaço e tempo), assim como a direcção e sentido do movimento; a combinação é o vector velocidade do ponto que se move. Em termos geométricos, é conveniente representar esse vector por uma flecha localizada na posição corrente do ponto sobre a trajectória, como se ilustra na figura 7.2. Embora a flecha represente a informação pretendida - a amplitude, direcção e sentido - é um objecto geométrico e não um objecto algébrico.

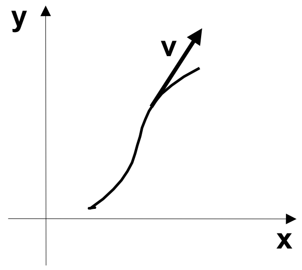

Figura 7.2: Vector no plano.

Vamos ver como a definição formal de um vector consegue combinar informação relativa à amplitude e à direcção/sentido das grandezas vectoriais.

Definição 7.1 - *Vector no plano*

Um vector no plano cartesiano é um par ordenado de números reais, da forma (a, b)*. Escrevemos* $\boldsymbol{v}(a, b)$ *e designamos* a *e* b *por componentes do vector* $\boldsymbol{v}$*.*

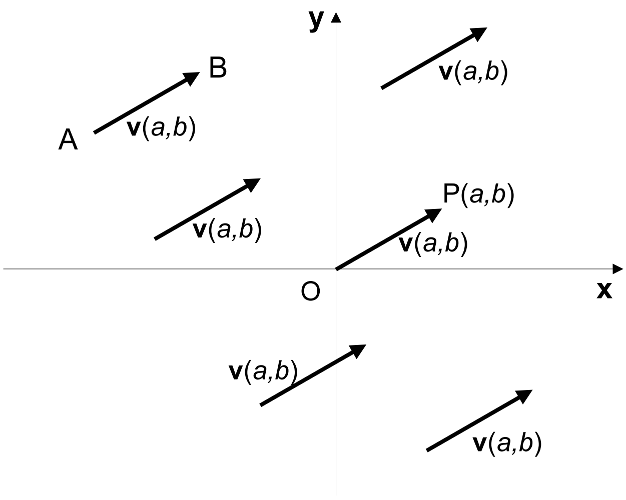

Figura 7.3: Vectores do plano.

Definição 7.2 - *Vector aplicado*

Vector ligado, vector aplicado ou segmento de recta orientado é todo o segmento de recta ao qual se atribui um determinado sentido, considerando um dos pontos extremos do segmento como origem e o outro como extremidade.

O segmento de recta orientado $\overrightarrow{\mathrm{OP}}$, da origem O para o ponto $\mathrm{P}(a, b)$, é uma representação geométrica do vector $\mathbf{v}$. Por esta razão, o vector $\mathbf{v}(a, b)$ é chamado vector de posição do ponto $\mathrm{P}(a, b)$. Notar porém que o segmento de recta orientado $\overrightarrow{\mathrm{AB}}$ é outra representação geométrica do mesmo vector $\mathbf{v}$. No entanto, este segmento orientado não é vector de posição de qualquer ponto do plano, dado que todos os vectores de posição têm a sua origem no ponto O.

A relação entre cada ponto do plano e o respectivo vector de posição é tão estreita que, em certas situações, é conveniente confundi-los deliberadamente, isto é, interpretar P e $\mathbf{v}$ como o mesmo objecto matemático; contudo, noutros contextos, a distinção entre eles é importante. Em particular, qualquer segmento de recta orientado com a mesma amplitude, direcção e sentido de $\overrightarrow{\mathrm{OP}}$ é um representação de $\mathbf{v}$, tal como é sugerido na figura 7.3.

Definição 7.3 - *Comprimento ou norma de um vector*

Comprimento do vector $\boldsymbol{v}(a, b)$, denotado por $\|\boldsymbol{v}\|$, é a amplitude associada ao vector sendo definida por

$$\|\boldsymbol{v}\| = \|(a, b)\| = \sqrt{(a^2 + b^2)}.$$

O único vector com comprimento nulo é o vector nulo $\mathbf{o}(0, 0)$; este vector não tem direcção específica.

Definição 7.4 - *Adição de vectores*

A soma de dois vectores $\boldsymbol{u}(u_1, u_2)$ e $\boldsymbol{v}(v_1, v_2)$ é o vector

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2).$$

Nas figuras 7.4 e 7.5 são apresentadas a regra do triângulo e a regra do paralelogramo, respectivamente. Estas duas regras podem ser consideradas como interpretações geométricas da adição de vectores.

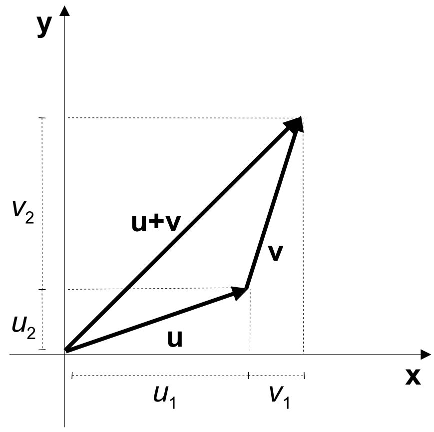

Figura 7.4: Regra do triângulo para a adição de vectores.

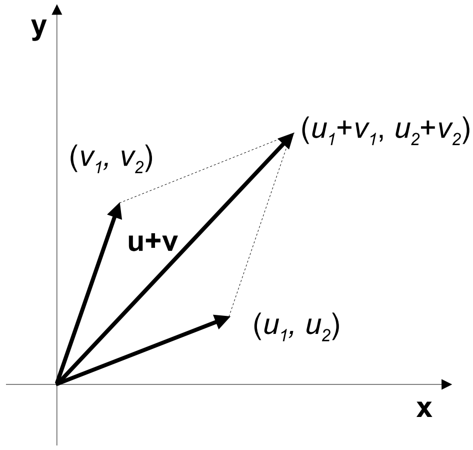

Figura 7.5: Regra do paralelogramo para a adição de vectores.

