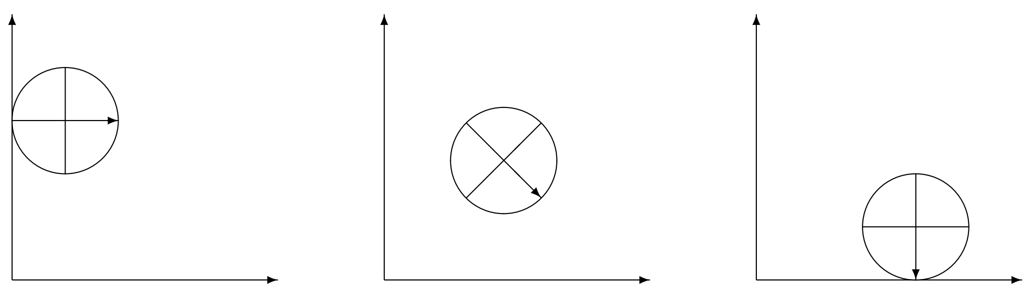
ÁLGEBRA LINEAR
e
Geometria Analítica
TEXTO TEÓRICO



Rosário Fernandes • Fátima Rodrigues
Departamento de Matemática
Faculdade de Ciências e Tecnologia
UNL
(2008/09)

Índice

1 SISTEMAS DE EQUAÇÕES LINEARES **1**

1.1 Introdução aos Sistemas de Equações Lineares 1
1.2 Solução e Conjunto-Solução de um Sistema 3
1.3 Interpretação Geométrica do Conjunto-Solução de um Sistema 4
1.4 Método de Eliminação de Gauss 6
1.5 Sistemas e Matrizes 8
1.6 Resolução de Sistemas por Transformações Elementares nas Linhas 9
1.7 Característica de uma Matriz 12
1.8 Discussão de um Sistema 14
1.9 Aplicações dos Sistemas de Equações Lineares 17
1.9.1 Cores dos Ecrãs de Televisão 17
1.9.2 Análise de Redes 19
1.9.3 Circuitos Eléctricos 20

2 OPERAÇÕES COM MATRIZES **23**

2.1 Definições Básicas e Exemplos de Matrizes 23
2.2 Operações com Matrizes 25
2.2.1 Soma de Matrizes e Produto de um Escalar por uma Matriz 26
2.2.2 Produto de uma Matriz por uma Matriz Coluna 27
2.2.3 Produto de Duas Matrizes 30
2.2.4 Transposta de uma Matriz 33
2.2.5 Inversa de uma Matriz 35
2.2.6 Potências de uma Matriz 36
2.3 Matrizes Elementares 37
2.4 Caracterização das Matrizes Invertíveis 39
2.5 Aplicações das Matrizes 41
2.5.1 Uma Aplicação à Robótica 41
2.5.2 Resolução de Sistemas 42

3 SUBESPAÇOS VECTORIAIS DE $\mathbb{R}^n$ **43**

3.1 Definição de Subespaço Vectorial de $\mathbb{R}^n$ 43
3.2 Subespaço Gerado 47
3.3 Dependência e Independência Linear 48
3.4 Aplicações 51
3.4.1 Som de Alta Fidelidade 51
3.5 Variedades Lineares 52

4 DETERMINANTES **55**
4.1 Definição de Determinante . . . 55
4.2 Teorema de Laplace . . . 57
4.3 Determinante e Transformações Elementares . . . 60
4.4 Outra Caracterização das Matrizes Invertíveis . . . 63
4.5 Sistemas de Cramer . . . 66
4.6 Determinante do Produto de Matrizes . . . 67
4.7 Interpretação Geométrica de Determinantes 2×2 . . . 70
4.8 Produto Externo e Produto Misto de Vectores de $\mathbb{R}^3$ . . . 71
4.9 Valores e Vectores Próprios de Matrizes . . . 75

5 APLICAÇÕES LINEARES **81**
5.1 Definições e Notações . . . 81
5.2 Definição de Aplicação Linear . . . 82
5.3 Matriz Canónica de uma Aplicação Linear . . . 85
5.4 Exemplos de Aplicações Lineares de $\mathbb{R}^2$ em $\mathbb{R}^2$ . . . 90
5.4.1 Rotações em torno da origem . . . 90
5.4.2 Reflexões através de rectas que passam pela origem . . . 92
5.4.3 Compressões e Expansões Horizontais e Verticais em $\mathbb{R}^2$ . . . 93
5.4.4 Alongamentos . . . 95
5.5 Núcleo e Imagem de uma Aplicação Linear . . . 96
5.6 Composição de Aplicações . . . 102
5.7 Composição de Aplicações e Matrizes Elementares . . . 105
5.8 Aplicações Lineares Invertíveis . . . 107

6 BASES **111**
6.1 Definição e Exemplos de Bases . . . 111
6.2 Dimensão de um Subespaço . . . 114
6.3 Alguns Resultados sobre Bases . . . 115
6.4 Coordenadas em Relação a uma Base . . . 121
6.5 Matriz de uma Aplicação Linear . . . 122
6.6 Matriz Mudança de Base . . . 124

7 DIAGONALIZAÇÃO **130**
7.1 Diagonalização de Matrizes Quadradas . . . 130
7.2 Classificação de Cónicas de $\mathbb{R}^2$ . . . 137
7.2.1 Formas Quadráticas de $\mathbb{R}^2$ . . . 137
7.2.2 Método para Classificar uma Cónica de $\mathbb{R}^2$ . . . 140

8 ESPAÇOS VECTORIAIS **147**
8.1 Conceitos principais . . . 147
8.2 Aplicações Lineares . . . 152

BIBLIOGRAFIA **155**

Capítulo 1

SISTEMAS DE EQUAÇÕES LINEARES

A resolução de sistemas de equações lineares e a interpretação geométrica das suas soluções constituem dois dos principais tópicos estudados em Álgebra Linear. Neste capítulo abordaremos um processo sistemático de resolução de sistemas de equações lineares e daremos alguns exemplos de aplicação dos sistemas.

1.1 Introdução aos Sistemas de Equações Lineares

Nesta primeira secção iremos introduzir alguma terminologia básica.

Uma equação da recta em $\mathbb{R}^2$ pode ser dada pela expressão

$$a_1x + a_2y = b,$$

em que a_1, a_2 não são ambos nulos, e uma equação geral do plano em $\mathbb{R}^3$, pela expressão

$$a_1x + a_2y + a_3z = b,$$

em que a_1, a_2, a_3 não são todos nulos. Estas duas equações são exemplos de equações lineares.

Definição 1.1 *Uma* **equação linear** *nas n variáveis $x_1, x_2, ..., x_n$ é uma equação da forma*

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

em que $a_1, a_2, ..., a_n$ são constantes (a que chamaremos **coeficientes***) não todas nulas e b é outra constante.*

No caso de $b = 0$, teremos a equação linear

$$a_1x_1 + a_2x_2 + ... + a_nx_n = 0,$$

que é denominada **equação linear homogénea***.*

Observação

1. Uma equação linear não envolve produto de variáveis, potências de variáveis (excepto a potência de expoente 1) nem variáveis como argumento de funções.

   As equações seguintes não são equações lineares:

$$\begin{array}{l} x + \cos y - 2arcsen\ z = 0; \\ xy - 3z = 5; \\ x^2 = 3; \\ \sqrt{x} + 4y = 2. \end{array}$$

   As equações seguintes são equações lineares:

$$\begin{array}{l} 5x + 2y = 2; \\ x_1 + 2x_2 + 4x_3 - x_4 = 0. \end{array}$$

2. As variáveis de uma equação linear são normalmente designadas pelas letras pequenas x, y, z, w, com ou sem índices.

Podemos agora definir sistema de equações lineares.

Definição 1.2 *A uma colecção finita de equações lineares chama-se* **sistema de equações lineares***, ou simplesmente* **sistema**. *As variáveis do sistema de equações lineares designam--se por* **incógnitas**.

Um sistema de m *equações lineares a* n *incógnitas,* $x_1, x_2, ..., x_n$*, pode escrever-se na forma*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n & = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n & = b_m \end{cases} \tag{1.1}$$

em que o coeficiente a_{ij} *está associado à i-ésima equação do sistema,*

$$a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n = b_i$$

e à j-ésima incógnita do sistema, x_j.

As constantes $b_1, b_2, ..., b_m$ *designam-se por* **termos independentes do sistema**.

Se no sistema (1.1) *tivermos* $b_1 = b_2 = ... = b_m = 0$*, dizemos que o sistema é* **homogéneo**.

Exemplo 1.3

$$\begin{cases} x_1 + x_2 + x_3 & = 2 \\ x_1 - x_2 + x_3 & = 0 \end{cases} \tag{1.2}$$

é um sistema de 2 equações lineares a 3 incógnitas, x_1, x_2, x_3.

1.2 Solução e Conjunto-Solução de um Sistema

Depois da definição de sistema de equações lineares, é importante sabermos identificar uma solução de um sistema e como classificar um sistema consoante o número de soluções do mesmo.

Definição 1.4 *Uma solução de um sistema de equações lineares nas incógnitas* $x_1, x_2, ..., x_n$ *é uma sequência de n números* $s_1, s_2, ..., s_n$ *tais que, substituindo em cada equação* $x_1, x_2, ..., x_n$ *por* $s_1, s_2, ..., s_n$*, respectivamente, tornam cada equação do sistema numa proposição verdadeira.*

Observação A solução nula é solução de qualquer sistema homogéneo.

Exemplo 1.5 *Usando o sistema do* Exemplo 1.3,

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

temos que

$$x_1 = 0, \ x_2 = 1, \ x_3 = 1$$

é uma solução do sistema, pois, $0 + 1 + 1 = 2$, $0 - 1 + 1 = 0$, *mas,*

$$x_1 = 1, \ x_2 = 0, \ x_3 = 1$$

não é solução do sistema, pois, $1 + 0 + 1 = 2$ *mas* $1 - 0 + 1 \neq 0$.

Definição 1.6 *O conjunto formado por todas as soluções de um sistema chama-se* **conjunto-solução**.

Considerando um sistema de equações lineares teremos 3 hipóteses para o seu conjunto-solução:

1. Se o sistema não tiver soluções, dizemos que o sistema é **impossível**. Neste caso, o conjunto-solução é vazio.

2. Se o sistema tiver uma, e uma só solução, dizemos que o sistema é **possível e determinado**. Neste caso, o conjunto-solução é constituído pela única solução do sistema.

3. Se o sistema tiver mais do que uma solução, dizemos que o sistema é **possível e indeterminado**.

1.3 Interpretação Geométrica do Conjunto-Solução de um Sistema

Se expressarmos uma solução $x_1 = s_1$, $x_2 = s_2$,..., $x_n = s_n$ do sistema (1.1), como o n-uplo ordenado $(s_1, s_2, ..., s_n)$ (por exemplo, a solução do sistema (1.2), atrás referida, $x_1 = 0$, $x_2 = 1$, $x_3 = 1$, pode ser expressa pelo terno $(0, 1, 1)$), podemos pensar nas soluções do sistema (1.1) como pontos de $\mathbb{R}^n$, abrindo a possibilidade a uma interpretação geométrica do conjunto-solução do sistema.

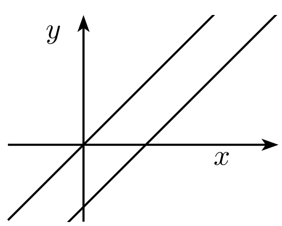
A intersecção de 2 rectas em $\mathbb{R}^2$ dá origem a um sistema de 2 equações lineares a 2 incógnitas:

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

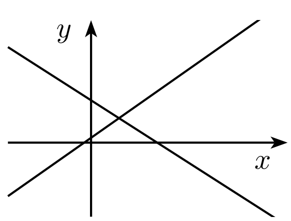
Cada uma destas 2 equações lineares representa uma recta do plano Oxy e, portanto, cada solução deste sistema corresponde a um ponto de intersecção destas rectas.

Assim, existem 3 possibilidades:

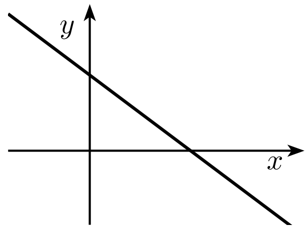
1. As rectas podem ser paralelas e distintas, sendo o sistema impossível (sem solução).



2. As rectas podem ser concorrentes (intersectam-se somente num ponto), sendo o sistema possível e determinado.



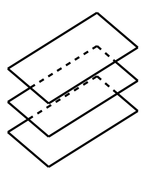
3. As rectas podem ser coincidentes e, neste caso, o sistema é possível e indeterminado (as soluções são todos os pontos da recta comum).



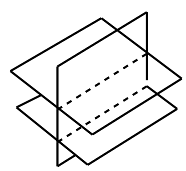
Se pensarmos na intersecção de 3 planos em $\mathbb{R}^3$, teremos o sistema

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

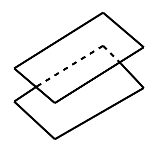
Como podemos ver existem 3 possibilidades: nenhuma solução, uma única solução ou uma infinidade de soluções.



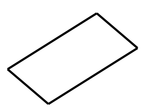
nenhuma solução
(3 planos paralelos)



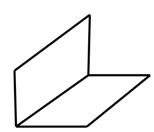
nenhuma solução
(2 planos paralelos)



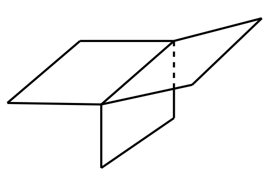
nenhuma solução
(2 coincidentes e
paralelos ao 3º)



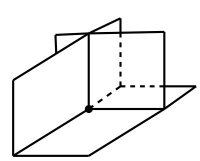
uma infinidade de soluções
(3 planos coincidentes)
(a intersecção é um plano)



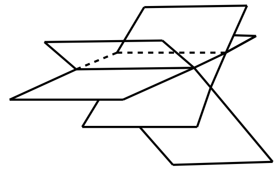
uma infinidade de soluções
(2 planos coincidentes)
(a intersecção é uma recta)



uma infinidade de soluções
(a intersecção é uma recta)



uma solução



nenhuma solução

1.4 Método de Eliminação de Gauss

Vejamos alguns exemplos de como determinar o conjunto-solução de um sistema.

1. Consideremos 3 planos em $\mathbb{R}^3$ de equações

$$3x + 3y + z = 4; \quad x + 2y - z = 0; \quad 2x + y + z = 3,$$

e determinemos a sua intersecção.
O sistema formado pelas 3 equações é

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

Multiplicando a primeira equação por $-\frac{1}{3}$ e adicionando-a à segunda, obtemos o sistema

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

Multiplicando a primeira equação por $-\frac{2}{3}$ e adicionando-a à terceira, temos

$$\begin{cases} 3x + 3y + z = 4 \\ y - \frac{4}{3}z = -\frac{4}{3} \\ -y + \frac{1}{3}z = \frac{1}{3}. \end{cases}$$

Somando a segunda equação à terceira, vem

$$\begin{cases} 3x + 3y + z = 4 \\ y - \frac{4}{3}z = -\frac{4}{3} \\ z = 1. \end{cases}$$

Resolvendo a partir da última equação e substituindo nas anteriores, obtemos a solução do sistema

$$x = 1, \quad y = 0, \quad z = 1.$$

Portanto, o conjunto-solução deste sistema é

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

Ou seja, a intersecção dos 3 planos em $\mathbb{R}^3$ é o ponto $(1,0,1)$.

Observação O método que utilizámos para resolver o sistema anterior é designado por **método de eliminação de Gauss**.

Pode acontecer que um sistema de equações lineares tenha determinados coeficientes e/ou termos independentes, que em vez de serem constantes, sejam variáveis.

2. Determinemos para que valores de a e b, as duas rectas em $\mathbb{R}^2$, cujas equações são

$$3x+3y=b \qquad \text{e} \qquad x+ay=\frac{1}{3},$$

coincidem.

Como vimos anteriormente, as rectas são coincidentes se o sistema, formado pelas suas equações, for possível e indeterminado. Vamos resolver o sistema

$$\begin{cases} 3x+3y = b \\ x+ay = \frac{1}{3} \end{cases},$$

utilizando o método de eliminação de Gauss.

Multiplicando a primeira equação por $-\frac{1}{3}$ e adicionando-a à segunda, temos

$$\begin{cases} 3x+3y = b \\ (a-1)y = -\frac{b}{3}+\frac{1}{3}. \end{cases}$$

Repare-se que se $a \neq 1$, olhando para a segunda equação, y tem um único valor. Substituindo esse valor de y na primeira equação, x terá, também, um único valor. Ou seja, nestas condições, o sistema era possível e determinado e as rectas concorrentes. Como queremos que elas sejam coincidentes, então $a=1$ e o sistema fica

$$\begin{cases} 3x+3y = b \\ 0 = -\frac{b}{3}+\frac{1}{3}. \end{cases}$$

Se $b \neq 1$, então o sistema é impossível (2^{a} equação) e as rectas paralelas. Portanto, $b=1$. Assim sendo, as rectas são coincidentes se $a=b=1$.

Vejamos como é o conjunto-solução do sistema neste caso. Se $a=b=1$, o sistema é

$$\begin{cases} 3x+3y = 1 \\ 0 = 0. \end{cases}$$

Então, $x=\frac{1}{3}-y$.

Assim, o conjunto-solução do sistema é

$$\left\{\left(\frac{1}{3}-y, y\right) : y \in \mathbb{R}\right\}.$$

1.5 Sistemas e Matrizes

Como vimos no último exemplo, além das incógnitas de um sistema podem surgir outras variáveis, o que, atendendo muitas vezes ao tamanho do sistema, dificulta a sua resolução. Um processo para contornar esta dificuldade é escrever o sistema (1.1) de m equações lineares a n incógnitas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n & = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n & = b_m \end{cases}$$

como um quadro com $m \times (n+1)$ números, dispostos por m linhas e $n+1$ colunas, da seguinte forma

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & ... & a_{mn} & b_m \end{array}\right]$$

que se designa por **matriz ampliada** do sistema e é denotada por $[A|B]$.

Observação Em Matemática chama-se **matriz** a qualquer quadro de números deste tipo.

A **matriz simples** do sistema é a matriz

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

a **matriz dos termos independentes** é a matriz

$$B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix},$$

e a **matriz das incógnitas** do sistema é a matriz

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

Exemplo 1.7 *O sistema* $\begin{cases} 3x + 3y = b \\ x + ay = \frac{1}{3} \end{cases}$ *tem associadas as matrizes:*

- *simples do sistema* $A = \begin{bmatrix} 3 & 3 \\ 1 & a \end{bmatrix}$

- *dos termos independentes* $B = \begin{bmatrix} b \\ \frac{1}{3} \end{bmatrix}$

- *das incógnitas* $X = \begin{bmatrix} x \\ y \end{bmatrix}$

- *ampliada do sistema* $[A|B] = \left[\begin{array}{cc|c} 3 & 3 & b \\ 1 & a & \frac{1}{3} \end{array}\right]$.

1.6 Resolução de Sistemas por Transformações Elementares nas Linhas

Como vimos, para resolver um sistema pelo método de eliminação de Gauss utilizamos três tipos de operações:

1. Multiplicamos uma equação por uma constante não nula;
2. trocamos duas equações de posição;
3. somamos um múltiplo de uma equação a outra equação.

Como as linhas da matriz ampliada de um sistema correspondem às equações do sistema associado, estas operações correspondem às operações nas linhas da matriz:

1. Multiplicar uma linha por uma constante não nula ;
2. trocar duas linhas;
3. somar um múltiplo de uma linha a outra linha.

Estas operações nas linhas da matriz são designadas por **transformações elementares nas linhas**.

Para simplificar usaremos as seguintes notações:

1. $l_i \to al_i$ para designar que multiplicámos a linha i pela constante, não nula, a.

2. $l_i \leftrightarrow l_j$ para designar que trocámos a linha i pela linha j.

3. $l_i \to (l_i + bl_j)$ para designar que somámos à linha i, a linha j depois de multiplicada por b.

Exemplo 1.8 1. *Do lado esquerdo iremos resolver o sistema e do lado direito iremos efectuar as mesmas transformações elementares sobre linhas.*

sistema	*matriz ampliada*
$\begin{cases} \quad y + 2z = \frac{5}{2} \\ 2x \quad + z = 2 \\ x - y + z = 1 \end{cases}$	$\left[\begin{array}{ccc\|c} 0 & 1 & 2 & \frac{5}{2} \\ 2 & 0 & 1 & 2 \\ 1 & -1 & 1 & 1 \end{array}\right]$
trocando a 1^{a} *com a* 3^{a} *equação*	$\downarrow l_1 \leftrightarrow l_3$
$\begin{cases} x - y + z = 1 \\ 2x \quad + z = 2 \\ \quad y + 2z = \frac{5}{2} \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 1 \\ 2 & 0 & 1 & 2 \\ 0 & 1 & 2 & \frac{5}{2} \end{array}\right]$
multiplicando a 1^{a} *equação por* -2 *e somando à* 2^{a}	$\downarrow l_2 \to (l_2 - 2l_1)$
$\begin{cases} x - y + z = 1 \\ \quad 2y - z = 0 \\ \quad y + 2z = \frac{5}{2} \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 1 \\ 0 & 2 & -1 & 0 \\ 0 & 1 & 2 & \frac{5}{2} \end{array}\right]$
multiplicando a 2^{a} *equação por* $-\frac{1}{2}$ *e somando à* 3^{a}	$\downarrow l_3 \to (l_3 - \frac{1}{2}l_2)$
$\begin{cases} x - y + z = 1 \\ \quad 2y - z = 0 \\ \quad \frac{5}{2}z = \frac{5}{2} \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & -1 & 1 & 1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & \frac{5}{2} & \frac{5}{2} \end{array}\right]$

Até aqui foi utilizado o método de eliminação de Gauss. Continuemos o processo, mas agora da última equação para a primeira. Este processo é conhecido por **método de eliminação de Gauss-Jordan**.

multiplicando a 3ª *equação por* $\frac{2}{5}$ $\qquad \downarrow l_3 \to \frac{2}{5} l_3$

$$\left\{\begin{array}{rrrl} x & -y & +z & = 1 \\ & 2y & -z & = 0 \\ & & z & = 1 \end{array}\right. \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

somando a 3ª *equação à* 2ª $\qquad \downarrow l_2 \to (l_2 + l_3)$

$$\left\{\begin{array}{rrrl} x & -y & +z & = 1 \\ & 2y & & = 1 \\ & & z & = 1 \end{array}\right. \qquad \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

multiplicando a 3ª *equação por* -1 *e somando à* 1ª $\qquad \downarrow l_1 \to (l_1 - l_3)$

$$\left\{\begin{array}{rrrl} x & -y & & = 0 \\ & 2y & & = 1 \\ & & z & = 1 \end{array}\right. \qquad \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

multiplicando a 2ª *equação por* $\frac{1}{2}$ $\qquad \downarrow l_2 \to \frac{1}{2} l_2$

$$\left\{\begin{array}{rrrl} x & -y & & = 0 \\ & y & & = \frac{1}{2} \\ & & z & = 1 \end{array}\right. \qquad \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

somando a 2ª *equação à* 1ª $\qquad \downarrow l_1 \to (l_2 + l_1)$

$$\left\{\begin{array}{rrrl} x & & & = \frac{1}{2} \\ & y & & = \frac{1}{2} \\ & & z & = 1 \end{array}\right. \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

A solução é $x = \frac{1}{2}$, $y = \frac{1}{2}$, $z = 1$, *ou em termos de conjunto-solução*

$$\left\{\left(\frac{1}{2}, \frac{1}{2}, 1\right)\right\}.$$

2. *Vamos resolver o sistema*

$$\left\{\begin{array}{rrrl} x & -y & +z & = 1 \\ 2x & & +z & = 2 \\ 3x & -y & +2z & = 3. \end{array}\right. \tag{1.3}$$

A matriz ampliada do sistema é

$$\left[\begin{array}{ccc|c}1 & -1 & 1 & 1\\ 2 & 0 & 1 & 2\\ 3 & -1 & 2 & 3\end{array}\right] \underset{l_2 \to (l_2 - 2l_1)}{\longrightarrow} \left[\begin{array}{ccc|c}1 & -1 & 1 & 1\\ 0 & 2 & -1 & 0\\ 3 & -1 & 2 & 3\end{array}\right] \underset{l_3 \to (l_3 - 3l_1)}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|c}1 & -1 & 1 & 1\\ 0 & 2 & -1 & 0\\ 0 & 2 & -1 & 0\end{array}\right] \underset{l_3 \to (l_3 - l_2)}{\longrightarrow} \left[\begin{array}{ccc|c}1 & -1 & 1 & 1\\ 0 & 2 & -1 & 0\\ 0 & 0 & 0 & 0\end{array}\right] \underset{l_2 \to \frac{1}{2}l_2}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|c}1 & -1 & 1 & 1\\ 0 & 1 & -\frac{1}{2} & 0\\ 0 & 0 & 0 & 0\end{array}\right] \underset{l_1 \to (l_1 + l_2)}{\longrightarrow} \left[\begin{array}{ccc|c}1 & 0 & \frac{1}{2} & 1\\ 0 & 1 & -\frac{1}{2} & 0\\ 0 & 0 & 0 & 0\end{array}\right]$$

e o sistema que corresponde a esta última matriz ampliada é

$$\begin{cases} x \quad +\frac{1}{2}z = 1 \\ \quad y -\frac{1}{2}z = 0 \\ \qquad 0 \quad = 0. \end{cases}$$

Porque x e y correspondem aos primeiros elementos não nulos da 1^{a} e 2^{a} linhas da matriz ampliada, dizemos que estas variáveis são **líderes**. *As outras variáveis (que neste caso é unicamente o z) dizem-se* **livres**.

Resolvendo o sistema em ordem às variáveis livres obtemos

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

Pelo que o conjunto-solução é

$$\left\{ \left(1 - \frac{1}{2}z, \frac{1}{2}z, z \right) : z \in \mathbb{R} \right\}.$$

1.7 Característica de uma Matriz

Quando resolvemos um sistema usando a sua matriz ampliada, o que fazemos é modificar a matriz inicial através de transformações elementares nas linhas da matriz, até obtermos uma matriz "especial". Estas matrizes "especiais"que se obtém, permitem-nos determinar mais facilmente o conjunto-solução do sistema.

Definição 1.9 *Se uma linha, de uma matriz, não é toda nula, chamamos* **pivot** *ao elemento não nulo, dessa linha, mais à esquerda.*

Definição 1.10 *Dizemos que uma matriz está em* **forma de escada***, abreviadamente* **f.e.***, se verificar as 2 condições seguintes:*

1. *Se houver uma linha nula na matriz, as outras linhas abaixo dela serão nulas;*

2. *Em duas linhas quaisquer, não nulas, o pivot da linha inferior ocorre mais à direita do que o pivot da linha superior.*

Exemplo 1.11 *As seguintes matrizes estão em forma de escada*

$$\begin{bmatrix} 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad ; \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad ; \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

As seguintes matrizes não estão em forma de escada

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

Definição 1.12 *Dizemos que uma matriz está em* **forma de escada reduzida***, abreviadamente* **f.e.r.***, se estiver em forma de escada com cada pivot igual a 1 e os restantes elementos de cada coluna, a que pertença um pivot, iguais a zero.*

Exemplo 1.13 *As seguintes matrizes estão em forma de escada reduzida*

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad ; \quad \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Observação

1. Se aplicarmos o método de eliminação de Gauss a uma determinada matriz, obtemos uma sua forma de escada. Se além deste, utilizarmos o método de eliminação de Gauss-Jordan obtemos a sua forma de escada reduzida.

2. Uma matriz pode dar origem a diversas suas formas de escada, mas só dá origem a uma sua forma de escada reduzida.

Proposição 1.14 *Seja C uma matriz. Qualquer matriz em forma de escada obtida de C, tem o mesmo número de linhas não nulas.*

Definição 1.15 *Seja C uma matriz. Ao número de linhas não nulas de qualquer sua forma de escada chama-se* **característica de C** *e denota-se por* $\mathbf{r(C)}$.

Exemplo 1.16 *Consideremos o sistema*

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

Vejamos as características das matrizes simples e ampliada do sistema.

Ora, a matriz ampliada é

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 2 \\ 0 & 2 & 1 & 0 \\ 1 & 3 & 0 & 3 \end{array}\right] \underset{l_3 \to (l_3 - l_1)}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 2 \\ 0 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{array}\right] \underset{l_3 \to (l_3 - l_2)}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & -1 & 2 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Assim, $r([A|B]) = 3$ (número de linhas não nulas) e $r(A) = 2$.

Observação 1. Se C é uma matriz com m linhas, da definição vem que $r(C) \leq m$.

2. Se A é a matriz simples de um dado sistema e $[A|B]$ é a matriz ampliada, então $r(A) \leq r([A|B])$.

3. Se C é uma matriz com n colunas, porque numa forma de escada de C não podem existir dois pivots na mesma coluna, então $r(C) \leq n$ (o número de linhas não nulas é o número de pivots, quando em forma de escada).

1.8 Discussão de um Sistema

Muitas vezes, dada a complexidade de certos sistemas ou porque podem tirar-se conclusões sem a solução do sistema, interessa saber se o sistema é possível e determinado, possível e indeterminado ou impossível. Uma forma de fazer este estudo é através da comparação das características das matrizes ampliada e simples do sistema, e do número de incógnitas. A este estudo chama-se **discussão do sistema**.

Teorema 1.17 *Dado um sistema com m equações a n incógnitas e sendo A a matriz simples do sistema e $[A|B]$ a matriz ampliada, então:*

1. *Se $r(A) < r([A|B])$, o sistema é impossível*

2. *Se $r(A) = r([A|B]) = n$, o sistema é possível e determinado*

3. *Se $r(A) = r([A|B]) < n$, o sistema é possível e indeterminado, sendo $n - r(A)$ o número de variáveis livres do sistema a que se chama* **grau de indeterminação do sistema**.

Demonstração Seja $[A'|B']$ uma forma de escada da matriz ampliada $[A|B]$. Então, A' está em forma de escada e é obtida a partir de A. Pela definição, $r([A|B]) = r([A'|B'])$ e $r(A) = r(A')$.

1. Se $r(A) < r([A|B])$, então $r(A') < r([A'|B'])$. Consequentemente, o sistema a que corresponde a matriz ampliada $[A'|B']$ terá, pelo menos, uma equação do tipo $0 = b'_j$ com $b'_j \neq 0$ (repare-se no exemplo anterior) e o sistema é impossível.

2. Se $r(A) = r([A|B]) = n$, então $r(A') = r([A'|B']) = n$. Neste caso, o sistema a que corresponde a matriz ampliada $[A'|B']$, pensando que $[A'|B']$ é a forma de escada reduzida de $[A|B]$, é

$$\begin{cases} x_1 & & & = b'_1 \\ & x_2 & & = b'_2 \\ & & & \vdots \\ & & x_n & = b'_n \end{cases}$$

(repare-se no Exemplo 1.8, 1)). Ou seja, o sistema é possível e determinado.

3. Sendo $[A'|B']$ a forma de escada reduzida de $[A|B]$, se $r(A) = r([A|B]) = s < n$, então o sistema a que corresponde a matriz ampliada $[A'|B']$, terá s variáveis líderes e portanto $n - s$ variáveis livres e será do tipo

$$\begin{cases} x_{i_1} & & & +\sum(\ ) = b'_1 \\ & x_{i_2} & & +\sum(\ ) = b'_2 \\ & & & \vdots \\ & & x_{i_s} & +\sum(\ ) = b'_n \end{cases}$$

em que $\sum(\ )$ designa, em cada equação, a soma que envolve as variáveis livres. Como neste caso, $s < n$, então $n - s \geq 1$, ou seja, o sistema é possível e indeterminado (repare-se no Exemplo 1.8, 2)). $\square$

Porque, num sistema homogéneo a matriz dos termos independentes é toda nula e transformações elementares nas linhas não alteram uma coluna nula, temos sempre $r(A) = r([A|B])$.

Corolário 1.18 *Dado um sistema homogéneo com m equações a n incógnitas e sendo A a matriz simples do sistema, então:*

1. *Se $r(A) = n$, o sistema é possível e determinado*

2. *Se $r(A) < n$, o sistema é possível e indeterminado, sendo $n - r(A)$ o número de variáveis livres do sistema a que se chama* **grau de indeterminação do sistema**.

Exemplo 1.19 *Façamos a discussão do sistema nas incógnitas x, y, z e nos parâmetros a e b.*

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

A matriz ampliada deste sistema é

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & a & 1 & 2 \\ 0 & 0 & a-1 & b \end{array}\right].$$

Vamos utilizar o Teorema *anterior para sabermos quando é que o sistema está em cada um dos casos:*

- *Se $a \neq 0$ e $a \neq 1$ então, a matriz $[A|B]$ está em forma de escada, pelo que $r(A) = 3 = r([A|B]) =$número de incógnitas. Pelo* Teorema*, nestas condições o sistema é possível e determinado.*

- *Se $a = 0$, a matriz ampliada é*

  $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & b \end{array}\right]$ *que não está em forma de escada.*

  Calculemos a sua característica

  $$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & b \end{array}\right] \underset{l_3 \to (l_3 + l_2)}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & b+2 \end{array}\right].$$

  Se $b \neq -2$, $r([A|B]) = 3$.

  Se $b = -2$, $r([A|B]) = 2$.

  Como $r(A) = 2$, temos :

  Se $a = 0$, $b \neq -2$, $r(A) = 2 < 3 = r([A|B])$ e o sistema é impossível.

  Se $a = 0$, $b = -2$, $r(A) = 2 = r([A|B]) < 3 =$número de incógnitas e o sistema é possível e indeterminado com grau de indeterminação $n - r(A) = 3 - 2 = 1$.

- *Se $a = 1$, a matriz ampliada é*

  $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & b \end{array}\right].$

  Pelo que, se $b \neq 0$, $r([A|B]) = 3$ e se $b = 0$, $r([A|B]) = 3$.

  Porque $r(A) = 2$, então

  Se $a = 1$, $b \neq 0$, $r(A) = 2 < 3 = r([A|B])$ e o sistema é impossível,

  Se $a = 1$, $b = 0$, $r(A) = 2 = r([A|B]) < 3 =$número de incógnitas e o sistema é possível e indeterminado com grau de indeterminação $n - r(A) = 3 - 2 = 1$.

Resumindo, o sistema é

- *possível e determinado se* $(a \neq 0$ *e* $a \neq 1)$,
- *possível e indeterminado se* $(a = 0$ *e* $b = -2)$ *ou* $(a = 1$ *e* $b = 0)$, *em qualquer dos casos com grau de indeterminação* 1,
- *impossível se* $(a = 0$ *e* $b \neq -2)$ *ou* $(a = 1$ *e* $b \neq 0)$.

1.9 Aplicações dos Sistemas de Equações Lineares

Ao longo desta secção veremos algumas aplicações dos sistemas de equações lineares aos problemas da vida quotidiana.

1.9.1 Cores dos Ecrãs de Televisão

As cores dos ecrãs de televisão são baseadas no **modelo** de cores **RGB** (red, green and blue). Neste modelo, as cores são criadas a partir das três cores: vermelho, verde e azul. Se identificarmos estas cores com os vectores de $\mathbb{R}^3$

$R = (1, 0, 0)$ (vermelho), $G = (0, 1, 0)$ (verde), $B = (0, 0, 1)$ (azul)

e adicionarmos estes vectores depois de cada um deles ter sido multiplicado por um escalar (um número real) entre 0 e 1, inclusivé, (estes escalares representam a percentagem de cada cor na mistura) obtemos todas as outras cores.

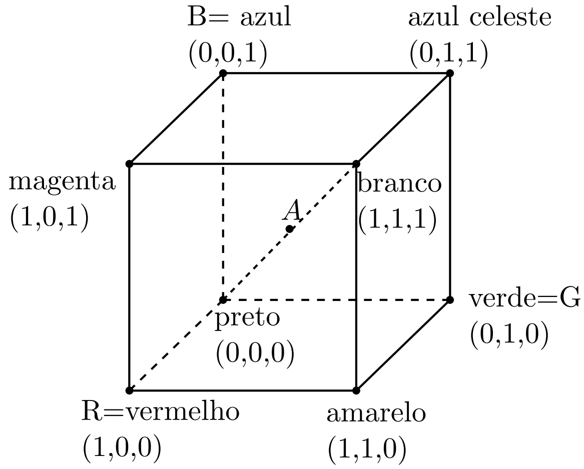
Definição 1.20 *Um vector* w *de* $\mathbb{R}^n$ *é* **combinação linear** *dos vectores* $v_1, v_2, ..., v_p$ *de* $\mathbb{R}^n$ *se* w *pode ser expresso na forma*

$$w = c_1 v_1 + c_2 v_2 + ... + c_p v_p$$

em que $c_1, c_2, ..., c_p \in \mathbb{R}$ *e são denominados* **coeficientes da combinação linear**.

Portanto, o que dissemos anteriormente, é que cada vector cor pode ser expresso como combinação linear de R,G,B, em que os coeficientes da combinação linear são números reais entre 0 e 1.

O conjunto de todas as cores pode ser representado pelo cubo:



Ao longo da diagonal entre o preto e o branco estão todas as tonalidades de cinzento. Se A (cor cinzenta) for o vector cor $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$, ele escreve-se como combinação linear dos vectores cores R,G,B da seguinte forma:

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}(1,0,0) + \frac{1}{2}(0,1,0) + \frac{1}{2}(0,0,1).$$

Vejamos se é possível escrever A como combinação linear dos vectores cores R,B e amarelo. Neste caso, o que pretendemos é encontrar coeficientes x, y, z tais que

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = x(1,0,0) + y(0,0,1) + z(1,1,0).$$

Atendendo ao produto de um escalar por um vector, soma de vectores e igualdade de vectores temos

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = (x+z, z, y).$$

Donde obtemos o sistema

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

que tem a solução $x = 0$, $y = \frac{1}{2}$, $z = \frac{1}{2}$. Assim a cor cinzenta A obtém-se à custa do amarelo e do azul sem necessitarmos do vermelho,

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}(0,0,1) + \frac{1}{2}(1,1,0).$$

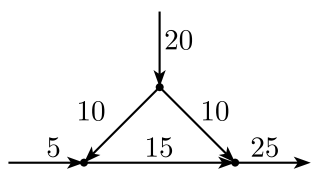
1.9.2 Análise de Redes

Rede é um conjunto de arcos ao longo dos quais "flui" alguma coisa. Por exemplo, arcos podem ser canos onde flui água, ruas de uma cidade onde flui o trânsito, fios eléctricos onde flui corrente eléctrica,... Estes arcos, na maioria das redes, encontram-se em pontos denominados vértices. Por exemplo, numa rede de canos de água, os vértices ocorrem quando se juntam três ou mais canos, na rede de trânsito, ocorrem quando há cruzamentos e na rede eléctrica quando se juntam três ou mais fios.

