Departamento de Matemática
Escola Superior de Tecnologia
Instituto Politécnico de Viseu

Notas de

ÁLGEBRA LINEAR E GEOMETRIA ANALÍTICA

Lurdes Sousa

Índice

Prefácio

Estas notas incluem uma síntese das matérias que habitualmente fazem parte do programa da disciplina de Álgebra Linear dos 1^os anos dos cursos de engenharia. Pretende--se que elas constituam para o aluno mais um elemento facilitador do estudo e estruturação dos assuntos tratados. Esta edição acrescenta à de 2000 o primeiro capítulo e algumas pequenas modificações nos restantes. Ademais, foram introduzidas algumas demonstrações, susceptíveis de serem abordadas num curso desta natureza.

I

Matrizes e sistemas de equações lineares

1. Sistemas de equações lineares, eliminação de Gauss e matrizes

Recordemos que, dado um sistema de equações lineares, se chega a um sistema equivalente por meio de cada um dos seguintes passos:

1. substituição de uma das equações pela que se obtém multiplicando cada um dos membros por um mesmo número não nulo;
2. substituição de uma das equações pela que se obtém somando-a membro a membro com uma equação obtida de outra como no passo 1;
3. troca da ordem das equações.

Antes de uma descrição precisa do método de Eliminação de Gauss, apresentam-se alguns exemplos onde ele é utilizado, que ajudarão a entender esta técnica. Na eliminação de Gauss do sistema usam-se os três princípios de equivalência enumerados atrás.

Exemplo 1. Consideremos o sistema seguinte:

$$\left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ -4x & -6y & +2z & +2v & = & 0 \\ 6x & +y & +z & +5v & = & 4 \end{array}\right.$$

Usando o princípio de equivalência 2 três vezes seguidas, tem-se:

$$\left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ -4x & -6y & +2z & +2v & = & 0 \\ 6x & +y & +z & +5v & = & 4 \end{array}\right. \quad \begin{array}{c} \Longleftrightarrow \\ E_2+2E_1 \\ E_3-3E_1 \end{array} \quad \left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ & -8y & +2z & +4v & = & 2 \\ & 4y & +z & +2v & = & 1 \end{array}\right.$$

$$\begin{array}{c} \Longleftrightarrow \\ \\ E_3+\frac{1}{2}E_2 \end{array} \quad \left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ & -8y & +2z & +4v & = & 2 \\ & & 2z & +4v & = & 2 \end{array}\right.$$

O último sistema tem uma apresentação que permite chegar facilmente às soluções usando sucessivamente substituição de variáveis, de baixo para cima, como descrito a seguir:

$$\left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ & -8y & +2z & +4v & = & 2 \\ & & 2z & +4v & = & 2 \end{array}\right. \Longleftrightarrow \left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ & -8y & +2z & +4v & = & 2 \\ & & 2z & & = & 2-4v \end{array}\right.$$

$$\Longleftrightarrow \left\{\begin{array}{rrrrcl} 2x & -y & & +v & = & 1 \\ & y & & & = & \frac{2-2(1-2v)-4v}{-8} \\ & & z & & = & 1-2v \end{array}\right. \Longleftrightarrow \left\{\begin{array}{rrrcl} x & & & = & \frac{1}{2}-\frac{1}{2}v \\ & y & & = & 0 \\ & & z & = & 1-2v \end{array}\right.$$

Este sistema é **possível**, pois tem solução. E é **indeterminado**, visto que tem várias (uma infinidade de) soluções. Tem uma variável livre, v, e três básicas, x, y e z. Claro que, utilizando outro processo de resolução, poder-se-ia obter uma diferente variável livre. Mas o número de variáveis livres e o número de variáveis básicas é independente do método de resolução utilizado. Como o número de variáveis livres do sistema é 1, dizemos que o seu **grau de indeterminação** é 1, ou ainda, que o sistema é **simplesmente indeterminado**.

Matrizes. Matriz dos coeficientes e matriz ampliada de um sistema. Consideremos o sistema do Exemplo 1; escrevendo os coeficientes das variáveix x, y, z e v na $1^{\underline{a}}$, $2^{\underline{a}}$, $3^{\underline{a}}$ e $4^{\underline{a}}$ colunas, respectivamente, obtém-se a seguinte matriz:

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

Esta matriz diz-se a **matriz dos coeficientes do sistema**. Se a esta matriz juntarmos, à direita, a coluna dos termos independentes do sistema,

$$\left[\begin{array}{rrrrcr} 2 & -1 & 0 & 1 & | & 1 \\ -4 & -6 & 2 & 2 & | & 0 \\ 6 & 1 & 1 & 5 & | & 4 \end{array}\right]$$

temos a chamada **matriz ampliada do sistema**.

Os passos da eliminação de Gauss do sistema do Exemplo 1 podem ser esquematizados, partindo da matriz ampliada, do modo seguinte:

$$\left[\begin{array}{rrrrcr} 2 & -1 & 0 & 1 & | & 1 \\ -4 & -6 & 2 & 2 & | & 0 \\ 6 & 1 & 1 & 5 & | & 4 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrcr} 2 & -1 & 0 & 1 & | & 1 \\ 0 & -8 & 2 & 4 & | & 2 \\ 0 & 4 & 1 & 2 & | & 1 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrcr} 2 & -1 & 0 & 1 & | & 1 \\ 0 & -8 & 2 & 4 & | & 2 \\ 0 & 0 & 2 & 4 & | & 2 \end{array}\right].$$

Uma **matriz** do tipo $m \times n$ (lei-se "m por n") é uma expressão constituída por mn números a_{11}, a_{12}, $\cdots$, a_{mn}, dispostos em linhas e colunas da seguinte forma:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & a_{m-1,3} & \cdots & a_{m-1,n-1} & a_{m-1,n} \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{m,n-1} & a_{mn} \end{bmatrix}.$$

Os números $a_{11},\ a_{12},\ \cdots,\ a_{21},\ a_{22},\ \cdots,$ dizem-se **entradas**, **elementos** ou **componentes** da matriz.

Uma matriz da forma acima pode representar-se abreviadamente por $[a_{ij}]_{1\leq i\leq m}^{1\leq j\leq n}$ ou $[a_{ij}]$ ou (a_{ij}).

Uma matriz com o mesmo número de linhas e colunas, digamos n, diz-se **quadrada** e de **ordem** n. Por exemplo,

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

é uma matriz quadrada de ordem 3.

Uma **matriz em escada de linhas** é uma matriz tal que por baixo do primeiro elemento não nulo de cada linha da matriz, e por baixo dos elementos anteriores da mesma linha, todas as entradas são nulas.

Numa matriz em escada de linhas, chama-se **pivô** ao primeiro elemento não nulo de cada linha.

Na eliminação de Gauss de uma matriz transformamos a matriz dada numa matriz em escada de linhas por meio de operações do seguinte tipo:

1. substituição de uma linha pela sua soma com o produto de um número por outra linha;

2. troca de linhas.

Numa matriz em escada por linhas, chama-se **característica** da matriz ao número de pivôs, o mesmo é dizer, ao número de linhas não nulas da matriz.

A **característica** de uma matriz A, denotada $\mathrm{car}(A)$, é, por definição, a característica da matriz em escada de linhas da matriz que se obtém através da eliminação de Gauss de A.

Uma matriz quadrada A de ordem n diz-se:

- **singular**, se $\mathrm{car}(A) < n$,
- **não singular**, se $\mathrm{car}(A) = n$.

Exemplo 2. Pretende-se determinar a característica da matriz

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

Para isso, fazemos a eliminação de Gauss da matriz. Começamos por utilizar o primeiro elemento da primeira linha, 1 ($\neq 0$), como pivô:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 3 & 1 \\ -4 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{bmatrix} \begin{array}{c} \longrightarrow \\ \\ L_3 - 2L_1 \\ L_4 + 4L_1 \\ \\ \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{bmatrix}$$

Como a segunda linha não tem nenhum elemento diferente de zero, passamos à terceira linha. O primeiro elemento não nulo da terceira linha é 3. Mas, escolhendo 3 como pivô, não é possível utilizar a técnica utilizada anteriormente para anular o -1 da quarta linha. Temos então de fazer uma troca de linhas, como ilustrado a seguir:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{bmatrix} \begin{array}{c} \\ \longrightarrow \\ \\ L_3 \leftrightarrow L_4 \\ \\ \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 2 & 2 \end{bmatrix}$$

A matriz obtida ainda não é uma matriz em escada de linhas; temos, portanto, de continuar a eliminação:

$$\begin{array}{c} \longrightarrow \\ \\ \\ \\ L_5 - \frac{2}{3}L_4 \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & 1 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & \frac{4}{3} \end{bmatrix}$$

Chegámos agora a uma matriz em escada de linhas, pelo que está feita a eliminação de Gaus da matriz dada. Olhando para a matriz obtida, verificamos que tem 4 linhas não nulas pelo que a caracteística da matriz dada é 4, abreviadamente, car(A)=4. Como a matriz A tem 5 colunas, concluimos que é singular.

Observação. Ao longo da eliminação de Gauss encontramo-nos repetidamente na seguinte situação: Temos uma matriz que tem uma linha da forma

$$0 \quad 0 \quad \ldots \quad 0 \quad a_{kt} \quad a_{k,t+1} \quad \ldots \quad a_{kn}$$

com $a_{kt} \neq 0$, que é escolhido para pivô, e uma linha abaixo dessa, como ilustrado a seguir

$$\begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{kt} & a_{k,t+1} & \dots & a_{kn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{rt} & a_{r,t+1} & \dots & a_{rn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix};$$

pretendemos anular o elemento na posição rt. A operação usada para esse efeito é a substituição da r-ésima linha pela que se obtém pondo

$$L_r - \frac{a_{rt}}{a_{kt}} L_k$$

(onde L_k e L_r são as linhas representadas acima).

Exemplo 3. Dado o sistema

$$\begin{cases} 2x & -y & & +v & = & 1 \\ -4x & -6y & +2z & +2v & = & 0 \\ 6x & -7y & +z & +5v & = & 2 \end{cases}$$

a sua matriz ampliada é:

$$\begin{bmatrix} 2 & -1 & 0 & 1 & | & 1 \\ -4 & -6 & 2 & 2 & | & 0 \\ 6 & -7 & 1 & 5 & | & 2 \end{bmatrix}.$$

A seguir faz-se a eliminação de Gauss da matriz ampliada do sistema, indicando por baixo das setas as operações que foram feitas sobre as linhas para passar de uma matriz a outra:

$$\begin{bmatrix} 2 & -1 & 0 & 1 & | & 1 \\ -4 & -6 & 2 & 2 & | & 0 \\ 6 & -7 & 1 & 5 & | & 2 \end{bmatrix} \begin{matrix} \longrightarrow \\ L_2 + 2L_1 \\ L_3 - 3L_1 \end{matrix} \begin{bmatrix} 2 & -1 & 0 & 1 & | & 1 \\ 0 & -8 & 2 & 4 & | & 2 \\ 0 & -4 & 1 & 2 & | & -1 \end{bmatrix}$$

$$\begin{matrix} \longrightarrow \\ \\ L_3 - \frac{1}{2}L_2 \end{matrix} \begin{bmatrix} 2 & -1 & 0 & 1 & | & 1 \\ 0 & -8 & 2 & 4 & | & 2 \\ 0 & 0 & 0 & 0 & | & -2 \end{bmatrix}$$

Concluimos que o sistema dado é equivalente a:

$$\begin{cases} 2x & -y & & +v & = & 1 \\ & -8y & +2z & +4v & = & 2 \\ & & & 0v & = & -2 \end{cases}$$

Atendendo à última equação, o sistema é **impossível**.

Denotemos por $[A|b]$ a matriz ampliada do sistema, onde A é a matriz dos coeficientes e b é a matriz coluna dos termos independentes. Denotemos por $[U|c]$ a matriz resultante da eliminação de Gauss de $[A|b]$. Temos então car(A)=car(U)=2 e car$([A|b])$=car$([U|c])$=3. Note-se que a impossibilidade do sistema tem a ver com o facto de a última linha de U ser nula, mas não ser nula a última linha de $[U|c]$, ou seja, tem a ver com facto de se ter car$(A) <$ car$([A|b])$.

Exemplo 4. Consideremos o sistema de equações

$$\left\{\begin{array}{rrrrcl} & x_2 & +2x_3 & -x_4 & = & 2 \\ 3x_1 & & -x_3 & +x_4 & = & 1 \\ -3x_1 & +x_2 & +3x_3 & +x_4 & = & 1 \\ 2x_1 & -2x_2 & +\frac{1}{3}x_3 & +x_4 & = & 1 \end{array}\right.$$

Eliminação de Gauss da matriz ampliada do sistema:

$$\left[\begin{array}{rrrr|r} 0 & 1 & 2 & -1 & 2 \\ 3 & 0 & -1 & 1 & 1 \\ -3 & 1 & 3 & 1 & 1 \\ 2 & -2 & \frac{1}{3} & 1 & 1 \end{array}\right] \underset{L_1 \leftrightarrow L_2}{\longrightarrow} \left[\begin{array}{rrrr|r} 3 & 0 & -1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 2 \\ -3 & 1 & 3 & 1 & 1 \\ 2 & -2 & \frac{1}{3} & 1 & 1 \end{array}\right]$$

$$\begin{array}{c} \longrightarrow \\ \\ L_3 + L_1 \\ L_4 - \frac{2}{3}L_1 \end{array} \left[\begin{array}{rrrr|r} 3 & 0 & -1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 2 \\ 0 & 1 & 2 & 2 & 2 \\ 0 & -2 & 1 & \frac{1}{3} & \frac{1}{3} \end{array}\right] \begin{array}{c} \longrightarrow \\ \\ L_3 - L_2 \\ L_4 + 2L_2 \end{array} \left[\begin{array}{rrrr|r} 3 & 0 & -1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 2 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 5 & -\frac{5}{3} & \frac{13}{3} \end{array}\right]$$

$$\begin{array}{c} \longrightarrow \\ \\ \\ L_3 \leftrightarrow L_4 \end{array} \left[\begin{array}{rrrr|r} 3 & 0 & -1 & 1 & 1 \\ 0 & 1 & 2 & -1 & 2 \\ 0 & 0 & 5 & -\frac{5}{3} & \frac{13}{3} \\ 0 & 0 & 0 & 3 & 0 \end{array}\right]$$

O sistema dado é equivalente a

$$\left\{\begin{array}{rrrrcl} 3x_1 & & -x_3 & +x_4 & = & 1 \\ & x_2 & +2x_3 & -x_4 & = & 2 \\ & & 5x_3 & -\frac{5}{3}x_4 & = & \frac{13}{3} \\ & & & 3x_4 & = & 0 \end{array}\right.$$

Resolvendo por substituição da última para a primeira equação, obtemos:

$$\begin{array}{rcl} x_1 & = & \frac{28}{45} \\ x_2 & = & \frac{4}{15} \\ x_3 & = & \frac{13}{15} \\ x_4 & = & 0 \end{array}$$

As variáveis básicas são as correspondentes às colunas que têm pivôs na matriz resultante da eliminação de Gauss da matriz dos coeficientes; neste caso: x_1, x_2, x_3, x_4.

Sendo todas as variáveis básicas (ou seja, não havendo variáveis livres), é claro que o sistema não poderia ter mais do que uma solução.

Observando os exemplos já dados, facilmente se concluem as propriedades seguintes relativamente à classificação de um sistema.

Classificação de um sistema mediante as características da matriz dos coeficientes e da matriz ampliada. Dado um sistema de m equações e n incógnitas, suponhamos que A é a matriz dos coeficientes do sistema (A é portanto uma matriz do tipo $m \times n$) e b é a matriz dos termos independentes do sistema. Assim,

$$[A|\, b]$$

é a matriz ampliada do sistema. Nestas condições, o sistema é:

- possível se e só se car(A)=car($[A|b]$);
- sendo possível, é
  - determinado, se e só se car(A) $= n$ (tendo-se que, nesse caso, todas as variáveis são básicas);
  - indeterminado, com indeterminação de grau k (com $k \neq 0$), se e só se car$(A) = n - k$. As variáveis básicas podem ser escolhidas como as variáveis correspondentes aos pivôs da matriz resultante da eliminação de Gauss da matriz dos coeficientes. O sistema tem então k variáveis livres.

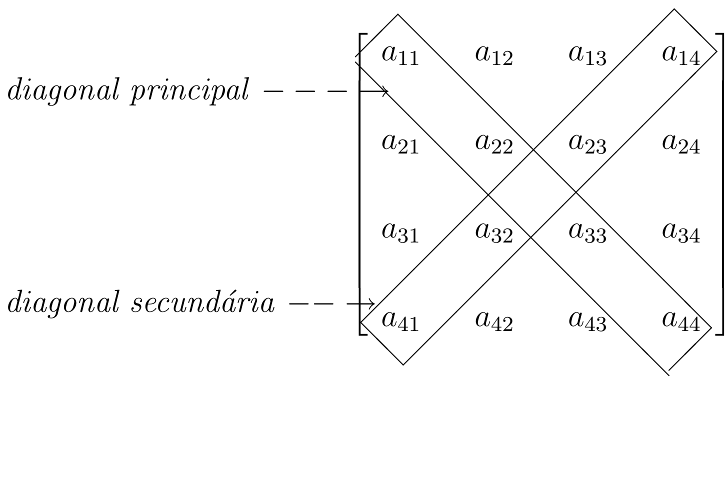
2. Matrizes especiais

As definições de matriz, elementos ou entradas de uma matriz, tipo de uma matriz, etc., foram já apresentadas na secção anterior. Viu-se também o que se entende por matriz em escada. A seguir dá-se conta de mais alguns tipos especiais de matrizes.

A **diagonal principal** (ou apenas **diagonal**) de uma matriz quadrada (a_{ij}) de ordem n é constituída pelos elementos

$$a_{11},\ a_{22},\ \ldots,\ a_{nn}\,.$$

Se pensarmos na representação de uma matriz quadrada como um quadrado, a diagonal principal corresponde efectivamente a uma das diagonais do quadrado. Os elementos que ficam situados na outra diagonal desse quadrado constituem a **diagonal secundária** da matriz. Na figura seguinte estão assinaladas as duas diagonais de uma matriz 4×4.



Uma **matriz nula** é uma matiz de entradas todas nulas. Por exemplo, a matriz nula 3×4 é

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

A matriz nula $m \times n$ designa-se usualmente por $O_{m \times n}$. Se $m = n$, pode escrever-se simplesmente O_n. Quando não há ambiguidade sobre o tipo da matriz, escreve-se simplesmente O.

A **matriz identidade** de ordem n é a matiz quadrada de ordem n cujas entradas na diagonal principal são todas 1 e fora da diagonal são nulas. Designa-se por I_n ou apenas I. Assim, a matriz identidade 4×4 é

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

Uma **matriz diagonal** é uma matriz quadrada cujas entradas fora da diagonal principal são todas nulas. É pois uma matriz da forma

$$\begin{bmatrix} \alpha_1 & 0 & 0 & ... & 0 & 0 \\ 0 & \alpha_2 & 0 & ... & 0 & 0 \\ 0 & 0 & \alpha_3 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \alpha_{n-1} & 0 \\ 0 & 0 & 0 & ... & 0 & \alpha_n \end{bmatrix}.$$

Uma matriz quadrada diz-se **triangular superior** se todas as suas entradas abaixo da diagonal principal forem nulas; assim:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & ... & 0 & a_{nn} \end{bmatrix}.$$

Analogamente uma matriz **triangular inferior** é uma matriz da forma

$$\begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 & 0 \\ a_{31} & a_{32} & a_{33} & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \ldots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{n,n-1} & a_{nn} \end{bmatrix}.$$

Uma **matriz linha** é uma matriz constituída por uma única linha. Por exemplo, $[1 \;\; 0 \;\; -1 \;\; 5]$. Uma **matriz coluna** é uma matriz constituída por uma única coluna. Por exemplo, $\begin{bmatrix} 1 \\ 0 \\ -1 \\ 5 \end{bmatrix}$.

3. Operações com matrizes

O conjunto de todas as matrizes $m \times n$ com entradas em $\mathbb{R}$ designa-se por $\mathcal{M}_{m\times n}(\mathbb{R})$. Analogamente, $\mathcal{M}_{m\times n}(\mathbb{C})$ designa o conjunto de todas as matrizes $m \times n$ com entradas no corpo complexo $\mathbb{C}$. Assim $\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 2 & 1 \end{bmatrix}$ é um elemento de $\mathcal{M}_{2\times 4}(\mathbb{R})$ (e também de $\mathcal{M}_{2\times 4}(\mathbb{C})$) e a matriz $\begin{bmatrix} 1 & 2-i & -1 \\ i & 2 & 1-i \\ i & -2i & \frac{1}{2}i \end{bmatrix}$ é um elemento de $\mathcal{M}_{3\times 3}(\mathbb{C})$. As matrizes com entradas em $\mathbb{R}$ dizem-se **reais** e aquelas cujas entradas são números complexos dizem-se **complexas**.

Adição. A adição de matrizes faz-se entre matrizes do mesmo tipo (reais ou complexas) e define-se do seguinte modo: Dadas matrizes $A = (a_{ij})$ e $B = (b_{ij})$ do tipo $m \times n$, a sua adição é a matriz soma $m \times n$ dada por

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

Por exemplo, $\begin{bmatrix} 1 & i & 1+2i \\ i & 1-i & -1 \\ 1 & 2-2i & 0 \end{bmatrix} + \begin{bmatrix} i & -i & -2i \\ i & 1+i & i \\ 0 & i & i \end{bmatrix} = \begin{bmatrix} 1+i & 0 & 1 \\ 2i & 2 & -1+i \\ 1 & 2-i & i \end{bmatrix}$.

É fácil verificar que a adição em $\mathcal{M}_{m\times n}(\mathbb{K})$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) goza das propriedades seguintes:

1. *Comutatividade:* $A + B = B + A$, para quaisquer $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$.
2. *Associatividade:* $(A+B)+C = A+(B+C)$, para quaisquer $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{K})$.

