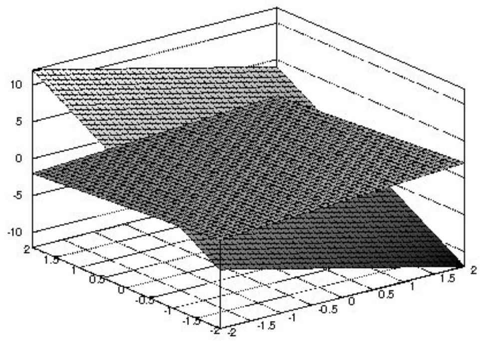
Introdução à Álgebra Linear



Pedro Patrício
Departamento de Matemática e Aplicações
Universidade do Minho
pedro@math.uminho.pt

Conteúdo

Capítulo 1

Introdução

Estes apontamentos pretendem complementar a matéria abordada na componente teórica das aulas de Álgebra Linear. Poderá aqui encontrar provas de resultados referidos nas aulas, bem como exemplos de aplicações e exercícios.

Capítulo 2

Cálculo Matricial

Ao longo deste documento, $\mathbb{K}$ designará o conjunto $\mathbb{C}$ dos números complexos ou $\mathbb{R}$ o dos números reais. Os elementos de $\mathbb{K}$ são denominados por escalares.

2.1 Notação matricial

Uma matriz do tipo $m \times n$, com m e n naturais, sobre $\mathbb{K}$ é uma tabela com mn elementos de $\mathbb{K}$, elementos esses dispostos em m linhas e n colunas:

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

Os elementos a_{ij} dizem-se os elementos, componentes ou entradas da matriz. A matriz diz-se do tipo $m \times n$ se tiver m linhas e n colunas. Usualmente as matrizes são denotadas por letras maiúsculas (A, B, C, $\ldots$).

O conjunto de todas as matrizes (do tipo) $m \times n$ sobre $\mathbb{K}$ representa-se por $\mathcal{M}_{m\times n}(\mathbb{K})$ ou por $\mathbb{K}^{m\times n}$, e o conjunto de todas as matrizes (finitas) sobre $\mathbb{K}$ por $\mathcal{M}(\mathbb{K})$.
$\mathbb{K}^m$ denota $\mathbb{K}^{m\times 1}$.

São exemplos de matrizes

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & -1 \end{bmatrix}, C = \begin{bmatrix} -2 & 1 & 0 & 6 \end{bmatrix}, D = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

Quando conveniente, escrevemos a matriz A da definição anterior como

$$[a_{ij}]_{m\times n},$$

, ou simplesmente como $[a_{ij}]_{m\times n}$, e referimos a_{ij} como o elemento (i,j) de A, isto é, o elemento que está na linha i e na coluna j de A. Iremos também usar a notação $(A)_{ij}$ para indicar o elemento na linha i e coluna j de A.

Duas matrizes $[a_{ij}], [b_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$ são iguais se $a_{ij} = b_{ij}$, para $i = 1, \ldots, m, j = 1, \ldots, n$. Ou seja, duas matrizes são iguais se têm o mesmo número de linhas e o mesmo número de colunas, e que os elementos na mesma linha e coluna são iguais.

Uma matriz do tipo m por n diz-se *quadrada* de ordem n se $m = n$, ou seja, se o número de linhas iguala o de colunas; diz-se *rectangular* caso contrário. Por exemplo, são quadradas as matrizes

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

e rectangulares as matrizes

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

Os *elementos diagonais* de $[a_{ij}]_{i,j=1,\ldots n}$ são $a_{11}, a_{22}, \ldots, a_{nn}$.

Por exemplo, os elementos diagonais de $\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$ são 1 e -2, e os da matriz $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & -3 \end{bmatrix}$ são 1 e 5.

Nos exemplos atrás apresentados, apenas a matriz A é quadrada, sendo as restantes rectangulares. Os elementos diagonais de A são $1, 3$.

Apresentamos, de seguida, alguns tipos especiais de matrizes.

1. Uma matriz do tipo $1 \times n$ diz-se matriz-linha e uma do tipo $n \times 1$ diz-se matriz-coluna.

2. Uma matriz diz-se *diagonal* se for da forma

   $$\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} = diag(d_1, d_2, \ldots, d_n),$$

   ou seja, o elemento (i, j) é nulo, se $i \neq j$. Portanto, uma matriz quadrada é diagonal se os únicos elementos possivelmente não nulos são os diagonais.

   Por exemplo, as matrizes $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ e $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ são matrizes diagonais.

3. A *matriz identidade* de ordem n, I_n, é a matriz diagonal de ordem n, com os elementos diagonais iguais a 1; ou seja,

   $$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Quando é clara a ordem da matriz identidade, escreveremos simplesmente I.

4. Uma matriz $A = [a_{ij}]$ diz-se *triangular superior* se $a_{ij} = 0$ quando $i > j$, e *triangular inferior* se $a_{ij} = 0$ quando $i < j$. Ou seja, são respectivamente triangulares superiores e inferiores as matrizes

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{mn} \end{bmatrix}, \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Por exemplo, as matrizes $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ e $\begin{bmatrix} 1 & -1 & -3 \\ 0 & 1 & -\frac{1}{3} \end{bmatrix}$ são triangulares superiores.

2.2 Operações matriciais

Vejamos agora algumas operações definidas entre matrizes, e algumas propriedades que estas satisfazem.

2.2.1 Soma e produto escalar

Sejam $A = [a_{ij}], B = [b_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $\alpha \in \mathbb{K}$.

1. A soma das matrizes A e B, denotada por $A+B$, é a matriz $m \times n$ cujo elemento (i,j) é $a_{ij} + b_{ij}$. Ou seja, $(A+B)_{ij} = (A)_{ij} + (B)_{ij}$.

2. O produto da matriz A pelo escalar α, notado por αA, é a matriz $m \times n$ cujo elemento (i,j) é $\alpha\, a_{ij}$. Ou seja, $(\alpha A)_{ij} = \alpha (A)_{ij}$.

Repare que a soma de duas matrizes, da mesma ordem, é feita elemento a elemento, e o produto escalar de uma matriz por $\alpha \in \mathbb{K}$ é de novo uma matriz da mesma ordem da dada, onde cada entrada surge multiplicada por α. Ou seja,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \vdots & & & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{bmatrix}$$

e

$$\alpha \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1m} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2m} \\ \vdots & & & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \dots & \alpha a_{nm} \end{bmatrix}.$$

Por exemplo,

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

e

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

Em particular, se $A = [a_{ij}]$ então $-A = [-a_{ij}]$.

De ora em diante, 0 representa uma qualquer matriz cujos elementos são nulos. A matriz nula do tipo $m \times n$ denota-se por $0_{m\times n}$.

Estas operações satisfazem as propriedades que de seguida se descrevem, onde $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $\alpha, \beta \in \mathbb{K}$:

1. A soma de matrizes é associativa: $(A + B) + C = A + (B + C)$.

2. A soma de matrizes é comutativa: $A + B = B + A$

3. A matriz nula é o elemento neutro da adição: $A + 0 = 0 + A = A$.

4. Existe o simétrico de cada matriz $A + (-A) = (-A) + A = 0$.

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

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

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

8. $1\,A = A$.

2.2.2 Produto

Resta-nos definir o produto matricial.

Seja $A = [a_{ij}]$ uma matriz $m \times p$ e $B = [b_{ij}]$ uma matriz $p \times n$. O produto de A por B, denotado por AB, é a matriz $m \times n$ cujo elemento (i, j) é $a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}$. Assim,

$$AB = \left[\sum_{k=1}^{p} a_{ik}b_{kj}\right]_{m\times n} \text{ e portanto } (AB)_{ij} = \sum_{k=1}^{p} (A)_{ik}(B)_{kj}.$$

Atente-se aos tipos de A e B na definição anterior.

Como exemplo, façamos o produto da matriz $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$ pela matriz $B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$.

Ora

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

Antes de fazermos referência a algumas propriedades, vejamos uma outra forma exprimir o produto de duas matrizes. Para tal, suponha que $X = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, sendo a primeira do tipo $1 \times n$ e a segunda do tipo $n \times 1$. Pelo que acabámos de referir, o produto de X por Y está bem definido, sendo a matriz produto do tipo 1×1, e portanto, um elemento de $\mathbb{K}$. Esse elemento é $x_1y_1 + x_2y_2 + \dots x_ny_n$. Voltemos agora ao produto de $A_{m\times p}$ por $B_{p\times n}$, e fixemos a linha i de A e a coluna j de B. Ou seja, a matriz linha $\begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{ip} \end{bmatrix}$ e a matriz coluna $\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{bmatrix}$. O produto da primeira pela segunda é o elemento de $\mathbb{K}$ dado por $a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$. Ora, este elemento não é mais nem menos que a entrada (i, j) da matriz produto AB. Ou seja, a entrada (i, j) de AB é o produto da linha i de A pela coluna j de B.

Vejamos algumas propriedades deste produto de matrizes, onde os tipos das matrizes $A, B, C, I, 0$ são tais que as operações indicadas estão definidas, e $\alpha \in \mathbb{K}$:

1. O produto de matrizes é associativo $(AB)C = A(BC)$;
2. O produto de matrizes é distributivo em relação à soma $A(B + C) = AB + AC, (A + B)C = AC + BC$;
3. A matriz identidade é o elemento neutro para o produto: $AI = A, IA = A$;
4. A matriz nula é o elemento absorvente para o produto: $0\,A = 0, A\,0 = 0$;
5. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Façamos a verificação da primeira igualdade de (2). A verificação de que as matrizes são do mesmo tipo fica ao cargo do leitor. Iremos apenas verificar que a entrada (i, j) de $A(B+C)$ iguala a entrada (i, j) de $AB + AC$. Ora, supondo que A tem p colunas, e portanto que B e C têm p linhas,

$$
\begin{aligned}
(A(B+C))_{ij} &= \sum_{k=1}^{p}(A)_{ik}((B)_{kj} + (C)_{kj}) \\
&= \sum_{k=1}^{p}((A)_{ik}(B)_{kj} + (A)_{ik}(C)_{kj}) \\
&= \sum_{k=1}^{p}(A)_{ik}(B)_{kj} + \sum_{k=1}^{p}(A)_{ik}(C)_{kj} \\
&= (AB)_{ij} + (AC)_{ij} = (AB + AC)_{ij}.
\end{aligned}
$$

Verifiquemos também a propriedade (3). Note-se que $(I)_{ii} = 1$ e $(I)_{ij} = 0$ se $i \neq j$. Ora $(AI)_{ij} = \sum_{k=1}^{p} (A)_{ik}(I)_{kj} = (A)_{ij}$.

É importante notar que o produto matricial <u>não é</u>, em geral, comutativo. Por exemplo, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. A lei do anulamento do produto também não é válida, em geral, no produto matricial. Por exemplo, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0$, sem que um dos factores seja nulo. Ou seja, $AB = 0 \not\Rightarrow (A = 0 \text{ ou } B = 0)$. De uma forma mais geral, $(AB = AC \text{ e } A \neq 0) \not\Rightarrow (B = C)$, já que, por exemplo, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & 3 \end{bmatrix}$.

Como exercício, calcule AB e BA, com $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

Como é fácil de observar, a soma de duas matrizes triangulares inferiores [resp. triangulares superiores] é de novo triangular inferior [resp. triangular superior]. O que se pode dizer em relação ao produto?

Teorema 2.2.1. *O produto de matrizes triangulares inferiores [resp. triangulares superiores] é de novo uma matriz triangular inferior [resp. triangular superior].*

Demonstração. Sejam A, B duas matrizes triangulares inferiores de tipo apropriado. Ou seja, $(A)_{ij}, (B)_{ij} = 0$, para $i < j$. Pretende-se mostrar que, para $i < j$ se tem $(AB)_{ij} = 0$. Ora, para $i < j$, e supondo que A tem p colunas, $(AB)_{ij} = \sum_{k=1}^{p} (A)_{ik}(B)_{kj} = \sum_{k=1}^{i} (A)_{ik}(B)_{kj} = 0$. □

Por vezes é conveniente considerar-se o produto matricial por blocos. Para tal, considere as matrizes A e B divididas em submatrizes

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

de forma conforme as operações descritas de seguida estejam definidas, então

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

De uma forma mais geral, se

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mp} \end{bmatrix}, B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{pn} & B_{pn} & \cdots & B_{pn} \end{bmatrix}$$

em que as submatrizes são tais que as operações seguintes estão bem definidas, então

$$AB = \begin{bmatrix} \sum_{k=1}^{p} A_{1k}B_{k1} & \sum_{k=1}^{p} A_{1k}B_{k2} & \cdots & \sum_{k=1}^{p} A_{1k}B_{kn} \\ \sum_{k=1}^{p} A_{2k}B_{k1} & \sum_{k=1}^{p} A_{2k}B_{k2} & \cdots & \sum_{k=1}^{p} A_{2k}B_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{p} A_{mk}B_{k1} & \sum_{k=1}^{p} A_{mk}B_{k2} & \cdots & \sum_{k=1}^{p} A_{mk}B_{kn} \end{bmatrix}.$$

Exercícios

1. Calcule $A + B$, se possível, onde

   (a) $A = \begin{bmatrix} 0 & 2 & -2 \\ -2 & -1 & -3 \\ 1 & 2 & 4 \\ -4 & 2 & -2 \end{bmatrix}$ e $B = \begin{bmatrix} 4 & 1 & 3 \\ -1 & -1 & -3 \\ 3 & 0 & -4 \\ 2 & 0 & 2 \end{bmatrix}$.

   (b) $A = \begin{bmatrix} -1 & -3 & -2 \\ -2 & 4 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} 4 & -4 & -4 \\ 3 & 0 & -1 \end{bmatrix}$.

   (c) $A = \begin{bmatrix} 3 & 3 & -2 \\ -3 & -2 & 2 \\ -2 & 0 & 2 \end{bmatrix}$ e $B = \begin{bmatrix} 4 & -2 & -2 \\ -1 & 4 & -3 \\ -4 & 0 & -4 \end{bmatrix}$

2. Calcule AB e BA, se possível, onde

   (a) $A = \begin{bmatrix} 2 & 1 & -2 \\ -3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 4 & -2 & -2 \\ -1 & 4 & -3 \\ -4 & 0 & -4 \end{bmatrix}$

   (b) $A = \begin{bmatrix} 1 & 1 & -4 \\ 3 & 4 & -3 \end{bmatrix}$ e $B = \begin{bmatrix} -3 & -1 \\ 4 & 3 \\ -1 & 0 \end{bmatrix}$

3. Considere as matrizes-linha $b = \begin{bmatrix} 0 & -1 & 12 & -23 \end{bmatrix}$ e $c = \begin{bmatrix} 3 & -4 & 0 & 3 \end{bmatrix}$.

   (a) Construa a matriz A cujas linhas são b e c.

   (b) Calcule $5b$.

   (c) Calcule $b + c$.

   (d) Calcule $\begin{bmatrix} b \\ b - c \\ c \end{bmatrix}$.

4. Considere as matrizes reais

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

(a) Calcule AB e BA e compare as respostas. O que pode inferir sobre o produto matricial?

(b) Faça o produto da linha i de A com a coluna j de B (fazendo i, j variar de 1 até 4), e compare o resultado com a entrada (i,j) de AB.

(c) Calcule $(AB)H$ e $A(BH)$ e compare os resultados. O resultado final ilustra que propriedade do produto?

(d) Calcule $(A+B)H$ e $AH+BH$ e compare os resultados. O resultado final ilustra que propriedade das operações matriciais?

5. Calcule as expressões seguintes:

(a) $\begin{bmatrix} 4 & 3 \\ 7 & 5 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$

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

(c) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -2 & 1 \\ 0 & 4 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}^3$

6. Calcule, se possível,

(a) $\begin{bmatrix} 1 & 14 & 20 \\ 8 & -25 & 9 \end{bmatrix} + \begin{bmatrix} 2 & -15 & 26 \\ 1 & 6 & -28 \end{bmatrix}$

(b) $\begin{bmatrix} -2 & 31 & 24 \\ -10 & 19 & 3 \end{bmatrix} + \begin{bmatrix} 18 & 22 & -29 \\ 15 & -9 & 17 \end{bmatrix}$

(c) $\begin{bmatrix} 9 & 4 \\ 4 & -19 \\ -4 & 1 \end{bmatrix} + \begin{bmatrix} -2 & 31 \\ 29 & 0 \\ 6 & -7 \end{bmatrix}$

(d) $\begin{bmatrix} 4 & 9 \\ 29 & 20 \\ 27 & 25 \end{bmatrix} + \begin{bmatrix} 22 & 18 & 29 \\ 1 & 8 & 22 \\ 14 & 0 & 23 \end{bmatrix}$

(e) $\begin{bmatrix} 27 & 1 & 10 \\ 30 & 30 & 4 \\ 6 & 14 & 23 \end{bmatrix} + \begin{bmatrix} 25 & 26 & 0 \\ 20 & 27 & 0 \\ 24 & 9 & 0 \end{bmatrix}$

(f) $3 \begin{bmatrix} 7 & -4 & 7 \\ 7 & 4 & -6 \end{bmatrix}$

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

(h) $\begin{bmatrix} 2 & 0 \\ 5 & 0 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 6 & 6 \\ 6 & 7 & 4 \end{bmatrix}$

(i) $\begin{bmatrix} 0 & 7 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 1 \\ 3 \end{bmatrix}$

(j) $\begin{bmatrix} 3 \\ 7 \end{bmatrix} \begin{bmatrix} 3 & -6 \end{bmatrix}$

7. Indique o valor lógico das afirmações seguintes, justificando:

   (a) Se A, B são matrizes quadradas das mesma ordem então $(AB)^n = A^n B^n$

   (b) Se A, B são matrizes quadradas das mesma ordem então $(A+B)^2 = A^2 + 2AB + B^2$

   (c) Se A, B são matrizes quadradas das mesma ordem então $A^2 - B^2 = (A+B)(A-B)$

8. Dadas matrizes diagonais D_1, D_2 quadradas com a mesma ordem, mostre que $D_1D_2 = D_2D_1$

2.2.3 Transposição

A *transposta* de uma matriz $A = [a_{ij}] \in \mathcal{M}_{m\times n}(\mathbb{K})$, é a matriz $A^T = [b_{ij}] \in \mathcal{M}_{n\times m}(\mathbb{K})$ cuja entrada (i, j) é a_{ji}, para $i = 1, \ldots, n, j = 1, \ldots, m$. Ou seja, $(A^T)_{ij} = (A)_{ji}$. A matriz é *simétrica* se $A^T = A$.

Como exemplo, a transposta da matriz $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ é a matriz $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, e a matriz $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ é uma matriz simétrica.

Repare que a coluna i de A^T é a linha i de A, e que uma matriz é simétrica se e só se for quadrada e forem iguais os elementos situados em posições simétricas relativamente à diagonal principal.

Exercícios

1. Indique A^T no caso de A ser

   (a) $\begin{bmatrix} 1 & 8 \\ 3 & 4 \\ 2 & 2 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 4 & 1 \\ 2 & 3 & 0 \\ 1 & 4 & 5 \end{bmatrix}$

2. Para as seguintes escolhas de A e de B, compare $(A+B)^T$ com A^T+B^T. O que pode inferir?

   (a) $A = \begin{bmatrix} 2 & 1 & -2 \\ -3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 4 & -2 & -2 \\ -1 & 4 & -3 \\ -4 & 0 & -4 \end{bmatrix}$

   (b) $A = \begin{bmatrix} 1 & 1 & -4 \\ 3 & 4 & -3 \end{bmatrix}$ e $B = \begin{bmatrix} -3 & -1 \\ 4 & 3 \\ -1 & 0 \end{bmatrix}^T$

3. Para $S = \begin{bmatrix} 0 & -2 & 0 & 0 \\ 2 & 2 & 0 & -2 \\ 0 & 0 & -3 & 4 \end{bmatrix}$, $X = \begin{bmatrix} 2 & -2 \\ 0 & -2 \\ -1 & -3 \\ -4 & 1 \end{bmatrix}$ e $U = \begin{bmatrix} 0 & 4 & -3 & 1 & 0 \\ -4 & -1 & -3 & -1 & 1 \end{bmatrix}$, verifique que $(SX)^T = X^T S^T$ e que $(XU)^T = U^T X^T$.

A transposição de matrizes goza das seguintes propriedades:

1. $\left(A^T\right)^T = A$;

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

3. $(\alpha A)^T = \alpha A^T$, para $\alpha \in \mathbb{K}$;

4. $(AB)^T = B^T A^T$;

5. $\left(A^k\right)^T = \left(A^T\right)^k$, $k \in \mathbb{N}$.

A afirmação (1) é válida já que $((A^T)^T)_{ij} = (A^T)_{ji} = (A)_{ij}$.

Para (2), $((A+B)^T)_{ij} = (A+B)_{ji} = (A)_{ji} + (B)_{ji} = (A^T)_{ij} + (B^T)_{ij}$.

Para (4), $((AB)^T)_{ij} = (AB)_{ji} = \sum_k (A)_{jk}(B)_{ki} = \sum_k (B)_{ki}(A)_{jk} = \sum_k (B^T)_{ik}(A^T)_{kj} = (B^T A^T)_{ij}$.

Para (5), a prova é feita por indução no expoente. Para $k=1$ a afirmação é trivialmente válida. Assumamos então que é válida para um certo k, e provemos que é válida para $k+1$. Ora $(A^{k+1})^T = (A^k A)^T =_{(4)} A^T (A^k)^T = A^T (A^T)^k = (A^T)^{k+1}$.

Exercícios

Considere a matriz real $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Considere ainda $e_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T, e_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T, e_3 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$.

1. Faça os produtos $Ae_1, Ae_2, Ae_3, e_1^T A, e_2^T A, e_3^T A$.

2. Compare $A(e_1+e_2)$ com $A\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$.
3. Preveja, e confirme, o resultado de
   (a) $A\begin{bmatrix} 2 & 0 & 0 \end{bmatrix}^T$
   (b) $A\begin{bmatrix} 0 & -1 & 0 \end{bmatrix}^T$
   (c) $A\begin{bmatrix} 2 & -1 & 0 \end{bmatrix}^T$
   (d) $A\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$
4. Mostre que, para $v \in \mathbb{R}^n$, se tem $v^Tv = 0$ se e só se $v = 0$.