Definição 7.5 - *Multiplicação de um vector por um escalar*

Se $\boldsymbol{u}(u_1, u_2)$ *é um vector e* c *é um número real, então o produto do vector* $\boldsymbol{u}$ *pelo escalar* c *é o vector*

$$c\boldsymbol{u} = (cu_1, cu_2).$$

Relativamente a esta operação é importante salientar que:

1. O comprimento do vector produto é $|c|$ vezes o comprimento do vector **u**.

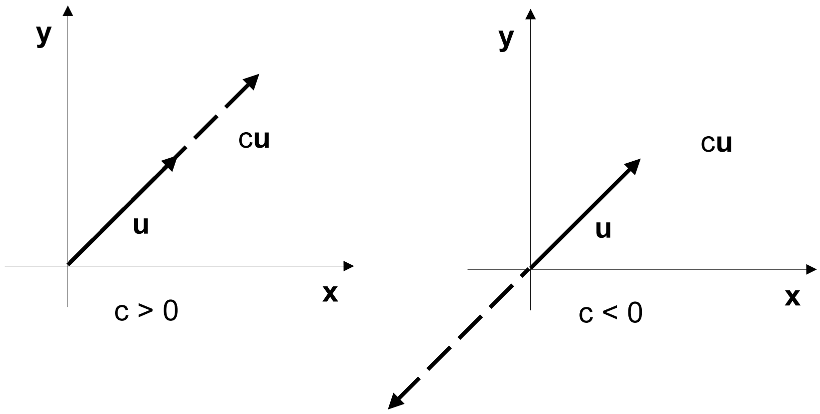

Figura 7.6: Interpretação geométrica da multiplicação escalar.

2. O simétrico do vector **u** é o vector **-u** $= (-u_1, -u_2)$; este novo vector tem o mesmo comprimento e direcção do vector **u**, mas tem sentido contrário. Logo, dois vectores não nulos **u** e **v** têm a mesma direcção se e só se $\mathbf{v} = c\mathbf{u}$; se $c > 0$, os dois vectores têm sentidos coincidentes, se $c < 0$ os vectores têm sentidos opostos .

Definição 7.6 - *Subtracção ou diferença de vectores*
A diferença de dois vectores $\boldsymbol{u}(u_1, u_2)$ *e* $\boldsymbol{v}(v_1, v_2)$ *é o vector*

$$\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (\text{-}\boldsymbol{v}) = (u_1 - v_1, u_2 - v_2).$$

Na figura 7.7 apresenta-se uma interpretação geométrica da diferença de dois vectores. Como pode ser observado, se **u** e **v** forem interpretados como vectores de posição de P e Q, respectivamente, então a diferença pode ser representada pelo vector que une os pontos Q e P.

Vectores unitários i e j

Um vector diz-se unitário se o seu comprimento é 1. Sendo $\mathbf{u}(u_1, u_2)$, então $\mathbf{a} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$ é o vector unitário com a mesma direcção e sentido de **u**, uma vez que

$$\|\mathbf{a}\| = \frac{1}{\|\mathbf{u}\|}\sqrt{(u_1^2) + (u_2^2)} = 1.$$

Dois vectores unitários particulares, $\mathbf{i}(1, 0)$ e $\mathbf{j}(0, 1)$ desempenham um papel muito especial. Com efeito, se $\mathbf{u}(u_1, u_2)$ então

$$\mathbf{u} = (u_1, 0) + (0, u_2) = u_1(1, 0) + u_2(0, 1) = u_1\mathbf{i} + u_2\mathbf{j}.$$

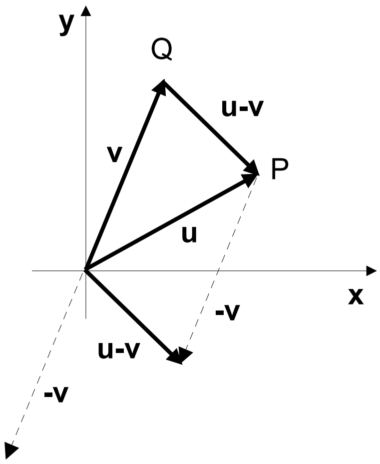

Figura 7.7: Interpretação geométrica da diferença de vectores

Assim, todo o vector é uma combinação linear dos vectores **i** e **j**. Como pode ser observado na figura 7.8, a combinação linear associada ao vector **u** pode ser interpretada geometricamente como uma decomposição do vector em duas componentes, uma horizontal e outra vertical, associadas às componentes segundo as direcções dos vectores unitários **i** e **j** , respectivamente.

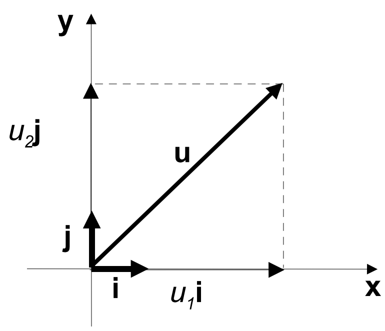

Figura 7.8: Decomposição do vector **u** nas suas componentes horizontal e vertical.

As vantagens deste tipo de representação são evidentes. Com efeito, se

$$\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} \quad \text{e} \quad \mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j},$$

então:

1. $\mathbf{u} + \mathbf{v} = (u_1\mathbf{i} + u_2\mathbf{j}) + (v_1\mathbf{i} + v_2\mathbf{j}) = (u_1 + v_1)\mathbf{i} + (u_2 + v_2)\mathbf{j}$

2. $c\mathbf{u} = c(u_1\mathbf{i} + u_2\mathbf{j}) = (cu_1)\mathbf{i} + (cu_2)\mathbf{j}$

A decomposição de uma grandeza vectorial nas suas componentes horizontal e vertical permite ainda que alguns problemas bidimensionais possam ser reduzidos à solução de dois problemas unidimensionais.

Perpendicularidade de vectores e produto interno

Dois vectores não nulos **u** e **v** são perpendiculares se, quando representados como vectores de posição de dois pontos P e Q, os segmentos orientados $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ forem perpendiculares.

Como pode ser constatado na figura 7.9, dois vectores **u** e **v** e a respectiva diferença formam um triângulo cujos vértices são os pontos O, P e Q. Se os dois vectores forem perpendiculares, tal como acontece no exemplo ilustrado na figura, pelo Teorema de Pitágoras verifica-se que $\|\mathbf{u}\|^2+\|\mathbf{v}\|^2 = \|\mathbf{u} - \mathbf{v}\|^2$. Nestas condições, é simples provar que para que a igualdade anterior seja verdadeira é necessário que $u_1v_1 + u_2v_2 = 0$. Esta combinação particular das componentes dos dois vectores é designada **produto interno** dos dois vectores.

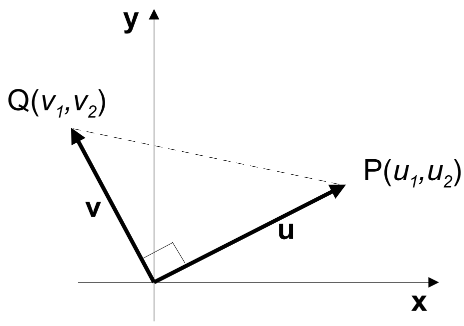

Figura 7.9: Perpendicularidade de vectores.