3. *Existência de elemento neutro:* $O_{m\times n}+A=A+O_{m\times n}$, para todo $A\in\mathcal{M}_{m\times n}(\mathbb{K})$.

4. *Existência de simétrico:* Para cada matriz $A=(a_{ij})$ a matriz $-A=(-a_{ij})$ satisfaz as igualdades $A+(-A)=-A+A=O$.

Multiplicação escalar. O produto de um escalar α (i.e., α é um número real ou complexo, consoante se esteja a considerar $\mathbb{K}=\mathbb{R}$ ou $\mathbb{K}=\mathbb{C}$) por uma matriz real $A=(a_{ij})$ do tipo $m\times n$ é a matriz $m\times n$

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

Por exemplo, $-\frac{i}{2}\begin{bmatrix} i & -2 & 4\\ 0 & 2-2i & 2\end{bmatrix}=\begin{bmatrix} \frac{1}{2} & i & -2i\\ 0 & -1-i & -i\end{bmatrix}$.

É imediato que, para quaisquer números α e β e quaisquer matrizes A e B do mesmo tipo, se verificam as seguintes igualdades:

1. $\alpha(\beta A)=(\alpha\beta)A$

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

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

Multiplicação de matrizes. Vamos ver agora como se define o produto entre matrizes.

O produto de uma matriz linha com n elementos por uma matriz coluna também com n elementos é uma matriz 1×1, ou seja, um número, que se obtém como indicado a seguir:

$$[a_1\ \ a_2\ \ ...\ \ a_n]\begin{bmatrix} b_1\\ b_1\\ \vdots\\ b_n\end{bmatrix}=a_1b_1+a_2b_2+...+a_nb_n\,.$$

O produto AB de duas matrizes só está definido se o número de colunas de A for igual ao número de linhas de B. Ou seja, se A é do tipo $m\times n$, B terá de ser do tipo $n\times p$ para algum p. Dadas A do tipo $m\times n$ e B do tipo $n\times p$, o **produto** AB, de A por B, é a matriz $m\times p$ tal que, para cada $i\in\{1,2,...,m\}$ e $j\in\{1,2,...,p\}$, o elemento na posição ij da matriz AB é igual ao produto da i-ésima linha de A pela j-ésima coluna de B. Por exemplo,

$$\begin{bmatrix} 2 & 1\\ -1 & 0\\ 3 & 4\end{bmatrix}\begin{bmatrix} 3 & -1\\ -2 & 5\end{bmatrix}=\begin{bmatrix} 2\times 3+1\times(-2) & 2\times(-1)+1\times 5\\ (-1)\times 3+0\times(-2) & (-1)\times(-1)+0\times 5\\ 3\times 3+4\times(-2) & 3\times(-1)+4\times 5\end{bmatrix}=\begin{bmatrix} 4 & 3\\ -3 & 1\\ 1 & 17\end{bmatrix}.$$

Por outras palavras, dadas $A=(a_{ij})$ do tipo $m\times n$ e $B=(b_{ij})$ do tipo $n\times p$, o **produto** AB, de A por B, é a matriz $m\times p$

$$AB = \begin{bmatrix} \sum_{k=1}^{n} a_{1k}b_{k1} & \sum_{k=1}^{n} a_{1k}b_{k2} & \dots & \sum_{k=1}^{n} a_{1k}b_{kp} \\ \sum_{k=1}^{n} a_{2k}b_{k1} & \sum_{k=1}^{n} a_{2k}b_{k2} & \dots & \sum_{k=1}^{n} a_{2k}b_{kp} \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^{n} a_{mk}b_{k1} & \sum_{k=1}^{n} a_{mk}b_{k2} & \dots & \sum_{k=1}^{n} a_{mk}b_{kp} \end{bmatrix}.$$

Usando uma notação abreviada,

$$AB = (\sum_{k=1}^{n} a_{ik}b_{kj}).$$

Repare-se que, utilizando esta definição, podemos representar um sistem da forma

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

pela equação matricial

$$Ax = b$$

onde A é a matriz dos coeficientes do sistema, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ e $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$.

Propriedades da multiplicação de matrizes.

1. *Distributividade (à esquerda e à direita da multiplicação relativamente à adição)*:

   (a) $A(B+C) = AB + AC$, para quaisquer matrizes A do tipo $m \times n$ e B e C do tipo $n \times p$;

   (b) $(A+B)C = AC + BC$, para quaisquer matrizes A e B do tipo $m \times n$ e C do tipo $n \times p$.

   Prova. (a) Sendo $A = (a_{ij})$, $B = (b_{jr})$ e $C = (c_{jr})$ dos tipos referidos acima, tem-se que:

$$\begin{aligned} A(B+C) &= (\sum_{k=1}^{n} a_{ik}(b_{kr} + c_{kr})) \\ &= (\sum_{k=1}^{n} (a_{ik}b_{kr} + a_{ik}c_{kr})) \\ &= (\sum_{k=1}^{n} a_{ik}b_{kr} + \sum_{k=1}^{n} a_{ik}c_{kr}) \\ &= AB + AC \end{aligned}$$

(b) prova-se de forma análoga.

A prova das duas igualdades seguintes, para quaisquer matrizes A e B tais que o produto AB está definido e quaisquer números α e β, é fácil e fica como exercício.

2. $(\alpha A)(\beta B) = (\alpha\beta)(AB)$

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

4. *Associatividade*: $A(BC) = (AB)C$, para quaisquer matrizes A $m \times n$, B $n \times p$ e C $p \times q$.

   Prova. Sendo $A = (a_{ij})$, $B = (b_{jr})$ e $C = (c_{rs})$ com $1 \leq i \leq m$, $1 \leq j \leq n$, $1 \leq r \leq p$ e $1 \leq s \leq q$, tem-se que, por definição:

$$AB = \left(\sum_{k=1}^{n} a_{ik}b_{kr}\right) \text{ e } AB \text{ é do tipo } m \times p\,;$$

$$BC = \left(\sum_{t=1}^{p} b_{jt}c_{ts}\right) \text{ e } BC \text{ é do tipo } p \times q\,.$$

   Vem, então:

$$\begin{aligned}
(AB)C &= \left(\sum_{t=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kt}\right)c_{ts}\right), && \text{por definição de produto de matrizes}\\
&= \left(\sum_{t=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kt}c_{ts}\right), && \text{usando a distributividade da multiplicação em relação à adição em } \mathbb{R} \text{ ou } \mathbb{C}\\
&= \left(\sum_{k=1}^{n}\sum_{t=1}^{p} a_{ik}b_{kt}c_{ts}\right), && \text{usando a comutatividade e associatividade da adição em } \mathbb{R} \text{ ou } \mathbb{C}\\
&= \left(\sum_{k=1}^{n}\left(a_{ik}\sum_{t=1}^{p} b_{kt}c_{ts}\right)\right), && \text{usando a distributividade da multiplicação em relação à adição em } \mathbb{R} \text{ ou } \mathbb{C}\\
&= A(BC), && \text{por definição do produto de matrizes}
\end{aligned}$$

Note-se que o produto de matrizes não é comutativo. Para o concluir, sugere-se o cálculo de, por exemplo, o produto de AB e BA para $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$.

Matrizes invertíveis. Uma matriz quadrada A de ordem n diz-se **invertível** se existir uma matriz quadrada B de ordem n tal que $AB = I$ e $BA = I$. Nesse caso, B diz-se **inversa** de A e designa-se por A^{-1}.

É de notar que a inversa de uma matriz, quando existe, é única. Com efeito, suponhamos que A, B_1 e B_2 são matrizes quadradas de ordem n tais que $AB_i = B_iA = I$, para $i = 1, 2$; então $B_1 = B_1I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2$.

Nem todas as matrizes têm inversa. Por exemplo, o produto da matriz $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ por qualquer outra dá uma matriz com a segunda linha com entradas todas nulas; logo é impossível encontrar uma inversa para esta matriz.

Fica como **exercício** chegar à seguinte propriedade:

Se A e B são matrizes quadradas da mesma ordem e invertíveis então AB também é invertível, tendo-se $(AB)^{-1} = B^{-1}A^{-1}$. Mais geralmente:

Propriedade. Se A_1, A_2, ..., A_n são matrizes quadradas da mesma ordem, todas invertíveis, então o produto $A_1A_2...A_n$ é invertível, tendo-se $(A_1A_2...A_n)^{-1} = A_n^{-1}...A_2^{-1}A_1^{-1}$.

Multiplicação por blocos. Consideremos a matriz

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

Podemos particioná-la em blocos do seguinte modo:

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

Se pusermos

$A_{11} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, $A_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ e $A_{22} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, podemos escrever

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Por vezes é vantajoso considerar as matrizes particionadas em blocos. Por exemplo, para matrizes de dimensões grandes, isso pode facilitar a multiplicação. É isso que vai ser descrito a seguir.

Sejam A e B matrizes do tipo $m \times n$ e $n \times p$, respectivamente, particionadas em blocos do modo seguinte

$$A = \begin{bmatrix} A_{11} & A_{12} & ... & A_{1s} \\ A_{21} & A_{22} & ... & A_{2s} \\ ... & .... & ... & ... \\ A_{r1} & A_{r2} & ... & A_{rs} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & B_{12} & ... & B_{1t} \\ B_{21} & B_{22} & ... & B_{2t} \\ ... & .... & ... & ... \\ B_{s1} & B_{s2} & ... & B_{st} \end{bmatrix}$$

e tais que, para $i = 1, ..., r$, $k = 1, ..., s$ e $j = 1, ..., t$, se verifica a igualdade

$$(\text{número de colunas de } A_{ik}) = (\text{número de linhas de } B_{kj}).$$

Esta igualdade assegura que o produto $A_{ik}B_{kj}$ está definido.

Note-se ainda que, olhadas por blocos, as matrizes A e B são do tipo $r \times s$ e $s \times t$, respectivamente, ou seja, o número de colunas de A é igual ao número de linhas de B.

Nas condições expostas, verifica-se então a igualdade

$$AB = \begin{bmatrix} \sum_{k=1}^{s} A_{1k}B_{k1} & \sum_{k=1}^{s} A_{1k}B_{k2} & \dots & \sum_{k=1}^{s} A_{1k}B_{kt} \\ \sum_{k=1}^{s} A_{2k}B_{k1} & \sum_{k=1}^{s} A_{2k}B_{k2} & \dots & \sum_{k=1}^{s} A_{2k}B_{kt} \\ \dots & \dots & \dots & \dots \\ \sum_{k=1}^{s} A_{rk}B_{k1} & \sum_{k=1}^{s} A_{rk}B_{k2} & \dots & \sum_{k=1}^{s} A_{rk}B_{kt} \end{bmatrix}$$

Note-se que, atendendo ao modo como as matrizes estão particionadas, fixados $i \in \{1, ..., r\}$ e $j \in \{1, ..., s\}$, as matrizes $A_{ik}B_{kj}$ são todas do mesmo tipo ($k = 1, ..., s$), logo a soma $\sum_{k=1}^{s} A_{ik}B_{kj}$ está definida.

Exemplo. Relativamente às matrizes

$$A = \left[\begin{array}{ccc|cc|cc} 1 & 0 & 0 & 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 3 & 4 & 0 & 0 \\ 0 & 0 & 1 & 5 & 6 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 2 & 2 & 0 & 0 \end{array}\right] \text{ e } B = \left[\begin{array}{cc|cccc|cc} 1 & 2 & 0 & 0 & 0 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & -1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 1 & -1 & 1 & -1 & 0 & 1 \\ \hline 6 & 0 & 6 & 6 & 6 & 6 & 1 & -2 \\ 0 & 6 & 1 & 1 & 1 & 1 & -1 & 2 \end{array}\right]$$

pode fazer-se o produto AB visto que o número de colunas de A é igual ao número de linhas de B (igual a 7). Por outro lado as duas matrizes estão particionadas de forma que é possível fazer o produto por blocos; para concluir isso, basta notar que as colunas de A estão agrupadas do mesmo modo que as linhas de B, especificamente: 3, 2 e 2. Repare-se que a partição das matrizes está feita de modo a fazer aparecer várias matrizes identidade e nulas para facilitar os cálculos. Temos assim que

$$AB = \begin{bmatrix} I & A_{12} & O \\ A_{21} & O & I \\ O & A_{32} & O \end{bmatrix} \begin{bmatrix} B_{11} & O & B_{13} \\ I & B_{22} & I \\ 6I & B_{32} & B_{33} \end{bmatrix} = \begin{bmatrix} B_{11} + A_{12} & A_{12}B_{22} & B_{13} + A_{12} \\ A_{21}B_{11} + 6I & B_{32} & A_{21}B_{13} + B_{33} \\ A_{31} & A_{32}B_{22} & A_{32} \end{bmatrix}.$$

A seguir apresenta-se parte do produto AB já feito. Fica como **exercício** completar a matriz com as quatro submatrizes que faltam.

$$AB = \begin{bmatrix} 2 & 0 & | & & & & & | & 2 & -1 \\ 5 & 5 & | & & & & & | & 4 & 5 \\ 8 & 6 & | & & & & & | & 6 & 7 \\ - & - & - & - & - & - & - & - & - & - \\ & & | & 6 & 6 & 6 & 6 & | & & \\ & & | & 1 & 1 & 1 & 1 & | & & \\ - & - & - & - & - & - & - & - & - & - \\ 2 & 2 & | & & & & & | & 2 & 2 \end{bmatrix}$$

Exemplos especiais de multiplicação por blocos. Sejam A e B matrizes $m \times n$ e $n \times p$, respectivamente. Designemos por

$$L_1,\ L_2,\ ...,\ L_m$$

as linhas de A, na ordem pela qual aparecem em A, e por

$$C_1,\ C_2,\ ...,\ C_p$$

as colunas de B, também na ordem pela qual estão em B.

Então:

$$AB = \begin{bmatrix} L_1 \\ L_2 \\ . \\ . \\ . \\ L_m \end{bmatrix} B = \begin{bmatrix} L_1B \\ L_2B \\ . \\ . \\ . \\ L_mB \end{bmatrix}.$$

e

$$AB = A\,[C_1 \ \ C_2 \ \ ... \ \ C_p] = [AC_1 \ \ AC_2 \ \ ... \ \ AC_p] .$$

4. Decomposição LU

Matrizes elementares. Relembremos que a linha i da matriz identidade tem zeros em todas as entradas excepto na i-ésima entrada que é igual a 1.

$$\begin{array}{c} [0\ \ 0\ \ ...\ \ 0\ \ 1\ \ 0\ \ ...\ \ 0] \\ i\text{-ésima entrada} \uparrow \end{array}$$

Fazendo o produto da i-ésima linha da matriz identidade por uma matriz $A = (a_{ij})$ e designando por L_1, L_2, ..., L_m, as linhas de A, tem-se

$$[0\ \ 0\ \ ...\ \ 0\ \ 1\ \ 0\ \ ...\ \ 0] \begin{bmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_m \end{bmatrix} = L_i = [a_{i1}\ \ a_{i2}\ \ ...\ \ a_{in}] .$$

Uma das passagens usada na eliminação de Gauss de uma matriz é substituir uma linha pela sua soma com um múltiplo de outra. Isso sugere a questão:

Dada uma matriz A cuja representação por linhas é

$$\begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{bmatrix},$$

qual a matriz cuja multiplicação por A tem o efeito de substituir a i-ésima linha de A pela sua soma com a j-ésima linha multiplicada pelo escalar α?

$$\left[(?)\right] \begin{bmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \\ L_m \end{bmatrix} = \begin{bmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ L_i + \alpha L_j \\ \vdots \\ L_m \end{bmatrix}$$

Note-se que, designando por e_1, e_2, ..., e_m as $1^{\underline{a}}$, $2^{\underline{a}}$, ... e m-ésima linhas da identidade de ordem m, respectivamente, temos que

$$\begin{aligned} (e_i + \alpha e_j)A &= e_i A + \alpha e_j A \\ &= L_i + \alpha L_j . \end{aligned}$$

Portanto, a matriz $\left[(?)\right]$ é $\begin{bmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ e_i + \alpha e_j \\ \vdots \\ e_m \end{bmatrix}$, ou seja, é a matriz

$$\begin{array}{cc}
 & \begin{array}{ccccccccc} & & & j & & i & & & \\ & & & \downarrow & & \downarrow & & & \end{array} \\
\begin{array}{c} \\ \\ \\ j \longrightarrow \\ \\ i \longrightarrow \\ \\ \\ \end{array} &
\begin{bmatrix}
1 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
0 & 0 & \ldots & 1 & \ldots & 0 & \ldots & 0 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
0 & 0 & \ldots & \alpha & \ldots & 1 & \ldots & 0 \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots \\
0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 1
\end{bmatrix}
\end{array}.$$

Esta matriz obtém-se da identidade substituindo a entrada nula na posição ij por α. Uma matriz deste tipo diz-se **matriz elementar** e designa-se por

$$E_{ij}(\alpha)\,.$$

Exemplos. A matriz elementar $E_{32}(-2)$ de ordem 4 é

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

e a matriz elementar $E_{41}(5)$ de ordem 4 é

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 5 & 0 & 0 & 1 \end{bmatrix}.$$

É fácil verificar que as matrizes elementares são invertíveis. Na verdade, como o efeito de multiplicar $E_{ij}(\alpha)$ por outra matriz qualquer é substituir a linha i pela sua soma com a linha j multiplicada por α, temos que

$$E_{ij}(\alpha)E_{ij}(-\alpha) = E_{ij}(\alpha)\begin{bmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ e_i - \alpha e_j \\ \vdots \\ e_m \end{bmatrix} = \begin{bmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ (e_i - \alpha e_j) + \alpha e_j \\ \vdots \\ e_m \end{bmatrix} = \begin{bmatrix} e_1 \\ \vdots \\ e_j \\ \vdots \\ e_i \\ \vdots \\ e_m \end{bmatrix} = I_m\,.$$

Analogamente se conclui que

$$E_{ij}(-\alpha)E_{ij}(\alpha) = I_m\,.$$

Por conseguinte,

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

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

Uma propriedade importante do produto de matrizes elementares. Um produto de matrizes elementares da forma

$$E_{i_1j_1}(\alpha_1)\, E_{i_2j_2}(\alpha_2) \ldots E_{i_sj_s}(\alpha_s)$$

onde $i_t > j_t$ $(t = 1, ..., s)$, $j_1 \leq j_2 \leq ... \leq j_s$ e $(i_t, j_t) \neq (i_{t'}, j_{t'})$ para $t \neq t'$ $(t, t' \in \{1, ..., s\})$, é igual à matriz que se obtém da matriz identidade substituindo cada entrada na posição i_tj_t $(t = 1, ..., s)$ por α_t.

(A demonstração deste facto fica como exercício, porventura só para os leitores mais audazes.)

Exemplos.

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

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

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

Matrizes elementares e eliminação de Gauss. Principiamos por um exemplo de aplicação do método de eliminação de Gauss:

$$A = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 4 & 3 & -2 & 0 \\ -2 & -3 & 0 & 2 \\ 6 & 1 & 0 & 1 \end{bmatrix} \xrightarrow[\substack{L_2 - 2L_1 \\ L_3 + L_1 \\ L_4 - 3L_1}]{\text{(a)}} A_1 = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & 1 & 2 & -2 \\ 0 & -2 & -2 & 3 \\ 0 & -2 & 6 & -2 \end{bmatrix} \xrightarrow[\substack{L_3 + 2L_2 \\ L_4 + 2L_2}]{\text{(b)}}$$

$$A_2 = \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 10 & -6 \end{bmatrix} \xrightarrow[L_4 - 5L_3]{\text{(c)}} \begin{bmatrix} 2 & 1 & -2 & 1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = U$$

Atendendo às passagens (a), (b) e (c) e às propriedades das matrizes elementares, temos, respectivamente:

$$\begin{aligned} &A_1 = E_{21}(-2)\, E_{31}(1)\, E_{41}(-3)\, A \\ &A_2 = E_{32}(2)\, E_{42}(2)\, A_1 \\ &U = E_{43}(-5)\, A_2 \end{aligned}$$

Portanto,

$$U = E_{43}(-5)\, E_{32}(2)\, E_{42}(2)\, E_{21}(-2)\, E_{31}(1)\, E_{41}(-3)\, A\,.$$

Pondo $B = E_{43}(-5)\, E_{32}(2)\, E_{42}(2)\, E_{21}(-2)\, E_{31}(1)\, E_{41}(-3)$, obtém-se

$$U = BA\,. \tag{1}$$

Como as matrizes elementares são invertíveis, a matriz B, que é um produto de matrizes elementares, é também invertível. Assim, da igualdade (1), vem

$$B^{-1}U = B^{-1}BA$$

e, finalmente,

$$A = LU$$

onde $L = B^{-1}$.

É agora fácil calcular a matriz L:

$$
\begin{aligned}
L = B^{-1} &= (E_{43}(-5)\,E_{32}(2)\,E_{42}(2)\,E_{21}(-2)\,E_{31}(1)\,E_{41}(-3))^{-1} \\
&= [E_{41}(-3)]^{-1}\,[E_{31}(1)]^{-1}\,[E_{21}(-2)]^{-1}\,[E_{42}(2)]^{-1}\,[E_{32}(2)]^{-1}\,[E_{43}(-5)]^{-1} \\
&= E_{41}(3)\,E_{31}(-1)\,E_{21}(2)\,E_{42}(-2)\,E_{32}(-2)\,E_{43}(5) \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & -2 & 1 & 0 \\ 3 & -2 & 5 & 1 \end{bmatrix}, \text{usando a propriedade descrita na subsecção anterior.}
\end{aligned}
$$

Vê-se assim que, para cada uma das operações do tipo "substituir a linha L_i por $L_i+\alpha L_j$" feitas ao longo da eliminação de Gauss da matriz A corresponde, na matriz L, o elemento $-\alpha$ na posição ij.