2.2.4 Invertibilidade

Uma matriz A quadrada de ordem n diz-se *invertível* se existir uma matriz B, quadrada de ordem n, para a qual

$$AB = BA = I_n.$$

Teorema 2.2.2. *Seja $A \in \mathcal{M}_n(\mathbb{K})$. Se existe uma matriz $B \in \mathcal{M}_n(\mathbb{K})$ tal que $AB = BA = I_n$ então ela é única.*

Demonstração. Se B e B' são matrizes quadradas, $n \times n$, para as quais

$$AB = BA = I_n = AB' = B'A$$

então

$$B' = B'I_n = B'(AB) = (B'A)B = I_nB = B.$$

□

A matriz B do teorema, caso exista, diz-se a *inversa* de A e representa-se por A^{-1}.

Por exemplo, a matriz $S = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ não é invertível. Por absurdo, suponha que existe T, de ordem 2, tal que $ST = I_2 = TS$. A matriz T é então da forma $\begin{bmatrix} x & y \\ z & w \end{bmatrix}$. Ora $ST = \begin{bmatrix} x & y \\ x & y \end{bmatrix}$, que por sua vez iguala I_2, implicando por sua vez $x = 1$ e $y = 0$, juntamente com $x = 0$ e $y = 1$.

Considere, agora, a matriz real de ordem 2 definida por $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. Esta matriz é invertível. Mais adiante, forneceremos formas de averiguação da invertibilidade de uma matriz, bem como algoritmos para calcular a inversa. Por enquanto, verifique que a inversa de A é a matriz $X = \begin{bmatrix} -3 & 2 \\ 2 & -1 \end{bmatrix}$. Ou seja, que $AX = XA = I_2$.

O que podemos afirmar sobre o produto de duas matrizes invertíveis? Será ele uma outra matriz invertível? Em caso afirmativo, como se relaciona a inversa da matriz produto face às inversas das matrizes? O Teorema seguinte responde a esta questão.

Teorema 2.2.3. *Dadas duas matrizes U e V de ordem n, então UV é invertível e*

$$(UV)^{-1} = V^{-1}U^{-1}.$$

Demonstração. Como

$$(UV)\left(V^{-1}U^{-1}\right) = U\left(VV^{-1}\right)U^{-1} = UI_nU^{-1} = UU^{-1} = I_n$$

e

$$\left(V^{-1}U^{-1}\right)(UV) = V^{-1}\left(U^{-1}U\right)V = V^{-1}I_nV = V^{-1}V = I_n,$$

segue que UV é invertível e a sua inversa é $V^{-1}U^{-1}$. $\square$

Ou seja, o produto de matrizes invertíveis é de novo uma matriz invertível, e iguala o produto das respectivas inversas por ordem inversa.

Exercícios

1. Seja A uma matriz invertível. Mostre que $\left(A^{-1}\right)^{-1} = A$.
2. Sejam C uma matriz invertível e $A = CBC^{-1}$. Mostre que A é invertível se e só se B é invertível.
3. Dada uma matriz invertível A, mostre que toda a potência de A é também invertível.

Duas matrizes A e B, do mesmo tipo, dizem-se *equivalentes*, e denota-se por $A \sim B$, se existirem matrizes U, V invertíveis para as quais $A = UBV$. Repare que se $A \sim B$ então $B \sim A$, já que se $A = UBV$, com U, V invertíveis, então também $B = U^{-1}AV^{-1}$. Pelo teorema anterior, se $A \sim B$ então A é invertível se e só se B é invertível.

Exercícios

Mostre que equivalência de matrizes é uma relação reflexiva, simétrica e transitiva, ou seja, $A \sim A$, $A \sim B \Rightarrow B \sim A$, e $(A \sim B \land B \sim C) \Rightarrow A \sim C$.

As matrizes A e B são *equivalentes por linhas* se existir U invertível tal que $A = UB$. É óbvio que se duas matrizes A e B são equivalentes por linhas, então são equivalentes, ou seja, $A \sim B$.

Se uma matriz U for invertível, então a sua transposta U^T também é invertível e $\left(U^T\right)^{-1} = \left(U^{-1}\right)^T$. A prova é imediata, bastando para tal verificar que $\left(U^{-1}\right)^T$ satisfaz as condições de inversa, seguindo o resultado pela unicidade.

Segue também pela unicidade da inversa que

$$\left(A^{-1}\right)^{-1} = A,$$

isto é, que a inversa da inversa de uma matriz é a própria matriz.

Vimos, atrás, que o produto de matrizes triangulares inferiores [resp. superiores] é de novo uma matriz triangular inferior [resp. superior]. O que podemos dizer em relação à inversa, caso exista?

Teorema 2.2.4. *Uma matriz quadrada triangular inferior [resp. superior] é invertível se e só se tem elementos diagonais não nulos. Neste caso, a sua inversa é de novo triangular inferior [resp. superior].*

Antes de efectuarmos a demonstração, vejamos a que se reduz o resultado para matrizes (quadradas) de ordem de 2, triangulares inferiores. Seja, então, $L = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$, que supusemos invertível. Portanto, existem $x, y, z, w \in \mathbb{K}$ para os quais $I_2 = L \begin{bmatrix} x & y \\ z & w \end{bmatrix}$, donde segue, em particular, que $a_{11}x = 1$, e portanto $a_{11} \neq 0$ e $x = \frac{1}{a_{11}}$. Assim, como $a_{11}y = 0$ e $a_{11} \neq 0$ tem-se que $y = 0$. Ou seja, a inversa é triangular inferior. Como $y = 0$, o produto da segunda linha de L com a segunda coluna da sua inversa é $a_{22}w$, que iguala $(I)_{22} = 1$. Portanto, $a_{22} \neq 0$ e $w = \frac{1}{a_{22}}$. O produto da segunda linha de L com a primeira coluna da sua inversa é $a_{21}\frac{1}{a_{11}} + a_{22}z$, que iguala $(I)_{21} = 0$. Ou seja, $z = -\frac{a_{21}}{a_{11}a_{22}}$.

Demonstração. A prova é feita por indução no número de linhas das matrizes quadradas.

Para $n = 1$ o resultado é trivial. Assuma, agora, que as matrizes de ordem n triangulares inferiores invertíveis são exactamente aquelas que têm elementos diagonais não nulos. Seja $A = [a_{ij}]$ uma matriz triangular inferior, quadrada de ordem $n + 1$. Particione-se a matriz por blocos da forma seguinte:

$$\left[\begin{array}{c|c} a_{11} & O \\ \hline b & \widetilde{A} \end{array}\right],$$

onde b é $n \times 1$, O é $1 \times n$ e $\widetilde{A}$ é $n \times n$ triangular inferior.

Por um lado, se A é invertível então existe $\left[\begin{array}{c|c} x & Y \\ \hline Z & W \end{array}\right]$ inversa de A, com $x_{1\times 1}, Y_{1\times n}, Z_{n\times 1}, W_{n\times n}$. Logo $a_{11}x = 1$ e portanto $a_{11} \neq 0$ e $x = \frac{1}{a_{11}}$. Assim, como $a_{11}Y = 0$ e $a_{11} \neq 0$ tem-se que $Y = 0$. O bloco $(2, 2)$ do produto é então $\widetilde{A}W$, que iguala I_n. Sabendo que $\left[\begin{array}{c|c} x & Y \\ \hline Z & W \end{array}\right] \left[\begin{array}{c|c} a_{11} & O \\ \hline b & \widetilde{A} \end{array}\right] = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & I_n \end{array}\right]$, tem-se que também $W\widetilde{A} = I_n$, e portanto $\widetilde{A}$ é invertível,

$n \times n$, com $(\widetilde{A})^{-1} = W$. Usando a hipótese de indução aplicada a $\widetilde{A}$, os elementos diagonais de $\widetilde{A}$, que são os elementos diagonais de A à excepção de a_{11} (que já mostrámos ser não nulo) são não nulos.

Reciprocamente, suponha que os elementos diagonais de A são não nulos, e portanto que os elementos diagonais de $\widetilde{A}$ são não nulos. A hipótese de indução garante-nos a invertibilidade de $\widetilde{A}$. Basta verificar que $\left[\begin{array}{c|c} \frac{1}{a_{11}} & 0 \\ \hline -\frac{1}{a_{11}}\widetilde{A}^{-1}b & \widetilde{A}^{-1} \end{array}\right]$ é a inversa de A. □

Para finalizar esta secção, e como motivação, considere a matriz $V = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Esta matriz é invertível, e $V^{-1} = V^T$ (verifique!). Este tipo de matrizes denominam-se por ortogonais. Mais claramente, uma *matriz ortogonal* é uma matriz (quadrada) invertível, e cuja inversa iguala a sua transposta. De forma equivalente, uma matriz A invertível diz-se ortogonal se $AA^T = A^T A = I$.

Teorema 2.2.5. *1. A inversa de uma matriz ortogonal é também ela ortogonal.*

2. O produto de matrizes ortogonais é de novo uma matriz ortogonal.

Demonstração. (1) Seja A uma matriz ortogonal, ou seja, para a qual a igualdade $A^T = A^{-1}$ é válida. Pretende-se mostrar que A^{-1} é ortogonal; ou seja, que $\left(A^{-1}\right)^{-1} = \left(A^{-1}\right)^T$. Ora $\left(A^{-1}\right)^T = \left(A^T\right)^{-1} = \left(A^{-1}\right)^{-1}$.

(2) Sejam A, B matrizes ortogonais. Em particular são matrizes invertíveis, e logo AB é invertível. Mais,

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

□

Exercícios

Considere a matriz $A = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$. Mostre que A é ortogonal.

A *transconjugada* de A é a matriz $A^* = \bar{A}^T$. Ou seja, $(A^*)_{ij} = \overline{(A)_{ji}}$. Esta diz-se *hermítica* (ou *hermitiana*) se $A^* = A$.

Sejam A, B matrizes complexas de tipo apropriado e $\alpha \in \mathbb{C}$. Então

1. $(A^*)^* = A$;

2. $(A + B)^* = A^* + B^*$;

3. $(\alpha A)^* = \bar{\alpha} A^*$;

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

5. $(A^n)^* = (A^*)^n$, para $n \in \mathbb{N}$;

A prova destas afirmações é análoga à que apresentámos para a transposta, e fica ao cuidado do leitor.

Uma *matriz unitária* é uma matriz (quadrada) invertível, e cuja inversa iguala a sua transconjugada. De forma equivalente, uma matriz A invertível diz-se unitária se $AA^* = A^*A = I$.

Teorema 2.2.6. *1. A inversa de uma matriz unitária é também ela unitária.*

2. O produto de matrizes unitárias é de novo uma matriz unitária.

Remetemos o leitor ao que foi referido no que respeitou as matrizes ortogonais para poder elaborar uma prova destas afirmações.

2.3 Um resultado de factorização de matrizes

2.3.1 Matrizes elementares

Nesta secção, iremos apresentar um tipo de matrizes que terão um papel relevante em resultados vindouros: as matrizes elementares. Estas dividem-se em três tipos:

$$a \neq 0; D_k(a) = \begin{array}{cc} \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & 0 & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & a & & \\ & & 0 & & & \ddots & \\ & & & & & & 1 \end{bmatrix} & \begin{array}{c} \\ \\ \\ \\ \leftarrow k \\ \\ \\ \end{array} \\ \begin{array}{c} \uparrow \\ k \end{array} & \end{array}$$

$$i \neq j; E_{ij}\left(a\right) = \begin{array}{cc} \begin{bmatrix} 1 & & & & & 0 & \\ & \ddots & & & & & \\ & & 1 & \cdots & a & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & 0 & & & \ddots & \\ & & & & & & 1 \end{bmatrix} & \begin{array}{c} \\ \\ \leftarrow i \\ \\ \\ \\ \\ \end{array} \\ \begin{array}{c} \uparrow \\ j \end{array} & \end{array}$$

$$
P_{ij} = \begin{array}{cc}
\begin{bmatrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 &&&& 1 &&& \\
&&&& 1 & \cdots &&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
&&& 1 &&&& 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{bmatrix} &
\begin{array}{c}
\\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\
\end{array} \\
\begin{array}{ccccccccccc}
&&& \uparrow &&&& \uparrow &&& \\
&&& i &&&& j &&&
\end{array} &
\end{array}
$$

Ou seja, as matrizes elementares de ordem n são obtidas da matriz identidade I_n fazendo:

- para $D_k(a)$, substituindo a entrada (k,k) por a;
- para $E_{ij}(a)$, substituindo a entrada (i,j) por a;
- para P_{ij}, trocando as linhas i e j (ou de outra forma, as colunas i e j).

É óbvio que $D_\ell(1) = E_{ij}(0) = P_{kk} = I_n$.

A primeira propriedade que interessa referir sobre estas matrizes é que são invertíveis. Mais, para $a, b \in \mathbb{K}, a \neq 0$,

$$(D_k(a))^{-1} = D_k\left(\frac{1}{a}\right)$$

$$(E_{ij}(b))^{-1} = E_{ij}(-b), \text{ para } i \neq j$$

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

A segunda, relevante para o que se segue, indica outro o modo de se obter as matrizes $D_k(a)$ e $E_{ij}(a)$ da matriz identidade, cujas linhas são denotadas por $l_1, l_2, \ldots, l_n$:

- para $D_k(a)$, substituindo a linha k por $a\, l_k$;
- para $E_{ij}(a)$, substituindo a linha i por $l_i + a\, l_j$.

Aplicando o mesmo raciocínio, mas considerando as colunas $c_1, c_2, \ldots, c_n$ da matriz identidade:

- para $D_k(a)$, substituindo a coluna k por $a\, c_k$;
- para $E_{ij}(a)$, substituindo a coluna j por $c_j + a\, c_i$.

O que sucede se, dada uma matriz A, a multiplicarmos à esquerda ou à direita[1] por uma matriz elementar? Vejamos com alguns exemplos, tomando

$$A = \begin{bmatrix} 4 & 2 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 4 \end{bmatrix}, P = P_{12}, E = E_{31}(-2), D = D_2\left(\frac{1}{2}\right).$$

Vamos determinar o produto $DEPA$. Calcularemos primeiro PA, a este produto fazemos a multiplicação, à esquerda, por E, e finalmente ao produto obtido a multiplicação por D, de novo à esquerda. Como exercício, verifique que $PA = \begin{bmatrix} 1 & 1 & 0 \\ 4 & 2 & 0 \\ 2 & -1 & 4 \end{bmatrix}$. Qual a relação entre A e PA? Repare que ocorreu uma troca da primeira e da segunda linha de A. Que por sinal foram as mesmas trocas que se efectuaram a I_3 de forma a obtermos P_{12}.

À matriz PA, multiplicamo-la, à esquerda, por E, para obtermos $EPA = \begin{bmatrix} 1 & 1 & 0 \\ 4 & 2 & 0 \\ 0 & -3 & 4 \end{bmatrix}$. Repare que a linha 3 de EPA foi obtida somando à terceira linha de PA o simétrico do dobro da sua linha 1. Finalmente $DEPA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & -3 & 4 \end{bmatrix}$.

Exercícios

1. Considere as matrizes 3×3 elementares $E = E_{3,1}(-2), P = P_{1,2}, D = D_2(2)$.
   (a) Descreva como foram obtidas à custa das linhas/colunas da matriz I_3.
   (b) Indique a inversa de cada uma delas.
   (c) Considere $A = \begin{bmatrix} 3 & 2 & -1 \\ 8 & 3 & 3 \\ -2 & -1 & -7 \end{bmatrix}$. Faça os produtos DA, EA, PA. Relacione-as com A. Recorde o que fez na alínea (a).
   (d) Repita a alínea anterior, mas agora com os produtos AD, AE, AP.

2. Considere as matrizes $E_{2,1}(-2), E_{3,1}(-1), E_{3,2}(2)$ do tipo 3×3. Considere ainda a matriz $A = \begin{bmatrix} 5 & 1 & -3 \\ 8 & 0 & -4 \\ 3 & 0 & 10 \end{bmatrix}$.
   (a) Relacione os produtos $E_{2,1}(-2)A, E_{3,1}(-1)A, E_{3,2}(2)A$ e os produtos $AE_{2,1}(-2)$, $AE_{3,1}(-1)$ e $AE_{3,2}(2)$ com A.

[1] Recorde que o produto matricial não é, em geral, comutativo, pelo que é relevante a distinção dos dois casos.

(b) Indique uma matriz P_1 tal que $P_1 A = \begin{bmatrix} 5 & 1 & -3 \\ 3 & 0 & 10 \\ 8 & 0 & -4 \end{bmatrix}$.

(c) Indique uma matriz P_2 tal que $AP_2 = \begin{bmatrix} 1 & 5 & -3 \\ 0 & 8 & -4 \\ 0 & 3 & 10 \end{bmatrix}$.

(d) Indique uma matriz D_1 tal que $D_1 A$ é a matriz obtida de A cuja segunda linha surge dividida por 2.

(e) Indique uma matriz D_2 tal que AD_2 é a matriz obtida de A cuja terceira coluna surge multiplicada por 4.

Uma *matriz permutação* de ordem n é uma matriz obtida de I_n à custa de trocas de suas linhas (ou colunas). Aqui entra o conceito de permutação. Uma permutação no conjunto $N_n = \{1, 2, \ldots, n\}$ é uma bijecção (ou seja, uma aplicação simultaneamente injectiva e sobrejectiva) de N_n em N_n. Uma permutação $\varphi : N_n \to N_n$ pode ser representada pela tabela $\begin{pmatrix} 1 & 2 & \cdots & n \\ \varphi(1) & \varphi(2) & \cdots & \varphi(n) \end{pmatrix}$. Para simplificar a escrita, é habitual omitir-se a primeira linha, já que a posição da imagem na segunda linha indica o (único) objecto que lhe deu origem.

Definição 2.3.1. *O conjunto de todas as permutações em N_n é denotado por S_n e denominado por* grupo simétrico.

Como exemplo, considere a permutação $\gamma = (2, 1, 5, 3, 4) \in S_5$. Tal significa que

$$\gamma(1) = 2, \gamma(2) = 1, \gamma(3) = 5, \gamma(4) = 3, \gamma(5) = 4.$$

Note que S_n tem $n! = n(n-1)(n-2)\ldots 2\cdot 1$ elementos. De facto, para $\gamma = (i_1, i_2, \ldots, i_n) \in S_n$, i_1 pode tomar n valores distintos. Mas i_2 apenas pode tomar um dos $n-1$ restantes, já que não se podem repetir elementos. E assim por diante. Obtemos então $n!$ permutações distintas.

Dada a permutação $\varphi = (i_1, i_2, \ldots, i_n) \in S_n$, se $1 \leq j < k \leq n$ e $i_j > i_k$ então $i_j > i_k$ diz-se uma *inversão* de φ. Na permutação $\gamma = (2, 1, 5, 3, 4)$ acima exemplificada existem três inversões, já que $\gamma(1) > \gamma(2), \gamma(3) > \gamma(4), \gamma(3) > \gamma(5)$. O *sinal* de uma permutação φ, denotado por $sgn(\varphi)$, toma o valor $+1$ caso o número de inversões seja par, e -1 caso contrário. Portanto, $sgn(\gamma) = -1$. As permutações com sinal $+1$ chamam-se *permutações pares* (e o conjunto por elas formado chama-se grupo alterno, A_n), e as cujo sinal é -1 denominam-se por *permutações ímpares*.

Uma *transposição* é uma permutação que fixa todos os pontos à excepção de dois. Ou seja, $\tau \in S_n$ é uma transposição se existirem i, j distintos para os quais $\tau(i) = j, \tau(j) = i$ e $\tau(k) = k$ para todo o k diferente de i e j. Verifica-se que toda a permutação φ se pode escrever como composição de transposições $\tau_1, \tau_2, \ldots, \tau_r$. Ou seja, $\varphi = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_r$.

Esta decomposição não é única, mas quaisquer duas decomposições têm a mesma paridade de transposições. Ou seja, se existe uma decomposição com um número par [resp. ímpar] de intervenientes, então qualquer outra decomposição tem um número par [resp. ímpar] de transposições. Mais, esse número tem a mesma paridade da do número de inversões. Por consequência, o sinal de qualquer transposição é -1. A permutação γ definida atrás pode-se decompor como $(2,1,5,3,4) = (2,1,3,4,5) \circ (1,2,5,4,3) \circ (1,2,4,3,5)$.

O conjunto das permutações S_n pode ser identificado com o conjunto das matrizes permutação de ordem n, em que a composição de permutação é de uma forma natural identificado com o produto de matrizes. A matriz permutação P associada à permutação γ é a matriz obtida de I_5 realizando as trocas de linhas segundo γ.

Para fácil compreensão, vejamos um exemplo. Considere-se a permutação $\gamma = (2,1,5,3,4)$ e a matriz P associada a γ. Ou seja, P é a matriz obtida de I_5 realizando as trocas de linhas segundo γ, e portanto $P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$. Na primeira linha de P surge a segunda de I_5, na segunda a primeira, na terceira a quinta de I_5, e assim por diante.

Toda a matriz permutação pode-se escrever como produto de matrizes da forma P_{ij}, tal como definidas atrás. Tal é consequência da existência de uma decomposição da permutação em transposições. Note que as transposições se identificam com as matrizes P_{ij}.

Voltemos ao exemplo acima, considerando as matrizes $P_{i,j}$ associadas às transposições na decomposição de γ enunciadas atrás. Ou seja, as matrizes $P_{1,2}$, $P_{3,5}$ e $P_{3,4}$. Verifica-se que $P = P_{1,2}P_{3,5}P_{3,4}$.

Operações elementares sobre as linhas de A são as que resultam pela sua multiplicação à esquerda por matrizes elementares. Ou seja, são operações elementares por linhas de uma matriz

- a troca de duas linhas,
- a multiplicação de uma linha por um escalar não nulo,
- a substituição de uma linha pela sua soma com um múltiplo de outra linha.

De forma análoga se definem as operações elementares sobre as colunas de uma matriz, sendo a multiplicação por matrizes elementares feita à direita da matriz. Na prática, tal resulta em substituir a palavra “linha” pela palavra “coluna” na descrição acima.