Definição 7.7 - *Produto interno*

O produto interno de dois vectores $\boldsymbol{u}(u_1, u_2)$ *e* $\boldsymbol{v}(v_1, v_2)$*, denotado por* $\boldsymbol{u}.\boldsymbol{v}$ *ou por* $\boldsymbol{u}|\boldsymbol{v}$*, é o número real* $u_1v_1 + u_2v_2$*.*

O produto interno, tal como foi definido antes, também é designado **produto escalar** dos dois vectores.

Teorema 7.1 - *Teste de perpendicularidade de dois vectores*

Dois vectores não nulos $\boldsymbol{u}(u_1, u_2)$ *e* $\boldsymbol{v}(v_1, v_2)$ *são perpendiculares se e só se* $u_1v_1 + u_2v_2 = 0$*, isto é se o seu produto interno for nulo.*

Propriedades do produto interno:

1. $\mathbf{u}.\mathbf{u} = \|\mathbf{u}\|^2$;
2. $\mathbf{u}.\mathbf{v} = \mathbf{v}.\mathbf{u}$;
3. $\mathbf{u}.(\mathbf{v} + \mathbf{w}) = \mathbf{u}.\mathbf{v} + \mathbf{u}.\mathbf{w}$;
4. $(a\mathbf{u}).\mathbf{v} = a(\mathbf{u}.\mathbf{v}) = \mathbf{u}.(a\mathbf{v})$.

O produto interno de dois vectores podem também ser calculado a partir dos comprimentos dos vectores e do ângulo entre as respectivas direcções (figura 7.10), tal como se indica no teorema a seguir.

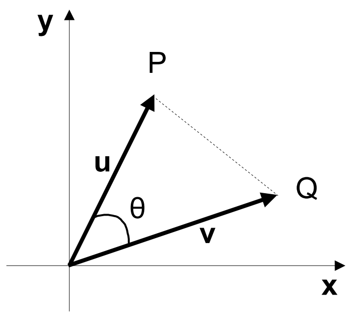

Figura 7.10: Interpretação geométrica do produto interno de dois vectores.

Teorema 7.2 *Sendo θ o ângulo entre os dois vectores $\boldsymbol{u}$ e $\boldsymbol{v}$, o respectivo produto interno é dado pela equação*

$$\boldsymbol{u}.\boldsymbol{v} = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \cos\theta \tag{7.1}$$

Projecção ortogonal de um vector

A equação (7.1), apresentada no teorema anterior, pode ser usada para determinar a componente de um vector $\mathbf{u}$ segundo a direcção de um outro vector $\mathbf{v}$.

Definição 7.8 - *Componente de um vector segundo a direcção de outro vector*
A componente de um vector $\boldsymbol{u}$ *segundo a direcção de outro vector* $\boldsymbol{v}$ *é um escalar numericamente igual ao comprimento da projecção ortogonal de* $\boldsymbol{u}$ *sobre a recta determinada por* $\boldsymbol{v}$*. Sendo* θ *o ângulo entre os dois vectores* $\boldsymbol{u}$ *e* $\boldsymbol{v}$*, este escalar é positivo se* $\theta < \pi/2$ *e negativo se* $\theta > \pi/2$*.*

A determinação geométrica da componente de um vector segundo a direcção de outro vector é ilustrada na figura 7.11.

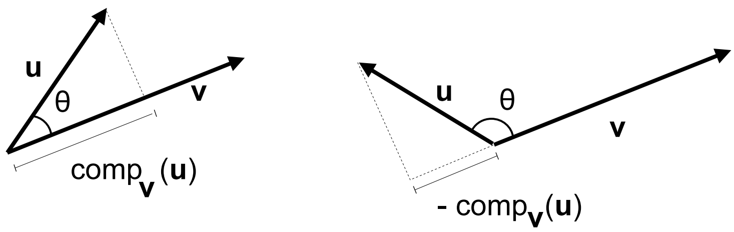

Figura 7.11: Componente de um vector $\mathbf{u}$ segundo a direcção de um vector $\mathbf{v}$.

Como pode ser observado na figura 7.11, a componente de um vector $\mathbf{u}$ segundo a direcção de um vector $\mathbf{v}$ pode ser calculada pela equação

$$\text{comp}_{\mathbf{v}}(\mathbf{u}) = \|\mathbf{u}\| \cos\theta = \frac{\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta}{\|\mathbf{v}\|} = \frac{\mathbf{u}.\mathbf{v}}{\|\mathbf{v}\|} \tag{7.2}$$

O vector $\mathbf{p}$, dado pela equação (7.3), tem a direcção de $\mathbf{v}$ e é chamado **projecção ortogonal do vector u segundo o vector v**.

$$\mathbf{p} = \text{comp}_{\mathbf{v}}(\mathbf{u}) \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{u}.\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} \tag{7.3}$$

Pontos e vectores no espaço

Todos os conceitos e definições que apresentámos na secção anterior relativas a pontos e vectores no plano podem ser estendidos se considerarmos agora como domínio o espaço tridimensional.

Pontos no espaço

Depois de ser definido um sistema de eixos coordenados no espaço, como o que se mostra na figura 7.12, é possível estabelecer uma correspondência biunívoca entre um ponto $\mathrm{P}(x, y, z)$ do espaço e o terno ordenado de números reais (x, y, z).

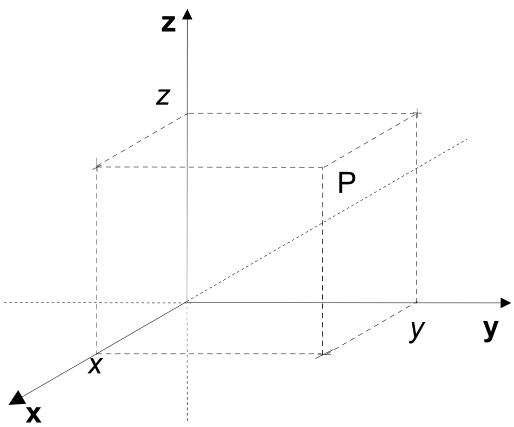

Figura 7.12: Correspondência entre um ponto do espaço e um terno de números reais.

Vectores no espaço

A discussão antes efectuada relativa aos vectores no plano pode agora ser repetida para os vectores no espaço. As definições e propriedades anteriormente apresentadas permanecem válidas, mas a diferença essencial reside no facto de um vector no espaço necessitar de três componentes para a sua representação.