De um modo geral, se A é uma matriz $m \times n$ e U é a matriz em escada de linhas que resulta da eliminação de Gauss de A, ao longo da qual não houve troca de linhas, então a chamada **decomposição** LU de A é dada por

$$\boxed{A = LU}$$

onde L é a matriz $m \times m$ que se obtém da matriz identidade substituindo, para cada operação do tipo "a linha L_i passa a $L_i + \alpha L_j$" feita ao longo da eliminação de Gauss, a entrada nula na posição ij por $-\alpha$.

No caso particular de A ser uma matriz quadrada, a matriz U resultante da sua eliminação de Gauss é triangular superior. Como L é triangular inferior, a decomposição LU de A é então o produto de duas matrizes triangulares, tomando também o nome de **factorização triangular** de A.

Matrizes de permutação. Seja $A = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{bmatrix}$ uma matriz $m \times n$, representada pelas respectivas linhas. Seja i_1, i_2, ..., i_m uma reordenação dos naturais 1, 2, ..., m.

Pretendemos determinar uma matriz cuja multiplicação por $\begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{bmatrix}$ dê $\begin{bmatrix} L_{i_1} \\ L_{i_2} \\ \vdots \\ L_{i_m} \end{bmatrix}$.

$$\Big[(?) \Big] A = \begin{bmatrix} L_{i_1} \\ L_{i_2} \\ \vdots \\ L_{i_m} \end{bmatrix}$$

Denotando de novo por e_1, e_2, ..., e_m, as linhas da matriz identidade de ordem m, tem-se que

$$\begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_m} \end{bmatrix} A = \begin{bmatrix} e_{i_1}A \\ e_{i_2}A \\ \vdots \\ e_{i_m}A \end{bmatrix} = \begin{bmatrix} L_{i_1} \\ L_{i_2} \\ \vdots \\ L_{i_m} \end{bmatrix} \qquad \text{(Justifique.)}$$

Portanto, podemos pôr $\Big[(?) \Big] = \begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_m} \end{bmatrix}$.

Exemplos.

1. $A = \begin{bmatrix} 0 & 1 & 2 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \underset{L_1 \leftrightarrow L_2}{\longrightarrow} \begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} = B$; $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = B$.

2. A matriz que multiplicada por $\begin{bmatrix} 1 & 2 & -1 & 3 \\ -1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}$ dá $\begin{bmatrix} 0 & 2 & 1 & 1 \\ 1 & 2 & -1 & 3 \\ -1 & 1 & 0 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}$ é $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$,

ou seja,

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

Uma **matriz de permutação** de ordem m é uma matriz cujas linhas são todas as linhas da identidade I_m colocadas por uma qualquer ordem.

Exemplos. Alguns exemplos de matrizes de permutação:

$$P=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix},\ Q=\begin{bmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{bmatrix},\ R=\begin{bmatrix}0&0&0&1\\1&0&0&0\\0&0&1&0\\0&1&0&0\end{bmatrix}.$$

Como acabamos de ver, o efeito da multiplicação de uma matriz de permutação por outra matriz é trocar as linhas dessa matriz. Qual o efeito do produto da matriz P por outra? E das matrizes Q e R?

Exercício. Determinar uma matriz L triangular inferior com 1's na diagonal principal, uma matriz em escadas de linhas U e uma matriz de permutação P tal que para a matriz

$$A=\begin{bmatrix}0&2&1\\1&1&2\\2&3&8\end{bmatrix} \text{ se tenha } PA=LU.$$

Resolução:

$$\begin{bmatrix}0&2&1\\1&1&2\\2&3&8\end{bmatrix}\underset{L_1\leftrightarrow L_2}{\longrightarrow}\begin{bmatrix}1&1&2\\0&2&1\\2&3&8\end{bmatrix}\underset{L_3-2L_1}{\longrightarrow}\begin{bmatrix}1&1&2\\0&2&1\\0&1&4\end{bmatrix}\underset{L_3-\frac{1}{2}L_2}{\longrightarrow}\begin{bmatrix}1&1&2\\0&2&1\\0&0&\frac{7}{2}\end{bmatrix}$$

Por um lado, tem-se:

$$\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&2&1\\1&1&2\\2&3&8\end{bmatrix}=\begin{bmatrix}1&1&2\\0&2&1\\2&3&8\end{bmatrix}.$$

Por outro, e de acordo com o já visto mais atrás nesta secção,

$$\begin{bmatrix}1&1&2\\0&2&1\\2&3&8\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\2&\frac{1}{2}&1\end{bmatrix}\begin{bmatrix}1&1&2\\0&2&1\\0&0&\frac{7}{2}\end{bmatrix}.$$

Logo

$$L=\begin{bmatrix}1&0&0\\0&1&0\\2&\frac{1}{2}&1\end{bmatrix},\ U=\begin{bmatrix}1&1&2\\0&2&1\\0&0&\frac{7}{2}\end{bmatrix},\ P=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}.$$

Inversa de uma matriz de permutação. Dada uma matriz A do tipo $m\times n$, chama-se **matriz transposta** de A, e denota-se por A^t, à matriz $n\times m$ que se obtém de A, "escolhendo para 1ª, 2ª, ..., m-ésima colunas de A^t, precisamente as 1ª, 2ª, ..., m-ésima linhas de A, respectivamente"[1]. Assim, temos, por exemplo, que

[1] A transposição de matrizes goza de propriedades interessantes que podem ser estudadas nas aulas teorico-práticas. Merecem também atenção as noções de conjugada e transconjugada de uma matriz complexa.

$$\begin{bmatrix} 1 & -1 & 0 \\ -2 & 0 & 2 \\ 3 & 0 & -3 \\ 1 & 1 & 1 \end{bmatrix}^t = \begin{bmatrix} 1 & -2 & 3 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & 2 & -3 & 1 \end{bmatrix}.$$

Mais precisamente, dada uma matriz $A = [a_{ij}]_{1\le i\le m}^{1\le j\le n}$, do tipo $m \times n$, a sua transposta é a matriz $n \times m$ dada por

$$A^t = [a_{ji}]_{1\le j\le n}^{1\le i\le m}.$$

É claro que se transpusermos uma matriz de permutação obtemos outra matriz de permutação. Por exemplo,

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

Seja

$$P = \begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_n} \end{bmatrix}$$

uma matriz de permutação, escrita por linhas. Portanto, e_{i_1}, e_{i_2}, ..., e_{i_n} são as linhas da identidade I_n, dispostas por uma certa ordem. A matriz transposta de P é, escrita por colunas,

$$P^t = \begin{bmatrix} e_{i_1}^t & e_{i_2}^t & \ldots & e_{i_n}^t \end{bmatrix}.$$

Vamos agora calcular os produtos PP^t e P^tP, utilizando multiplicação por blocos.

$$PP^t = \begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_n} \end{bmatrix} \begin{bmatrix} e_{i_1}^t & e_{i_2}^t & \ldots & e_{i_n}^t \end{bmatrix} = \begin{bmatrix} e_{i_1}e_{i_1}^t & e_{i_1}e_{i_2}^t & \ldots & e_{i_1}e_{i_n}^t \\ e_{i_2}e_{i_1}^t & e_{i_2}e_{i_2}^t & \ldots & e_{i_2}e_{i_n}^t \\ \vdots & \vdots & \ldots & \vdots \\ e_{i_n}e_{i_1}^t & e_{i_n}e_{i_2}^t & \ldots & e_{i_n}e_{i_n}^t \end{bmatrix}$$

Ora,

$$e_{i_r}e_{i_s}^t = \overset{\substack{i_r\text{-ésima posição} \\ \downarrow}}{[0 \ldots 0\ 1\ 0 \ldots 0]} \begin{array}{r} \\ \\ \\ i_s\text{-ésima posição} \longrightarrow \\ \\ \\ \end{array} \begin{bmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix} = \begin{cases} 1 \text{ se } r = s \\ 0 \text{ se } r \neq s \end{cases}$$

Logo

$$PP^t = I_n\,. \tag{2}$$

$$P^tP = \begin{bmatrix} e^t_{i_1} & e^t_{i_2} & \dots & e^t_{i_n} \end{bmatrix} \begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_n} \end{bmatrix} = \sum_{k=1}^{n} e^t_{i_k} e_{i_k}$$

e, como para cada k,

$$e^t_{i_k} e_{i_k} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [0 \dots 0\ 1\ 0 \dots 0] = \begin{array}{rc} & \begin{array}{c} i_k\text{-ésima coluna} \\ \downarrow \end{array} \\ i_k\text{-ésima linha} \rightarrow & \begin{bmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \end{array},$$

conclui-se que

$$P^tP = I_n \tag{3}$$

De (2) e (3), conclui-se que a transposta de P é a sua inversa. Em conclusão: Toda a matriz de permutação P é invertível e tem-se $P^t = P$.

Exemplos. A inversa de $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ é $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

A inversa de $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ é $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

5. Método de Gauss-Jordan. Cálculo da inversa de uma matriz

A resolução de um sistema pelo método de Gauss-Jordan compreende três fases:

1. Eliminação de Gauss da matriz ampliada do sistema.

   (Só interessa passar à fase seguinte se o sistema for possível.)

2. A partir da matriz em escada por linhas obtida em 1, chegar a uma matriz que, não só por baixo, mas também por cima dos pivôs, tem todas as entradas nulas, por meio de adequadas operações sucessivas sobre as linhas.

3. Multiplicação de cada linha não nula da matriz obtida em 2 pelo inverso do número que é o pivô dessa linha.

Ao tirar da matriz final o sistema correspondente, obtemos imediatamente as soluções do sistema.

Exemplo. O sistema de equações lineares nas variáveis x_1, x_2, x_3, x_4 e x_5,

$$\begin{aligned}
&2x_1 + 3x_2 - 2x_3 + x_4 - 5x_5 = 3\\
&2x_1 - 3x_3 + 3x_4 + x_5 = 4\\
&-2x_1 + 3x_2 + 3x_3 - 4x_4 - 6x_5 = -4\\
&4x_1 + 3x_2 - 3x_3 + 2x_4 - 6x_5 = 5
\end{aligned}$$

tem como matriz ampliada

$$\left[\begin{array}{rrrrrcr} 2 & 3 & -2 & 1 & -5 & | & 3\\ 2 & 0 & -3 & 3 & 1 & | & 4\\ -2 & 3 & 3 & -4 & -6 & | & -4\\ 4 & 3 & -3 & 2 & -6 & | & 5 \end{array}\right].$$

Vamos fazer a eliminação de Gauss-Jordan desta matriz.
Primeiro fazemos a eliminação de Gauss:

$$\left[\begin{array}{rrrrrcr} 2 & 3 & -2 & 1 & -5 & | & 3\\ 2 & 0 & -3 & 3 & 1 & | & 4\\ -2 & 3 & 3 & -4 & -6 & | & -4\\ 4 & 3 & -3 & 2 & -6 & | & 5 \end{array}\right] \begin{array}{c} \longrightarrow \\ L_2 - L_1 \\ L_3 + L_1 \\ L_4 - 2L_1 \end{array} \left[\begin{array}{rrrrrcr} 2 & 3 & -2 & 1 & -5 & | & 3\\ 0 & -3 & -1 & 2 & 6 & | & 1\\ 0 & 6 & 1 & -3 & -11 & | & -1\\ 0 & -3 & 1 & 0 & 4 & | & -1 \end{array}\right]$$

$$\begin{array}{c} \longrightarrow \\ \\ L_3 + 2L_2 \\ L_4 - L_2 \end{array} \left[\begin{array}{rrrrrcr} 2 & 3 & -2 & 1 & -5 & | & 3\\ 0 & -3 & -1 & 2 & 6 & | & 1\\ 0 & 0 & -1 & 1 & 1 & | & 1\\ 0 & 0 & 2 & -2 & -2 & | & -2 \end{array}\right] \begin{array}{c} \longrightarrow \\ \\ \\ L_4 + 2L_3 \end{array} \left[\begin{array}{rrrrrcr} 2 & 3 & -2 & 1 & -5 & | & 3\\ 0 & -3 & -1 & 2 & 6 & | & 1\\ 0 & 0 & -1 & 1 & 1 & | & 1\\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{array}\right]$$

Como a característica da matriz dos coeficientes é igual à característica da matriz ampliada, o sistema é possível. Além disso, como essa característica é igual a 3 e a matriz dos coeficientes tem 5 colunas, vamos obter 3 variáveis básicas e 2 variáveis livres.

Passa-se agora à segunda fase: Primeiro anulam-se as entradas acima do pivô da última linha não nula, i.e., da 3ª linha:

$$\left[\begin{array}{rrrrrcr} 2 & 3 & -2 & 1 & -5 & | & 3\\ 0 & -3 & -1 & 2 & 6 & | & 1\\ 0 & 0 & -1 & 1 & 1 & | & 1\\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{array}\right] \begin{array}{c} \longrightarrow \\ L_1 - 2L_3 \\ L_2 - L_3 \\ \\ \\ \end{array} \left[\begin{array}{rrrrrcr} 2 & 3 & 0 & -1 & -7 & | & 1\\ 0 & -3 & 0 & 1 & 5 & | & 0\\ 0 & 0 & -1 & 1 & 1 & | & 1\\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{array}\right]$$

Em seguida anulam-se as entradas acima do pivô da 2ª linha:

$$\underset{L_1+L_2}{\longrightarrow} \left[\begin{array}{ccccc|c} 2 & 0 & 0 & 0 & -2 & 1 \\ 0 & -3 & 0 & 1 & 5 & 0 \\ 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Finalizada a 2ª fase, passamos agora à parte final do método de eliminação de Gauss-Jordan:

$$\begin{array}{c} \longrightarrow \\ \frac{1}{2}L_1 \\ -\frac{1}{3}L_2 \\ -L_3 \\ \end{array} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -1 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{3} & -\frac{5}{3} & 0 \\ 0 & 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Esta matriz corresponde ao sistema

$$\begin{array}{l} x_1 - x_5 = \frac{1}{2} \\ x_2 - \frac{1}{3}x_4 - \frac{5}{3}x_5 = 0 \\ x_3 - x_4 - x_5 = -1 \end{array}$$

que é equivalente a

$$\begin{array}{l} x_1 = x_5 + \frac{1}{2} \\ x_2 = \frac{1}{3}x_4 + \frac{5}{3}x_5 \\ x_3 = x_4 + x_5 - 1 \end{array}$$

sendo x_1, x_2 e x_3 variáveis básicas e x_4 e x_5 variáveis livres.

Note-se que, quando o sistema dado é possível determinado, a sua solução única aparece, no final da eliminação de Gauss-Jordan, na coluna dos termos independentes.

Exemplo. Resolução do sistema

$$\begin{array}{l} x_1 + x_2 = 2 \\ x_1 - x_2 = 0 \end{array}$$

pelo método de Gauss-Jordan:

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 1 & -1 & 0 \end{array}\right] \underset{L_2-L_1}{\longrightarrow} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & -2 & -2 \end{array}\right] \underset{L_1+\frac{1}{2}L_2}{\longrightarrow} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & -2 & -2 \end{array}\right]$$

$$\underset{-\frac{1}{2}L_2}{\longrightarrow} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right]$$

Conclui-se que o sistema dado é equivalente a

$$\begin{array}{l} x_1 = 1 \\ x_2 = 1 \end{array}.$$

Portanto, a coluna dos termos independentes da última matriz ampliada dá-nos a solução do sistema.

Aplicação do método de Gauss-Jordan ao cálculo da inversa de uma matriz invertível. Seja A uma matriz quadrada de ordem n. Pretendemos saber se A é invertível e, nesse caso, determinar A^{-1}.

Por definição, uma matriz B de ordem n é inversa de A se $AB = I$ e $BA = I$. Vamos começar por ver como determinar uma matriz B de ordem n tal que $AB = I$ (caso tal matriz B exista).

Ora, determinar uma tal matriz B é equivalente a determinar matrizes coluna $n \times 1$

$$C_1,\ C_2,\ \ldots\ C_n$$

tais que

$$A[C_1\ \ C_2\ \ \ldots\ \ C_n] = I,$$

ou seja,

$$[AC_1\ \ AC_2\ \ \ldots AC_n] = I,$$

isto é,

$$AC_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ AC_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ \ldots,\ AC_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

O objectivo é pois determinar C_j, para cada $j \in \{1, 2, \ldots, n\}$, tal que

$$\begin{array}{r} AC_j = \\ \\ \\ j\text{-ésima linha} \rightarrow \\ \\ \\ \\ \end{array} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Ou seja, para cada $j \in \{1, 2, \ldots, n\}$, pretendemos determinar a(s) solução(s) do sistema, de n equações e n incógnitas,

$$\begin{array}{r} Ax = \\ \\ \\ j\text{-ésima linha} \rightarrow \\ \\ \\ \\ \end{array} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \qquad (S_j)$$

Consequentemente, existe uma matriz B de ordem n tal que $AB = I$ se e só se os n sistemas (S_1), (S_2), ..., (S_j), ..., (S_n) são possíveis; nesse caso, B é a matriz cujas colunas são as soluções dos sistemas (S_1), (S_2), ..., (S_j), ..., (S_n), por esta ordem.

A propriedade seguinte é importante para o que se segue.

Propriedade 1. Se A e B são matrizes $n \times n$ tais que $AB = I$ então também $BA = I$ e, consequentemente, $B = A^{-1}$.

Prova. Como se verá no Exercício da Secção 11, temos sempre a seguinte desigualdade respeitante à característica: $\text{car}(AB) \leq \text{car}(A)$. Usando este facto, e tendo em conta que A tem n colunas, pelo que $\text{car}(A) \leq n$, obtemos

$$n = \text{car}(I) = \text{car}(AB) \leq \text{car}A \leq n$$

donde se conclui que $\text{car}(A) = n$.

Por outro lado, como $AB = I$, temos a igualdade $ABA = AI$, ou ainda, usando a propriedade distributiva,

$$A(BA - I) = 0\,. \tag{4}$$

Mas, como $\text{car}(A) = n$, o sistema homogéneo $Ax = 0$ é possível determinado, admitindo a solução nula como única solução. Por conseguinte, a igualdade (4) obriga a que $BA - I = 0$, ou seja, $BA = I$, como pretendido.

Note-se que uma leitura atenta da prova acabada de fazer facilmente levará à conclusão da póxima propriedade (Justifique).

Propriedade 2. Uma matriz quadrada de ordem n é invertível se e só se $\text{car}A = n$.

Sabemos mais: Se $\text{car}A = n$ então a sua inversa é a matriz cujas $1^{\underline{a}}$, $2^{\underline{a}}$, ... e n-ésima colunas são as soluções dos sistemas (S_1), (S_2), ..., (S_n), respectivamente.

Claro que, para resolver esses sistemas, podemos decidir por vários métodos. Vamos ver que, usando o método de Gauss-Jordan, podemos optar por uma resolução simultânea dos n sistemas que nos leva à inversa da matriz A, quando ela existe, de uma forma simples e relativamente económica. Para melhor compreensão do processo, vamos ilustrá-lo com um exemplo:

Seja $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 2 \\ -1 & 1 & -2 \end{bmatrix}$. Para determinar a inversa de A, caso exista, vamos resolver os correspondentes sistemas (S_1), (S_2) e (S_3).

Neste caso, o primeiro sistema é, na forma matricial, o seguinte:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 2 \\ -1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{S_1}$$

Façamos a eliminação de Gauss-Jordan da matriz ampliada do sistema:

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

$$\underset{\substack{L_1+\frac{6}{5}L_3\\ L_2-\frac{14}{5}L_3}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 0 & \frac{2}{5} \\ 0 & -6 & 0 & -\frac{8}{5} \\ 0 & 0 & -\frac{5}{2} & -\frac{1}{2} \end{array}\right] \longrightarrow L_1+\tfrac{1}{3}L_2 \left[\begin{array}{ccc|c} 1 & 20 & 0 & -\frac{2}{15} \\ 0 & -6 & 0 & -\frac{8}{5} \\ 0 & 0 & -\frac{5}{2} & -\frac{1}{2} \end{array}\right]$$

$$\underset{\substack{-\frac{1}{6}L_2\\ -\frac{2}{5}L_3}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{2}{15} \\ 0 & 1 & 0 & \frac{4}{15} \\ 0 & 0 & 1 & \frac{1}{5} \end{array}\right]$$

Consequentemente, o sistema (S_1) é equivalente a

$$\begin{aligned} x_1 &= -\tfrac{2}{15} \\ x_2 &= \tfrac{4}{15} \\ x_3 &= \tfrac{1}{5} \end{aligned}$$

ou seja, o sistema (S_1) é possível determinado, tendo $(-\frac{2}{15}, \frac{4}{15}, \frac{1}{5})$ como solução única.

Neste momento, como $\mathrm{car}(A) = 3$, sendo A de ordem 3, sabemos já que os outros sistemas, (S_2) e (S_3), são também possíveis e determinados; logo A tem inversa.