Considere a matriz $A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 4 & 2 \\ -1 & 0 & 1 \end{bmatrix}$. Em primeiro lugar, e efectuando operações elementares nas linhas de A, tentaremos obter zeros por debaixo da entrada $(A)_{11}$. Ou seja, pretendemos obter algo como $\begin{bmatrix} 2 & 4 & 6 \\ 0 & ? & ? \\ 0 & ? & ? \end{bmatrix}$. Substitua-se a segunda linha, l_2, pela sua soma com o simétrico de metade da primeira. Ou seja,

$$\begin{bmatrix} 2 & 4 & 6 \\ 1 & 4 & 2 \\ -1 & 0 & 1 \end{bmatrix} \xrightarrow[l_2 \leftarrow l_2 - \frac{1}{2} l_1]{} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

Tal corresponde a multiplicar à esquerda a matriz A por $E_{21}(-\frac{1}{2}) = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Façamos o mesmo raciocínio para a terceira linha:

$$\begin{bmatrix} 2 & 4 & 6 \\ 1 & 4 & 2 \\ -1 & 0 & 1 \end{bmatrix} \xrightarrow[l_2 \leftarrow l_2 - \frac{1}{2} l_1]{} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix} \xrightarrow[l_3 \leftarrow l_3 + \frac{1}{2} l_1]{} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ 0 & 2 & 4 \end{bmatrix}$$

Tal correspondeu a multiplicar o produto obtido no passo anterior, à esquerda, por $E_{31}(\frac{1}{2})$. Ou seja, e até ao momento, obteve-se

$$E_{31}(\frac{1}{2})E_{21}(-\frac{1}{2})A = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ 0 & 2 & 4 \end{bmatrix} = B.$$

Todos os elementos na primeira coluna de B, à excepção de $(B)_{11}$, são nulos. Concentremo-nos agora na segunda coluna, e na segunda linha. Pretendem-se efectuar operações elementares nas linhas de B por forma a obter uma matriz da forma $\begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ 0 & 0 & ? \end{bmatrix}$. Para tal,

$$\begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ 0 & 2 & 4 \end{bmatrix} \xrightarrow[l_3 \leftarrow l_3 - l_2]{} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ 0 & 0 & 5 \end{bmatrix} = U.$$

Ou seja, multiplicou-se B, à esquerda, pela matriz $E_{32}(-1)$. Como $B = E_{31}(\frac{1}{2})E_{21}(-\frac{1}{2})A$ e $E_{32}(-1)B = U$ podemos concluir que

$$E_{32}(-1)E_{31}(\frac{1}{2})E_{21}(-\frac{1}{2})A = U = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & -1 \\ 0 & 0 & 5 \end{bmatrix}$$

Repare que U é uma matriz triangular superior, e que neste exemplo tem elementos diagonais não nulos, e portanto é uma matriz invertível. Como as matrizes elementares são invertíveis e $(E_{32}(-1)E_{31}(\frac{1}{2})E_{21}(-\frac{1}{2}))^{-1}U = A$, segue que a matriz A é também ela invertível. Note ainda que $(E_{32}(-1)E_{31}(\frac{1}{2})E_{21}(-\frac{1}{2}))^{-1} = E_{21}(\frac{1}{2})E_{31}(-\frac{1}{2})E_{32}(1)$. A estratégia descrita acima aplicada à matriz A é denominada por *algoritmo de eliminação de Gauss*. O resultado final foi a factorização $A = LU$, onde U é uma matriz triangular superior (veremos mais adiante que de facto pertence a uma subclasse desse tipo de matrizes) e L é uma matriz invertível triangular inferior (por ser a inversa de produto de matrizes invertíveis triangulares inferiores). Nem sempre é possível percorrer estes passos do algoritmo, para uma matriz dada arbitrariamente. Veremos, na próxima secção, que modificações se realizam na estratégia apresentada acima por forma a que se garanta algum tipo de factorização.

O exemplo escolhido foi, de facto, simples na aplicação. Alguns passos podem não ser possíveis, nomeadamente o primeiro. Repare que o primeiro passo envolve uma divisão (no nosso caso, dividimos a linha 1 por $(A)_{11}$). A propósito, os elementos-chave na divisão, ou de forma mais clara, o primeiro elemento não nulo da linha a que vamos tomar um seu múltiplo denomina-se por *pivot*. Ora esse pivot tem que ser não nulo. E se for nulo? Nesse caso, trocamos essa linha por outra mais abaixo que tenha, nessa coluna, um elemento não nulo. E se *todos* forem nulos? Então o processo terminou para essa coluna e consideramos a coluna seguinte. Apresentamos dois exemplos, um para cada um dos casos descritos:

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 2 \\ -3 & 2 & 9 \end{bmatrix} ; \begin{bmatrix} 0 & 1 & 1 \\ 0 & 6 & 7 \\ 0 & 1 & -2 \end{bmatrix} .$$

No primeiro caso, a troca da primeira linha pela linha dois ou três resolve o problema. No segundo caso, aplicamos a estratégia a partir da segunda coluna. Recorde que a troca da linha i pela linha j é uma operação elementar de linhas que corresponde à multiplicação, à esquerda, por P_{ij}.

Apresentamos, de seguida, o *algoritmo de eliminação de Gauss* de uma forma mais formal.

2.3.2 O Algoritmo de Eliminação de Gauss

O *Algoritmo de Eliminação de Gauss*, (*abrev.* AEG), segue os passos que em baixo se descrevem:

Seja A uma matriz $m \times n$ não nula.

1. Assuma que $(A)_{11} \neq 0$. Se tal não acontecer, então troque-se a linha 1 com uma linha i para a qual $(A)_{i1} \neq 0$. Ou seja, multiplique A, à esquerda, por P_{1i}. Para simplificar a notação, A denotará tanto a matriz original como a obtida por troca de duas das suas linhas. A $(A)_{11}$ chamamos *pivot* do algoritmo. Se todos os elementos da primeira coluna são nulos, use 2.

2. Se a estratégia indicada no passo 1 não for possível (ou seja, os elementos da primeira coluna são todos nulos), então aplique de novo o passo 1 à submatriz obtida de A retirando a primeira coluna.

3. Para $i = 2, \ldots, m$, e em A, substitua a linha i pela sua soma com um múltiplo da linha 1 por forma a que o elemento obtido na entrada $(i, 1)$ seja 0. Tal corresponde a multiplicar a matriz A, à esquerda, por $E_{i1}\left(-\frac{(A)_{i1}}{(A)_{11}}\right)$.

4. Repita os passos anteriores à submatriz da matriz obtida pelos passos descritos, a que se retirou a primeira linha e a primeira coluna.

Após se aplicar o passo 3 em todas as linhas e na primeira coluna, e supondo que $(A)_{11} \neq 0$, a matriz que se obtem tem a forma seguinte:

$$\begin{bmatrix} (A)_{11} & (A)_{12} & (A)_{13} & (A)_{1n} \\ 0 & ? & ? & ? \\ 0 & ? & ? & ? \\ \vdots & ? & ? & ? \\ 0 & ? & ? & ? \end{bmatrix}.$$

Ou seja, e por operações elementares de linhas, podemos obter de A uma matriz com a forma $\left[\begin{array}{c|c} (A)_{11} & * \\ \hline 0 & \widetilde{A} \end{array}\right]$. O algoritmo continua agora aplicado à matriz $\widetilde{A}$ segundo os passos 1, 2 e 3. Note que as operações elementares operadas nas linhas de $\widetilde{A}$ são também elas operações elementares realizadas nas linhas de $\left[\begin{array}{c|c} (A)_{11} & * \\ \hline 0 & \widetilde{A} \end{array}\right]$. As operações elementares efectuadas em $\widetilde{A}$ dão origem a uma matriz da forma $\left[\begin{array}{c|c} (\widetilde{A})_{11} & * \\ \hline 0 & \widetilde{\widetilde{A}} \end{array}\right]$, onde consideramos $(\widetilde{A})_{11} \neq 0$. Essas operações elementares aplicadas às linhas de $\left[\begin{array}{c|c} (A)_{11} & * \\ \hline 0 & \widetilde{A} \end{array}\right]$ dão lugar à matriz $\left[\begin{array}{cc|c} (A)_{11} & \dots & (A)_{1m} \\ 0 & (\widetilde{A})_{11} & * \\ \hline 0 & 0 & \widetilde{\widetilde{A}} \end{array}\right]$. Note que se supôs que as entradas (i, i) são não nulas, ou que existe uma troca conveniente de linhas por forma a se contornar essa questão. Como é óbvio, tal pode não ser possível. Nesse caso aplica-se o passo 2. Ou seja, e quando tal acontece, tal corresponde à não existência de pivots em colunas consecutivas. Como exemplo, considere a matriz $M = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$. Multiplicando esta matriz, à esquerda, por $E_{31}(-\frac{1}{2})E_{21}(-1)$, ou seja, substiuindo a linha 2 pela sua soma com o simétrico da linha 1, e a linha 3 pela sua soma com metade do simétrico da linha 1, obtemos a matriz $M_2 = \left[\begin{array}{c|ccc} 2 & 2 & 2 & 2 \\ \hline 0 & 0 & 0 & -2 \\ 0 & 0 & -1 & 0 \end{array}\right]$.

Aplicamos agora o algoritmo à submatriz $\widetilde{M} = \begin{bmatrix} 0 & 0 & -2 \\ 0 & -1 & 0 \end{bmatrix}$. Note que a esta submatriz teremos que aplicar (2) por impossibilidade de se usar (1); de facto, não há elementos não nulos na primeira coluna de $\widetilde{M}$. Seja, então, $\widetilde{M_2}$ a matriz obtida de $\widetilde{M}$ a que retirou a primeira coluna; ou seja, $\widetilde{M_2} = \begin{bmatrix} 0 & -2 \\ -1 & 0 \end{bmatrix}$. É necessário fazer a troca das linhas por forma a obtermos um elemento não nulo que terá as funções de pivot. Essa troca de linhas é uma operação elementar também na matriz original $M_2 = \left[\begin{array}{cc|cc} 2 & 2 & 2 & 2 \\ \hline 0 & 0 & 0 & -2 \\ 0 & 0 & -1 & 0 \end{array}\right]$. Tal corresponde a multiplicá-la, à esquerda, por P_{23}. Repare que, sendo os elementos nas linhas 2 e 3 e nas colunas 1 e 2 nulos, a troca das linhas de facto apenas altera as entradas que estão simultaneamente nas linhas envolvidas e nas entradas à direita do novo pivot. Obtemos, assim, a matriz $\begin{bmatrix} 2 & 2 & 2 & 2 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$. A matriz obtida tem uma particularidade: debaixo de cada pivot todos os elementos são nulos.

Como foi referido, a matriz obtida por aplicação dos passos descritos no Algoritmo de Eliminação de Gauss tem uma forma muito particular. De facto, debaixo de cada pivot todos os elementos são nulos. A esse tipo de matriz chamamos *matriz escada (de linhas)*. Uma matriz $A = [a_{ij}]$ é matriz escada (de linhas) se

(i) se $a_{ij} \neq 0$ com $a_{ik} = 0$, para $k < j$, então $a_{lk} = 0$ se $k \leq j$ e $l > i$;

(ii) as linhas nulas surgem depois de todas as outras.

Sempre que o contexto o permita, diremos matriz escada para significar matriz escada de linhas.

A matriz $U = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ é uma matriz escada (de linhas) que se obteve de M por aplicação dos passos (1)–(4). É óbvio que uma matriz escada é triangular superior, mas o recíproco não é válido em geral. Como exemplo, considere a matriz $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$.

Teorema 2.3.2 (Factorização $PA = LU$)**.** *Dada uma matriz A, existem matrizes P permutação, L triangular inferior com 1's na diagonal principal e U matriz escada para as quais $PA = LU$.*

Ou seja, a matriz A iguala $P^{-1}LU$. Portanto, toda a matriz é equivalente por linhas a uma matriz escada de linhas.

Antes de procedermos à prova deste resultado, abrimos um parênteses para apresentarmos dois exemplos que servem de motivação ao lema que se segue.

Considere a matriz $A = \begin{bmatrix} 0 & 3 & -2 \\ -1 & 3 & 0 \\ 1 & 3 & -5 \end{bmatrix}$. A troca da primeira com a segunda linhas dá origem à matriz $\widetilde{A} = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 3 & -2 \\ 1 & 3 & -5 \end{bmatrix}$, a qual, e usando o AEG descrito atrás, satisfaz $E_{32}(-2)E_{31}(1)\widetilde{A} = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 3 & -2 \\ 0 & 0 & -1 \end{bmatrix}$. Ou seja, existem matrizes P permutação, L triangular inferior com 1's na diagonal e U matriz escada para as quais $PA = LU$. Para tal, basta tomar $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $L = (E_{32}(-2)E_{31}(1))^{-1} = E_{31}(-1)E_{32}(2)$, e $U = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 3 & -2 \\ 0 & 0 & -1 \end{bmatrix}$.

Considere agora a matriz $M = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$. Ora $E_{31}(-1)M = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$, o que força a troca da segunda pela terceira linha. Obtemos, assim, $P_{23}E_{31}(-1)M = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, que é uma matriz escada. Neste caso, como se obtêm as matrizes P, L, U do teorema? Ao contrário do exemplo anterior, a realização matricial das operações elementares por linhas do AEG não nos fornece, de forma imediata, essa factorização. No entanto, poder-se-ia escrever $E_{31}(-1)M = P_{23}\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, já que $P_{23}^{-1} = P_{23}$, e portanto $M = E_{31}(1)P_{23}\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, pois $E_{31}(-1)^{-1} = E_{31}(1)$. Note que $E_{31}(1)P_{23} \neq P_{23}E_{31}(1)$. Não obstante, repare que $E_{31}(1)P_{23} = P_{23}E_{21}(1)$, donde $M = P_{23}E_{21}(1)\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, e portanto $PA = LU$, com $P = P_{23}, L = E_{21}(1)$ e $U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Lema 2.3.3. *Para $i, k, l > j$, e para todo o $a \in \mathbb{K}$, é válida a igualdade $E_{ij}(a)P_{kl} = P_{kl}E_{lj}(a)$.*

Demonstração. Se $k \neq i$, então a igualdade é óbvia.

Suponha que $k = i$. Pretende-se mostrar que $E_{ij}(a)P_{il} = P_{il}E_{lj}(a)$, com $i, l > j$. Sendo $P_{il}E_{lj}(a)$ a matriz obtida de $E_{lj}(A)$ trocando as linhas i e l, e visto a linha l de $E_{lj}(a)$ ser

$$\begin{array}{ccccccccccc} [0 & \cdots & 0 & a & 0 & \cdots & 0 & 1 & 0 & \cdots & 0] \\ & & & \uparrow & & & & \uparrow & & & \\ & & & j & & & & l & & & \end{array}$$

então a linha i de $P_{il}E_{lj}(a)$ é

$$\begin{array}{ccccccccccc} [0 & \cdots & 0 & a & 0 & \cdots & 0 & 1 & 0 & \cdots & 0] \\ & & & \uparrow & & & & \uparrow & & & \\ & & & j & & & & l & & & \end{array}.$$

$E_{ij}(a)P_{il}$ é a matriz obtida de P_{il} a que à linha i se somou a linha j de P_{il} multiplicada por a. Sendo a linha i de P_{il}

$$\begin{array}{ccccccccc} [0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0] \\ & & & & & \uparrow & & & \\ & & & & & l & & & \end{array}$$

e a linha j de P_{il}, e já que $j < i, l$,

$$\begin{array}{ccccccccc} [0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0] \\ &&&&& \uparrow &&& \\ &&&&& j &&& \end{array}$$

segue que a linha i de $E_{ij}(a)P_{il}$ é a soma

$$\begin{array}{cccccccccccccccccccc} & [0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0]+ & a\,[0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0] \\ &&&&&& \uparrow &&&&&&&&& \uparrow &&& \\ &&&&&& l &&&&&&&&& j &&& \\ = & [0 & \cdots & 0 & a & 0 & \cdots & 0 & 1 & 0 & \cdots & 0] &&&&&&& \\ &&&& \uparrow &&&& \uparrow &&&&&&&&&& \\ &&&& j &&&& l &&&&&&&&&& \end{array}$$

Para $k \neq i$, a linha k de $E_{ij}(a)P_{il}$ é a linha k de P_{il}, sendo esta a linha k da matriz identidade se $k \neq l$, ou a linha i da identidade se $k = l$. Por sua vez, a linha k de $P_{il}E_{lj}(a)$ é a linha k da ientidade se $k \neq l$, ou é a linha i de I_n se $k = l$. □

Demonstração do teorema 2.3.2. A prova segue da aplicação do algoritmo de eliminação de Gauss, fazendo-se uso do lema para se obter a factorização da forma $U = PLA$, onde os pivots do algoritmo são o primeiro elemento não nulo de cada linha (não nula) de U. □

Exercícios

1. Considere a matriz $A = \begin{bmatrix} 8 & 2 & 3 \\ 4 & 3 & 2 \\ 1 & -2 & 1 \end{bmatrix}$.
   (a) Calcule $B = E_{21}(-\frac{1}{2})A$.
   (b) Indique uma matriz elementar da forma $E_{ij}(\alpha)$ tal que $C = E_{ij}(\alpha)B$ seja uma matriz com as entradas $(2,1)$ e $(3,1)$ nulas.
   (c) Indique uma matriz elementar E tal que EC é uma matriz triangular superior.
   (d) Indique uma matriz invertível K triangular inferior tal que KA é triangular superior.
   (e) Mostre existe uma matriz triangular superior U e L triangular inferior invertível para as quais $A = LU$.
   (f) Conclua que a matriz A é invertível.

2. Considere a matriz $A = \begin{bmatrix} 2 & 4 & 3 \\ -1 & 4 & 0 \\ 3 & 1 & 1 \end{bmatrix}$.
   (a) Indique uma matriz invertível K triangular inferior tal que KA é triangular superior.
   (b) Mostre existe uma matriz triangular superior U e L triangular inferior invertível para as quais $A = LU$.

(c) Conclua que a matriz A é invertível.

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

   (a) Indique uma matriz P tal que $PA = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

   (b) Indique uma matriz invertível K triangular inferior tal que KPA é triangular superior.

   (c) Mostre existe uma matriz triangular superior U e L triangular inferior invertível para as quais $PA = LU$.

   (d) Conclua que a matriz A é invertível.

4. Considere matrizes 3×3.

   (a) Mostre que $E_{31}(2)P_{23} = P_{23}E_{21}(2)$.

   (b) Mostre que $E_{32}(1)P_{13} = P_{13}E_{12}(1)$.

   (c) Indique uma matriz permutação P e uma matriz elementar da forma $E_{ij}(\alpha)$ para as quais $E_{21}(-3)P_{23} = PE_{ij}(\alpha)$.

   (d) Indique uma matriz permutação P e uma matriz elementar da forma $E_{ij}(\alpha)$ para as quais $P_{32}E_{21}(-1) = E_{ij}(\alpha)P$.

5. Considere a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ -1 & 1 & 2 \end{bmatrix}$.

   (a) Indique uma matriz Y, à custa de produtos de matrizes elementares, tal que YA é triangular superior.

   (b) Deduza que A é invertível.

   (c) Factorize $Y = K\tilde{P}$, onde $\tilde{P}$ é uma matriz permutação e K é triangular inferior.

   (d) Mostre existe uma matriz permutação P, uma triangular superior U e L triangular inferior invertível para as quais $PA = LU$.

6. Encontre uma factorização da forma $PA = LU$ para $A = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & -1 & 0 & 2 \\ 1 & 0 & 0 & 1 \end{bmatrix}$.

A *característica* de uma matriz A, denotada por $\mathrm{car}(A)$, por $c(A)$ ou ainda por $\mathrm{rank}(A)$, é o número de linhas não nulas na matriz escada U obtida por aplicação do Algoritmo de Eliminação de Gauss. Ou seja, e sabendo que toda a linha não nula de U tem exactamente 1 pivot que corresponde ao primeiro elemento não nulo da linha, a característica de A é o número de pivots no algoritmo (ainda que o último possa não ser usado, por exemplo, no caso de estar na última linha). Note ainda que $\text{car}(A) = \text{car}(U)$. Por exemplo, $\text{car}\begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} = 3$, já que a matriz escada obtida desta tem 3 linhas não nulas.

Uma matriz quadrada A de ordem n diz-se *não-singular* se $\text{car}(A) = n$. Ou seja, A é não-singular se forem usados n pivots no algoritmo de eliminação de Gauss. Uma matriz é *singular* se não for não-singular.

Teorema 2.3.4. *As matrizes não-singulares são exactamente as matrizes invertíveis.*

Demonstração. Seja A uma matriz quadrada, e U a matriz escada obtida de A por Gauss.

Por um lado, se A é invertível, e como $A \sim U$, segue que U é invertível, quadrada. Como U é triangular superior, não pode ter linhas nulas caso contrário teria um elemento diagonal nulo, o que contraria a invertibilidade de U.