Na figura 7.13 mostra-se o vector $\mathbf{u}$ determinado pelo ponto $\mathrm{P}(x, y, z)$, e que é o seu vector de posição. Usando a fórmula da distância, o comprimento do vector $\mathbf{u}$ é $\|\mathbf{u}\| = \|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}$.

Se considerarmos dois pontos arbitrários $\mathrm{A}(a_1, a_2, a_3)$ e $\mathrm{B}(b_1, b_2, b_3)$, a distância entre estes dois pontos é o comprimento do vector $\overrightarrow{\mathrm{AB}}$, com origem no ponto A e

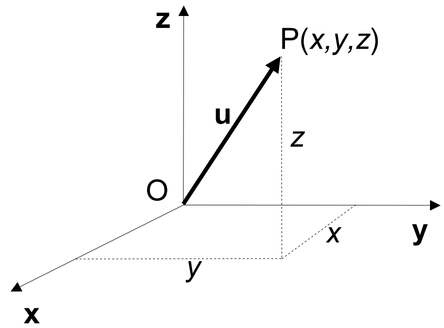

Figura 7.13: Vector de posição do ponto P de coordenadas (x,y,z).

que termina no ponto B (figura 7.14). Assim, podemos escrever que

$$\vec{\mathrm{AB}} = (b_1 - a_1, b_2 - a_2, b_3 - a_3)$$

e que o respectivo comprimento é

$$\left\| \vec{AB} \right\| = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}$$

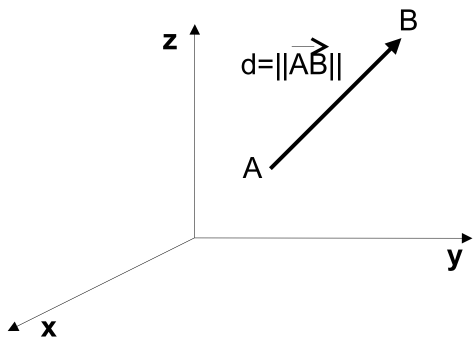

Figura 7.14: Distância entre dois pontos A e B.

À semelhança do caso bidimensional, a adição de dois vectores $\mathbf{u}(u_1, u_2, u_3)$ e $\mathbf{v}(v_1, v_2, v_3)$ é o vector

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3).$$

Porque **u** e **v** coexistem no mesmo plano (que, no caso geral, não é o plano xy), a adição de vectores obedece à mesma lei do paralelogramo, apresentada no caso bidimensional.

A multiplicação escalar de um número real c por um vector **u** é o vector

$$c\mathbf{u} = (cu_1, cu_2, cu_3).$$

Vectores unitários básicos

Os vectores unitários básicos, também designados por vectores coordenados, são agora $\mathbf{i}(1,0,0)$, $\mathbf{j}(0,1,0)$ e $\mathbf{k}(0,0,1)$.

À semelhança do que acontece no plano, qualquer vector do espaço pode ser expresso como uma combinação linear dos três vectores unitários básicos antes definidos. Assim,

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

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

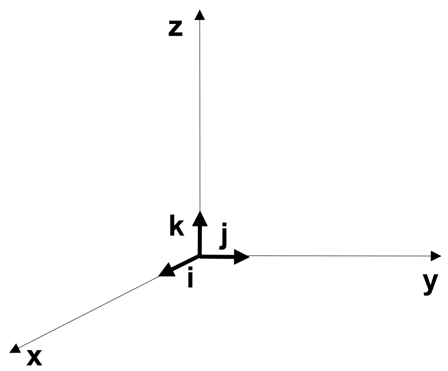

Figura 7.15: Vectores unitários básicos no espaço tridimensional.

Produto interno

O produto interno dos vectores $\mathbf{u}(u_1, u_2, u_3)$ e $\mathbf{v}(v_1, v_2, v_3)$ é o número real

$$\mathbf{u}.\mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta. \tag{7.4}$$

No caso dos dois vectores serem perpendiculares, o respectivo produto interno é nulo.

O ângulo θ na equação (7.4) é o ângulo entre os vectores $\mathbf{u}$ e $\mathbf{v}$, e pode ser calculado a partir desta equação. Com efeito,

$$\cos\theta = \frac{\mathbf{u}.\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}. \tag{7.5}$$

Os **ângulos directores** de um vector não nulo $\mathbf{u}$ são os ângulos α, β e γ, que o vector faz com os vectores unitários, $\mathbf{i}$, $\mathbf{j}$ e $\mathbf{k}$, respectivamente. Os cosenos deste ângulos são vulgarmente designados por **cosenos directores**.

A projecção ortogonal de um vector $\mathbf{u}$ sobre um vector $\mathbf{v}$, definida pela equação (7.6), é calculada a partir da determinação da componente de $\mathbf{u}$ segundo a direcção de $\mathbf{v}$ (equação (7.7)). Tal como foi definido no caso dos vectores no plano,

$$\mathbf{p} = \text{comp}_{\mathbf{v}}\left(\mathbf{u}\right)\frac{\mathbf{v}}{\left\|\mathbf{v}\right\|} = \frac{\mathbf{u}.\mathbf{v}}{\left\|\mathbf{v}\right\|^2}\mathbf{v} \tag{7.6}$$

$$\text{comp}_{\mathbf{v}}\left(\mathbf{u}\right) = \left\|\mathbf{u}\right\|\cos\theta = \left\|\mathbf{u}\right\|\left\|\mathbf{v}\right\|\cos\theta / \left\|\mathbf{v}\right\| = \mathbf{u}.\mathbf{v}/\left\|\mathbf{v}\right\| \tag{7.7}$$

Produto vectorial

Para o conjunto dos vectores do espaço está definida um nova operação cujo resultado é também um vector do espaço, denominada **produto vectorial** ou **produto externo**.

Definição 7.9 - *Produto vectorial de dois vectores*
O produto vectorial dos vectores $\boldsymbol{u}(u_1, u_2, u_3)$ *e* $\boldsymbol{v}(v_1, v_2, v_3)$ *é o vector*

$$\boldsymbol{u} \times \boldsymbol{v} = (u_2v_3 - u_3v_2,\ u_3v_1 - u_1v_3,\ u_1v_2 - u_2v_1).$$

Um processo alternativo para a determinação do resultado do produto vectorial consiste no cálculo do determinante a seguir indicado, através do desenvolvimento laplaciano aplicado à 1ª linha. Assim,

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

O vector que resulta de um produto vectorial é sempre perpendicular ao plano definido pelos dois vectores operados, e o seu comprimento depende do ângulo entre estes vectores, bem como dos respectivos módulos. Com efeito, se o ângulo entre os vectores $\mathbf{u}$ e $\mathbf{v}$ for θ , então

$$\left\|\mathbf{u} \times \mathbf{v}\right\| = \left\|\mathbf{u}\right\|\left\|\mathbf{v}\right\|\sin\theta.$$

Estes dois factos são ilustrados nas figuras 7.16 e 7.17. Nesta última figura também se ilustra o facto do comprimento do produto interno ser numericamente igual à área do paralelogramo determinado pelos dois vectores.