A $1^{\underline{a}}$ coluna da matriz inversa de A é então a solução do sistema (S_1), ou seja,

$$\begin{bmatrix} -\frac{2}{15} \\ \frac{4}{15} \\ \frac{1}{15} \end{bmatrix}.$$

Resolvendo os sistemas

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 2 \\ -1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \qquad (S_2)$$

e

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 2 \\ -1 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \qquad (S_3)$$

obtemos, respectivamente, as $2^{\underline{a}}$ e $3^{\underline{a}}$ colunas da matriz inversa de A. São elas:

$$2^{\underline{a}}\text{ coluna}: \begin{bmatrix} \frac{7}{15} \\ \frac{1}{15} \\ -\frac{1}{5} \end{bmatrix}; \quad 3^{\underline{a}}\text{ coluna}: \begin{bmatrix} \frac{4}{15} \\ \frac{7}{15} \\ -\frac{2}{5} \end{bmatrix}.$$

Considerando os resultados dos três sistemas, temos que

$$A^{-1} = \begin{bmatrix} -\frac{2}{15} & \frac{7}{15} & \frac{4}{15} \\ \frac{4}{15} & \frac{1}{15} & \frac{7}{15} \\ \frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \end{bmatrix}.$$

Note-se agora que, tendo em conta que os três sistemas têm a mesma matriz dos coeficientes, as operações feitas sobre as linhas ao longo da eliminação de Gauss-Jordan das matrizes ampliadas dos três sistemas são as mesmas em cada um deles. Assim, podem resolver-se os três sistemas todos ao mesmo tempo fazendo a eliminação de Gauss-Jordan da matriz

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & 0 & 2 & 0 & 1 & 0 \\ -1 & 1 & -2 & 0 & 0 & 1 \end{array}\right].$$

Temos então:

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & 0 & 2 & 0 & 1 & 0 \\ -1 & 1 & -2 & 0 & 0 & 1 \end{array}\right] \underset{\substack{L_2 - 3L_1 \\ L_3 + L_1}}{\longrightarrow} \left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -6 & -7 & -3 & 1 & 0 \\ 0 & 3 & 1 & 1 & 0 & 1 \end{array}\right] \underset{L_3 + \frac{1}{2}L_2}{\longrightarrow}$$

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -6 & -7 & -3 & 1 & 0 \\ 0 & 0 & -\frac{5}{2} & -\frac{1}{2} & \frac{1}{2} & 1 \end{array}\right] \underset{\substack{L_1 + \frac{6}{5}L_3 \\ L_2 - \frac{14}{5}L_3}}{\longrightarrow} \left[\begin{array}{rrr|rrr} 1 & 2 & 0 & \frac{2}{5} & \frac{3}{5} & \frac{6}{5} \\ 0 & -6 & 0 & -\frac{8}{5} & -\frac{2}{5} & -\frac{14}{5} \\ 0 & 0 & -\frac{5}{2} & -\frac{1}{2} & \frac{1}{2} & 1 \end{array}\right]$$

$$\underset{L_1 + \frac{1}{3}L_2}{\longrightarrow} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{2}{15} & \frac{7}{15} & \frac{4}{15} \\ 0 & -6 & 0 & -\frac{8}{5} & -\frac{2}{5} & -\frac{14}{5} \\ 0 & 0 & -\frac{5}{2} & -\frac{1}{2} & \frac{1}{2} & 1 \end{array}\right] \underset{\substack{-\frac{1}{6}L_2 \\ -\frac{2}{5}L_3}}{\longrightarrow} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{2}{15} & \frac{7}{15} & \frac{4}{15} \\ 0 & 1 & 0 & \frac{4}{15} & \frac{1}{15} & \frac{7}{15} \\ 0 & 0 & 1 & \frac{1}{5} & -\frac{1}{5} & -\frac{2}{5} \end{array}\right]$$

Na última matriz, a submatriz da direita tem como colunas as soluções dos sistemas (S_1), (S_2) e (S_3). Ela é, portanto, a inversa de A.

De um modo geral, dada uma matriz A quadrada de ordem n, para determinar B tal que $AB = I$, podemos resolver os sistemas (S_1), (S_2), ..., (S_n), descritos atrás para o caso geral, todos ao mesmo tempo, de forma análoga ao feito no final do exemplo anterior.

Esquematizando:

$$[A|I] \longrightarrow \underset{\text{eliminação de Gauss-Jordan}}{\ldots\ldots\ldots\ldots\ldots\ldots} \longrightarrow [I|A^{-1}]$$

O processo acabado de descrever é o do cálculo da inversa de uma matriz pelo método de Gauss-Jordan.

II

Espaços vectoriais

6. Espaços vectoriais

Um **espaço vectorial** sobre $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$) consiste num conjunto não vazio $\mathcal{V}$ equipado com duas operações, a adição

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$$

(que a cada dois elementos $\mathbf{u}$ e $\mathbf{v}$ de $\mathcal{V}$ faz corresponder um elemento $\mathbf{u} + \mathbf{v}$ de $\mathcal{V}$) e a multiplicação escalar

$$. : \mathbb{K} \times \mathcal{V} \to \mathcal{V}$$

(que a cada dois elementos $\alpha \in \mathbb{K}$ e $\mathbf{v} \in \mathcal{V}$ faz corresponder um elemento $\alpha\mathbf{v}$ de $\mathcal{V}$),

satisfazendo os seguintes axiomas:

(V1) Comutatividade da adição: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, para $\mathbf{u}, \mathbf{v} \in \mathcal{V}$;

(V2) Associatividade da adição: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$, para $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$;

(V3) Existência de zero: existe um elemento de $\mathcal{V}$, designado por $\vec{\mathbf{0}}$, ou apenas por $\mathbf{0}$, tal que $\mathbf{v} + \mathbf{0} = \mathbf{v}$, para todo $\mathbf{v} \in \mathcal{V}$;

(V4) Existência de simétrico: para todo o elemento $\mathbf{v}$ de $\mathcal{V}$, existe um elemento $\mathbf{u}$ tal que $\mathbf{u} + \mathbf{v} = \mathbf{0}$ (este elemento $\mathbf{u}$ é habitualmente denotado por $-\mathbf{v}$);

(V5) Associatividade mista: $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$, para todo $\alpha, \beta \in \mathbb{K}$ e todo $\mathbf{v} \in \mathcal{V}$;

(V6) Distributividade da multiplicação escalar em relação à adição em $\mathbb{K}$: $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$, para todo $\alpha, \beta \in \mathbb{K}$ e todo $\mathbf{v} \in \mathcal{V}$;

(V7) Distributividade da multiplicação escalar em relação à adição em $\mathcal{V}$: $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$, para todo $\alpha \in \mathbb{K}$ e todo $\mathbf{u}, \mathbf{v} \in \mathcal{V}$;

(V8) Existência de identidade: $1\mathbf{v} = \mathbf{v}$, para todo $\mathbf{v} \in \mathcal{V}$.

Quando $\mathbb{K} = \mathbb{R}$, dizemos que $\mathcal{V}$ é um **espaço vectorial real**; se $\mathbb{K} = \mathbb{C}$, dizemos que $\mathcal{V}$ é um **espaço vectorial complexo**. Os elementos de um espaço vectorial dizem-se **vectores** e os elementos de $\mathbb{K}$ são designados por **escalares**. Em particular, $\vec{\mathbf{0}}$ diz-se o **vector nulo** de $\mathcal{V}$.

Exemplos.

1. $\mathbb{R}^n$ com a adição e a multiplicação por um real usuais é um espaço vectorial real.
2. $\mathbb{C}^n$ com a adição e a multiplicação por um complexo usuais é um espaço vectorial complexo.
3. O conjunto de matrizes $\mathcal{M}_{m\times n}(\mathbb{R})$ com a adição e a multiplicaçãoescalar por um número real é um espaço vectorial real.
4. Analogamente, $\mathcal{M}_{m\times n}(\mathbb{C})$ é um espaço vectorial complexo.
5. O conjunto de todas as sucessões de números reais com a adição e a multiplicação escalar definidas componente a componente, i.e., $(x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}} = (x_n + y_n)_{n\in\mathbb{N}}$ e $\alpha(x_n)_{n\in\mathbb{N}} = (\alpha x_n)_{n\in\mathbb{N}}$, é um espaço vectorial real. Este espaço costuma designar-se por $\mathbb{R}^{\mathbb{N}}$.
6. O conjunto das funções reais de variável real definidas num conjunto $S \subseteq \mathbb{R}$ com a adição e a multiplicão por um número real usuais, i.e., $(f+g)(x) = f(x)+g(x)$ e $(\alpha f)(x) = \alpha f(x)$, é um espaço vectorial real.
7. O conjunto dos polinómios reais com as operações usuais é um espaço vectorial real.

A seguir enumeram-se algumas propriedades que se deduzem directamente da definição de espaço vectorial.

Propriedades. Sendo v e w vectores quaisquer de um espaço vectorial $\mathcal{V}$ e α e β escalares, verificam-se as propriedades seguintes:

1. $\alpha(v - w) = \alpha v - \alpha w$;
2. $(\alpha - \beta)v = \alpha v - \beta v$;
3. $\alpha 0 = \mathbf{0}$;
4. $0v = \mathbf{0}$;
5. $\alpha v = \mathbf{0} \Longrightarrow \alpha = 0$ ou $v = \mathbf{0}$;
6. $(\alpha v = \alpha w$ e $\alpha \neq 0) \Longrightarrow v = w$;
7. $(\alpha v = \beta v$ e $v \neq \mathbf{0}) \Longrightarrow \alpha = \beta$.

Segue-se a demonstração das propriedades 2 e 4 anteriores. A prova das outras fica como exercício.

Prova de 2. $\alpha v = (\alpha - \beta + \beta)v = (\alpha - \beta)v + \beta v$, pelo axioma $(V6)$. Agora, somando o simétrico de βv a ambos os membros da igualdade $\alpha v = (\alpha - \beta)v + \beta v$, obtém-se $\alpha v - \beta v = (\alpha - \beta)v + \beta v - \beta v = (\alpha - \beta)v + \mathbf{0} = (\alpha - \beta)v$.

Prova de 4. Usando a propriedade 2, temos $0v = (0 - 0)v = 0v - 0v = \mathbf{0}$.

7. Subespaços vectoriais

Um subconjunto $\mathcal{W}$ de um espaço vectorial $\mathcal{V}$ diz-se um **subespaço vectorial** de $\mathcal{V}$ se $\mathcal{W}$ for um espaço vectorial para as operações de adição e multiplicação escalar definidas em $\mathcal{V}$.

Exemplos.

1. $\{(x, 0) : x \in \mathbb{R}\}$ é um subespaço vectorial de $\mathbb{R}^2$.

2. O conjunto de todas as matrizes reais $n \times n$ triangulares superiores (bem como o de todas as triangulares inferiores) é um subespaço vectorial de $\mathcal{M}_{n\times n}(\mathbb{R})$.

3. O conjunto de todas as matrizes complexas $n \times n$ triangulares superiores é um subespaço vectorial de $\mathcal{M}_{n\times n}(\mathbb{C})$. O mesmo acontece com as triangulares inferiores.

4. O conjunto de todas as funções reais de variável real contínuas num intervalo $[a, b]$ é um subespaço vectorial do espaço de todas as funções reais de variável real definidas no intervalo $[a, b]$, com as operações usuais. Este subespaço será designado por $C[a, b]$.

5. O conjunto de todos os polinómios reais de grau menor ou igual a n é um subespaço vectorial do espaço de todos os polinómios reais.

Teorema 1. Um subconjunto $\mathcal{W}$ de um espaço vectorial $\mathcal{V}$ é um subespaço vectorial de $\mathcal{V}$ se e só se:

1. $\mathcal{W} \neq \emptyset$;

2. $\mathcal{W}$ é fechado para a adição e a multiplicação escalar definidas em $\mathcal{V}$.

Demonstração. Se $\mathcal{W}$ é um subespaço vectorial de $\mathcal{V}$, então, por definição, é imediato que as condições 1 e 2 se verificam.

Reciprocamente, suponhamos que $\mathcal{W}$ é um subconjunto de $\mathcal{V}$ satisfazendo as condições 1 e 2. Antes de mais, por 2, $\mathcal{W}$ é fechado para a adição e a multiplicação escalar. É necessário agora assegurar a verificação dos axiomas de $(V1)$ a $(V8)$ da definição de espaço vectorial. Quanto aos axiomas $(V1)$ e $(V2)$ e os de $(V5)$ a $(V8)$, é óbvio que, sendo eles verdadeiros em $\mathcal{V}$, também o são em $\mathcal{W}$. Resta então provar a veracidade dos axiomas $(V3)$ e $(V4)$ para $\mathcal{W}$. Para isso, é suficiente mostrar que $\mathbf{0} \in \mathcal{W}$ e que se $u \in \mathcal{W}$ também $-u \in \mathcal{W}$. Como $\mathcal{W} \neq \emptyset$, existe algum vector v de $\mathcal{V}$ em $\mathcal{W}$. Mas então, como $\mathcal{W}$ é fechado para a multiplicação escalar, $0v \in \mathcal{W}$ e, atendendo a 4 das Propriedades enunciadas na secção anterior, o vector nulo $\mathbf{0}$ pertence a $\mathcal{W}$. Seja agora $u \in \mathcal{W}$. Como $-u = (-1)u$, atendendo às propriedades 2 e 4 da mesma secção, e $\mathcal{W}$ é fechado para a multiplicação escalar, conclui-se que $-u \in \mathcal{W}$.

Para concluir o teorema a seguir basta mostrar que a verificação das condições 1 e 2 desse teorema é equivalente à ocorrência das condições 1 e 2 do Teorema 1.

Teorema 2. Um subconjunto $\mathcal{W}$ de um espaço vectorial $\mathcal{V}$ é um subespaço vectorial de $\mathcal{V}$ se e só se:

1. $\mathbf{0} \in \mathcal{W}$;
2. $\mathcal{W}$ é fechado para a adição e a multiplicação escalar definidas em $\mathcal{V}$.

8. Subespaço gerado por um conjunto. Conjuntos geradores

Sejam $v_1, v_2, ..., v_n$ vectores de um espaço $\mathcal{V}$. Um vector $v \in \mathcal{V}$ diz-se uma **combinação linear** dos vectores $v_1, v_2, ..., v_n$ se existirem escalares $\alpha_1, \alpha_2, ..., \alpha_n$ tais que

$$v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n.$$

Neste caso os escalares $\alpha_1, \alpha_2, ..., \alpha_n$ dizem-se **coeficientes da combinação linear**. Seja $\mathcal{V}$ um espaço vectorial e S um subconjunto de $\mathcal{V}$. Consideremos o subconjunto $\mathcal{L}(S)$ de $\mathcal{V}$ constituído por todas as combinações lineares de elementos de S, i.e.,

$$\mathcal{L}(S) = \{\alpha_1 s_1 + \alpha_2 s_2 + ... + \alpha_k s_k \,:\, s_1, s_2, ..., s_k \in S,\ \alpha_1, \alpha_2, ..., \alpha_k \text{ escalares},\ k \in \mathbb{N}\}.$$

Não é difícil verificar que $\mathcal{L}(S)$ é um subespaço vectorial de $\mathcal{V}$. Se $S = \emptyset$, convenciona-se $\mathcal{L}(S) = \{0\}$. O subespaço $\mathcal{L}(S)$ diz-se o **subespaço vectorial de $\mathcal{V}$ gerado por S**. Na verdade, ele é o "menor" subespaço de $\mathcal{V}$ que contém S.

Se S é finito, digamos, $S = \{s_1, s_2, ..., s_n\}$, então é claro que

$$\mathcal{L}(S) = \{\alpha_1 s_1 + \alpha_2 s_2 + ... + \alpha_n s_n \ : \ \alpha_1, \alpha_2, ..., \alpha_n \text{ escalares}\}.$$

Dizemos que um subconjunto não vazio S do espaço vectorial $\mathcal{V}$ é um **conjunto gerador de** $\mathcal{V}$, ou que S **gera** $\mathcal{V}$, ou que os vectores de S **geram** $\mathcal{V}$, se qualquer vector de $\mathcal{V}$ se pode escrever como combinação linear de elementos de S, ou seja, $\mathcal{V} = \mathcal{L}(S)$.

Exemplos.

1. $S_1 = \{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ é um conjunto gerador de $\mathbb{R}^3$ (justifique);
2. $S_2 = \{(1,0,0),\ (0,1,0),\ (0,0,1), (1,1,1)\}$ é também um conjunto gerador de $\mathbb{R}^3$ (justifique);
3. $S_3 = \{(3,2,1),\ (0,3,2),\ (0,0,3)\}$ é outro conjunto gerador de $\mathbb{R}^3$ (verifique);
4. $S_4 = \{(1,0,1),\ (3,0,2)\}$ não é um conjunto gerador de $\mathbb{R}^3$ (porquê?).

Proposição. Sejam $v_1, v_2, ..., v_n$ vectores de um espaço vectorial.

1. Dados um escalar $\alpha \neq 0$ e $i \in \{1, ..., n\}$, os vectores $v_1, v_2, ..., v_n$ geram $\mathcal{V}$ se e só se $v_1, ..., \alpha v_i, ..., v_n$ gerarem $\mathcal{V}$.
2. Dados um escalar α e $i, j \in \{1, ..., n\}$, com $i \neq j$, os vectores $v_1, v_2, ..., v_n$ geram $\mathcal{V}$ se e só se $v_1, ..., v_j, ..., v_i + \alpha v_j, ..., v_n$ gerarem $\mathcal{V}$.

Demonstração. 1. Suponhamos que $v_1, v_2, ..., v_n$ geram $\mathcal{V}$ e seja u um qualquer vector de $\mathcal{V}$. Por hipótese, existem escalares β_1, β_2, ..., β_n tais que

$$u = \beta_1 v_1 + ... + \beta_i v_i + ... + \beta_n v_n\,.$$

Como $\alpha \neq 0$, existe o escalar $\frac{\beta_i}{\alpha}$ e temos

$$u = \beta_1 v_1 + ... + \frac{\beta_i}{\alpha}(\alpha v_i) + ... + \beta_n v_n\,,$$

ou seja, u escreve-se como combinação linear de $v_1, ..., \alpha v_i, ..., v_n$.

Reciprocamente, suponhamos agora que os vectores $v_1, ..., \alpha v_i, ..., v_n$ geram $\mathcal{V}$. Dado $u \in \mathcal{V}$, sejam γ_1, γ_2, ..., γ_n tais que

$$u = \gamma_1 v_1 + ... + \gamma_i(\alpha v_i) + ... + \gamma_n v_n\,.$$

Então

$$u = \gamma_1 v_1 + ... + (\gamma_i \alpha) v_i + ... + \gamma_n v_n$$

é uma combinação linear dos vectores $v_1, ..., v_i, ..., v_n$.

2. Fica como exercício.

9. Dependência e independência linear

Seja $\mathcal{V}$ um espaço vectorial e sejam $v_1, v_2, ..., v_n$ vectores de $\mathcal{V}$. Dizemos que **os vectores $v_1, v_2, ..., v_n$ são linearmente independentes**, ou que **o conjunto $\{v_1, v_2, ..., v_n\}$ é linearmente independente**, se a única maneira de escrever o vector nulo como combinação linear de $v_1, v_2, ..., v_n$ é com todos os escalares iguais a zero, isto é, se

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \mathbf{0} \Rightarrow \alpha_1 = \alpha_2 = ... = \alpha_n = 0.$$

Se os vectores $v_1, v_2, ..., v_n$ não são linearmente independentes, diz-se que são **linearmente dependentes** ou que **o conjunto $\{v_1, v_2, ..., v_n\}$ é linearmente dependente**

Exemplos. Os conjuntos S_1, S_3 e S_4 dos Exemplos da Secção 8 são linearmente independentes no espaço vectorial $\mathbb{R}^3$; mas S_2 não é linearmente independente. (Justifique.)

Dependência e independência linear em espaços $\mathbb{R}^n$ e $\mathbb{C}^n$. Sejam $v_1, v_2, ..., v_t \in \mathbb{K}^m$ (com $\mathbb{K} = \mathbb{R}$, $\mathbb{C}$). Podemos considerar os v_i's como matrizes $m \times 1$; temos então:

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_t v_t = \mathbf{0} \Leftrightarrow [v_1 \ v_2 \ ... \ v_t] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_t \end{bmatrix} = \mathbf{0} \tag{5}$$

Consequentemente, $v_1, v_2, ..., v_t$ são linearmente independentes se e só se o sistema homogéneo (5) é determinado, ou seja, se e só se $\text{car}([v_1 \ v_2 \ ... \ v_t]) = t$. Portanto, nestes casos, para analisar a dependência ou independência linear dos vectores, podemos usar a eliminação de Gauss.

Observação. Atendendo à definição de dependência linear, é imediato que:

1. Dados os vectores $v_1, v_2, ..., v_n$, a sua dependência (ou independência) linear não depende da ordem pela qual são considerados.

2. Um vector v é linearmente independente se e só se $v \neq \mathbf{0}$.

Proposição 1. Seja $\mathcal{V}$ um espaço vectorial e $v_1, v_2, ..., v_n$ vectores de $\mathcal{V}$. Então $v_1, v_2, ..., v_n$ são linearmente dependentes se e só se um dos vectores se pode escrever como combinação linear dos restantes.

Demonstração. Admitamos que os vectores $v_1, v_2, ..., v_n$ são linearmente dependentes, isto é, existem escalares $\alpha_1, \alpha_2, ..., \alpha_n$ não todos nulos tais que

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \mathbf{0}\,.$$

Sabemos que um dos escalares é diferente de zero, seja ele α_i (para algum i de 1 a n). Então obtém-se (justifique):

$$v_i = -\frac{\alpha_1}{\alpha_i} v_1 - ... - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - ... - \frac{\alpha_n}{\alpha_i} v_{i-1} v_n\,.$$

Reciprocamente, suponhamos que, para algum $j \in \{1, 2, ..., n\}$, existem escalares α_1, ..., α_{j-1}, α_{j+1}, ..., α_n tais que

$$v_j = \alpha_1 v_1 + ... + \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + ... + \alpha_n v_n\,.$$

Então

$$\mathbf{0} = \alpha_1 v_1 + ... + \alpha_{j-1} v_{j-1} + (-1) v_j + \alpha_{j+1} v_{j+1} + ... + \alpha_n v_n\,,$$

ou seja, o vector nulo escreve-se como combinação linear de $v_1, v_2, ..., v_n$ com os escalares não todos nulos (visto que o esclar que está a multiplicar por v_j é -1). Logo os vectores são linearmente dependentes.

Corolário 1. Se algum dos vectores $v_1, v_2, ..., v_n$ de um espaço vectorial é nulo, então $v_1, v_2, ..., v_n$ são linearmente dependentes.

Corolário 2. Se $v_1, v_2, ..., v_n$ são vectores de um espaço vectorial e para alguns $i, j \in \{1, ..., n\}$, com $i \neq j$, $v_i = v_j$, então $v_1, v_2, ..., v_n$ são linearmente dependentes.

A demonstração da proposição seguinte é um exercício simples.