Por outro lado, se A é não-singular então U não tem linhas nulas. Como cada coluna de U tem no máximo 1 pivot, e existem n linhas e n pivots, então cada linha tem exactamente 1 pivot. Ou seja, os elementos diagonais de U são não nulos. Como U é triangular superior, segue que U é invertível, e portanto A é invertível visto $A \sim U$. $\square$

Teorema 2.3.5. *Se A é uma matriz não-singular, então existe uma matriz P permutação tal que PA é factorizável, de forma única, como $PA = LU$, onde L é triangular inferior com 1's na diagonal e U é uma matriz triangular superior com elementos diagonais não nulos.*

Demonstração. A existência de tal factorização é consequência do teorema 2.3.2. Repare que, sendo a matriz não singular, tal significa que os pivots estão presentes em todas as colunas de U. Assim, os elementos diagonais de U são os pivots, sendo estes não nulos. Resta-nos provar a unicidade. Para tal, considere as matrizes L_1, L_2 triangulares inferiores com 1's na diagonal, e as matrizes U_1, U_2 triangulares superiores com elementos diagonais diferentes de zero, matrizes essas que satisfazem $PA = L_1U_1 = L_2U_2$. Portanto, $L_1U_1 = L_2U_2$, o que implica, e porque L_1, U_2 são invertíveis (porquê?), que $U_1U_2^{-1} = L_1^{-1}L_2$. Como L_1, U_2 são, respectivamente, triangulares inferior e superior, então L_1^{-1} e U_2^{-1} são também triangulares inferior e superior, respectivamente. Recorde que sendo a diagonal de L_1 constituída por 1's, então a diagonal da sua inversa tem também apenas 1's. Daqui segue que $L_1^{-1}L_2$ é triangular inferior, com 1's na diagonal, e que $U_1U_2^{-1}$ é triangular superior. Sendo estes dois produtos iguais, então $L_1^{-1}L_2$ é uma matriz diagonal, com 1's na diagonal; ou seja, $L_1^{-1}L_2 = I$, e portanto $L_1 = L_2$. Tal leva a que $L_1U_1 = L_1U_2$, o que implica, por multiplicação à esquerda por L_1^{-1}, que $U_1 = U_2$. $\square$

Exercícios

1. Considere a matriz $A = \begin{bmatrix} 1 & -3 & 1 \\ 2 & -4 & 2 \\ 2 & 2 & -3 \end{bmatrix}$.
   (a) Explique como se obteve $E_{2,1}(-2)$ à custa das linhas de I_3, e como se obteve $E_{2,1}(-2)A$ à custa das linhas de A.
   (b) Efectue os passos seguintes do Algoritmo de Eliminação de Gauss para obter a matriz escada U equivalente por linhas a A.
   (c) Use as matrizes elementares de (a) para construir a matriz V tal que $VA = U$. Diga por que razão V é invertível.
   (d) Use a alínea anterior para determinar L triangular inferior tal que $A = LU$.
   (e) Indique a característica da matriz A. Diga, justificando, se a matriz é invertível.
2. Considere a matriz $B = \begin{bmatrix} 4 & 2 & 2 & -2 \\ 2 & 0 & 5 & 1 \\ -2 & -3 & 5 & 4 \end{bmatrix}$.
   (a) Efectue os passos do Algoritmo de Eliminação de Gauss para obter a matriz escada U equivalente por linhas a B. Identifique os *pivots*.
   (b) Use as matrizes elementares de (a) para construir a matriz V tal que $VB = U$. Diga por que razão V é invertível.
   (c) Encontre a decomposição LU de B.
   (d) Indique a característica da matriz B.
3. Considere a matriz $C = \begin{bmatrix} 1 & 2 & 3 \\ -2 & -1 & -5 \\ -1 & 4 & -3 \\ 2 & 1 & 1 \end{bmatrix}$.
   (a) Efectue os passos do Algoritmo de Eliminação de Gauss para obter a matriz escada U equivalente por linhas a C.
   (b) Use as matrizes elementares de (a) para construir a matriz V tal que $VC = U$. Diga por que razão V é invertível.
   (c) Indique a característica da matriz C.
4. Considere a matriz $G = \begin{bmatrix} 0 & 3 & -2 \\ -1 & 3 & 0 \\ 2 & 3 & -5 \end{bmatrix}$.
   (a) Efectue os passos do Algoritmo de Eliminação de Gauss para obter a matriz escada U equivalente por linhas a G.
   (b) Use as matrizes elementares de (a) para construir uma matriz V tal que $VG = U$. Diga por que razão V é invertível. Verifique <u>se</u> V é triangular inferior.

(c) Indique a característica da matriz G.

5. Calcule a característica das matrizes seguintes, fazendo uso do Algoritmo de Eliminação de Gauss.

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

$$D = \begin{bmatrix} 4 & 3 & 1 \\ 1 & 1 & 0 \\ 2 & 3 & 1 \\ 3 & 6 & -2 \end{bmatrix}, E = \begin{bmatrix} 0 & 1 & 6 & -5 \\ -3 & 2 & 9 & -1 \\ -2 & 1 & 4 & 1 \\ -5 & 2 & 7 & 5 \end{bmatrix}, F = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \\ 2 & 6 & 11 \\ 0 & -1 & -3 \end{bmatrix}.$$

6. Determine $k \in \mathbb{R}$ por forma que a característica da matriz

$$F = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -1 & 2 & 1 \\ 1 & 2 & 1 & k \end{bmatrix}$$

seja inferior a 3.

2.4 Determinantes

2.4.1 Definição

Considere a matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e suponha que $a \neq 0$. Aplicando o AEG, obtemos a factorização $\begin{bmatrix} 1 & 0 \\ -\frac{c}{a} & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & -\frac{bc}{a} + d \end{bmatrix}$. Ou seja, a matriz A é equivalente por linhas à matriz $U = \begin{bmatrix} a & b \\ 0 & -\frac{bc}{a} + d \end{bmatrix}$, que é uma matriz triangular superior. Recorde que A é invertível se e só se U for invertível. Ora, a matriz U é invertível se e só se $-\frac{bc}{a} + d \neq 0$, ou de forma equivalente, se $ad - bc \neq 0$. Portanto, A é invertível se e só se $ad - bc \neq 0$.

Este caso simples serve de motivação para introduzir a noção de determinante de uma matriz.

Na definição que se apresenta de seguida, S_n indica o grupo simétrico (ver Definição 2.3.1).

Definição 2.4.1. *Seja A uma matriz quadrada de ordem n. O* determinante *de A, denotado por* $\det A$ *ou* $|A|$*, é o escalar definido por*

$$\sum_{\sigma \in S_n} sgn(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Vejamos o que resulta da fórmula quando consideramos matrizes 2×2 e matrizes 3×3.

Seja $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Neste caso, o grupo simétrico S_2 tem apenas as permutações $\sigma_1 = (1\,2)$ e $\sigma_2 = (2\,1)$, sendo que $sgn(\sigma_1) = 1$ e que $sgn(\sigma_2) = -1$. Recorde que $\sigma_1(1) = 1, \sigma_1(2) = 2, \sigma_2(1) = 2$ e $\sigma_2(2) = 1$. Obtemos, então, $|A| = a_{11}a_{22} - a_{12}a_{21}$.

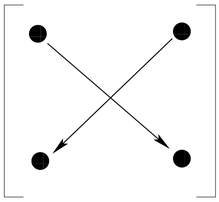


Figura 2.1: Esquema do cálculo do determinante de matrizes de ordem 2

Seja agora $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Recorde que S_3 tem 6 elementos. No quadro seguinte, indicamos, respectivamente, a permutação $\sigma \in S_3$, o seu sinal, e o produto $a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n\sigma(n)}$.

Permutação $\sigma \in S_3$	$sgn(\sigma)$	$a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n\sigma(n)}$
$(1\,2\,3)$	$+1$	$a_{11}a_{22}a_{33}$
$(2\,3\,1)$	$+1$	$a_{12}a_{23}a_{31}$
$(3\,1\,2)$	$+1$	$a_{13}a_{21}a_{32}$
$(1\,3\,2)$	-1	$a_{11}a_{23}a_{32}$
$(2\,1\,3)$	-1	$a_{12}a_{21}a_{33}$
$(3\,2\,1)$	-1	$a_{13}a_{22}a_{31}$

Obtemos, assim,

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

Para fácil memorização, pode-se recorrer ao esquema apresentado de seguida.

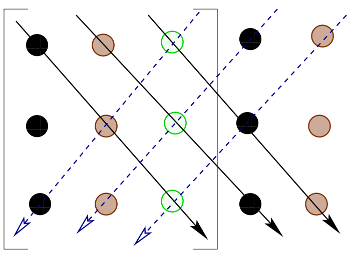


Figura 2.2: Esquema do cálculo do determinante de matrizes de ordem 3, ou a Regra de Sarrus

2.4.2 Propriedades

São consequência da definição os resultados que de seguida apresentamos, dos quais omitimos a demonstração.

Teorema 2.4.2. *Seja A uma matriz quadrada.*

1. *Se A tem uma linha ou uma coluna nula então $|A| = 0$.*
2. $|A| = |A^T|$.
3. *Se A é triangular (inferior ou superior) então* $|A| = \prod_{i=1,\ldots,n} (A)_{ii}$.
4. $|P_{ij}| = -1, |D_k(a)| = a, |E_{ij}(a)| = 1,$ *com* $i \neq j$.

Daqui segue que $|I_n| = 1$. Segue também que dada uma matriz triangular (inferior ou superior) que esta é invertível se e só se tiver determinante não nulo. Mais adiante, apresentaremos um resultado que generaliza esta equivalência para matrizes quadradas não necessariamente triangulares.

Teorema 2.4.3. *Dada uma matriz A quadrada, $a \in \mathbb{K}$,*

1. $|D_i(a)A| = a|A| = |AD_i(a)|$;
2. $|P_{ij}A| = |AP_{ij}| = -|A|$;
3. $|E_{ij}(a)A| = |A| = |AE_{ij}(a)|$.

Como $|D_i(A)| = a, |P_{ij}| = -1$ e $|E_{ij}(a)| = 1$, segue que $|D_i(a)A| = |D_i(a)||A|$, $|P_{ij}A| = |P_{ij}||A|$ e que $|E_{ij}(a)A| = |E_{ij}(a)||A|$. Repare ainda que, se A é $n \times n$, é válida a igualdade $|\alpha A| = \alpha^n|A|$, já que $\alpha A = \prod_{i=1}^n D_i(\alpha)A$. De forma análoga, dada uma matriz diagonal D com elementos diagonais $d_1, d_2, \ldots, d_n$, tem-se $|DA| = d_1 d_2 \cdots d_n |A| = |D||A|$.

Corolário 2.4.4. *Uma matriz com duas linhas/colunas iguais tem determinante nulo.*

Demonstração. Se a matriz tem duas linhas iguais, digamos i e j, basta subtrair uma à outra, que corresponde a multiplicar à esquerda pela matriz $E_{ij}(-1)$. A matriz resultante tem uma linha nula, e portanto tem determinante zero. Para colunas iguais, basta aplicar o mesmo raciocínio a A^T. $\square$

O corolário anterior é passível de ser generalizado considerando não linhas iguais, mas tal que uma linha se escreva como soma de múltiplos de outras linhas. O mesmo se aplica a colunas.

Corolário 2.4.5. *Tem determinante nulo uma matriz que tenha uma linha que se escreve como a soma de múltiplos de outras das suas linhas.*

Demonstração. Suponha que a linha i, ℓ_i, de uma matriz A se escreve como a soma de múltiplos de outras das suas linhas, ou seja, que $\ell_i = \sum_{j\in J} \alpha_j \ell_j = \alpha_{j_1}\ell_{j_1} + \alpha_{j_2}\ell_{j_2} + \cdots + \alpha_{j_s}\ell_{j_s}$. A linha i de $E_{ij_1}(-\alpha_{j_1})A$ é a matriz obtida de A substituindo a sua linha i por $\ell_i - \alpha_{j_1}\ell_{j_1} = \alpha_{j_2}\ell_{j_2} + \cdots + \alpha_{j_s}\ell_{j_s}$. Procedemos ao mesmo tipo de operações elementares por forma a obtermos uma matriz cuja linha i é nula. Como o determinante de cada uma das matrizes obtidas por operação elementar de linhas iguala o determinante de A, e como a última matriz tem uma linha nula, e logo o seu determinante é zero, segue que $|A| = 0$. □

Corolário 2.4.6. *Seja U a matriz obtida da matriz quadrada A por Gauss. Então $|A| = (-1)^r|U|$, onde r indica o número de trocas de linhas no algoritmo.*

Sabendo que uma matriz é invertível se e só se a matriz escada associada (por aplicação de Gauss) é invertível, e que esta sendo triangular superior é invertível se e só se os seus elementos diagonais são todos não nulos, segue que, e fazendo uso de resultados enunciados e provados anteriormente,

Corolário 2.4.7. *Sendo A uma matriz quadrada de ordem n, as afirmações seguintes são equivalentes:*

1. *A é invertível;*

2. *$|A| \neq 0$;*

3. *$\operatorname{car}(A) = n$;*

4. *A é não-singular.*

Portanto, uma matriz com duas linhas/colunas iguais não é invertível. Mais, uma matriz que tenha uma linha que se escreva como soma de múltiplos de outras das suas linhas não é invertível.

Teorema 2.4.8. *Seja A e B matrizes $n \times n$.*

$$|AB| = |A||B|.$$

Demonstração. Suponha que A é invertível.

Existem matrizes elementares $E_1, \ldots, E_s$ e uma matriz escada (de linhas) U tal que $A = E_1E_2 \ldots E_sU$. Ora existem também $E_{s+1}, \ldots, E_r$ matrizes elementares, e U_1 matriz escada de linhas para as quais $U^T = E_{s+1} \ldots E_rU_1$. Note que neste último caso se pode pressupor que não houve trocas de linhas, já que os pivots do AEG são os elementos diagonais de U já que U^T é triangular inferior, que são não nulos por A ser invertível. Ora U_1 é então uma matriz triangular superior que se pode escrever como produto de matrizes triangulares inferiores, e portanto U_1 é uma matriz diagonal. Seja $D = U_1$. Resumindo, $A = E_1E_2 \ldots E_s(E_{s+1} \ldots E_rD)^T = E_1E_2 \ldots E_sDE_r^TE_{r-1}^T \ldots E_{s+1}^T$. Recorde que, dada uma matriz elementar E, é válida $|EB| = |E||B|$. Então,

$$\begin{aligned}
|AB| &= |E_1E_2 \dots E_s D E_r^T E_{r-1}^T \dots E_{s+1}^T B| \\
&= |E_1||E_2 \dots E_s D E_r^T E_{r-1}^T \dots E_{s+1}^T B| \\
&= |E_1||E_2||E_3 \dots E_s D E_r^T E_{r-1}^T \dots E_{s+1}^T B| \\
&= \cdots \\
&= |E_1||E_2||E_3| \dots |E_s||D||E_r^T||E_{r-1}^T| \dots |E_{s+1}^T||B| \\
&= |E_1E_2E_3 \dots E_s D E_r^T E_{r-1}^T \dots E_{s+1}^T||B| \\
&= |A||B|.
\end{aligned}$$

Se A não é invertível, e portanto $|A| = 0$, então AB não pode ser invertível, e portanto $|AB| = 0$. □

Como $|I_n| = 1$, segue do teorema anterior a relação entre o determinante uma matriz invertível com o da sua inversa.

Corolário 2.4.9. *Se A é uma matriz invertível então*

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

Recorde que para que uma matriz A seja invertível exige-se a existência de uma outra X para a qual $AX = I_n = XA$. O resultado seguinte mostra que se pode prescindir da verificação de uma das igualdades.

Corolário 2.4.10. *Seja A uma matriz $n \times n$. São equivalentes:*

1. *A é invertível*
2. *existe uma matriz X para a qual $AX = I_n$*
3. *existe uma matriz Y para a qual $YA = I_n$*

Nesse caso, $A^{-1} = X = Y$.

Demonstração. As equivalências são imediatas, já que se $AX = I_n$ então $1 = |I_n| = |AX| = |A||X|$ e portanto $|A| \neq 0$.

Para mostrar que $A^{-1} = X$, repare que como $AX = I_n$ então A é invertível, e portanto $A^{-1}AX = A^{-1}$, donde $X = A^{-1}$. □

Faça a identificação dos vectores $(a, b) \in \mathbb{R}^2$ com as matrizes coluna $\begin{bmatrix} a \\ b \end{bmatrix}$. O produto interno usual $(u_1, u_2) \cdot (v_1, v_2)$ em $\mathbb{R}^2$ pode ser encarado como o produto matricial $\begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Ou seja, $u \cdot v = u^T v$. Esta identificação e noção pode ser generalizada de forma trivial para $\mathbb{R}^n$. Dois vectores u e v de $\mathbb{R}^n$ dizem-se ortogonais, $u \perp v$, se $u \cdot v = u^T v = 0$. A norma usual em $\mathbb{R}^n$ é definida por $\|u\| = \sqrt{u \cdot u}$, com $u \in \mathbb{R}^n$

Corolário 2.4.11. *Seja A uma matriz real $n \times n$ com colunas $c_1, c_2, \ldots, c_n$. Então A é ortogonal se e só se $c_i \perp c_j = 0$ se $i \neq j$, e $\|c_i\| = 1$, para $i, j = 1, \ldots, n$.*

Demonstração. Condição suficiente: Escrevendo $A = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix}$, temos que

$$I_n = A^T A = \begin{bmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_n^T \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}.$$

Como o elemento (i, j) de $\begin{bmatrix} c_1^T \\ c_2^T \\ \vdots \\ c_n^T \end{bmatrix} \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$ é $c_i^T c_j$, obtemos o resultado.

Condição necessária: Ora $c_i^T c_j = 0$ se $i \neq j$, e $c_i^T c_i = 1$ é o mesmo que $A^T A = I_n$, e pelo corolário anterior implica que A é invertível com $A^{-1} = A^T$, pelo que A é ortogonal. □

Ou seja, as colunas das matrizes ortogonais são ortogonais duas a duas. O mesmo se pode dizer acerca das linhas, já que a transposta de uma matriz ortogonal é de novo uma matriz ortogonal.

2.4.3 Teorema de Laplace

Dada uma matriz A, quadrada de ordem n, denota-se por $A(i|j)$ a submatriz de A obtida por remoção da sua linha i e da sua coluna j.

Definição 2.4.12. *Seja $A = [a_{ij}]$ uma matriz quadrada.*

1. *O* complemento algébrico *de a_{ij}, ou* cofactor *de a_{ij}, denotado por A_{ij}, está definido por*

$$A_{ij} = (-1)^{i+j} |A(i|j)|$$

2. *A* matriz adjunta *é a transposta da matriz dos complementos algébricos*

$$Adj(A) = [A_{ij}]^T.$$

Teorema 2.4.13 (Teorema de Laplace I)**.** *Para $A = [a_{ij}]$, $n \times n$, $n > 1$, então, e para $k = 1, \ldots, n$,*

$$\begin{aligned} |A| &= \sum_{j=1}^{n} a_{kj} A_{kj} \\ &= \sum_{j=1}^{n} a_{jk} A_{jk} \end{aligned}$$

O teorema anterior é o caso especial de um outro que enunciaremos de seguida. Para tal, é necessário introduzir mais notação e algumas definições (cf. [10]).

Seja A uma matriz $m \times n$. Um *menor de ordem p de A*, com $1 \leq p \leq \min\{m, n\}$, é o determinante de uma submatriz $p \times p$ de A, obtida de A eliminando $m - p$ linhas e $n - p$ colunas de A.

Considere duas sequências crescentes de números

$$1 \leq i_1 < i_2 < \cdots < i_p \leq m, \ 1 \leq j_1 < j_2 < \cdots < j_p \leq n,$$

e o determinante da submatriz de A constituída pelas linhas $i_1, i_2, \ldots i_p$ e pelas colunas $j_1, j_2, \ldots, j_p$. Este determinante vai ser denotado por $A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix}$. Ou seja,

$$A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix} = |\,[a_{i_k j_k}]_{k=1,\ldots p}\,|.$$

Paralelamente, podemos definir os *menores complementares* de A como os determinantes das submatrizes a que se retiraram linhas e colunas. Se A for $n \times n$,

$$A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix}^c$$

denota o determinante da submatriz de A após remoção das linhas $i_1, i_2, \ldots i_p$ e das colunas $j_1, j_2, \ldots, j_p$ de A. O *cofactor complementar* está definido como

$$A^c\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix} = (-1)^s A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix}^c,$$

onde $s = (i_1 + i_2 + \cdots i_p) + (j_1 + j_2 + \cdots j_p)$.

O caso em que $p = 1$ coincide com o exposto no início desta secção.

Teorema 2.4.14 (Teorema de Laplace II). *Sejam $A = [a_{ij}]$, $n \times n$, $1 \leq p \leq n$. Para qualquer escolha de p linhas $i_1, i_2, \ldots, i_p$ de A, ou de p colunas $j_1, j_2, \ldots, j_p$ de A,*

$$|A| \ = \ \sum_j A\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix} A^c\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix}$$

onde a soma percorre todos os menores referentes à escolha das linhas [resp. colunas].

Para finalizar, apresentamos um método de cálculo da inversa de uma matriz não singular.

Teorema 2.4.15. *Se A é invertível então*

$$A^{-1} = \frac{Adj(A)}{|A|}.$$

Exercícios

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

   (a) Troque duas linhas de A e compare o determinante da matriz obtida com $|A|$. Repita o exercício fazendo trocas de colunas.

   (b) Substitua uma linha/coluna de A pela linha nula e calcule o determinante da matriz obtida.

   (c) Multiplique uma linha por um escalar não nulo e compare o determinante da matriz obtida com $|A|$. Repita o exercício multiplicando uma coluna por um escalar não nulo.

   (d) A uma linha de A some-lhe outra multiplicada por um escalar não nulo. Compare o determinante da matriz obtida com $|A|$. Repita o exercício fazendo a operação elementar por colunas.

   (e) Encontre uma factorização $PA = LU$. Calcule $\det(U)$ e compare com $\det(A)$.

   (f) O que pode conjecturar sobre o valor de $|-A|$? Teste a validade da sua conjectura.

   (g) Verifique que $|A^T| = |A|$.

   (h) Calcule a matriz dos complementos algébricos, a adjunta e a inversa (caso exista) da matriz dada.

2. Considere a matriz $A = \begin{bmatrix} -1 & 2 & -5 \\ 0 & -6 & 6 \\ -4 & -1 & 0 \end{bmatrix}$.

   (a) Encontre uma factorização $PA = LU$.

   (b) Calcule $\mathrm{car}(A)$.

   (c) Calcule $\det(A)$.

   (d) Calcule a matriz dos complementos algébricos, a adjunta e a inversa (caso exista) da matriz dada.

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

   (a) Encontre uma factorização $PA = LU$.

   (b) Calcule $\mathrm{car}(A)$.

   (c) Calcule $\det(A)$.

   (d) Calcule a matriz dos complementos algébricos, a adjunta e a inversa (caso exista) da matriz dada.

4. Considere a matriz $A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 4 & 1 & 1 \end{bmatrix}$.

   (a) Encontre uma factorização $PA = LU$.

   (b) Calcule $\mathrm{car}(A)$.

   (c) Calcule $\det(A)$.

   (d) Calcule a matriz dos complementos algébricos, a adjunta e a inversa (caso exista) da matriz dada.

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

   (a) encontre uma factorização $PA = LU$,

   (b) calcule $\mathrm{car}(A)$,

   (c) calcule $\det(A)$.

   (d) calcule a matriz dos complementos algébricos, a adjunta e a inversa (caso exista).