A maioria das redes tem três propriedades básicas:

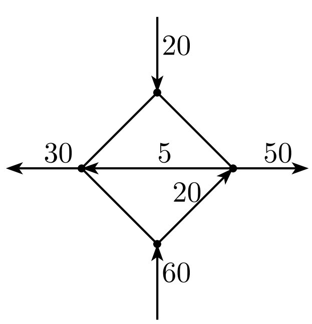
1. O fluxo num arco não pode mudar de sentido.
2. A taxa de fluxo que termina num vértice é igual à que sai do vértice.
3. A taxa de fluxo que sai da rede é igual à que entra na rede, sendo a taxa de fluxo de um arco uma medida numérica.

Exemplo 1.21 *Consideremos a seguinte rede de canos de água*

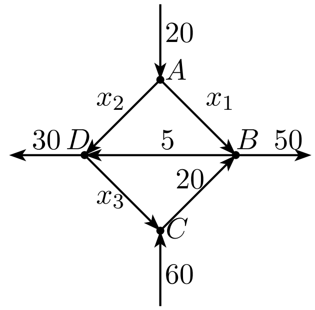


onde a medida dos arcos é em litros por minuto. Esta rede tem 3 *vértices. A taxa de fluxo que entra na rede é* $20 + 5 = 25$ *e a que sai da rede é* 25*.*

Exemplo 1.22 *A seguinte figura representa uma rede de trânsito com indicação de algumas taxas de fluxo nos arcos. Vamos encontrar as outras taxas de fluxo nos arcos e o sentido desse fluxo nos arcos.*



Vamos escolher sentidos arbitrários para os arcos que não o têm. Se não estiver bem escolhido o sentido, o seu valor virá negativo.



Como, nos vértices a taxa de fluxo que termina em cada um é igual à que começa, temos

$$\begin{aligned} 20 &= x_1 + x_2 && (\textit{vértice } A) \\ x_1 + 20 &= 50 + 5 && (\textit{vértice } B) \\ x_3 + 60 &= 20 && (\textit{vértice } C) \\ 5 + x_2 &= 30 + x_3 && (\textit{vértice } D) \end{aligned}$$

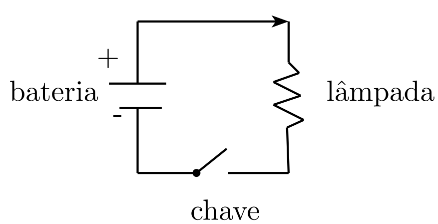
Estas condições produzem o sistema

$$\begin{cases} x_1 + x_2 = 20 \\ x_1 = 35 \\ x_3 = -40 \\ x_2 - x_3 = 25. \end{cases}$$

A solução do sistema é $x_1 = 35$, $x_2 = -15$, $x_3 = -40$, *ou seja, os sentidos de* x_2 *e* x_3 *estão incorrectos.*

1.9.3 Circuitos Eléctricos

Um circuito eléctrico é normalmente composto por uma fonte de energia eléctrica (bateria) e um elemento que dissipa energia eléctrica (lâmpada).



Iremos considerar que num circuito eléctrico, quando a chave está fechada, a corrente eléctrica flui do pólo positivo da bateria para a lâmpada e vai até ao pólo negativo.

A bateria cria uma tensão eléctrica que é medida em volts(V). A resistência é o valor que a lâmpada reduz à tensão eléctrica e é medida em ohms(Ω). A taxa de fluxo das cargas eléctricas num fio eléctrico (corrente eléctrica) é medida em amperes(A) e é designada por intensidade da corrente. Estes valores estão relacionados pela Lei de Ohm.

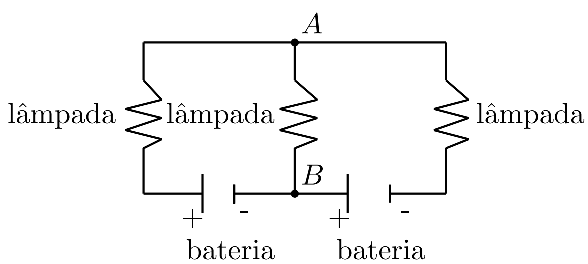
Lei de Ohm

Se uma corrente de I amperes passa por uma lâmpada com uma resistência de R ohms, então resulta uma diferença de potencial (quebra de voltagem) de E volts que é o produto da corrente pela resistência, ou seja,

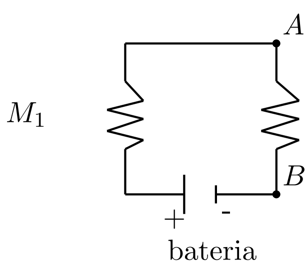
E=IR.

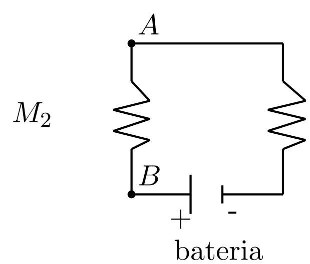
Numa rede eléctrica chamamos malha a uma sucessão de arcos que começa e termina no mesmo vértice.

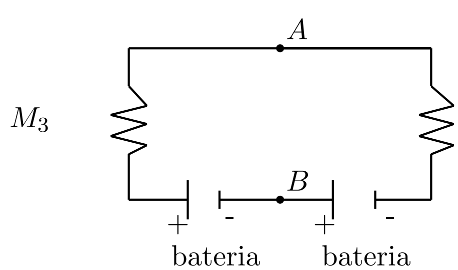
Na rede



temos 2 vértices A, B e as malhas







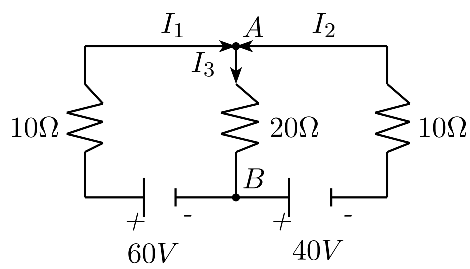
Lei das Tensões de Kirchhoff

Em qualquer malha, a soma dos aumentos de voltagem é igual à soma das quebras de voltagem.

Lei das Correntes de Kirchhoff

A soma das correntes que terminam num vértice é igual à soma das correntes que começam no vértice.

Exemplo 1.23 *Determine as correntes I_1, I_2 e I_3 do circuito*



Atendendo às leis de Kirchhoff

$$
\begin{array}{rrcl}
\textit{vértice } A & I_1 + I_2 & = & I_3 \\
\textit{malha à esquerda } (M_1) & 60 & = & 10I_1 + 20I_3 \\
 & & (\textit{lei de Ohm}) & \\
 & (\textit{aumento de voltagem}) & & (\textit{quebra de voltagem}) \\
\textit{malha à direita } (M_2) & 40 + 10I_2 + 20I_3 & = & 0 \\
\textit{malha externa } (M_3) & 60 + 40 + 10I_2 & = & 10I_1
\end{array}
$$

Donde resulta o sistema

$$
\left\{
\begin{array}{llll}
I_1 & +I_2 & -I_3 & = \ 20 \\
10I_1 & & +20I_3 & = \ 60 \\
 & -10I_2 & -20I_3 & = \ 40 \\
10I_1 & -10I_2 & & = 100.
\end{array}
\right.
$$

A solução deste sistema é $I_1 = \frac{26}{5}A$, $I_2 = -\frac{24}{5}A$, $I_3 = \frac{2}{5}A$. *Como* I_2 *é negativo, o sentido da corrente é o oposto ao indicado.*

Capítulo 2

OPERAÇÕES COM MATRIZES

No capítulo anterior utilizámos as matrizes para simplificar o estudo e a resolução dos sistemas de equações lineares. Neste capítulo consideraremos as matrizes como entes matemáticos e definiremos algumas operações matriciais e as suas principais propriedades.

2.1 Definições Básicas e Exemplos de Matrizes

A notação que temos usado para um vector de $\mathbb{R}^n$ é

$$v = (v_1, \ldots, v_n).$$

Atendendo a que um vector de $\mathbb{R}^n$ é uma lista de n números (as componentes do vector) ordenados, qualquer notação que tenha as componentes do vector pela mesma ordem é uma outra forma de escrever o vector. Por exemplo, podemos escrever o vector v como a matriz

$$\begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}$$

(matriz linha) ou

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

(matriz coluna).

Definição 2.1 **Matriz** *é um quadro de números, denominados* **entradas da matriz***, dispostos em m linhas e n colunas, e por isso diz-se que a* **matriz é de tamanho** $m \times n$ *(ou de tipo $m \times n$).*

Vejamos alguns exemplos de tipos de matrizes:

- Como já o dissemos, uma **matriz linha** é uma matriz com uma única linha, portanto de tamanho $1 \times n$ (ou, de tipo $1 \times n$).
- Uma **matriz coluna** é uma matriz de tipo $m \times 1$.
- Uma matriz diz-se **quadrada** se o número das suas linhas for igual ao número das suas colunas. Neste caso, se for de tipo $n \times n$, dizemos que é uma matriz quadrada de **ordem** n. Por exemplo, $\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$ é uma matriz quadrada de ordem 2, ou simplesmente matriz de ordem 2.

Notação 2.2 *A notação usual para designar uma matriz* A *de tipo* $m \times n$ (*repare-se que em primeiro lugar aparece o número de linhas da matriz*) *é*

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

em que a_{ij} *é o elemento que ocupa* **a posição** (i,j) (*ou* **a entrada** (i,j)) *da matriz* A*, ou seja,* a_{ij} *encontra-se na linha* i *e coluna* j *de* A*. Também se usa a notação* $(A)_{ij}$ *para designar o elemento que ocupa a posição* (i,j) *da matriz* A*.*

Quando queremos uma notação mais compacta para a matriz A*, escrevemos* $A = [a_{ij}]$*.*

Definição 2.3 *Seja* $A = [a_{ij}]$ *uma matriz de ordem* n*.*

Os elementos $a_{11}, a_{22}, \dots, a_{nn}$ *formam a* **diagonal principal** *da matriz* A*.*

Dizemos que A *é uma matriz* **triangular superior** *se as entradas abaixo da diagonal principal são nulas, isto é, se* A *é da forma*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Dizemos que A *é uma matriz* **triangular inferior** *se as entradas acima da diagonal principal são nulas, isto é, se* A *é da forma*

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Dizemos que A é uma **matriz diagonal** *se for triangular inferior e triangular superior, isto é, se A é da forma*

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Dizemos que A é uma **matriz escalar** *se for diagonal com $a_{11} = a_{22} = \ldots = a_{nn}$ (as entradas da diagonal principal são iguais).*

Entre as matrizes escalares há a destacar a

- **matriz nula**, *matriz na qual qualquer entrada da matriz é nula. A notação usual da matriz nula de ordem n é 0_n;*
- **matriz identidade**, *matriz que tem $a_{11} = a_{22} = \ldots = a_{nn} = 1$. A notação usual da matriz identidade de ordem n é I_n.*

Exemplo 2.4 *A matriz $A = \begin{bmatrix} 2 & 3 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ é triangular superior.*

A matriz $B = \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ é triangular inferior.

A matriz $C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ é diagonal.

Em qualquer uma delas, os elementos $2, -1, 0$ formam a sua diagonal principal.

As matrizes $E = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$; $0_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$; $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ são todas escalares sendo a segunda a matriz nula de ordem 2 e a terceira a matriz identidade de ordem 2.

Definição 2.5 *Duas matrizes A e B dizem-se iguais se tiverem o mesmo tamanho e as entradas correspondentes iguais. Neste caso, escrevemos $A = B$.*

2.2 Operações com Matrizes

Como veremos ao longo desta secção, muitas das operações que se efectuam com matrizes não são mais do que generalizações das respectivas operações em $\mathbb{R}$. No entanto, há certas operações que se efectuam nas matrizes e que não têm significado em $\mathbb{R}$.

2.2.1 Soma de Matrizes e Produto de um Escalar por uma Matriz

Como sabemos somar 2 vectores e multiplicar um escalar por um vector, e já vimos na secção anterior como representar um vector na forma matricial, estamos em condições de definir as operações de soma de matrizes e produto de um escalar por uma matriz.

Definição 2.6 *Sejam $A = [a_{ij}]$ uma matriz de tipo $m \times n$ e α um escalar. O produto αA é a matriz de tipo $m \times n$ que se obtém multiplicando cada entrada de A por α, isto é,*

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

Definição 2.7 *Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes de tipo $m \times n$. A soma da matriz A com a matriz B, que se denota por $A + B$, é a matriz de tipo $m \times n$ cuja entrada (i, j) é o elemento $a_{ij} + b_{ij}$, isto é,*

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij}.$$

Exemplo 2.8 *Consideremos as matrizes*

$$A = \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -1 & 1 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 5 & 2 \\ 2 & 2 \\ 0 & -1 \end{bmatrix}.$$

As duas matrizes são de tipo 3×2, pelo que podemos somá-las

$$A + B = \begin{bmatrix} 2+5 & -1+2 \\ 3+2 & 0+2 \\ -1+0 & 1-1 \end{bmatrix} = \begin{bmatrix} 7 & 1 \\ 5 & 2 \\ -1 & 0 \end{bmatrix}.$$

Já a matriz

$$3A = \begin{bmatrix} 3 \times 2 & 3 \times (-1) \\ 3 \times 3 & 3 \times 0 \\ 3 \times (-1) & 3 \times 1 \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ 9 & 0 \\ -3 & 3 \end{bmatrix}.$$

Relativamente às operações atrás definidas temos as seguintes propriedades:

Proposição 2.9 *Sejam A, B, C matrizes de tipo $m \times n$ e α, β escalares. Tem-se:*

1. $A + B = B + A$ *(comutatividade da adição)*
2. $A + (B + C) = (A + B) + C$ *(associatividade da adição)*
3. $(\alpha\beta)A = \alpha(\beta A) = \beta(\alpha A)$

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

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

6. $(-\alpha)A = -(\alpha A) = \alpha(-A)$ *em que* $-A$ *significa* $(-1)A$.

Demonstração Vamos demonstrar a propriedade 4..

Como $A = [a_{ij}]$, $B = [b_{ij}]$ são matrizes de tipo $m \times n$, então $\alpha(A+B)$ e $\alpha A + \alpha B$ são de tipo $m \times n$. Atendendo às definições,

$$(\alpha(A+B))_{ij} = \alpha(A+B)_{ij} = \alpha(a_{ij} + b_{ij}).$$

Usando a propriedade distributiva do produto em relação à adição de números, vem que

$$(\alpha(A+B))_{ij} = \alpha a_{ij} + \alpha b_{ij}.$$

Usando as definições, temos

$$(\alpha(A+B))_{ij} = (\alpha A)_{ij} + (\alpha B)_{ij} = (\alpha A + \alpha B)_{ij}.$$

□

2.2.2 Produto de uma Matriz por uma Matriz Coluna

Vejamos em primeiro lugar como se define o produto de uma matriz por uma matriz coluna.

Em certas ocasiões é conveniente pensarmos numa matriz como uma sequência de matrizes linha ou matrizes coluna, que podem ser traduzidas em vectores.

Por exemplo, se

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

então A pode ser subdividida em matrizes coluna

$$A = \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} C_1 & C_2 & C_3 & C_4 \end{bmatrix}$$

em que $C_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \end{bmatrix}$ é a j-ésima coluna de A, com $j = 1, 2, 3, 4$.

Ou subdividir A em matrizes linha

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ - & - & - & - \\ a_{21} & a_{22} & a_{23} & a_{24} \\ - & - & - & - \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix}$$

em que $R_i = \left[\, a_{i1} \;\; a_{i2} \;\; a_{i3} \;\; a_{i4} \,\right]$ é a i-ésima linha de A, com $i = 1, 2, 3$.

Já vimos que um sistema de m equações lineares a n incógnitas, (1.1),

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n & = b_1 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n & = b_m \end{cases}$$

tem 3 matrizes associadas:

-a matriz simples do sistema, $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$,

-a matriz dos termos independentes, $B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$,

-a matriz das incógnitas, $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

Vamos definir o produto AX por forma a que o sistema (1.1) possa ser escrito pela expressão matricial $AX = B$.

Ora (1.1) na forma matricial é

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

(recorde-se que duas matrizes são iguais se forem do mesmo tipo e as correspondentes entradas forem iguais).

Atendendo à soma de matrizes e ao produto de um escalar por uma matriz, a matriz anterior pode ser escrita como

$$x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Concluímos que

$$AX = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Definição 2.10 *Se A é uma matriz $m\times n$ e X é uma matriz coluna $n\times 1$, então o produto AX é uma matriz coluna $m\times 1$ que resulta da combinação linear das matrizes coluna de A, $C_1, C_2, \ldots, C_n$, tendo como coeficientes (da combinação linear) as entradas de X, isto é,*

$$AX = \begin{bmatrix} C_1 & C_2 & \ldots & C_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1C_1 + x_2C_2 + \ldots + x_nC_n.$$

Exemplo 2.11

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

$$\begin{bmatrix} -2 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} -12 \\ 4 \\ -4 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -8 \\ 7 \\ -5 \end{bmatrix}.$$

Vejamos agora duas propriedades que envolvem os três tipos de operações definidos:

Proposição 2.12 *Sejam A uma matriz $m\times n$, B e D matrizes coluna $n\times 1$ e α um escalar. Tem-se:*

1. $A(\alpha B) = \alpha(AB) = (\alpha A)B$
2. $A(B + D) = AB + AD.$

Demonstração Sejam $C_1, \ldots, C_n$ as matrizes coluna de A e $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$, $D = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$.

Então,

$$\alpha B = \begin{bmatrix} \alpha b_1 \\ \vdots \\ \alpha b_n \end{bmatrix} \qquad \text{e} \qquad B + D = \begin{bmatrix} b_1 + d_1 \\ \vdots \\ b_n + d_n \end{bmatrix}.$$

Assim,

$$A(\alpha B) = \begin{bmatrix} C_1 & C_2 & \ldots & C_n \end{bmatrix} \begin{bmatrix} \alpha b_1 \\ \vdots \\ \alpha b_n \end{bmatrix} = (\alpha b_1)C_1 + (\alpha b_2)C_2 + \ldots + (\alpha b_n)C_n =$$

porque α e b_i são escalares, $i = 1, \ldots, n$

$$= \alpha(b_1C_1) + \alpha(b_2C_2) + \ldots + \alpha(b_nC_n) = \alpha(b_1C_1 + b_2C_2 + \ldots + b_nC_n) = \alpha(AB).$$

Da mesma forma,

$$A(B+D) = (b_1 + d_1)C_1 + \ldots + (b_n + d_n)C_n =$$

porque b_i, d_i são escalares, $i = 1, \ldots, n$,

$$= b_1C_1 + d_1C_1 + \ldots + b_nC_n + d_nC_n =$$

usando a comutativa da soma de matrizes

$$= (b_1C_1 + \ldots + b_nC_n) + (d_1C_1 + \ldots + d_nC_n) = AB + AD.$$

□

2.2.3 Produto de Duas Matrizes

Vejamos agora como se efectua o produto de duas matrizes A e B quando B tem mais do que uma coluna. Para que o produto AB esteja definido, exigimos que a condição da associatividade

$$(AB)X = A(BX)$$

seja verdadeira para qualquer matriz coluna X, de tipo $n \times 1$.

Pelo que vimos anteriormente, se X tem n linhas, para o produto BX estar definido, B terá n colunas. Portanto, B é matriz de tipo $s \times n$. O que implica que BX é matriz coluna de tipo $s \times 1$. Mas então, para o produto $A(BX)$ estar definido, a matriz A deve ter s colunas. Consequentemente, mostrámos que para se conseguir efectuar o produto, o **número de colunas** de A tem de ser igual ao **número de linhas** de B.

Se as colunas de B forem $B_1, \ldots, B_n$, então

$$BX = \begin{bmatrix} B_1 & B_2 & \ldots & B_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1B_1 + x_2B_2 + \ldots + x_nB_n$$

e usando a Proposição anterior,

$$A(BX) = A(x_1B_1 + x_2B_2 + \ldots + x_nB_n) = A(x_1B_1) + A(x_2B_2) + \ldots + A(x_nB_n)$$

$$= x_1(AB_1) + x_2(AB_2) + \ldots + x_n(AB_n).$$

Como queremos que $(AB)X = A(BX)$, vem que

$$(AB)X = x_1(AB_1) + \ldots + x_n(AB_n)$$

e podemos considerar que as colunas de AB são as matrizes coluna

$$AB_1, \ldots, AB_n.$$

Definição 2.13 *Sejam A uma matriz $m \times s$ e B uma matriz $s \times n$ cujas matrizes coluna são $B_1, \ldots, B_n$, o produto AB é a matriz $m \times n$ tal que $AB = [AB_1| \ldots |AB_n]$.*

Observação 1. Repare-se que o número de colunas de A é igual ao número de linhas de B para podermos efectuar o produto AB.

2. A matriz AB tem tantas linhas quantas as da matriz A e tantas colunas quantas as da matriz B.

Exemplo 2.14 *Sendo* $A = \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$ *e* $B = \begin{bmatrix} 1 & 2 \\ -1 & 2 \\ 2 & 0 \\ 2 & 0 \end{bmatrix}$*, o número de colunas de A é 4 e é o mesmo que o número de linhas de B. Portanto podemos efectuar o produto AB. Pela definição teremos de calcular o produto de A por cada coluna de B. Ora*

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + (-1) \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

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

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

Assim,

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

Observação Muitas vezes o produto AB e BA estão definidos, mas pode acontecer que $AB \neq BA$:

1. Se A for de tipo $n \times s$ e B de tipo $s \times n$, então AB é matriz de tipo $n \times n$ e BA é de tipo $s \times s$. Se $s \neq n$ então $AB \neq BA$ (para duas matrizes serem iguais têm de ser de mesmo tipo e ...)

2. Se A e B forem matrizes de ordem n, por exemplo $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, matrizes de ordem 2, temos $AB = [AB_1 \ AB_2]$ sendo B_1, B_2 as matrizes coluna de B. Ora

$$AB_1 = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$AB_2 = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = (-1) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

Pelo que $AB = \begin{bmatrix} 1 & -2 \\ 0 & 2 \end{bmatrix}$.

Calculando BA temos que $BA = \begin{bmatrix} 1 & -3 \\ 0 & 2 \end{bmatrix}$. Donde, $AB \neq BA$.

Vejamos agora como encontrar uma entrada específica do produto de 2 matrizes, sem termos de calcular toda a matriz produto. Sejam $A = [a_{ij}]$ uma matriz de tipo $m \times s$ e $B = [b_{ij}]$ uma matriz de tipo $s \times n$. Sabemos que é possível efectuar o produto AB, que será uma matriz de tipo $m \times n$. E se quisermos saber qual a entrada (i, j) de AB, isto é, qual o elemento, $(AB)_{ij}$, que ocupa a i-ésima linha e a j-ésima coluna de AB?

Sendo $B_1, \ldots, B_n$ as matrizes coluna de B, a j-ésima matriz coluna de AB é

$$\begin{aligned} AB_j &= \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1s} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{ms} \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{sj} \end{bmatrix} = \\ &= b_{1j} \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \ldots + b_{sj} \begin{bmatrix} a_{1s} \\ \vdots \\ a_{ms} \end{bmatrix}. \end{aligned}$$

Ora a entrada (i, j) de AB é a $(AB_j)_{i1}$, pelo que

$$\begin{aligned} (AB)_{ij} &= b_{1j}a_{i1} + \ldots + b_{sj}a_{is} = \\ &= a_{i1}b_{1j} + \ldots + a_{is}b_{sj} = [a_{i1} \ldots a_{is}] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{sj} \end{bmatrix}. \end{aligned}$$

Isto é, $(AB)_{ij}$ é o produto da i-ésima matriz linha de A pela j-ésima matriz coluna de B.

Exemplo 2.15 *Sendo* $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \end{bmatrix}$ *e* $B = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ -2 & 2 \end{bmatrix}$, *a entrada* $(2, 1)$ *de* AB *é*

$$(AB)_{21} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} = 1 \times 0 + (-1) \times 1 + (-2) \times (-1) = 1.$$

Temos as seguintes propriedades:

Proposição 2.16 *Sejam A, B, C três matrizes de tipo adequado por forma a que as operações indicadas se possam efectuar, e α um escalar. Tem-se:*

1. $A(BC) = (AB)C$ *(associatividade da multiplicação)*
2. $A(B + C) = AB + AC$ *(distributividade)*
3. $(B + C)A = BA + CA$ *(distributividade)*
4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$

Demonstração Demonstremos 1. O que iremos mostrar é que a j-ésima matriz coluna de $A(BC)$ é igual à j-ésima matriz coluna de $(AB)C$. Seja C_j a j-ésima matriz coluna de C. Então a j-ésima matriz coluna de $(AB)C$ é $(AB)C_j$.

Como BC_j é a j-ésima coluna de BC, então $A(BC_j)$ é a j-ésima coluna de $A(BC)$.

Mas a multiplicação de matrizes foi criada para tornar $(AB)X = A(BX)$ verdadeiro para qualquer matriz coluna X apropriada, então

$$(AB)C_j = A(BC_j).$$

□

2.2.4 Transposta de uma Matriz

Vejamos uma operação com matrizes que não tem paralelo com a álgebra dos números reais.

Definição 2.17 *Seja A uma matriz $m \times n$, então a* **matriz transposta de** *A, denotada por A^T, é a matriz $n \times m$ que se obtém de A transformando as linhas de A em colunas de A^T, isto é,*

$$(A^T)_{ij} = (A)_{ji}.$$

Exemplo 2.18 *A transposta da matriz $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 4 \end{bmatrix}$ é a matriz*

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

Relativamente à operação de transposição, temos:

Proposição 2.19 *Sejam A, B matrizes de tipo adequado por forma a que as operações indicadas se possam efectuar, e α um escalar. Tem-se:*

1. $(A^T)^T = A$

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

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

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

Demonstração Demonstremos a propriedade 4. Sendo $A = [a_{ij}]$ uma matriz de tipo $m \times s$ e $B = [b_{ij}]$ uma matriz de tipo $s \times n$, vem que

$$\begin{aligned} ((AB)^T)_{ij} &= (AB)_{ji} \text{ (definição de matriz transposta)} \\ &= [a_{j1} \dots a_{js}] \begin{bmatrix} b_{1i} \\ \vdots \\ b_{si} \end{bmatrix}. \end{aligned}$$

Atendendo a que a linha i de B^T é a coluna i de B

$$(B^T A^T)_{ij} = [b_{1i} \dots b_{si}] \begin{bmatrix} a_{j1} \\ \vdots \\ a_{js} \end{bmatrix}$$

obtemos que $(AB)^T = B^T A^T$, uma vez que AB é matriz $m \times n$, $(AB)^T$ é matriz $n \times m$ e $B^T A^T$ é matriz $n \times m$. □

Definição 2.20 *Uma matriz A diz-se* **simétrica** *se $A^T = A$, e* **hemi-simétrica** (*ou* **anti-simétrica**) *se $A^T = -A$.*

Exemplo 2.21 *A matriz $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$ é simétrica e a matriz $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ é hemi-simétrica. Já a matriz $C = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ não é simétrica nem hemi-simétrica.*

Proposição 2.22 *Sejam A, B matrizes simétricas de tipo adequado por forma a que as operações indicadas se possam efectuar, e α um escalar. Tem-se:*

1. *A^T é simétrica*

2. *$A + B$ é simétrica*

3. *αA é simétrica*

4. *AB é simétrica se, e só se, $AB = BA$.*

2.2.5 Inversa de uma Matriz

Nos números reais, cada número não nulo a tem um inverso para a multiplicação, isto é, existe a^{-1}(real) tal que $aa^{-1} = a^{-1}a = 1$. Vejamos uma definição análoga para matrizes.

Definição 2.23 *Seja A uma matriz quadrada de ordem n. Dizemos que A é* **invertível** *se existir uma matriz B tal que*

$$AB = BA = I_n.$$

Observação

1. Para que o produto AB e BA estejam definidos, B tem que ser uma matriz quadrada de ordem n.

2. Se $AB = BA = I_n$, sendo A uma matriz quadrada de ordem n, então, também temos $BA = AB = I_n$. Portanto, B é invertível.

Vejamos que a definição introduzida está bem formulada.

Proposição 2.24 *Seja A uma matriz quadrada de ordem n, invertível. Então, existe uma única matriz B tal que*

$$AB = BA = I_n.$$

Demonstração A existência de uma matriz B nestas condições está garantida pela definição de matriz invertível. Suponhamos que B, C são duas matrizes tais que

$$AB = BA = I_n \quad , \qquad AC = CA = I_n.$$

Então, porque $I_nD = D = DI_n$, qualquer que seja D de ordem n,

$$C = I_nC \underset{\substack{\uparrow \\ BA=I_n}}{=} (BA)C \underset{\substack{\uparrow \\ \text{associativa}}}{=} B(AC) \underset{\substack{\uparrow \\ AC=I_n}}{=} BI_n = B.$$

Ou seja, $B = C$ e só existe uma matriz que verifica a condição. $\square$

Definição 2.25 *Seja A uma matriz invertível de ordem n. À única matriz B tal que $AB = BA = I_n$ chamamos* **inversa** *de A e denotamo-la por A^{-1}.*

Nem todas as matrizes são invertíveis:

Definição 2.26 *Seja A uma matriz quadrada de ordem n. Se A não é invertível dizemos que A é* **singular**.

Exemplo 2.27 1. *Sendo* $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}$ *então* $AB = BA = I_2$, *pelo que* B *é a inversa de* A *e* A *é a inversa de* B.

2. *Se* $A = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$, *qualquer que seja* $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ *vem que*

$$A \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = b_{11} \begin{bmatrix} 2 \\ 0 \end{bmatrix} + b_{21} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2b_{11} + b_{21} \\ 0 \end{bmatrix}.$$

$$A \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} = b_{12} \begin{bmatrix} 2 \\ 0 \end{bmatrix} + b_{22} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2b_{12} + b_{22} \\ 0 \end{bmatrix}.$$

Donde

$$AB = \begin{bmatrix} 2b_{11} + b_{21} & 2b_{12} + b_{22} \\ 0 & 0 \end{bmatrix} \neq I_2.$$

Ou seja, A *é singular.*

Proposição 2.28 *Sejam* A *e* B *matrizes invertíveis de ordem* n, *então* AB *é invertível e*

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

Demonstração Porque

$$(AB)(B^{-1}A^{-1}) \underset{\substack{\uparrow \\ \text{associativa}}}{=} A(BB^{-1})A^{-1} \underset{\substack{\uparrow \\ B^{-1} \text{ inversa de } B}}{=} AI_nA^{-1} = AA^{-1} \underset{\substack{\uparrow \\ A^{-1} \text{ inversa de } A}}{=} I_n$$

e da mesma forma

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

podemos afirmar, atendendo à unicidade da inversa, que AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$. □

2.2.6 Potências de uma Matriz

Sendo A uma matriz quadrada de ordem n, definimos as potências inteiras, não negativas, de A por

$$\begin{aligned} A^0 &= I_n \\ A^k &= A^{k-1}A, \quad \text{para } k \in \mathbb{N}. \end{aligned}$$

Se A é invertível, definimos as potências inteiras negativas de A por

$$A^{-k} = (A^{-1})^k, \quad \text{para } k \in \mathbb{N}.$$

As leis usuais das potências são válidas para matrizes, isto é,

$$\begin{aligned} (A^r)^s &= (A^{rs}) \\ A^r.A^s &= A^{r+s}. \end{aligned}$$

Proposição 2.29 *Sejam A uma matriz invertível de ordem n, α um escalar não nulo e $k \in \mathbb{N}_0$. Então:*

1. A^{-1} *é invertível e* $(A^{-1})^{-1} = A$.
2. A^k *é invertível e* $(A^k)^{-1} = (A^{-1})^k = A^{-k}$.
3. αA *é invertível e* $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$.

Exemplo 2.30 *Sabendo que a inversa de* $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$ *é* $A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}$, *vamos determinar a inversa de* A^2. *Como* $(A^2)^{-1} = (A^{-1})^2$ *vem que* $(A^2)^{-1} = \begin{bmatrix} \frac{1}{4} & -\frac{3}{4} \\ 0 & 1 \end{bmatrix}$.

Não necessitámos de calcular A^2 e depois a sua inversa, porque usámos a Proposição 2.29.

2.3 Matrizes Elementares

No capítulo anterior falámos nas transformações elementares que se podem efectuar nas linhas de uma matriz:
1. Multiplicar uma linha por uma constante não nula;
2. trocar duas linhas;
3. somar um múltiplo de uma linha a outra linha.

Definição 2.31 *Chamamos* **matriz elementar** *a qualquer matriz que se obtém da matriz identidade efectuando uma única transformação elementar.*

Exemplo 2.32 1. *A matriz* $\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$ *é matriz elementar porque foi obtida da matriz I_2 através da transformação elementar*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{l_2 \to -2l_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}.$$

2. *A matriz* $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, *que foi obtida da matriz I_3 efectuando a transformação elementar*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{l_1 \leftrightarrow l_2}{\longrightarrow} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

é uma matriz elementar.

3. *A matriz* $\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, *é elementar porque foi obtida da matriz* I_3 *efectuando a transformação elementar*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underset{l_1 \to (l_1 - 2l_2)}{\longrightarrow} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As matrizes elementares são importantes porque podem ser usadas para efectuar transformações elementares nas linhas de uma matriz, se multiplicadas à esquerda por essa matriz.

Proposição 2.33 *Sejam* A *uma matriz* $m \times n$ *e* E *uma matriz elementar de ordem* m. *Se efectuarmos em* A *a mesma transformação elementar que transforma* I_m *em* E, *obtemos* EA.

Exemplo 2.34 *Considere a matriz*

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

Se efectuar a transformação elementar nas linhas de A

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 1 & 1 \end{bmatrix} \underset{l_2 \to (l_2 - 2l_1)}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -5 & -7 \end{bmatrix}.$$

Se fizermos a mesma transformação elementar nas linhas de I_2, *temos*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{l_2 \to (l_2 - 2l_1)}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = E.$$

Vejamos se EA *dá o mesmo resultado*

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

Como se vê, a multiplicação por E, *soma* -2 *vezes a primeira linha à segunda linha de* A.

Proposição 2.35 *Seja* E *uma matriz elementar de ordem* n. *Então* E *é invertível, a sua inversa é matriz elementar e*

1. *Se* $I_n \underset{l_i \to \alpha l_i}{\longrightarrow} E$, *com* $\alpha \neq 0$, *então* $I_n \underset{l_i \to \frac{1}{\alpha} l_i}{\longrightarrow} E^{-1}$.

2. *Se* $I_n \underset{l_i \leftrightarrow l_j}{\longrightarrow} E$, *então* $I_n \underset{l_j \leftrightarrow l_i}{\longrightarrow} E^{-1}$.

3. *Se* $I_n \underset{l_i \rightarrow (l_i + \alpha l_j)}{\longrightarrow} E$, *então* $I_n \underset{l_i \rightarrow (l_i - \alpha l_j)}{\longrightarrow} E^{-1}$.

Exemplo 2.36 *Sendo* $E = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *a matriz elementar que se obteve da matriz* I_3 *efectuando a transformação elementar* $l_1 \rightarrow (l_1 - 2l_2)$, *então pela* Proposição 2.35, E^{-1} *é a matriz elementar que se obtém de* I_3 *efectuando a transformação elementar* $l_1 \rightarrow (l_1 + 2l_2)$, *ou seja,*

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

Pode confirmar-se este resultado efectuando os produtos EE^{-1} *e* $E^{-1}E$, *e vendo que são iguais a* I_3.

2.4 Caracterização das Matrizes Invertíveis

A partir da definição é muitas vezes complicado saber se uma matriz é ou não invertível. Vejamos processos que nos permitam conhecer a invertibilidade de uma matriz.

Teorema 2.37 *Seja* A *uma matriz invertível de ordem* n. *As afirmações seguintes são equivalentes:*

1. A *é invertível*

2. $r(A) = n$

3. *A forma de escada reduzida de* A *é* I_n

4. A *pode escrever-se como produto de matrizes elementares.*

Demonstração Podemos demonstrar a equivalência entre as 4 afirmações estabelecendo a cadeia de implicações

$$1. \Rightarrow 2. \Rightarrow 3. \Rightarrow 4. \Rightarrow 1.$$

1. $\Rightarrow$ 2. Suponhamos que A é invertível e seja B uma matriz em forma de escada obtida a partir de A. Como B é obtida de A através de uma sequência de transformações elementares nas linhas, então existe uma sequência de matrizes elementares (cada uma delas corresponde a uma transformação elementar efectuada nas linhas de A), $E_1, \ldots, E_k$ tais que

$$B = E_k \ldots E_1 A.$$

Porque toda a matriz elementar é invertível e A também é invertível, então B é invertível. Consequentemente, por definição, B não pode ter linhas nulas, e $r(A) = r(B) = n$.

2. $\Rightarrow$ 3. Se $r(A) = n$, então qualquer matriz em forma de escada obtida a partir de A terá característica n. Sendo B' a sua matriz em forma de escada reduzida, então $r(B') = n$. Porque A tem ordem n, então $B' = I_n$.

3. $\Rightarrow$ 4. Se a forma de escada reduzida de A é I_n, então existe uma sequência de matrizes elementares, $E_1, \ldots, E_k$ tais que

$$I_n = E_k \ldots E_1 A.$$

Como as matrizes elementares são invertíveis, então

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

Cada E_i^{-1} é uma matriz elementar, pelo que temos 4.

4. $\Rightarrow$ 1. Se A pode escrever-se como produto de matrizes elementares e, porque, as matrizes elementares são invertíveis e o produto de matrizes invertíveis é invertível, então A é invertível. $\square$

Pelo Teorema anterior vemos que se A é uma matriz invertível de ordem n, a forma de escada reduzida de A é I_n, sendo esta obtida através do produto de matrizes elementares $E_1, \ldots, E_k$. Isto é,

$$I_n = E_k \ldots E_1 A.$$

Daqui vem que

$$A^{-1} = E_k \ldots E_1 = (E_k \ldots E_1) I_n.$$

Mas isto significa que a mesma sequência de transformações elementares que permitiu obter I_n de A, permite obter A^{-1} de I_n. Esquematicamente, se pensarmos na matriz

$$[A|I_n]$$

e efectuarmos a sequência de transformações elementares que permitem obter I_n de A, obtemos a matriz

$$[I_n|A^{-1}].$$

Exemplo 2.38 1. *Consideremos a matriz* $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$. *Vejamos se A é invertível e caso seja, calculemos a sua inversa.*

Vamos usar o algoritmo anterior,

$$\left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ -2 & 2 & 0 & 1 \end{array}\right] \underset{l_2 \to (l_2 + 2l_1)}{\longrightarrow} \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{array}\right].$$

Como obtivemos uma linha de zeros do lado esquerdo, A não é invertível (A é singular).

2. *Consideremos a matriz* $A = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$ *e determinemos, caso exista, a sua inversa.*

Pelo algoritmo temos

$$\left[\begin{array}{ccc|ccc} 2 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \underset{l_1 \to \frac{1}{2}l_1}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \underset{l_3 \to (l_3 + l_1)}{\longrightarrow}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right] \underset{l_1 \to (l_1 - l_2)}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right].$$

Assim, A é invertível (do lado esquerdo aparece I_3) e

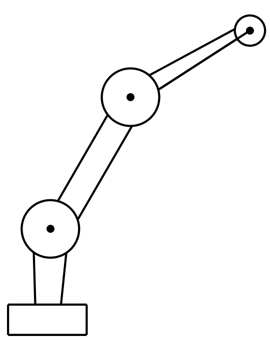
$$A^{-1} = \begin{bmatrix} \frac{1}{2} & -1 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

2.5 Aplicações das Matrizes

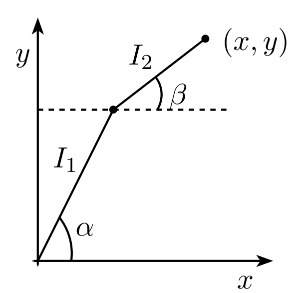
Vejamos algumas aplicações das matrizes a problemas práticos que nos podem surgir.