Uma consequência importante das relações antes apontadas é o facto do produto interno de dois vectores paralelos ser sempre nulo.

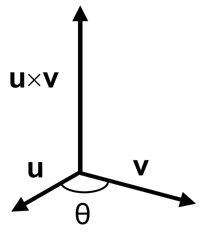

Figura 7.16: Perpendicularidade do produto vectorial.

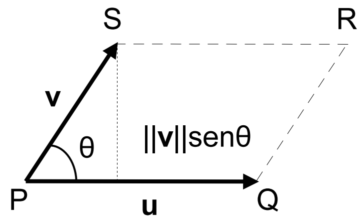

Figura 7.17: Interpretação geométrica do comprimento do produto vectorial.

Propriedades do produto vectorial:

1. $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$; (anti-comutatividade)
2. $(c\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (c\mathbf{v}) = c(\mathbf{u} \times \mathbf{v})$;
3. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$;
4. $\mathbf{u}.(\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}).\mathbf{w}$;
5. $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u}.\mathbf{w})\mathbf{v} - (\mathbf{u}.\mathbf{v})\mathbf{w}$.
6. $\mathbf{u} \times \mathbf{u} = \mathbf{o}$;

Relações entre produto interno e produto vectorial

Sendo $\mathbf{u}$ e $\mathbf{v}$ dois vectores do espaço, o produto interno dos dois vectores está relacionado com o respectivo produto vectorial pela expressão a seguir indicada.

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u}.\mathbf{v})^2$$

Produto triplo escalar

A partir das duas operações entre vectores antes referidas, é possível definir uma nova operação cujo resultado, à semelhança do produto interno, é um número real. Assim, o produto triplo escalar de três vectores $\mathbf{u}(u_1, u_2, u_3)$, $\mathbf{v}(v_1, v_2, v_3)$ e $\mathbf{w}(w_1, w_2, w_3)$ é o número real calculado pela expressão

$$\mathbf{u}.(\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}).\mathbf{w}.$$

O cálculo do produto triplo escalar pode ser efectuado recorrendo às definições de produto interno e produto vectorial antes apresentadas. Atendendo a que

$$\mathbf{v} \times \mathbf{w} = (v_2w_3 - v_3w_2,\ v_3w_1 - v_1w_3,\ v_1w_2 - v_2w_1)$$

então

$$\mathbf{u}.(\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

O valor absoluto do produto triplo escalar pode ser interpretado geometricamente como sendo igual ao volume V do paralelepípedo determinado pelos vectores $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$, tal como se mostra na figura 7.18. Como já vimos anteriormente, $A = \|\mathbf{v} \times \mathbf{w}\|$ é a área da base do paralelepípedo; a altura h pode ser calculada como $h = \|\mathbf{u}\| \, \|\cos\theta\|$, onde θ é o ângulo entre os vectores $\mathbf{u}$ e $\mathbf{v} \times \mathbf{w}$. Assim, podemos escrever que

$$V = \|\mathbf{u}.(\mathbf{v} \times \mathbf{w})\|\,.$$

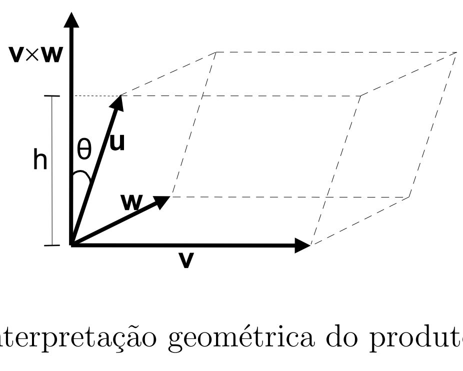

Figura 7.18: Interpretação geométrica do produto triplo escalar.

7.2 Rectas no plano

Uma recta no plano fica completamente definida depois de serem conhecidos dois dos seus pontos (P_0 e P_1, na figura 7.19), ou um dos seus pontos e um vector com a direcção da recta (P_0 ou P_1 e $\mathbf{v}$, na figura 7.19). O primeiro caso, da recta definida por dois dos seus pontos, reduz-se ao segundo se atendermos a que o vector $v = \overrightarrow{OP}_1 - \overrightarrow{OP}_0$ tem a direcção da recta.

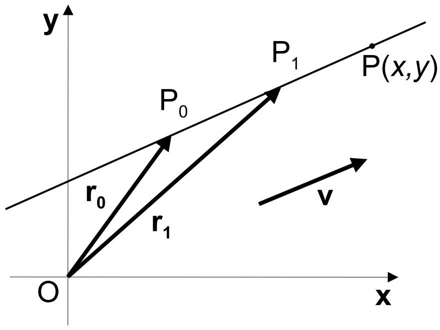

Figura 7.19: Definição de uma recta no plano.

Como pode ser observado na figura 7.19, se $P(x, y)$ for um ponto arbitrário da recta, então podemos escrever

$$\overrightarrow{P_0P} = t\mathbf{v}$$

ou ainda

$$\overrightarrow{P_0P} = \mathbf{r} - \mathbf{r}_0$$

onde $\mathbf{r}$ e $\mathbf{r}_0$ são os vectores de posição dos pontos P e P_0, respectivamente, e t é um parâmetro real. Assim, igualando as duas expressões antes apresentadas, podemos escrever

$$\mathbf{r} - \mathbf{r}_0 = t\mathbf{v}. \tag{7.8}$$

A equação (7.8), denominada **equação vectorial da recta no plano**, também pode ser apresentada na forma

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}. \tag{7.9}$$

Atendendo a que $\mathbf{r} = (x, y)$, $\mathbf{r}_0 = (x_0, y_0)$ e $\mathbf{v} = (a, b)$, a equação (7.9) pode ser reescrita como

$$(x, y) = (x_0, y_0) + t(a, b).$$

Decompondo a equação anterior obtemos as duas equações independentes nas variáveis x e y apresentadas a seguir, genericamente designadas por **equações paramétricas da recta no plano**.

$$\begin{cases} x = x_0 + ta \\ y = y_0 + tb \end{cases} \tag{7.10}$$

Se o parâmetro t for eliminado no sistema de equações (7.10) obtemos a **equação reduzida da recta no plano** que se mostra na equação (7.11).

$$y = mx + d \tag{7.11}$$

$$\text{com} \quad m = \frac{b}{a} \quad \text{e} \quad d = y_0 - \frac{b}{a}x_0.$$

7.3 Rectas no espaço

Tal como acontece com uma recta no plano, uma recta no espaço também pode ser definida por dois quaisquer dos seus pontos (P_0 e P_1). Em alternativa, uma recta pode ser especificada de forma completa se for conhecido um dos seus pontos P_0 e um vector $\mathbf{v}$ que determina a respectiva direcção.