6. Calcule o determinante, a matriz dos complementos algébricos, a adjunta e a inversa (caso exista) das matrizes

   (a) $\begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 2 \\ 7 & 1 & 1 \end{bmatrix}$.

   (b) $\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 1 \\ 2 & 3 & 3 \end{bmatrix}$.

   (c) $\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 5 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

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

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

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

7. Calcule o determinante das matrizes seguintes:

(a) $\begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}$ (d) $\begin{bmatrix} 6 & 9 \\ 8 & 12 \end{bmatrix}$

(e) $\begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$ (f) $\begin{bmatrix} n+1 & n \\ n & n-1 \end{bmatrix}$ (g) $\begin{bmatrix} a+b & a-b \\ a-b & a+b \end{bmatrix}$ (h) $\begin{bmatrix} 1 & \mathrm{i} \\ 1 & \imath \end{bmatrix}$

(i) $\begin{bmatrix} a & c+d\,\mathrm{i} \\ c-d\,\mathrm{i} & b \end{bmatrix}$ (j) $\begin{bmatrix} a+b\,\mathrm{i} & b \\ 2a & a-b\,\mathrm{i} \end{bmatrix}$ (k) $\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$ (l) $\begin{bmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{bmatrix}$

8. Se A é uma matriz simétrica, mostre que $\det(A+B) = \det\left(A+B^T\right)$, para qualquer matriz B com a mesma ordem de A.

9. Uma matriz A é anti-simétrica se $A^T = -A$. Mostre que, para $A \in \mathcal{M}_n(\mathbb{K})$ com n ímpar e A anti-simétrica, se tem $\det A = 0$.

Capítulo 3

Sistemas de equações lineares

3.1 Formulação matricial

Uma *equação linear* em n variáveis $x_1, \ldots, x_n$ sobre $\mathbb{K}$ é uma equação da forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

onde $a_1, a_2, \ldots, a_n, b \in \mathbb{K}$. Um sistema de equações lineares é um conjunto finito de equações lineares que é resolvido simultaneamente. Ou seja, que se pode escrever da forma

$$\left\{ \begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n & = & b_2 \\ & \cdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array} \right. \qquad (1)$$

Este tipo de sistema pode ser representado na forma matricial

$$Ax = b,$$

com

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

A é a *matriz do sistema*, x é a *coluna das incógnitas* e b é a *coluna dos termos independentes*, também denominado por segundo membro do sistema.

De ora em diante não faremos distinção entre o sistema de equações lineares e a sua formulação matricial $Ax = b$.

Neste capítulo, vamo-nos debruçar sobre a resolução deste tipo de equação. Dizemos que v é solução de $Ax = b$ se $Av = b$, ou seja, quando v é uma realização possível para a coluna das incógnitas. Iremos ver em que condições a equação tem solução, e como se podem determinar. Entende-se por *resolver o sistema* $Ax = b$ encontrar o conjunto (ainda que vazio) de todas as realizações possíveis para a coluna das incógnitas. O sistema diz-se *impossível* ou inconsistente se o conjunto é vazio e possível ou consistente caso contrário. Neste último caso, diz-se que é *possível determinado* se existir apenas um e um só elemento no conjunto das soluções, e *possível indeterminado* se for possível mas existirem pelo menos duas soluções distintas[1]. Entende-se por *classificar o sistema* a afirmação em como ele é impossível, possível determinada ou possível indeterminado.

Um caso particular da equação $Ax = b$ surge quando $b = 0$. Ou seja, quando a equação é da forma $Ax = 0$. O sistema associado a esta equação chama-se *sistema homogéneo*. Repare que este tipo de sistema é sempre possível. De facto, o vector nulo (ou seja, a coluna nula) é solução. Ao conjunto das soluções de $Ax = 0$ chamamos *núcleo*[2] de A, e é denotado por $N(A)$ ou ainda por $\ker(A)$. Ou seja, para A do tipo $m \times n$,

$$N(A) = \ker(A) = \{x \in \mathbb{K}^n : Ax = 0_{m\times 1}\}.$$

Pelo que acabámos de referir, e independentemente da escolha de A, o conjunto $\ker(A)$ é sempre não vazio já que $0_{n\times 1} \in \ker(A)$.

Outro caso relevante no estudo da equação $Ax = b$ surge quando a matriz A é invertível. Neste caso, multiplicando ambos os membros de $Ax = b$, à esquerda, por A^{-1}, obtemos $A^{-1}Ax = A^{-1}b$, e portanto $x = A^{-1}b$. Ou seja, a equação é possível determinada, sendo $A^{-1}b$ a sua única solução.

3.2 Resolução de $Ax = b$

Nesta secção, vamos apresentar uma forma de resolução da equação $Ax = b$, fazendo uso da factorização $PA = LU$ estudada atrás. Vejamos de que forma essa factorização é útil no estudo da equação.

Considere a equação $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$. O sistema associado escreve-se como $\begin{cases} x_1 + 2x_2 + 3x_3 = 7 \\ \quad\;\; 4x_2 + 5x_3 = 8 \\ \qquad\qquad 6x_3 = 9 \end{cases}$. Calculando o valor de x_3 pela última equação, este é substituido na segunda equação para se calcular o valor de x_2, que por sua vez são usados na primeira equação para se obter x_1. Procedeu-se à chamada *substituição inversa* para se calcular a única (repare que a matriz dada é invertível) solução do sistema. Em que condições se pode usar a substituição inversa? Naturalmente quando a matriz dada é triangular superior com elementos diagonais não nulos. Mas também noutros casos. Considere a equação matricial $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$. A matriz do sistema não é quadrada, mas o método da susbstituição inversa pode ainda ser aplicado. O sistema associado é

[1] Veremos, mais adiante, que se existem duas soluções distintas então existe uma infinidade delas.

[2] Iremos também usar a denominação *espaço nulo* de A.

$\begin{cases} x_1 + 2x_2 + 3x_3 = 5 \\ 4x_3 = 6 \end{cases}$, donde $x_3 = \frac{3}{2}$, e x_1 dependerá do valor de x_2. A solução geral do sistema é $(x_1, x_2, x_3) = (5 - \frac{9}{2} - 2x_2, x_2, \frac{3}{2}) = (5 - \frac{9}{2}, 0, \frac{3}{2}) + x_2(-2, 1, 0)$. Mais à frente veremos qual a importância de escrevermos a solução na última forma apresentada. É fácil constatar que a substituição inversa é aplicável desde que a matriz do sistema seja uma matriz escada de linhas. A estatégia na resolução da equação irá, portanto, passar pela matriz escada obtida por Gauss, para depois se aplicar a substituição inversa. Desde que o sistema seja possível, claro.

Considere o sistema $Ax = b$ e a factorização $PA = LU$. Ou seja, $U = L^{-1}PA$. Recorde que $L^{-1}P$ reflecte as operações elementares efectuadas nas linhas de A por forma a se obter a matriz escada, percorrendo os passos do AEG. Multiplique ambos os membros de $Ax = b$, à esquerda, por $L^{-1}P$ para obter $L^{-1}PA = L^{-1}Pb$. Como $U = L^{-1}PA$ tem-se que $Ux = L^{-1}Pb$, e daqui podemos aplicar a substituição inversa... depois de se determinar o termo independente $L^{-1}Pb$. Recorde que $L^{-1}P$ reflecte as operações elementares efectuadas nas linhas de A, de modo que para se obter $L^{-1}Pb$ basta efectuar essas mesmas operações elementares, pela mesma ordem, nas linhas de b. Por forma a simplificar o raciocínio e evitar possíveis enganos, esse processo pode ser efectuado *ao mesmo tempo* que aplicamos o AEG nas linhas de A. Consideramos, para esse efeito, a *matriz aumentada do sistema* $\left[\begin{array}{c|c} A & b \end{array}\right]$, aplicamos o AEG para se obter a matriz $\left[\begin{array}{c|c} U & c \end{array}\right]$, onde U é matriz escada de linhas e $c = L^{-1}Pb$. Se o sistema for possível, aplica-se a substituição inversa a $Ux = c$.

As soluções de $Ax = b$ são *exactamente* as mesmas de $Ux = c$, e por este facto dizem-se equações equivalentes, e os sistemas associados são equivalentes. De facto, se v é solução de $Ax = b$ então $Av = b$, o que implica, por multiplicação à esquerda por $L^{-1}P$ que $L^{-1}PAv = L^{-1}Pb$, ou seja, que $Uv = c$. Por outro lado, se $Uv = c$ então $LUv = Lc$ e portanto $PAv = Lc$. Ora $c = L^{-1}Pb$, e portanto $Lc = Pb$. Obtemos então $PAv = Pb$. Como P é invertível, segue que $Av = b$ e v é solução de $Ax = b$.

Visto determinar as soluções de $Ax = b$ é o mesmo que resolver $Ux = c$, interessa-nos, então classificar este último.

Como exemplo, considere a equação $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$. A segunda equação do sistema associado reflecte a igualdade $0 = 5$, o que é impossível. A equação é impossível já que não tem soluções. A matriz aumentada associada à equação é $\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 5 \end{array}\right]$. Repare que a característica da matriz A é 1 enquanto que a característica da matriz aumentada $[A|b]$ é 2.

Como é fácil verificar, a característica da matriz a que se acrescentou linhas ou colunas é não inferior à característica da matriz inicial. Por consequência, $\operatorname{car}(A) \leq \operatorname{car}([A|b])$.

Teorema 3.2.1. *A equação matricial $Ax = b$ é consistente se e só se* $\operatorname{car}(A) = \operatorname{car}\left(\left[\begin{array}{c|c} A & b \end{array}\right]\right)$.

Demonstração. Considere $PA = LU$ e $c = L^{-1}Pb$. A equação $Ax = b$ é equivalente à equação $Ux = c$, e portanto $Ax = b$ tem solução se e só se $Ux = c$ tem solução. Tal equivale a dizer que o número de linhas nulas de U iguala o número de linhas nulas de $[U|c]$. De facto, o número sendo o mesmo, por substituição inversa é possível obter uma solução de $Ux = c$, e caso o número seja distinto então obtemos no sistema associado a igualdade $0 = c_i$, para algum $c_i \neq 0$, o que torna $Ux = c$ impossível. Se o número de linhas nulas de U iguala o de $[U|c]$ então o número de linhas não nulas de U iguala o de $[U|c]$. $\square$

Como exemplo, considere a equação matricial $Ax = b$ onde $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & \frac{1}{2} \end{bmatrix}$ e $b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. A equação é consistente se e só se $\mathrm{car}(A) = \mathrm{car}([A|b])$. Ora $A = LU$ com $L = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}$ e $U = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, e portanto $\mathrm{car}(A) = 1$. A matriz escada obtida da matriz aumentada $\left[\begin{array}{c|c} A & b \end{array}\right]$ é $\left[\begin{array}{ccc|c} 2 & 2 & 1 & -1 \\ 0 & 0 & 0 & \frac{3}{2} \end{array}\right]$ Ora a caraterística da matriz aumentada é 2, pelo que $Ax = b$ é inconsistente.

Dada a equação $A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = b$, considere $U \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = c$ equivalente à primeira fazendo uso da factorização $PA = LU$ da forma habitual. A incógnita x_i diz-se *incógnita básica* se a **coluna** i de U tem pivot. Uma incógnita diz-se *livre* se não for básica. A *nulidade* de A, $\mathrm{nul}(A)$, é o número de incógnitas livres na resolução de $Ax = 0$.

Na equação $Ax = b$, com $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & -1 \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, obtemos a decomposição $A = LU$, com

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

e

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

Repare que $\mathrm{car}(A) = 2$. Ora $2 = \mathrm{car}(A) \leq \mathrm{car}([A|b]) \leq 2$, já que a característica de uma matriz é não superior ao seu número de linhas **e** ao seu número de colunas. Segue que $\mathrm{car}([A|b]) = 2$. A equação $Ax = b$ é, portanto, consistente. Façamos, então, a classificação das incógnitas x_1, x_2, x_3 em livres e em básicas. Atente-se à matriz escada de linhas U apresentada atrás. As colunas 1 e 3 têm como pivots, respectivamente, 2 e $-\frac{3}{2}$. As incógnitas x_1 e x_3 são *básicas*. Já x_2 é *livre* pois a coluna 2 de U não tem pivot.

Qual o interesse neste tipo de classificação das incógnitas? A explicação é feita à custa do exemplo anterior. A equação $Ax = b$ é equivalente à equação $Ux = c$, com $U =$

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

Com os dados fornecidos, a matriz escada seguindo os passos do AEG do exemplo iguala

$$\left[\begin{array}{c|c} U & c \end{array}\right] = \left[\begin{array}{ccc|c} 2 & 2 & 1 & -1 \\ 0 & 0 & -\frac{3}{2} & \frac{3}{2} \end{array}\right].$$

Podemos, agora, aplicar o método da substituição inversa para obter as soluções da equação. Esse método é aplicado da seguinte forma:

1. obtem-se o valor das **incógnitas básicas** x_i no sentido sul$\rightarrow$norte,

2. as **incógnitas livres** comportam-se como se de termos independentes se tratassem.

Para conveniência futura, a solução é apresentada na forma

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} ? \\ ? \\ \vdots \\ ? \end{bmatrix} + x_{i_1} \begin{bmatrix} ? \\ ? \\ \vdots \\ ? \end{bmatrix} + x_{i_2} \begin{bmatrix} ? \\ ? \\ \vdots \\ ? \end{bmatrix} + \dots x_{i_k} \begin{bmatrix} ? \\ ? \\ \vdots \\ ? \end{bmatrix}$$

onde x_{i_ℓ} são as incógnitas livres.

Voltando ao exemplo, recorde que se obteve a equação equivalente à dada

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

Resolvendo a última equação correspondente, obtemos o valor da incógnita básica x_3. De facto, $-\frac{3}{2}x_3 = \frac{3}{2}$ implica $x_3 = -1$. Na equação $2x_1 + 2x_2 + x_3 = -1$, o valor de x_3 é conhecido (bastando-nos, portanto, fazer a substituição) e a incógnita x_2 é livre, comportando-se então como termo independente. Já x_1 é básica, e resolve-se a equação em relação a esta. Obtemos $x_1 = \frac{-2x_2}{2} = -x_2$. Para cada escolha de x_2 obtemos outo valor para x_1. A solução geral é da forma

$$(x_1, x_2, x_3) = (-x_2, x_2, -1) = (0, 0, -1) + x_2(-1, 1, 0),$$

onde x_2 varia livremente em $\mathbb{K}$.

Num sistema possível, a existência de incógnitas livres confere-lhe a existência de várias soluções, e portanto o sistema é possível indeterminado. Ora, se o número de incógnitas é n e se k delas são básicas, então as restantes $n - k$ são livres. Recorde que o número de incógnitas iguala o número de colunas da matriz do sistema, e que a característica de uma matriz é igual ao número de pivots. Existindo, no máximo, um pivot por coluna, e como o número das colunas com pivots é igual ao número de incógnitas básicas, segue que a característica da matriz é igual ao número de incógnitas básicas. A existência de incógnitas livres é equivalente ao facto de existirem colunas sem pivot, ou seja, do número de colunas ser estritamente maior que a característica da matriz. De facto, as incógnitas livres são, em número, igual ao número de colunas sem pivot.

Teorema 3.2.2. *A equação consistente $Ax = b$, onde A é $m \times n$, tem uma única solução se e só se* $\mathrm{car}(A) = n$.

Corolário 3.2.3. *Um sistema possível de equações lineares com menos equações que incógnitas é indeterminado.*

Recorde que o número de incógnitas livres é o número de colunas sem pivot na resolução de um sistema possível $Ax = b$. Por outro lado, a nulidade de A, $\mathrm{nul}(A)$, é o número de incógnitas livres que surgem na resolução de $Ax = 0$. Recorde ainda que a característica de A, $\mathrm{car}(A)$, é o número de pivots na implementação de Gauss, que por sua vez é o número de colunas com pivot, que iguala o número de incógnitas básicas na equação $Ax = 0$. Como o número de colunas de uma matriz iguala o número de incógnitas equação $Ax = 0$, e estas se dividem em básicas e em livres, correspondendo em número a, respectivamente, $\mathrm{car}(A)$ e $\mathrm{nul}(A)$, temos o resultado seguinte:

Teorema 3.2.4. *Para A matriz $m \times n$,*

$$n = \mathrm{car}(A) + \mathrm{nul}(A).$$

O resultado seguinte descreve as soluções de uma equação possível $Ax = b$ à custa do sistema homogéneo associado (ou seja, $Ax = 0$) e de uma solução particular v de $Ax = b$.

Teorema 3.2.5. *Sejam $Ax = b$ uma equação consistente e v uma solução particular de $Ax = b$. Então w é solução de $Ax = b$ se e só se existir $u \in N(A)$ tal que $w = v + u$.*

Demonstração. Suponha v, w soluções de $Ax = b$. Pretende-se mostrar que $w - v \in N(A)$, ou seja, que $A(w - v) = 0$. Ora $A(w - v) = Aw - Av = b - b = 0$. Basta, portanto, tomar $u = w - v$.

Reciprocamente, assuma v solução de $Ax = b$ e u solução de $Ax = 0$. Pretende-se mostrar que $w = v + u$ é solução de $Ax = b$. Para tal, $Aw = A(v + u) = Av + Au = b + 0 = b$. $\square$

Ou seja, conhecendo o conjunto das soluções de $Ax = 0$ e uma solução particular de $Ax = b$, conhece-se o conjunto das soluções de $Ax = b$.

Exemplo. Considere a equação matricial $Ax = b$, com $A = \begin{bmatrix} -12 & -1 & -8 \\ 6 & -5 & 0 \\ 9 & -2 & 4 \end{bmatrix}$ e $b = \begin{bmatrix} -1 \\ 5 \\ 3 \end{bmatrix}$. O sistema é consistente, já que $\mathrm{car}(A) = \mathrm{car}([A|b]) = 2$. De facto, a matriz escada de linhas $[U|c]$ obtida de $[A|b]$ após aplicação do AEG é $\left[\begin{array}{ccc|c} -12 & -1 & -8 & -1 \\ 0 & -\frac{11}{2} & -4 & \frac{9}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$. Sendo a característica de A igual a 2 e tendo a matriz 3 colunas, então existe uma, e uma só, incógnita livre na resolução de $Ax = b$. Façamos, então, a divisão das incógnitas em livres e básicas. Observando as colunas da matriz escada de linhas U, se $x = (x_1, x_2, x_3)$, as incógnitas básicas são x_1 e x_2, enquanto que x_3 é incógnita livre, já que a única coluna de U que não tem pivot é a terceira.

Como vimos do resultado anterior, conhecendo uma solução particular de $Ax = b$, digamos, v, e conhecendo $N(A)$, ou seja, o conjunto das soluções de $Ax = 0$, então as soluções de $Ax = b$ são da forma $v + u$, onde $u \in N(A)$. Uma solução particular pode ser encontrada tomando a incógnita livre como zero. Ou seja, considerando $x_3 = 0$. A substituição inversa fornece o valor das incógnitas básicas x_1, x_2. Obtemos $x_2 = \frac{\frac{9}{2}}{-\frac{11}{2}} = -\frac{9}{11}$ e $x_1 = \frac{-1-(-1)x_2}{-12} = \frac{1}{66}$.

Resta-nos determinar $N(A)$, ou seja, resolver o sistema homogéneo $Ax = 0$. Para tal, podemos fazer uso da matriz escada U encontrada atrás, e resolvemos o sistema $Ux = 0$ em relação às incógnitas básicas x_1, x_2, tratando a incógnita x_3 como se de um termo independente se tratasse. As soluções serão da forma $x = x_3 u$.

Considere, agora, a matriz $A = \begin{bmatrix} 2 & 2 & 2 & 0 \\ 1 & 1 & 2 & 2 \end{bmatrix}$. Esta matriz tem característica 2, e a matriz escada de linhas obtida de A após aplicação do AEG é $U = \begin{bmatrix} 2 & 2 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$. A nulidade de A é 2, já que existem 2 incógnitas livres na resolução de $A \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T = 0_{2\times 1}$. As incógnitas livres são as correspondentes às colunas de U que não têm pivot; no caso, x_2 e x_4. O sistema associado à equação $Ux = 0$ é $\begin{cases} 2x_1 + 2x_2 + 2x_3 = 0 \\ x_3 + 2x_4 = 0 \end{cases}$. Resolvendo o sistema em relação às incógnitas básicas x_1, x_3, e por substituição inversa, obtemos $x_3 = -2x_4$, que por sua vez fornece, substituindo na primeira equação, $x_1 = \frac{1}{2}(-2x_2 + 4x_4) = -x_2 + 2x_4$. Ou seja, a solução geral de $Ax = 0$ é

$$(x_1, x_2, x_3, x_4) = (-x_2 + 2x_4, x_2, -2x_4, x_4) = x_2(-1, 1, 0, 0) + x_4(2, 0, -2, 1).$$

3.3 Algoritmo de Gauss-Jordan

A aplicação do Algoritmo de Eliminação de Gauss na resolução de um sistema de equações lineares reduz o problema inicial a um outro, equivalente ao dado (ou seja, com o mesmo conjunto de soluções) onde a matriz associada ao sistema é escada de linhas. Tal permite a utilização do algoritmo da substituição inversa por forma a se encontrar (caso existam) as soluções para o problema. Nesse método, o valor das incógnitas básicas era encontrado à custa das incógnitas livres e dos termos independentes, bem como do valor das incógnitas básicas encontrado no passo anterior, no sentido sul$\to$norte do vector das incógnitas. Recorde que no AEG a estratégia tinha como objectivo, por operações elementares de linhas, obter zeros por debaixo de cada pivot, estratégia essa implementada no sentido NW$\to$SE. Este raciocínio pode ser estendido a obterem-se zeros por cima dos pivots, no sentido SW$\to$NE, por operações elementares de linhas. De facto, neste processo estão ausentes trocas de linhas, já que os pivots usados neste novo processo são que estiveram envolvidos na fase inicial correspondente ao AEG. O resultado final será uma matriz constituída pelos pivots, tendo estes zeros por debaixo e por cima. Ou seja, se se dividir cada linha não nula pelo pivot respectivo, obtemos uma matriz da forma, a menos de trocas de colunas, uma matriz da forma $\left[\begin{array}{c|c} I_k & M \\ \hline 0 & 0 \end{array}\right]$, podendo os blocos nulos não existir. A este método (à excepção da possível troca de colunas) é denominado o algoritmo de Gauss-Jordan, e a matriz obtida diz-se que está na *forma canónica (reduzida) de linhas*, ou na *forma normal (ou canónica) de Hermite*. Ou seja, a matriz $H = [h_{ij}]$, $m \times n$, obtida satisfaz:

1. H é triangular superior,
2. h_{ii} é ou 0 ou 1,
3. se $h_{ii} = 0$ então $h_{ik} = 0$, para cada k tal que $1 \leq k \leq n$,
4. se $h_{ii} = 1$ então $h_{ki} = 0$ para cada $k \neq i$.