2.5.1 Uma Aplicação à Robótica

Consideremos o seguinte robot industrial simplificado, que consiste num braço e num antebraço que podem ser girados independentemente por ângulos α e β e que podem ser colocados independentemente com comprimentos I_1 e I_2.



robot



Para ângulos α e β fixos, quais deveriam ser os comprimentos do braço e do antebraço para poder colocar a ponta do robot no ponto (x, y)?

Como é conhecido,

$$\begin{cases} x = I_1 \cos\alpha + I_2 \cos\beta \\ y = I_1 sen\ \alpha + I_2 sen\ \beta \end{cases}$$

Estamos perante um sistema de 2 equações a 2 incógnitas. Este sistema na forma matricial é

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\alpha & \cos\beta \\ sen\ \alpha & sen\ \beta \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}.$$

Porque a matriz simples do sistema é

$$\begin{bmatrix} \cos\alpha & \cos\beta \\ sen\ \alpha & sen\ \beta \end{bmatrix}$$

e $\cos\alpha\ sen\ \beta - sen\ \alpha \cos\beta = sen\ (\beta - \alpha)$, mostra-se que $r(A) = 2$ se, e só se, $\beta - \alpha$ não é múltiplo de π.

Pelo Teorema 2.37, se $\beta - \alpha$ não é múltiplo de π, então A é invertível e como

$$\begin{bmatrix} x \\ y \end{bmatrix} = A \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}, \text{ vem que } \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = A^{-1} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Mas,

$$A^{-1} = \frac{1}{sen\ (\beta - \alpha)} \begin{bmatrix} sen\ \beta & -\cos\beta \\ -sen\ \alpha & \cos\alpha \end{bmatrix}$$

donde,

$$I_1 = \frac{sen\ \beta}{sen\ (\beta - \alpha)} x - \frac{\cos\beta}{sen\ (\beta - \alpha)} y$$

e

$$I_2 = -\frac{sen\ \alpha}{sen\ (\beta - \alpha)} x + \frac{\cos\alpha}{sen\ (\beta - \alpha)} y.$$

2.5.2 Resolução de Sistemas

O processo que foi utilizado no exemplo anterior, pode ser usado para qualquer sistema de n equações a n incógnitas desde que seja conhecido que a matriz simples do sistema é invertível. Isto é, se o sistema de n equações a n incógnitas, estiver na forma matricial $AX = B$, sendo A a matriz simples, X a matriz das incógnitas e B a matriz dos termos independentes e se a matriz A for invertível, então a única solução do sistema (sistema possível e determinado) é

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

Capítulo 3

SUBESPAÇOS VECTORIAIS DE $\mathbb{R}^n$

Neste capítulo iremos estender as noções de recta e plano de $\mathbb{R}^n$, que passam na origem, a outros tipos de objectos geométricos em $\mathbb{R}^n$, os subespaços vectoriais.

3.1 Definição de Subespaço Vectorial de $\mathbb{R}^n$

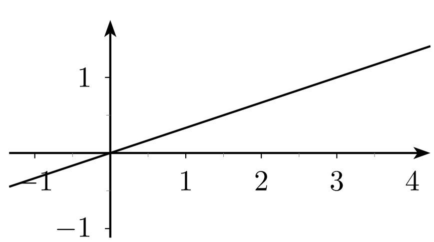
Dado um vector não nulo de $\mathbb{R}^n$, $v = (v_1, v_2, \ldots, v_n)$, a **equação vectorial da recta** que passa pela origem e tem vector director $v \neq 0$ é (estamos a omitir o ponto $(0, \ldots, 0)$)

$$(x_1, x_2, \ldots, x_n) = \lambda(v_1, v_2, \ldots, v_n), \qquad \lambda \in \mathbb{R}.$$

Se pensarmos num plano de $\mathbb{R}^n$, para escrevermos uma sua equação vectorial teremos de ter dois vectores $u = (u_1, u_2, \ldots, u_n)$ e $v = (v_1, v_2, \ldots, v_n)$ de $\mathbb{R}^n$ (que não tenham a mesma direcção), sendo uma **equação vectorial do plano** que passa pela origem,

$$(x_1, x_2, \ldots, x_n) = \lambda(u_1, u_2, \ldots, u_n) + \mu(v_1, v_2, \ldots, v_n), \qquad \lambda, \mu \in \mathbb{R}.$$

Exemplo 3.1 1. *A equação* $(x_1, x_2) = \lambda(3, 1)$, $\lambda \in \mathbb{R}$, *define a equação vectorial da recta de* $\mathbb{R}^2$ *que passa pela origem e tem vector director* $(3, 1)$.



2. *A equação* $(x_1, x_2, x_3) = \lambda_1(2, 0, 1) + \lambda_2(-2, 0, -1)$, *com* $\lambda_1, \lambda_2 \in \mathbb{R}$, *não define um plano de* $\mathbb{R}^3$ *pois os vectores* $(2, 0, 1)$ *e* $(-2, 0, -1)$ *têm a mesma direcção. A equação anterior é equivalente à equação*

$$(x_1, x_2, x_3) = \lambda(2, 0, 1) \qquad \lambda \in \mathbb{R},$$

que define uma recta de $\mathbb{R}^3$.

Notação 3.2 *Para simplificar, se denotarmos um ponto de* $\mathbb{R}^n$ *por uma das letras* $x, y, \ldots,$ *(com ou sem índice) e cada vector de* $\mathbb{R}^n$ *por uma das letras* $u, v, \ldots,$ *(com ou sem índice) (recorde-se que cada ponto de* $\mathbb{R}^n$ *e cada vector de* $\mathbb{R}^n$ *tem* n *coordenadas), as equações anteriores podem ser escritas na forma*

$$\begin{array}{lll} x = \lambda v, & \lambda \in \mathbb{R} & (\textit{equação da recta de } \mathbb{R}^n) \\ x = \lambda_1 u + \lambda_2 v, & \lambda_1, \lambda_2 \in \mathbb{R} & (\textit{equação do plano de } \mathbb{R}^n). \end{array}$$

Observação Tenha em atenção que na forma resumida,

$$x = \lambda v, \qquad \lambda \in \mathbb{R},\ v \neq 0$$

pode designar uma recta de $\mathbb{R}^2$ ou uma recta de $\mathbb{R}^3$ ou, mais geralmente, de $\mathbb{R}^n$, consoante o vector v for um vector de $\mathbb{R}^2$ ou de $\mathbb{R}^3$ ou de $\mathbb{R}^n$.

Exemplo 3.3 *Se* $v = (3, 1)$, *a equação da recta de* $\mathbb{R}^2$ *que passa pela origem é*

$$(x_1, x_2) = \lambda(3, 1), \qquad \lambda \in \mathbb{R},$$

ou resumidamente

$$x = \lambda v, \qquad \lambda \in \mathbb{R}.$$

Se $v = (3, 1, 2)$ *(vector de* $\mathbb{R}^3$*), a recta de* $\mathbb{R}^3$ *que passa pela origem é*

$$(x_1, x_2, x_3) = \lambda(3, 1, 2), \qquad \lambda \in \mathbb{R},$$

ou resumidamente

$$x = \lambda v, \qquad \lambda \in \mathbb{R}.$$

Os planos e as rectas de $\mathbb{R}^n$, que passam pela origem têm duas propriedades muito interessantes:

1. Consideremos o plano de $\mathbb{R}^n$ cuja equação vectorial é

$$x = \lambda_1 v_1 + \lambda_2 v_2, \qquad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Designemos este plano por W.

Se x_1, x_2 são dois elementos do plano W, então existem $\alpha_1, \alpha_2 \in \mathbb{R}$,

$$x_1 = \alpha_1 v_1 + \alpha_2 v_2$$

e existem $\beta_1, \beta_2 \in \mathbb{R}$,

$$x_2 = \beta_1 v_1 + \beta_2 v_2.$$

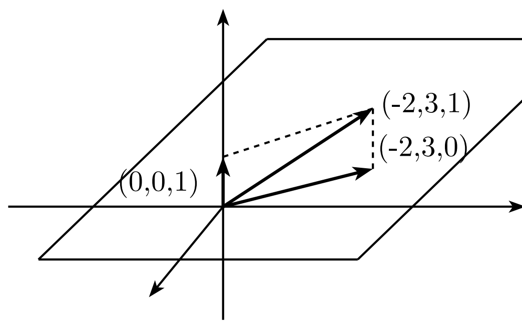
Mas então,

$$x_1 + x_2 = \alpha_1 v_1 + \alpha_2 v_2 + \beta_1 v_1 + \beta_2 v_2 = (\alpha_1 + \beta_1) v_1 + (\alpha_2 + \beta_2) v_2.$$

Porque a soma de reais é um real, $x_1 + x_2$ é um elemento de W.

Por exemplo no plano W de equação

$$x = \lambda_1 (1, 0, 1) + \lambda_2 (-2, 3, 0),$$



os elementos $x_1 = (1, 0, 1)$ e $x_2 = (-2, 3, 0)$ estão em W e o elemento $x_1 + x_2 = (-2, 3, 1)$ também, e é a soma de (0,0,1) com (-2,3,0) ("Regra do Paralelogramo")

2. Com o mesmo plano de $\mathbb{R}^n$, W, e o mesmo elemento x_1, se k for um escalar então

$$k x_1 = k(\alpha_1 v_1 + \alpha_2 v_2) = (k\alpha_1) v_1 + (k\alpha_2) v_2,$$

isto é, $k x_1$ é um elemento de W.

Observação Como não fizemos restricções a λ_1 e a λ_2, a equação de W poderia ser uma equação da recta (caso em que um dos escalares é nulo) ou simplesmente a origem (caso em que os escalares são ambos nulos). Assim, tanto os planos como as rectas, ou a origem, verificam as 2 propriedades anteriores. Quanto à propriedade 1., quando verificada por um subconjunto não vazio de $\mathbb{R}^n$, W, diz-se que W é fechado para a **soma de 2 quaisquer elementos de** W. Quando W verifica 2., diz-se que W é fechado para o **produto de um escalar por um elemento qualquer de** W.

Definição 3.4 *Um subconjunto não vazio de* $\mathbb{R}^n$*,* W*, diz-se um* **subespaço vectorial de** $\mathbb{R}^n$*, ou simplesmente, subespaço de* $\mathbb{R}^n$ *se:*

1. $\forall x_1, x_2 \in W, \qquad x_1 + x_2 \in W$

2. $\forall k \in \mathbb{R}, \forall x_1 \in W, \qquad kx_1 \in W.$

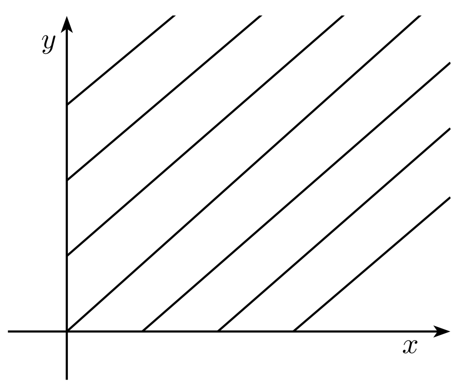
Já observámos que os planos de $\mathbb{R}^n$ que passam pela origem, as rectas de $\mathbb{R}^n$ que passam pela origem, o conjunto formado só pelo ponto zero de $\mathbb{R}^n$, $\{0_{\mathbb{R}^n}\}$, são todos subespaços vectoriais de $\mathbb{R}^n$ (ao último chama-se subespaço nulo). Um outro subespaço de $\mathbb{R}^n$ é o próprio $\mathbb{R}^n$.

Proposição 3.5 *Se W é um subespaço de $\mathbb{R}^n$, então $0_{\mathbb{R}^n} \in W$.*

Demonstração Porque W é um subespaço de $\mathbb{R}^n$, então $W \neq \emptyset$. Seja x_1 um elemento de W. Atendendo à propriedade 2. da definição de subespaço, $0x_1 \in W$. Mas $0x_1 = 0_{\mathbb{R}^n}$, isto é, $0_{\mathbb{R}^n} \in W$. $\square$

Observação Na demonstração anterior quando escrevemos $0x_1 = 0_{\mathbb{R}^n}$, estão presentes 2 zeros. Do lado esquerdo da igualdade é o escalar 0 (0 elemento de $\mathbb{R}$) e do lado direito $0_{\mathbb{R}^n}$ ($(0, 0, \ldots, 0)$ elemento de $\mathbb{R}^n$).

Exemplo 3.6 *Sendo W o conjunto dos elementos (x_1, x_2)de $\mathbb{R}^2$ tais que $x_1 \geq 0$, $x_2 \geq 0$ (os elementos do primeiro quadrante do plano Oxy),*



facilmente se vê que W é fechado para a soma de 2 elementos quaisquer dele. Mas, $(1, 0) \in W$e no entanto,

$$(-1)(1, 0) = (-1, 0) \notin W,$$

isto é, W não verifica 2. Logo, W não é subespaço de $\mathbb{R}^2$.

3.2 Subespaço Gerado

Vejamos mais subespaços de $\mathbb{R}^n$. Sejam $v_1, v_2, \ldots, v_s$ vectores de $\mathbb{R}^n$ e W o conjunto de todas as combinações lineares destes vectores (Definição 1.20), isto é,

$$W = \{x \in \mathbb{R}^n : x = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_s v_s, \lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R}\}.$$

W é um subconjunto não vazio de $\mathbb{R}^n$, pois se fizermos $\lambda_1 = \lambda_2 = \ldots = \lambda_s = 0$ obtemos

$$0v_1 + 0v_2 + \ldots + 0v_s = 0_{\mathbb{R}^n}.$$

Isto é, $0_{\mathbb{R}^n}$ é elemento de W. Fazendo contas análogas às feitas para o plano de $\mathbb{R}^n$, concluímos que

1. Dadas duas combinações lineares de $v_1, v_2, \ldots, v_s$, a sua soma é outra combinação linear de $v_1, v_2, \ldots, v_s$.

2. Multiplicando uma combinação linear de $v_1, v_2, \ldots, v_s$ por um escalar, obtemos outra combinação linear de $v_1, v_2, \ldots, v_s$.

Donde, W é subespaço de $\mathbb{R}^n$. E assim temos o resultado:

Teorema 3.7 *Se $v_1, v_2, \ldots, v_s$ são vectores de $\mathbb{R}^n$, então o conjunto de todas as combinações lineares*

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_s v_s, \qquad \lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R}$$

é um subespaço de $\mathbb{R}^n$.

Temos assim a possibilidade de construção de um subespaço, a partir de um conjunto de vectores de $\mathbb{R}^n$.

Definição 3.8 *O subespaço W de $\mathbb{R}^n$ de todas as combinações lineares dos vectores $v_1, v_2, \ldots, v_s$ de $\mathbb{R}^n$ é denominado por* **subespaço gerado** *por $v_1, v_2, \ldots, v_s$ e denotado por*

$$W = \langle v_1, v_2, \ldots, v_s \rangle.$$

Exemplo 3.9 1. *A recta de $\mathbb{R}^2$ que passa pela origem*

$$(x_1, x_2) = \lambda(3, 1), \qquad \lambda \in \mathbb{R},$$

pode ser expressa como o subespaço de $\mathbb{R}^2$ gerado pelo vector $(3, 1)$,

$$W = \langle (3, 1) \rangle.$$

2. *Como já vimos* $\langle 0_{\mathbb{R}^n} \rangle$ *(subespaço gerado pelo vector nulo de* $\mathbb{R}^n$*) é* $\{0_{\mathbb{R}^n}\}$*, ou seja,* $\langle 0_{\mathbb{R}^n} \rangle = \{0_{\mathbb{R}^n}\}$ *(subespaço nulo).*

3. *Sejam* $e_1 = (1,0,0)$*,* $e_2 = (0,1,0)$*,* $e_3 = (0,0,1)$ *vectores de* $\mathbb{R}^3$ *(estes vectores designam-se por* **vectores canónicos** *de* $\mathbb{R}^3$*). Como,*

   $$(x,y,z) = x(1,0,0) + y(0,1,0) + z(0,0,1) = xe_1 + ye_2 + ze_3$$

   qualquer que seja o elemento (x,y,z) *de* $\mathbb{R}^3$*, então,*

   $$\langle e_1, e_2, e_3 \rangle = \mathbb{R}^3.$$

   Mais geralmente, se $e_1, e_2, \ldots, e_n$ *forem os* **vectores canónicos** *de* $\mathbb{R}^n$*, então*

   $$\langle e_1, e_2, \ldots, e_n \rangle = \mathbb{R}^n.$$

Mostrámos que as possibilidades para os subespaços gerados de $\mathbb{R}^2$ são:

- subespaço nulo
- rectas que passam pela origem
- $\mathbb{R}^2$

e as possibilidades para os subespaços gerados de $\mathbb{R}^3$ são:

- subespaço nulo
- rectas que passam pela origem
- planos que passam pela origem
- $\mathbb{R}^3$

3.3 Dependência e Independência Linear

Como já mencionámos, dada a equação vectorial de $\mathbb{R}^n$

$$x = \lambda_1 v_1 + \lambda_2 v_2, \qquad \lambda_1, \lambda_2 \in \mathbb{R},$$

ela pode definir:

- um plano se os vectores v_1 e v_2 não tiverem a mesma direcção,
- uma recta se v_1 e v_2 tiverem a mesma direcção e forem não nulos,
- o ponto zero se v_1 e v_2 forem o vector nulo.

Assim, vemos que as propriedades geométricas do subespaço de $\mathbb{R}^n$

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_s v_s, \qquad \lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R},$$

são afectadas pelas interligações dos vectores $v_1, v_2, \ldots, v_s$ de $\mathbb{R}^n$.

Definição 3.10 *Dizemos que um conjunto, não vazio, $S = \{v_1, v_2, \ldots, v_s\}$ de vectores de $\mathbb{R}^n$ é* **linearmente independente** *se os únicos escalares que satisfazem a equação*

$$c_1 v_1 + c_2 v_2 + \ldots + c_s v_s = 0_{\mathbb{R}^n}$$

são $c_1 = 0$, $c_2 = 0$,..., $c_s = 0$.

Se existirem escalares, não todos nulos, que satisfaçam esta equação, dizemos que o conjunto é **linearmente dependente**.

Observação Apesar da definição anterior ser aplicada a um conjunto não vazio, muitas vezes dizemos que os vectores $v_1, v_2, \ldots, v_s$ são linearmente independentes ou dependentes, para nos referirmos a que o conjunto $S = \{v_1, v_2, \ldots, v_s\}$ é linearmente independente ou dependente.

Proposição 3.11 *Um conjunto S de vectores de $\mathbb{R}^n$ que contenha o vector nulo é linearmente dependente.*

Demonstração Suponhamos que $S = \{0_{\mathbb{R}^n}, v_2, \ldots, v_s\}$. Porque,

$$1.0_{\mathbb{R}^n} + 0v_2 + \ldots + 0v_s = 0_{\mathbb{R}^n}$$

e o escalar que está associado ao vector nulo é $1 \neq 0$, então S é linearmente dependente. □

Proposição 3.12 *Um conjunto $S = \{v_1, v_2, \ldots, v_s\}$ de $\mathbb{R}^n$ com dois ou mais vectores é linearmente dependente se, e só se, pelo menos um dos vectores de S se escreve como combinação linear dos outros vectores de S.*

Demonstração Suponhamos que S é linearmente dependente. Então, existem escalares $c_1, c_2, \ldots, c_s$, não todos nulos, tais que

$$c_1 v_1 + c_2 v_2 + \ldots + c_s v_s = 0_{\mathbb{R}^n}.$$

Suponhamos, sem perda de generalidade, que $c_1 \neq 0$. Então,

$$v_1 = \left(-\frac{c_2}{c_1}\right) v_2 + \ldots + \left(-\frac{c_s}{c_1}\right) v_s.$$

Ou seja, v_1 escreve-se como combinação linear de $v_2, \ldots, v_s$.

Reciprocamente, suponhamos, sem perda de generalidade, que v_1 se escreve como combinação linear de $v_2, \ldots, v_s$. Assim, existem escalares $d_2, \ldots, d_s$ tais que

$$v_1 = d_2 v_2 + \ldots + d_s v_s.$$

Mas isto implica que

$$1.v_1 + (-d_2)v_2 + \ldots + (-d_s)v_s = 0_{\mathbb{R}^n}.$$

Como o escalar associado a v_1 é $1 \neq 0$, temos que S é linearmente dependente. $\square$

Um processo para determinarmos se um conjunto $S = \{v_1, v_2, \ldots, v_s\}$, de vectores de $\mathbb{R}^n$, é linearmente independente é através de um sistema homogéneo de equações lineares. Consideramos a matriz

$$A = \begin{bmatrix} v_1 & v_2 & \ldots & v_s \end{bmatrix}$$

de tipo $n \times s$, cujas colunas são os vectores de S. Usando esta matriz, podemos reescrever a equação

$$c_1 v_1 + c_2 v_2 + \ldots + c_s v_s = 0_{\mathbb{R}^n}$$

na forma matricial

$$\begin{bmatrix} v_1 & v_2 & \ldots & v_s \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(sistema homogéneo cuja matriz simples do sistema é A).

O problema de determinar se S é linearmente independente reduz-se a determinar se o sistema anterior tem somente a solução nula (se o sistema tiver mais do que uma solução então, S é linearmente dependente).

Exemplo 3.13 *Vejamos se os vectores $v_1 = (1, 2, 0)$, $v_2 = (0, 1, 2)$, $v_3 = (1, 3, 2)$ são linearmente independentes e, caso não sejam, escrevamos o vector nulo de $\mathbb{R}^3$ como combinação linear, não nula, de v_1, v_2, v_3.*

Como já dissemos, vamos ver se o sistema homogéneo

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

só tem a solução nula. Ora, como a matriz ampliada de um sistema homogéneo tem a coluna, correspondente à matriz dos termos independentes, nula e transformações elementares nas linhas não alteram uma coluna nula, não necessitamos de colocar essa coluna para resolver o sistema. Temos

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 0 & 2 & 2 \end{bmatrix} \underset{l_2 \to (l_2 - 2l_1)}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \underset{l_3 \to (l_3 - 2l_2)}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

O sistema correspondente a esta última matriz é

$$\begin{cases} c_1 \quad +c_3 = 0 \\ \quad c_2 + c_3 = 0 \\ \qquad 0 = 0. \end{cases}$$

Tomando $c_3 = -1$ *temos a solução* $c_1 = 1$, $c_2 = 1$, $c_3 = -1$, *ou seja,*

$$1.v_1 + 1.v_2 - 1.v_3 = 0_{\mathbb{R}^3}$$

é uma combinação linear, não nula, do vector nulo.

Não foi mencionado, mas para afirmarmos que os vectores são linearmente dependentes, bastava reparar que $r(A) = 2 < 3 =$*número de incógnitas do sistema, ou seja, o sistema é possível e indeterminado. Logo os vectores são linearmente dependentes.*

Já dissemos que se A é uma matriz $n \times m$ com $m > n$, então, $r(A) < m$. Pelo que:

Proposição 3.14 *Seja* S *um conjunto de* $\mathbb{R}^n$ *com* m *vectores em que* $m > n$. *Então* S *é linearmente dependente.*

Exemplo 3.15 *Consideremos os vectores de* $\mathbb{R}^4$, $v_1 = (1,2,3,4)$, $v_2 = (0,1,0,0)$, $v_3 = (0,0,1,0)$, $v_4 = (1,1,0,0)$, $v_5 = (-1,3,2,0)$. *Pela* Proposição *anterior, como* v_1, v_2, v_3, v_4, v_5 *são* 5 *vectores de* $\mathbb{R}^4$, *então são linearmente dependentes.*

3.4 Aplicações

Vejamos uma aplicação do conceito de independência linear.

3.4.1 Som de Alta Fidelidade

O som de alta fidelidade pode ser gravado digitalmente, através de uma onda sonora à taxa de 44100 vezes por segundo. Assim, um segmento de 10 segundos pode ser representado por um vector de $\mathbb{R}^{441000}$. Um técnico de som num festival de jazz planeia gravar vectores de som com 2 microfones, um vector de som s, através de um microfone colocado perto do saxofone e um vector de som g, através de um microfone colocado perto da guitarra. Já no estúdio de som, ele faz a mistura, criando uma combinação linear dos 2 vectores de som, que produz o efeito sonoro desejado. Suponhamos que cada microfone além de captar o som do instrumento próximo, também grava uma pequena quantidade de som do outro instrumento, de forma que os reais vectores de som gravados são

$u = s + 0,05g$ para o saxofone,
$v = g + 0,12s$ para a guitarra.

Qual a combinação linear de u e v que recria a mistura $\frac{1}{2}(s+g)$?

Pretendemos encontrar escalares λ_1 e λ_2 tais que

$$\lambda_1 u + \lambda_2 v = \frac{1}{2}(s+g)$$

($\frac{1}{2}(s+g)$ como combinação linear de u e v). Porque $u = s+0,05g$ e $v = g+0,12s$, vem que

$$\lambda_1(s+0,05g) + \lambda_2(g+0,12s) = \frac{1}{2}(s+g)$$

ou seja,

$$\left(\lambda_1 + 0,12\lambda_2 - \frac{1}{2}\right) s + \left(0,05\lambda_1 + \lambda_2 - \frac{1}{2}\right) g = 0.$$

Como s e g são vectores linearmente independentes (cada instrumento emite o seu som),

$$\lambda_1 + 0,12\lambda_2 - \frac{1}{2} = 0 \qquad 0,05\lambda_1 + \lambda_2 - \frac{1}{2} = 0.$$

Ou seja, o sistema

$$\begin{cases} \lambda_1 + 0,12\lambda_2 = \frac{1}{2} \\ 0,05\lambda_1 + \lambda_2 = \frac{1}{2}. \end{cases}$$

Pelo que,

$$\lambda_1 = \frac{0,45}{0,994}, \qquad \lambda_2 = \frac{0,443}{0,994}.$$

3.5 Variedades Lineares

A equação vectorial de uma recta de $\mathbb{R}^2$ que passa na origem é

$$x = \lambda_1 v_1 \quad \lambda_1 \in \mathbb{R}.$$

em que v_1 é um vector não nulo. Já a equação

$$x = x_0 + \lambda_1 v_1 \quad \lambda_1 \in \mathbb{R}.$$

é a equação vectorial de uma recta de $\mathbb{R}^2$, paralela à anterior e que passa no ponto x_0. Para obtermos a 2ª recta, fazemos uma translacção da primeira, por x_0.

Mais geralmente, sendo $v_1, v_2, \ldots, v_s$ vectores de $\mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, o conjunto dos pontos de $\mathbb{R}^n$ que verificam a equação

$$x = x_0 + \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_s v_s \quad \lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R},$$

pode ser visto como a translação, por x_0 do subespaço

$$W = \langle v_1, v_2, \ldots, v_s \rangle$$

que se denota por $x_0 + \langle v_1, v_2, \ldots, v_s \rangle$ e se chama **variedade linear de** $\mathbb{R}^n$.

Exemplo 3.16 *Já vimos que a intersecção dos* 3 *planos de* $\mathbb{R}^3$ *de equações gerais* (Exemplo 1.8.2)

$$x - y + z = 1, \qquad 2x + z = 2, \qquad 3x - y + 2z = 3$$

é a recta de $\mathbb{R}^3$ (*equação vectorial*)

$$(x, y, z) = (1, 0, 0) + \lambda\left(-\frac{1}{2}, \frac{1}{2}, 1\right), \qquad \lambda \in \mathbb{R}.$$

(*Recorde-se que construímos o sistema*

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

que teve como conjunto-solução

$$\left\{\left(1 - \frac{1}{2}z, \frac{1}{2}z, z\right) : z \in \mathbb{R}\right\},$$

e cada elemento deste conjunto por ser escrito como

$$\left(1 - \frac{1}{2}z, \frac{1}{2}z, z\right) = (1, 0, 0) + z\left(-\frac{1}{2}, \frac{1}{2}, 1\right).)$$

Se pensarmos na intersecção dos 3 *planos que são paralelos aos dados e que passam pela origem, teremos o sistema homogéneo*

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

O conjunto-solução deste sistema é

$$\left\{\left(-\frac{1}{2}z, \frac{1}{2}z, z\right) : z \in \mathbb{R}\right\},$$

ou seja, a recta de $\mathbb{R}^3$ *que passa pela origem e tem equação vectorial*

$$(x, y, z) = \lambda\left(-\frac{1}{2}, \frac{1}{2}, 1\right), \qquad \lambda \in \mathbb{R}.$$

Pelo que foi explicado anteriormente, o conjunto-solução é um subespaço de $\mathbb{R}^3$, *que é*

$$\left\langle\left(-\frac{1}{2}, \frac{1}{2}, 1\right)\right\rangle$$

e o conjunto-solução dos 3 *planos iniciais é uma variedade linear que resulta da translação deste subespaço por* $x_0 = (1, 0, 0)$, *ou seja,*

$$(1, 0, 0) + \left\langle\left(-\frac{1}{2}, \frac{1}{2}, 1\right)\right\rangle.$$

Repare que x_0 *é uma solução do sistema inicial.*

Proposição 3.17 *Se $AX = 0$ é um sistema homogéneo, então o seu conjunto-solução é um subespaço de $\mathbb{R}^n$, onde n é o número de incógnitas do sistema.*

Proposição 3.18 *Se $AX = B$ é um sistema possível, não homogéneo e se W é o subespaço (conjunto-solução) solução do sistema $AX = 0$, então o conjunto-solução de $AX = B$ é a variedade linear*

$$x_0 + W$$

onde x_0 é uma solução particular de $AX = B$.

Como encontrar os vectores que geram o subespaço solução de um sistema homogéneo?

Suponhamos que o conjunto-solução de determinado sistema homogéneo de $\mathbb{R}^4$ é

$$\{(t_1 + t_2, 2t_2, t_2 - t_3, t_1) : t_1, t_2, t_3 \in \mathbb{R}\}.$$

Porque cada elemento deste conjunto se escreve como combinação linear de vectores de $\mathbb{R}^4$

$$\begin{aligned}(t_1 + t_2, 2t_2, t_2 - t_3, t_1) &= (t_1, 0, 0, t_1) + (t_2, 2t_2, t_2, 0) + (0, 0, -t_3, 0)\\ &= t_1(1, 0, 0, 1) + t_2(1, 2, 1, 0) + t_3(0, 0, -1, 0),\end{aligned}$$

então o subespaço solução é

$$\langle(1, 0, 0, 1), (1, 2, 1, 0), (0, 0, -1, 0)\rangle.$$

Capítulo 4

DETERMINANTES

Ao longo deste capítulo veremos uma outra maneira de determinar se uma matriz quadrada é ou não invertível, aprenderemos um processo de resolver sistemas de n equações a n incógnitas, calcularemos áreas de paralelogramos (em $\mathbb{R}^2$ e em $\mathbb{R}^3$) e volumes de paralelepípedos (em $\mathbb{R}^3$) e por fim tentaremos encontrar soluções para o seguinte problema:

"Dada uma matriz quadrada de ordem n, A, quais os valores de λ, se existirem, para os quais existe uma matriz, não nula, X de tipo $n \times 1$ tal que

$$(\lambda I_n - A)X = 0?\text{"}$$

Todos estes assuntos se resolvem usando "**DETERMINANTES**".

4.1 Definição de Determinante

Além da definição de determinante de uma matriz quadrada, ao longo desta secção veremos exemplos do cálculo do determinante de certas matrizes.

Notação 4.1 *Sejam A uma matriz de ordem n $(n \geq 2)$ e $i, j \in \{1, \ldots, n\}$, denotamos por*

$$A(i|j)$$

a matriz que resulta de A retirando a linha i e a coluna j.

Observação Repare-se que $A(i|j)$ é uma matriz quadrada de ordem $n-1$.

Exemplo 4.2 *Sendo* $A = \begin{bmatrix} 1 & 0 & 2 \\ 4 & -1 & 3 \\ -3 & -2 & 5 \end{bmatrix}$*, então*

$$A(1|2) = \begin{bmatrix} 4 & 3 \\ -3 & 5 \end{bmatrix}, \qquad A(3|1) = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}.$$

Definição 4.3 *Seja $A = [a_{ij}]$ uma matriz quadrada de ordem n. Chamamos* **determinante** *de A, e denotamos por $\det A$ ou $|A|$, ao número*

- *Se $n = 1$, então*
$$\det A = a_{11}.$$
- *Se $n > 1$, então*
$$\det A = a_{11}(-1)^{1+1} \det A(1|1) + \ldots + a_{1n}(-1)^{1+n} \det A(1|n),$$
ou seja, se $n > 1$,
$$\det A = \sum_{i=1}^{n} a_{1i}(-1)^{1+i} \det A(1|i).$$

Exemplo 4.4 1. *Se A for uma matriz de ordem 2, isto é,*
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
então,
$$\det A = a_{11}(-1)^{1+1} \det A(1|1) + a_{12}(-1)^{1+2} \det A(1|2).$$
Porque $A(1|1) = a_{22}$ e $A(1|2) = a_{21}$ vem que
$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$
(*Uma forma de fixar quanto é o determinante de uma matriz de ordem 2, é pensar que é o produto dos elementos da diagonal principal menos o produto dos elementos da outra diagonal, isto é, se*
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
então
$$\det A = a_{11}a_{22} - a_{12}a_{21}.)$$

ATENÇÃO: *Esta regra só é válida para matrizes de ordem* 2!!!

2. *Se A for uma matriz de ordem 3, isto é,*
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$
então,
$$\begin{aligned} \det A &= a_{11}(-1)^{1+1} \det A(1|1) + a_{12}(-1)^{1+2} \det A(1|2) + a_{13}(-1)^{1+3} \det A(1|3) \\ &= a_{11}(-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \\ &= (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{31}a_{22}). \end{aligned}$$

Para fixar esta expressão, recorramos à **REGRA de SARRUS***:*

coloquemos as 3 *colunas da matriz* A *e repitamos as* 2 *primeiras colunas, isto é,*

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Desenhando todas as diagonais possíveis, com 3 *elementos cada, obtemos o esquema da figura. O determinante da matriz é a soma do produto dos* 3 *elementos de cada diagonal cuja seta vai da esquerda para a direita, menos a soma do produto dos* 3 *elementos de cada diagonal cuja seta vai da direita para a esquerda.*

ATENÇÃO: *Esta regra só é válida para matrizes de ordem* 3!!!

4.2 Teorema de Laplace

Como se vê, se a expressão do determinante de uma matriz de ordem 3 envolve 6 parcelas, então a do determinante de uma matriz de ordem 4 terá 24 parcelas, o que torna o processo muitas vezes complicado e moroso. Vejamos, então, outros processos para calcular o determinante.

Definição 4.5 *Seja* A *uma matriz de ordem* n*, com* $n \geq 2$*. Chamamos* **complemento algébrico** *do elemento da posição* (i,j) *de* A*, e denotamos por* $\widehat{A}_{ij}$*, o elemento*

$$\widehat{A}_{ij} = (-1)^{i+j} \det A(i|j).$$

Exemplo 4.6 *Sendo* $A = \begin{bmatrix} 1 & 0 & 2 \\ 4 & -1 & 3 \\ -3 & -2 & 5 \end{bmatrix}$*, então*

$$\widehat{A}_{12} = (-1)^{1+2} \det A(1|2) = (-1)^{1+2} \begin{vmatrix} 4 & 3 \\ -3 & 5 \end{vmatrix} = -(20+9) = -29,$$

$$\widehat{A}_{31} = (-1)^{3+1} \det A(3|1) = (-1)^{3+1} \begin{vmatrix} 0 & 2 \\ -1 & 3 \end{vmatrix} = 0 + 2 = 2.$$

Teorema 4.7 (Teorema de Laplace) *Sejam* $A = [a_{ij}]$ *uma matriz quadrada de ordem* n*,* $n \geq 2$ *e* $k \in \{1, \ldots, n\}$*, então*

1. $\det A = a_{k1}\widehat{A}_{k1} + \ldots + a_{kn}\widehat{A}_{kn} = \sum_{j=1}^{n} a_{kj}\widehat{A}_{kj}$.

2. $\det A = a_{1k}\widehat{A}_{1k} + \ldots + a_{nk}\widehat{A}_{nk} = \sum_{i=1}^{n} a_{ik}\widehat{A}_{ik}$.

Observação Em 1., do Teorema de Laplace, faz-se o desenvolvimento do cálculo do determinante, através da linha k e em 2., o desenvolvimento é feito através da coluna k.

Na prática, quando estamos a calcular o determinante de uma matriz, escolhemos a linha ou a coluna da matriz que tenha maior número de zeros, pois teremos menos parcelas no cálculo do determinante, através do Teorema de Laplace.

A definição de determinante de uma matriz quadrada A é o desenvolvimento do determinante, usando o Teorema de Laplace, através da primeira linha da matriz A.

Exemplo 4.8 *Se* $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & -2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$,
como a segunda coluna e a terceira linha de A têm duas posições iguais a zero, vamos escolher uma delas para calcularmos o determinante de A.

Através da segunda coluna, pela expressão 2. *do* Teorema 4.7 *temos*

$$\begin{aligned} \det A &= 0\widehat{A}_{12} - 2\widehat{A}_{22} + 0\widehat{A}_{32} = -2(-1)^{2+2} \det A(2|2) \\ &= -2(-1)^{2+2} \begin{vmatrix} 1 & -1 \\ 0 & 3 \end{vmatrix} = -2(3-0) = -6 \end{aligned}$$

Para comprovarmos que o valor do determinante da matriz A é o mesmo se utilizarmos o Teorema de Laplace com qualquer uma das linhas ou colunas de A, vamos calcular o determinante de A, através da terceira linha de A.

Assim, usando a expressão 1. *do Teorema de Laplace, temos*

$$\begin{aligned} \det A &= 0\widehat{A}_{31} + 0\widehat{A}_{32} + 3\widehat{A}_{33} = 3(-1)^{3+3} \det A(3|3) \\ &= 3(-1)^{3+3} \begin{vmatrix} 1 & 0 \\ 2 & -2 \end{vmatrix} = 3(-2-0) = -6 \end{aligned}$$

Se tivessemos utilizado a definição de determinante, teríamos o desenvolvimento através da primeira linha de A, ou seja,

$$\begin{aligned} \det A &= 1\widehat{A}_{11} + 0\widehat{A}_{12} - 1\widehat{A}_{13} = 1(-1)^{1+1} \det A(1|1) - 1(-1)^{1+3} \det A(1|3) \\ &= 1(-1)^{1+1} \begin{vmatrix} -2 & 4 \\ 0 & 3 \end{vmatrix} - 1(-1)^{1+3} \begin{vmatrix} 2 & -2 \\ 0 & 0 \end{vmatrix} = 1(-6-0) - 1(0-0) = -6 \end{aligned}$$

Como consequência do Teorema de Laplace surge o seguinte corolário.