Todas as equações para rectas no espaço assumem formas idênticas às que foram indicadas anteriormente no caso de rectas no plano. Como pode ser observado na figura 7.20, se considerarmos um ponto arbitrário na recta $\mathrm{P}(x, y, z)$ com vector de posição $\mathbf{r}$, um ponto $\mathrm{P}_0(x_0, y_0, z_0)$ com vector de posição $\mathbf{r}_0$, e um vector $\mathbf{v}$ paralelo à recta, a forma da **equação vectorial da recta no espaço** é a que se exprime na equação (7.12). Tal como se verificou no caso das rectas no plano, esta equação pode assumir a forma indicada na equação (7.13) se os vectores forem especificados através das suas componentes e t um parâmetro real.

$$\mathbf{r} - \mathbf{r}_0 = t\mathbf{v} \tag{7.12}$$

$$(x, y, z) = (x_0, y_0, z_0) + t(a, b, c). \tag{7.13}$$

O sistema de equações que pode ser derivado da equação vectorial (equação (7.13)), formado pelas **equações paramétricas da recta no espaço**, é apresentado a seguir.

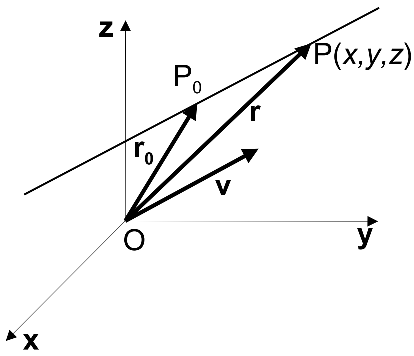

Figura 7.20: Recta no espaço.

$$\begin{cases} x = x_0 + ta \\ y = y_0 + tb \\ z = z_0 + tc \end{cases} \tag{7.14}$$

Se nas equações paramétricas de uma recta no espaço eliminarmos o parâmetro t, obtemos o resultado que se mostra na equação (7.15). Esta equação é denominada **equação cartesiana da recta no espaço**[1].

$$\text{Se } a \neq 0, \quad \begin{cases} y - y_0 = \frac{b}{a}(x - x_0) \\ z - z_0 = \frac{c}{a}(x - x_0) \end{cases} \tag{7.15}$$

Quando todas as componentes do vector de direcção são não nulas, também é possível obter as **equações simétricas da recta no espaço**, que se condensam na equação (7.16).

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \tag{7.16}$$

Se uma das componentes do vector de direcção da recta $\mathbf{v}$ for nula, por exemplo $c = 0$ mas $a \neq 0$ e $b \neq 0$, as equações simétricas assumem a forma

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}, \quad z = z_0 \text{ (a recta existe num plano paralelo ao plano xy).}$$

[1]Podem ser obtidos sistemas de equações equivalentes no caso de $b \neq 0$, ou de $c \neq 0$

Posição relativa de rectas no espaço

Dadas duas rectas no espaço, L_1 e L_2, um problema interessante é determinar a respectiva posição relativa, e em particular verificar se as rectas são paralelas.

Se as rectas estiverem definidas na forma vectorial, basta verificar a proporcionalidade dos respectivos vectores de direcção para avaliar se as rectas são paralelas. Sejam L_1 e L_2 duas rectas do espaço, com vectores de direcção $\mathbf{v}_1(a_1, b_1, c_1)$ e $\mathbf{v}_2(a_2, b_2, c_2)$, respectivamente, e sejam t e s dois parâmetros reais. Consideremos que as equações paramétricas das rectas são:

$$\text{recta } L_1 \begin{cases} x = x_1 + ta_1 \\ y = y_1 + tb_1 \\ z = z_1 + tc_1 \end{cases}$$

$$\text{recta } L_2 \begin{cases} x = x_2 + sa_2 \\ y = y_2 + sb_2 \\ z = z_2 + sc_2 \end{cases}$$

Neste caso, se:

1. $\mathbf{v}_1 = k\mathbf{v}_2$, as duas rectas são **paralelas**;

2. $\mathbf{v}_1 \neq k\mathbf{v}_2$, as duas rectas não são paralelas, sendo possível tentar encontrar um ponto de intersecção por resolução do sistema de equações que se obtém igualando, duas a duas, as equações paramétricas correspondentes das duas rectas, como se indica a seguir.

$$\begin{cases} x_1 + ta_1 = x_2 + sa_2 \\ y_1 + tb_1 = y_2 + sb_2 \\ z_1 + tc_1 = z_2 + sc_2 \end{cases}$$

   Se o sistema for possível e determinado, o ponto de intersecção das duas rectas pode ser obtido substituindo a solução obtida para o parâmetro t na equação de L_1, ou para o parâmatro s na equação de L_2. Se o sistema for impossível, diz-se que as rectas são **enviesadas** (existem em planos paralelos).

7.4 Planos no espaço

Um plano P pode ser determinado por:

1. três pontos não colineares;
2. um ponto e dois vectores não paralelos;
3. um ponto e um vector perpendicular ao plano.

Analisemos com mais detalhe cada uma das situações identificadas.

Plano definido por três pontos não colineares

Atendendo a que a partir de três pontos não colineares é trivial definir dois vectores não paralelos, esta situação é equivalente à definição de um plano usando um ponto e dois vectores não paralelos.

Plano definido por um ponto e dois vectores não paralelos

Consideremos que são conhecidos:

- um ponto do plano P_0;
- dois vectores do plano $\mathbf{v}_1$ e $\mathbf{v}_2$.

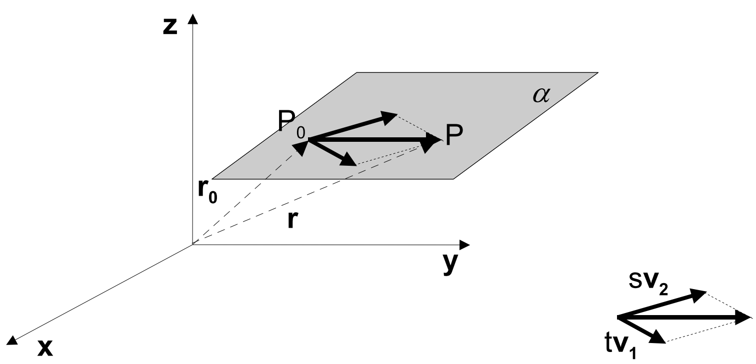

Figura 7.21: Definição de um plano a partir de um ponto e dois vectores.