Proposição 2. Sejam $v_1, v_2, ..., v_n$ vectores de um espaço vectorial.

1. Dados um escalar $\alpha \neq 0$ e $i \in \{1, ..., n\}$, os vectores $v_1, v_2, ..., v_n$ são linearmente independentes se e só se $v_1, ..., \alpha v_i, ..., v_n$ o forem.
2. Dados um escalar α e $i, j \in \{1, ..., n\}$, com $i \neq j$, os vectores $v_1, v_2, ..., v_n$ são linearmente independentes se e só se $v_1, ..., v_j, ..., v_i + \alpha v_j, ..., v_n$ o forem.

10. Bases e dimensão

Seja $\mathcal{V}$ um espaço vectorial diferente de $\{0\}$. Um conjunto B de vectores de $\mathcal{V}$ diz-se uma **base** de $\mathcal{V}$ se for simultaneamente um conjunto gerador de $\mathcal{V}$ e linearmente independente.

Uma importante caracterização de base de um espaço vectorial é a seguinte:

Teorema 1. Seja $\mathcal{V}$ um espaço vectorial e B um subconjunto de $\mathcal{V}$. Então B é uma base de $\mathcal{V}$ se e só se qualquer vector de $\mathcal{V}$ se pode exprimir de forma única como combinação linear de vectores de B.

Demonstração. Tendo em conta a definição de base, é claro que basta mostrar que um conjunto gerador B de $\mathcal{V}$ é linearmente independente se e só se nenhum vector de $\mathcal{V}$ se pode escrever como combinação linear de B de duas formas distintas. Apesar de o teorema ser válido independentemente de B ser ou não finito, e a demonstração ser essencialmente a mesma nos dois casos, vamos fazê-la para B finito, com o intuito de simplificar a exposição. Seja então $B = \{v_1, v_2, ..., v_n\}$ um conjunto gerador de $\mathcal{V}$ e suponhamos que é linearmente independente. Se

$$v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n$$

então

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + ... + (\alpha_n - \beta)v_n = \mathbf{0}\,.$$

Como B é linearmente independente, conclui-se que tem de ser

$$\alpha_1 - \beta_1 = 0;\ \alpha_2 - \beta_2 = 0;\ ...;\ \alpha_n \beta_n = 0\,,$$

ou seja,

$$\alpha_1 = \beta_1;\ \alpha_2 = \beta_2;\ ...;\ \alpha_n = \beta_n\,.$$

Suponhamos agora que todo o vector de $\mathcal{V}$ se escreve de forma única como combinação linear de B. Então, em particular, o vector nulo tem uma única maneira de se exprimir como combinação linear de B, ou seja, $\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = \mathbf{0}$ implica $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$, ou seja, B é linearmente independente.

Se $\mathcal{V}$ tem uma base finita, ou $\mathcal{V} = \{0\}$, dizemos que $\mathcal{V}$ é um espaço de **dimensão finita**. Caso contrário, tem **dimensão infinita**.

O teorema a seguir (cuja demonstração se omite) dá conta de uma propriedade muito importante dos espaços vectoriais.

Teorema 2. Se um espaço vectorial $\mathcal{V}$ de dimensão finita tem uma base com n vectores, então qualquer base de $\mathcal{V}$ tem n vectores.

O próximo teorema assegura que é possível encontrar uma base para qualquer espaço vectorial de dimensão finita.

Teorema 3. Seja $\mathcal{V}$ um espaço vectorial de dimensão n. Então qualquer conjunto com $m < n$ vectores linearmente independente pode ser ampliado de modo a obter-se uma base de $\mathcal{V}$.

Atendendo aos dois teoremas anteriores, faz sentido considerar a seguinte definição:

Se $\mathcal{V}$ é um espaço de dimensão finita, a **dimensão de** $\mathcal{V}$ é o número de vectores que cada base de $\mathcal{V}$ tem. Escrevemos $\dim \mathcal{V}{=}n$ para significar que toda a base de $\mathcal{V}$ tem n vectores.

Exemplos.

1. Os conjuntos S_1 e S_3 dos Exemplos da Secção 8 são bases de $\mathbb{R}^3$. (Porquê?) Portanto $\dim \mathbb{R}^3{=}3$.

2. Mais geralmente, para cada $n \in \mathbb{N}$, o espaço vectorial real $\mathbb{R}^n$ tem dimensão n. De facto, é fácil concluir que

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

   é uma base de $\mathbb{R}^n$. É a chamada **base canónica**.

3. Para cada $n \in \mathbb{N}$, o espaço vectorial complexo $\mathbb{C}^n$ tem dimensão n. O conjunto

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

   constitui a **base canónica** de $\mathbb{C}^n$.

4. Para cada $n \in \mathbb{N}$, o espaço vectorial real $\mathbb{C}^n$ tem dimensão $2n$. O conjunto

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

   é uma base do espaço vectorial real $\mathbb{C}^n$.

Usando os Teorema 2 e Teorema 3 anteriores, é fácil chegar às seguintes propriedades importantes dos espaços vectorias de dimensão finita:

Proposição. Seja $\mathcal{V}$ um espaço vectorial de dimensão n. Então:

1. O número máximo de vectores que um conjunto linearmente independente em $\mathcal{V}$ pode ter é n.
2. Todo o conjunto linearmente independente em $\mathcal{V}$ com n vectores é uma base de $\mathcal{V}$.
3. Qualquer conjunto gerador de $\mathcal{V}$ tem pelo menos n vectores.
4. Todo o conjunto gerador de $\mathcal{V}$ com n vectores é uma base de $\mathcal{V}$.

Enuncia-se a seguir uma consequência simples de 1 da proposição anterior.

Propriedade. Se $\mathcal{W}$ é um subespaço de $\mathcal{V}$ então $\dim\mathcal{W} \leq \dim\mathcal{V}$.

11. Espaços associados a uma matriz

Espaço nulo de uma matriz. Dada uma matriz A $m \times n$ de entradas em $\mathbb{K}$ (com $\mathbb{K} = \mathbb{R}, \mathbb{C}$), consideremos o seguinte subconjunto de $\mathbb{K}^n$ (onde identificamos $\mathcal{M}_{n\times 1}(\mathbb{K})$ com $\mathbb{K}^n$):

$$\{x \in \mathbb{K}^n : Ax = 0\}.$$

É fácil verificar que, na verdade, ele é um subespaço vectorial de $\mathbb{K}^n$. A este subespaço chamamos **espaço nulo de** A e designamo-lo por $N(A)$.

Espaços das colunas e das linhas de uma matriz. Seja A uma matriz $m \times n$ de entradas em $\mathbb{K}$ (com $\mathbb{K} = \mathbb{R}, \mathbb{C}$).

O **espaço das colunas de** A, que se designa por $C(A)$, é o subespaço de $\mathbb{K}^m$ gerado pelas colunas de A.

O espaço das colunas de A é dado por

$$C(A) = \{Ax \,:\, x \in \mathbb{K}^n\},$$

ou seja, $C(A)$ é o conjunto de todos os $b \in \mathbb{K}^m$ para os quais o sistema

$$Ax = b$$

é possível (Justifique).

O **espaço das linhas de** A, que se designa por $L(A)$, é o subespaço de $\mathcal{M}_{1\times n}(\mathbb{K}^n)$ gerado pelas linhas de A.

Observação. Quando fazemos a eliminação de Gauss de uma matriz, utilizamos dois tipos de operações sobre as linhas:

- troca de linhas,
- adição a uma linha do produto de um escalar por outra linha.

Portanto, atendendo aos resultados enunciados nas Secções 8 e 9, podemos concluir que:

1. A dependência ou independência linear das linhas de uma matriz não se altera ao longo da eliminação de Gauss da matriz. As k primeiras linhas da matriz inicial são linearmente independentes se e só se as correspondentes linhas da matriz final o forem, respeitando as trocas de linhas se as houver.

2. O espaço das linhas de uma matriz não se altera ao longo da eliminação de Gauss, isto é, todas as matrizes obtidas ao longo da eliminação de Gaus de uma matriz têm o mesmo espaço das linhas.

Quanto ao espaço das colunas:

1. A dependência ou independência linear das colunas de uma matriz não se altera ao longo da eliminação de Gauss da matriz.

   Mas:

2. Em geral, as matrizes obtidas ao longo da eliminação de Gauss de uma matriz têm espaços das colunas diferentes.

Um facto que usaremos muitas vezes é o seguinte:

Se U é uma matriz em escada de linhas então:

1. As suas linhas não nulas são linearmente independentes; portanto, formam uma base para o espaço das linhas.
2. As colunas que contêm os pivôs são linearmente independentes. Estas colunas formam uma base do espaço das colunas de U.

Como consequência temos que: car(U)=dim$L(U)$=dim$C(U)$.

Tendo em conta que a dependência ou independência das linhas e das colunas não se altera ao longo da eliminação de Gauss de uma matriz, concluimos que também para qualquer matriz A se verificam as igualdades:

$$\boxed{\text{car}(A)=\dim L(A)=\dim C(A)}$$

Exercício. Mostrar que se A é uma matriz $m \times n$ e B é uma matriz $n \times p$, então car$(AB) \leq$car$A$.

Resolução. Tem-se que $C(AB) \subseteq C(A)$, porque se $z \in C(AB)$, z é da forma $(AB)x$; mas então, $z = A(Bx)$, pelo que pertence a $C(A)$. Como $C(AB)$ é um subespaço de $C(A)$, tem-se que dim$(C(AB)) \leq$dim$(C(A))$, ou seja, car$(AB) \leq$car(A).

Chama-se **nulidade de uma matriz** A, e designa-se por nul(A), à dimensão do espaço nulo da matriz A. Tem-se que:

$$\begin{aligned} \mathrm{nul}(A) &= \text{número de incógnitas livres do sistema homogéneo } Ax=0 \\ &= (\text{n.}^{\underline{\text{o}}} \text{ de incógns. do sistema } Ax=b) - (\text{n.}^{\underline{\text{o}}} \text{ de incógns. básicas do sistema } Ax=b) \\ &= n - \mathrm{car}(A), \text{ onde } n \text{ é o número de colunas de } A. \end{aligned}$$

Verifica-se, então, a igualdade:

$$\boxed{\mathrm{nul}(A) \;=\; n \text{ - } \mathrm{car}(A)}$$

Grande parte dos espaços vectoriais com que vamos trabalhar na prática são de dimensão finita; são espaços $\mathbb{K}^n$, onde $\mathbb{K}$ é $\mathbb{R}$ ou $\mathbb{C}$. Para o estudo da dependência ou independência linear bem como de outras questões referentes a estes espaços, podemos utilizar o já aprendido sobre matrizes.

De acordo com o que for conveniente, o espaço $\mathbb{K}^n$ pode ser identificado com o espaço das matrizes coluna $n \times 1$ com elementos em $\mathbb{K}$ ou com o espaço das matrizes linha $1 \times n$ com elementos em $\mathbb{K}$. As propriedades dos espaços das linhas e das colunas de uma matriz podem ser úteis no estudo de espaços $\mathbb{K}^n$.

Exemplo. Consideremos a matriz real

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

Façamos a eliminação de Gauss de A:

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

Podemos então afirmar que o espaço das linhas de A tem como base

$$\{[1 \;\; -1 \;\; 0],\; [0 \;\; 4 \;\; 2]\}$$

ou ainda, atendendo a que não trocámos linhas ao longo da eliminação,

$$\{[1 \;\; -1 \;\; 0],\; [2 \;\; 2 \;\; 2]\}.$$

Uma base para o espaço das colunas de A é

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\},$$

isto é, é constituída pelas colunas da matriz A que correspondem às colunas da matriz U que têm os pivôs. Mas, atenção, $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix} \right\}$ não é uma base para o espaço das colunas de A, embora o seja do espaço das colunas de U. Com efeito, é claro que não é possível escrever a primeira coluna de A, $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, como combinação linear dos vectores $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}$.

A eliminação de Gauss feita acima pode dar resposta a várias outras questões. Por exemplo, consideremos as matrizes

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

Suponhamos que queremos saber se b pertence ou não ao espaço das colunas de B. Como já vimos,

$$b \in C(B) \iff Bx = b \text{ é um sistema possível.}$$

Ora A não é mais do que a matriz ampliada do sistema $Bx = b$. Então, olhando para o resultado da eliminação de Gauss de A, vemos que car(B)=car($[B \,|\, b]$) e, portanto, o sistema é possível, pelo que $b \in C(B)$.

Podíamos agora estar interessados em escrever $b = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$ como combinação linear das colunas $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ e $\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$. Notemos que, pelo que foi visto, estas duas matrizes coluna constituem uma base de $C(B)$ e, portanto, a combinação linear procurada é única. Na verdade, ela é determinada pela solução do sistema $Bx = b$ (que, claro está, é possível determinado). Resolvendo o sistema obtém-se, para $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$,

$$Bx = b \iff x_1 = x_2 = 1/2\,.$$

Portanto,

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

III

Aplicações lineares

12. Aplicações lineares

Sejam $\mathcal{V}$ e $\mathcal{W}$ dois espaços vectoriais sobre $\mathbb{R}$ ou $\mathbb{C}$. Uma aplicação $f : \mathcal{V} \to \mathcal{W}$ diz-se uma **aplicação linear** ou uma **transformação linear** se preservar a adição e a multiplicação escalar, isto é,

(i) $f(u + v) = f(u) + f(v)$ para todo $u, v \in \mathcal{V}$;

(ii) $f(\alpha v) = \alpha f(v)$ para todo o escalar α e todo o $v \in \mathcal{V}$.

Exemplos.

1. $g : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida por $g(x_1, x_2, x_3) = x_1 + x_2 + x_3$ é linear. (Verifique.)
2. $f : \mathbb{C}^2 \longrightarrow \mathbb{C}^3$ definida por $f(z_1, z_2) = (iz_1, 0, z_2)$ é linear. (Verifique.)
3. Se A é uma matriz $m \times n$ sobre $\mathbb{R}$ (ou $\mathbb{C}$), a aplicação $a : \mathcal{M}_{n\times 1}(\mathbb{R}) \longrightarrow \mathcal{M}_{m\times 1}(\mathbb{R})$ (ou $a : \mathcal{M}_{n\times 1}(\mathbb{C}) \longrightarrow \mathcal{M}_{m\times 1}(\mathbb{C})$) definida por $a(x) = Ax$ é uma transformação linear. (Verifique.)
4. A aplicação $h : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por $h(x_1, x_2) = (x_1^2, x_2)$ não é linear. Basta notar que, por exemplo,

   $h(2, 0) + h(1, 0) = (2^2, 0) + (1, 0) = (5, 0) \neq (9, 0) = h(3, 0) = h((2, 0) + (1, 0)).$
5. A aplicação $j : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida por $j(x_1, x_2, x_3) = (x_1, x_2 + x_3 + 1)$ não é linear. (Porquê?)

Uma propriedade importante que ilustra bem a força da noção de base de um espaço vectorial é a seguinte:

> **Teorema.** Dados espaços vectoriais $\mathcal{V}$ e $\mathcal{W}$, uma base $\{v_1, v_2, ..., v_n\}$ de $\mathcal{V}$ e n vectores $w_1, w_2, ..., w_n$ de $\mathcal{W}$, existe uma e uma só aplicação linear $f : \mathcal{V} \to \mathcal{W}$ tal que $f(v_j) = w_j$, $j = 1, 2, ..., n$.

Portanto, toda a aplicação linear $f : \mathcal{V} \to \mathcal{W}$ fica definida pelas imagens por f dos vectores de uma base de $\mathcal{V}$. Na verdade, seja f dada pelos seus valores numa base de $\mathcal{V}$, como em cima; dado $v \in \mathcal{V}$, existem escalares únicos $\alpha_1, \alpha_2, ..., \alpha_n$ tais que $v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$. Então, atendendo a que f é linear, vem:

$$\begin{aligned} f(v) &= f(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) \\ &= f(\alpha_1 v_1) + f(\alpha_2 v_2) + ... + f(\alpha_n v_n) \\ &= \alpha_1 f(v_1) + \alpha_2 f(v_2) + ... + \alpha_n f(v_n) \\ &= \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n. \end{aligned}$$

13. Matriz de uma aplicação linear

Uma importante consequência do teorema da secção anterior é a seguinte:

> Toda a transformação linear entre espaços vectoriais finitos pode ser representada por uma matriz.

Vamos ver como se faz essa representação.

Sejam $\mathcal{V}$ e $\mathcal{W}$ dois espaços vectoriais reais ou complexos, $B = (v_1, v_2, ..., v_n)$ e $C = (w_1, w_2, ..., w_m)$ bases ordenadas de, respectivamente, $\mathcal{V}$ e $\mathcal{W}$, e $f : \mathcal{V} \to \mathcal{W}$ uma aplicação linear. A **matriz da aplicação linear f relativamente às bases B e C** é a matriz

$$M(f; B, C) = \begin{bmatrix} \alpha_{11} & \alpha_{12} & ... & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & ... & \alpha_{2n} \\ ... & ... & ... & ... \\ \alpha_{m1} & \alpha_{m2} & ... & \alpha_{mn} \end{bmatrix}$$

onde a primeira coluna é constituída pelas coordenadas de $f(v_1)$ relativamente à base C, a segunda coluna tem as coordenadas de $f(v_2)$ relativamente a C, e assim sucessivamente, mais precisamente:

$$f(v_1) = \alpha_{11}w_1 + \alpha_{21}w_2 + ... + \alpha_{m1}w_m$$

$$f(v_2) = \alpha_{12}w_1 + \alpha_{22}w_2 + ... + \alpha_{m2}w_m$$

$$... \quad ... \quad ... \quad ... \quad ... \quad ... \quad ... \quad ... \quad ...$$

$$f(v_n) = \alpha_{1n}w_1 + \alpha_{2n}w_2 + ... + \alpha_{mn}w_m$$

Sejam $f : \mathcal{V} \to \mathcal{W}$ e B e C nas condições descritas acima. Se β_1, β_2, ..., β_n são as coordenadas de um vector $v \in \mathcal{V}$ relativamente à base B, então as coordenadas γ_1, γ_2, ..., γ_m de $f(v)$ relativamente à base C são dadas pela igualdade:

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & ... & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & ... & \alpha_{2n} \\ ... & ... & ... & ... \\ \alpha_{m1} & \alpha_{m2} & ... & \alpha_{mn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_m \end{bmatrix}$$

Com efeito, seja $v = \beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n$; então:

$$\begin{aligned} f(v) &= f(\beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n) \\ &= \beta_1 f(v_1) + \beta_2 f(v_2) + ... + \beta_n f(v_n) \\ &= \beta_1 (\sum_{i=1}^{m} \alpha_{i1} w_i) + \beta_2 (\sum_{i=1}^{m} \alpha_{i2} w_i) + ... + \beta_n (\sum_{i=1}^{m} \alpha_{in} w_i) \\ &= \sum_{j=1}^{n} \beta_j \sum_{i=1}^{m} \alpha_{ij} w_i \\ &= \sum_{i=1}^{m} (\sum_{j=1}^{n} \alpha_{ij} \beta_j) w_i \end{aligned} \tag{6}$$

A igualdade $f(v) = \sum_{i=1}^{m} (\sum_{j=1}^{n} \alpha_{ij} \beta_j) w_i$ mostra que $\left(\sum_{j=1}^{n} \alpha_{1j}\beta_j,\ \sum_{j=1}^{n} \alpha_{2j}\beta_j,\ ...,\ \sum_{j=1}^{n} \alpha_{mj}\beta_j) \right)$ são as coordenadas de $f(v)$ relativamente à base ordenada $(w_1, w_2, ..., w_m)$. Como, por outro lado,

$$M(f; B, C) \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} \alpha_{1j}\beta_j \\ \sum_{j=1}^{n} \alpha_{1j}\beta_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{1j}\beta_j \end{bmatrix},$$

fica provada a igualdade (6).

14. Núcleo e imagem

Seja $f : \mathcal{V} \to \mathcal{W}$ uma aplicação linear.
O **núcleo de** f é definido por:

$$N(f) = \{v \in \mathcal{V} \,:\, f(v) = 0\}.$$

A **imagem de** f é dada por:

$$\mathrm{Im}(f) = \{f(v) \,:\, v \in \mathcal{V}\}.$$

$N(f)$ é um subespaço vectorial de $\mathcal{V}$ e $\mathrm{Im}(f)$ é um subespaço vectorial de $\mathcal{W}$ (Verifique).

Exemplos.
1. Seja $f : \mathbb{R}^3 \to \mathbb{R}^2$ a aplicação linear definida por $f(x, y, z) = (x - 2y, z)$. Então

$$\begin{aligned}(x, y, z) \in N(f) &\Leftrightarrow f(x, y, z) = (0, 0, 0)\\ &\Leftrightarrow (x - 2y, z) = (0, 0, 0)\\ &\Leftrightarrow x - 2y = 0 \wedge z = 0\\ &\Leftrightarrow x = 2y \wedge z = 0\end{aligned}$$

Portanto, $N(f) = \{(2y, y, 0) \,:\, y \in \mathbb{R}\}$.

2. Seja A uma matriz $m \times n$ com elementos em $\mathbb{R}$ e consideremos a aplicação linear

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & v & \mapsto & Av \end{array}.$$

Então $N(f) = N(A)$ e $\mathrm{Im}(f) = C(A)$.

Proposição. Uma aplicação linear $f : \mathcal{V} \to \mathcal{W}$ é injectiva se e só se $N(f) = \{0\}$.

Demonstração. Se f é injectiva, tem-se

$$x \in N(f) \Leftrightarrow f(x) = \mathbf{0} \Leftrightarrow f(x) = f(\mathbf{0}) \Leftrightarrow x = \mathbf{0}\,,$$

pelo que $N(f) = \{\mathbf{0}\}$. Reciprocamente, se $N(f) = \{\mathbf{0}\}$, então, para quaisquer vectores x e y de $\mathcal{V}$,

$$f(x) = f(y) \Leftrightarrow f(x) - f(y) = \mathbf{0} \Leftrightarrow f(x - y) = \mathbf{0} \Leftrightarrow x - y = \mathbf{0} \Leftrightarrow x = y\,,$$

pelo que f é injectiva.