Corolário 4.9 *Seja A uma matriz, quadrada de ordem n, com uma linha ou uma coluna nula, então* $\det A = 0$.

Proposição 4.10 *Seja A uma matriz, quadrada de ordem n, triangular superior (respectivamente, triangular inferior), então o determinante de A é o produto dos elementos da diagonal principal.*

Demonstração A demonstração deste resultado faz-se por indução em n.

Seja $A = [a_{ij}]$ uma matriz triangular superior de ordem n.

Para $n = 1$, temos $\det A = a_{11}$, que é o produto dos elementos da sua diagonal principal.

Suponhamos que o resultado se verifica para qualquer matriz triangular superior de ordem $k-1$, em que $k-1 \geq 1$. Consideremos uma matriz $A = [a_{ij}]$ de ordem k.

Recorde-se que se A é triangular superior então

$$a_{k1} = 0, \ldots, a_{k,k-1} = 0.$$

Se fizermos o desenvolvimento do determinante de A, pela linha k, usando o Teorema de Laplace (Teorema 4.7), temos

$$\det A = a_{kk}\widehat{A}_{kk} = a_{kk}(-1)^{k+k}\det A(k|k).$$

Dado que $A(k|k)$ é triangular superior, mas de ordem $k-1$, pela hipótese de indução temos que

$$\det A(k|k) = a_{11}a_{22}\ldots a_{k-1,k-1}.$$

Assim,

$$\det A = a_{kk}\widehat{A}_{kk} = a_{11}a_{22}\ldots a_{k-1,k-1}a_{kk}.$$

Usando o Princípio de Indução, concluímos o resultado. □

Outro resultado que se demonstra usando o Princípio de Indução é o seguinte:

Proposição 4.11 *Seja A uma matriz quadrada de ordem n, então,*

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

Demonstração Seja $A = [a_{ij}]$ uma matriz de ordem n.

Se $n = 1$, então $A^T = a_{11} = A$, pelo que $\det A^T = \det A$.

Suponhamos que o resultado é verdadeiro para qualquer matriz de ordem $k-1$, com $k-1 \geq 1$, e seja $A = [a_{ij}]$ uma matriz de ordem k.

Porque as linhas de A^T são as colunas de A, fazendo o desenvolvimento do determinante de A^T pela primeira linha, temos

$$\begin{aligned}\det A^T &= (A^T)_{11}\widehat{A^T}_{11} + \ldots + (A^T)_{1k}\widehat{A^T}_{1k} \\ &= a_{11}\widehat{A^T}_{11} + \ldots + a_{k1}\widehat{A^T}_{1k}.\end{aligned}$$

Como

$$\begin{aligned}
(A^T)_{1i} &= (-1)^{1+i}\det A^T(1|i)\\
&= (-1)^{1+i}\det(A(i|1))^T\\
&= (-1)^{1+i}\det A(i|1) &&\text{(por hipótese de indução)}\\
&= \widehat{A}_{i1}
\end{aligned}$$

temos que,

$$\begin{aligned}
\det A^T &= a_{11}\widehat{A}_{11} + \ldots + a_{k1}\widehat{A}_{k1}\\
&= \det A &&\text{(desenvolvimento do determinante pela 1}^{\text{a}}\text{ coluna de } A).
\end{aligned}$$

Usando o Princípio de Indução, concluímos o resultado. □

4.3 Determinante e Transformações Elementares

O Teorema seguinte mostra as alterações que as transformações elementares nas linhas de uma matriz de ordem n, provocam ao seu determinante.

Teorema 4.12 *Seja A uma matriz de ordem n. Então*

1. *Se B é a matriz que resulta de A multiplicando uma linha de A pelo escalar k,*

$$\det B = k\det A.$$

2. *Se B é a matriz que resulta de A trocando duas linhas de A,*

$$\det B = -\det A.$$

3. *Se B é a matriz que resulta de A somando um múltiplo de uma linha de A a uma outra,*

$$\det B = \det A.$$

Demonstração

1. Suponhamos que é a linha i de A que foi multiplicada por k para obtermos B. Fazendo o desenvolvimento do determinante de B pela linha i (Teorema de Laplace) e partindo do princípio que $A = [a_{ij}]$ temos

$$\det B = (ka_{i1})\widehat{B}_{i1} + \ldots + (ka_{in})\widehat{B}_{in}.$$

Porque as outras linhas de B são as de A, vem que

$$\det B = k(a_{i1}\widehat{A}_{i1} + \ldots + a_{in}\widehat{A}_{in}) = k\det A.$$

2. Suponhamos, sem perda de generalidade, que trocámos as linhas i e j, com $i < j$, de A para obtermos B, ou seja,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \qquad \text{e} \qquad \begin{array}{r} \\ \\ \\ B = \\ \text{linha } {\scriptstyle j\to} \\ \\ \\ \end{array} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Fazendo o desenvolvimento do determinante de B (Teorema de Laplace) pela linha j obtemos,

$$\det B = a_{i1}\widehat{B}_{j1} + \ldots + a_{in}\widehat{B}_{jn}.$$

Para cada $k \in \{1, \ldots, n\}$, $\widehat{B}_{jk} = (-1)^{j+k} \det B(j|k)$ ou seja, sendo

$$\begin{array}{r} \\ \\ \text{linha } {\scriptstyle i\to} \\ {}^kC = B(j|k) = \\ \\ \\ \\ \\ \end{array} \begin{bmatrix} a_{11} & \cdots & a_{1,k-1} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{j1} & \cdots & a_{j,k-1} & a_{j,k+1} & \cdots & a_{jn} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & & a_{i-1,k-1} & a_{i-1,k+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & & a_{i+1,k-1} & a_{i+1,k+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & & a_{n,k-1} & a_{n,k+1} & \cdots & a_{nn} \end{bmatrix}$$

e fazendo o desenvolvimento pela linha i (Teorema de Laplace), obtemos

$$\det B(j|k) = a_{j1}\widehat{{}^kC}_{i1} + \ldots + a_{j,k-1}\widehat{{}^kC}_{i,k-1} + a_{j,k+1}\widehat{{}^kC}_{ik} + \ldots + a_{jn}\widehat{{}^kC}_{i,n-1}$$

onde,

$$\widehat{{}^kC}_{il} = (-1)^{i+l} \det(B(j|k))(i|l) \qquad \text{para } l \in \{1, \ldots, n-1\}.$$

Mas a matriz

$$(B(j|k))(i|l) = (A(i|k))(j-1|l),$$

pois quando retiramos a linha i à matriz A, a linha j de A passa a ser a linha $j-1$ desta matriz.

Assim,

$$\begin{aligned} \widehat{{}^kC}_{il} &= (-1)^{i+l} \det(A(i|k))(j-1|l) \\ &= (-1)^{i+l-(j-1+l)}(-1)^{j-1+l} \det(A(i|k))(j-1|l) \\ &= (-1)^{i+l-(j-1+l)}\widehat{{}^kD}_{j-1,l} \\ &= (-1)^{i-j+1}\widehat{{}^kD}_{j-1,l} \end{aligned}$$

sendo

$$\widehat{{}^kD}_{j-1,l} = (-1)^{j-1+l} \det(A(i|k))(j-1|l) \qquad \text{e} \qquad {}^kD = A(i|k),$$

para $k \in \{1, \ldots, n\}$. Donde,

$$\begin{aligned}
\det B(j|k) &= a_{j1}{}^k\widehat{C}_{i1} + \ldots + a_{j,k-1}{}^k\widehat{C}_{i,k-1} + a_{j,k+1}{}^k\widehat{C}_{ik} + \ldots + a_{jn}{}^k\widehat{C}_{i,n-1} \\
&= a_{j1}(-1)^{i-j+1}{}^k\widehat{D}_{j-1,1} + \ldots + a_{j,k-1}(-1)^{i-j+1}{}^k\widehat{D}_{j-1,k-1} + \\
&\quad + a_{j,k+1}(-1)^{i-j+1}{}^k\widehat{D}_{j-1,k} + \ldots + a_{jn}(-1)^{i-j+1}{}^k\widehat{D}_{j-1,n-1} \\
&= (-1)^{i-j+1} \det A(i|k)
\end{aligned}$$

e portanto,

$$\widehat{B}_{jk} = (-1)^{j+k-i-k}(-1)^{i-j+1}\widehat{A}_{ik} = -\widehat{A}_{ik},$$

$$\begin{aligned}
\det B &= a_{i1}\widehat{B}_{j1} + \ldots + a_{in}\widehat{B}_{jn} \\
&= a_{i1}(-\widehat{A}_{i1}) + \ldots + a_{in}(-\widehat{A}_{in}) \\
&= -\det A
\end{aligned}$$

3. Suponhamos que para obter B efectuámos a transformação elementar $l_i \longrightarrow (l_i + kl_j)$. Então, sendo $A = [a_{ij}]$, se efectuarmos o desenvolvimento do determinante de B pela sua linha i (Teorema de Laplace) temos

$$\begin{aligned}
\det B &= (a_{i1} + ka_{j1})\widehat{B}_{i1} + \ldots + (a_{in} + ka_{jn})\widehat{B}_{in} \\
&= (a_{i1}\widehat{B}_{i1} + \ldots + a_{in}\widehat{B}_{in}) + k(a_{j1}\widehat{B}_{i1} + \ldots + a_{jn}\widehat{B}_{in}).
\end{aligned}$$

Como retirando a linha i de B obtemos uma matriz igual à matriz A retirando a linha i, vem que

$$\det B = (a_{i1}\widehat{A}_{i1} + \ldots + a_{in}\widehat{A}_{in}) + k(a_{j1}\widehat{A}_{i1} + \ldots + a_{jn}\widehat{A}_{in}).$$

Sendo C a matriz que se obtém de A substituindo a sua linha i, pela sua linha j, então

$$\det B = \det A + k \det C.$$

Repare-se que C é a matriz,

$$C = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \vdots & & \vdots \\
a_{j1} & a_{j2} & \cdots & a_{jn} \\
\vdots & \vdots & & \vdots \\
a_{j1} & a_{j2} & \cdots & a_{jn} \\
\vdots & \vdots & & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\begin{array}{l}
\\ \\ \leftarrow \\ \\ \leftarrow \\ \\ \\
\end{array}
\text{linhas } i \text{ e } j$$

que tem as linhas i e j iguais.

Afirmação 4.13 *Se C é uma matriz de ordem n, com as linhas i e j iguais ($i \neq j$), então* $\det C = 0$.

Demonstração Se D for a matriz que se obtém de C trocando a linha i com a linha j, então de 2.,

$$\det D = -\det C.$$

Mas $D = C$, pelo que $\det D = \det C$ e consequentemente,

$$\det C = 0.$$

$\square$

Usando esta afirmação, temos que $\det B = \det A$. Ou seja, a condição 3. verifica-se. $\square$

Este Teorema dá-nos um processo de calcularmos o determinante de uma matriz.

Através de transformações elementares, conseguimos obter uma matriz triangular e como sabemos calcular o determinante de uma matriz triangular (Proposição 4.10), usando o Teorema anterior temos o determinante da matriz inicial.

Exemplo 4.14 *Se* $A = \begin{bmatrix} 0 & 2 & 3 \\ 1 & 0 & 2 \\ 2 & 4 & 6 \end{bmatrix}$, *então vamos calcular o determinante de* A *usando o* Teorema 4.12 *e depois a* Proposição 4.10.

$$|A| = \begin{vmatrix} 0 & 2 & 3 \\ 1 & 0 & 2 \\ 2 & 4 & 6 \end{vmatrix} \underset{\substack{\uparrow \\ \textit{trocámos} \\ \textit{linhas, por } 2.}}{\overset{=}{\scriptstyle l_1 \leftrightarrow l_3}} - \begin{vmatrix} 2 & 4 & 6 \\ 1 & 0 & 2 \\ 0 & 2 & 3 \end{vmatrix} \underset{\substack{\uparrow \\ \textit{mult. uma linha} \\ \textit{por } \frac{1}{2}, \textit{ por } 1.}}{\overset{=}{\scriptstyle l_1 \to \frac{1}{2} l_1}} -2 \begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 0 & 2 & 3 \end{vmatrix}$$

$$\underset{\substack{\uparrow \\ \textit{por } 3.}}{\overset{=}{\scriptstyle l_2 \to (l_2 - l_1)}} -2 \begin{vmatrix} 1 & 2 & 3 \\ 0 & -2 & -1 \\ 0 & 2 & 3 \end{vmatrix} \underset{\substack{\uparrow \\ \textit{por } 3.}}{\overset{=}{\scriptstyle l_3 \to (l_3 + l_2)}} -2 \begin{vmatrix} 1 & 2 & 3 \\ 0 & -2 & -1 \\ 0 & 0 & 2 \end{vmatrix} \underset{\substack{\uparrow \\ \textit{determinante de} \\ \textit{matriz triang.}}}{=} -2(-4) = 8$$

4.4 Outra Caracterização das Matrizes Invertíveis

No capítulo 2, vimos diversas caracterizações das matrizes invertíveis. Usando o determinante iremos conhecer um outro processo, de verificação da invertabilidade de uma matriz.

Teorema 4.15 *Seja* A *uma matriz quadrada.*
A é invertível se, e só se, $\det A \neq 0$.

Demonstração Suponhamos que A tem ordem n e que é invertível. Então pelo Teorema 2.37, a forma de escada reduzida de A é I_n. Atendendo ao Teorema 4.12 (I_n foi obtida a partir de A efectuando um número finito de transformações elementares em que nenhuma delas é multiplicar uma linha pelo escalar zero) e porque $\det I_n \neq 0$, então $\det A \neq 0$.

Reciprocamente, se $\det A \neq 0$ e se R é a forma de escada reduzida de A, então pelo Teorema 4.12, $\det R \neq 0$. Pelo Corolário 4.9, R não tem linhas nulas, pelo que $R = I_n$. Pelo Teorema 2.37, A é invertível. □

Vejamos um processo de calcular a inversa de uma matriz invertível, usando o determinante.

Definição 4.16 *Seja* $A = [a_{ij}]$ *uma matriz quadrada de ordem* n*, com* $n \geq 2$*.*

Chamamos **matriz dos complementos algébricos** *de* A *à matriz dos seus complementos algébricos e denotamo-la por* $\widehat{A}$*.*

Chamamos **adjunta** *de* A *à matriz transposta da matriz* $\widehat{A}$ *e denotamo-la por adj* A*, isto é,*

$$(adj\ A)_{ij} = \widehat{A}_{ji}.$$

Exemplo 4.17 *Se* $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 3 \\ -2 & 0 & 1 \end{bmatrix}$ *então para calcularmos a matriz adjunta de* A*, temos de calcular a matriz dos complementos algébricos de* A*. Ora*

$$\widehat{A}_{11} = (-1)^{1+1} \det A(1|1) = (-1)^{1+1} \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} = (2-0) = 2$$

$$\widehat{A}_{12} = (-1)^{1+2} \det A(1|2) = (-1)^{1+2} \begin{vmatrix} -1 & 3 \\ -2 & 1 \end{vmatrix} = -(-1+6) = -5$$

$$\widehat{A}_{13} = (-1)^{1+3} \det A(1|3) = (-1)^{1+3} \begin{vmatrix} -1 & 2 \\ -2 & 0 \end{vmatrix} = (0+4) = 4$$

$$\widehat{A}_{21} = (-1)^{2+1} \det A(2|1) = (-1)^{2+1} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -(1-0) = -1$$

$$\widehat{A}_{22} = (-1)^{2+2} \det A(2|2) = (-1)^{2+2} \begin{vmatrix} 0 & 0 \\ -2 & 1 \end{vmatrix} = (0-0) = 0$$

$$\widehat{A}_{23} = (-1)^{2+3} \det A(2|3) = (-1)^{2+3} \begin{vmatrix} 0 & 1 \\ -2 & 0 \end{vmatrix} = -(0+2) = -2$$

$$\widehat{A}_{31} = (-1)^{3+1} \det A(3|1) = (-1)^{3+1} \begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} = (3-0) = 3$$

$$\widehat{A}_{32} = (-1)^{3+2} \det A(3|2) = (-1)^{3+2} \begin{vmatrix} 0 & 0 \\ -1 & 3 \end{vmatrix} = -(0-0) = 0$$

$$\widehat{A}_{33} = (-1)^{3+3} \det A(3|3) = (-1)^{3+3} \begin{vmatrix} 0 & 1 \\ -1 & 2 \end{vmatrix} = (0+1) = 1$$

Assim, pela definição, porque

$$\widehat{A} = \begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} & \widehat{A}_{13} \\ \widehat{A}_{21} & \widehat{A}_{22} & \widehat{A}_{23} \\ \widehat{A}_{31} & \widehat{A}_{32} & \widehat{A}_{33} \end{bmatrix} = \begin{bmatrix} 2 & -5 & 4 \\ -1 & 0 & -2 \\ 3 & 0 & 1 \end{bmatrix}$$

e porque
$$(\widehat{A})^T = adj\ A,$$
vem que
$$adj\ A = \begin{bmatrix} 2 & -1 & 3 \\ -5 & 0 & 0 \\ 4 & -2 & 1 \end{bmatrix}.$$

Teorema 4.18 *Seja A uma matriz quadrada de ordem n. Então*

1. $A.adj\ A = (\det A)I_n$.

2. *Se A é invertível, então* $A^{-1} = \frac{1}{\det A} adj\ A$.

Demonstração

1. Sendo $A = [a_{ij}]$ então $adj\ A = [\widehat{A}_{ij}]^T = [\widehat{A}_{ji}]$. Usando o facto de $(A.adj\ A)_{ij}$ ser o produto da i-ésima matriz linha de A pela j-ésima matriz coluna de $adj\ A$, vem que o elemento da posição (i,i) de $A.adj\ A$ é
$$a_{i1}\widehat{A}_{i1} + \ldots + a_{in}\widehat{A}_{in}.$$
Pelo Teorema de Laplace isto é o desenvolvimento do $\det A$.
O elemento da posição (i,j) de $A.adj\ A$ é (em que $i \neq j$)
$$a_{i1}\widehat{A}_{j1} + \ldots + a_{in}\widehat{A}_{jn}.$$
Pelo Teorema de Laplace isto é o desenvolvimento do determinante da matriz C que se obtém de A substituindo a linha j de A, pela sua linha i. Atendendo à Afirmação 4.13, porque C tem duas linhas iguais, $\det C = 0$. Donde, obtemos o resultado.

2. Se A é invertível, existe A^{-1}. Por 1.,
$$A.adj\ A = (\det A)I_n.$$
Donde,
$$A^{-1}(A.adj\ A) = A^{-1}(\det A)$$
ou seja, porque pelo Teorema 4.15, $\det A \neq 0$, então
$$A^{-1} = \frac{1}{\det A} adj\ A.$$

□

Exemplo 4.19 *Usando a matriz do* Exemplo 4.17, *porque* $\det A = -5 \neq 0$, *então A é invertível e pelo* Teorema 4.18,
$$A^{-1} = -\frac{1}{5}\begin{bmatrix} 2 & -1 & 3 \\ -5 & 0 & 0 \\ 4 & -2 & 1 \end{bmatrix}.$$

4.5 Sistemas de Cramer

Dado um sistema de n equações a n incógnitas, ele pode escrever-se na forma matricial por

$$AX = B.$$

Se A for invertível, então $r(A) = r([A|B]) = n$ e o sistema é possível e determinado (estes sistemas são chamados **Sistemas de Cramer**). O Teorema seguinte diz-nos como calcular a única solução deste sistema usando determinantes.

Teorema 4.20 (**Regra de Cramer**) *Seja $AX = B$ um sistema de n equações a n incógnitas, tal que, A é invertível. A única solução do sistema é*

$$x_1 = \frac{\det A_1}{\det A}, \qquad \ldots \qquad , x_n = \frac{\det A_n}{\det A}$$

em que A_j é a matriz que resulta substituindo a j-ésima coluna de A por B.

Demonstração Sendo $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$, como A é invertível então, de $AX = B$, obtemos $A^{-1}(AX) = A^{-1}B$, ou seja, $X = A^{-1}B$. Porque

$$A^{-1}B = \left(\frac{1}{\det A} adj\ A\right) B = \frac{1}{\det A}\left((adj\ A)B\right),$$

então

$$x_j = \frac{1}{\det A}\left((adj\ A)B\right)_{j1} = \frac{1}{\det A}(b_1\widehat{A}_{1j} + \ldots + b_n\widehat{A}_{nj}).$$

Pelo Teorema de Laplace,

$$x_j = \frac{1}{\det A}\det A_j = \frac{\det A_j}{\det A}.$$

□

Exemplo 4.21 *O sistema* $\begin{cases} 3x + 3y + z = 4 \\ x + 2y - z = 0 \\ 2x + y + z = 3 \end{cases}$ *é tal que a matriz simples é*

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

e a matriz dos termos independentes é

$$B = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}$$

Porque $\det A = -3 \neq 0$, *então* A *é invertível. Pelo* Teorema 4.20*,*

$$x = \frac{\begin{vmatrix} 4 & 3 & 1 \\ 0 & 2 & -1 \\ 3 & 1 & 1 \end{vmatrix}}{-3} = 1, \qquad y = \frac{\begin{vmatrix} 3 & 4 & 1 \\ 1 & 0 & -1 \\ 2 & 3 & 1 \end{vmatrix}}{-3} = 0, \qquad z = \frac{\begin{vmatrix} 3 & 3 & 4 \\ 1 & 2 & 0 \\ 2 & 1 & 3 \end{vmatrix}}{-3} = 1.$$

4.6 Determinante do Produto de Matrizes

Observemos, em primeiro lugar, o que acontece ao determinante de uma matriz, quando esta é multiplicada à esquerda por uma matriz elementar.

Pelo Teorema 4.12 concluímos o seguinte lema.

Lema 4.22 *Seja E uma matriz elementar de ordem n e B uma matriz de ordem n.*

1. *Se E se obtém multiplicando uma linha de I_n por k, então*
$$\det E = k \qquad e \qquad \det(EB) = \det E . \det B.$$
2. *Se E se obtém trocando duas linhas de I_n, então*
$$\det E = -1 \qquad e \qquad \det(EB) = \det E . \det B.$$
3. *Se E se obtém somando um múltiplo de uma linha de I_n a outra linha, então*
$$\det E = 1 \qquad e \qquad \det(EB) = \det E . \det B.$$

Lema 4.23 *Sejam A e B matrizes de ordem n.*

1. *Se $AB = I_n$ ou $BA = I_n$, então A e B são invertíveis.*
2. *Se AB é invertível então A e B são invertíveis.*

Demonstração

1. Suponhamos que $AB = I_n$. Para mostrar que A é invertível, vamos ver que o sistema homogéneo $BX = 0$ é possível e determinado. Seja X uma solução do sistema $BX = 0$, então
$$X = I_n X \underset{\substack{\uparrow \\ AB = I_n}}{=} (AB)X = A(BX) = A0 = 0,$$
isto é, a única solução do sistema é a solução nula. Pelo Corolário 1.18, $r(B) = n$. Então B é invertível. Assim,
$$A = AI_n = A(BB^{-1}) = (AB)B^{-1} = I_n B^{-1} = B^{-1},$$
isto é, A é invertível.

2. Se AB é invertível, então

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

Por 1., A é invertível. Mas também,

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

Por 1., B é invertível. □

Teorema 4.24 *Sejam A e B duas matrizes de ordem n. Então,*

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

Demonstração Se A não é invertível, pelo Lema anterior, AB não é invertível. Porque $\det A = 0$, $\det(AB) = 0$, vem que

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

Se A é invertível, então A pode escrever-se como produto de matrizes elementares, ou seja,

$$A = E_1 E_2 \ldots E_s.$$

Assim, usando o Lema 4.22

$$\begin{aligned}
\det(AB) &= \det(E_1((E_2 \ldots E_s)B)) = \det E_1. \det((E_2 \ldots E_s)B) \\
&= \ldots = \det E_1. \det E_2 \ldots \det E_s. \det B = \det(E_1 E_2 \ldots E_s). \det B \\
&= \det A. \det B.
\end{aligned}$$

□

Corolário 4.25 *Seja A uma matriz de ordem n. Se A é invertível então*

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

Demonstração Porque A é invertível, $A^{-1}A = I_n$. Usando o Teorema anterior temos que

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

Pela Proposição 4.10,

$$\det I_n = 1.$$

Donde,

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

ou seja,

$$\det A^{-1}.\det A = 1.$$

Mas se A é invertível, pelo Teorema 4.15, $\det A \neq 0$. Logo,

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

□

Observação Não se pense que se A e B são matrizes de ordem n então,

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

Por exemplo, se

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

então,

$$\det(A+B) = \det \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

e

$$\det A + \det B = -2 + 0 = -2.$$

Portanto, neste caso temos que $\det(A+B) = 0 \neq -2 = \det A + \det B$.

Mas não se pense, por outro lado, que temos sempre

$$\det(A+B) \neq \det A + \det B.$$

Pode acontecer, em certas ocasiões, que sendo C e D matrizes de ordem n,

$$\det(C+D) = \det C + \det D.$$

Por exemplo, se

$$C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \text{e} \qquad D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

então,

$$\det(C+D) = \det \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

e

$$\det C + \det D = -1 + 1 = 0.$$

Portanto, neste caso temos que $\det(C+D) = 0 = \det C + \det D$.

4.7 Interpretação Geométrica de Determinantes 2×2

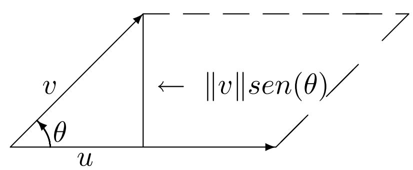
Nesta secção veremos a que corresponde, numa visão geométrica, o determinante de uma matriz de ordem 2.

Proposição 4.26 *Se A é uma matriz de ordem 2, então $|\det A|$ é a área do paralelogramo determinado pelos 2 vectores de $\mathbb{R}^2$ que representam as matrizes colunas de A.*

Demonstração Se $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ então os vectores, de $\mathbb{R}^2$ que correspondem às colunas de A, serão

$$u = (a_{11}, a_{21}) \qquad \text{e} \qquad v = (a_{12}, a_{22}).$$

Vamos supor que os vectores não são múltiplos um do outro. Pois caso sejam múltiplos um do outro, então a área do paralelogramo é zero e também $\det A = 0$.



Sabemos que a área do paralelogramo é

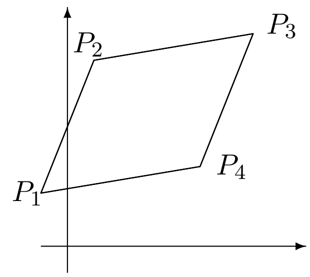
$$\text{área } = \text{base} \times \text{altura} = \|u\|\|v\|sen(\theta)$$

em que θ é o ângulo formado por u e v. Assim,

$$\begin{aligned}
(\text{área})^2 &= \|u\|^2\|v\|^2 sen^2(\theta) = \|u\|^2\|v\|^2(1 - cos^2(\theta)) \\
&= (\|u\|^2\|v\|^2) - (\|u\|^2\|v\|^2 cos^2(\theta)) \\
&= \|u\|^2\|v\|^2 - (u|v)^2 \\
&= (a_{11}^2 + a_{21}^2)(a_{12}^2 + a_{22}^2) - (a_{11}a_{12} + a_{21}a_{22})^2 \\
&= a_{11}^2 a_{22}^2 + a_{21}^2 a_{12}^2 - 2(a_{11}a_{12})(a_{21}a_{22}) \\
&= (a_{11}a_{22} - a_{21}a_{12})^2 \\
&= (\det A)^2.
\end{aligned}$$

Donde, $|\det A|$ =área. □

Exemplo 4.27 *A área do paralelogramo de vértices $P_1 = (-1, 2)$, $P_2 = (1, 7)$, $P_3 = (7, 8)$, $P_4 = (5, 3)$.*



Pela figura vemos que $\overrightarrow{P_1P_2} = (2, 5)$ *e* $\overrightarrow{P_1P_4} = (6, 1)$ *formam dois lados adjacentes do paralelogramo. Assim, pela Proposição anterior,*

$$\text{área} = \left| \begin{vmatrix} 2 & 6 \\ 5 & 1 \end{vmatrix} \right| = |2 - 30| = | - 28| = 28.$$

Porque a área de um triângulo é metade da área do paralelogramo, definido pelos 2 vectores, podemos utilizar a Proposição 4.26, para determinar a área de um triângulo.

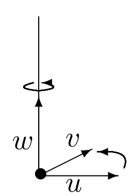
Exemplo 4.28 *A área do triângulo de vértices* $A = (-5, 4)$*,* $B = (3, 2)$*,* $C = (-2, -3)$ *pode ser calculada tendo em conta que* $\overrightarrow{AB} = (8, -2)$ *e* $\overrightarrow{AC} = (3, -7)$ *formam 2 lados do triângulo. Assim, pelo que mencionámos anteriormente,*

$$\text{área} = \frac{1}{2} \left| \begin{vmatrix} 8 & 3 \\ -2 & -7 \end{vmatrix} \right| = \frac{1}{2} | - 56 + 6| = 25.$$

4.8 Produto Externo e Produto Misto de Vectores de $\mathbb{R}^3$

Um problema que surge no estudo de movimentos rotacionais no espaço tridimensional é o de encontrar o eixo de rotação de um objecto que está a girar e identificar se a rotação é horária ou anti-horária, a partir de um ponto específico do eixo de rotação.

Consideremos dois vectores u e v como representado na figura.



Façamos o vector u girar até ser colinear com o vector v. O eixo desta rotação tem de ser perpendicular ao plano definido por u e v, em particular o vector w serve para identificar a orientação do eixo de rotação. Ou seja, o vector que for escolhido para eixo de rotação terá de ser perpendicular a u e a v e conterá a informação sobre o sentido da rotação.

Definição 4.29 *Sejam $u = (u_1, u_2, u_3)$ e $v = (v_1, v_2, v_3)$ vectores de $\mathbb{R}^3$. Designa-se por* **produto externo** (*ou produto vectorial*) *de u por v e denota-se por $u \times v$, o vector de $\mathbb{R}^3$ definido por*

$$u \times v = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1).$$

Como "mnemónica", expressemos o vector $u \times v$ como combinação linear dos vectores canónicos de $\mathbb{R}^3$, $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$.

Porque

$$\begin{aligned} u \times v &= \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right) \\ &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} e_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} e_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} e_3 \\ &= \begin{vmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}. \end{aligned}$$

Exemplo 4.30 *Sendo $u = (1, 0, 2)$, $v = (0, -1, 0)$ então*

$$\begin{aligned} u \times v &= \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & 2 \\ 0 & -1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ -1 & 0 \end{vmatrix} e_1 - \begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} e_2 + \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} e_3 \\ &= 2e_1 - 0e_2 - e_3 = (2, 0, -1) \end{aligned}$$

$$\begin{aligned} v \times u &= \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & -1 & 0 \\ 1 & 0 & 2 \end{vmatrix} = \begin{vmatrix} -1 & 0 \\ 0 & 2 \end{vmatrix} e_1 - \begin{vmatrix} 0 & 0 \\ 1 & 2 \end{vmatrix} e_2 + \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} e_3 \\ &= -2e_1 + 0e_2 + e_3 = (-2, 0, 1) \end{aligned}$$

Proposição 4.31 *Sejam u, v, t vectores de $\mathbb{R}^3$ e α um escalar. Entâo:*

1. $u \times v = -(v \times u)$
2. $(u \times v)|u = (u \times v)|v = 0$
3. $\alpha(u \times v) = (\alpha u) \times v = u \times (\alpha v)$
4. $(u + v) \times t = (u \times t) + (v \times t)$
5. $t \times (u + v) = (t \times u) + (t \times v)$

Teorema 4.32 *Sejam u e v vectores não nulos de $\mathbb{R}^3$ e seja θ o ângulo formado por u e v.*

1. $\|u \times v\| = \|u\|\|v\|sen(\theta)$

2. *a área do paralelogramo de lados adjacentes u e v é*

$$\|u \times v\|.$$

Demonstração

1. Como $0 \leq \theta \leq \pi$ vem que $sen(\theta) = \sqrt{1 - cos^2(\theta)}$. Assim,

$$\begin{aligned}
\|u\|\|v\|sen(\theta) &= \|u\|\|v\|\sqrt{1 - cos^2(\theta)} = \|u\|\|v\|\sqrt{1 - \frac{(u|v)^2}{\|u\|^2\|v\|^2}} \\
&= \sqrt{\|u\|^2\|v\|^2 - (u|v)^2} \\
&= \sqrt{(u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)} \\
&= \sqrt{(u_2v_3 - v_2u_3)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2} \\
&= \|u \times v\|.
\end{aligned}$$

2. A área do paralelogramo de lados adjacentes u e v é

$$\text{área} = base \times altura \underset{\substack{\uparrow \\ \text{como já vimos}}}{=} \|u\|\|v\|sen(\theta) = \|u \times v\|.$$

□

Exemplo 4.33 *A área do triângulo de $\mathbb{R}^3$ de vértices $P_1 = (2, 2, 0)$, $P_2 = (-1, 0, 2)$, $P_3 = (0, 4, 3)$ é metade da área do paralelogramo de lados adjacentes $\overrightarrow{P_1P_2} = (-3, -2, 2)$ e $\overrightarrow{P_1P_3} = (-2, 2, 3)$.*

Como

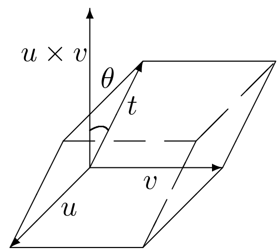
$$\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3} = \begin{vmatrix} e_1 & e_2 & e_3 \\ -3 & -2 & 2 \\ -2 & 2 & 3 \end{vmatrix} = -10e_1 + 5e_2 - 10e_3 = (-10, 5, -10)$$

vem que a área do triângulo é

$$\frac{1}{2}\|\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}\| = \frac{1}{2}\sqrt{100 + 25 + 100} = \frac{15}{2}.$$

Dados 3 vectores u, v, t de $\mathbb{R}^3$, que têm o mesmo ponto inicial, eles definem um paralelepípedo cujo volume é dado pelo produto da área da sua base (paralelogramo definido por u e v) pela sua altura h. Já vimos que $\|u \times v\|$ nos dá a área da base (Teorema 4.32) e que $u \times v$ é perpendicular a u e a v (Proposição 4.31). Sendo θ o ângulo formado pelos vectores $u \times v$ e t, a altura do paralelepípedo é

$$h = \|t\||cos(\theta)|.$$



Assim, o volume do paralelepípedo é

$$\|u \times v\|\|t\||cos(\theta)| = |\|u \times v\|\|t\|cos(\theta)| = |(u \times v)|t|.$$

Definição 4.34 *Dados 3 vectores u, v, t de $\mathbb{R}^3$, chamamos* **produto misto** *dos vectores u, v e t ao número real*

$$(u \times v)|t.$$

Como "mnemónica", se $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ e $t = (t_1, t_2, t_3)$ então

$$\begin{aligned}(u \times v)|t &= \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right) |(t_1, t_2, t_3) \\ &= t_1 \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} - t_2 \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} + t_3 \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \\ &= \begin{vmatrix} t_1 & t_2 & t_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.\end{aligned}$$

Exemplo 4.35 *Determine k por forma a que o paralelepípedo de lados $u = (2, k, 2)$, $v = (0, 4, -2)$, $t = (5, -4, 0)$ tenha volume igual a 4.*

Porque o volume é

$$\begin{aligned}|(u \times v)|t| &= | \begin{vmatrix} 5 & -4 & 0 \\ 2 & k & 2 \\ 0 & 4 & -2 \end{vmatrix} | \\ &= | - 10k - 40 - 16| = | - 10k - 56| = 4,\end{aligned}$$

então,

$$-(-10k - 56) = 4 \qquad ou \qquad (-10k - 56) = 4.$$

Donde,

$$k = -5.2 \qquad ou \qquad k = -6.$$

4.9 Valores e Vectores Próprios de Matrizes

Nesta secção vamos estudar o seguinte problema:

"Se A é uma matriz de ordem n, para que valores do escalar λ, se houver, existem matrizes não nulas, X, de tipo $n \times 1$ (matrizes coluna) tais que $AX = \lambda X$?"

Definição 4.36 *Sejam A uma matriz de ordem n e λ um escalar:*

1. *λ diz-se um* **valor próprio** *de A, se existir uma matriz coluna X, $n \times 1$, não nula, tal que $AX = \lambda X$.*

2. *Se λ é um valor próprio de A, cada matriz coluna, não nula, X, tal que $AX = \lambda X$ é denominada* **vector próprio** *de A associado a λ.*

Exemplo 4.37 *Sendo $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$, temos que*

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

Pelo que, 2 é valor próprio de A e $\begin{bmatrix} 4 \\ -4 \end{bmatrix}$ é vector próprio de A associado ao 2.

A forma mais fácil de determinar os valores próprios de uma matriz é pensar que a equação

$$AX = \lambda X$$

é equivalente a

$$(\lambda I_n - A)X = 0 \qquad \text{(sistema homogéneo)}.$$

Sabemos que um sistema homogéneo, se for possível e indeterminado, tem soluções diferentes da solução nula, ou seja, se a matriz simples do sistema não for invertível. Mas isto significa que

$$\det(\lambda I_n - A) = 0.$$

Definição 4.38 *Seja A uma matriz de ordem n. Chama-se* **polinómio característico** *de A ao polinómio*

$$p_A(\lambda) = \det(\lambda I_n - A).$$

Usando o que acabámos de mencionar, podemos concluir o seguinte Teorema.

Teorema 4.39 *Sejam A uma matriz de ordem n e λ um escalar. As seguintes afirmações são equivalentes:*

1. *λ é valor próprio de A.*
2. $\det(\lambda I_n - A) = 0$.
3. *λ é raíz do polinómio característico de A.*
4. *O sistema homogéneo $(\lambda I_n - A)X = 0$ é possível e indeterminado.*

Exemplo 4.40 *Sendo $A = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$, vejamos quais os seus valores próprios.*

O polinómio característico de A é

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_2 - A) = \begin{vmatrix} \lambda - 1 & -1 \\ 2 & \lambda + 2 \end{vmatrix} = (\lambda - 1)(\lambda + 2) + 2 \\ &= \lambda^2 + \lambda - 2 + 2 = \lambda^2 + \lambda = \lambda(\lambda + 1) \end{aligned}$$

As raízes do polinómio característico são

$$\lambda = 0 \qquad\qquad \lambda = -1$$

que pelo Teorema 4.39 *são os valores próprios de A.*

Definição 4.41 *Sejam A uma matriz de ordem n e α um valor próprio de A, designamos por* **multiplicidade algébrica** *do valor próprio α e denotamos por*

$$ma(\alpha),$$

o número de vezes que α aparece como raíz de $p_A(\lambda)$ (polinómio característico de A).