Para um ponto arbitrário do plano, $P(x, y, z)$, é possível escrever (ver figura 7.21)

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}_1 + s\mathbf{v}_2 \tag{7.17}$$

A equação (7.17) é conhecida como a **equação vectorial do plano no espaço**. Se usarmos para os vectores a representação em coordenadas obtemos a equação

$$(x, y, z) = (x_0, y_0, z_0) + t(a_1, b_1, c_1) + s(a_2, b_2, c_2)$$

a partir da qual podemos estabelecer as **equações paramétricas do plano**, representadas pelo sistema 7.18.

$$\begin{cases} x = x_0 + ta_1 + sa_2 \\ y = y_0 + tb_1 + sb_2 \\ z = z_0 + tc_1 + sc_2 \end{cases} \tag{7.18}$$

A equação cartesiana pode ser obtida eliminando os parâmetros t e s nas equações paramétricas. Porém, os valores de x, y e z necessitam de ser condicionados por forma a que as três equações sejam compatíveis. Para isso basta que, para cada par de parâmetros (t, s), o sistema das três equações seja possível e determinado, isto é, que o determinante característico da terceira equação seja nulo. Assim,

$$\begin{vmatrix} x - x_0 & a_1 & a_2 \\ y - y_0 & b_1 & b_2 \\ z - z_0 & c_1 & c_2 \end{vmatrix} = 0$$

Desenvolvendo o determinante ao longo da primeira coluna obtemos

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \tag{7.19}$$

com

$$A = \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}, \quad B = -\begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}, \quad C = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

Finalmente, obtemos a **equação geral do plano** na forma que se mostra na equação (7.20).

$$Ax + By + Cz + D = 0, \quad \text{com} \quad D = -Ax_0 - By_0 - Cz_0. \tag{7.20}$$

Plano definido por um ponto e um vector normal

Como já foi referido antes, um plano também pode ser especificado de forma completa por um dos seus pontos P_0 e por um vector $\mathbf{n}(a, b, c)$ cuja direcção é normal ao plano, tal como se ilustra na figura 7.22.

Nestas condições, um ponto $\mathrm{P}(x, y, z)$ arbitrário pertence ao plano se o vector $\overrightarrow{\mathrm{P}_0\mathrm{P}}$ for perpendicular a $\mathbf{n}$. Atendendo ao que vimos anteriormente relativamente

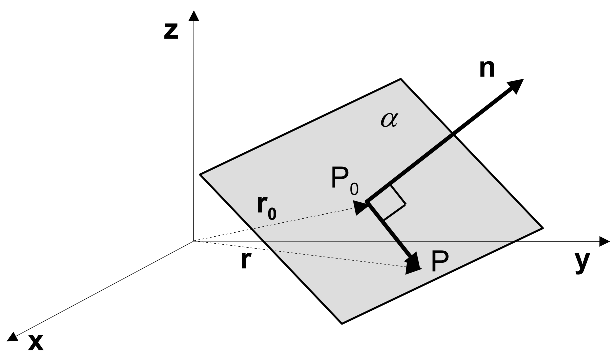

Figura 7.22: Definição de um plano a partir de um ponto e de um vector normal ao plano.

à condição de perpendicularidade de dois vectores, se $\mathbf{r}_0$ for o vector de posição do ponto P_0 e $\mathbf{r}$ o vector de posição do ponto P, então podemos escrever que o produto interno dos vectores $\mathbf{r} - \mathbf{r}_0$ e $\mathbf{n}$ deve ser nulo.

$$\mathbf{n}.(\mathbf{r} - \mathbf{r}_0) = 0 \tag{7.21}$$

A equação (7.21) é geralmente designada por **equação normal do plano**. Se nesta equação substituirmos os vectores pelas suas componentes é possível obter novamente a equação do plano na sua forma cartesiana. Com efeito,

$$(a, b, c).(x - x_0, y - y_0, z - z_0) = a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

Se compararmos as expressões apresentadas para a equações cartesiana e para a equação normal de um plano, é possível extrair alguns conclusões que se resumem a seguir:

1. Os coeficientes A, B e C das variáveis x, y e z na forma cartesiana da equação do plano são as componentes de um vector normal ao plano.

2. A equação cartesiana do plano é linear nas respectivas variáveis x, y e z.

3. Qualquer equação linear nas variáveis x, y e z representa um plano desde que os seus coeficientes não sejam todos simultaneamente nulos.

4. Se o plano for especificado por dois dos seus vectores não paralelos, $\mathbf{v}_1$ e $\mathbf{v}_2$, o respectivo produto vectorial, $\mathbf{v}_1 \times \mathbf{v}_2$ é um vector normal ao plano definido pelos dois vectores.

Intersecção de planos

Intersecção de dois planos

Sejam $\mathbf{n}_1$ e $\mathbf{n}_2$ vectores normais a cada um dos planos. Se os dois vectores forem paralelos, os dois planos também são paralelos, podendo eventualmente ser coincidentes. No caso dos vectores normais não serem paralelos, os planos intersectam-se segundo uma recta cuja equação pode ser calculada resolvendo o sistema constituído pelas equações dos dois planos.

Intersecção de três planos

A resolução deste problema passa, como no caso anterior dos dois planos, pela avaliação da posição relativa dos três vectores normais, agora designados por $\mathbf{n}_1$, $\mathbf{n}_2$ e $\mathbf{n}_3$, e pela resolução do sistema constituído pelas equações dos planos. Devem, no entanto, ser distinguidas algumas situações:

1. Se os três vectores normais são paralelos, os planos são paralelos. A avaliação da coincidência entre planos pode ser realizada com recurso à resolução do sistema das três equações dos planos:

   (a) três planos coincidentes: o sistema é possível e duplamente indeterminado (as três equações dos planos são proporcionais);

   (b) dois planos coincidentes: o sistema é impossível (duas das equações dos planos são proporcionais, mas a terceira tem termo independente não proporcional);

   (c) três planos paralelos distintos: o sistema é impossível (os termos independentes das equações não verificam a relação de proporcionalidade que necessariamente existe entre os coeficientes das variáveis).