Repare que só são realizadas operações elementares nas linhas da matriz. A aplicação deste método na resolução de uma sistema de equações lineares permite obter, de forma imediata, o valor das incógnitas básicas. Apesar deste método parecer mais atractivo que o de Gauss (ou suas variantes), em geral é menos eficiente do ponto de vista computacional.

Exemplo. Considere a equação $Ax = b$, com $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$ e $b = \begin{bmatrix} 3 \\ -1 \\ -3 \end{bmatrix}$. A matriz aumentada correspondente ao sistema de equações lineares é $[A|b] = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 2 & 0 & 1 & -1 \\ 0 & 0 & -1 & -3 \end{array}\right]$ e a forma normal de Hermite de $[A|b]$ é $\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$. A forma canónica apresentada atrás fornece-nos uma solução para o sistema, no caso $(-2, -2, 3)$.

No que se segue, mostramos como se aplica o algoritmo de Gauss-Jordan para inverter matrizes.

Seja A uma matriz $n \times n$, não-singular. Ou seja, invertível. De forma equivalente, existe uma única matriz X tal que $AX = I_n$. Denotemos a matriz X, que pretendemos obter, à custa das suas colunas: $X = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}$. Pela forma como está definido o produto matricial, e tomando e_i como a i-ésima coluna da matriz I_n, a igualdade $AX = I_n$ pode-se escrever como $\begin{bmatrix} AX_1 & AX_2 & \cdots & AX_n \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix}$. Surgem-nos, então, n equações matriciais:

$$AX_1 = e_1,\ AX_2 = e_2,\ \ldots,\ AX_n = e_n.$$

Como A é invertível, cada uma destas equações é consistente e tem apenas uma solução. A solução de $AX_j = e_j$ é a coluna j da matriz X inversa de A que pretendemos calcular. Poder-se-ia aplicar a estratégia de Gauss a cada uma destas equações, ou seja, à matriz aumentada $\left[\begin{array}{c|c} A & e_j \end{array}\right]$. Como a matriz do sistema é a mesma, as operações elementares envolvidas seriam as mesmas para as n equações. Essas operações elementares podem ser efectuadas simultaneamente, considerando a matriz aumentada $n \times 2n$

$$\left[\begin{array}{c|cccc} & 1 & 0 & \cdots & 0 \\ & 0 & 1 & & \vdots \\ A & & \cdots & & \\ & 0 & \cdots & 0 & 1 \end{array}\right].$$

Sendo a matriz invertível, a matriz escada de linhas U obtida de A por aplicação do AEG tem elementos diagonais não nulos, que são os pivots que surgem na implementação do algoritmo. Aplicando Gauss-Jordan (ou seja, no sentido SE→NW, criando zeros por cima dos pivots que se vão considerando), obtemos uma matriz da forma $\left[\begin{array}{cccc|cccc} d_1 & 0 & \dots & 0 & & & & \\ 0 & d_2 & & 0 & & & & \\ \vdots & & & \vdots & Y_1 & Y_2 & \cdots & Y_n \\ 0 & & 0 & d_n & & & & \end{array}\right]$.

Dividindo a linha i por d_i, para $i = 1, ..., n$, obtém-se a matriz

$$\left[\begin{array}{c|cccc} I_n & \tilde{X}_1 & \tilde{X}_2 & \cdots & \tilde{X}_n \end{array}\right].$$

Ora tal significa que $\tilde{X}_i$ é a solução de $AX = e_i$. Ou seja, o segundo bloco da matriz aumentada indicada atrás não é mais que a inversa da matriz A. Isto é, Gauss-Jordan forneceu a matriz $\left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$.

3.4 Regra de Cramer

A regra de Cramer fornece-nos um processo de cálculo da solução de uma equação consistente $Ax = b$ quando A é invertível, e portanto a solução é única.

Dada a equação $Ax = b$, onde A é $n \times n$ não-singular, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ e b é do tipo $n \times 1$, denote-se por $A^{(i)}$ a matriz obtida de A substituindo a coluna i de A pela coluna b.

Teorema 3.4.1 (Regra de Cramer). *Nas condições do parágrafo anterior, a única solução de $Ax = b$ é dada por*

$$x_i = \frac{|A^{(i)}|}{|A|}.$$

Exemplo. Para aplicar a regra de Cramer, considere-se $A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 1 & 0 \\ -2 & 1 & -2 \end{bmatrix}$ e $b = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$. Como $|A| = -3 \neq 0$, a matriz A é invertível, e portanto $Ax = b$ é uma equação consistente com uma única solução. Definamos as matrizes $A^{(1)}, A^{(2)}, A^{(3)}$ como no teorema anterior. Alguns cálculos mostram que $\det(A^{(1)}) = -4$, $\det(A^{(2)}) = 1$ e $\det(A^{(3)}) = 3$. Aplicamos, de seguida, a regra de Cramer para obtermos $x_1 = \frac{4}{3}$, $x_2 = -\frac{1}{3}$ e $x_3 = -1$.

Exercícios

1. Resolva os seguintes sistemas de equações lineares:

   (a) $\begin{cases} 3x_1 - x_2 + x_3 = -1 \\ 9x_1 - 2x_2 + x_3 = -9 \\ 3x_1 + x_2 - 2x_3 = -9 \end{cases}$

   (b) $\begin{cases} 2x_1 + x_2 + 2x_3 + x_4 = 5 \\ 4x_1 + 3x_2 + 7x_3 + 3x_4 = 8 \\ -8x_1 - x_2 - x_3 + 3x_4 = 4 \\ 6x_1 + x_2 + 2x_3 - x_4 = 1 \end{cases}$

   (c) $\begin{cases} 2x - 3y = -1 \\ -4x - 6y = -2 \\ 12x - 18y = -6 \end{cases}$

   (d) $\begin{cases} -x_1 - x_2 + 2x_3 = -5 \\ -3x_1 - x_2 + 7x_3 = -22 \\ x_1 - 3x_2 - x_3 = 10 \end{cases}$

   (e) $\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_3 = 2 \end{cases}$

   (f) $\begin{cases} x_1 + x_2 + x_3 = 2 \\ 2x_1 + x_2 + x_3 = 3 \\ 3x_1 + x_2 + x_3 = 4 \end{cases}$

2. Determine $\beta \in \mathbb{R}$ de modo que o sistema

$$\begin{cases} \beta x - y + \beta z = 0 \\ -2\beta y - 2z = 0 \\ x - y + \beta z = 0 \end{cases}$$

   seja determinado.

3. Considere a matriz $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \end{bmatrix}$.

   (a) Encontre uma factorização $PA = LU$, onde P é matriz permutação, L é invertível triangular inferior e U é escada de linhas.

   (b) Resolva a equação $Ax = b$, onde $b = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

4. Determine todas as matrizes $X \in \mathcal{M}_{3\times 4}(\mathbb{R})$ tais que $AX = 0$, com
$$A = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 5 & -6 \\ 2 & -3 & 6 \end{bmatrix}.$$

5. Encontre os valores do parâmetro $k \in \mathbb{R}$ para os quais o sistema
$$\begin{cases} x - 2y + 3z & = & 1 \\ 2x + ky + 6z & = & 6 \\ -x + 3y + (k-3)z & = & 0 \end{cases}$$
seja
   (a) Possível determinado;
   (b) Possível indeterminado;
   (c) Impossível.

6. Considere a matriz $A = \begin{bmatrix} 0 & 0 & -1 & -1 \\ -1 & 1 & 0 & 1 \\ 2 & 2 & -1 & -3 \end{bmatrix}$. Resolva, usando o algoritmo de eliminação de Gauss, a equação matricial $Ax = (-2, -1, 0)$.

7. Considere a matriz real
$$A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 2 & 0 \\ 2 & 3 & -2 \end{bmatrix}.$$
   (a) Mostre que a matriz A é invertível e calcule a sua inversa fazendo uso do algoritmo de Gauss-Jordan.
   (b) Recorrendo à regra de Cramer, resolva
$$Ax = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

8. Resolva $AX = B$, com
$$A = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}.$$

9. Para $U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, mostre que U é invertível e calcule U^{-1} pelo algoritmo de Gauss-Jordan.

10. Seja $B = \begin{bmatrix} -1 & 0 & 0 \\ -2 & -1 & 0 \\ 5 & 2 & 1 \end{bmatrix}$.

(a) Mostre que B é invertível e faça uso do algoritmo de Gauss-Jordan para calcular a inversa de B.

(b) Use a regra de Cramer para determinar a única solução de $Bx = (0, 1, 0)$.

11. Para $U = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 0 \\ 2 & 1 & 4 \end{bmatrix}$, mostre que U é invertível e calcule U^{-1} pelo algoritmo de Gauss-Jordan.

12. Usando o algoritmo de Gauss-Jordan, calcule, se possível, a inversa de

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

Capítulo 4

Os espaços vectoriais $\mathbb{K}^n$

Neste capítulo estudaremos um caso muito particular de uma importante classe de estruturas algébricas, denominada por espaços vectoriais.

Tal como nos resultados apresentados anteriormente, $\mathbb{K}$ denota $\mathbb{R}$ ou $\mathbb{C}$.

4.1 Definição e exemplos

Considere-se o conjunto $\mathbb{K}^n = \{(x_1, \ldots, x_n) : x_1, \ldots, x_n \in \mathbb{K}\}$. Faça-se, ainda, a identificação dos elementos de $\mathbb{K}^n$ com o conjunto das matrizes coluna $n \times 1$ com entradas em $\mathbb{K}$. Os elementos de $\mathbb{K}^n$ serão denominados como vectores de $\mathbb{K}^n$.

Relembrando as propriedades do produto escalar e da soma de matrizes, e particularizando-as para o caso das matrizes coluna $n \times 1$, obtêm-se as seguintes propriedades:

1. Fecho da adição: $\forall_{x,y \in \mathbb{K}^n}, x + y \in \mathbb{K}^n$;
2. Fecho da multiplicação por escalares: $\forall_{x \in \mathbb{K}^n}, \alpha \in \mathbb{K}, \alpha x \in \mathbb{K}^n$;
3. Comutatividade da adição: $\forall_{x,y \in \mathbb{K}^n}, x + y = y + x$;
4. Associatividade da adição: $\forall_{x,y,z \in \mathbb{K}^n}, x + (y + z) = (x + y) + z$;
5. Existência de zero: existe um elemento de $\mathbb{K}^n$ (a matriz 0 de tipo $n \times 1$), tal que $x + 0 = x$, para $x \in \mathbb{K}^n$;
6. Existência de simétricos: $\forall_{x \in \mathbb{K}^n}, x + (-1)\, x = 0$;
7. Associatividade da multiplicação por escalares: $\forall_{\alpha,\beta \in \mathbb{K}, x \in \mathbb{K}^n}, \alpha\,(\beta x) = (\alpha\beta)\, x$;
8. Distributividade: $\alpha\,(x + y) = \alpha x + \alpha y$ e $(\alpha + \beta)\, x = \alpha x + \beta x$, para $x, y \in \mathbb{K}^n$ e $\alpha, \beta \in \mathbb{K}$;
9. Existência de identidade: $1x = x$, para todo $x \in \mathbb{K}^n$.

Definição 4.1.1. *Seja $W \subseteq \mathbb{K}^n$. Então W é um* subespaço *de $\mathbb{K}^n$ se as condições seguintes forem satisfeitas:*

1. $W \neq \emptyset$;

2. $u, v \in W \Rightarrow u + v \in W$;

3. $v \in W, \alpha \in \mathbb{K} \Rightarrow \alpha v \in W$.

Observe que se W é subespaço de $\mathbb{K}^n$ então **necessariamente** $(0, 0, \ldots, 0) \in W$.

Como exemplo, considere o subconjunto S de $\mathbb{R}^2$ dado por $S = \{(x, y) : y = 2x\}$. O conjunto é obviamente não vazio, já que $(0, 0) \in S$. Tomemos, agora, e de forma arbitrária, dois elementos, u e v, se S. Então $u = (x_1, 2x_1)$ e $v = (x_2, 2x_2)$, para algum $x_1, x_2 \in \mathbb{R}$, donde $u + v = (x_1 + x_2, 2x_1 + 2x_2) = (x_1 + x_2, 2(x_1 + x_2) \in S$. Finalmente, e para $\alpha \in \mathbb{R}$, tem-se $\alpha u = (\alpha x_1, 2\alpha x_1) \in S$.

No entanto, $T = \left\{(x, x^3) \in \mathbb{R}^2 : x \in \mathbb{R}\right\}$ não é subespaço de $\mathbb{R}^2$, isto apesar de $(0, 0) \in T$. De facto, e apesar $(1, -1), (2, 8) \in T$, a sua soma não é elemento de S.

De ora em diante, sempre que nos referirmos a um *espaço vectorial* V este é um subespaço de $\mathbb{K}^n$, dizendo-se que V é um espaço vectorial real ou complexo, consoante $\mathbb{K}$ é $\mathbb{R}$ ou $\mathbb{C}$. Note-se que, em particular, $\mathbb{K}^n$ é um espaço vectorial.

Dizemos que um subconjunto não vazio W de um espaço vectorial V é subespaço vectorial de V se para quaisquer escolhas de $u, v \in W$ e de $\alpha \in \mathbb{K}$ se tem $u + v \in W$ e $\alpha v \in W$. Mostra-se facilmente que um subespaço vectorial é também um subespaço de $\mathbb{K}^n$. Não faremos, por consequência, distinção entre subespaço e subespaço vectorial.

Exercícios

Diga quais dos conjuntos seguintes são subespaços vectoriais do espaço vectorial real $\mathbb{R}^4$:

1. $W_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 - x_2 = 0 \text{ e } x_2 + 2x_4 = 0\}$.

2. $W_2 = \{(0, a, b, -1) : a, b \in \mathbb{R}\}$.

3. $W_3 = \{(0, 0, 0, 0), (0, 0, 0, 1)\}$.

4. $W_4 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_2 \in \mathbb{Q}\}$.

4.2 Independência linear

Sejam V um espaço vectorial e $\{v_i\}_{i \in I} \subseteq V, \{\alpha_i\}_{i \in I} \subseteq \mathbb{K}$. Se

$$v = \sum_{i=1}^{n} \alpha_i v_i,$$

diz-se que v é uma *combinação linear* dos vectores $v_1, \ldots, v_n$. Neste caso, dizemos que v se pode escrever como *combinação linear* de $v_1, \ldots, v_n$.

Definição 4.2.1 (Conjunto linearmente independente)**.** *Um conjunto não vazio* $\{v_i\}_{i\in I} \subseteq V$ *diz-se* linearmente independente *se*

$$\sum_{i\in I} \alpha_i v_i = 0 \Longrightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Um conjunto diz-se linearmente dependente se não for linearmente independente.

Por abuso de linguagem, tomaremos, em algumas ocasiões, vectores linearmente independentes para significar que o conjunto formado por esses vectores é linearmente independente.

O conceito de dependência e independência linear é usualmente usado de duas formas.

(i) Dado um conjunto não vazio $\{v_i\}$ de n vectores linearmente dependentes, então é possível escrever o vector nulo como combinação linear não trivial de $v_1, \ldots, v_n$. Ou seja, existem escalares $\alpha_1, \ldots, \alpha_n$, algum ou alguns dos quais não nulos, tais que

$$0 = \sum_{i=n}^{n} \alpha_i v_i.$$

Seja α_k um coeficiente não nulo dessa combinação linear. Então

$$v_k = \sum_{i=1, i\neq k}^{n} \left(-\alpha_k^{-1}\alpha_i\right) v_i.$$

Concluindo, dado um conjunto de vectores linearmente dependentes, então pelo menos um desses vectores é uma combinação linear (não trivial) dos outros vectores.

(ii) Dado um conjunto não vazio $\{v_i\}$ de n vectores linearmente independentes, da relação

$$0 = \sum_{i=n}^{n} \alpha_i v_i$$

podemos concluir de forma imediata e óbvia que $\alpha_1 = \cdots = \alpha_n = 0$. Esta implicação será muito útil ao longo desta disciplina.

Como exemplo, consideremos em $\mathbb{R}^3$ os vectores $\alpha = (1,1,0), \beta = (1,0,1), \gamma = (0,1,1), \delta = (1,1,1)$. Estes quatro vectores são linearmente dependentes (pois $\alpha + \beta + \gamma - 2\delta = 0$), apesar de quaisquer três deles serem linearmente independentes.

Teorema 4.2.2. *Sejam* $v_1, \ldots, v_n$ *elementos linearmente independentes de um espaço vectorial* V*. Sejam ainda* $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{K}$ *tais que*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Então $\alpha_i = \beta_i$*, para todo* $i = 1, \ldots, n$*.*

Demonstração. Se $\alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_n v_n$ então

$$(\alpha_1 - \beta_1)\, v_1 + \cdots + (\alpha_n - \beta_n)\, v_n = 0,$$

pelo que, usando o facto de $v_1, \ldots, v_n$ serem linearmente independentes, se tem $\alpha_i - \beta_i = 0$, para todo $i = 1, \ldots, n$. □

O resultado anterior mostra a *unicidade* da escrita de um vector como combinação linear de elementos de um conjunto linearmente independente, caso essa combinação linear exista.

Teorema 4.2.3. *Seja A um subconjunto não vazio de um espaço vectorial V. Então o conjunto de todas as combinações lineares de elementos de A é um subespaço vectorial de V.*

Demonstração. Seja A' o conjunto de todas as combinações lineares de elementos de A. A' é obviamente não vazio visto $A \neq \emptyset$. Sejam $u, v \in A'$. Ou seja,

$$u = \sum_{i \in I} \alpha_i a_i, \;\; v = \sum_{j \in J} \beta_j a_j,$$

para alguns $\alpha_i, \beta_j \in \mathbb{K}$, com $a_i \in A$. Note-se que

$$u + v = \sum_{i \in I} \alpha_i a_i + \sum_{j \in J} \beta_j a_j$$

e portanto $u + v$ é assim uma combinação linear de elementos de A – logo, $u + v \in A'$. Para $\kappa \in \mathbb{K}$, temos que $\kappa u = \sum_{i \in I}^{n} \kappa \alpha_i a_i$ e portanto $\kappa u \in A'$. □

Tendo em conta o teorema anterior, podemos designar o conjunto das combinações lineares dos elementos de A como o *espaço gerado* por A. Este espaço vectorial (subespaço de V) denota-se por $\langle A \rangle$.

Quando o conjunto A está apresentado em extensão, então não escrevemos as chavetas ao denotarmos o espaço gerado por esse conjunto. Por exemplo, se $A = \{v_1, v_2, v_3\}$, então $\langle A \rangle$ pode-se escrever como $\langle v_1, v_2, v_3 \rangle$. Por notação, $\langle \emptyset \rangle = \{0\}$.

É importante referir os resultados que se seguem, onde V indica um espaço vectorial.

1. Os vectores não nulos $v_1, v_2, \ldots, v_n \in V$ são linearmente independentes se e só se, para cada k, $v_k \notin \langle v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n \rangle$.

2. Sejam $A, B \subseteq V$.

   (a) Se $A \subseteq B$ então $\langle A \rangle \subseteq \langle B \rangle$.

   (b) $\langle A \rangle = \langle \langle A \rangle \rangle$.

   (c) $\langle A \rangle$ é o menor (para a relação de ordem $\subseteq$) subespaço de V que contém A.

4.3 Bases de espaços vectoriais

Definição 4.3.1. *Seja V um espaço vectorial.*

Um conjunto $\mathcal{B}$ linearmente independente tal que $\langle \mathcal{B} \rangle = V$ é chamado de base *de V.*

Teorema 4.3.2. *Todo o espaço vectorial tem uma base.*

De ora em diante, apenas consideraremos espaços vectoriais finitamente gerados. Por vezes faremos referência à base $v_1, v_2, \ldots, v_n$ para indicar que estamos a considerar a base $\{v_1, v_2, \ldots, v_n\}$.

Definição 4.3.3. *Uma base ordenada $\mathcal{B} = \{v_1, \ldots, v_m\}$ é uma base de V cujos elementos estão dispostos por uma ordem fixa*[1]. *Chamam-se* componentes *ou* coordenadas *de $u \in V$ na base $\{v_1, \ldots, v_m\}$ aos coeficientes escalares $\alpha_1, \ldots, \alpha_m$ da combinação linear*

$$u = \sum_{k=1}^{m} \alpha_k v_k.$$

As coordenadas de u na base $\mathcal{B}$ são denotadaspor

$$(u)_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Recordemos que, se $\mathcal{B} = \{v_1, \ldots, v_m\}$ é uma base de V, em particular são linearmente independentes, e portanto dado $v \in V$, os coeficientes de v na base $\mathcal{B}$ são únicos.

Teorema 4.3.4. *Se um espaço vectorial tem uma base com um número finito n de elementos, então todas as bases de V têm n elementos.*

Demonstração. Seja V um espaço vectorial e $v_1, \ldots, v_n$ uma base de V. Seja $w_1, \ldots, w_m$ outra base de V com m elementos.

Como $v_1, \ldots, v_n$ é base de V, existem $\alpha_{ji} \in \mathbb{K}$ para os quais

$$w_i = \sum_{j=1}^{n} \alpha_{ji} v_j.$$

[1] De uma forma mais correcta, $\mathcal{B}$ não deveria ser apresentado como conjunto, mas sim como um n-uplo: $(v_1, \ldots, v_m)$. Comete-se assim um abuso de notação, tendo em mente que a notação escolhida indica a ordem dos elementos da base pelos quais foram apresentados.