Como um polinómio de grau n, com coeficientes reais tem n raízes (pode acontecer algumas destas raízes serem números complexos que não são números reais), então sendo $\alpha_1, \alpha_2, \ldots, \alpha_s$ os valores próprios de A, dois a dois distintos,

$$ma(\alpha_1) + ma(\alpha_2) + \ldots + ma(\alpha_s) = \text{ordem de } A = n.$$

Exemplo 4.42 *Sendo $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, o polinómio característico de A é*

$$p_A(\lambda) = \lambda^2 + 1,$$

cujas raízes são

$$\lambda = i \qquad e \qquad \lambda = -i.$$

Ou seja,

$$p_A(\lambda) = (\lambda + i)(\lambda - i).$$

Os valores próprios de A verificam

$$ma(i) = 1 \qquad e \qquad ma(-i) = 1$$

e

$$ma(i) + ma(-i) = 2 = ordem\ de\ A.$$

Definição 4.43 *Sejam A uma matriz de ordem n e α um valor próprio de A. Chamamos* **subespaço próprio** *de A associado ao valor próprio α ao conjunto solução do sistema homogéneo*

$$(\alpha I_n - A)X = 0.$$

Este conjunto é denotado por M_α.

Observações 4.44 *Como já vimos* (Exemplo 4.42) *os valores próprios de uma matriz com todas as entradas reais, podem ser números complexos que não são números reais.*

1. *Se o valor próprio α de A for um número real, então as entradas da matriz $(\alpha I_n - A)$ são todas números reais. Assim, como já vimos, o conjunto-solução do sistema homogéneo $(\alpha I_n - A)X = 0$ é um subespaço.*
2. *Se o valor próprio α de A for um número complexo que não é um número real, então as entradas da matriz $(\alpha I_n - A)$ não serão todas números reais. Assim, ainda não sabemos se o conjunto-solução do sistema homogéneo $(\alpha I_n - A)X = 0$ terá algum nome especial.*
3. *No caso anterior, temos ainda o problema de resolver o sistema homogéneo $(\alpha I_n - A)X = 0$, para encontrarmos o subespaço próprio da matriz, em que certas entradas da matriz simples do sistema são números complexos que não são números reais. Mas o processo de resolução deste sistema é em tudo análogo ao dado para sistemas em que todas as entradas são números reais.*

Exemplo 4.45 *Com a matriz do* Exemplo 4.42, $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ *que tem os valores próprios*

$$i \qquad e \qquad -i,$$

calculemos M_i.

Ora sendo $\lambda = i$ vem

$$(iI_2 - A)X = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Como

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \underset{l_1 \to -il_1}{\longrightarrow} \begin{bmatrix} 1 & i \\ 1 & i \end{bmatrix} \underset{l_2 \to (l_2 - l_1)}{\longrightarrow} \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

temos,

$$x_1 + ix_2 = 0,$$

ou seja, $x_1 = -ix_2$, pelo que

$$M_i = \{(-ix_2, x_2) : \ x_2 \in \mathbb{R}\}.$$

Exemplo 4.46 *A matriz $A = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$* (Exemplo 4.40) *tem os valores próprios*

$$0 \qquad\qquad e \qquad\qquad -1,$$

cada um com multiplicidade algébrica 1.

Determinemos os subespaços M_0 e M_{-1}.

Ora sendo $\lambda = 0$ vem

$$M_0 = \left\{(x_1, x_2) \in \mathbb{R}^2 : \ (0I_2 - A) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}.$$

Porque

$$(0I_2 - A)X = \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

vem que

$$\begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix} \underset{l_2 \to (l_2 + 2l_1)}{\longrightarrow} \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix} \underset{l_1 \to -l_1}{\longrightarrow} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

e assim,

$$x_1 + x_2 = 0,$$

ou seja, $x_1 = -x_2$, pelo que

$$M_0 = \{(x_1, x_2) \in \mathbb{R}^2 : \ x_1 = -x_2\} = \{(-x_2, x_2) : \ x_2 \in \mathbb{R}\} =< (-1, 1) > .$$

Para $\lambda = -1$ temos

$$M_{-1} = \left\{(x_1, x_2) \in \mathbb{R}^2 : \ (-I_2 - A) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}.$$

Porque

$$(-I_2 - A)X = \begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

vem que

$$\begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix} \underset{l_2 \to (l_2+l_1)}{\longrightarrow} \begin{bmatrix} -2 & -1 \\ 0 & 0 \end{bmatrix} \underset{l_1 \to -\frac{1}{2}l_1}{\longrightarrow} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}.$$

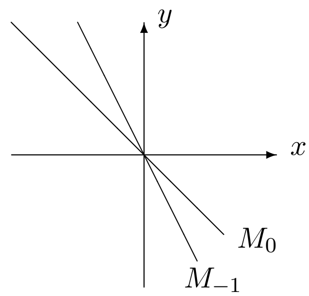
Donde,

$$x_1 + \frac{1}{2}x_2 = 0,$$

ou seja, $x_1 = -\frac{1}{2}x_2$, *e*

$$M_{-1} = \left\{ (x_1, x_2) \in \mathbb{R}^2 : \ x_1 = -\frac{1}{2}x_2 \right\} = \left\{ \left(-\frac{1}{2}x_2, x_2 \right) : \ x_2 \in \mathbb{R} \right\} = \left\langle \left(-\frac{1}{2}, 1 \right) \right\rangle.$$

Repare que qualquer um destes subespaços define uma recta que passa pela origem.



Capítulo 5

APLICAÇÕES LINEARES

Dada uma variável y que depende de uma variável x de tal modo que, para cada valor de x temos um único valor de y, dizemos que y é uma função ou **aplicação** de x.

As aplicações serão denotadas pelas letras $f,\ g,\ h, \ldots$, com ou sem índices.

Assim, supondo que uma aplicação f é um programa de computador que para cada entrada (input) x, produz uma única saída (output) y, temos o esquema

Dependendo da aplicação, as entradas e/ou saídas podem ser números, vectores, matrizes,...

5.1 Definições e Notações

O conjunto das entradas de uma aplicação f chama-se **domínio** de f.

Dado um elemento do domínio de f, denotamos por $f(x)$ a saída correspondente, e chamamos **imagem de** x **por** f.

O conjunto de todas as saídas dos elementos do domínio de f chama-se **imagem de** f ou contradomínio de f e denota-se por $Im\ f$.

Iremos estudar aplicações f cujo domínio é $\mathbb{R}^n$ e cuja imagem é um subconjunto de $\mathbb{R}^m$.

Neste caso escreveremos

$$\begin{aligned} f:\ \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ X &\mapsto Y \end{aligned}$$

para dizer que, a cada elemento X de $\mathbb{R}^n$, pela aplicação f, corresponde a imagem Y de $\mathbb{R}^m$.

O conjunto $\mathbb{R}^m$ é designado por **conjunto de chegada** de f.

Exemplo 5.1 1. *Seja*

$$\begin{aligned} f:\ \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x,y) &\mapsto (x+y, x^2, y^2) \end{aligned}$$

Como, para cada elemento (x,y) de $\mathbb{R}^2$, existe um único elemento de $\mathbb{R}^3$, que é $(x+y, x^2, y^2)$, tal que

$$f(x,y) = (x+y, x^2, y^2),$$

então f é uma aplicação.

O elemento $(1,0,0)$ pertence a $\mathbb{R}^3$ (conjunto de chegada de f) mas não pertence ao contradomínio de f pois não existe nenhum (x,y) em $\mathbb{R}^2$ tal que

$$f(x,y) = (1,0,0)$$

ou seja, tal que

$$x+y=1, \qquad x^2=0, \qquad y^3=0.$$

2. *Seja*

$$\begin{aligned} f:\ \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto z \end{aligned}$$

tal que $z^2 = x$.

Porque $f(1,0) = \pm 1$, então não existe um único elemento de $\mathbb{R}$ que é imagem de $(1,0)$ por f, portanto, f não é aplicação.

5.2 Definição de Aplicação Linear

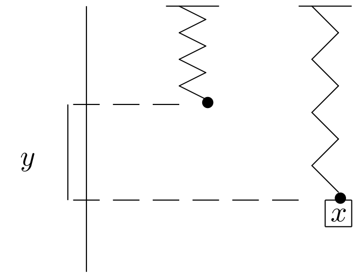
Dizemos que uma variável y é directaments proporcional a uma variável x, se existe uma constante k tal que

$$y = kx.$$

Um exemplo deste conceito é a **Lei de Hooke** que diz:

"Um peso de x unidades suspenso de uma mola, com rigidez k, alonga a mola, a partir do seu comprimento natural, por uma quantidade y que é directamente proporcional a x, isto é, $y = kx$."

Esquematicamente,



Escrevendo $f(x) = kx$, reparamos que esta equação verifica 2 propriedades:

1. Se colocarmos um peso múltiplo de x, αx, suspenso da mola com a rigidez k, então o alongamento da mola será α vezes o alongamento que sofreu a mola com o peso x, isto é,
$$f(\alpha x) = k(\alpha x) = \alpha(kx) = \alpha f(x).$$

2. Se colocarmos dois pesos x_1 e x_2 suspensos da mola, o alongamento que provocam na mola é a soma dos alongamentos que provoca cada peso, isto é,
$$f(x_1 + x_2) = k(x_1 + x_2) = kx_1 + kx_2 = f(x_1) + f(x_2).$$

Este exemplo leva-nos à definição seguinte:

Definição 5.2 *Seja $f :\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação. Dizemos que f é* **aplicação linear** *se*

1. $\forall x \in \mathbb{R}^n,\ \forall \alpha \in \mathbb{R},\ f(\alpha x) = \alpha f(x),$
2. $\forall x_1, x_2 \in \mathbb{R}^n,\ f(x_1 + x_2) = f(x_1) + f(x_2).$

Observação As duas propriedades anteriores podem escrever-se numa única propriedade
$$\forall x_1, x_2 \in \mathbb{R}^n,\ \forall \alpha_1, \alpha_2 \in \mathbb{R},\ f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2).$$

Exemplo 5.3 1. *Sendo*
$$\begin{aligned} f :\ \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\mapsto (x + y, x^2, y^2) \end{aligned}$$
a aplicação do Exemplo 5.1, 1.*, temos*
$$f(3(1, 0)) = f(3, 0) = (3, 9, 0)$$
e
$$3f(1, 0) = 3(1, 1, 0) = (3, 3, 0).$$

Portanto, $f(3(1,0)) \neq 3f(1,0)$. *Logo não se verifica a propriedade* 1. *da* Definição 5.2, *ou seja,*

$$f(\alpha(x,y)) \neq \alpha f(x,y)$$

para algum $\alpha \in \mathbb{R}$ *e* $(x,y) \in \mathbb{R}^2$ *e portanto,* f *não é aplicação linear.*

2. *A aplicação*

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

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

é aplicação linear porque

i) $\forall (x,y)\mathbb{R}^2, \ \forall \alpha \in \mathbb{R}$,

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

ii) $\forall (x,y),(z,w)\mathbb{R}^2$,

$$f((x,y)+(z,w)) = f(x+z,y+w) = (y+w,x+z) = (y,x)+(w,z) = f(x,y)+f(z,w).$$

Ou seja, verificam-se as propriedades 1. *e* 2. *da* Definição 5.2.

Esta aplicação é uma **reflexão** *através da recta* $y = x$ *de* $\mathbb{R}^2$.

Teorema 5.4 *Se* $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *é uma aplicação linear, então*

1. $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.

2. $f(-x) = -f(x), \ \forall x \in \mathbb{R}^n$.

Observação Repare-se que em 1. do Teorema 5.4, $0_{\mathbb{R}^n}$ é o vector nulo de $\mathbb{R}^n$ e $0_{\mathbb{R}^m}$ é o vector nulo de $\mathbb{R}^m$.

Demonstração

1. Atendendo a que

$$0.f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m},$$

a que

$$f(0.0_{\mathbb{R}^n}) = f(0_{\mathbb{R}^n})$$

e a que f é aplicação linear, temos

$$f(0_{\mathbb{R}^n}) = f(0.0_{\mathbb{R}^n}) = 0f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}.$$

2. Porque
$$\forall x \in \mathbb{R}^n,\ x - x = 0_{\mathbb{R}^n},$$
de 1. e da linearidade de f, temos
$$0_{\mathbb{R}^m} = f(0_{\mathbb{R}^n}) = f(x + (-x)) = f(x) + f(-x).$$
Então, $f(-x) = -f(x)$. $\square$

Observação Muitas vezes teremos de considerar aplicações nas quais o domínio D é um subespaço de $\mathbb{R}^n$, em vez de todo o $\mathbb{R}^n$. Exactamente, como na definição de linearidade que demos, dizemos que uma aplicação $f:\ D \longrightarrow \mathbb{R}^m$ é linear se satisfaz as propriedades 1. e 2.. O Teorema anterior também é válido para estas aplicações lineares.

5.3 Matriz Canónica de uma Aplicação Linear

Sejam $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear e $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$,..., $e_n = (0, 0, \ldots, 1)$ os vectores canónicos de $\mathbb{R}^n$.

Sabemos que cada vector de $\mathbb{R}^n$, $(x_1, x_2, \ldots, x_n)$, pode escrever-se como combinação linear dos vectores $e_1, e_2, \ldots, e_n$ da seguinta forma,
$$(x_1, x_2, \ldots, x_n) = x_1e_1 + x_2e_2 + \ldots + x_ne_n.$$
Atendendo à linearidade de f,
$$f(x_1, x_2, \ldots, x_n) = x_1f(e_1) + x_2f(e_2) + \ldots + x_nf(e_n), \qquad (*)$$
ou seja, a imagem de qualquer vector de $\mathbb{R}^n$, por f, pode escrever-se como combinação linear dos vectores $f(e_1), f(e_2), \ldots, f(e_n)$.

Definição 5.5 *Sejam $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear e $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, ..., $e_n = (0, 0, \ldots, 1)$ os vectores canónicos de $\mathbb{R}^n$. Chamamos* **matriz canónica** *de f, e denotamo-la por M_f, a matriz cujas colunas são $f(e_1), f(e_2), \ldots, f(e_n)$.*

Exemplo 5.6 1. *Considerando a reflexão através da recta $y = x$ de $\mathbb{R}^2$* (Exemplo 5.3, 2.,
$$f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x, y) \ \mapsto\ (y, x)$$
porque
$$f(e_1) = f(1, 0) = (0, 1),$$

$$f(e_2) = f(0,1) = (1,0),$$

então

$$M_f = \begin{array}{c} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \begin{array}{cc} \uparrow & \uparrow \\ f(e_1) & f(e_2) \end{array} \end{array}$$

2. *Seja*

$$f: \ \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

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

uma aplicação linear. Então

$$f(1,0,0) = (1,0,1),$$

$$f(0,1,0) = (1,0,0),$$

$$f(0,0,1) = (0,-1,1),$$

pelo que,

$$M_f = \begin{array}{c} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix} \\ \begin{array}{ccc} \nearrow & \uparrow & \nwarrow \\ f(e_1) & f(e_2) & f(e_3) \end{array} \end{array}$$

Como já o dissemos (Capítulo 2), os vectores de $\mathbb{R}^n$ e de $\mathbb{R}^m$ podem ser representados por matrizes colunas ($n \times 1$, no primeiro caso e $m \times 1$ no segundo caso). Visto assim, $(*)$ será

$$x_1 f(e_1) + x_2 f(e_2) + \ldots + x_n f(e_n) = M_f \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Se X representar a matriz coluna $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, então a matriz coluna Y, cujo vector coluna é $f(x_1, x_2, \ldots, x_n)$ é

$$Y = M_f X \qquad\qquad (**)$$

Isto significa que se nos derem a matriz canónica, M_f, de uma aplicação linear, $f: \ \mathbb{R}^n \longrightarrow \mathbb{R}^m$, através da igualdade $(**)$ podemos obter:

1. *a imagem de um vector de* $\mathbb{R}^n$ *por* f

   Basta construir a matriz coluna X desse vector e através do produto

   $$M_f X = Y$$

   obtemos a matriz coluna Y, cujo vector de $\mathbb{R}^m$ corresponde à imagem do vector inicial por f.

2. *a expressão da aplicação* f

   Usamos o processo descrito em 1., sendo a matriz coluna X, uma matriz de variáveis.

3. *informação sobre se um dado vector de* $\mathbb{R}^m$ *pertence à* $Im\ f$

   Basta construir a matriz coluna Y desse vector e discutir e/ou resolver o sistema $M_f X = Y$.

Exemplo 5.7 *Seja* $f:\ \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ *a aplicação linear tal que*

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

A imagem do vector $(1,1,0)$ *por* f *é,*

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

o vector $f(1,1,0) = (2,0)$.

A expressão da aplicação f *é,*

$$M_f \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y+z \\ -x+y+z \end{bmatrix},$$

dada por

$$\begin{aligned} f:\ \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x,y,z) &\mapsto (2y+z, -x+y+z). \end{aligned}$$

O vector $(1,1) \in Im\ f$ *porque o sistema*

$$M_f X = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

ou seja,

$$\begin{bmatrix} 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$[A|B] = \left[\begin{array}{ccc|c} 0 & 2 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{array}\right] \underset{l_1 \leftrightarrow l_2}{\longrightarrow} \left[\begin{array}{ccc|c} -1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{array}\right],$$

é tal que $r(A) = 2 = r([A|B]) < 3 =$ número de incógnitas. Portanto, o sistema é possível e indeterminado. Uma solução do sistema é,

$$\left[\begin{array}{ccc|c} -1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{array}\right] \underset{l_1 \to -l_1}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & -1 & -1 & -1 \\ 0 & 2 & 1 & 1 \end{array}\right] \underset{l_2 \to \frac{1}{2}l_2}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & -1 & -1 & -1 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right] \underset{l_1 \to (l_1 + l_2)}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right]$$

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

$x = 0$, $y = 0$, $z = 1$. Ou seja, o vector $(0, 0, 1)$ de $\mathbb{R}^3$ é tal que $f(0, 0, 1) = (1, 1)$.

Observação É fácil mostrar que se $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ é uma aplicação tal que $f(x_1, \ldots, x_n) = a_1x_1 + \ldots + a_nx_n$, com $(x_1, \ldots x_n) \in \mathbb{R}^n$ (ou seja, $a_1x_1 + \ldots + a_nx_n = 0$ é uma equação linear), então f é uma aplicação linear.

Por outro lado se $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma aplicação linear cuja matriz canónica é $M_f = [a_{ij}]$, então, $f(x_1, \ldots, x_n)$, com $(x_1, \ldots x_n) \in \mathbb{R}^n$, é a matriz coluna

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = M_f \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix}.$$

Daqui vem (igualdade de matrizes) que

$$\begin{array}{l} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{array},$$

e cada

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0,\ i = 1, \ldots, m$$

é uma equação linear.

Portanto, uma aplicação

$$\begin{aligned} f :\ & \mathbb{R}^n \longrightarrow \mathbb{R}^m \\ (x_1, \ldots, x_n) & \mapsto f(x_1, \ldots, x_n) = (y_1, \ldots, y_m) \end{aligned}$$

é linear se, e só se, cada $y_i = 0$ é uma equação linear, com $i = 1, \ldots, m$.

Exemplo 5.8 1. *A aplicação*

$$\begin{aligned} f :\ & \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \\ (x_1, x_2) & \mapsto (x_1x_2, x_2, 0) \end{aligned}$$

não é aplicação linear, porque sendo

$$(y_1, y_2, y_3) = f(x_1, x_2) = (x_1x_2, x_2, 0),$$

então $y_1 = x_1x_2$ *e* $x_1x_2 = 0$ *não é uma equação linear.*

2. *A aplicação*

$$\begin{aligned} f: \ \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ (x_1, x_2, x_3) &\mapsto (x_1 + x_2, x_3) \end{aligned}$$

é aplicação linear, porque se

$$(y_1, y_2) = f(x_1, x_2, x_3) = (x_1 + x_2, x_3),$$

então $y_1 = x_1 + x_2$ *e* $y_2 = x_3$ *em que as equações*

$$x_1 + x_2 = 0 \qquad e \qquad x_3 = 0$$

são equações lineares.

3. *A aplicação*

$$\begin{aligned} f: \ \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x_1, x_2, x_3) &\mapsto (x_1 + x_2, x_3 - 1, x_1) \end{aligned}$$

não é aplicação linear, porque sendo

$$(y_1, y_2, y_3) = f(x_1, x_2, x_3) = (x_1 + x_2, x_3 - 1, x_1),$$

então $y_2 = x_3 - 1$ *e a equação*

$$x_3 - 1 = 0$$

não é uma equação linear.

Proposição 5.9 *Seja* A *uma matriz de tipo* $m \times n$*. Então, existe uma, e uma só, aplicação linear* $f: \ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *cuja matriz canónica,* M_f*, é* A*.*

Demonstração Seja $A = [a_{ij}]$. Então, se $x_1, \ldots, x_n \in \mathbb{R}$,

$$A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Porque o produto de duas matrizes dá uma única matriz, podemos falar na aplicação

$$\begin{aligned} f: \ \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x_1, \ldots, x_n) &\mapsto (y_1, \ldots, y_m) \end{aligned}$$

em que $y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n,\ i = 1, \ldots, m$.

Pela observação anterior, porque cada $y_i = 0$ é uma equação linear, $i = 1, \ldots, m$, então f é aplicação linear. Por construção, $M_f = A$.

Suponhamos que existem duas aplicações lineares

$$f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m \qquad \text{e} \qquad g:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

tais que $M_f = A$ e $M_g = A$. Assim sendo, se $(x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$[f(x_1, x_2, \ldots, x_n)] = M_f \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = M_g \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [g(x_1, x_2, \ldots, x_n)].$$

Donde, $f = g$ e a aplicação linear é única. $\square$

5.4 Exemplos de Aplicações Lineares de $\mathbb{R}^2$ em $\mathbb{R}^2$

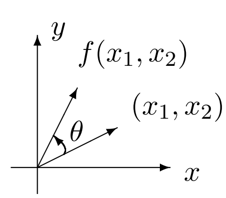
Nesta secção iremos estudar algumas aplicações lineares de $\mathbb{R}^2$ em $\mathbb{R}^2$.

5.4.1 Rotações em torno da origem

Seja θ um ângulo. Consideremos a aplicação

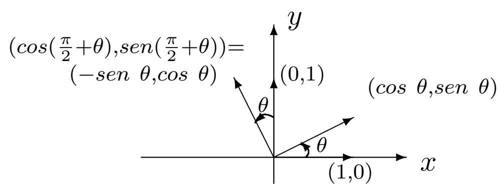
$$f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

que faz girar cada vector (x_1, x_2) de $\mathbb{R}^2$, em torno da origem, segundo um ângulo de amplitude θ, no sentido anti-horário



É fácil verificar, usando a soma de vectores (regra do paralelogramo), que a aplicação f é linear. Denotemos por R_θ a matriz canónica desta aplicação f.

Como $e_1 = (1, 0)$ e $e_2 = (0, 1)$



vem que $f(e_1) = (cos\ \theta, sen\ \theta)$, $f(e_2) = (-sen\ \theta, cos\ \theta)$.
Donde,

$$R_\theta = \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix}.$$

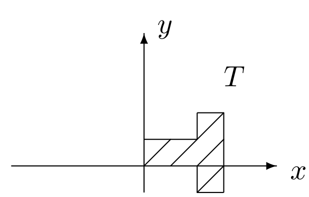
Exemplo 5.10 *Se $\theta = \frac{\pi}{2}$, então a matriz canónica da rotação com ângulo $\frac{\pi}{2}$, de $\mathbb{R}^2$ em $\mathbb{R}^2$, é*

$$R_{\frac{\pi}{2}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

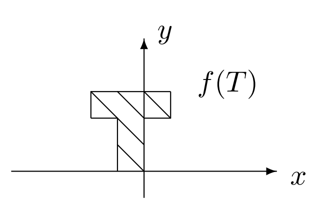
e

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x_1, x_2) \mapsto (-x_2, x_1)$$

A rotação de $\frac{\pi}{2}$ da figura T



é a figura

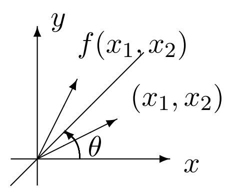


5.4.2 Reflexões através de rectas que passam pela origem

Consideremos a aplicação

$$f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

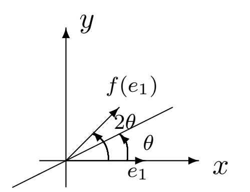
que reflecte cada vector de $\mathbb{R}^2$, através de uma recta que passa pela origem e faz um ângulo de amplitude θ com o eixo dos xx positivo.

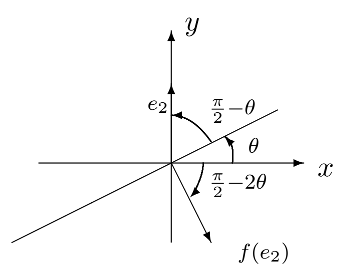


Facilmente se prova que f é aplicação linear. Se designarmos por H_θ a matriz canónica desta aplicação, vem que

$$f(e_1) = f(1,0) = (cos\ 2\theta, sen\ 2\theta)$$

$$f(e_2) = f(0,1) = \left(cos\left(\frac{\pi}{2} - 2\theta\right), -sen\left(\frac{\pi}{2} - 2\theta\right)\right) = (sen\ 2\theta, -cos\ 2\theta)$$





pelo que

$$H_\theta = \begin{bmatrix} cos\ 2\theta & sen\ 2\theta \\ sen\ 2\theta & -cos\ 2\theta \end{bmatrix}.$$

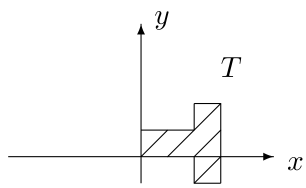
Exemplo 5.11 *Se for a reflexão no eixo dos yy, então* $\theta = \frac{\pi}{2}$,

$$H_{\frac{\pi}{2}} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

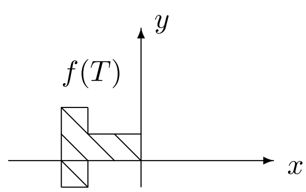
e

$$\begin{aligned} f:\ \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x_1, x_2) &\mapsto (-x_1, x_2) \end{aligned}$$

A reflexão no eixo dos yy da figura T



é a figura



5.4.3 Compressões e Expansões Horizontais e Verticais em $\mathbb{R}^2$

Sendo

$$f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

a aplicação linear definida por

$$f(x_1, x_2) = (kx_1, x_2)$$

com $k \in \mathbb{R}_0^+ \setminus \{1\}$, dizemos que,

- f é uma compressão na direcção de x, de razão k, no plano xy, se $0 \leq k < 1$.
- f é uma expansão na direcção de x, de razão k, no plano xy, se $k > 1$.

Estas são as compressões e as expansões horizontais.

Se

$$f(x_1, x_2) = (x_1, kx_2)$$

com $k \in \mathbb{R}_0^+ \setminus \{1\}$, dizemos que,

- f é uma compressão na direcção de y, de razão k, no plano xy, se $0 \leq k < 1$.
- f é uma expansão na direcção de y, de razão k, no plano xy, se $k > 1$.

Estas são as compressões e as expansões verticais.

A matriz canónica de f é

- $$M_f = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix},$$
  se f é uma compressão ou uma expansão horizontal.
- $$M_f = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix},$$
  se f é uma compressão ou uma expansão vertical.

Observação No caso de $k = 1$, teriamos a aplicação linear

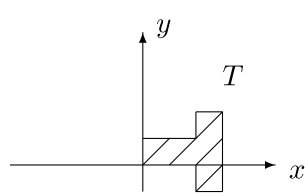
$$f(x_1, x_2) = (x_1, x_2)$$

que tem como matriz canónica

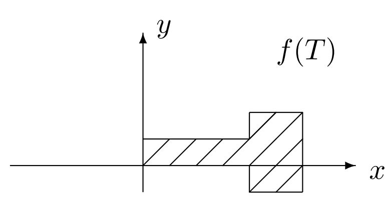
$$M_f = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

ou seja, estamos perante uma rotação em torno da origem com um ângulo de amplitude 0.

Exemplo 5.12 *Sendo $f(x_1, x_2) = (2x_1, x_2)$, então f é uma expansão horizontal de razão 2 e a imagem de T*



é a figura



5.4.4 Alongamentos

Sendo

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

a aplicação linear definida por

$$f(x_1, x_2) = (x_1 + kx_2, x_2)$$

com $k \in \mathbb{R}$, esta aplicação translada um ponto (x_1, x_2), do plano xy, paralelamente ao eixo dos xx, por uma quantidade kx_2. Dizemos que, f é um alongamento na direcção de x, de razão k.

Se

$$f(x_1, x_2) = (x_1, kx_1 + x_2)$$

com $k \in \mathbb{R}$, esta aplicação translada um ponto (x_1, x_2), do plano xy, paralelamente ao eixo dos yy, por uma quantidade kx_1. Dizemos que, f é um alongamento na direcção de y, de razão k.

A matriz canónica de f é

- $$M_f = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix},$$

  se f é um alongamento na direcção de x.

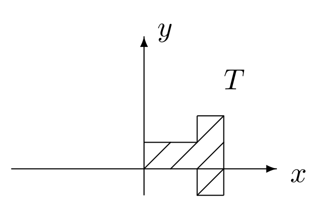
- $$M_f = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix},$$

  se f é um alongamento na direcção de y.

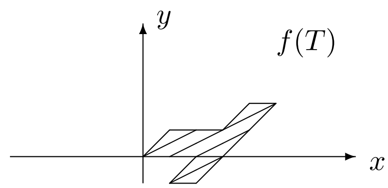
Exemplo 5.13 *Sendo $f(x_1, x_2) = (x_1 + x_2, x_2)$, então f é um alongamento na direcção de x e*

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

A imagem de T



através de f é a figura



5.5 Núcleo e Imagem de uma Aplicação Linear

Seja $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear. Já vimos que $Im\ f$ é o conjunto das imagens, por f, dos elementos de $\mathbb{R}^n$, isto é,

$$Im\ f = \{f(x_1, \ldots, x_n) :\ (x_1, \ldots, x_n) \in \mathbb{R}^n\}.$$

No caso de S ser um subconjunto de $\mathbb{R}^n$, define-se

$$f(S) = \{f(x_1, \ldots, x_n) :\ (x_1, \ldots, x_n) \in S\}.$$

Usando esta última definição,

$$f(\mathbb{R}^n) = Im\ f.$$

Assim, se

$$(x_1, x_2, \ldots, x_n) = t(v_1, v_2, \ldots, v_n),$$

em que $t \in \mathbb{R}$ e $(v_1, v_2, \ldots, v_n)$ é um elemento não nulo de $\mathbb{R}^n$, é uma recta de $\mathbb{R}^n$, a sua imagem por f é o conjunto dos elementos de $\mathbb{R}^m$ que verificam

$$f(x_1, x_2, \ldots, x_n) = f(t(v_1, v_2, \ldots, v_n)) = tf(v_1, v_2, \ldots, v_n), \qquad t \in \mathbb{R}.$$

Ou seja, a imagem por f de uma recta de $\mathbb{R}^n$ é

- o vector nulo de $\mathbb{R}^m$, se $f(v_1, v_2, \ldots, v_n) = 0_{\mathbb{R}^m}$.
- uma recta de $\mathbb{R}^m$, se $f(v_1, v_2, \ldots, v_n) \neq 0_{\mathbb{R}^m}$.

Temos pois que distinguir estas duas situações.

Definição 5.14 *Seja $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear.*

Chama-se **núcleo** *de f e denota-se por $Nuc\ f$ ou $Ker\ f$, o conjunto dos elementos de $\mathbb{R}^n$ que têm por imagem, através de f, o vector nulo de $\mathbb{R}^m$, isto é*

$$Nuc\ f = \{(x_1, \ldots, x_n) \in \mathbb{R}^n :\ f(x_1, \ldots, x_n) = 0_{\mathbb{R}^m}\}.$$

Observação Sendo $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear, repare-se que os conjuntos

$$Nuc\ f = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \ f(x_1, \ldots, x_n) = 0_{\mathbb{R}^m}\}$$

e

$$Im\ f = \{f(x_1, \ldots, x_n) : \ (x_1, \ldots, x_n) \in \mathbb{R}^n\}$$

são distintos.

Enquanto que $Im\ f$ é um subconjunto de $\mathbb{R}^m$, $Nuc\ f$ é um subconjunto de $\mathbb{R}^n$.

Mesmo no caso em que $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ é uma aplicação linear, $Nuc\ f$ e $Im\ f$ são distintos pois o $Nuc\ f$ é constituido pelos elementos de $\mathbb{R}^n$ que são transformados por f no zero, enquanto que $Im\ f$ é constituido pelas imagens dos elementos de $\mathbb{R}^n$ por f.

Exemplo 5.15 *Consideremos a aplicação*

$$\begin{aligned} f: \ \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x_1, x_2) &\mapsto (x_1 - x_2, -2x_1 + 2x_2) \end{aligned}$$

Atendendo à aplicação, facilmente se demonstra que f é aplicação linear. Por definição,

$$Nuc\ f = \{(x_1, x_2) \in \mathbb{R}^2 : \ f(x_1, x_2) = (0,0)\}.$$

Porque $f(x_1, x_2) = (x_1 - x_2, -2x_1 + 2x_2)$, de $f(x_1, x_2) = (0,0)$ resulta o sistema homogéneo

$$\begin{cases} x_1 - x_2 = 0 \\ -2x_1 + 2x_2 = 0 \end{cases}$$

Resolvendo, obtemos $x_1 - x_2 = 0$, pelo que

$$\begin{aligned} Nuc\ f &= \{(x_1, x_2) \in \mathbb{R}^2 : \ x_1 - x_2 = 0\} \\ &= \{(x_1, x_2) \in \mathbb{R}^2 : \ x_1 = x_2\} \\ &= \{(x_1, x_1) : \ x_1 \in \mathbb{R}\} \\ &= \{x_1(1,1) : \ x_1 \in \mathbb{R}\} \\ &= < (1,1) > . \end{aligned}$$

Por outro lado,

$$\begin{aligned} Im\ f &= \{f(x_1, x_2) \in \mathbb{R}^2 : \ (x_1, x_2) \in \mathbb{R}^2\} \\ &= \{(x_1 - x_2, -2x_1 + 2x_2) : \ x_1, x_2 \in \mathbb{R}\} \\ &= \{(x_1, -2x_1) + (-x_2, 2x_2) : \ x_1, x_2 \in \mathbb{R}\} \\ &= \{x_1(1,-2) + x_2(-1,2) : \ x_1, x_2 \in \mathbb{R}\} \\ &= < (1,-2), (-1,2) > . \end{aligned}$$

Observação Se $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma aplicação linear e se M_f é a matriz canónica de f, então o núcleo de f é o conjunto solução do sistema homogéneo

$$M_f X = 0.$$

Usando esta Observação e a Proposição 3.17, temos o seguinte resultado.

Proposição 5.16 *Seja $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear. Então o núcleo de f é um subespaço de $\mathbb{R}^n$.*

Proposição 5.17 *Sejam $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear e S um subespaço de $\mathbb{R}^n$, então $f(S)$ é um subespaço de $\mathbb{R}^m$.*

Demonstração Por definição,

$$f(S) = \{f(u): \ u \in S\} \subseteq \mathbb{R}^m$$

e porque

$$f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m},$$

então $f(S)$ é um subconjunto não vazio de $\mathbb{R}^m$.

Sejam r e v dois vectores de $f(S)$. Então, existem u e s em S tais que

$$r = f(u) \qquad \text{e} \qquad v = f(s).$$

Assim,

$$r + s = f(u) + f(v) \underset{\substack{\uparrow \\ \text{linearidade de } f}}{=} f(u+v).$$

Como S é subespaço de $\mathbb{R}^n$, $u+s$ pertence a S. Isto implica que $r+v$ seja um elemento de $f(S)$.

Seja k um escalar. Então,

$$kr = kf(u) \underset{\substack{\uparrow \\ \text{linearidade de } f}}{=} f(ku).$$

Como S é subespaço de $\mathbb{R}^n$, ku pertence a S e consequentemente kr é um elemento de $f(S)$. Portanto, $f(S)$ é subespaço de $\mathbb{R}^m$. $\square$

Desta Proposição resulta o seguinte resultado.

Proposição 5.18 *Seja $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear, então $Im\ f = f(\mathbb{R}^n)$ é um subespaço de $\mathbb{R}^m$.*

Recorde-se que se A e B são conjuntos e $f: A \longrightarrow B$ é uma aplicação, dizemos que

- f é **injectiva** se
$$\forall x, x' \in A, \ f(x) = f(x') \Longrightarrow x = x'.$$
ou seja, f é injectiva se elementos diferentes de A têm imagens diferentes.

- f é **sobrejectiva** se
$$\forall y \in B,\ \exists x \in A,\ f(x) = y$$
isto é, f é sobrejectiva se qualquer elemento de B é imagem, por f, de algum elemento de A.

- f é **bijectiva** se f é injectiva e sobrejectiva, isto é,
$$\forall y \in B,\ \exists^1 x \in A:\ f(x) = y.$$

Vejamos uma outra caracterização de injectividade quando f é uma aplicação linear.

Proposição 5.19 *Seja* $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação linear. Então*
$$f \text{ é injectiva se, e só se, } Nuc\ f = \{0_{\mathbb{R}^n}\}.$$

Demonstração Suponhamos que f é injectiva e seja $u \in Nuc\ f$. Então,
$$f(u) = 0_{\mathbb{R}^m} = f(0_{\mathbb{R}^n}).$$
Como f é injectiva, $u = 0_{\mathbb{R}^n}$. Donde,
$$Nuc\ f = \{0_{\mathbb{R}^n}\}.$$

Reciprocamente, suponhamos que $Nuc\ f = \{0_{\mathbb{R}^n}\}$ e vejamos que f é injectiva.
Se x e $x' \in \mathbb{R}^n$ forem tais que
$$f(x) = f(x') \qquad \text{então} \qquad f(x) - f(x') = 0_{\mathbb{R}^m}.$$
Mas f é aplicação linear, pelo que
$$f(x - x') = 0_{\mathbb{R}^m}.$$
Mas isto significa que $x - x' \in Nuc\ f = \{0_{\mathbb{R}^n}\}$, ou seja, $x - x' = 0_{\mathbb{R}^n}$ e consequentemente, $x = x'$. □

Observação Sendo $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear e M_f a matriz canónica de f, então,

- f é injectiva se, e só se, $Nuc\ f = \{0_{\mathbb{R}^n}\}$ (Proposição 5.19) se, e só se, o sistema homogéneo $M_f X = 0$ é possível e determinado (só tem a solução nula).

- f é sobrejectiva se, e só se, o sistema $M_f X = B$ é possível, qualquer que seja a matriz coluna B cujo vector está em $\mathbb{R}^m$.

Observação Se $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma aplicação linear e M_f é a matriz canónica de f (matriz $m \times n$) , então,

- Se $n > m$, o sistema homogéneo $M_f X = 0$ tem mais incógnitas (o número de colunas de M_f é n) do que equações (o número de linhas de M_f é m). Assim,

$$r(M_f) \leq m < n$$

e o sistema é possível e indeterminado. Logo, f não é injectiva.