15. Espaços isomorfos

A demonstração da proposição seguinte é um exercício fácil.

Proposição 1. Se $f : \mathcal{V} \to \mathcal{W}$ é uma aplicação linear bijectiva, então $f^{-1} : \mathcal{W} \to \mathcal{V}$ também é linear.

Uma aplicação linear bijectiva diz-se um **isomorfismo** entre espaços vectoriais. Dois espaços vectoriais dizem-se **isomorfos** se existir entre eles algum isomorfismo.

Proposição 2. Dois espaços vectoriais de dimensão finita sobre o mesmo $\mathbb{K}$ são isomorfos se e só se tiverem a mesma dimensão.

Demonstração. Para mostrar que se $\dim(\mathcal{V}) = \dim(\mathcal{W}) = n$ então $\mathcal{V}$ é isomorfo a $\mathcal{W}$, basta considerar uma base $\{v_1, v_2, ..., v_n\}$ de $\mathcal{V}$ e uma base $\{w_1, w_2, ..., w_n\}$ de $\mathcal{W}$ e verificar que a aplicação linear determinada pelas igualdades $f(v_1) = w_1, f(v_2) = w_2, ..., f(v_n) = w_n$ é um isomorfismo. Reciprocamente, dado um isomorfismo $g : \mathcal{V} \to \mathcal{W}$ e uma base $\{v_1, v_2, ..., v_n\}$ de $\mathcal{V}$, pode provar-se (fica como exercício) que $\{f(v_1), f(v_2), ..., f(v_n)\}$ é uma base de $\mathcal{W}$, pelo que $\dim(\mathcal{W}) = \dim(\mathcal{V}) = n$.

16. Matriz de mudança de base

Sejam $B = (u_1, u_2, ..., u_n)$ e $B' = (v_1, v_2, ..., v_n)$ duas bases ordenadas de um espaço vectorial $\mathcal{V}$. Suponhamos que, para cada $j \in \{1, 2, ..., n\}$,

$$v_j = \sum_{i=1}^{n} a_{ij} u_i \,.$$

Então, a matriz $n \times n$

$$M(B, B') = (a_{ij})$$

diz-se a **matriz de mudança da base B para a base B'**. De notar que

$$M(B, B') = M(\mathrm{id}_{\mathcal{V}}, B', B).$$

Usando então o que vimos na Secção 13, podemos concluir o seguinte:

Seja $X = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$ a matriz coluna das coordenadas de um vector $v \in \mathcal{V}$ relativamente à base B e seja $Y = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$ constituída pelas coordenadas de v relativamente à base B'. Então

$$M(B, B')Y = X\,.$$

Uma matriz de mudança de base é sempre invertível e tem-se

$$(M(B, B'))^{-1} = M(B', B)\,.$$

Portanto,

$$Y = M(B', B)X\,.$$

IV

Determinantes

17. Formas determinantes

Seja $\mathcal{M}_{n\times n}(\mathbb{K})$ o espaço das matrizes de ordem n de entradas em $\mathbb{K}$. Dada uma matriz qualquer $A = (a_{ij})$ em $\mathcal{M}_{n\times n}(\mathbb{K})$, escrevemos

$$A = [A_1\ A_2\ ...\ A_n]$$

onde $A_1, A_2, ..., A_n$ representam as colunas de A.

Uma aplicação $D : \mathcal{M}_{n\times n}(\mathbb{K}) \longrightarrow \mathbb{K}$ diz-se uma **forma determinante** se, para quaisquer $i, j \in \{1, ..., n\}$, satisfaz os seguintes axiomas:

(D1) $D([A_1\ ...\ A_i + B_i\ ...\ A_n]) = D([A_1\ ...\ A_i\ ...\ A_n]) + D[A_1\ ...\ B_i\ ...\ A_n]);$

(D2) $D([A_1\ ...\alpha A_i\ ...\ A_n]) = \alpha D([A_1\ ...\ A_i\ ...\ A_n])$ (para qualquer escalar α);

(D3) $D([A_1\ ...\ A_i\ ...\ A_j\ ...\ A_n]) = -D([A_1\ ...\ A_j\ ...\ A_i\ ...\ A_n]);$

(D4) $D(I_n) = 1.$

As condições *(D1)* e *(D2)* significam que D é uma aplicação linear na coluna i $(i = 1, ..., n)$ quando as outras colunas se mantêm fixas, i.e., D é n-linear.

Observação. Verifica-se facilmente que se em cima substituirmos a condição *(D3)* por

(D3′) $D(A) = 0$ sempre que A tem duas colunas iguais,

obtemos uma definição equivalente de forma determinante. Ou seja, uma aplicação $D : \mathcal{M}_{n\times n}(\mathbb{K}) \longrightarrow \mathbb{K}$ é uma forma determinante se e só se satisfizer *(D1)*, *(D2)*, *(D3′)* e *(D4)*.

Teorema. Para cada $n \in \mathbb{N}$, existe uma e uma só forma determinante $D : \mathcal{M}_{n\times n}(\mathbb{K}) \longrightarrow \mathbb{K}$. Esta forma determinante é definida, para cada $A = (a_{ij})$, por

$$D(A) = \sum_{\sigma \in S_n} (-1)^{s(\sigma)} a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

onde S_n é o conjunto de todas as aplicações bijectivas do conjunto $\{1, 2, ..., n\}$ em si próprio e $s(\sigma) =$ n.º de pares (i, j) tais que $1 \leq i < j \leq n$ e $\sigma(i) > \sigma(j)$.

Este teorema é essencial para o que se vai seguir neste capítulo. Omite-se a demonstração, mas sugere-se como exercício a prova para o caso $n = 2$ (ou $n = 3$) mediante o seguinte procedimento:

Para provar a unicidade da forma determinante, caso exista, começa-se por ter em conta que

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}e_1 + a_{21}e_2 & a_{12}e_1 + a_{22}e_2 \end{bmatrix}$$

onde e_1 e e_2 são as colunas da matriz identidade, e desenvolve-se

$$D\left(\begin{bmatrix} a_{11}e_1 + a_{21}e_2 & a_{12}e_1 + a_{22}e_2 \end{bmatrix}\right)$$

partindo do princípio de que os axiomas $(D1)$, $(D2)$, $(D3')$ e $(D4)$ são válidos. Obter-se-á deste modo que, caso exista, a forma determinante tem de ser definida pela igualdade

$$D\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{22} - a_{12}a_{21} \,. \tag{7}$$

Repare-se que $a_{11}a_{22} - a_{12}a_{21} = (-1)^{s(\sigma)} a_{1\sigma(1)} a_{2\sigma(2)} + (-1)^{s(\sigma')} a_{1\sigma'(1)} a_{2\sigma'(2)}$, onde σ e σ' são as duas aplicações bijectivas possíveis no conjunto $\{1, 2\}$, ou seja, $\sigma = \mathrm{id}_{\{1,2\}}$ e σ' troca 1 com 2.

Para concluir sobre a existência de uma forma determinante para $n = 2$, basta verificar que a igualdade (7) satisfaz efectivamente os 4 axiomas.

Isto dará uma ideia da demonstração no caso geral.

18. Determinante de uma matriz

Para cada matriz quadrada A, chamamos a $D(A)$ o **determinante de** A e designamo-lo por $\det A$ ou $|A|$.

Determinantes de ordem 2. Aplicando o teorema anterior ao caso $n = 2$, temos que

$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

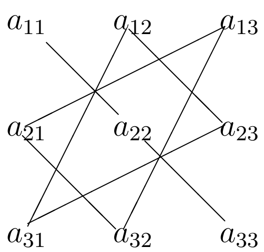
Determinantes de ordem 3. Aplicando o teorema anterior ao caso $n = 3$, obtemos uma soma de 3! parcelas. Uma regra que facilita a determinação dessas 6 parcelas é a chamada Regra de Sarrus que se descreve a seguir:

Regra de Sarrus(só para determinantes de ordem 3)**.** Consideremos uma matriz de ordem 3:

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

Para calcular o determinante desta matriz:

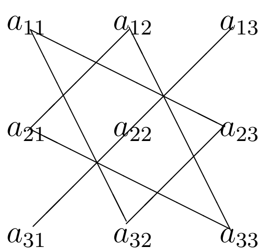
1º) Marcamos a diagonal principal e os dois triângulos que têm um dos lados paralelos a essa diagonal, como ilustrado na figura



O produto dos elementos da diagonal e os dois produtos cujos factores são os três vértices de cada um dos triângulos formam as três parcelas afectadas pelo sinal $+$:

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \tag{8}$$

2º) Procedendo de forma análoga relativamente à diagonal secundária



obtemos as parcelas afectadas pelo sinal $-$:

$$a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32}\,. \tag{9}$$

Somando (8) com o simétrico de (9), obtemos o determinante da matriz A, ou seja,

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

Fórmula de Laplace. Uma manipulação adequada do exposto no Teorema da Secção 17 permite chegar à Fórmula de Laplace, que aqui se descreve para dois casos:

1. **Fórmula de Laplace segundo a linha i:**

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

onde A_{ij} é a matriz $(n-1) \times (n-1)$ que se obtém de A suprimindo a linha i e a coluna j. Chama-se a A_{ij} o **menor-ij** da matriz A e chama-se a $(-1)^{i+j} \det A_{ij}$ o **cofactor-ij**, ou **complemento algébrico** ij, de A.

2. **Fórmula de Laplace segundo a coluna j:**

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

Exemplo. Seja

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

Desenvolvendo segundo a 2.ª coluna, obtemos:

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

$$= 0 + 0 - 2 \times (8 - (-8)) = -32.$$

Uma forma mais rápida de determinar o valor dos $(-1)^{i+j}$ que figuram na Fórmula de Laplace é ter em conta que o sinal de $(-1)^{i+j}$ é o sinal que se obtém na posição ij da matriz preenchendo cada posição com os sinais $+$ e $-$, alternando-os quer nas linhas quer nas colunas, e começando com o sinal $+$ na posição 1,1. Assim às posições na matriz 3×3 correspondem os seguintes sinais:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

A seguir listamos algumas propriedades importantes dos determinantes; de notar que algumas delas não são mais do que a repetição dos axiomas da definição de forma determinante apresentada no início deste capítulo.

Propriedades. Sejam A e B matrizes quadradas de ordem n.

1. $\det(A^t) = \det A$.
2. Se A é uma matriz com uma coluna ou uma linha nula, então $\det(A) = 0$.
3. Seja $A = [A_1 \dots A_j + B_j \dots A_n]$ uma representação de A indicando as suas colunas. Então $\det(A) = \det([A_1 \dots A_j \dots A_n]) + ([A_1 \dots B_j \dots A_n])$. A mesma propriedade se aplica às linhas.
4. Seja $A = [A_1 \dots \alpha A_j \dots A_n]$ uma representação de A indicando as suas colunas. Então $\det(A) = \alpha \det([A_1 \dots A_j \dots A_n])$. A mesma propriedade se aplica às linhas.
5. Se B é uma matriz obtida de A por meio da troca de duas colunas (respectivamente, duas linhas) entre si, então $\det(B) = -\det(A)$.
6. $\det(AB) = \det A \det B$.
7. O determinante de uma matriz triangular (superior ou inferior) é igual ao produto dos elementos da diagonal principal.
8. O $\det A$ não se altera quando se adiciona a uma coluna (respectivamente, uma linha) de A uma combinação linear das outras colunas (respectivamente, linhas).
9. As colunas (respectivamente, linhas) de uma matriz quadrada A são linearmente dependentes se e só se $\det A = 0$. Portanto:

$$A \text{ é invertível} \Leftrightarrow A \text{ é não singular} \Leftrightarrow \det A \neq 0.$$

Cálculo de um determinante usando Eliminação de Gauss. As propriedades 5., 7. e 8. acabadas de enunciar permitem calcular o determinante de uma matriz usando a eliminação de Gauss. Por exemplo:

$$\begin{vmatrix} 0 & 1 & 2 \\ 2 & 2 & 2 \\ 4 & 1 & 2 \end{vmatrix} = - \begin{vmatrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 4 & 1 & 2 \end{vmatrix} = - \begin{vmatrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & -3 & -2 \end{vmatrix} = - \begin{vmatrix} 2 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{vmatrix} = -2 \times 1 \times 4 = -8.$$

Observação. Foram descritas várias técnicas de cálculo de determinantes. Na prática (se nada for dito em contrário) usa-se muitas vezes uma mistura dessas técnicas, de acordo com o que for conveniente.

19. Aplicação dos determinantes à resolução de sistemas e ao cálculo da inversa de uma matriz

Cálculo da inversa. Dada uma matriz $A = (a_{ij})$ $n \times n$, chama-se **matriz adjunta de** A, e denota-se por $\operatorname{adj} A$, à matriz $n \times n$

$$\operatorname{adj} A = \left((-1)^{i+j} \det A_{ij}\right) ,$$

onde A_{ij} se define como atrás, na descrição da Fórmula de Laplace feita na Secção 18. Ou seja, as entradas de $\operatorname{adj} A$ são os complementos algébricos de A. Se A for invertível, $\det A \neq 0$ e a inversa de A é dada por

$$A^{-1} = \frac{(\operatorname{adj} A)^t}{\det A} .$$

Regra de Cramer. Seja A uma matriz não singular. Então a solução única do sistema de equações lineares $Ax = b$, com $x = [x_1\ x_2\ ...\ x_n]^t$, é dada por

$$x_j = \frac{\det C_j}{\det A} ,$$

onde C_j é a matriz que se obtém de A substituindo a coluna j de A pela matriz coluna b.

V

Valores próprios e vectores próprios

20. Valores próprios e vectores próprios

Seja $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ (com $\mathbb{K}$ igual a $\mathbb{R}$ ou $\mathbb{C}$). Um escalar λ diz-se um **valor próprio de** A se existir um vector não nulo $v \in \mathbb{K}^n$ tal que $Av = \lambda v$. Neste caso dizemos que v é um **vector próprio** de A.

Exemplo. $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$; logo 2 é um valor próprio da matriz $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ e $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é um seu vector próprio.

Observação. Um vector próprio é, por definição, não nulo.

Seja λ um valor próprio de A. O subespaço de $\mathbb{K}^n$,

$$\mathcal{E}(\lambda) = \{v \in \mathbb{K}^n \mid Av = \lambda v\}$$

diz-se **espaço próprio de** A associado a λ. (Mostre que $\mathcal{E}(\lambda)$ é efectivamente um subespaço vectorial de $\mathbb{K}^n$.)

Exemplo. Consideremos a matriz do exemplo anterior, $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. Como vimos, 2 é um valor próprio de A. Vamos determinar o espaço próprio associado a 2:

$$Av = 2v \iff (A-2I)v = 0 \iff \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{cases} -x_1 + x_2 = 0 \\ 0 = 0 \end{cases}$$
$$\iff \begin{cases} x_2 = x_1 \\ x_1 \in \mathbb{K} \end{cases};$$

Então

$$\mathcal{E}(2) = \{(x_1, x_1) \mid x_1 \in \mathbb{K}\} = \mathcal{L}(\{(1,1)\})\,.$$

Cálculo dos valores próprios de uma matriz. Dada uma matriz A $n\times n$, pretendemos encontrar os valores próprios de A, isto é, os escalares λ para os quais existe algum $v \neq 0$ tal que $Av = \lambda v$. Ora $Av = \lambda v \Leftrightarrow (A - \lambda I)v = 0$. E

$$(A - \lambda I)v = 0$$

é afinal um sistema quadrado homogéneo cuja matriz dos coeficientes é $A - \lambda I$. Existirá uma solução $v \neq 0$ do sistema se e só se $\mathrm{car}(A - \lambda I) < n$, se e só se a matriz $A - \lambda I$ é singular, se e só se

$$|A - \lambda I| = 0\,. \tag{10}$$

Portanto, determinar os valores próprios de A é determinar os valores de λ que satisfazem a equação (10).

À expressão $|A - \lambda I|$ chamamos **polinómio característico de** A e à equação (10) chamamos **equação característica de** A.

Pode provar-se a seguinte propriedade sobre vectores próprios:

Teorema 1. Seja A uma matriz quadrada. Se u_1, u_2, ..., u_k são vectores próprios de A associados a valores próprios todos distintos, então são linearmente independentes.

Corolário. Uma matriz $n \times n$ não tem mais do que n valores próprios.

A recíproca do teorema anterior é falsa como podemos concluir pelo exemplo seguinte:

Consideremos a matriz real $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Tem-se que $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, pelo que 1 é valor próprio de A.

Calculemos $\mathcal{E}(1)$:

$$Ax = x \Longleftrightarrow (A - I)x = 0 \Longleftrightarrow \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x = 0 \Longleftrightarrow \left\{ \begin{array}{l} x_1 = x_2 \\ x_2, x_3 \text{ livres} \end{array} \right. .$$

Então

$$\mathcal{E}_1 = \{(x_2, x_2, x_3) \,|\, x_2, x_3 \in \mathbb{R}\}\,.$$

Portanto, os vectores (1,1,0) e (0,0,1) pertencem a $\mathcal{E}(1)$, estando então ambos associados ao mesmo valor próprio 1, e são linearmente independentes.

Duas matrizes A e B dizem-se **semelhantes** se existir uma matriz invertível S tal que $A = SBS^{-1}$.

Uma matriz A diz-se **diagonizável** se for semelhante a uma matriz diagonal, ou seja, se existir uma matriz diagonal D e uma matriz invertível S tais que $A = SDS^{-1}$. Nesse caso S diz-se uma **matriz diagonalizante** de A.

Teorema 2. Uma matriz A $n \times n$ é diagonalizável se e só se A tem n vectores próprios linearmente independentes.

Demonstração. Com efeito, $A = SDS^{-1}$ se e só se $AS = SD$ com S invertível. Sejam

$$S = [v_1 \ \ v_2 \ \ ... \ \ v_n] \quad \text{e} \quad D = \operatorname{diag}(\lambda_1,\, \lambda_2,\, ...\, \lambda_n)\,.$$

Temos então

$$\begin{aligned} AS = SD \ &\Leftrightarrow A[v_1 \ \ v_2 \ \ ... \ \ v_n] = [v_1 \ \ v_2 \ \ ... \ \ v_n] \operatorname{diag}(\lambda_1,\, \lambda_2,\, ...\, \lambda_n) \\ &\Leftrightarrow Av_i = \lambda_i v_i, \ \ i = 1, 2, ..., n\,. \end{aligned}$$

Por conseguinte:

Dizer que v_1, v_2, ..., v_n são n vectores próprios de A linearmente independentes, associados aos valores próprios, respectivamente, λ_1, λ_2, ... λ_n, significa exactamente que $A = SDS^{-1}$ com $S = [v_1 \quad v_2 \quad ... \quad v_n]$ e $D = \operatorname{diag}(\lambda_1,\, \lambda_2,\, ...\, \lambda_n)$.

Diagonalização de uma matriz. Pelo que ficou dito, podemos concluir que, dada uma matriz $n \times n$, três situações podem acontecer:

1. Existem n valores próprios distintos e, portanto, existem n vectores próprios linearmente independentes, pelo que a matriz é diagonalizável.

2. Existem m valores próprios distintos com $m < n$, mas existem n vectores próprios linearmente independentes e, portanto, a matriz é diagonalizável.

3. Existem m valores próprios distintos, $m < n$, e não existe um conjunto com n vectores próprios que seja linearmente independente e, portanto, a matriz não é diagonalizável.

Seja A uma matriz $n \times n$ cujo polinómio característico $P(\lambda)$ tem n raízes, eventualmente com algumas repetidas. Isto é, tal que

$$P(\lambda) = k(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_k)^{m_k} \quad \text{com} \quad m_1 + m_2 + ... + m_k = n\,. \qquad (11)$$

É o que acontece sempre, se calcularmos as raízes em $\mathbb{C}$, visto que todo o polinómio em $\mathbb{C}$ pode ser decomposto na forma (11). Chama-se **multiplicidade algébrica** de um valor próprio λ_i à sua multiplicidade como raiz do polinómio característico. Assim na equação (11), λ_1 tem multiplicidade algébrica m_1, λ_2 tem multiplicidade algébrica m_2, etc. Chama-se **multiplicidade geométrica** do valor próprio λ_i à dimensão do espaço próprio associado a λ_i.

Designemos por $m.a.(\lambda_i)$ a multiplicidade algébrica de λ_i e por $m.g.(\lambda_i)$ a multiplicidade geométrica de λ_i. Se o polinómio característico de A admite uma decomposição da forma (11), teremos uma das seguintes três situações, correspondentes às três situações descritas atrás:

1. A matriz tem n valores próprios Neste caso é imediato que $m.g.(\lambda_i) = m.a.(\lambda_i)$=1 para todo o valor próprio λ_i de A. A matriz é diagonalizável.

2. Para cada valor próprio λ_i de A, $m.g.(\lambda_i) = m.a.(\lambda_i) = m_i$. A matriz é diagonalizável.

3. Algum dos valores próprios tem multiplicidade geométrica inferior à multiplicidade algébrica, i.e., para algum λ_i, $m.g.(\lambda_i) < m.a.(\lambda_i)$. Neste caso, a matriz não é diagonalizável.