Note-se que

$$\begin{aligned}
\sum_{i=1}^{m} x_i w_i = 0 \quad &\Leftrightarrow \quad \sum_{i=1}^{m} x_i \sum_{j=1}^{n} \alpha_{ji} v_j = 0 \\
&\Leftrightarrow \quad \sum_{i=1}^{m} \sum_{j=1}^{n} x_i \alpha_{ji} v_j = 0 \\
&\Leftrightarrow \quad \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_i \alpha_{ji} \right) v_j = 0 \\
&\Leftrightarrow \quad \sum_{i=1}^{m} x_i \alpha_{ji} = 0, \text{ para todo } j \\
&\Leftrightarrow \quad \left[\begin{array}{c} \alpha_{ji} \end{array} \right] \left[\begin{array}{c} x_1 \\ \vdots \\ x_m \end{array} \right] = 0
\end{aligned}$$

e que

$$\sum_{i=1}^{m} x_i w_i = 0 \Leftrightarrow x_1 = x_2 = \cdots = x_m = 0.$$

Portanto,

$$\left[\begin{array}{c} \alpha_{ji} \end{array} \right] \left[\begin{array}{c} x_1 \\ \vdots \\ x_m \end{array} \right] = 0$$

é um sistema *determinado*, pelo que

$$m = \operatorname{car} \left(\left[\begin{array}{c} \alpha_{ji} \end{array} \right] \right) \leq n.$$

Trocando os papéis de $\langle v_1, \dots, v_n \rangle$ e de $\langle w_1, \dots, w_m \rangle$, obtemos $n \leq m$. Logo, $n = m$. □

Definição 4.3.5. *Seja V um espaço vectorial. Se existir uma base de V com n elementos, então diz-se que V tem dimensão n, e escreve-se* $\dim V = n$. *V tem dimensão nula,* $\dim V = 0$ *se $V = \{0\}$.*

Corolário 4.3.6. *Seja V um espaço vectorial com* $\dim V = n$. *Para $m > n$, qualquer conjunto de m elementos de V é linearmente dependente.*

Demonstração. A demonstração segue a do teorema anterior. □

Teorema 4.3.7. *Seja V um espaço vectorial com* $\dim V = n$.

1. *Se $v_1, \dots, v_n$ são linearmente independentes em V, então $v_1, \dots, v_n$ formam uma base de V.*

2. *Se $\langle v_1, \dots, v_n \rangle = V$, então $v_1, \dots, v_n$ formam uma base de V.*

Demonstração. (1) Basta mostrar que $\langle v_1, \ldots, v_n\rangle = V$. Suponhamos, por absurdo, que $v_1, \ldots, v_n$ são linearmente independentes, e que $\langle v_1, \ldots, v_n\rangle \subsetneq V$. Ou seja, existe $0 \neq w \in V$ para o qual $w \notin \langle v_1, \ldots, v_n\rangle = V$. Logo, $v_1, \ldots, v_n, w$, são linearmente independentes, pelo que em V existem $n+1$ elementos linearmente independentes, o que contradiz o corolário anterior.
(2) Basta mostrar que $v_1, \ldots, v_n$ são linearmente independentes. Suponhamos que $v_1, \ldots, v_n$ são linearmente dependentes e que $A = \{v_1, \ldots, v_n\}$. Então pelo menos um deles é combinação linear dos outros. Ou seja, existe v_k tal que $v_k \in \langle v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\rangle$. Se $v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$ não forem linearmente independentes, então repetimos o processo até obtermos $B \subsetneq A$ linearmente independente. Vamos mostrar que $\langle B\rangle = \langle A\rangle$, recordando que $\langle A\rangle = V$. Seja $C = A \setminus B$; isto é, C é o conjunto dos elementos que se retiraram a A de forma a obter o conjunto linearmente independente B. Portanto,

$$v_i \in C \Rightarrow v_i = \sum_{v_j \in B} \beta_{ij} v_j.$$

Seja então $v \in V = \langle A\rangle$. Ou seja, existem α_i's para os quais

$$\begin{aligned} v &= \sum_{v_i \in A} \alpha_i v_i \\ &= \sum_{v_i \in B} \alpha_i v_i + \sum_{v_i \in C} \alpha_i v_i \\ &= \sum_{v_i \in B} \alpha_i v_i + \sum_i \alpha_i \sum_{v_j \in B} \beta_{ij} v_j \\ &= \sum_{v_i \in B} \alpha_i v_i + \sum_i \sum_{v_j \in B} \alpha_i \beta_{ij} v_j \in \langle B\rangle. \end{aligned}$$

Portanto, B é uma base de V com $m < n$ elementos, o que é absurdo. □

Corolário 4.3.8. *Sejam V um espaço vectorial e W_1, W_2 subespaços vectoriais de V. Se $W_1 \subseteq W_2$ e $\dim W_1 = \dim W_2$ então $W_1 = W_2$*

Demonstração. Se $W_1 \subseteq W_2$ e ambos são subespaços de V então W_1 é subespaço de W_2. Seja $\mathcal{B} = \{w_1, \ldots, w_r\}$ uma base de W_1, com $r = \dim W_1$. Segue que $\mathcal{B}$ é linearmente independente em W_2. Como $r = \dim W_1 = \dim W_2$, temos um conjunto linearmente inpedente com r elementos. Por (1) do teorema, $\mathcal{B}$ é base de W_2, o portanto $W_1 = \langle \mathcal{B}\rangle = W_2$. □

Corolário 4.3.9. *Seja V um espaço vectorial e A um conjunto tal que $\langle A\rangle = V$. Então existe $B \subseteq A$ tal que B é base de V.*

Demonstração. A demonstração segue o mesmo raciocínio da demonstração de (2) do teorema anterior. □

4.4 Núcleo e espaço das colunas de uma matriz

Faremos agora a interpretação dos conceitos apresentados anteriormente à custa de matrizes sobre $\mathbb{K}$.

Repare que as colunas de I_n formam uma base de $\mathbb{K}^n$, pelo que $\dim \mathbb{K}^n = n$, se considerarmos os escalares em $\mathbb{K}$. Mostre-se que de facto geram $\mathbb{K}^n$. Se se denotar por e_i a coluna i de I_n, é imediato verificar que $(x_1, x_2, \ldots, x_n) = \sum_{i=1}^n x_i e_i$. Por outro lado, $\sum_{i=1}^n \alpha_i e_i = 0$ implica $(\alpha_1, \alpha_2, \ldots, \alpha_n) = (0, 0, \ldots, 0)$, e portanto $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. O conjunto $\{e_i\}_{i=1,\ldots,n}$ é chamado *base canónica* de $\mathbb{K}^n$.

Teorema 4.4.1. *Sejam $v_1, \cdots, v_n \in \mathbb{K}^m$ e $A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}_{m\times n}$ (as colunas de A são os vectores $v_i \in \mathbb{K}^m$). Então $\{v_1, \cdots, v_n\}$ é linearmente independente se e só se $car(A) = n$.*

Demonstração. Consideremos a equação $Ax = 0$, com $x = [x_1\, x_2\, \cdots\, x_n]^T$. Ou seja, consideremos a equação

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

Equivalentemente,

$$x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = 0.$$

Ou seja, a independência linear de $v_1, \ldots, v_n$ é equivalente a $N(A) = \{0\}$ (isto é, 0 ser a única solução de $Ax = 0$). Recorde que $Ax = 0$ é possível determinado se e só se $car(A) = n$. $\square$

Com base no teorema anterior, os vectores

$$u = (1, 2, 3, 3); v = (2, 0, 1, -1); w = (0, 0, -1, -3)$$

são linearmente independentes. Tal é equivalente a mostrar que

$$\text{car} \begin{bmatrix} u & v & w \end{bmatrix} = \text{car} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 0 \\ 3 & 1 & -1 \\ 3 & -1 & -3 \end{bmatrix} = 3.$$

Para $y = (1, -6, -7, -11)$, os vectores u, v, y não são linearmente independentes, já que $\text{car} \begin{bmatrix} u & v & y \end{bmatrix} = 2$.

Teorema 4.4.2. *Dados $v_1, \ldots, v_m \in \mathbb{K}^n$, seja A a matriz $A = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix} \in \mathcal{M}_{n\times m}(\mathbb{K})$ cujas colunas são $v_1, \ldots, v_m$. Então $w \in \langle v_1, \ldots, v_m \rangle$ se e só se $Ax = w$ tem solução.*

Demonstração. Escrevendo $Ax = w$ como

$$[v_1 \ldots v_m]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = w,$$

temos que $Ax = w$ tem solução se e só se existirem $x_1, x_2, \ldots, x_m \in \mathbb{K}$ tais que

$$x_1v_1 + x_2v_2 + \cdots + x_mv_m = w,$$

isto é, $w \in \langle v_1, \ldots, v_m \rangle$. $\square$

Definição 4.4.3. *Ao subespaço* $CS(A) = \{Ax : x \in \mathbb{K}^n\}$ *de* $\mathbb{K}^m$ *chamamos* imagem *de* A*, ou* espaço das colunas *de* A*. Por vezes,* $CS(A)$ *é denotado também por* $R(A)$ *e por* $Im(A)$*. O espaço das colunas da* A^T *designa-se por* espaço das linhas *de* A *e denota-se por* $RS(A)$*.*

Considerando u, v, w, y como no exemplo anterior, vamos verificar se $y \in \langle u, v, w \rangle$. Para $A = \begin{bmatrix} u & v & w \end{bmatrix}$, tal é equivalente a verificar se $Ax = y$ tem solução. Ou seja, se $\mathrm{car}A = \mathrm{car}\left(\begin{bmatrix} A & | & y \end{bmatrix}\right)$. Aplicando o AEG, deduzimos que $\mathrm{car}(A) = 3 = \mathrm{car}(\begin{bmatrix} A & | & y \end{bmatrix})$.

Já o vector $(0,0,0,1)$ não é combinação linear de u, v, w, ou seja, $(0,0,0,1) \notin \langle u, v, w \rangle$. De facto, $\mathrm{car}(A) \neq \mathrm{car}\left(\left[\begin{array}{c|c} & 0 \\ A & 0 \\ & 0 \\ & 1 \end{array}\right]\right)$.

Vejamos qual a razão de se denominar "espaço das colunas de A" a $CS(A)$. Escrevendo $A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ através das colunas de A, pela forma como o produto de matrizes foi definido, obtemos

$$A\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n.$$

O teorema anterior afirma que $b \in CS(A)$ (i.e., $Ax = b$ é possível) se e só se b for um elemento do espaço *gerado* pelas colunas de A.

A classificação de sistemas de equações lineares como impossível, possível determinado ou possível indeterminado, ganha agora uma nova perspectiva geométrica.

Por exemplo, consideremos a equação matricial $A[x\,y\,z]^T = b$, com $A = \begin{bmatrix} 2 & 4 & -8 \\ 1 & 2 & -4 \\ 2 & 3 & 5 \end{bmatrix}$ e $b = \begin{bmatrix} 14 \\ 7 \\ 10 \end{bmatrix}$. O sistema é possível, já que $car(A) = car([A\,b])$, mas é indeterminado pois

$car(A) < 3$.

As colunas de A, que geram $CS(A)$, não são linearmente independentes. Como $Ax = b$ é possível temos que $b \in CS(A)$, mas não sendo as colunas linearmente independentes, b não se escreverá de forma única como combinação linear das colunas de A. O sistema de equações tem como soluções as realizações *simultâneas* das equações $2x+4y-8z = 14$, $x+2y-4z = 7$ e $2x + 3y + 5z = 10$. Cada uma destas equações representa um plano de $\mathbb{R}^3$, e portanto as soluções de $Ax = b$ são exactamente os pontos de $\mathbb{R}^3$ que estão na intersecção destes planos.

No entanto, o sistema $Ax = c = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$ é impossível, já que $car(A) = 2 \neq 3 = car([A\,c])$. A intersecção dos planos dados pelas equações do sistema é vazia.

Considere agora $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}$ e $b = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. O facto de $Ax = b$ ser impossível (compare a característica de A com a de $[A\,b]$) significa que $b \notin CS(A)$. Ora $CS(A) = \langle (1,1,-1),(1,0,1) \rangle$, ou seja, $CS(A)$ é o conjunto dos pontos de $\mathbb{R}^3$ que se escrevem da forma

$$(x,y,z) = \alpha(1,1,-1) + \beta(1,0,1) = (\alpha+\beta, \alpha, -\alpha+\beta),\ \alpha,\beta \in \mathbb{R}.$$

Com alguns cálculos, podemos encontrar a equação que define $CS(A)$. Recorde que se pretende encontrar os elementos $\begin{bmatrix} x & y & z \end{bmatrix}^T$ para os quais existem $\alpha,\beta \in \mathbb{R}$ tais que

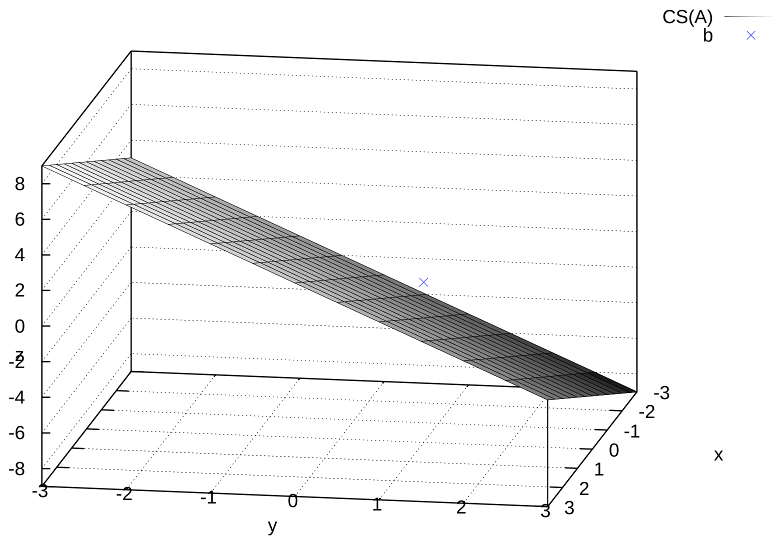
$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Usando o método que foi descrito na parte sobre resolução de sistemas lineares,

$$\left[\begin{array}{cc|c} 1 & 1 & x \\ 1 & 0 & y \\ -1 & 1 & z \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 1 & x \\ 0 & -1 & y-x \\ 0 & 0 & z-x+2y \end{array}\right].$$

Como o sistema tem que ter soluções α,β, somos forçados a ter $z = x - 2y$.

Ora $Ax = b$ é impossível, pelo que $b \notin CS(A)$. Ou seja, b não é um ponto do plano gerado pelas colunas de A.



Se A for invertível, então $CS(A) = \mathbb{K}^n$ (neste caso, tem-se necessariamente $m = n$). De facto, para $x \in \mathbb{K}^n$, podemos escrever $x = A(A^{-1}x)$, pelo que, tomando $y = A^{-1}x \in \mathbb{K}^n$, temos $x = Ay \in CS(A)$. Portanto,

$$\mathbb{K}^n \subseteq CS(A) \subseteq \mathbb{K}^n.$$

Se A, B são matrizes reais para as quais AB existe, temos a inclusão $CS(AB) \subseteq CS(A)$. De facto, se $b \in CS(AB)$ então $ABx = b$, para algum x. Ou seja, $A(Bx) = b$, pelo que $b \in CS(A)$.

Se B for invertível, então $CS(AB) = CS(A)$. Esta igualdade fica provada se se mostrar que $CS(A) \subseteq CS(AB)$. Para $b \in CS(A)$, existe x tal que $b = Ax = A(BB^{-1})x = (AB)B^{-1}x$, e portanto $b \in CS(AB)$.

Recordemos, ainda, que para A matriz real $m \times n$, existem matrizes P, L, U permutação, triangular inferior com 1's na diagonal (e logo invertível) e escada, respectivamente, tais que

$$PA = LU.$$

Ou seja,

$$A = P^{-1}LU.$$

Finalmente, e a comprovação deste facto fica ao cargo do leitor, as linhas não nulas de U, matriz obtida de A por aplicação do método de eliminação de Gauss, são *linearmente independentes*.

Para A, P, L, U definidas atrás,

$$RS(A) = CS(A^T) = CS(U^T(P^{-1}L)^T) = CS(U^T) = RS(U).$$

Ou seja, o espaço das linhas de A e o das linhas de U são o mesmo, e uma base de $RS(A)$ são as linhas não nulas de U enquanto elementos de $\mathbb{K}^n$. Temos, então,

$$RS(A) = RS(U) \text{ e } \dim RS(A) = car(A)$$

Seja QA a forma normal de Hermite de A. Portanto, existe uma matriz permutação P_{erm} tal que $QAP_{\text{erm}} = \left[\begin{array}{c|c} I_r & M \\ \hline 0 & 0 \end{array}\right]$, onde $r = \text{car}(A)$. Repare que $CS(QA) = CS(QAP_{\text{erm}})$, já que o conjunto gerador é o mesmo (ou ainda, porque P_{erm} é invertível). As primeiras r colunas de I_m formam uma base de $CS(QAP_{\text{erm}}) = CS(QA)$, e portanto $\dim CS(QA) = r$. Pretendemos mostrar que $\dim CS(A) = \text{car}(A) = r$. Para tal, considere o lema que se segue:

Lema 4.4.4. *Seja Q uma matriz $n \times n$ invertível e $v_1, v_2, \ldots, v_r \in \mathbb{K}^n$. Então $\{v_1, v_2, \ldots, v_r\}$ é linearmente independente se e só se $\{Qv_1, Qv_2, \ldots, Qv_r\}$ é linearmente independente.*

Demonstração. Repare que $\sum_{i=1}^r \alpha_i Q v_i = 0 \Leftrightarrow Q\left(\sum_{i=1}^r \alpha_i v_i\right) = 0 \Leftrightarrow \sum_{i=1}^r \alpha_i v_i = 0.$ □

Usando o lema anterior,

$$\dim CS(A) = \dim CS(QA) = r = \text{car}(A).$$

Sendo U a matriz escada de linhas obtida por Gauss, U é equivalente por linhas a A, e portanto $\dim CS(U) = \dim CS(A) = \text{car}(A)$.

Considere os vectores de $\mathbb{R}^3$:

$$u = (1, 0, -2), v = (2, -2, 0), w = (-1, 3, -1).$$

Estes formam uma base de $\mathbb{R}^3$, já que $CS(\begin{bmatrix} u & v & w \end{bmatrix}) = \mathbb{R}^3$. Esta igualdade é válida já que $CS(\begin{bmatrix} u & v & w \end{bmatrix}) \subseteq \mathbb{R}^3$ e $\text{car}(\begin{bmatrix} u & v & w \end{bmatrix}) = \dim CS(\begin{bmatrix} u & v & w \end{bmatrix}) = 3$. Fica ao cargo do leitor verificar que, para $A = \begin{bmatrix} u & v & w \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 3 \\ -2 & 0 & -1 \end{bmatrix}$ se tem $\text{car}(A) = 3$.

Já os vectores u, v, q, com $q = (-5, 6, -2)$, não são uma base de $\mathbb{R}^3$. De facto,

$$\text{car}\left(\begin{bmatrix} u & v & q \end{bmatrix}\right) = \text{car}\left(\begin{bmatrix} 1 & 2 & -5 \\ 0 & -2 & 6 \\ -2 & 0 & -2 \end{bmatrix}\right) = 2,$$

e portanto $\dim CS(\begin{bmatrix} u & v & q \end{bmatrix}) = 2 \neq 3 = \dim \mathbb{R}^3$. As colunas da matriz não são linearmente independentes, e portanto não são uma base do espaço das colunas da matriz $\begin{bmatrix} u & v & q \end{bmatrix}$.

A questão que se coloca aqui é: **como obter uma base para** $CS(A)$?

Suponha que V é a matriz escada de linhas obtida da matriz A^T. Recorde que $RS(A^T) = RS(V)$, e portanto $CS(A) = CS(V^T)$. Portanto, e considerando a matriz $A = [u\,v\,q]$ do exemplo anterior, basta-nos calcular uma matriz escada de linhas V associada a A^T. Por exemplo, $V^T = \begin{bmatrix} -5 & 0 & 0 \\ 6 & \frac{6}{5} & 0 \\ -2 & -\frac{12}{5} & 0 \end{bmatrix}$. As duas primeiras colunas de V^T formam uma base de $CS(A)$.

Em primeiro lugar, verifica-se que as r colunas de U com pivot, digamos $u_{i_1}, u_{i_2}, \ldots, u_{i_r}$ são linearmente independentes pois $\begin{bmatrix} u_{i_1} & u_{i_2} & \ldots & u_{i_r} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} = 0$ é possível determinado.

Em segundo lugar, vamos mostrar que as colunas de A correspondentes às colunas de U com pivot são também elas linearmente independentes. Para tal, alertamos para a igualdade $U \begin{bmatrix} e_{i_1} & \ldots & e_{i_r} \end{bmatrix} = \begin{bmatrix} u_{i_1} & u_{i_2} & \ldots & u_{i_r} \end{bmatrix}$, onde e_{i_j} indica a i_j-ésima coluna de I_n. Tendo $U = L^{-1}PA$, e como $\begin{bmatrix} u_{i_1} & u_{i_2} & \ldots & u_{i_r} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} = 0$ é possível determinado, segue que, pela invertibilidade de $L^{-1}P$, a equação $A \begin{bmatrix} e_{i_1} & \ldots & e_{i_r} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} = 0$ admite apenas a solução nula. Mas $A \begin{bmatrix} e_{i_1} & \ldots & e_{i_r} \end{bmatrix}$ é a matriz constituída pelas colunas $i_1, i_2, \ldots, i_r$ de A, pelo que estas são linearmente independentes, em número igual a $r = \text{car}(A)$. Visto $\dim CS(A) = r$, essas colunas constituem de facto uma base de $CS(A)$.

Seja A a matriz do exemplo anterior:

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

Vamos agora descrever esta segunda forma de encontrar uma base de $CS(A)$. Como já vimos, $\text{car}(A) = 2$, pelo que as colunas de A não formam uma base de $CS(A)$ pois não são linearmente independentes, e $\dim CS(A) = 2$. Façamos a decomposição $PA = LU$, trocando a primeira pela terceira linha, obtendo a matriz escada de linhas $U = \begin{bmatrix} -2 & 0 & -2 \\ 0 & -2 & 6 \\ 0 & 0 & 0 \end{bmatrix}$. Uma base possível para $CS(A)$ são as **colunas de A** correspondendo às colunas de U que **têm pivot**. No caso, a primeira e a segunda colunas de A formam uma base de $CS(A)$.