- Se $m > n$, então a matriz em forma de escada reduzida obtida de M_f terá, pelo menos, uma linha nula. Sendo A a matriz em forma de escada reduzida obtida de M_f, então existem matrizes elemntares $E_1, \ldots, E_k$ tais que

$$E_1 \ldots E_k M_f = A.$$

Como as matrizes elementares são invertíveis,

$$M_f = E_k^{-1} \ldots E_1^{-1} A.$$

Se B for a matriz coluna $m \times 1$, com todas as entradas iguais a zero à excepção da entrada $(m, 1)$ que será igual a 1, então o sistema $AX = B$ é impossível (recorde-se que a m-ésima linha de A é nula). Mas isto implica que o sistema

$$M_f X = (E_k^{-1} \ldots E_1^{-1}) AX = (E_k^{-1} \ldots E_1^{-1}) B$$

seja impossível. Logo, f não é sobrejectiva.

Proposição 5.20 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ uma aplicação linear. Então*

f é injectiva se, e só se, f é sobrejectiva.

Demonstração Suponhamos que f é injectiva e seja M_f a sua matriz canónica. Pela observação, o sistema $M_f X = 0$ é possível e determinado. Isto implica que $r(M_f) = n$, donde M_f é invertível. Assim, o sistema $M_f X = B$ é possível e determinado, qualquer que seja a matriz coluna B. Donde, f é sobrejectiva.

Reciprocamente, se f é sobrejectiva, o sistema $M_f X = B$ é possível, qualquer que seja a matriz coluna B de tipo $m \times 1$. Se $r(M_f) < n$, então a matriz em forma de escada reduzida obtida de M_f teria, pelo menos, uma linha nula. Fazendo um raciocínio análogo ao feito na observação anterior (caso $m > n$), concluíriamos que existiria uma matriz coluna B' tal que $M_f X = B'$ era um sistema impossível. Como isto não acontece, $r(M_f) = n$ e M_f é invertível. Portanto, o sistema $M_f X = 0$ é possível e determinado e f é injectiva. $\square$

Exemplo 5.21 *A aplicação linear*

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

$$(x, y, z) \mapsto (x + y, z, 0)$$

não é sobrejectiva, porque a matriz canónica de f é

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

e o sistema

$$M_f X = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

é impossível. Pela Proposição 5.20, *f também não é injectiva.*

Proposição 5.22 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear. Então,*

1. *Se $W =< v_1, \ldots, v_k >$ é subespaço de $\mathbb{R}^n$, então*

$$f(W) =< f(v_1), \ldots, f(v_k) > .$$

2. *Se f é injectiva e S é um conjunto de vectores linearmente independentes, de $\mathbb{R}^n$, então, $f(S)$ é um conjunto de vectores linearmente independentes de $\mathbb{R}^m$.*

Demonstração

1. Pela Proposição 5.17, $f(W)$ é subespaço de $\mathbb{R}^m$. Porque $v_1, \ldots, v_k \in W$, então $f(v_1), \ldots, f(v_k) \in f(W)$ e

$$< f(v_1), \ldots, f(v_k) >\subseteq f(W).$$

Seja $z \in f(W)$. Então, existe $u \in W$ tal que $f(u) = z$. Sendo

$$u = \alpha_1 v_1 + \ldots + \alpha_k v_k$$

com $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, então

$$z = f(u) = f(\alpha_1 v_1 + \ldots + \alpha_k v_k) \underset{\substack{\uparrow \\ \text{linearidade de } f}}{=} \alpha_1 f(v_1) + \ldots + \alpha_k f(v_k).$$

Ou seja, $z \in< f(v_1), \ldots, f(v_k) >$ e

$$f(W) =< f(v_1), \ldots, f(v_k) > .$$

2. Suponhamos que $S = \{v_1, \ldots, v_r\}$ é lineramente independente.

Então, $f(S) = \{f(v_1), \ldots, f(v_r)\}$. Se

$$\alpha_1 f(v_1) + \ldots + \alpha_r f(v_r) = 0_{\mathbb{R}^m},$$

então, por linearidade,

$$f(\alpha_1 v_1 + \ldots + \alpha_r v_r) = 0_{\mathbb{R}^m},$$

ou seja,

$$\alpha_1 v_1 + \ldots + \alpha_r v_r \in Nuc\ f.$$

Porque f é injectiva, $Nuc\ f = \{0_{\mathbb{R}^n}\}$ e

$$\alpha_1 v_1 + \ldots + \alpha_r v_r = 0_{\mathbb{R}^n}.$$

Mas S é linearmente independente, então

$$\alpha_1 = 0, \ldots, \alpha_r = 0$$

e $f(S)$ é linearmente independente. □

5.6 Composição de Aplicações

Se A, B e C forem conjuntos, $f: A \longrightarrow B$ e $g: B \longrightarrow C$ forem duas aplicações, como o conjunto de chegada de f é igual ao domínio de g, podemos definir a **composição** de g após f, que se designa por $g \circ f$, e é a aplicação

$$\begin{aligned} g \circ f:\ A &\longrightarrow C \\ x &\mapsto g \circ f(x) = g(f(x)) \end{aligned}$$

Proposição 5.23 *Se $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e $g: \mathbb{R}^m \longrightarrow \mathbb{R}^k$ forem duas aplicações lineares, então a aplicação $g \circ f: \mathbb{R}^n \longrightarrow \mathbb{R}^k$ é uma aplicação linear.*

Demonstração Vejamos que $g \circ f: \mathbb{R}^m \longrightarrow \mathbb{R}^k$ é linear.

1. Se $x \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$, então,

$$g \circ f(\alpha x) = g(f(\alpha x)) \underset{\substack{\uparrow \\ f\ linear}}{=} g(\alpha f(x)) \underset{\substack{\uparrow \\ g\ linear}}{=} \alpha g(f(x)) = \alpha(g \circ f(x)).$$

2. Se $x_1, x_2 \in \mathbb{R}^n$, então,

$$g \circ f(x_1 + x_2) = g(f(x_1 + x_2)) \underset{\substack{\uparrow \\ f\ linear}}{=} g(f(x_1) + f(x_2)) =$$

$$\underset{\substack{\uparrow \\ g\ linear}}{=} g(f(x_1)) + g(f(x_2)) = (g \circ f(x_1)) + (g \circ f(x_2)).$$

Logo, $g \circ f$ é uma aplicação linear. $\square$

Sejam $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e $g: \mathbb{R}^m \longrightarrow \mathbb{R}^k$ duas aplicações lineares, M_f e M_g as respectivas matrizes canónicas. Podemos definir a aplicação linear

$$g \circ f: \mathbb{R}^n \longrightarrow \mathbb{R}^k.$$

Nestas condições, uma pergunta pode ser colocada:

Qual a relação que existe entre as matrizes canónicas $M_{g \circ f}$, M_f e M_g?

Atendendo à definição de composição de aplicações, se e_i for um vector canónico de $\mathbb{R}^n$, então

$$g \circ f(e_i) = g(f(e_i)).$$

Porque o vector $g(f(e_i))$ pode ser escrito na forma matricial por

$$M_g[f(e_i)]$$

onde $[f(e_i)]$ é a matriz coluna $m \times 1$, com as componentes de $f(e_i)$ e o vector $f(e_i)$ pode ser escrito como

$$M_f[e_i]$$

em que $[e_i]$ é a matriz coluna $n \times 1$, com as componentes de e_i, vem que

$$M_g M_f[e_i]$$

é a forma matricial de $g(f(e_i))$. Porque a forma matricial de $g \circ f(e_i)$ é

$$M_{g \circ f}[e_i]$$

e $g \circ f(e_i) = g(f(e_i))$, então,

$$M_{g \circ f} = M_g M_f.$$

Acabámos de demonstrar o seguinte resultado:

Proposição 5.24 *Sejam $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e $g: \mathbb{R}^m \longrightarrow \mathbb{R}^k$ duas aplicações lineares cujas matrizes canónicas são M_f e M_g. Então, a matriz canónica de $g \circ f: \mathbb{R}^n \longrightarrow \mathbb{R}^k$ é obtida por*

$$M_{g \circ f} = M_g M_f.$$

Exemplo 5.25 1. *Sejam $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ a rotação em torno da origem, segundo o ângulo de amplitude θ e $g: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ a rotação em torno da origem, mas segundo o ângulo de amplitude ρ.*

As matrizes canónicas de f e g são

$$R_\theta = \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix}, \qquad R_\rho = \begin{bmatrix} cos\ \rho & -sen\ \rho \\ sen\ \rho & cos\ \rho \end{bmatrix}.$$

A composição $g \circ f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ é a aplicação linear que tem a matriz canónica (Proposição 5.24)

$$M_{g \circ f} = R_\rho R_\theta = \begin{bmatrix} cos\ \rho & -sen\ \rho \\ sen\ \rho & cos\ \rho \end{bmatrix} \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix} = \begin{bmatrix} cos\ (\rho + \theta) & -sen\ (\rho + \theta) \\ sen\ (\rho + \theta) & cos\ (\rho + \theta) \end{bmatrix}$$

ou seja, $g \circ f$ é a rotação em torno da origem segundo o ângulo de amplitude $(\rho + \theta)$.

Repare que neste caso, se tivessemos calculado $M_{f \circ g}$ teríamos obtido uma matriz igual a $M_{g \circ f}$.

2. *Sejam $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ e $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ duas reflexões no eixo dos yy. As matrizes canónicas de f e g são iguais à matriz*

$$H_{\frac{\pi}{2}} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Pela Proposição 5.24,

$$M_{g \circ f} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2,$$

pelo que $g \circ f$ não é uma reflexão através de uma recta que passe pela origem, é sim, uma rotação segundo o ângulo de amplitude 0 (*ou também se poderia dizer que é um alongamento de razão* 0).

3. *Sejam $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ a rotação em torno da origem, segundo o ângulo de amplitude θ e $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ uma reflexão no eixo dos yy. As matrizes canónicas de f e g são respectivamente*

$$M_f = R_\theta = \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix}, \qquad M_g = H_{\frac{\pi}{2}} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Pela Proposição 5.24,

$$M_{g \circ f} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix} = \begin{bmatrix} -cos\ \theta & sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix} = \begin{bmatrix} cos\ (\pi - \theta) & sen\ (\pi - \theta) \\ sen\ (\pi - \theta) & -cos\ (\pi - \theta) \end{bmatrix},$$

pelo que $g \circ f$ é uma reflexão através de uma recta que passa pela origem e faz um ângulo de amplitude $\frac{(\pi - \theta)}{2}$ com o eixo dos xx positivo.

Por outro lado, pela Proposição 5.24,

$$M_{f \circ g} = \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -cos\ \theta & -sen\ \theta \\ -sen\ \theta & cos\ \theta \end{bmatrix} = \begin{bmatrix} cos\ (\pi + \theta) & sen\ (\pi + \theta) \\ sen\ (\pi + \theta) & -cos\ (\pi + \theta) \end{bmatrix},$$

pelo que $f \circ g$ é uma reflexão através de uma recta que passa pela origem e faz um ângulo de amplitude $\frac{(\pi + \theta)}{2}$ com o eixo dos xx positivo.

Portanto neste caso, as aplicações $g \circ f$ e $f \circ g$ são distintas.

5.7 Composição de Aplicações e Matrizes Elementares

Se uma matriz é invertível então pode escrever-se como produto de matrizes elementares (Teorema 2.37). Repare que as matrizes elementares de ordem 2 são:

1. $$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$$

   com $k > 0$, $k \neq 1$, que corresponde a uma compressão ou a uma expansão horizontal.

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

   com $k > 0$, $k \neq 1$, que corresponde a uma compressão ou a uma expansão vertical.

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

   que corresponde a uma reflexão no eixo dos yy.

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

   que corresponde a uma reflexão no eixo dos xx.

5. $$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$$

   com $k < 0$, $k \neq -1$, que corresponde a $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -k & 0 \\ 0 & 1 \end{bmatrix}$, em que $-k > 0$ e $-k \neq 1$, ou seja, uma compressão ou a uma expansão horizontal seguida de uma reflexão no eixo dos yy.

6. $$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

   com $k < 0$, $k \neq -1$, que corresponde a $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -k \end{bmatrix}$, em que $-k > 0$ e $-k \neq 1$, ou seja, uma compressão ou a uma expansão vertical seguida de uma reflexão no eixo dos xx.

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

   que corresponde a uma reflexão através da recta $y = x$.

8.
$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

que corresponde a um alongamento na direcção de x, de razão k.

9.
$$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

que corresponde a um alongamento na direcção de y, de razão k.

Assim:

Proposição 5.26 *Se A for uma matriz de ordem 2 invertível, então a aplicação linear $f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, cuja matriz canónica é A, é uma composição de alongamentos, compressões e expansões nas direcções dos eixos ordenados e reflexões nos eixos ordenados e através da recta $y = x$.*

Exemplo 5.27 *A matriz* $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ *é invertível pois* $\det\ A = 3 + 2 = 5 \neq 0$.

Para obtermos a forma de escada reduzida de A teremos de efectuar as transformações elementares

$$\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \underset{l_2 \to (l_2 + l_1)}{\longrightarrow} \begin{bmatrix} 1 & 2 \\ 0 & 5 \end{bmatrix} \underset{l_2 \to \frac{1}{5} l_2}{\longrightarrow} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \underset{l_1 \to (l_1 - 2l_2)}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

que corresponde às matrizes elementares

$$E_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

Então,

$$A = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

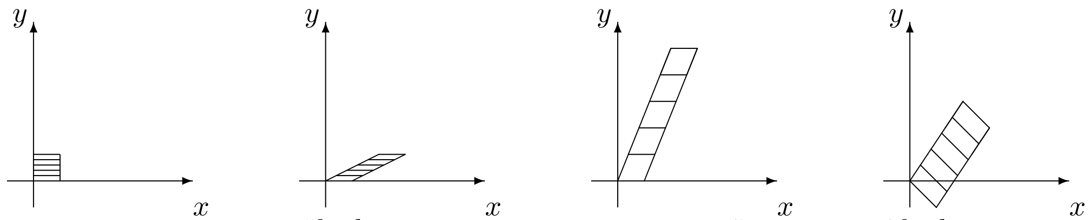
Se $f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ for a aplicação linear cuja matriz canónica é A, então

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ -x_1 + 3x_2 \end{bmatrix}$$

isto é, $f(x_1, x_2) = (x_1 + 2x_2, -x_1 + 3x_2)$.

Usando a Proposição 5.26, *f é uma composição de aplicações que resulta de efectuar um alongamento na direcção de x, de razão 2, uma expansão na direcção de y, de razão 5 e um alongamento na direcção de y, de razão -1.*

(**ATENÇÃO***: a composição de aplicações lê-se da direita para a esquerda.*)



1^o *alongamento* *expansão* 2^o *alongamento*

5.8 Aplicações Lineares Invertíveis

Vejamos o que acontece às aplicações lineares que são invertíveis.

Definição 5.28 *Sejam A, B conjuntos e $f: A \longrightarrow B$ uma aplicação. Dizemos que f é* **invertível** *se existir uma aplicação $g: B \longrightarrow A$ tal que*

$$f \circ g = id_B \qquad , \qquad g \circ f = id_A$$

em que

$$\begin{array}{rl} id_B: & B \longrightarrow B \\ & x \mapsto x \end{array} \qquad , \qquad \begin{array}{rl} id_A: & A \longrightarrow A \\ & x \mapsto x \end{array}$$

Observação

1. $f: A \longrightarrow B$ é uma aplicação invertível se, e só se, f é bijectiva.
2. Se f é invertível então a aplicação g tal que $g \circ f = id$, $f \circ g = id$ é única. Esta aplicação chama-se **inversa** de f e denota-se por f^{-1}.
3. Se $f: A \longrightarrow B$ é injectiva, então $f: A \longrightarrow Im\ f$ é bijectiva, pelo que existe $f^{-1}: Im\ f \longrightarrow A$.

Proposição 5.29 *Se $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ é uma aplicação linear injectiva, então*

$$f^{-1}: Im\ f \longrightarrow \mathbb{R}^n$$

é uma aplicação linear.

Demonstração Atendendo à observação, $f: \mathbb{R}^n \longrightarrow Im\ f$ é uma aplicação linear bijectiva. Então existe a aplicação

$$f^{-1}: \ Im\ f \longrightarrow \mathbb{R}^n$$

$$f(x) = y \to x$$

Vejamos que f^{-1} é linear.

Sejam $y_1, y_2 \in Im\ f$, então se α_1 e $\alpha_2 \in \mathbb{R}$,

$$f(f^{-1}(\alpha_1 y_1 + \alpha_2 y_2)) = f \circ f^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 y_1 + \alpha_2 y_2$$

e

$$f(\alpha_1 f^{-1}(y_1) + \alpha_2 f^{-1}(y_2)) \underset{\substack{\uparrow \\ f \text{ linear}}}{=} \alpha_1(f \circ f^{-1}(y_1)) + \alpha_2(f \circ f^{-1}(y_2)) = \alpha_1 y_1 + \alpha_2 y_2$$

Consequentemente,

$$f(f^{-1}(\alpha_1 y_1 + \alpha_2 y_2)) = f(\alpha_1 f^{-1}(y_1) + \alpha_2 f^{-1}(y_2)).$$

Como f é injectiva,

$$f^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 f^{-1}(y_1) + \alpha_2 f^{-1}(y_2),$$

ou seja, f^{-1} é linear. □

Teorema 5.30 *Seja $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ uma aplicação linear. Se a matriz canónica, M_f, de f é invertível então*

1. *f é invertível.*
2. *a matriz canónica de f^{-1} é $M_{f^{-1}}$, isto é $M_{f^{-1}} = M_f^{-1}$.*

Demonstração

1. Se M_f é invertível então, pela Observação que está a seguir à Proposição 5.19, f é injectiva e sobrejectiva. Então, f é invertível.
2. Porque $f \circ f^{-1} = id_{\mathbb{R}^n}$ e a matriz canónica de $id_{\mathbb{R}^n}$ é I_n, temos que

$$M_f M_{f^{-1}} = M_{f \circ f^{-1}} = I_n.$$

   Consequentemente, $M_{f^{-1}} = M_f^{-1}$. □

Exemplo 5.31 *Seja* $f:\ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *a rotação em torno da origem, segundo o ângulo de amplitude* θ. *a matriz canónica de* f *é*

$$R_\theta = \begin{bmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{bmatrix}.$$

Porque $\det R_\theta = cos^2\theta + sen^2\theta = 1 \neq 0$, *então* R_θ *é invertível. Pelo* Teorema 5.30, f *é invertível e a matriz canónica de* f^{-1} *é*

$$R_\theta^{-1} = \begin{bmatrix} cos\ \theta & sen\ \theta \\ -sen\ \theta & cos\ \theta \end{bmatrix} = \begin{bmatrix} cos\ (2\pi - \theta) & -sen\ (2\pi - \theta) \\ sen\ (2\pi - \theta) & cos\ (2\pi - \theta) \end{bmatrix}.$$

Ou seja, f^{-1} *é a rotação em torno da origem, segundo o ângulo de amplitude* $(2\pi - \theta)$.

Capítulo 6

BASES

Dado um subespaço $W =< v_1, \ldots, v_k >$ de $\mathbb{R}^n$, se o conjunto $S = \{v_1, \ldots, v_k\}$ é linearmente dependente, existe, pelo menos, um vector de S que é combinação linear dos outros vectores (Proposição 3.12). Suponhamos, sem perda de generalidade, que v_1 é combinação linear de $v_2, \ldots, v_k$, isto é,

$$v_1 = a_2 v_2 + \ldots + a_k v_k, \qquad \text{com } a_2, \ldots, a_k \in \mathbb{R}.$$

Seja u um vector de W, então

$$u = b_1 v_1 + b_2 v_2 + \ldots + b_k v_k, \qquad \text{com } b_1, b_2, \ldots, b_k \in \mathbb{R}.$$

Mas então,

$$\begin{aligned} u &= b_1(a_2 v_2 + \ldots + a_k v_k) + b_2 v_2 + \ldots + b_k v_k \\ &= (b_1 a_2 + b_2) v_2 + \ldots + (b_k a_k + b_k) v_k \end{aligned}$$

isto é, $u \in< v_2, \ldots, v_k >$. Portanto,

$$W \subseteq< v_2, \ldots, v_k > .$$

Facilmente se vê que $< v_2, \ldots, v_k >\subseteq W =< v_1, \ldots, v_k >$. Ou seja,

$$W =< v_2, \ldots, v_k > .$$

Podemos repetir este processo até termos um conjunto de vectores que gera W e é linearmente independente.

6.1 Definição e Exemplos de Bases

Comecemos pela definição de base.

Definição 6.1 *Um conjunto de vectores de um subespaço W de $\mathbb{R}^n$ é uma* **base** *de W se verifica as duas condições seguintes:*

1. *É linearmente independente.*
2. *Gera W.*

Exemplo 6.2 1. *Já vimos que $\mathbb{R}^n =< e_1, e_2, \ldots, e_n >$ em que $e_1, e_2, \ldots, e_n$ são os vectores canónicos de $\mathbb{R}^n$. Como o conjunto formado por estes vectores é linearmente independente, então*

$$\{e_1, e_2, \ldots, e_n\}$$

é uma base de $\mathbb{R}^n$, chamada **base canónica de** $\mathbb{R}^n$.

2. *Sendo $W =< (1,2,2),(0,1,1),(2,4,4) >$, porque*

$$(2,4,4) = 2(1,2,2)$$

então

$$W =< (1,2,2),(0,1,1) > .$$

Estes dois vectores que geram W são linearmente independentes, pois se

$$\alpha_1(1,2,2) + \alpha_2(0,1,1) = (0,0,0)$$

então

$$(\alpha_1, 2\alpha_1 + \alpha_2, 2\alpha_1 + \alpha_2) = (0,0,0).$$

Donde,

$$\alpha_1 = 0 \qquad e \qquad 2\alpha_1 + \alpha_2 = 0.$$

Portanto,

$$\alpha_1 = 0 \qquad e \qquad \alpha_2 = 0.$$

Então, o conjunto formado por estes dois vectores é uma base de W.

3. *O conjunto*

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

é uma base de $\mathbb{R}^3$ porque se

$$\alpha_1(1,1,1) + \alpha_2(0,1,1) + \alpha_3(0,0,1) = (0,0,0)$$

então

$$(\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3) = (0,0,0).$$

Donde,

$$\alpha_1 = 0, \qquad \alpha_1 + \alpha_2 = 0, \qquad \alpha_1 + \alpha_2 + \alpha_3 = 0.$$

Portanto,

$$\alpha_1 = 0, \qquad \alpha_2 = 0, \qquad \alpha_3 = 0.$$

Ou seja, o conjunto formado pelos 3 vectores é linearmente independente.

Vejamos que também gera $\mathbb{R}^3$*. Seja* $(x, y, z) \in \mathbb{R}^3$*. Mostremos que existem escalares* $\alpha_1,\ \alpha_2,\ \alpha_3 \in \mathbb{R}$ *tais que*

$$\alpha_1(1,1,1) + \alpha_2(0,1,1) + \alpha_3(0,0,1) = (x, y, z).$$

Mas isto implica que

$$(\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3) = (x, y, z).$$

Donde,

$$\alpha_1 = x, \qquad \alpha_1 + \alpha_2 = y, \qquad \alpha_1 + \alpha_2 + \alpha_3 = z.$$

Portanto,

$$\alpha_1 = x, \qquad \alpha_2 = y - x, \qquad \alpha_3 = z - y.$$

Ou seja, qualquer vector de $\mathbb{R}^3$ *pode escrever-se como combinação linear destes 3 vectores. Logo, o conjunto formado pelos 3 vectores é uma base de* $\mathbb{R}^3$*.*

4. *O conjunto*

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

não é uma base de $\mathbb{R}^3$ *porque se*

$$\alpha_1(1,1,1) + \alpha_2(0,1,1) + \alpha_3(1,0,0) = (0,0,0)$$

então

$$(\alpha_1 + \alpha_3, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2) = (0,0,0).$$

Donde,

$$\alpha_1 + \alpha_3 = 0 \qquad e \qquad \alpha_1 + \alpha_2 = 0.$$

Portanto,

$$\alpha_3 = -\alpha_1 \qquad e \qquad \alpha_2 = -\alpha_1.$$

Ou seja, o conjunto formado pelos 3 vectores é linearmente dependente.

5. *O conjunto*

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

não é uma base de $\mathbb{R}^3$ *porque* $(0,2,0) \in \mathbb{R}^3$ *e vejamos que não existem escalares* $\alpha_1,\ \alpha_2 \in \mathbb{R}$ *tais que*

$$\alpha_1(1,1,1) + \alpha_2(0,1,1) = (0,2,0).$$

Isto implica que

$$(\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2) = (0,2,0).$$

Donde,

$$\alpha_1 = 0, \qquad \alpha_1 + \alpha_2 = 2, \qquad \alpha_1 + \alpha_2 = 0.$$

Portanto, um sistema impossível. Ou seja, o vector $(0,2,0)$ *de* $\mathbb{R}^3$ *não se escreve como combinação linear destes 2 vectores. Logo, o conjunto formado pelos 2 vectores não gera* $\mathbb{R}^3$*.*

Convenção 6.3 *No caso de $W =< 0_{\mathbb{R}^n} >$, convenciona-se que a sua base é o conjunto vazio, $\emptyset$.*

6.2 Dimensão de um Subespaço

Uma pergunta pode colocar-se; "Será possível arranjarmos duas bases de um mesmo subespaço vectorial com um número diferente de vectores?"

O seguinte teorema responde a esta pergunta.

Teorema 6.4 *Todas as bases de um subespaço de $\mathbb{R}^n$ têm o mesmo número de vectores.*

Demonstração Seja W um subespaço de $\mathbb{R}^n$.

Se $W =< 0_{\mathbb{R}^n} >= \{0_{\mathbb{R}^n}\}$ então, por convenção, a única base é o $\emptyset$.

Suponhamos que $W \neq \{0_{\mathbb{R}^n}\}$ e que $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_k\}$ e $\mathcal{B}_2 = \{s_1, s_2, \ldots, s_m\}$ são duas bases de W. Suponhamos, sem perda de generalidade, que $k > m$.

Como $\mathcal{B}_2$ gera W, cada vector de $\mathcal{B}_1$ escreve-se como combinação linear dos vectores de $\mathcal{B}_2$. Isto é,

$$\begin{aligned}
v_1 &= a_{11}s_1 + a_{12}s_2 + \ldots + a_{1m}s_m \\
v_2 &= a_{21}s_1 + a_{22}s_2 + \ldots + a_{2m}s_m \\
&\vdots \\
v_k &= a_{k1}s_1 + a_{k2}s_2 + \ldots + a_{km}s_m
\end{aligned}$$

Se considerarmos o sistema homogéneo

$$\begin{bmatrix} a_{11} & a_{21} & \ldots & a_{k1} \\ a_{12} & a_{22} & \ldots & a_{k2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \ldots & a_{km} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

como $m < k$, o sistema é possível e indeterminado (Corolário 1.18), ou seja, existem escalares $c_1, c_2, \ldots, c_k$, não todos nulos, tais que

$$\begin{cases} c_1a_{11} + c_2a_{21} + \ldots + c_ka_{k1} = 0 \\ c_1a_{12} + c_2a_{22} + \ldots + c_ka_{k2} = 0 \\ \vdots \\ c_1a_{1m} + c_2a_{2m} + \ldots + c_ka_{km} = 0. \end{cases}$$

Mas então,

$$\begin{aligned}
c_1v_1 + \ldots + c_kv_k &= c_1(a_{11}s_1 + a_{12}s_2 + \ldots + a_{1m}s_m) + \ldots + c_k(a_{k1}s_1 + a_{k2}s_2 + \ldots + a_{km}s_m) \\
&= (c_1a_{11} + c_2a_{21} + \ldots + c_ka_{k1})s_1 + \ldots + (c_1a_{1m} + c_2a_{2m} + \ldots + c_ka_{km})s_m \\
&= 0.
\end{aligned}$$

O que significa que os vectores $v_1, v_2, \ldots, v_k$ são linearmente dependentes. Impossível pois $\mathcal{B}_2$ é uma base de W. Logo, $k \leq m$. Fazendo o mesmo raciocínio mas trocando os papéis de $\mathcal{B}_1$ e $\mathcal{B}_2$ obtemos que $k = m$. $\square$

Definição 6.5 *Se W é um subespaço de $\mathbb{R}^n$, chamamos* **dimensão** *de W e denotamos por*

$$\dim W,$$

ao número de vectores de qualquer sua base.

Exemplo 6.6

1. *Se $W = \{0_{\mathbb{R}^n}\}$ então $\dim W = 0$.*
2. *Uma recta de $\mathbb{R}^n$ que passa pela origem tem dimensão 1.*
3. *Um plano de $\mathbb{R}^n$ que passa pela origem tem dimensão 2.*
4. *$\mathbb{R}^n$ tem dimensão n.*
5. *O subespaço $W =< (1,2,2), (0,1,1), (2,4,4) >$ do* Exemplo 6.2.2*, tem como sua base o conjunto*
$$\{(1,2,2), (0,1,1)\},$$
pelo que
$$\dim W = 2.$$

6.3 Alguns Resultados sobre Bases

Nesta secção iremos ver a importância de determinarmos uma base de um subespaço.

Teorema 6.7 *Se $S = \{v_1, \ldots, v_k\}$ é uma base do subespaço W de $\mathbb{R}^n$, então qualquer vector u de W se escreve de forma única como combinação linear dos vectores de S.*

Demonstração como $W =< v_1, \ldots, v_k >$, se $u \in W$, então existem escalares $\alpha_1, \alpha_2, \ldots, \alpha_k$ tais que

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k.$$

Suponhamos que u se escreve de outra forma como combinação linear dos vectores $v_1, v_2, \ldots, v_k$, isto é, existem escalares $\beta_1, \beta_2, \ldots, \beta_k$ tais que

$$u = \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_k v_k.$$

Então,

$$0 = u - u = (\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \ldots + (\alpha_k - \beta_k)v_k.$$

Porque S é linearmente independente,

$$\alpha_1 - \beta_1 = 0,\ \alpha_2 - \beta_2 = 0,\ \ldots,\ \alpha_k - \beta_k = 0$$

ou seja,

$$\alpha_1 = \beta_1,\ \alpha_2 = \beta_2,\ \ldots,\ \alpha_k = \beta_k.$$

E a forma de escrever u como combinação linear de $v_1, \ldots, v_k$ é única. □

Já vimos que se W é um subespaço de $\mathbb{R}^n$ e se S é um conjunto de vectores de W que o geram, então conseguimos encontrar uma base de W que é um subconjunto de S. Vejamos agora como construir uma base de W, que contenha um subconjunto de vectores de W linearmente independentes.

Teorema 6.8 *Um conjunto de $\mathbb{R}^n$ com mais do que n vectores é linearmente dependente.*

Demonstração Seja $S = \{v_1, \ldots, v_m\}$ um conjunto de $\mathbb{R}^n$ com $m > n$ vectores. Consideremos a matriz A cujas colunas são os vectores $v_1, \ldots, v_m$. Porque

$$r(A) \leq \min\{m, n\} = n \neq m,$$

então S é linearmente dependente. □

Teorema 6.9 *Sejam W um subespaço de $\mathbb{R}^n$ de dimensão k e S um conjunto de vectores linearmente independentes, contido em W. Então, existe uma base de W que contém os vectores de S.*

Demonstração Se

$$S = \{v_1, \ldots, v_r\}$$

não é uma base de W, então existe $s_1 \in W$ que não se escreve como combinação linear dos vectores de S. Mas então, o conjunto

$$T_1 = \{v_1, \ldots, v_r, s_1\}$$

é linearmente independente. Se T_1 não é uma base de W, repetimos o processo anterior.

Este processo tem fim porque $\dim \mathbb{R}^n = n$, ou seja, um conjunto com n vectores linearmente independente tem que gerar $\mathbb{R}^n$ (Teorema 6.8) e portanto qualquer vector de W, sendo vector de $\mathbb{R}^n$, se escreve como combinação linear destes n vectores. Então, conseguimos encontrar uma base de W, nas condições do enunciado. □

Corolário 6.10 1. *Qualquer subespaço de $\mathbb{R}^n$ tem uma base e a sua dimensão é menor ou igual a n.*

2. *Qualquer conjunto linearmente independente com n vectores de $\mathbb{R}^n$ é uma sua base.*

3. *Qualquer conjunto com n vectores que gere $\mathbb{R}^n$ é uma sua base.*

Exemplo 6.11 1. *Sendo $S = \{(1,1,1), (-2,-2,3)\}$ um conjunto de vectores linearmente independente de $\mathbb{R}^3$, determinemos uma base de $\mathbb{R}^3$ que contenha os vectores de S.*

Sabemos que $\dim \mathbb{R}^3 = 3$*, portanto teremos de encontrar um vector de $\mathbb{R}^3$, que junto com os vectores de S forme um conjunto de vectores linearmente independente.*

Se este vector for o vector (a,b,c), queremos determinar a,b,c por forma a que $S' = \{(1,1,1), (-2,-2,3), (a,b,c)\}$ seja linearmente independente, ou seja, que o sistema

$$\begin{bmatrix} 1 & -2 & a \\ 1 & -2 & b \\ 1 & 3 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

seja possível e determinado. Porque

$$\begin{bmatrix} 1 & -2 & a \\ 1 & -2 & b \\ 1 & 3 & c \end{bmatrix} \underset{l_2 \to (l_2 - l_1)}{\longrightarrow} \begin{bmatrix} 1 & -2 & a \\ 0 & 0 & b-a \\ 1 & 3 & c \end{bmatrix} \underset{l_3 \to (l_3 - l_1)}{\longrightarrow} \begin{bmatrix} 1 & -2 & a \\ 0 & 0 & b-a \\ 0 & 5 & c-a \end{bmatrix} \underset{l_2 \leftrightarrow l_3}{\longrightarrow} \begin{bmatrix} 1 & -2 & a \\ 0 & 5 & c-a \\ 0 & 0 & b-a \end{bmatrix}$$

a matriz simples do sistema tem característica 3 (*que é o número de incógnitas*) *se* $b - a \neq 0$. *Isto verifica-se por exemplo, se* $a = 0,\ b = 1,\ c = 0$. *Neste caso,*

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

é uma base de $\mathbb{R}^3$ que contém os vectores de S.

2. *Seja*

$$W = \{(x,y,z,w) \in \mathbb{R}^4 :\ x + y + w = 0\}$$

um subconjunto de $\mathbb{R}^4$.

Facilmente se vê que W é um subespaço de $\mathbb{R}^4$ e que

$$W = < (1,-1,0,0), (0,0,1,0), (0,-1,0,1) > .$$

Vejamos como conseguimos obter uma base de W que contenha os vectores do conjunto

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

Ora uma base de W é o conjunto $T = \{(1,-1,0,0), (0,0,1,0), (0,-1,0,1)\}$. Portanto, $\dim W = 3$. *Os elementos de S pertencem a W. Atendendo ao* Teotema 6.4, *uma base de W tem* 3 *elementos. Vejamos um vector de T que não pertença a $< S >$. Por exemplo o vector* $(0,0,1,0)$. *Ora*

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & -2 & 0 \\ 2 & 5 & 1 \\ -3 & 1 & 0 \end{bmatrix} \underset{l_4 \to (l_4 + l_1)}{\longrightarrow} \begin{bmatrix} 3 & 1 & 0 \\ 0 & -2 & 0 \\ 2 & 5 & 1 \\ 0 & 2 & 0 \end{bmatrix} \underset{l_4 \to (l_4 - l_2)}{\longrightarrow} \begin{bmatrix} 3 & 1 & 0 \\ 0 & -2 & 0 \\ 2 & 5 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Daqui, facilmente se conclui que a característica desta matriz é 3*, pelo que os* 3 *vectores são linearmente independentes. Assim, o conjunto*

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

é uma base de W *nas condições.*

Até aqui estudámos só subespaços de $\mathbb{R}^n$. Mas pode acontecer que W e V sejam dois subespaços de $\mathbb{R}^n$ e que $V \subseteq W$. Nestas condições dizemos que V é **subespaço** de W.

Teorema 6.12 *Sejam* W *e* V *dois subespaços de* $\mathbb{R}^n$*, sendo* $V \subseteq W$*. Então:*

1. $0 \leq \dim V \leq \dim W \leq n$.
2. $\dim V = \dim W$ *se, e só se,* $V = W$.

Demonstração

1. Seja $\mathcal{B}$ uma base de V. Então $\mathcal{B}$ é um conjunto linearmente independente de W. Pelo Teorema anterior, existe uma base de W, que contém os vectores de $\mathcal{B}$. Pela definição de dimensão,
$$0 \leq \dim V \leq \dim W \leq n.$$
2. Se $\dim V = \dim W$, então sendo $\mathcal{B}$ uma base de V, como $\mathcal{B}$ é um conjunto linearmente independente de W, podemos afirmar que existe uma base de W, $\mathcal{B}_1$, que contém os vectores de $\mathcal{B}$.

   Mas $\dim W = \dim V =$número de vectores de $\mathcal{B}$ e $\dim W =$número de vectores de $\mathcal{B}_1$. Então,
$$\mathcal{B}_1 = \mathcal{B}.$$
   Porque $V =< \mathcal{B} >$ e $W =< \mathcal{B}_1 >$, então $W = V$.

   Se $V = W$ é evidente que $\dim V = \dim W$. □

O teorema que vamos abordar é conhecido como Teorema das dimensões. Ele relaciona as dimensões do núcleo e da imagem de uma aplicação linear f, com a dimensão do conjunto de partida da aplicação f.

Teorema 6.13 *Seja* $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação linear. Então,*

$$\dim Nuc\ f + \dim Im\ f = n.$$

Demonstração Porque $Nuc\ f$ é um subespaço de $\mathbb{R}^n$,

$$\dim Nuc\ f \leq n.$$

Como $\mathbb{R}^n =< e_1, e_2, \ldots, e_n >$ em que $e_1, e_2, \ldots, e_n$ são os vectores canónicos de $\mathbb{R}^n$, então (Proposição 5.22) $Im\ f =< f(e_1), f(e_2), \ldots, f(e_n) >$. Donde,

$$\dim Im\ f \leq n.$$

1^o **caso)** Se $f = 0$, então $Nuc\ f = \mathbb{R}^n$ e $Im\ f = \{0_{\mathbb{R}^m}\}$. Pelo que o Teorema se verifica.

2^o **caso)** Suponhamos que $f \neq 0$. Seja $\mathcal{B} = \{v_1, \ldots, v_j\}$ uma base de $Nuc\ f$ e seja $\mathcal{B}_1 = \{v_1, \ldots, v_j, \ldots, v_n\}$ uma base de $\mathbb{R}^n$ que contem os vectores de $\mathcal{B}$. Porque $f \neq 0$, $\mathcal{B}_1 \setminus \mathcal{B} \neq \emptyset$. Atendendo à Proposição 5.22,

$$f(\mathbb{R}^n) = Im\ f =< f(v_1), \ldots, f(v_j), \ldots, f(v_n) > .$$

Como $f(v_i) = 0_{\mathbb{R}^m}$ se $v_i \in \mathcal{B}$, então

$$Im\ f =< f(v_{j+1}), \ldots, f(v_n) > .$$

Vejamos que estes vectores são linearmente independentes.