Terminamos este capítulo com um interesante teorema sobre a equação característica de uma matriz:

Teorema de Caley-Hamilton. Toda a matriz A $n \times n$ satifaz a sua equação característica. Isto é, se $P(\lambda) = a_o + a_1\lambda + a_2\lambda^2 + ... + a_n\lambda^n$ é o polinómio característico de A, então

$$a_o + a_1A + a_2A^2 + ... + a_nA^n = 0\,.$$

VI
Espaços com produto interno

21. Produto interno

Seja $\mathcal{V}$ um espaço vectorial real. Um **produto interno** em $\mathcal{V}$ é uma operação que a cada par de vectores u, v de $\mathcal{V}$ associa um escalar $\langle u, v\rangle$, satisfazendo os axiomas seguintes:

(I1) Simetria: $\langle u, v\rangle = \langle v, u\rangle$;

(I2) Linearidade: $\langle \alpha_1 u_1 + \alpha_2 u_2, v\rangle = \alpha_1\langle u_1, v\rangle + \alpha_2\langle u_2, v\rangle$;

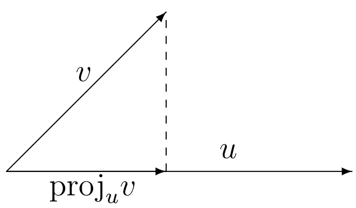
(I3) Positividade: $\langle u, u\rangle \geq 0$; $\langle u, u\rangle = 0$ se e só se $u = 0$.

Os espaços vectoriais reais com produto interno chamam-se **espaços euclidianos**.

Num espaço com produto interno:

- A **norma** de um vector u de $\mathcal{V}$ denota-se por $\|u\|$ e é o número real não negativo $\|u\| = \sqrt{\langle u, u\rangle}$.

- A **projecção ortogonal de** v **sobre** u é o vector $\quad \operatorname{proj}_u v = \dfrac{\langle v, u\rangle}{\|u\|^2}\, u.$



- O **ângulo** entre dois vectores u e v não nulos é $(\widehat{u,v}) = \arccos\dfrac{\langle u, v\rangle}{\|u\|\,\|v\|}$.[2]

- u é **ortogonal** a v, e escreve-se $u \perp v$ se e só se $\langle u, v\rangle = 0$.

[2]Esta definição tem sentido, porque num espaço com produto interno, para quaisquer vectores u e v não nulos verifica-se sempre que $\dfrac{|\langle u, v\rangle|}{\|u\|\,\|v\|} \leq 1$, como decorre da Desigualdade de Schwarz apresentada mais à frente.

Exemplos.

1. À semelhança do que se passa em $\mathbb{R}^2$ e $\mathbb{R}^3$, em qualquer espaço $\mathbb{R}^n$ $(n \geq 1)$ a operação que a cada dois vectores $u = (a_1, a_2, ..., a_n)$ e $v = (b_1, b_2, ..., b_n)$ faz corresponder o real
$$\langle u, v\rangle = a_1b_1 + a_2b_2 + ... + a_nb_n.$$
satisfaz *(I1)*, *(I2)* e *(I3)*, sendo pois os $\mathbb{R}^n$ espaços com produto interno. É este o produto interno que se considerará sempre nestes espaços, a menos que seja dito algo em contrário.

   A norma de um vector u é então dada por
$$\|u\| = \sqrt{a_1^2 + a_2^2 + ... + a_n^2}.$$

2. Consideremos o espaço vectorial real das funções reais contínuas no intervalo $[a, b]$ com a adição e a multiplicação escalar usuais. Pondo, para cada f e g em $C[a, b]$,
$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$
obtemos um produto interno em $C[a, b]$. É sempre este o produto interno que consideraremos nos espaços deste tipo, salvo alguma indicação em contrário.

3. Todo o subespaço vectorial $\mathcal{W}$ de um espaço euclidiano $\mathcal{V}$ é ele próprio um espaço euclidiano, visto que a restrição do produto interno aos vectores de $\mathcal{W}$ é um produto interno em $\mathcal{W}$.

Enuncia-se a seguir, sem demonstração, um teorema de consequências algébricas, geométricas e analíticas muito importantes.

Desigualdade de Schwarz. Se $\mathcal{V}$ é um espaço vectorial com produto interno então a *desigualdade de Schwarz*
$$|\langle u, v\rangle| \leq \|u\| \, \|v\| \tag{12}$$
verifica-se para quaisquer vectores u e v de $\mathcal{V}$. A igualdade acontece se e só se os vectores u e v são linearmente dependentes.

Observação. De seguida vamos referir algumas das consequências da desigualdade de Schwarz.

1. Atendendo à definição de produto interno usual num espaço $\mathbb{R}^n$, a desigualdade de Schwarz afirma que quaisquer $2n$ números reais x_1, x_2, ...,x_n, y_1, y_2, ...,y_n verificam a desigualdade

$$(\sum_{i=1}^{n} x_i y_i)^2 \leq (\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i^2).$$

2. No espaço $C[a,b]$ das funções contínuas no intervalo $[a,b]$ a desigualdade de Schwarz assegura que

$$\left| \int_a^b f(x)g(x)dx \right|^2 \leq \int_a^b (f(x))^2 dx \int_a^b (g(x))^2 dx.$$

3. Num espaço com produto interno define-se *distância* $d(u,v)$ entre dois vectores u e v por

$$d(u,v) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}.$$

A distância d verifica as três propriedades seguintes, sendo a terceira uma consequência da Desigualdade de Schwarz:

(i) $d(u,v) = d(v,u)$;
(ii) $d(u,v) \geq 0; d(u,v) = 0$ se e só se $u = v$;
(iii) $d(u,v) \leq d(u,w) + d(w,v)$.

22. Ortogonalidade

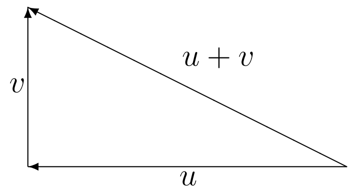
Entre as bases do espaço vectorial $\mathbb{R}^3$, a chamada base canónica revela-se, em muitas situações, "melhor" do que as outras. Note-se que se designarmos por e_1, e_2 e e_3 os vectores desta base, eles verificam as seguintes propriedades:

(1) $\langle e_i, e_j \rangle = 0$ para $i \neq j$
(2) $\|e_i\| = 1$, $i = 1, 2, 3$.

A propriedade (1) traduz o facto de os três vectores serem perpendiculares dois a dois e (2) explicita que cada um deles tem norma igual a 1. Exprimimos estas duas propriedades dizendo que $\{e_1, e_2, e_3\}$ constitui uma base ortonormada. Claro que não é a única em $\mathbb{R}^3$, é possível encontrar muitas outras. Como é sabido, no estudo dos espaços vectoriais é fundamental a noção de base. Quando interpretamos geometricamente uma base de $\mathbb{R}^2$ ou de $\mathbb{R}^3$, pensamos num sistema de eixos de coordenadas. E geralmente trata-se de um sistema de eixos perpendiculares entre si, mais do que isso, um sistema ortonormado. Isto é, quando escolhemos uma base, temos tendência a escolher uma base ortonormada. Realmente, esta possibilita-nos fazer facilmente uma leitura algébrica de construções geométricas. Vamos ver que o conceito de ortogonalidade entre vectores está presente e tem um papel fundamental em qualquer espaço com produto interno. Tal como em $\mathbb{R}^3$, de um modo geral, num espaço com produto interno, as bases ortonormadas têm propriedades especiais que se podem revelar muito úteis nos cálculos algébricos e na sua interpretação geométrica.

Recorde-se que dois vectores u e v se dizem ortogonais se $\langle u, v\rangle = 0$.
Observe-se que a relação de ortogonalidade entre dois vectores é simétrica – isto é, $\langle u, v\rangle = 0$ se e só se $\langle v, u\rangle = 0$.

O célebre Teorema de Pitágoras afirma que "A soma dos quadrados dos catetos de um triângulo rectângulo é igual ao quadrado da hipotenusa". Supondo que os vectores u e v formam os catetos dum triângulo rectângulo, como ilustrado na figura



temos então que

$$\|u\|^2 + \|v\|^2 = \|u + v\|^2.$$

Fica como exercício provar que, esta propriedade se verifica em qualquer espaço euclidiano, de acordo com o enunciado a seguir.

Teorema de Pitágoras. Num espaço euclidiano, dois vectores u e v são ortogonais se e só se satisfazem a igualdade

$$\|u\|^2 + \|v\|^2 = \|u + v\|^2.$$

Num espaço com produto interno:

1. Um conjunto de vectores diz-se **ortogonal** se os vectores forem ortogonais dois a dois.

2. Um conjunto **ortonormado** é um conjunto ortogonal cujos vectores são todos unitários, i.e., de norma igual a 1.

Sejam u e v dois quaisquer vectores não nulos e perpendiculares em $\mathbb{R}^2$. É imediata a verificação geométrica de que eles são linearmente independentes – não é possível exprimir nenhum deles como combinação linear do outro. Facilmente nos apercebemos de que o mesmo é válido em $\mathbb{R}^3$ para dois ou três vectores não nulos perpendiculares dois a dois. O teorema seguinte assenta que esta é uma propriedade geral de todos os espaços com produto interno de dimensão finita.

Teorema 1. Num espaço com produto interno todo o conjunto ortogonal de vectores não nulos é linearmente independente.

Demonstração. Seja $v_1, v_2, ..., v_n$ vectores não nulos de um conjunto ortogonal; então $\langle v_i, v_j\rangle = 0$ se $i \neq j$ e $\langle v_i, v_i\rangle \neq 0$ para $i, j = 1, 2, ...n$. Sejam $\alpha_1, \alpha_2, ..., \alpha_n$ escalares tais que

$$\sum_{j=1}^{n} \alpha_j v_j = 0.$$

Pretendemos mostrar que os α_i's são todos nulos. Ora, para cada i, temos

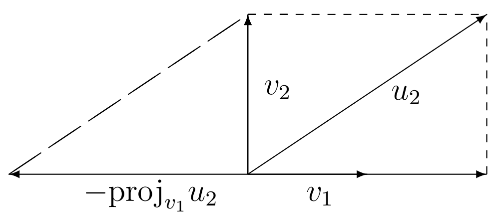
$$0 = \langle \sum_{j=1}^{n} \alpha_j v_j, v_i\rangle = \sum_{j=1}^{n} \alpha_j \langle v_j, v_i\rangle = \alpha_i \langle v_i, v_i\rangle;$$

como $\langle v_i, v_i\rangle \neq 0$, conclui-se que $\alpha_i = 0$.

Como consequência do teorema anterior, num espaço com produto interno de dimensão n todo o conjunto ortogonal de n vectores não nulos é uma base do espaço; em particular, todo o conjunto ortonormado de n vectores é uma base.

Põe-se agora a questão: Dado um espaço com produto interno, é sempre possível determinar uma base ortonormada para esse espaço? Notemos antes de mais que a questão essencial é saber se é possível encontrar uma base ortogonal porque a partir dela é fácil obter uma ortonormada – basta dividir cada um dos vectores da base ortogonal pela sua norma.

A figura a seguir é uma abordagem geométrica para o caso em que, a partir de uma base $\{u_1, u_2\}$, se pretende encontrar uma base ortogonal da forma $\{v_1, v_2\}$ com $v_1 = u_1$ e $v_2 = u_2 - \alpha v_1$.



Na verdade, como ilustrado na figura, $\alpha v_1 = \text{proj}_{v_1} u_2$.

Um cálculo simples levar-nos-ía à mesma conclusão:

$\langle u_2 - \alpha v_1, v_1\rangle = 0 \Longleftrightarrow \langle u_2, v_1\rangle - \alpha\langle v_1, v_1\rangle = 0 \Longleftrightarrow \alpha = \frac{\langle u_2, v_1\rangle}{\|v_1\|^2}$.

Mais geralmente, alguns cálculos levam sem dificuldade à verificação de que a técnica descrita a seguir - o chamado método de ortogonalização de Gram-Schmidt - nos fornece sempre a uma base ortogonal:

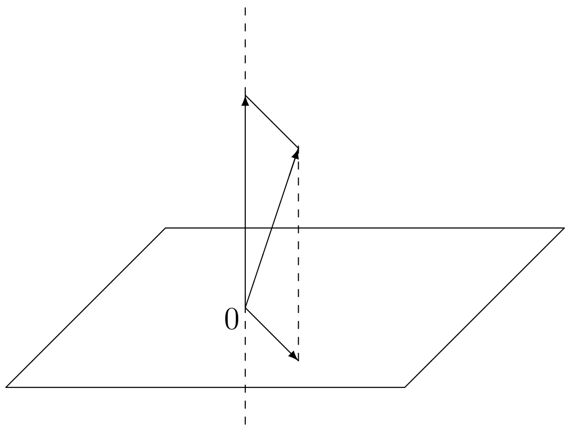
Método de ortogonalização de Gram-Schmidt. Seja $\mathcal{V}$ um espaço com produto interno e $\{u_1, u_2, ..., u_n\}$ uma base de $\mathcal{V}$. Defina-se

$$\begin{aligned} v_1 &= u_1 \\ v_j &= u_j - \sum_{i=1}^{j-1} \frac{\langle u_j, v_i \rangle}{\|v_i\|^2} v_i, \ j = 2, ...n \end{aligned}$$

Então, para cada j, os vectores $v_1, v_2, ..., v_j$ são ortogonais dois a dois.

23. Projecções ortogonais sobre subespaços

Consideremos $\mathbb{R}^3$ como o espaço de todos os vectores do espaço ordinário com origem num ponto O fixado previamente. O plano $\mathcal{W}$ contendo esse ponto pode ser identificado com o subespaço de todos os vectores contidos no plano. O conjunto de todos os vectores ortogonais a $\mathcal{W}$ é constituído por todos aqueles que ficam sobre a recta perpendicular ao plano no ponto O. A esta recta chamamos complemento ortogonal de $\mathcal{W}$. O espaço $\mathbb{R}^3$ aparece assim "partido" em dois subespaços ortogonais entre si, à custa dos quais cada vector de $\mathbb{R}^3$ é expresso de forma única por meio da soma de dois vectores ortogonais, conforme ilustrado na figura junta.



Este fenómeno generaliza-se a outros espaços com produto interno.

Seja $\mathcal{V}$ um espaço com produto interno e $\mathcal{W}$ um subespaço de $\mathcal{V}$. O conjunto de todos os vectores de $\mathcal{V}$ ortogonais a cada um dos vectores de $\mathcal{W}$ diz-se o **complemento ortogonal** de $\mathcal{W}$ e designa-se por $\mathcal{W}^\perp$. A prova das propriedades seguintes fica como exercício.

1. O complemento ortogonal de um subespaço de um espaço com produto interno é também um subespaço.
2. Se $\mathcal{W}$ é um subespaço de um espaço $\mathcal{V}$ com produto interno e $\{w_1, w_2, ..., w_k\}$ é uma base de $\mathcal{W}$, então um vector v de $\mathcal{V}$ pertence a $\mathcal{W}^\perp$ se e só se v é ortogonal a todos os w_i, $i = 1, 2, ..., k$.
3. Para qualquer subespaço $\mathcal{W}$ de um espaço com produto interno, $\mathcal{W} \subseteq (\mathcal{W}^\perp)^\perp$.

Seja $\mathcal{V}$ um espaço vectorial e $\mathcal{U}$ e $\mathcal{W}$ subespaços vectoriais de $\mathcal{V}$. Diz-se que $\mathcal{V}$ é a **soma directa** de $\mathcal{U}$ com $\mathcal{W}$, e escreve-se $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, se cada elemento de $\mathcal{V}$ se escreve de forma única como soma de um vector de $\mathcal{U}$ com um vector de $\mathcal{W}$.

Teorema 1. Se $\mathcal{W}$ é um subespaço de dimensão finita de um espaço $\mathcal{V}$ com produto interno, então $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$.

Seja $\mathcal{V}$ um espaço com produto interno e $\mathcal{W}$ um subespaço de $\mathcal{V}$ de dimensão finita. Seja $v \in \mathcal{V}$ e $v = v_\mathcal{W} + v_{\mathcal{W}^\perp}$ a decomposição de v na soma de um vector $v_\mathcal{W}$ de $\mathcal{W}$ com um vector $v_{\mathcal{W}^\perp}$ de $\mathcal{W}^\perp$. Então $v_\mathcal{W}$ diz-se a **projecção ortogonal de** v **em** $\mathcal{W}$ e escreve-se $\text{proj}_\mathcal{W}(v) = v_\mathcal{W}$. Analogamente se diz que $v_{\mathcal{W}^\perp}$ é a projecção ortogonal de v sobre $\mathcal{W}^\perp$.

Teorema 2. Seja $\{w_1, w_2, ...w_k\}$ uma base ortonormada de $\mathcal{W}$. Então, para cada $v \in \mathcal{V}$,

$$\text{proj}_\mathcal{W}(v) = v_\mathcal{W} = \sum_{i=1}^{k} \langle v, w_i \rangle w_i$$

O teorema seguinte afirma que, se $\mathcal{W}$ é um subespaço de dimensão finita de um espaço $\mathcal{V}$ com produto interno e v é um elemento de $\mathcal{V}$, então o elemento de $\mathcal{W}$ mais próximo de v é a sua projecção ortogonal sobre $\mathcal{W}$ (ver definição de distância no final da Observação da Secção 21).

Teorema 3. Seja $\mathcal{W}$ um subespaço de dimensão finita de um espaço $\mathcal{V}$ com produto interno e seja v um elemento de $\mathcal{V}$. Se $v_{\mathcal{W}}$ é a projecção ortogonal de v sobre $\mathcal{W}$, tem- se

$$\|v - v_{\mathcal{W}}\| \leq \|v - w\|$$

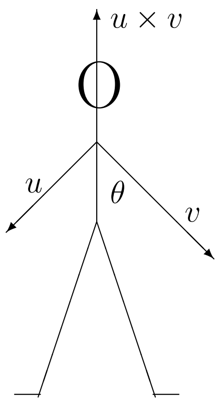
para todo $w \in \mathcal{W}$. A igualdade verifica-se se e só se $w = v_{\mathcal{W}}$.

24. Produto externo

Sejam u e v dois vectores de $\mathbb{R}^3$. Chamamos **produto externo** de u por v, e denota-se por $u \times v$, ao vector de $\mathbb{R}^3$, definido do seguinte modo:

· A norma de $u \times v$ é dada por: $\|u \times v\| = \|u\|\|v\| \sin(\hat{u,v})$;

· Se $\|u \times v\| \neq 0$, $u \times v$ tem direção perpendicular aos vectores u e v; quanto ao sentido de $u \times v$, pode ser descrito intuitivamente do seguinte modo: Supondo que u corresponde ao braço direito e v corresponde ao braço esquerdo, com o ângulo $\theta = (\hat{u,v})$ como representado na figura, $u \times v$ dirige-se no sentido da cabeça.



Uma mnemónica para o cálculo do produto externo de dois vectores é a seguinte: Sejam $u = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ e $v = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$; então

$$u \times v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

Mais precisamente, obtém-se $u \times v$ fazendo o desenvolvimento de Laplace do determinante acima, segundo a primeira linha. Deste modo:

$$u \times v = (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}\,.$$

O **produto misto** dos vectores u, v e w é igual ao produto interno de $u \times v$ por w; pode denotar-se por $u \times v|w$. Se $u = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $v = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ e $w = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$, o produto misto de u por v por w é dado pelo determinante:

$$u \times v|w = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

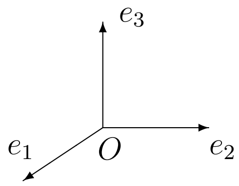
Propriedades. Para quaisquer vectores u, v e w de $\mathbb{R}^3$ verificam-se as seguintes propriedades:

1. $u \times v = -v \times u$.
2. $(u + v) \times w = u \times w + v \times w$ e $(\alpha u) \times v = \alpha(u \times v)$.
3. $u \times v$ é ortogonal a u e v
4. $u \times v = 0$ se e só se u e v são linearmente dependentes.
5. Se u e v são linearmente independentes, então os três vectores u, v, e $u \times v$ são linearmente independentes.
6. A área do paralelogramo que tem como dois dos lados os vectores linearmente independentes u e v é igual a $\|u \times v\|$.
7. O volume de um paralelepípipedo que tem como três das suas arestas os vectores linearmente independentes u, v e w é igual a $|u \times v|w|$.

VII
Geometria analítica

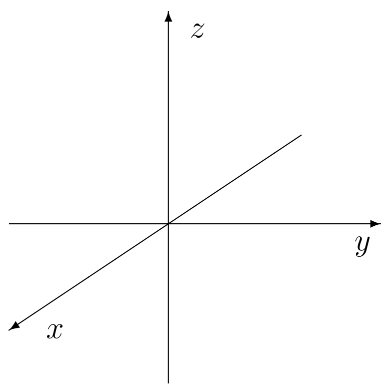
25. Pontos e referencial canónico

Consideremos um ponto O fixado no espaço ordinário e três vectores e_1, e_2 e e_3 que constituem uma base ortonormada no espaço dos vectores com origem em O, como representado na figura



Identificamos e_1 com (1,0,0);
e_2 com (0,1,0);
e_3 com (0,0,1).

O referencial ortonormado $(O; e_1, e_2, e_3)$ diz-se um referencial canónico. Associado a este referencial canónico temos, como é bem conhecido, o seguinte referencial



constituído pelo ponto O, a que chamamos origem do referencial, e três rectas perpendiculares entre si e orientadas como indicado na figura que são chamadas eixo dos xx, eixo dos yy e eixo dos zz. Referimo-nos ao plano que contém os eixos dos xx e dos yy como sendo o plano XOY. De forma análoga, falamos do plano XOZ e do plano YOZ.

Recordemos ainda que um ponto P do espaço fica determinado pelas suas coordenadas (a, b, c) onde

· a corresponde à projecção ortogonal sobre o eixo dos xx; (é, em valor absoluto, igual à distância de P ao plano YOZ.)

· b corresponde à projecção ortogonal sobre o eixo dos yy; (é, em valor absoluto, igual à distância de P ao plano ZOX.)

· c corresponde à projecção ortogonal sobre o eixo dos zz; (é, em valor absoluto, igual à distância de P ao plano XOY.)

Escrevemos $P(a, b, c)$ ou apenas (a, b, c) para denotar o ponto de coordenadas (a, b, c). Note-se que as coordenadas de um ponto P são precisamente as coordenadas do vector $\overrightarrow{OP}$ relativamente à base (e_1, e_2, e_3).