2. Se os três vectores normais não são paralelos, os planos não são paralelos. Nestas circunstâncias podem ocorrer as seguintes situações:

   (a) os planos intersectam-se num ponto: o sistema de equações é possível e determinado e a respectiva solução fornece as coordenadas do ponto de intersecção;

   (b) os planos intersectam-se segundo uma recta: o sistema é possível e simplesmente indeterminado;

(c) os três planos são paralelos a uma determinada direcção: o sistema é impossível, não se verificando proporcionalidade entre os coeficientes de pelos menos duas das equações dos planos.

Intersecção de uma recta com um plano

As equações cartesianas de uma recta no espaço são as equações de dois planos que nela se cruzam. Deste modo, determinar a intersecção de uma recta com um plano equivale a procurar a intersecção de três planos, problema que já foi abordado na secção anterior. Contudo, se a recta estiver expressa na sua forma paramétrica, o problema pode ser simplificado, bastando substituir as equações da recta na equação do plano, resolvendo de seguida a nova equação que se obtém para determinar o valor do parâmetro correspondente ao ponto de intersecção pretendido.

7.5 Problemas métricos sobre a recta e o plano

Ao contrário do que sucedeu nas secções anteriores, vamos agora tratar de questões que envolvem distâncias e ângulos e para as quais é indispensável estabelecer uma métrica. Nesta secção vamos utilizar sempre a métrica euclidiana, à semelhança do que aconteceu na secção 7.1.

Distâncias

Distância de dois pontos

A distância entre os pontos $\mathrm{P}_1(x_1, y_1, z_1)$ e $\mathrm{P}_2(x_2, y_2, z_2)$ é definida na equação 7.22.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{7.22}$$

Distância de um ponto a uma recta

Para determinar a distância d de um ponto P_0 a uma recta L_1 basta observar que sendo A um ponto arbitrário da recta e $\mathbf{u}$ um vector com a direcção da recta, o produto vectorial do vector $\overrightarrow{\mathrm{AP}_0}$ com o vector $\mathbf{u}$ permite obter o valor pretendido (figura 7.23).

Atendendo a que

$$\left\|\mathbf{u}\times \overrightarrow{\mathrm{AP}_0}\right\| = \|\mathbf{u}\| \left\|\overrightarrow{\mathrm{AP}_0}\right\| \sin\theta = \|\mathbf{u}\|\, d$$

podemos concluir que

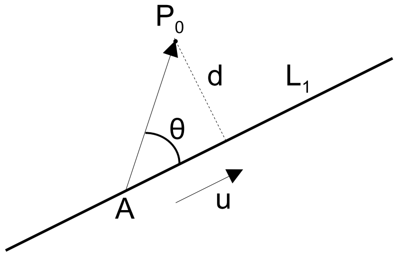

Figura 7.23: Distância de um ponto a uma recta.

$$d = \frac{\left\| \mathbf{u} \times \overrightarrow{\mathrm{AP}_0} \right\|}{\|\mathbf{u}\|} \ .$$

Distância de um ponto a um plano

A distância de um ponto P_0 a um plano α é a distância entre P_0 e a sua projecção ortogonal sobre o plano (ponto P' na figura 7.24), ou seja, a distância entre P_0 e o ponto de intersecção do plano com a recta normal ao plano que contém P_0.

A equação a seguir constitui um processo alternativo de cáculo desta distância. Sendo $\mathbf{n}$ um vector normal ao plano e A um ponto qualquer do plano, a distância é numericamente igual ao comprimento da projecção do vector $\overrightarrow{\mathrm{PA}}$ segundo a direcção do vector normal.

$$d = \frac{\left\| \mathbf{n} . \overrightarrow{\mathrm{PA}} \right\|}{\|\mathbf{n}\|} \ .$$

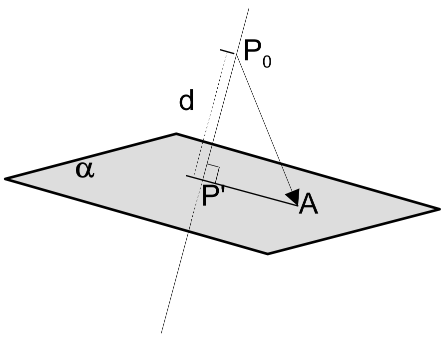

Figura 7.24: Distância de um ponto a um plano.

Um caso particular importante do cálculo da distância de um ponto a um plano é a determinação da distância de um plano à origem do sistema de referência, isto é ao ponto $(0,0,0)$.

Distância entre dois planos

O cálculo da distância entre dois planos só faz sentido quando os dois planos são paralelos. Nestas circunstâncias, a distância é calculada entre os pontos de intersecção de cada um dos planos com um recta que lhes é perpendicular. Considerando a definição de distância de um plano à origem, então a distância entre os dois planos é a diferença das respectivas distâncias à origem.

Ângulos

Ângulo de uma recta com um plano

O ângulo θ de uma recta com um plano é calculado a partir do ângulo entre o vector normal ao plano e o vector de direcção da recta (figura 7.25). No caso deste ângulo ser agudo, θ é o seu complementar; se for obtuso, θ é o complementar do suplementar do ângulo calculado.

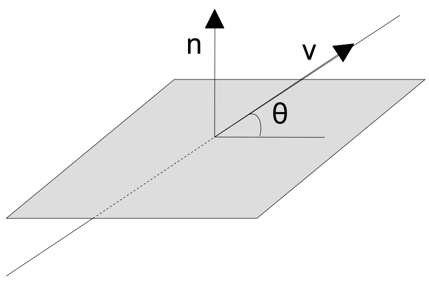

Figura 7.25: Ângulo entre uma recta e um plano.

Ângulo entre dois planos

O ângulo θ entre dois planos (figura 7.26) é igual ao ângulo entre os respectivos vectores normais, ou o seu suplementar quando este for obtuso.

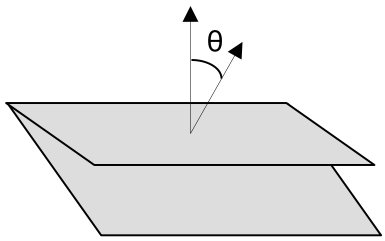

Figura 7.26: Ângulo entre dois planos.

Bibliografia

[1] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra, 2nd edition.* Prentice-Hall International Editions, 1989.

[2] Edward M. Landesman and Magnus R. Hestenes. *Linear Algebra for Mathematics, Science and Engineering.* Prentice-Hall International Editions, 1992.

[3] F. R. Dias Agudo. *Introdução à Álgebra Linear e Geometria Analítica.* Escolar Editora, 1996.

[4] C. H. Edwards and David E. Penney. *Calculus and Analytic Geometry, 3rd edition.* Prentice Hall International, 1990.