Se existissem escalares $a_{j+1}, \ldots, a_n$, não todos nulos, tais que

$$a_{j+1} f(v_{j+1}) + \ldots + a_n f(v_n) = 0_{\mathbb{R}^m},$$

porque f é aplicação linear,

$$f(a_{j+1} v_{j+1} + \ldots + a_n v_n) = 0_{\mathbb{R}^m}.$$

Mas isto significava que

$$a_{j+1} v_{j+1} + \ldots + a_n v_n \in Nuc\ f =< v_1, \ldots, v_j > .$$

Então, existiriam escalares $b_1, \ldots, b_j$ tais que

$$a_{j+1} v_{j+1} + \ldots + a_n v_n = b_1 v_1 + \ldots + b_j v_j.$$

Donde,

$$b_1 v_1 + \ldots + b_j v_j + (-a_{j+1}) v_{j+1} + \ldots + (-a_n) v_n = 0_{\mathbb{R}^n},$$

com algum dos $a_{j+1}, \ldots, a_n$ não nulo. Mas isto é impossível pois $\{v_1, \ldots, v_j, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$, logo linearmente independente. Assim, $\{f(v_{j+1}), \ldots, f(v_n)\}$ é linearmente independente e portanto uma base de $Im\ f$. Consequentemente,

$$\dim Im\ f = n - j.$$

Porque $\dim Nuc\ f = j$, temos que

$$\dim Nuc\ f + \dim Im\ f = n.$$

□

Vejamos agora um processo mais cómodo para determinar uma base a partir de um conjunto de geradores de um subespaço de $\mathbb{R}^n$.

Dado um conjunto $S = \{v_1, \ldots, v_k\}$ de geradores do subespaço W de $\mathbb{R}^n$,

- Construamos a matriz A cujas colunas são os vectores de S.
- Reduzamos a matriz A a uma sua forma de escada A'.
- Fixemos em A as colunas que correspondem às colunas de A' com pivots.
- Construamos com cada uma dessas colunas de A, o correspondente vector de $\mathbb{R}^n$.

Estes vectores formam uma base de W.

Observação É fácil ver que qualquer coluna de A' é combinação linear das colunas de pivots de A'.

Prova-se que, uma coluna de A' é combinação linear das colunas de pivots de A' se, e só se, a correspondente coluna de A é combinação linear das colunas de A que correspondem às colunas de pivots de A'.

Usando estes dois argumentos, temos a justificação da última afirmação do processo para determinar uma base de W.

Exemplo 6.14 *Seja* $W =< v_1, v_2, v_3, v_4, v_5 >$ *o subespaço de* $\mathbb{R}^4$ *tal que*

$$v_1 = (1, -2, 0, 3),\ v_2 = (2, -5, -3, 6),\ v_3 = (0, 1, 3, 0),$$

$$v_4 = (2, -1, 4, -7),\ v_5 = (5, -8, 1, 2).$$

Então,

$$A = \begin{array}{c} \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \\ \begin{array}{ccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ v_1 & v_2 & v_3 & v_4 & v_5 \end{array} \end{array} \begin{array}{c} \longrightarrow \\ l_2 \to (l_2 + 2l_1) \end{array} \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & -1 & 1 & 3 & 2 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \begin{array}{c} \longrightarrow \\ l_4 \to (l_4 - 3l_1) \end{array}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & -1 & 1 & 3 & 2 \\ 0 & -3 & 3 & 4 & 1 \\ 0 & 0 & 0 & -13 & -13 \end{bmatrix} \begin{array}{c} \longrightarrow \\ l_3 \to (l_3 - 3l_2) \end{array} \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & -1 & 1 & 3 & 2 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & -13 & -13 \end{bmatrix} \begin{array}{c} \longrightarrow \\ l_4 \to (l_4 - \frac{13}{5}l_3) \end{array} \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & -1 & 1 & 3 & 2 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = A'$$

em A'*, as colunas de pivots são a primeira, a segunda e a quarta. Estas colunas correspondem aos vectores (em* A*)* v_1, v_2, v_4*, então*

$$W =< v_1, v_2, v_4 > \qquad e \qquad \{v_1, v_2, v_4\} \text{ é uma base de } W.$$

6.4 Coordenadas em Relação a uma Base

Sabemos escrever um vector de $\mathbb{R}^n$ como combinação linear dos vectores canónicos de $\mathbb{R}^n$. Nesta secção veremos que os coeficientes desta combinação linear são distintos se mudarmos os vectores canónicos para outros vectores que formem uma base de $\mathbb{R}^n$.

Definição 6.15 *Seja $\mathcal{B} = \{v_1, v_2, \ldots, v_k\}$ uma base ordenada (isto é, a ordem dos vectores em $\mathcal{B}$ é fixa) do subespaço W de $\mathbb{R}^n$. Se a única forma de escrever $u \in W$, como combinação linear dos vectores de $\mathcal{B}$ é*

$$u = a_1 v_1 + a_2 v_2 + \ldots + a_k v_k,$$

dizemos que as **coordenadas** *de u na base $\mathcal{B}$ são*

$$a_1, a_2, \ldots, a_k$$

ou que,

$$(a_1, a_2, \ldots, a_k)$$

é o $k-$uplo de coordenadas de u na base $\mathcal{B}$. Quando não houver lugar a confusão, diremos simplesmente coordenadas de u na base $\mathcal{B}$.

Observação Normalmente escrevemos

$$x = (x_1, x_2, \ldots, x_n)$$

para designar um vector de $\mathbb{R}^n$. Como $\{e_1, e_2, \ldots, e_n\}$ é a base canónica de $\mathbb{R}^n$ e

$$(x_1, x_2, \ldots, x_n) = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n,$$

então $(x_1, x_2, \ldots, x_n)$ são as coordenadas de x na base canónica de $\mathbb{R}^n$.

Se nada for dito, sempre que nos derem um vector $(x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$, suporemos que são as coordenadas do vector na base canónica de $\mathbb{R}^n$.

Exemplo 6.16 *Seja $W =< (1, 2, 0), (0, 1, 2) >$ um subespaço de $\mathbb{R}^3$. Facilmente se vê que*

$$\mathcal{B} = \{(1, 2, 0), (0, 1, 2)\}$$

é uma base de W (o conjunto $\mathcal{B}$ é linearmente independente e gera W).

Vejamos se $u = (2, 0, -8)$ é um vector de W e, se for, quais as suas coordenadas na base $\mathcal{B}$.

O que pretendemos é encontrar, se existirem, escalares a_1 e a_2 tais que

$$u = (2, 0, -8) = a_1(1, 2, 0) + a_2(0, 1, 2) = (a_1, 2a_1 + a_2, 2a_2).$$

Donde,

$$\begin{cases} a_1 = 2 \\ 2a_1 + a_2 = 0 \\ 2a_2 = -8 \end{cases}$$

ou seja, $a_1 = 2$, $a_2 = -4$. Portanto,

$$(2, -4)$$

é o 2−uplo de coordenadas de u na base $\mathcal{B}$.

Observação Se no Exemplo anterior pedissem as coordenadas de u na base $\{(0, 1, 2), (1, 2, 0)\}$ (repare que esta base tem os vectores de $\mathcal{B}$ trocados) de W, seria o 2−uplo

$$(-4, 2).$$

ATENÇÃO: É muito importante a ordenação da base.

6.5 Matriz de uma Aplicação Linear

Na secção 5.3 aprendemos a construir a matriz canónica de uma aplicação linear. Aqui veremos que a matriz canónica não é mais do que a matriz da aplicação linear relativamente às bases canónicas.

Definição 6.17 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear e sejam $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base de $\mathbb{R}^n$ e $\mathcal{B}' = \{v'_1, \ldots, v'_m\}$ uma base de $\mathbb{R}^m$. Chamamos* **matriz de f em relação às bases $\mathcal{B}$ e $\mathcal{B}'$**, *e denotamo-la por*

$$M(f; \mathcal{B}, \mathcal{B}'),$$

a matriz $A = [a_{ij}]$ de tipo $m \times n$, cuja i−ésima coluna é o m−uplo de coordenadas de $f(v_i)$ na base $\mathcal{B}'$, $i = 1, \ldots, n$. Portanto,

$$f(v_i) = a_{1i}v'_1 + \ldots + a_{mi}v'_m.$$

Exemplo 6.18 1. *Seja*

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\mapsto (x + 2y, 3y, -x + y) \end{aligned}$$

uma aplicação linear e sejam $\mathcal{B} = \{(2, 1), (0, 3)\}$ uma base de $\mathbb{R}^2$ e $\mathcal{B}' = \{(1, 1, 1), (0, 1, 0), (0, 0, 1)\}$ uma base de $\mathbb{R}^3$. Construamos a $M(f; \mathcal{B}, \mathcal{B}')$.

Ora

$$f(2,1) = (4,3,-1) = 4(1,1,1) - 1(0,1,0) - 5(0,0,1)$$
$$f(0,3) = (6,9,3) = 6(1,1,1) + 3(0,1,0) - 3(0,0,1)$$

Assim,

$$M(f;\mathcal{B},\mathcal{B}') = \begin{bmatrix} 4 & 6 \\ -1 & 3 \\ -5 & -3 \end{bmatrix}$$

Repare que a matriz canónica de f *(matriz de* f *em relação à base canónica de* $\mathbb{R}^2$ *e à base canónica de* $\mathbb{R}^3$*) é*

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

pois

$$f(1,0) = (1,0,-1)$$

e

$$f(0,1) = (2,3,1).$$

2. *Considere a aplicação linear*

$$id_{\mathbb{R}^2} : \ \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$(x,y) \ \mapsto \ (x,y)$$

e sejam $\mathcal{B} = \{(2,1),(0,3)\}$ *uma base de* $\mathbb{R}^2$ *e* $\mathcal{B}_1$ *a base canónica de* $\mathbb{R}^2$*. Então,*

$$M(id_{\mathbb{R}^2};\mathcal{B},\mathcal{B}_1) = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$$

pois

$$id_{\mathbb{R}^2}(2,1) = (2,1) = 2(1,0) + 1(0,1)$$

e

$$id_{\mathbb{R}^2}(0,3) = (0,3) = 0(1,0) + 3(0,1),$$

e

$$M(id_{\mathbb{R}^2};\mathcal{B}_1,\mathcal{B}) = \begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

pois

$$id_{\mathbb{R}^2}(1,0) = (1,0) = \frac{1}{2}(2,1) - \frac{1}{6}(0,3)$$

e

$$id_{\mathbb{R}^2}(0,1) = (0,1) = 0(2,1) + \frac{1}{3}(0,3).$$

6.6 Matriz Mudança de Base

Sejam $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear, $\mathcal{B}$ e $\mathcal{B}_1$ bases de $\mathbb{R}^n$ e $\mathcal{B}'$ uma base de $\mathbb{R}^m$. Duas perguntas se podem colocar:

- Qual a relação entre $M(id_{\mathbb{R}^n}; \mathcal{B}, \mathcal{B}_1)$ e $M(id_{\mathbb{R}^n}; \mathcal{B}_1, \mathcal{B})$?
- Qual a relação entre $M(f; \mathcal{B}, \mathcal{B}')$ e M_f?

São estas duas perguntas que iremos estudar a seguir.

Definição 6.19 *Sejam $\mathcal{B}$ e $\mathcal{B}_1$ duas bases de $\mathbb{R}^n$. Chamamos* **matriz de mudança de base** $(\mathcal{B}, \mathcal{B}_1)$ *à matriz*

$$M(id_{\mathbb{R}^n}; \mathcal{B}, \mathcal{B}_1).$$

Sejam $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear, $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base de $\mathbb{R}^n$ e $\mathcal{B}' = \{v'_1, \ldots, v'_m\}$ uma base de $\mathbb{R}^m$. Sendo u um vector de $\mathbb{R}^n$, vejamos como calcular $f(u)$ usando $M(f; \mathcal{B}, \mathcal{B}') = [a_{ij}]$.

Seja $(b_1, \ldots b_n)$ o $n-$uplo de coordenadas de u na base $\mathcal{B}$ (base de $\mathbb{R}^n$). Então,

$$u = b_1 v_1 + \ldots + b_n v_n.$$

Atendendo à Definição 6.17

$$\begin{aligned} f(v_1) &= a_{11} v'_1 + \ldots + a_{m1} v'_m \\ &\vdots \\ f(v_n) &= a_{1n} v'_1 + \ldots + a_{mn} v'_m. \end{aligned}$$

Então, usando a linearidade de f temos que

$$\begin{aligned} f(u) &= f(b_1 v_1 + \ldots + b_n v_n) \\ &= b_1 f(v_1) + \ldots + b_n f(v_n) \\ &= b_1(a_{11} v'_1 + \ldots + a_{m1} v'_m) + \ldots + b_n(a_{1n} v'_1 + \ldots + a_{mn} v'_m) \\ &= (b_1 a_{11} + \ldots + b_n a_{1n}) v'_1 + \ldots + (b_1 a_{m1} + \ldots + b_n a_{mn}) v'_m \end{aligned}$$

ou seja, as coordenadas de $f(u)$ na base $\mathcal{B}'$ são

$$(\alpha_1, \ldots, \alpha_m)$$

em que

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = M(f; \mathcal{B}, \mathcal{B}') \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

O que acabámos de demonstrar foi o resultado seguinte:

Proposição 6.20 *Sejam* $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *uma aplicação linear,* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *uma base de* $\mathbb{R}^n$ *e* $\mathcal{B}' = \{v'_1, \ldots, v'_m\}$ *uma base de* $\mathbb{R}^m$*. Se* $(b_1, \ldots, b_n)$ *for o* $n-$*uplo de coordenadas de um vector* u *de* $\mathbb{R}^n$*, na base* $\mathcal{B}$*, então o* $m-$*uplo de coordenadas de* $f(u)$ *na base* $\mathcal{B}'$ *é* $(\alpha_1, \ldots, \alpha_m)$ *tal que*

$$M(f; \mathcal{B}, \mathcal{B}') \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}.$$

Exemplo 6.21 *Seja* $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *a aplicação linear que tem como matriz relativamente às bases* $\mathcal{B} = \{(1,1), (0,1)\}$ *de* $\mathbb{R}^2$ *e* $\mathcal{B}' = \{(1,1,1), (0,1,1), (0,0,1)\}$ *de* $\mathbb{R}^3$*,*

$$M(f; \mathcal{B}, \mathcal{B}') = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Determinemos $f(2,-1)$*.*

Como nada dissemos, supomos que $(2,-1)$ *são as coordenadas do vector de* $\mathbb{R}^2$ *na sua base canónica.*

Porque

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

então

$$(2,-3)$$

são as coordenadas de $(2,-1)$ *na base* $\mathcal{B}$*.*

Ora, usando a Proposição 6.20*,*

$$M(f; \mathcal{B}, \mathcal{B}') \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ -1 \end{bmatrix}.$$

Donde,

$$(-4, 2, -1)$$

são as coordenadas de $f(2,-1)$ *na base* $\mathcal{B}'$*. Assim,*

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

Corolário 6.22 *Sejam* $\mathcal{B}$ *e* $\mathcal{B}_1$ *duas bases de* $\mathbb{R}^n$*. Se* $(b_1, \ldots, b_n)$ *for o o* $n-$*uplo de coordenadas de um vector* u *de* $\mathbb{R}^n$*, na base* $\mathcal{B}$*, então o* $n-$*uplo de coordenadas de* u *na base* $\mathcal{B}_1$ *é* $(\alpha_1, \ldots, \alpha_n)$ *tal que*

$$M(id_{\mathbb{R}^n}; \mathcal{B}, \mathcal{B}_1) \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Observação Usando as hipóteses do Corolário 6.22, temos que

$$M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B})M(id_{\mathbb{R}^n};\mathcal{B},\mathcal{B}_1)\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B})\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = I_n \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Porque esta igualdade se obtém qualquer que seja $(b_1,\ldots,b_n)$, então,

$$M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B})M(id_{\mathbb{R}^n};\mathcal{B},\mathcal{B}_1) = I_n.$$

Pelo Lema 4.23,

$$M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B}) = M(id_{\mathbb{R}^n};\mathcal{B},\mathcal{B}_1)^{-1}.$$

Estamos agora em condições de responder à pergunta sobre a relação que existe entre as matrizes da mesma aplicação linear, mas relativamente a bases distintas.

Teorema 6.23 *Sejam $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ uma aplicação linear, $\mathcal{B}$ e $\mathcal{B}_1$ duas bases de $\mathbb{R}^n$ e $\mathcal{B}'$ e $\mathcal{B}'_1$ duas bases de $\mathbb{R}^m$. Então,*

$$M(f;\mathcal{B}_1,\mathcal{B}'_1) = M(id_{\mathbb{R}^m};\mathcal{B}',\mathcal{B}'_1)M(f;\mathcal{B},\mathcal{B}')M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B}).$$

Demonstração Queremos mostrar que

$$M(f;\mathcal{B}_1,\mathcal{B}'_1) = M(id_{\mathbb{R}^m};\mathcal{B}',\mathcal{B}'_1)M(f;\mathcal{B},\mathcal{B}')M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B}) \qquad (*)$$

ou seja, que a matriz da aplicação linear $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ relativamente às bases $\mathcal{B}_1$ de $\mathbb{R}^n$ e $\mathcal{B}'_1$ de $\mathbb{R}^m$ e a matriz da mesma aplicação linear f mas relativamente às bases $\mathcal{B}$ de $\mathbb{R}^n$ e $\mathcal{B}'$ de $\mathbb{R}^m$ se relacionam.

Construamos o diagrama destas aplicações da seguinta forma:

No topo do diagrama colocamos a aplicação $f:\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e mencionamos que é relativamente às bases $\mathcal{B}_1$ de $\mathbb{R}^n$ e $\mathcal{B}'_1$ de $\mathbb{R}^m$ (bases da matriz que surge do lado esquerdo da igualdade $(*)$). Isto é,

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "f"] & \mathbb{R}^m \\ \mathcal{B}_1 \arrow[r, mapsto] & \mathcal{B}'_1 \end{tikzcd}$$

Na parte final do diagrama colocamos a mesma aplicação do topo do diagrama mas relativamente às bases em que surge a matriz desta aplicação do lado direito da igualdade $(*)$. Isto é,

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\quad f \quad} & \mathbb{R}^m \\ \mathcal{B}_1 & \longmapsto & \mathcal{B}'_1 \end{array}$$

$$\begin{array}{ccc} \mathcal{B} & \longmapsto & \mathcal{B}' \\ \mathbb{R}^n & \xrightarrow[\quad f \quad]{} & \mathbb{R}^m \end{array}$$

Do lado direito do diagrama colocamos a aplicação identidade de $\mathbb{R}^m$, em que o sentido da aplicação é do final do diagrama para o seu topo. Do lado esquerdo do diagrama colocamos a aplicação identidade de $\mathbb{R}^n$, em que o sentido da aplicação é do topo do diagrama para o seu final. Isto é,

$$\begin{tikzcd} \mathbb{R}^n \ (\mathcal{B}_1) \arrow[r, "f"] \arrow[d, "id_{\mathbb{R}^n}"'] & \mathbb{R}^m \ (\mathcal{B}'_1) \\ \mathbb{R}^n \ (\mathcal{B}) \arrow[r, "f"'] & \mathbb{R}^m \ (\mathcal{B}') \arrow[u, "id_{\mathbb{R}^m}"'] \end{tikzcd}$$

Seja u um vector de $\mathbb{R}^n$ em que $(b_1, \ldots, b_n)$ é o $n-$uplo de coordenadas de u na base $\mathcal{B}$ e $(c_1, \ldots, c_n)$ é o $n-$uplo de coordenadas de u na base $\mathcal{B}_1$. Sejam $(\alpha_1, \ldots, \alpha_m)$ o $m-$uplo de coordenadas de $f(u)$ na base $\mathcal{B}'$ e $(\beta_1, \ldots, \beta_m)$ o $m-$uplo de coordenadas de $f(u)$ na base $\mathcal{B}'_1$.

Usando a Proposição 6.20, temos que

$$M(f; \mathcal{B}_1, \mathcal{B}'_1) \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}.$$

Por outro lado, usando o Corolário 6.22 e a Proposição 6.20,

$$M(id_{\mathbb{R}^m}; \mathcal{B}', \mathcal{B}'_1) M(f; \mathcal{B}, \mathcal{B}') M(id_{\mathbb{R}^n}; \mathcal{B}_1, \mathcal{B}) \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} =$$

$$M(id_{\mathbb{R}^m}; \mathcal{B}', \mathcal{B}'_1) M(f; \mathcal{B}, \mathcal{B}') \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} =$$

$$M(id_{\mathbb{R}^m}; \mathcal{B}', \mathcal{B}_1') \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}.$$

Donde o resultado. $\square$

Exemplo 6.24 1. *Determinemos a matriz da aplicação linear*

$$f: \ \mathbb{R}^3 \longrightarrow \mathbb{R}^2$$

em relação às bases $\mathcal{B}_1 = \{(1,0,0),(0,0,1),(0,1,0)\}$ *de* $\mathbb{R}^3$ *e* $\mathcal{B}_1' = \{(0,1),(1,1)\}$ *de* $\mathbb{R}^2$*, sabendo que*

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

Vamos construir o diagrama como descrito na demonstração do Teorema 6.23.

$$\begin{tikzcd}
\mathbb{R}^3 \arrow[rrr, "f"] \arrow[ddd, "id_{\mathbb{R}^3}"'] & & & \mathbb{R}^2 \\
& \mathcal{B}_1 \arrow[r, mapsto] & \mathcal{B}_1' & \\
& b.c. \arrow[r, mapsto] & b.c. & \\
\mathbb{R}^3 \arrow[rrr, "f"'] & & & \mathbb{R}^2 \arrow[uuu, "id_{\mathbb{R}^2}"']
\end{tikzcd}$$

(*em que b.c. designa a base canónica*) *temos que,*

$$M(f; \mathcal{B}_1, \mathcal{B}_1') = M(id_{\mathbb{R}^2}; b.c., \mathcal{B}_1') M_f M(id_{\mathbb{R}^3}; \mathcal{B}_1, b.c.).$$

Porque

$$M(id_{\mathbb{R}^2}; b.c., \mathcal{B}_1') = M(id_{\mathbb{R}^2}; \mathcal{B}_1', b.c.)^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$$

temos que

$$M(f; \mathcal{B}_1, \mathcal{B}_1') = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}.$$

2. *Determinemos a matriz da aplicação linear*

$$f: \ \mathbb{R}^2 \longrightarrow \mathbb{R}^3$$

em relação às bases $\mathcal{B} = \{(0,1),(1,1)\}$ *de* $\mathbb{R}^2$ *e à base canónica de* $\mathbb{R}^3$*, sabendo que a matriz canónica de* f *é*

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

Vamos construir o diagrama como descrito na demonstração do Teorema 6.23.

$$\begin{tikzcd}
\mathbb{R}^2 \arrow[rrr, "f"] \arrow[ddd, "id_{\mathbb{R}^2}"'] & \mathcal{B} \arrow[r, mapsto] & b.c. & \mathbb{R}^3 \\
& & & \\
& & & \\
\mathbb{R}^2 \arrow[rrr, "f"'] & b.c. \arrow[r, mapsto] & b.c. & \mathbb{R}^3 \arrow[uuu, "id_{\mathbb{R}^3}"']
\end{tikzcd}$$

(*em que b.c. designa a base canónica*) *temos que,*

$$M(f;\mathcal{B},b.c.) = M(id_{\mathbb{R}^3};b.c.,b.c.)M_f M(id_{\mathbb{R}^2};\mathcal{B},b.c.).$$

Porque

$$M(id_{\mathbb{R}^3};b.c.,b.c.) = I_3$$

temos que

$$M(f;\mathcal{B},b.c.) = I_3 \begin{bmatrix} 1 & 2 \\ -5 & 0 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & -5 \\ 4 & 7 \end{bmatrix}.$$

Capítulo 7

DIAGONALIZAÇÃO

Neste capítulo voltamos a abordar o tema dos valores e vectores próprios de uma matriz de ordem n e veremos uma aplicação, à geometria, destes conceitos.

7.1 Diagonalização de Matrizes Quadradas

Dada uma aplicação linear f de $\mathbb{R}^n$ em $\mathbb{R}^n$, usando matrizes mudança de base, podemos em certas ocasiões determinar uma base de $\mathbb{R}^n$, na qual a aplicação linear f tenha como sua matriz, uma matriz diagonal.

Definição 7.1 *Sejam A e B duas matrizes de ordem n. Dizemos que A* **é semelhante a** *B, se existe uma matriz invertível P, de ordem n, tal que*

$$P^{-1}AP = B.$$

Observação Se A é semelhante a B, existe P invertível tal que $P^{-1}AP = B$. Mas então, $PBP^{-1} = A$ e podemos dizer que B é semelhante a A. Muitas vezes diz-se, simplesmente que A e B **são semelhantes**.

Exemplo 7.2 *As matrizes $A = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}$ são semelhantes, porque a matriz $P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ é invertível e*

$$P^{-1}AP = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix} = B.$$

Teorema 7.3 *Sejam A e B duas matrizes de ordem n. A e B são semelhantes se, e só se, existem bases em relação às quais as matrizes representam a mesma aplicação linear.*

Demonstração Suponhamos que as duas matrizes representam a mesma aplicação linear $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$, sendo

$$A = M(f;\mathcal{B},\mathcal{B}) \qquad \text{e} \qquad B = M(f;\mathcal{B}_1,\mathcal{B}_1)$$

em que $\mathcal{B}$ e $\mathcal{B}_1$ são duas bases de $\mathbb{R}^n$. Pelo Teorema 6.23,

$$B = M(id_{\mathbb{R}^n};\mathcal{B},\mathcal{B}_1)AM(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B}).$$

Porque $M(id_{\mathbb{R}^n};\mathcal{B},\mathcal{B}_1) = M(id_{\mathbb{R}^n};\mathcal{B}_1,\mathcal{B})^{-1}$, então A e B são semelhantes.

Reciprocamente, se $B = P^{-1}AP$ para alguma matriz P de ordem n, seja $\mathcal{B}$ a base de $\mathbb{R}^n$ tal que

$$P = M(id_{\mathbb{R}^n};\mathcal{B},b.c.)$$

em que b.c. é a base canónica de $\mathbb{R}^n$. Pela Proposição 5.9, seja $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ a aplicação linear tal que $M_f = A$. Então,

$$B = M(id_{\mathbb{R}^n};b.c.,\mathcal{B})M_f M(id_{\mathbb{R}^n};\mathcal{B},b.c.).$$

Pelo Teorema 6.23,

$$B = M(f;\mathcal{B},\mathcal{B}),$$

ou seja, A e B representam a mesma aplicação linear f. □

Exemplo 7.4 *Usando o* Exemplo 7.2, *em que* $A = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$ *e* $B = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}$ *e supondo que A é a matriz canónica de uma aplicação linear de $\mathbb{R}^2$ em $\mathbb{R}^2$, então a aplicação linear é*

$$\begin{aligned} f: \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\mapsto (x-y,-x+3y) \end{aligned}$$

Pelo Teorema 7.3, *porque A e B são semelhantes, existe uma base $\mathcal{B}$ de $\mathbb{R}^2$, tal que*

$$B = M(f;\mathcal{B},\mathcal{B}).$$

pela demonstração deste Teorema, *a base $\mathcal{B}$ determina-se usando uma matriz P tal que $B = P^{-1}AP$.*

Pelo Exemplo 7.2, *se* $P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, *então $B = P^{-1}AP$.*

Pelo que, sendo $P = M(id_{\mathbb{R}^n};\mathcal{B},b.c.)$, vem que

$$\mathcal{B} = \{(1,0),(2,1)\}$$

é a base de $\mathbb{R}^2$ em relação à qual B representa a aplicação f.

Proposição 7.5 *Matrizes semelhantes têm*

1. *o mesmo determinante.*
2. *o mesmo polinómio característico e portanto os mesmos valores próprios com as mesmas multiplicidades algébricas.*

Definição 7.6 *Seja A uma matriz de ordem n. A diz-se* **diagonalizável** *se A é semelhante a uma matriz diagonal, isto é, se existe uma matriz P invertível e uma matriz diagonal D tais que*

$$A = P^{-1}DP.$$

A matriz P diz-se a matriz **diagonalizante** *de A.*

Observação Na definição anterior poderíamos dizer que A é diagonalizável se existe uma matriz P invertível e uma matriz diagonal D tais que

$$PAP^{-1} = D.$$

Proposição 7.7 *Sejam A uma matriz de ordem n diagonalizável, P uma matriz invertível e $D = \begin{bmatrix} d_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_n \end{bmatrix}$ uma matriz diagonal tais que $P^{-1}DP = A$. Então,*

1. *cada d_i, $i = \{1, \dots, n\}$, é um valor próprio de A.*
2. *a matriz P^{-1} tem como suas colunas, n vectores próprios de A, linearmente independentes, cada um deles associado,respectivamente, a $d_1, \dots, d_n$.*

Demonstração Porque $P^{-1}DP = A$ então $AP^{-1} = P^{-1}D$. Seja C_1 a primeira coluna de P^{-1} (P^{-1} é invertível, então C_1 não é uma coluna nula). Mas a primeira coluna de $P^{-1}D$ é d_1C_1, então,

$$AC_1 = d_1C_1,$$

ou seja, C_1 é um vector próprio de A associado ao valor próprio d_1.

O mesmo acontece com as outras colunas de P^{-1}.

Atendendo ao processo descrito depois do Teorema 6.13 e porque P^{-1} é invertível, podemos afirmar que a matriz P^{-1} tem como suas colunas, n vectores próprios de A, linearmente independentes, cada um deles associado, respectivamente, a $d_1, \dots, d_n$. □

Definição 7.8 *Seja A uma matriz de ordem n e α um valor próprio de A, chamamos* **multiplicidade geométrica** *de α, e denotamos por*

$$mg(\alpha),$$

à dimensão do subespaço próprio associado ao valor próprio α, isto é,

$$mg(\alpha) = \dim M_\alpha.$$

Proposição 7.9 *Sejam A uma matriz de ordem n e α um valor próprio de A. Então,*

$$1 \leq mg(\alpha) \leq ma(\alpha).$$

Demonstração Suponhamos que $mg(\alpha) = s$ e seja $\{v_1, \ldots, v_s\}$ uma base de M_α, temos $s \geq 1$ pois, por definição de vector próprio, M_α não é o subespaço nulo . Pelo Teorema 6.9, existem vectores $u_{s+1}, \ldots, u_n$ de $\mathbb{R}^n$ tais que

$$\mathcal{B} = \{v_1, \ldots, v_s, u_{s+1}, \ldots, u_n\}$$

é uma base de $\mathbb{R}^n$. Pela Proposição 5.9, seja f a aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}^n$ tal que $M_f = A$.

Porque $Av_i = \alpha v_i$, $i = 1, \ldots, s$, então

$$f(v_i) = \alpha v_i,$$

$i = 1, \ldots, s$. Consequentemente,

$$B = M(f;\mathcal{B},\mathcal{B}) = \begin{array}{cc} & \overline{\quad s \quad} \\ s & \left[\begin{array}{ccc|c} \alpha & \ldots & 0 & \\ \vdots & \ddots & \vdots & * \\ 0 & \ldots & \alpha & \\ \hline & & & \\ & 0 & & C \\ & & & \end{array}\right] \end{array}$$

Como B é semelhante a A (Teorema 7.3) e α é valor próprio de B com multiplicidade algébrica $\geq s$, então

$$ma(\alpha) \geq s = mg(\alpha).$$

□

Proposição 7.10 *Se $v_1, \ldots, v_k$ são vectores próprios de A associados aos valores próprios distintos $\alpha_1, \ldots \alpha_k$, então o conjunto $\{v_1, \ldots, v_k\}$ é linearmente independente.*

Demonstração Suponhamos que $S = \{v_1, \ldots, v_k\}$ é linearmente dependente e seja $S' = \{v_{i_1}, \ldots, v_{i_r}\}$ um subconjunto de S, linearmente independente, maximal. Porque $S \backslash S' \neq \emptyset$, seja $v_j \in S \backslash S'$. Então v_j é combinação linear dos vectores de S', ou seja, existem $a_{i_1}, \ldots, a_{i_r}$ escalares tais que

$$v_j = a_{i_1} v_{i_1} + \ldots + a_{i_r} v_{i_r}.$$

Se pensarmos nos vectores como matrizes coluna, então

$$Av_j = A(a_{i_1} v_{i_1} + \ldots + a_{i_r} v_{i_r}) = a_{i_1} Av_{i_1} + \ldots + a_{i_r} Av_{i_r}.$$

Donde, $\alpha_j v_j = a_{i_1} \alpha_{i_1} v_{i_1} + \ldots + a_{i_r} \alpha_{i_r} v_{i_r}$, ou seja,

$$0 = a_{i_1}(\alpha_{i_1} - \alpha_j) v_{i_1} + \ldots + a_{i_r}(\alpha_{i_r} - \alpha_j) v_{i_r}.$$

Como $\alpha_1, \ldots \alpha_k$ são distintos e S' é linearmente independente então $a_{i_1} = \ldots = a_{i_r} = 0$ e $v_j = 0$. Mas isto é impossível pois v_j é vector próprio de A, logo não nulo. □

Teorema 7.11 *Seja A uma matriz de ordem n. A é diagonalizável se, e só se, a soma das multiplicidades geométricas dos valores próprios de A é n.*

Demonstração Suponhamos que A é diagonalizável, então existe uma matriz P invertível tal que

$$A = P^{-1}DP$$

em que D é matriz diagonal. Sendo $D = \begin{bmatrix} d_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & d_n \end{bmatrix}$ então $d_1, \ldots, d_n$ são os valores próprios de D, logo os de A.

Sejam $v_1, \ldots, v_n$ as colunas da matriz P^{-1} (já sabemos que são vectores próprios de A), então

$$P^{-1}D = \begin{bmatrix} d_1v_1 & d_2v_2 & \ldots & d_nv_n \end{bmatrix}.$$

De $A = P^{-1}DP$ vem que $AP^{-1} = P^{-1}D$ e daqui sai que

$$A[v_i] = [d_iv_i]$$

ou seja, v_i é vector próprio de A associado ao valor próprio d_i.

Como P^{-1} é invertível, então $\{v_1, \ldots, v_n\}$ é uma base de $\mathbb{R}^n$, formada por vectores próprios de A. Se $d_{i_1}, \ldots, d_{i_r}$ são os valores próprios distintos de A, porque

$$ma(d_{i_j}) \geq mg(d_{i_j})$$

(Proposição 7.9), então

$$\sum_{j=1}^{r} mg(d_{i_j}) \leq \sum_{j=1}^{r} ma(d_{i_j}) = n$$

ou seja, o conjunto formado pelos vectores de uma base de cada M_{α_i} tem no máximo n vectores. Como $\mathcal{B} = \{v_1, \ldots, v_n\}$ tem n vectores, então

$$\sum_{j=1}^{r} mg(d_{i_j}) = n.$$

Reciprocamente, sejam $\alpha_1, \ldots \alpha_k$ os valores próprios de A e seja

$$\mathcal{B}_i = \{v_{i_1}, \ldots, v_{i_{r_i}}\}$$

uma base do subespaço próprio M_{α_i}, $i = 1, \ldots, k$. Consideremos o conjunto $\mathcal{B}$ formado pelos vectores de cada $\mathcal{B}_i$, $i = 1, \ldots, k$.

Se $\mathcal{B}$ fosse linearmente dependente, existiriam escalares, $a_{11}, \ldots, a_{1r_1}, \ldots, a_{k1}, \ldots, a_{kr_k}$, não todos nulos, tais que

$$0 = a_{11}v_{11} + \ldots + a_{1r_1}v_{1r_1} + \ldots + a_{k1}v_{k1} + \ldots + a_{kr_k}v_{kr_k}.$$

Sendo

$$t_i = a_{i1}v_{i1} + \ldots + a_{ir_i}v_{ir_i}$$

então t_i pertence a M_{α_i}. Donde,

$$0 = t_1 + \ldots + t_k.$$

Os vectores $t_1, \ldots, t_k$ pertencem a subespaços próprios distintos, então, pela Proposição 7.10, $t_1 = 0, \ldots, t_k = 0$. Mas então,

$$0 = t_i = a_{i1}v_{i1} + \ldots + a_{ir_i}v_{ir_i}$$

e porque $\mathcal{B}_i$ é uma base de M_{α_i}, $a_{i1} = \ldots = a_{ir_i} = 0$, para $i = 1, \ldots, k$. Ou seja, temos uma situação impossível. Portanto, $\mathcal{B}$ é linearmente independente.

Porque $\sum_{i=1}^{k} mg(\alpha_i) = n$, então $\mathcal{B}$ tem n vectores. Consequentemente, $\mathcal{B}$ é uma base de $\mathbb{R}^n$.

Se $P^{-1} = \begin{bmatrix} v_{11} \cdots v_{1r_1} \cdots v_{k1} \cdots v_{kr_k} \end{bmatrix}$, então

$$\begin{aligned} AP^{-1} &= \begin{bmatrix} \alpha_1 v_{11} \cdots \alpha_1 v_{1r_1} \cdots \alpha_k v_{k1} \cdots \alpha_k v_{kr_k} \end{bmatrix} \\ &= P^{-1} \begin{bmatrix} \alpha_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \alpha_k \end{bmatrix}. \end{aligned}$$

Donde, $A = P^{-1}DP$ com $D = \begin{bmatrix} \alpha_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \alpha_k \end{bmatrix}$. □

Exemplo 7.12 *Vejamos se* $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$ *é diagonalizável e se for, determinemos uma matriz diagonalizante.*

Calculemos os valores próprios de A:

$$\det(\lambda I_3 - A) = \begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} =$$

$$= \lambda(\lambda - 2)(\lambda - 3) + 2(\lambda - 2) = (\lambda - 2)(\lambda^2 - 3\lambda + 2) = (\lambda - 2)^2(\lambda - 1).$$

Então 2 *e* 1 *são os valores próprios de A e*

$$ma(2) = 2, \qquad ma(1) = 1.$$

Vejamos os subespaços próprios. Por definição,

$$M_2 = \{(x,y,z) \in \mathbb{R}^3 : (2I_3 - A)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$$

$$2I_3 - A = \begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \underset{\substack{l_2 \to (l_2 + \frac{1}{2}l_1) \\ l_3 \to (l_3 + \frac{1}{2}l_1)}}{\longrightarrow} \begin{bmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$M_2 = \{(x,y,z) \in \mathbb{R}^3 : x = -z\} = \langle (-1,0,1),(0,1,0) \rangle.$$

Como $\{(-1,0,1),(0,1,0)\}$ *é linearmente independente, pois*

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

então $\{(-1,0,1),(0,1,0)\}$ *é uma base de* M_2*.*

$$M_1 = \{(x,y,z) \in \mathbb{R}^3 : (I_3 - A)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$$

$$I_3 - A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -1 & -1 \\ -1 & 0 & -2 \end{bmatrix} \underset{\substack{l_2 \to (l_2 + l_1) \\ l_3 \to (l_3 + l_1)}}{\longrightarrow} \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$M_1 = \{(x,y,z) \in \mathbb{R}^3 : x = -2z, \quad y = z\} = \langle (-2,1,1) \rangle.$$

Como $(-2,1,1) \neq (0,0,0)$ *então é linearmente independente, logo* $\{(-2,1,1)\}$ *é uma base de* M_1*. Então,* $mg(2) = 2$*,* $mg(1) = 1$*. Assim,*

$$mg(2) + mg(1) = 3 = \text{ordem A}$$

e A é diagonalizável (Teorema 7.11)*. Sendo*

$$P^{-1} = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \qquad \text{(}\textit{matriz formada pelos vectores das bases de } M_2 \textit{ e } M_1\text{)}$$

então, P diagonaliza A.