A distância entre dois pontos $P_1(x_1, y_1, z_1)$ e $P_2(x_2, y_2, z_2)$ é dada por

$$d(P_1, P_2) = \| \overrightarrow{P_1P_2} \| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}\,.$$

O ponto médio do segmento de $[P_1P_2]$ é o ponto de coordenadas

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2}\right).$$

26. Planos

Um **plano** (no espaço ordinário) é todo o conjunto de pontos X definido por

$$X = P + \alpha\mathbf{u} + \beta\mathbf{v}, \quad \alpha, \beta \in \mathbb{R} \tag{13}$$

onde P é um dado ponto do espaço e $\mathbf{u}$ e $\mathbf{v}$ são vectores de $\mathbb{R}^3$ linearmente independentes. Portanto, o plano representado por (13) contém o ponto P e é paralelo aos vectores $\mathbf{u}$ e $\mathbf{v}$.

Pondo

$$\begin{aligned} X &= (x, y, z) \\ P &= (p_1, p_2, p_3) \\ \mathbf{u} &= (u_1, u_2, u_3) \\ \mathbf{v} &= (v_1, v_2, v_3) \end{aligned}$$

podemos escrever (13) na forma seguinte

$$(x, y, z) = (p_1, p_2, p_3) + \alpha(u_1, u_2, u_3) + \beta(v_1, v_2, v_3)\,, \quad \alpha, \beta \in \mathbb{R}\,. \tag{14}$$

A uma representação do plano da forma (13) ou (14) chamamos **equação vectorial do plano**.

De (14), obtemos as equações

$$\begin{cases} x = p_1 + \alpha u_1 + \beta v_1 \\ y = p_2 + \alpha u_2 + \beta v_2 \\ z = p_3 + \alpha u_3 + \beta v_3 \end{cases} \qquad \alpha, \beta \in \mathbb{R} \tag{15}$$

conhecidas pelas **equações paramétricas do plano**.

Uma **equação cartesiana** de um conjunto $\mathcal{C}$ de pontos do espaço é uma equação da forma $f(x, y, z) = 0$ tal que $\mathcal{C} = \{(x, y, z) \mid f(x, y, z) = 0\}$.

Pretendemos agora encontrar uma equação cartesiana do plano. Consideremos a equação (14) do plano definido pelo ponto $P(p_1, p_2, p_3)$ e pelas direcções dos vectores $\mathbf{u} = (u_1, u_2, u_3)$ e $\mathbf{v} = (v_1, v_2, v_3)$. Podemos escrevê-la na forma

$$(x - p_1, y - p_2, z - p_3) = \alpha(u_1, u_2, u_3) + \beta(v_1, v_2, v_3), \quad \alpha, \beta \in \mathbb{R}$$

ou seja, (x, y, z) é um ponto do plano se e só se o vector $(x - p_1, y - p_2, z - p_3)$ se pode escrever como combinação linear dos vectores (u_1, u_2, u_3) e (v_1, v_2, v_3). Como estes dois vectores são linearmente independentes, concluimos que (x, y, z) é um ponto do plano se e só se os vectores $(x - p_1, y - p_2, yz - p_3)$, (u_1, u_2, u_3) e (v_1, v_2, v_3) são linearmente dependentes, ou ainda, se e só se

$$\begin{vmatrix} x - p_1 & y - p_2 & z - p_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0 \tag{16}$$

Assim, a equação (16) é uma equação do plano que contém P e é paralelo a $\mathbf{u}$ e $\mathbf{v}$. Desenvolvendo o determinante, obtemos a equação:

$$(u_2 v_3 - u_3 v_2)x + (u_3 v_1 - u_1 v_3)y + (u_1 v_2 - u_2 v_1)z - (p_1 k_1 + p_2 k_2 + p_3 k_3) = 0 \tag{17}$$

com k_1, k_2 e k_3 constantes. A equação (17) é, portanto, da forma

$$Ax + By + Cz + D = 0\,. \tag{18}$$

Toda a equação do tipo (18), com A, B e C não conjuntamente nulos representa um plano e diz-se uma **equação geral do plano**.

Até aqui, tratámos de um plano definido por um ponto e dois vectores linearmente independentes. Um plano pode também ser definido por três pontos não colineares. É evidente que as equações do plano já estudadas podem ser facilmente obtidas tendo em conta que se A, B e C são três pontos não colineares, então os vectores $\vec{AB}$ e $\vec{AC}$ são vectores linearmente independentes e o plano que contém os pontos A, B e C fica definido pelo ponto A e pelos vectores $\vec{AB}$ e $\vec{AC}$.

Seja Π o plano que contém o ponto $P(p_1, p_2, p_3)$ e é perpendicular ao vector $\mathbf{w} = (w_1, w_2, w_3)$. Então um ponto de coordenadas (x, y, z) pertence a Π se e só se os vectores $\mathbf{w}$ e $\vec{PX}$ são ortogonais, ou seja, o seu produto interno é igual a zero, i.e.,

$$w_1(x - p_1) + w_2(y - p_2) + w_3(z - p_3) = 0$$

ou, equivalentemente,

$$w_1x + w_2y + w_3z - (w_1p_1 + w_2p_2 + w_3p_3) = 0.$$

Obtivemos, portanto, uma equação geral do plano. Notemos que os coeficientes de x, y e z são precisamente as coordenadas do vector w que é perpendicular (ou normal) ao plano. De um modo geral, se

$$Ax + By + Cz + D = 0$$

é a equação geral de um plano, então (A, B, C) são as coordenadas de um vector normal a esse plano.

Posições relativas de dois planos. Dados dois planos, um dos três casos seguintes acontece:

(1) são coincidentes;

(2) são paralelos não coincidentes;

(3) intersectam-se segundo uma recta.

Sejam Π e Γ dois planos de equações vectoriais

$$\begin{aligned} X &= P + \alpha\mathbf{u} + \beta\mathbf{v} \qquad \alpha, \beta \in \mathbb{R} \\ \text{e} \quad X &= P' + \alpha\mathbf{u}' + \beta\mathbf{v}' \qquad \alpha, \beta \in \mathbb{R}\,, \end{aligned}$$

respectivamente. Então Π e Γ são paralelos se e só se $\{u, v\}$ e $\{u', v'\}$ geram o mesmo subespaço de $\mathbb{R}^3$, abreviadamente, $\mathcal{L}(\{u, v\}) = \mathcal{L}(\{u', v'\})$. Assim, os dois planos são:

(1) coincidentes, se $\mathcal{L}(\{u, v\}) = \mathcal{L}(\{u', v'\})$ e têm algum ponto em comum;

(2) paralelos não coincidentes se $\mathcal{L}(\{u, v\}) = \mathcal{L}(\{u', v'\})$ e não têm nenhum ponto comum;

(3) intersectam-se segundo uma recta, se $\mathcal{L}(\{u, v\}) \neq \mathcal{L}\{u', v'\})$.

Sejam agora Π e Γ representados pelas suas equações gerais,

$$\begin{aligned} & Ax + By + Cz + D = 0 \\ \text{e} \quad & A'x + B'y + C'z + D' = 0\,, \end{aligned}$$

respectivamente.

Então Π e Γ são paralelos se e só se os vectores (A, B, C) e (A', B', C') são linearmente dependentes, i.e., se e só se

$$\operatorname{car}\left(\begin{bmatrix} A & B & C \\ A' & B' & C' \end{bmatrix}\right) = 1\,; \tag{19}$$

neste caso, são:

(1) coincidentes, se, além de (19), também

$$\operatorname{car}\left(\begin{bmatrix} A & B & C & D \\ A' & B' & C' & D' \end{bmatrix}\right) = 1\,;$$

(2) paralelos não coincidentes, se, além de (19),

$$\operatorname{car}\left(\begin{bmatrix} A & B & C & D \\ A' & B' & C' & D' \end{bmatrix}\right) = 2\,,$$

o que signfica que o sistema

$$\begin{cases} Ax + By + Cz + D = 0 \\ A'x + B'y + C'z + D' = 0 \end{cases} \tag{20}$$

é impossível, ou seja, a intersecção dos dois planos é vazia.

(3) Os planos Π e Γ intersectam-se segundo uma recta se o conjunto de soluções do sistema (20) é uma recta, ou seja, o sistema é possível indeterminado com grau de indeterminação 1, ou ainda,

$$\operatorname{car}\left(\begin{bmatrix} A & B & C \\ A' & B' & C' \end{bmatrix}\right) = 2\,.$$

27. Rectas

Como acabámos de ver, quando, para as equações de dois planos

$$\begin{array}{l} Ax + By + Cz + D = 0 \\ A'x + B'y + C'z + D' = 0 \end{array},$$

temos que

$$\operatorname{car}\left(\begin{bmatrix} A & B & C \\ A' & B' & C' \end{bmatrix}\right) = 2\,,$$

então aquelas duas equações definem uma recta, precisamente a recta de intersecção dos dois planos. Por outro lado, dada uma recta, é sempre possível representá-la por um sistema de duas equações daquela forma.

Vamos ver outras formas de representar uma recta. Sejam P um ponto do espaço e $\mathbf{u}$ um vector não nulo. Então o conjunto de pontos X que constituem a recta que contém P e é paralela a $\mathbf{u}$ é definido por

$$X = P + \alpha\mathbf{u} \quad (\alpha \in \mathbb{R})\,. \tag{21}$$

Considerando

$$\begin{array}{l} X = (x, y, z) \\ P = (p_1, p_2, p_3) \\ \mathbf{u} = (u_1, u_2, u_3) \end{array}$$

podemos escrever a equação (21) na forma

$$(x, y, z) = (p_1, p_2, p_3) + \alpha(u_1, u_2, u_3)\,, \qquad \alpha \in \mathbb{R}\,. \tag{22}$$

Uma equação da forma (21) ou (22) diz-se uma **equação vectorial da recta**.

De (22), obtém-se ainda:

$$\begin{cases} x = p_1 + \alpha u_1 \\ y = p_2 + \alpha u_2 \\ z = p_3 + \alpha u_3 \end{cases} \qquad \alpha \in \mathbb{R}\,. \tag{23}$$

Dizemos então que (23) são as **equações paramétricas da recta**.

Note-se que se uma recta for dada por dois dos seus pontos, digamos P e Q, podemos cair no caso acabado de estudar, considerando o ponto P e a direcção do vector $\overrightarrow{PQ}$.

Posições relativas de uma recta e um plano. Sejam r uma recta de equação vectorial

$$X = P + \alpha\mathbf{u} \quad (\alpha \in \mathbb{R})\,,$$

e Π um plano de equação vectorial

$$X = Q + \alpha\mathbf{v} + \beta\mathbf{w} \quad (\alpha, \beta \in \mathbb{R})\,.$$

Então:

((1)) r é estritamente paralela ao plano Π se $\mathbf{u} \in \mathcal{L}(\{\mathbf{v}, \mathbf{w}\})$ mas $P \notin \Pi$;

((2)) r está contida no plano Π se $\mathbf{u} \in \mathcal{L}(\{\mathbf{v}, \mathbf{w}\})$ e $P \in \Pi$;

((3)) a recta r intersecta o plano Π num (único) ponto se $\mathbf{u} \notin \mathcal{L}(\{\mathbf{v}, \mathbf{w}\})$ (e, portanto, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ é um conjunto linearmente independente).

Consideremos agora uma recta r definida cartesianamente por

$$\begin{array}{l} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{array},$$

e um plano Π definido por

$$Ax + By + Cz + D = 0\,.$$

Analisando a relação entre a posição relativa da recta e do plano e a classificação do sistema de equações

$$\begin{array}{l} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \\ Ax + By + Cz + D = 0 \end{array}$$

concluimos que as situações ((1)), ((2)) e ((3)) descritas atrás correspondem a, respectivamente,

((1)) $\operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_3 \\ A & B & C \end{bmatrix}\right) = 2 \quad \text{e car}\left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A & B & C & D \end{bmatrix}\right) = 3$ (justifique);

((2)) $\operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_3 \\ A & B & C \end{bmatrix}\right) = \operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A & B & C & D \end{bmatrix}\right) = 2$ (justifique);

((3)) $\operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A & B & C & D \end{bmatrix}\right) = 3$ (justifique).

Posições relativas de duas rectas. Sejam r e s duas rectas definidas, respectivamente, pelas equações vectoriais:

$$\begin{array}{rll} & X = P + \alpha\mathbf{u} & \alpha \in \mathbb{R}\,, \\ \text{e} & X = Q + \alpha\mathbf{v} & \alpha \in \mathbb{R}\,. \end{array}$$

Então:

[1] r e s são coincidentes se $\mathcal{L}(\{\mathbf{u}\}) = \mathcal{L}(\{\mathbf{v}\})$ (ou seja, se $\mathbf{u}$ paralelo a $\mathbf{v}$) e $P \in s$;

[2] r e s são estritamente paralelas se $\mathcal{L}(\{\mathbf{u}\}) = \mathcal{L}(\{\mathbf{v}\})$ e $P \notin s$;

[3] r e s são concorrentes (ou seja, intersectam-se num (único) ponto), se $\mathcal{L}(\{\mathbf{u}\}) \neq \mathcal{L}(\{\mathbf{v}\})$ e têm algum ponto em comum.

[4] r e s são enviezadas, i.e., não paralelas e não concorrentes, se $\mathcal{L}(\{\mathbf{u}\}) \neq \mathcal{L}(\{\mathbf{v}\})$ e não se intersectam.

Sejam agora r e s dadas pelas representações cartesianas seguintes:

$$r: \quad \begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases}$$

$$s: \quad \begin{cases} A_3x + B_3y + C_3z + D_3 = 0 \\ A_4x + B_4y + C_4z + D_4 = 0 \end{cases}$$

O estudo do sistema constituído pelas quatro equações permite concluir que as situações [1], [2] e [3] consideradas atrás correspondem a, respectivamente:

$$[1]\ \operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \\ A_4 & B_4 & C_4 \end{bmatrix}\right) = \left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{bmatrix}\right) = 2;$$

$$[2]\ \operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \\ A_4 & B_4 & C_4 \end{bmatrix}\right) = 2 \ \text{ e } \ \operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{bmatrix}\right) = 3;$$

$$[3]\ \operatorname{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \\ A_4 & B_4 & C_4 \end{bmatrix}\right) = \left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{bmatrix}\right) = 3;$$

$$[4]\ \mathrm{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \\ A_4 & B_4 & C_4 \end{bmatrix}\right) = 3 \ \text{ e } \ \mathrm{car}\left(\begin{bmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \\ A_3 & B_3 & C_3 & D_3 \\ A_4 & B_4 & C_4 & D_4 \end{bmatrix}\right) = 4.$$

28. Distâncias e ângulos

Sejam F e G dois conjuntos de pontos no espaço ordinário. A **distância** entre F e G, usualmente denotada por

$$d(F,G)$$

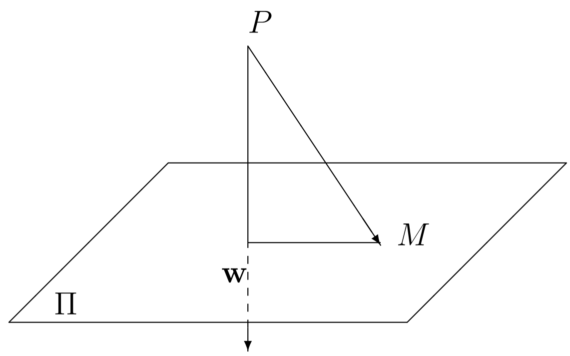
é o ínfimo das distâncias entre um ponto de F e um ponto de G, ou seja,

$$d(F,G) = \inf\{d(A,B) : A \in F,\ B \in G\}\,.$$

Distância de um ponto a um plano. Seja Π um plano de equação geral

$$Ax + By + Cz + D = 0$$

e seja P um ponto qualquer. Consideremos a figura



onde $\mathbf{w}$ é o vector perpendicular a Π de coordenadas (A,B,C). Então:

$$d(P,\Pi) = \left\| \mathrm{proj}_{\mathbf{w}}\ \overrightarrow{PM} \right\| \tag{24}$$

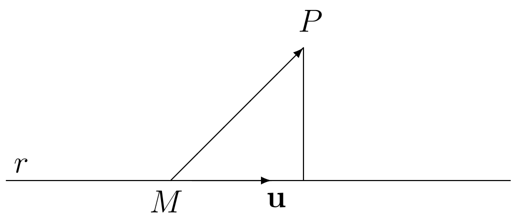
onde M é um ponto arbitrário do plano Π. Sejam (x_o, y_o, z_o) as coordenadas de P; então de (24) deduz-se que

$$d(P,\Pi) = \frac{|Ax_o + By_o + Cz_o + D|}{\sqrt{A^2+B^2+C^2}}\,.$$

Distância de um ponto a uma recta. Sejam r uma recta de equação vectorial

$$X = M + \alpha\mathbf{u}\,, \quad \alpha \in \mathbb{R}\,,$$

e P um ponto. Consideremos a figura



Tem-se

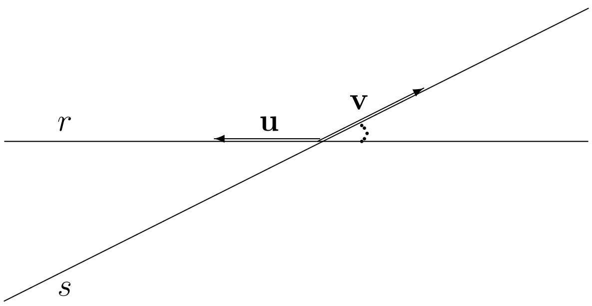
$$\frac{d(P,r)}{d(M,P)} = |\sin(\mathbf{u}\hat{,}\ \overrightarrow{MP})|\,.$$

Desta igualdade, usando a definição de produto externo, obtemos

$$d(P,r) = \frac{\|\mathbf{u}\times \overrightarrow{MP}\,\|}{\|\mathbf{u}\|}\,.$$

Ângulo entre duas rectas. Sejam r e s duas rectas com vectores directores $\mathbf{u}$ e $\mathbf{v}$, respectivamente. O ângulo entre r e s é dado por

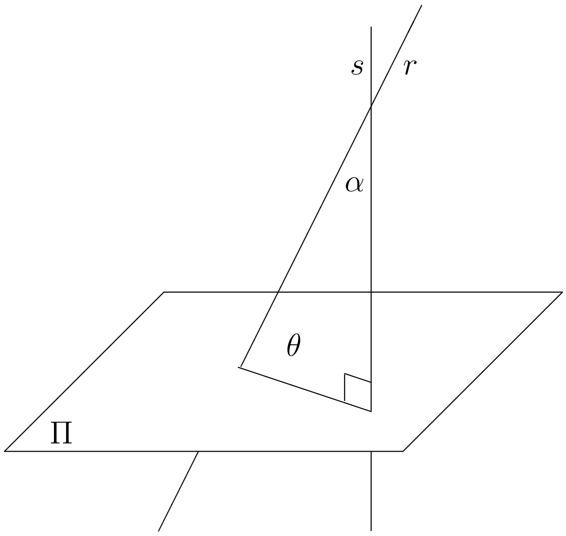
$$(r\hat{,}s) = \min\{(\mathbf{u}\hat{,}\mathbf{v}), \pi - (\mathbf{u}\hat{,}\mathbf{v})\}\,.$$



Consequentemente,

$$\cos(r\hat{,}s) = |\cos(\mathbf{u}\hat{,}\mathbf{v})| = \frac{|\langle\mathbf{u},\mathbf{v}\rangle|}{\|\mathbf{u}\|\|\mathbf{v}\|}\,.$$

Ângulo entre uma recta e um plano. O ângulo de uma recta r com um plano Π, é o complementar do ângulo formado por r com uma recta perpendicular a Π. Sejam $\mathbf{u}$ um vector director de r e $\mathbf{w}$ um vector normal a Π. Seja ainda s uma recta com a direcção de $\mathbf{w}$. Na figura a seguir, θ representa o ângulo entre r e Π.



Por conseguinte,

$$\begin{aligned}\theta &= \arcsin(\cos(\tfrac{\pi}{2}-\theta))\ ^{1}\\ &= \arcsin(\cos\alpha)\\ &= \arcsin(|\cos(\hat{\mathbf{u},\mathbf{w}})|\\ &= \arcsin\left(\tfrac{|\langle \mathbf{u},\mathbf{w}\rangle|}{\|\mathbf{u}\|\|\mathbf{w}\|}\right).\end{aligned}$$

Ou seja,

$$\sin(\hat{r,\Pi}) = \frac{|\langle \mathbf{u},\mathbf{w}\rangle|}{\|\mathbf{u}\|\|\mathbf{w}\|}.$$

Ângulo entre dois planos. O ângulo formado por dois planos Π_1 e Π_2 é igual ao ângulo formado por duas rectas r_1 e r_2 tais que $r_1 \perp \Pi_1$ e $r_2 \perp \Pi_2$. Logo, se w_1 é um vector normal a Π_1 e w_2 é um vector normal a Π_2, então:

$$(\hat{\Pi_1,\Pi_2}) = \min\{(\hat{w_1,w_2}),\ \pi - (\hat{w_1,w_2})\}$$

ou ainda,

$$\cos(\hat{\Pi_1,\Pi_2}) = \frac{|\langle \mathbf{w_1},\mathbf{w_2}\rangle|}{\|\mathbf{w_1}\|\|\mathbf{w_2}\|}.$$

[1] Usando a fórmula trigonométrica $\cos(\pi/2-\theta) = \sin\theta$.

Bibliografia

F. R. Dias Agudo, *Introdução à Álgebra Linear e Geometria Analítica*, Escolar Editora

E. Giraldes, V. H. Fernandes e M. H. Santos, *Álgebra Linear e Geometria Analítica*, McGraw-Hill

P. R. Halmos, *Finite-dimensional Vector Spaces*, Springer-Verlag

L. T. Magalhães, *Álgebra Linear como Introdução à Matemática Aplicada*, Texto Editora

L. Sousa, *Ortogonalidade*, Escola Superior de Tecnologia do Instituto Politécnico de Viseu, 1997, Lição no âmbito de provas públicas para professor coordenador