Observação No exemplo anterior,

$$PAP^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Se tivessemos escolhido a matriz $Q^{-1} = \begin{bmatrix} -1 & -2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (matriz formada pelos vectores das bases de M_2 e M_1, com uma ordem diferente da que tem P^{-1}) então, Q também diagonalizaria A, mas $QAQ^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(Repare a que subespaço pertence cada vector que é uma coluna de Q^{-1} e de P^{-1}).

7.2 Classificação de Cónicas de $\mathbb{R}^2$

Nesta secção vamos ver uma aplicação da diagonalização de matrizes.

7.2.1 Formas Quadráticas de $\mathbb{R}^2$

Além da definição de forma quadrática vamos aprender a diagonalizá-las.

Definição 7.13 *Chama-se* **forma quadrática** *de* $\mathbb{R}^2$ *a toda a aplicação do tipo*

$$\begin{array}{rccl} Q: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & Ax^2 + Bxy + Cy^2. \end{array}$$

Dizemos que Q *é uma* **forma quadrática diagonal** *se*

$$Q(x,y) = Ax^2 + Cy^2.$$

Observação $Q(x,y)$ pode ser escrita na forma matricial da seguinte forma

$$Q(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

A matriz $M = \begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}$ é a **matriz associada à forma quadrática** Q.

Repare que Q é forma quadrática diagonal se, e só se, a matriz associada à forma quadrática Q for diagonal.

Observação Repare que se M é uma matriz associada a uma forma quadrática então M é uma matriz simétrica.

Exemplo 7.14 *Sendo*

$$\begin{array}{rccl} Q: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & x^2 + 3xy - y^2 \end{array}$$

então a matriz associada à forma quadrática Q é

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

Vejamos como se diagonaliza uma forma quadrática Q cuja matriz associada é

$$M = \begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}.$$

-Em primeiro lugar calculemos os valores próprios de M.

Eles são obtidos através do polinómio característico de M que é,

$$p_M(\lambda) = \begin{vmatrix} \lambda - A & -\frac{B}{2} \\ -\frac{B}{2} & \lambda - C \end{vmatrix} = (\lambda - A)(\lambda - C) - \frac{B^2}{4} =$$

$$= \lambda^2 - (A+C)\lambda + (AC - \frac{B^2}{4}).$$

Então, $p_M(\lambda) = 0$ se

$$\lambda = \frac{A + C \pm \sqrt{(A+C)^2 - 4AC + B^2}}{2} = \frac{A + C \pm \sqrt{(A-C)^2 + B^2}}{2}.$$

Porque $(A-C)^2 + B^2 \geq 0$, quaisquer que sejam A, B e C reais, então os valores próprios de M são números reais.

1ºCaso Se $(A-C)^2 + B^2 = 0$, então $A = C$, $B = 0$. Pelo que M é diagonal.

2ºCaso Se $(A-C)^2 + B^2 > 0$, então M tem 2 valores próprios λ_1 e λ_2, distintos. Porque a dimensão de qualquer subespaço próprio de M é maior ou igual a 1, pelo Teorema 7.11, M é diagonalizável.

-Agora determinemos uma matriz diagonalizante de M, no 2º caso:

Sendo $\lambda_1 \neq \lambda_2$, os dois valores próprios de M e se

$$M_{\lambda_1} = \langle (x_1, y_1) \rangle, \quad M_{\lambda_2} = \langle (x_2, y_2) \rangle$$

consideremos os vectores

$$(x_1', y_1') = \left(\frac{x_1}{\sqrt{x_1^2 + y_1^2}}, \frac{y_1}{\sqrt{x_1^2 + y_1^2}} \right) \quad \text{de } M_{\lambda_1}, \text{ com norma } 1$$

$$(x_2', y_2') = \left(\frac{x_2}{\sqrt{x_2^2 + y_2^2}}, \frac{y_2}{\sqrt{x_2^2 + y_2^2}} \right) \quad \text{de } M_{\lambda_2}, \text{ com norma } 1.$$

Usando a Proposição 7.10, $\mathcal{B} = \{(x_1', y_1'), (x_2', y_2')\}$ é uma base de $\mathbb{R}^2$, formada por vectores próprios de M, com norma 1.

Sendo $S^{-1} = \begin{bmatrix} x_1' & x_2' \\ y_1' & y_2' \end{bmatrix}$, então S diagonaliza M.

Repare que $S^{-1} = M(id_{\mathbb{R}^2}; \mathcal{B}, b.c.)$, pelo que S^{-1} pode ser considerada a matriz mudança da base $\mathcal{B}$ para a base b.c. (transforma cada vector escrito na base $\mathcal{B}$, para a base canónica).

-A matriz S^{-1} anterior é tal que $(S^{-1})^T S^{-1} = I_2$.

$$(S^{-1})^T S^{-1} = \begin{bmatrix} x_1' & y_1' \\ x_2' & y_2' \end{bmatrix} \begin{bmatrix} x_1' & x_2' \\ y_1' & y_2' \end{bmatrix} = \begin{bmatrix} {x_1'}^2 + {y_1'}^2 & x_1'x_2' + y_1'y_2' \\ x_1'x_2' + y_1'y_2' & {x_2'}^2 + {y_2'}^2 \end{bmatrix}$$

e

$$\begin{aligned}(\lambda_1 \begin{bmatrix} x_1' & y_1' \end{bmatrix}) \begin{bmatrix} x_2' \\ y_2' \end{bmatrix} &= (\lambda_1 \begin{bmatrix} x_1' \\ y_1' \end{bmatrix})^T \begin{bmatrix} x_2' \\ y_2' \end{bmatrix} = (M \begin{bmatrix} x_1' \\ y_1' \end{bmatrix})^T \begin{bmatrix} x_2' \\ y_2' \end{bmatrix} \\ &= \begin{bmatrix} x_1' & y_1' \end{bmatrix} (M \begin{bmatrix} x_2' \\ y_2' \end{bmatrix}) = \lambda_2 \begin{bmatrix} x_1' & y_1' \end{bmatrix} \begin{bmatrix} x_2' \\ y_2' \end{bmatrix}.\end{aligned}$$

Temos que,

$$(\lambda_1 - \lambda_2) \begin{bmatrix} x_1' & y_1' \end{bmatrix} \begin{bmatrix} x_2' \\ y_2' \end{bmatrix} = 0.$$

Mas $\lambda_1 \neq \lambda_2$, pelo que, $x_1'x_2' + y_1'y_2' = \begin{bmatrix} x_1' & y_1' \end{bmatrix} \begin{bmatrix} x_2' \\ y_2' \end{bmatrix} = 0$. Por outro lado, os vectores de $\mathcal{B}$ têm norma 1, ou seja,

$$\begin{aligned} 1 &= ||(x_1', y_1')||^2 = {x_1'}^2 + {y_1'}^2 \\ 1 &= ||(x_2', y_2')||^2 = {x_2'}^2 + {y_2'}^2 \end{aligned}$$

Portanto, $(S^{-1})^T S^{-1} = I_2$, ou seja, $(S^{-1})^T = S$. Logo,

$$(S^{-1})^T M S^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Obtemos, assim, a forma quadrática diagonal de Q que é

$$Q(x', y') = \lambda_1 {x'}^2 + \lambda_2 {y'}^2$$

(forma quadrática na base $\mathcal{B}$).

Observação Repare que se (x', y') são as coordenadas na base $\mathcal{B}$ de um elemento de $\mathbb{R}^2$ que tem coordenadas, na base canónica de $\mathbb{R}^2$, (x, y), então

$$\begin{aligned} Q(x', y') &= \begin{bmatrix} x' & y' \end{bmatrix} (S^{-1})^T M S^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} = (S^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix})^T M S^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} \\ &= \begin{bmatrix} x \\ y \end{bmatrix}^T M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} M \begin{bmatrix} x \\ y \end{bmatrix} = Q(x, y). \end{aligned}$$

Portanto, o que fazemos é uma rotação das rectas de $\mathbb{R}^2$

$$y = 0 \quad ((x, y) = \lambda(1, 0),\ \lambda \in \mathbb{R}), \quad x = 0 \quad ((x, y) = \lambda(0, 1),\ \lambda \in \mathbb{R})$$

para as rectas

$$(x, y) = \lambda(x_1', y_1'),\ \lambda \in \mathbb{R}, \quad (x, y) = \lambda(x_2', y_2'),\ \lambda \in \mathbb{R}.$$

Exemplo 7.15 1. *Diagonalizemos a forma quadrática*

$$Q(x, y) = 2xy.$$

A matriz associada a Q é $M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ que tem os valores próprios distintos $\lambda_1 = 1$, $\lambda_2 = -1$. Então

$$M' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

é uma matriz diagonal semelhante a M (Outra matriz diagonal semelhante a M é a matriz $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$). Mas M' é a matriz associada à forma quadrática

$$Q(x', y') = (x')^2 - (y')^2.$$

2. *Determinemos qual a base de $\mathbb{R}^2$ em que a forma quadrática anterior é a forma quadrática*

$$Q(x', y') = (x')^2 - (y')^2.$$

Para isso necessitamos de determinar os seus subespaços próprios que são

$$M_1 = \langle (1,1) \rangle, \quad M_{-1} = \langle (-1,1) \rangle.$$

Consideremos

$$(x_1', y_1') = \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}} \right) \qquad \text{vector de } M_1 \text{ de norma 1}$$

$$(x_2', y_2') = \left(-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}} \right) \qquad \text{vector de } M_{-1} \text{ de norma 1.}$$

Sendo

$$S^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \qquad (\text{rotação de } \theta = \tfrac{\pi}{4})$$

então,

$$(S^{-1})^T M S^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

e uma forma quadrática diagonal de Q é

$$Q(x', y') = x'^2 - y'^2$$

na base $\mathcal{B} = \left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \right\}$).

7.2.2 Método para Classificar uma Cónica de $\mathbb{R}^2$

Depois de definir cónica de $\mathbb{R}^2$, podemos esquematizar o processo que nos permite classificar as cónicas.

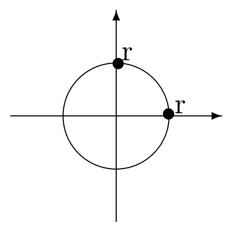
Definição 7.16 *Chama-se* **cónica** *de* $\mathbb{R}^2$ *ao conjunto de elementos de* $\mathbb{R}^2$ *que satisfazem uma equação cartesiana da forma*

$$Ax^2 + Bxy + Cy^2 + Dx + Ey = F$$

com A, B *ou* C *não nulo.*

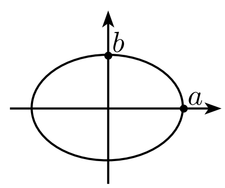
Uma cónica define, sempre, uma das seguintes figuras geométricas :

1. **Circunferência**



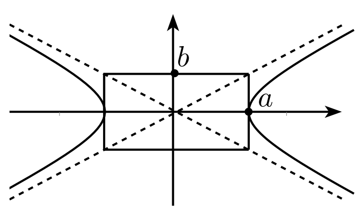
$$x^2 + y^2 = r^2 \quad \text{ou} \quad \frac{x^2}{r^2} + \frac{y^2}{r^2} = 1$$

2. **Elipse**

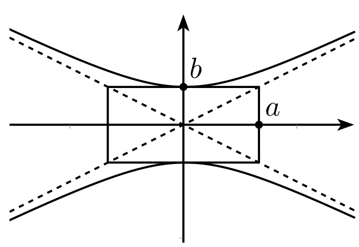


$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

3. **Hipérbole**

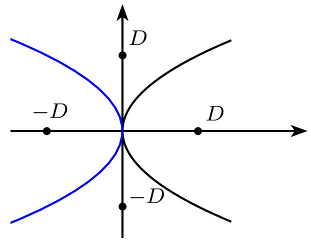


$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

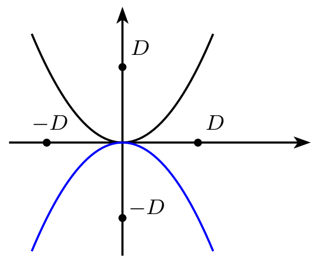


$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$$

4. **Parábola** passando pela origem $\qquad D > 0$



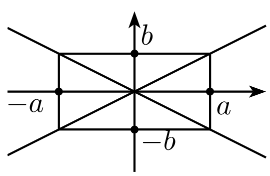
$$y^2 - Dx = 0 \qquad (\ y^2 + Dx = 0 \ )$$



$$x^2 - Dy = 0 \qquad (\ x^2 + Dy = 0 \ )$$

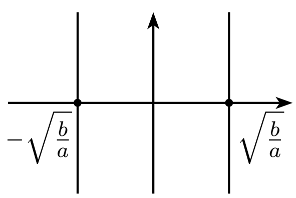
5. **Cónicas degeneradas**

a) **Duas rectas concorrentes** $\qquad$ (hipérbole degenerada)



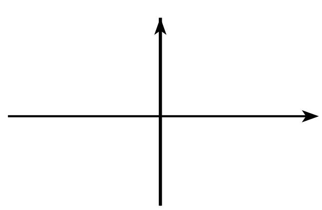
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$$

b) **Duas rectas paralelas** $\qquad$ (parábola degenerada)



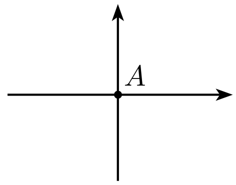
$$ax^2 - b = 0$$

c) **Uma recta** $\qquad$ (parábola degenerada)



$$x^2 = 0$$

d) **Um ponto** (elipse degenerada)



$$ax^2 + by^2 = 0, \quad a, b > 0$$

e) **Conjunto vazio** (elipse ou parábola degenerada)

$$ax^2 + by^2 + r^2 = 0, \quad a, b, r > 0$$

A classificação de uma cónica de $\mathbb{R}^2$, reduz-se a uma mudança de coordenadas de forma a que a equação assuma um dos tipos conhecidos. Isto é conseguido,

- por rotação dos eixos ordenados por forma a alinhá-los com os eixos da cónica correspondente (diagonalização da forma quadrática),
- seguida de uma translação destinada a colocar a origem no centro da cónica.

A rotação de coordenadas faz-se diagonalizando a forma quadrática $Ax^2 + Bxy + Cy^2$, ou seja, se λ_1 e λ_2 forem os valores próprios de $\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}$, obtemos uma equação cartesiana do tipo

$$\lambda_1(x')^2 + \lambda_2(y')^2 + D'x' + E'y' = F'$$

Se fizermos a translação de coordenadas

$$x'' = x' - x_0 \ , \qquad y'' = y' - y_0$$

onde (x_0, y_0) são as coordenadas do centro da cónica na base em que a forma quadrática é diagonal, obtemos uma equação cartesiana

$$\lambda_1(x'')^2 + \lambda_2(y'')^2 = F''.$$

Exemplo 7.17 *Considere a equação cartesiana*

$$2xy + x + y = 0.$$

No Exemplo 7.15, *diagonalizámos a forma quadrática*

$$2xy$$

e obtivemos a matriz

$$S^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \qquad \left(\theta = \frac{\pi}{4}\right)$$

e a forma quadrática diagonal

$$x'^2 - y'^2.$$

De $S^{-1}\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$ *temos que*

$$\begin{cases} x = \frac{1}{\sqrt{2}}(x' - y') \\ y = \frac{1}{\sqrt{2}}(x' + y'). \end{cases}$$

A equação cartesiana fica

$$(x')^2 - (y')^2 + \frac{2}{\sqrt{2}}x' = 0$$

ou seja,

$$(x')^2 - (y')^2 + \frac{2}{\sqrt{2}}x' + \frac{1}{2} - \frac{1}{2} = 0.$$

Porque

$$(x')^2 + \frac{2}{\sqrt{2}}x' + \frac{1}{2} = (x' + \frac{1}{\sqrt{2}})^2$$

vem que

$$\left(x' + \frac{1}{\sqrt{2}}\right)^2 - (y')^2 = \frac{1}{2}.$$

Se fizermos a translação de coordenadas

$$x'' = x' + \frac{1}{\sqrt{2}}, \qquad\qquad y'' = y'$$

obtemos a equação cartesiana

$$(x'')^2 - (y'')^2 = \frac{1}{2}$$

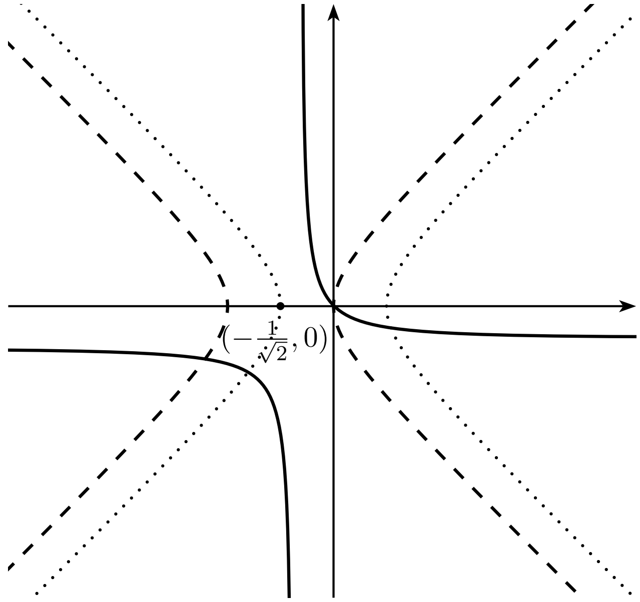
ou seja,

$$\frac{(x'')^2}{\frac{1}{2}} - \frac{(y'')^2}{\frac{1}{2}} = 1 \qquad \textit{(equação da hipérbole).}$$

Portanto, $\left(-\frac{1}{\sqrt{2}}, 0\right)$ *são as coordenadas do centro da hipérbole na base*

$$\mathcal{B} = \{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\}.$$

Esquematicamente: a pontilhado [.....], preto, temos representada a hipérbole nas variáveis (x'', y'')*, a tracejado [- - - -], verde, temos a hipérbole de centro* $\left(-\frac{1}{\sqrt{2}}, 0\right)$ *resultado de termos efectuado uma translação de coordenadas e finalmente, com traço continuo, azul, depois de uma rotação de amplitude* $\frac{\pi}{4}$*, obtemos a hipérbole inicial.*


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

Capítulo 8

ESPAÇOS VECTORIAIS

Como último capítulo deste manual e a título de resumo, iremos generalizar os conceitos e resultados que introduzimos para $\mathbb{R}^n$, a outros conjuntos.

8.1 Conceitos principais

Comecemos pela definição que vai uniformizar os conjuntos que podem generalizar $\mathbb{R}^n$.

Definição 8.1 *Seja E um conjunto não vazio e $\mathbb{K}$ o conjunto dos números reais, $\mathbb{R}$, ou o conjunto dos números complexos, $\mathbb{C}$. Suponhamos definidas duas operações:*

- *uma, designada por* **adição** *em E, que associa a cada par (u, v) de elementos de E um, e um só, elemento de E que é representado por $u + v$;*
- *outra, designada por* **multiplicação externa**, *que a cada $\alpha \in \mathbb{K}$ e a cada $u \in E$ associa um, e um só, elemento de E que é representado por $\alpha.u$ ou αu.*

Dizemos que E, com estas duas operações, é um **espaço vectorial sobre** $\mathbb{K}$, *ou que, $(E, +, .)$ é um espaço vectorial sobre $\mathbb{K}$, se*

1. *A* **adição** *em E verifica*

   A1) *a operação $+$ é associativa*

   A2) *a operação $+$ é comutativa*

   A3) *existe elemento neutro para a operação $+$*

   A4) *todo o elemento de E tem oposto para a operação $+$*

2. *A* **multiplicação externa** *verifica*

   M1) $\forall \alpha \in \mathbb{K}$, $\forall u, v \in E$, $\qquad \alpha(u + v) = \alpha u + \alpha v$

M2) $\forall \alpha, \beta \in \mathbb{K}, \forall u \in E, \qquad (\alpha + \beta)u = \alpha u + \beta u$

M3) $\forall \alpha, \beta \in \mathbb{K}, \forall u \in E, \qquad (\alpha\beta)u = \alpha(\beta u)$

M4) $\forall u \in E, \qquad 1.u = u$ \qquad (*sendo 1 o número real*)

Definição 8.2 *Seja $(E, +, .)$ um espaço vectorial sobre $\mathbb{K}$.*
Aos elementos de E chamamos **vectores**.
Aos elementos de $\mathbb{K}$ chamamos **escalares**.
Se $\mathbb{K} = \mathbb{R}$, dizemos que E é um **espaço vectorial real**.
Se $\mathbb{K} = \mathbb{C}$, dizemos que E é um **espaço vectorial complexo**.

Exemplo 8.3

1. *Como podemos observar, a definição de espaço vectorial é baseada nas propriedades das operações em $\mathbb{R}^n$, então $\mathbb{R}^n$ é um espaço vectorial real.*

2. *Uma forma de generalizar $\mathbb{R}^n$ é pensarmos no conjunto de todas as sequências infinitas de números reais, isto é, elementos da forma*

$$(u_1, u_2, \ldots, u_n, \ldots)$$

em que $u_1, u_2, \ldots, u_n, \ldots$ são números reais. Este conjunto é designado por $\mathbb{R}^\infty$.

Assim como em $\mathbb{R}^n$, somamos dois elementos desta forma ou multiplicamos um real por um elemento destes, pelo que podemos afirmar que $(\mathbb{R}^\infty, +, .)$ é um espaço vectorial real.

3. *Designando por $M_{n\times m}(\mathbb{R})$ o conjunto das matrizes de tipo $n \times m$ com entradas reais e considerando a adição de matrizes e o produto de um número real por uma matriz, podemos dizer que $(M_{n\times m}(\mathbb{R}), +, .)$ é um espaço vectorial real.*

4. *Sendo $\mathbb{R}_n[x]$ o conjunto de todos os polinómios na variável x, com coeficientes reais, de grau menor ou igual a n, com $n \in \mathbb{N}_0$, isto é,*

$$\mathbb{R}_n[x] = \{a_n x^n + \ldots + a_1 x + a_0 : a_n, \ldots, a_1, a_0 \in \mathbb{R}\},$$

e definindo as operações

$$\forall (a_n x^n + \ldots + a_1 x + a_0),\ (b_n x^n + \ldots + b_1 x + b_0) \in \mathbb{R}_n[x], \qquad \forall \alpha \in \mathbb{R}$$

$$(a_n x^n + \ldots + a_1 x + a_0) + (b_n x^n + \ldots + b_1 x + b_0) = (a_n + b_n)x^n + \ldots + (a_1 + b_1)x + (a_0 + b_0)$$

$$\alpha(a_n x^n + \ldots + a_1 x + a_0) = (\alpha a_n)x^n + \ldots + (\alpha a_1)x + (\alpha a_0)$$

temos que $\mathbb{R}_n[x]$ é um espaço vectorial real.

5. *Sendo* $\mathbb{R}[x]$ *o conjunto de todos os polinómios na variável* x*, com coeficientes reais (sem restrição de grau) e com as operações generalizadas de* $\mathbb{R}_n[x]$*, então,* $(\mathbb{R}[x], +, .)$ *é um espaço vectorial real.*

Todos os exemplos anteriores, alterando em cada caso o conjunto E *e/ou o conjunto* $\mathbb{K}$*, dão origem a outros espaços vectoriais:*

6. $(\mathbb{C}^n, +, .)$ *é um espaço vectorial real.*
7. $(\mathbb{C}^\infty, +, .)$ *é um espaço vectorial real.*
8. $(M_{n\times m}(\mathbb{C}), +, .)$ *é um espaço vectorial real.*
9. $(\mathbb{C}_n[x], +, .)$ *é um espaço vectorial real.*
10. $(\mathbb{C}[x], +, .)$ *é um espaço vectorial real.*

Mas também,

11. $(\mathbb{C}^n, +, .)$ *é um espaço vectorial complexo.*
12. $(\mathbb{C}^\infty, +, .)$ *é um espaço vectorial complexo.*
13. $(M_{n\times m}(\mathbb{C}), +, .)$ *é um espaço vectorial complexo.*
14. $(\mathbb{C}_n[x], +, .)$ *é um espaço vectorial complexo.*
15. $(\mathbb{C}[x], +, .)$ *é um espaço vectorial complexo.*

O que não é verdade é que algum dos espaços vectoriais reais 1. a 5., seja um espaço vectorial complexo.

Por exemplo, em 1., com $n = 3$, temos $(1, 2, -1) \in \mathbb{R}^3$ e $i \in \mathbb{C}$ mas

$$i(1, 2, -1) = (i, 2i, -i) \notin \mathbb{R}^3,$$

isto é, $\mathbb{R}^3$ não é "fechado"para a multiplicação de um número complexo (que não é real) por um vector de $\mathbb{R}^3$.

Teorema 8.4 *Sejam* $(E, +, .)$ *um espaço vectorial sobre* $\mathbb{K}$*,* $u, v \in E$ *e* $\alpha \in \mathbb{K}$*. Então:*

1. $\alpha 0_E = 0_E$*;*
2. $0_\mathbb{K} u = 0_E$*;*
3. $(-\alpha)u = \alpha(-u) = -(\alpha u)$*;*
4. $\alpha u = 0_E \Rightarrow \alpha = 0_\mathbb{K}$ *ou* $u = 0_E$*.*

Demonstração

1. Temos, por M1), $\alpha 0_E = \alpha(0_E + 0_E) = \alpha 0_E + \alpha 0_E$. Então, $0_E = \alpha 0_E - \alpha 0_E = \alpha 0_E + \alpha 0_E - \alpha 0_E = \alpha 0_E$.

3. Vejamos que $(-\alpha)u = -(\alpha u)$.

   Porque, por M2), $(-\alpha)u + \alpha u = (-\alpha + \alpha)u = 0_{\mathbb{K}} u = 0_E$, então temos o resultado.

4. Vejamos que se $\alpha u = 0_E$, então $\alpha = 0_{\mathbb{K}}$ ou $\alpha \neq 0_{\mathbb{K}}$.

   Se $\alpha = 0_{\mathbb{K}}$, temos o resultado.

   Se $\alpha \neq 0_{\mathbb{K}}$, então existe α^{-1} e

$$\alpha^{-1}(\alpha u) = \alpha^{-1} 0_E = 0_E.$$

   Por M3), $\alpha^{-1}(\alpha u) = (\alpha^{-1}\alpha)u = 1u = u$, por M4).

   Donde, $u = 0_E$. □

Definição 8.5 *Se W é um subconjunto não vazio de um espaço vectorial E sobre $\mathbb{K}$ e se W, com as duas operações definidas em E, é um espaço vectorial sobre $\mathbb{K}$, dizemos que W é um* **subespaço vectorial** *de E, ou simplesmente, um subespaço de E.*

Observação Repare que as propriedades A1) a A4) e M1) a M4) não dependem de W, então podemos estabelecer o próximo resultado.

Teorema 8.6 *Sejam $(E, +, .)$ um espaço vectorial sobre $\mathbb{K}$ e W um subconjunto não vazio de E. Então, W é subespaço de E se, e só se,*

1. $\forall u, v \in W, \qquad u + v \in W$

2. $\forall \alpha \in \mathbb{K}, \forall u \in W, \qquad \alpha u \in W.$

Exemplo 8.7 1. $(\mathbb{R}^n, +, .)$ *é um subespaço do espaço vectorial de* **real** $(\mathbb{C}^n, +, .)$.

2. $(\mathbb{R}^n, +, .)$ *não é um subespaço do espaço vectorial de* **complexo** $(\mathbb{C}^n, +, .)$ (*a condição* 2. *do* Teorema 8.6 *não se verifica*).

Observação As definições de combinação linear, independência linear, conjunto gerador e base são análogas às dadas em $\mathbb{R}^n$.

Exemplo 8.8 1. *As matrizes*

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

formam uma base do espaço vectorial real $M_{2\times 2}(\mathbb{R})$ (*chamada base canónica de* $M_{2\times 2}(\mathbb{R})$), *pois qualquer matriz de* $M_{2\times 2}(\mathbb{R})$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = aE_1 + bE_2 + cE_3 + dE_4$$

e se

$$\alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3 + \alpha_4 E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

usando a igualdade de matrizes, temos que, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$.

Então, $\dim M_{2\times 2}(\mathbb{R}) = 4$ *e* $M_{2\times 2}(\mathbb{R}) = \langle E_1, E_2, E_3, E_4\rangle$.

2. *O espaço vectorial real* $\mathbb{R}_2[x]$ *é gerado pelos polinómios*

$$1, x, x^2$$

(*estes três polinómios formam a base canónica de* $\mathbb{R}_2[x]$), *pois qualquer polinómio de* $\mathbb{R}_2[x]$ *é da forma*

$$a_2x^2 + a_1x + a_0.1$$

(*escreve-se como combinação linear de* $1, x, x^2$).

Se $\alpha_2x^2 + \alpha_1x + \alpha_0.1 = 0 = 0x^2 + 0x + 0.1$, *então*

$$\alpha_2 = \alpha_1 = \alpha_0 = 0.$$

Assim, $\mathbb{R}_2[x] = \langle x^2, x, 1\rangle$ *e* $\dim \mathbb{R}_2[x] = 3$.

3. *O espaço vectorial* **complexo** $\mathbb{C}^2$ *é gerado pelos vectores*

$$e_1 = (1, 0),\ \ e_2 = (0, 1)$$

pois qualquer vector de $\mathbb{C}^2$

$$(a_1 + b_1i, a_2 + b_2i) = (a_1 + b_1i)(1, 0) + (a_2 + b_2i)(0, 1)$$

(*é combinação linear de* e_1 *e* e_2).

Se $\alpha_1(1, 0) + \alpha_2(0, 1) = (0, 0)$ *então* $(\alpha_1, \alpha_2) = (0, 0)$ *e* $\alpha_1 = \alpha_2 = 0$.

Neste caso, $\mathbb{C}^2 = \langle (1, 0), (0, 1)\rangle$ *e* $\dim \mathbb{C}^2 = 2$.

4. *Vejamos agora* $\mathbb{C}^2$ *como espaço vectorial* **real**.

Neste caso, $\mathbb{C}^2$ *é gerado pelos vectores*

$$v_1 = (1, 0), v_2 = (i, 0), v_3 = (0, 1), v_4 = (0, i)$$

pois

$$(a_1 + b_1 i, a_2 + b_2 i) = a_1 v_1 + b_1 v_2 + a_2 v_3 + b_2 v_4.$$

Se $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 = (0,0)$, *então*

$$(\alpha_1 + \alpha_2 i, \alpha_3 + \alpha_4 i) = (0,0).$$

Porque $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ *são reais, então* $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$.

Donde,

$\mathbb{C}^2 = \langle (1,0), (i,0), (0,1), (0,i) \rangle$ *e* $\dim \mathbb{C}^2 = 4$.

Observação A respeito da dimensão de um espaço vectorial, ter sempre presente que

Definição 8.9 *Um espaço vectorial E é de* **dimensão finita** *se tem uma base com um número finito de vectores e é de* **dimensão infinita** *caso contrário.*

Exemplo 8.10 *O espaço vectorial* $\mathbb{R}[x]$ *é de* **dimensão infinita**. *Vejamos como se demonstra esta afirmação.*

Se $\mathbb{R}[x]$ fosse de dimensão finita, existia uma base $\mathcal{B}$ de $\mathbb{R}[x]$ com um número finito de polinómios. Seja n o maior grau dos polinómios de $\mathcal{B}$. Então, $x^{n+1} \notin \mathcal{B}$ e não se escreve como combinação linear dos elementos de $\mathcal{B}$. No entanto, $x^{n+1} \in \mathbb{R}[x]$. □

Os resultados de $\mathbb{R}^n$ que envolvem a sua dimensão, são válidos para espaços vectoriais de dimensão finita. Por exemplo,

Teorema 8.11 *Todas as bases de um espaço vectorial de dimensão finita têm o mesmo número de vectores.*

8.2 Aplicações Lineares

As definições e os resultados apresentados no capítulo 5 (Aplicações Lineares) podem ser aplicados a um espaço vectorial arbitrário. No entanto, temos que ter em atenção que sempre que estejam presentes dois espaços vectoriais, eles têm que ser espaços vectoriais sobre o mesmo conjunto $\mathbb{K}$. Por exemplo,

Definição 8.12 *Sejam E e E' dois espaços vectoriais sobre o mesmo conjunto $\mathbb{K}$ e $f : E \longrightarrow E'$ uma aplicação. Então, f diz-se* **aplicação linear** *se*

1. $\forall \alpha \in \mathbb{K}, \forall u \in E, \quad f(\alpha u) = \alpha f(u)$;

2. $\forall u, u' \in E, \quad f(u + u') = f(u) + f(u')$.

Definição 8.13 *Sejam E e E' dois espaços vectoriais sobre o mesmo conjunto $\mathbb{K}$ e $f : E \longrightarrow E'$ uma aplicação linear. Dizemos que E é isomorfo a E' se f é bijectiva.*

Observação Se E é isomorfo a E', então existe uma aplicação $f : E \longrightarrow E'$, linear e bijectiva. Mas atendendo à Proposição 5.29,

$$f^{-1} : E' \longrightarrow E$$

é aplicação linear e bijectiva. Então, E' é **isomorfo** a E. Portanto, podemos dizer simplesmente que E e E' são isomorfos e denotamos este facto por $E \cong E'$.

Exemplo 8.14 *Seja*

$$\begin{array}{rccc} f: & \mathbb{R}_2[x] & \longrightarrow & \mathbb{R}^3 \\ & a_2x^2 + a_1x + a_0 & \longmapsto & (a_2, a_1, a_0). \end{array}$$

Facilmente se prova que f é aplicação linear bijectiva. Então $\mathbb{R}_2[x] \cong \mathbb{R}^3$.

Teorema 8.15 *Qualquer espaço vectorial real de dimensão n é isomorfo a $\mathbb{R}^n$.*

Demonstração Seja E um espaço vectorial real de dimensão n e seja $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma sua base.

Porque cada vector u de E se escreve, de uma única forma, como combinação linear dos vectores de $\mathcal{B}$, então existem reais $\alpha_1, \alpha_2, \ldots, \alpha_n$ tais que

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n.$$

Portanto, podemos construir a aplicação

$$\begin{array}{rccc} f: & E & \longrightarrow & \mathbb{R}^n \\ & u = \alpha_1 v_1 + \ldots + \alpha_n v_n & \longmapsto & (\alpha_1, \ldots, \alpha_n). \end{array}$$

Para mostrarmos que $E \cong \mathbb{R}^n$, teremos de mostrar que

1. f é linear (exercício)

2. f é injectiva

3. f é sobrejectiva.

Vejamos então:

2. Sejam $u = \alpha_1 v_1 + \ldots + \alpha_n v_n$ e $u' = \beta_1 v_1 + \ldots + \beta_n v_n$ dois vectores de E tais que $f(u) = f(u')$.

Então, $(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n)$, ou seja,

$$\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$$

e $u = u'$. Portanto, f é injectiva.

3. Seja $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$. Porque E é um espaço vectorial real e $\mathcal{B}$ é uma base de E, o vector

$$u = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

pertence a E. Por definição, $f(u) = (\alpha_1, \ldots, \alpha_n)$. Logo, f é sobrejectiva. □

Este Teorema afirma que dado um espaço vectorial real de dimensão n, podemos pensar nele como se fosse o espaço vectorial $\mathbb{R}^n$ e tirar todas as conclusões.

Exemplo 8.16 *Seja*

$$\begin{aligned} f : M_{2\times 2}(\mathbb{R}) &\longrightarrow \mathbb{R}_4[x] \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} &\longmapsto (a+d)x^4 + (c-a)x^3 + bx. \end{aligned}$$

Porque

$$\begin{aligned} M_{2\times 2}(\mathbb{R}) &\cong \mathbb{R}^4 \\ E_i &\longmapsto e_i \end{aligned}$$

e

$$\begin{aligned} \mathbb{R}_4[x] &\cong \mathbb{R}^5 \\ x^i &\longmapsto e_{i+1} \end{aligned}$$

(Exemplo 8.8.1) *podemos pensar na aplicação*

$$\begin{aligned} g : \quad \mathbb{R}^4 &\longrightarrow \mathbb{R}^5 \\ (a,b,c,d) &\longmapsto (a+d, c-a, 0, b, 0). \end{aligned}$$

Porque g é linear, então f é linear.

$$\begin{aligned} Nuc\ g &= \{(a,b,c,d) \in \mathbb{R}^4 : g(a,b,c,d) = (0,0,0,0,0)\} \\ &= \{(a,b,c,d) \in \mathbb{R}^4 : (a+d, c-a, 0, b, 0) = (0,0,0,0,0)\} \\ &= \{(a,b,c,d) \in \mathbb{R}^4 : a = -d,\ c = a,\ b = 0\} \\ &= \{(a,0,a,-a) : a \in \mathbb{R}\} = \langle (1,0,1,-1) \rangle \end{aligned}$$

então,

$$Nuc\ f = \left\langle \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \right\rangle.$$

Temos também, pela Proposição 5.22*,*

$$\begin{aligned} Im\ g &= \langle g(1,0,0,0), g(0,1,0,0), g(0,0,1,0), g(0,0,0,1) \rangle = \\ &= \langle (1,-1,0,0,0), (0,0,0,1,0), (0,1,0,0,0), (1,0,0,0,0) \rangle \\ &= \langle (0,0,0,1,0), (0,1,0,0,0), (1,0,0,0,0) \rangle \end{aligned}$$

pois $(1,-1,0,0,0) = 1.(1,0,0,0,0) - 1.(0,1,0,0,0)$, *então*

$$Im\ f = \langle x, x^3, x^4 \rangle.$$

Observação Como já vimos, o espaço vectorial **real** $(\mathbb{C}^2, +, .)$ tem dimensão 4. Então pelo Teorema anterior,

$$\mathbb{C}^2 \cong \mathbb{R}^4.$$

Bibliografia

[1] HOWARD ANTON e ROBERT C.BUSBY, *Álgebra Linear Contemporânea*, Bookman, 2006.

[2] HOWARD ANTON e CHRIS RORRES, *Elementary Linear Algebra-Applications version*, 8th Edition, John Wiley and Sons,Inc., 2000.

[3] HOWARD ANTON e CHRIS RORRES, *Álgebra Linear com Aplicações*, 8ª Edição, Bookman, 2001.

[4] T.S.BLYTH e E.F.ROBERTSON, *Basic Linear Algebra*, Springer-Verlag, 1998.

[5] EMÍLIA GIRALDES, VÍTOR HUGO FERNANDES e M. PAULA MARQUES SMITH, *Álgebra Linear e Geometria Analítica* , McGraw-Hill de Portugal, 1995.

[6] LUÍS T. MAGALHÃES, *Álgebra Linear como Introdução à Matemática Aplicada*, 9ª Edição, Texto Editora, 2004.

[7] A. MONTEIRO, *Álgebra Linear e Geometria Analítica* , McGraw-Hill de Portugal, 2001.