Finalmente, como $car(A^T) = \dim CS(A^T) = \dim RS(A) = car(A)$, temos a igualdade

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

Repare que $N(A) = N(U)$ já que $Ax = 0$ se e só se $Ux = 0$. Na resolução de $Ux = 0$, é feita a separação das incógnitas em básicas e em livres. Recorde que o número destas últimas é denotado por $\mathrm{nul}(A)$. Na apresentação da solução de $Ax = 0$, obtemos, pelo algoritmo para a resolução da equação somas de vectores, cada um multiplicado por uma das incógnitas livres. Esses vectores são geradores de $N(A)$, e são em número igual a $n-r$, onde $r = \mathrm{car}(A)$. Queremos mostrar que $\mathrm{nul}(A) = \dim N(A)$. Seja QA a forma normal de Hermite de A; existe P permutação tal que $QAP = \left[\begin{array}{c|c} I_r & M \\ \hline 0 & 0 \end{array}\right] = H_A$, tendo em mente que $r \leq m, n$. Como Q é invertível, segue que $N(QA) = N(A)$. Sendo H_A a matriz obtida de QA fazendo trocas convenientes de colunas, tem-se $\mathrm{nul}(H_A) = \mathrm{nul}(QA) = \mathrm{nul}(A)$. Definamos a matriz quadrada, de ordem n, $G_A = \left[\begin{array}{c|c} I_r & M \\ \hline 0 & 0 \end{array}\right]$. Como $H_A G_A = H_A$ segue que $H_A(I_n - G) = 0$, e portanto as colunas de $I_n - G$ pertencem a $N(H_A)$. Mas $I_n - G = \left[\begin{array}{c|c} 0 & M \\ \hline 0 & I_{n-r} \end{array}\right]$ e as suas últimas $n-r$ colunas são linearmente independentes (já que $\mathrm{car}\left(\begin{bmatrix} M \\ I_{n-r} \end{bmatrix}\right) = \mathrm{car}\left(\begin{bmatrix} I_{n-r} \\ M \end{bmatrix}\right) = \mathrm{car}\left(\begin{bmatrix} I_{n-r} \\ M \end{bmatrix}^T\right) = n-r$). Logo, $\dim N(A) = \dim N(H_A) \geq n-r$. Pelo que vimos atrás, $\dim N(A) = \dim N(U) \leq n-r$. Segue das duas desigualdades que

$$\mathrm{nul}(A) = \dim N(A).$$

Como $n = \mathrm{car}(A) + \mathrm{nul}(A)$, obtemos, finalmente,

$$n = \dim CS(A) + \dim N(A).$$

Considere o subespaço W de $\mathbb{R}^3$ gerado pelos vectores

$$(1,2,1), (2,-3,-1), (3,1,2), (4,1,2), (5,0,4).$$

Como temos 5 vectores de um espaço de dimensão 3, eles são necessariamente linearmente dependentes. Qual a dimensão de W? W é o espaço das colunas da matriz A, cujas colunas são os vectores dados:

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

Ora $\dim CS(A) = \mathrm{car}(A)$. Calcule a matriz U escada de linhas aplicando o AEG, e verifique que as colunas de U com pivots são a primeira, a segunda e a terceira. Como $\mathrm{car}(A) = 3$ então

$\dim W = 3$. Ora $W \subseteq \mathbb{R}^3$ e têm a mesma dimensão, pelo que $W = \mathbb{R}^3$. Ou seja, as colunas de A geram $\mathbb{R}^3$. As colunas de A que formam uma base para W são aquelas correspondentes às colunas de U que têm pivot; neste caso, as três primeiras de U. Uma base $\mathcal{B}$ para W é o conjunto formado pelos vectores $v_1 = (1,2,1), v_2 = (2,-3,-1), v_3 = (3,1,2)$. Vamos agora calcular as coordenadas de $b = (0,-2,-2)$ nesta base. Tal corresponde a resolver a equação $\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} x = b$. A única solução é o vector $(1,1,-1)$, que é o vector das coordenadas de b na base v_1, v_2, v_3. Ou seja, $(0,-2,-2)_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

Vamos agora apresentar alguns resultados importantes que se podem deduzir facilmente à custa de $car(A) + \text{nul}(A) = n$, onde A é uma matriz $m \times n$. Pressupõe-se que B é uma matriz tal que AB existe.

1. $car(AB) \leq car(A)$.
   Como vimos na secção anterior, $CS(AB) \subseteq CS(A)$, pelo que $\dim CS(AB) \leq \dim CS(A)$.

2. Se B é invertível então $car(A) = car(AB)$.

3. $N(B) \subseteq N(AB)$.
   Se $b \in N(B)$ então $Bb = 0$. Multiplicando ambos os lados, à esquerda, por A obtemos $ABb = 0$, pelo que $b \in N(AB)$.

4. $\text{nul}(B) \leq \text{nul}(AB)$.

5. $N(A^T A) = N(A)$.
   Resta mostrar que $N(A^T A) \subseteq N(A)$. Se $x \in N(A^T A)$ então $A^T Ax = 0$. Multiplicando ambos os lados, à esquerda, por x^T obtemos $x^T A^T Ax = 0$, pelo $(Ax)^T Ax = 0$. Seja $(y_1, \ldots, y_n) = y = Ax$. De $y^T y = 0$ obtemos $y_1^2 + y_2^2 + \ldots y_n^2 = 0$. A soma de reais não negativos é zero se e só se cada parcela é nula, pelo que cada $y_i^2 = 0$, e portanto $y_i = 0$. Ou seja, $y = 0$, donde segue que $Ax = 0$, ou seja, que $x \in N(A)$.

6. $\text{nul}(A^T A) = \text{nul}(A)$.

7. $car(A^T A) = car(A) = car(AA^T)$.
   De $car(A) + \text{nul}(A) = n = car(A^T A) + \text{nul}(A^T A)$ e $\text{nul}(A^T A) = \text{nul}(A)$ segue que $car(A^T A) = car(A)$. Da mesma forma, $car(A^T) = car(AA^T)$. Como $car(A) = car(A^T)$, obtemos $car(A) = car(AA^T)$.

8. Se $car(A) = n$ então $A^T A$ é invertível.
   $A^T A$ é uma matriz $n \times n$ com característica igual a n, pelo que é uma matriz não-singular, logo invertível.

Exercícios

1. Considere, no espaço vectorial real $\mathbb{R}^3$, os vectores

   $v_1 = (1,-1,1), \quad v_2 = (2,1,-2), \quad u_1 = (-1,0,1), \quad u_2 = (1,0,0), \quad u_3 = (1,0,1).$

   Verifique se

   (a) $(1,-4,5)$ é combinação linear de v_1, v_2.

   (b) $(1,2,1)$ é combinação linear de v_1, v_2.

   (c) $(3,0,2)$ é combinação linear de u_1, u_2, u_3.

   (d) $(0,2,1)$ é combinação linear de u_1, u_2, u_3.

2. Verifique se $(2,5,-3) \in \langle (1,4,-2), (-2,1,3) \rangle$.

3. Determine α, β de forma a que $(1,1,\alpha,\beta) \in \langle (1,0,2,1), (1,-1,2,2) \rangle$.

4. Diga quais dos seguintes conjuntos de vectores são linearmente dependentes:

   (a) $\{(0,1,1,0), (-1,0,1,1), (1,1,0,-1)\}$ no espaço vectorial real $\mathbb{R}^4$.

   (b) $\{(1,2,1), (-2,3,1)\}$ no espaço vectorial real $\mathbb{R}^3$.

   (c) $\{(0,1), (1,2), (2,3)\}$ no espaço vectorial real $\mathbb{R}^2$.

5. Considere os seguintes subespaços vectoriais do espaço vectorial real $\mathbb{R}^3$:

   $$V_1 = \{(x,y,z) \in \mathbb{R}^3 : z = 0\}, \qquad V_2 = \{(x,y,z) \in \mathbb{R}^3 : y+z = 0, y-z = 0\},$$
   $$V_3 = \{(x,y,z) \in \mathbb{R}^3 : x-y = 0, 2y+z = 0\}, \quad V_4 = \{(x,y,z) \in \mathbb{R}^3 : x+y = 0\}.$$

   Indique a dimensão e uma base para cada um deles.

6. Considere, no espaço vectorial real $\mathbb{R}^4$, os subespaços

   $U = \{(a_1,a_2,a_3,a_4) \in \mathbb{R}^4 : a_1 - a_4 = 0, a_4 - a_3 = 0\}$

   $W_1 = \{(b_1,b_2,b_3,b_4) \in \mathbb{R}^4 : b_2 + 2b_3 = 0, b_1 + 2b_3 - b_4 = 0\}$

   $W_2 =< (1,1,1,0), (-1,1,0,1), (1,3,2,1), (-3,1,-1,2) >.$

   (a) Diga, justificando, se $\{(1,1,1,1), (0,1,0,0), (1,0,0,1)\}$ é uma base de U.

   (b) Determine uma base de i. W_1. ii. W_2.

7. Considere os seguintes vectores do espaço vectorial real $\mathbb{R}^3$:

   $$v_1 = (\alpha,6,-1),\ v_2 = (1,\alpha,-1),\ v_3 = (2,\alpha,-3).$$

   (a) Determine os valores do parâmetro real α para os quais o conjunto $\{v_1, v_2, v_3\}$ é uma base de $\mathbb{R}^3$.

   (b) Para um dos valores de α determinados na alínea anterior, calcule as coordenadas do vector $v = (-1, 1.2)$ em relação à base $\{v_1, v_2, v_3\}$.

8. Considere os seguintes elementos de $\mathbb{R}^3$:

$$v_1 = (1,0,2), v_2 = (1,-1,1), v_3 = (0,-1,1), v_4 = (1,-\frac{1}{2},\frac{3}{2}).$$

Verifique se $\langle v_1, v_2\rangle = \langle v_3, v_4\rangle$.

9. Considere os elementos de $\mathbb{R}^3$: $v_1 = (2,-3,1), v_2 = (0,1,2), v_3 = (1,1,-2)$.

   (a) Mostre que são uma base de $\mathbb{R}^3$.

   (b) Determine as coordenadas de $(3,2,1)$ relativamente a esta base.

10. Mostre que os vectores $(a,b),(c,d)$ são uma base de $\mathbb{R}^2$ se e só se $ad - bc \neq 0$.

11. Considere os seguintes subespaços de $\mathbb{R}^4$:

$$\{(x_1,x_2,x_3,x_4) : x_1 + x_2 = 0\}, \{(x_1,x_2,x_3,x_4) : x_2 = 0\},$$

$$\{(x_1,x_2,x_3,x_4) : x_1 + x_2 + x_3 + x_4 = 0\}.$$

Para cada um deles, determine a dimensão e indique uma base.

12. Considere os seguintes subespaços de $\mathbb{R}^4$:

$$F = \{(x_1,x_2,x_3,x_4) : x_1 = x_3 \wedge x_4 = 2x_2\}, G = \langle(1,0,1,0),(0,2,0,1),(-1,2,-1,1)\rangle.$$

Determine a dimensão e indique uma base para F e para G.

13. Encontre uma base para o espaço das colunas das matrizes seguintes:

   (a) $\begin{bmatrix} 2 & 8 & -2 \\ 1 & -17 & 6 \\ 9 & 6 & 1 \end{bmatrix}$

   (b) $\begin{bmatrix} 3 & -4 & 0 & 4 \\ 2 & 4 & -2 & 0 \\ 3 & -2 & 4 & -4 \end{bmatrix}$.

   (c) $\begin{bmatrix} 0 & -7 & 3 & -8 & -1 \\ -1 & 6 & -8 & -2 & 3 \\ 0 & 0 & 0 & 5 & -8 \end{bmatrix}$

14. Indique, justificando convenientemente, o valor lógico da seguinte afirmação:

`‘‘Se as colunas da matriz quadrada` A `são linearmente independentes, então as colunas de` A^2 `são também elas linearmente independentes.’’`

15. Seja $A \in \mathcal{M}_n(\mathbb{R})$. Mostre que

   (a) se $A^2 = A$ e $car(A) = n$ então $A = I_n$;

   (b) se $A^2 = A$ então $CS(A) \cap N(A) = \{0\}$.

4.5 Uma aplicação

Como motivação para o que se segue, suponha que se quer encontrar (caso exista) a recta r de $\mathbb{R}^2$ que incide nos pontos $(-2,-5),(0,-1),(1,1)$. Sendo a recta não vertical, terá uma equação da forma $y = mx + c$, com $m, c \in \mathbb{R}$. Como r incide nos pontos indicados, então necessariamente

$$-5 = m \cdot (-2) + c,\ -1 = m \cdot 0 + c,\ 1 = m \cdot 1 + c.$$

A formulação matricial deste sistema de equações lineares (nas incógnitas m e c) é

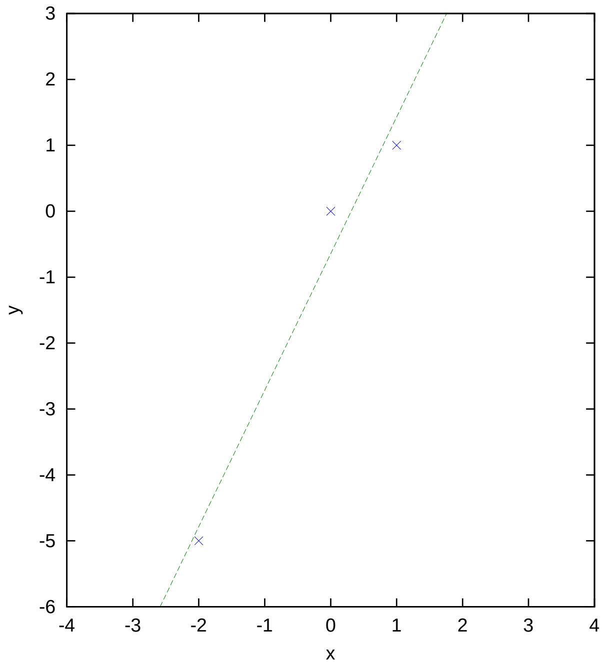
$$\begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \\ 1 \end{bmatrix}.$$

O sistema é possível determinado, pelo que a existência da recta e a sua unicidade está garantida. A única solução é $(m, c) = (2, 1)$ e portanto a recta tem equação $y = 2x - 1$.

No entanto, se considerarmos como dados os pontos $(-2,-5),(0,0),(1,1)$, facilmente chegaríamos à conclusão que não existe uma recta incidente nos três pontos. Para tal, basta mostrar que o sistema de equações dado pelo problema (tal como fizemos no caso anterior) é impossível. Obtemos a relação

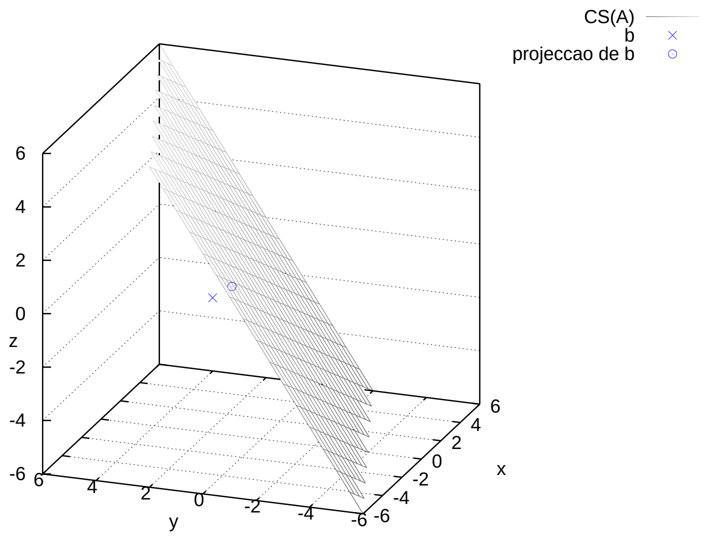
$$b \notin CS(A),$$

onde $A = \begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ e $b = \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$. Suponha que os pontos dados correspondem a leituras de uma certa experiência, pontos esses que, teoricamente, deveriam ser colineares. Ou seja, em algum momento houve um desvio da leitura em relação ao que se esperaria. Desconhece-se qual ou quais os pontos que sofreram incorrecções. Uma solução seria a de negligenciar um dos pontos e considerar os outros dois como correctos. É imediato concluir que este raciocínio pode levar a conclusões erróneas. Por exemplo, vamos pressupor que é o primeiro dado que está incorrecto (o ponto $(-2,-5)$). A rectas que passa pelos pontos $(0,0),(1,1)$ tem como equação $y = x$. Ora se o erro esteve efectivamente na leitura do ponto $(0,0)$ (que deveria ser $(0,-1)$) então o resultado correcto está bastante distante do que obtivémos. O utilizador desconhece qual (ou quais, podendo haver leituras incorrectas em todos os pontos) dos dados sofreu erros. Geometricamente, a primeira estratégia corresponde a eliminar um dos pontos e traçar a recta que incide nos outros dois. Uma outra que, intuitivamente, parece a mais indicada, será a de, de alguma forma e com mais ou menos engenho, traçar uma recta que se tente *aproximar o mais possível* de todos os pontos, ainda que não incida em nenhum deles!



Vamos, de seguida, usar todo o engenho que dispomos para encontrar a recta que se aproxima o mais possível dos pontos $(-2,-5),(0,0),(1,1)$.

Sabendo que $b \notin CS(A)$, precisamos de encontrar $b' \in CS(A)$ por forma a que b' seja o ponto de $CS(A)$ mais próximo de b. Ou seja, pretendemos encontrar $b' \in CS(A)$ tal que $d(b,b') = \min_{c\in CS(A)} d(c,b)$, onde $d(u,v) = \|u-v\|$. O ponto b' é o de $CS(A)$ que minimiza a distância a b. Este ponto b' é único e é tal que $b-b'$ é ortogonal a todos os elementos de $CS(A)$. A b' chamamos *projecção ortogonal* de b sobre (ou ao longo) de $CS(A)$, e denota-se por $proj_{CS(A)}b$.



Apresentamos, de seguida, uma forma fácil de cálculo dessa projecção, quando as colunas de A são linearmente independentes. Neste caso, A^TA é invertível e a projecção de b sobre $CS(A)$ é dada por

$$b' = A(A^TA)^{-1}A^Tb.$$

Pretendemos agora encontrar x por forma a que $Ax = b'$, ou seja, x por forma a que a distância de Ax a b seja a menor possível. Repare que se $Ax = b$ é impossível, então essa distância será, seguramente, não nula. A equação $Ax = b'$ é sempre possível, já que $b' = A(A^TA)^{-1}A^Tb \in CS(A)$; ou seja, b' escreve-se como Aw, para algum w (bastando tomar $w = (A^TA)^{-1}A^Tb$). No entanto, o sistema pode ser indeterminado, e nesse caso poderá interessar, de entre todas as soluções possíveis, a que tem norma mínima. O que acabámos por expôr, de uma forma leve e ingénua, denomina-se o *método dos mínimos quadrados*, e a x solução de $Ax = b'$ de norma minimal, denomina-se a solução no sentido dos mínimos quadrados de norma minimal.

Exercícios

1. Calcule a projecção ortogonal do vector $(2, -1, 1)$ sobre o espaço gerado por $(1, 1, 1), (0, 1, 3)$.

2. Para $A = \begin{bmatrix} 1 & -3 & 2 \\ 4 & 10 & -1 \end{bmatrix}^T$ e $b = \begin{bmatrix} 5 & 7 & 10 \end{bmatrix}^T$, determine a solução no sentido dos mínimos quadrados de $Ax = b$.

3. Determine a solução no sentido dos mínimos quadrados de

(a) $\begin{cases} x_1 &= 1 \\ x_2 &= 1 \\ x_1 + x_2 &= 0 \end{cases}$

(b) $\begin{cases} x_1 + 2x_2 &= 1 \\ 2x_1 + 5x_2 &= 0 \\ 3x_1 + 7x_2 &= 2 \end{cases}$

4. Considere a matriz $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$.

   (a) Calcule a projecção ortogonal de $b = \begin{bmatrix} 4 & 5 & -1 & 4 \end{bmatrix}^T$ sobre $CS(A)$.

   (b) O que pode dizer sobre o sistema $Ax = b$?

Ao invés de procurarmos a recta que melhor se adequa aos dados disponíveis, podemos procurar o polinómio de segundo, terceiro, etc, graus. Se os dados apresentados forem pontos de $\mathbb{R}^3$, podemos procurar o plano que minimiza as somas das distâncias dos pontos a esse plano. E assim por diante, desde que as funções que definem a curva ou superfície sejam lineares nos parâmetros. Por exemplo, $ax^2 + bx + c = 0$ não é uma equação linear em x mas é-o em a e b.

Exemplo 4.5.1. O exemplo que de seguida apresentamos baseia-se no descrito em [3, pag.58]

Suponha que se está a estudar a cinética de uma reacção enzimática que converte um substrato S num produto P, e que essa reacção segue a equação de Michaelis-Menten,

$$r = \frac{k_2[E]_0[S]}{K_m + [S]},$$

onde

1. $[E]_0$ indica concentração enzimática original adicionada para iniciar a reacção, em gramas de E por litro,

2. r é o número de gramas de S convertido por litro por minuto (ou seja, a velocidade da reacção),

3. k_2 é o número de gramas de S convertido por minuto por grama de E.

Depois de se efectuar uma série de experiências, obtiveram-se os dados apresentados na tabela seguinte, referentes à taxa de conversão de gramas de S por litro por minuto:

$[S]$ g s/l	$[E]_0 = 0.005$ g_E/l	$[E]_0 = 0.01$ g_E/l
1.0	0.055	0.108
2.0	0.099	0.196
5.0	0.193	0.383
7.5	0.244	0.488
10.0	0.280	0.569
15.0	0.333	0.665
20.0	0.365	0.733
30.0	0.407	0.815

Re-escrevendo a equação de Michaelis-Menten como

$$\frac{[E]_0}{r} = \frac{K_m}{k_2}\frac{1}{[S]} + \frac{1}{k_2},$$

obtemos um modelo linear

$$y = b_1 x + b_0$$

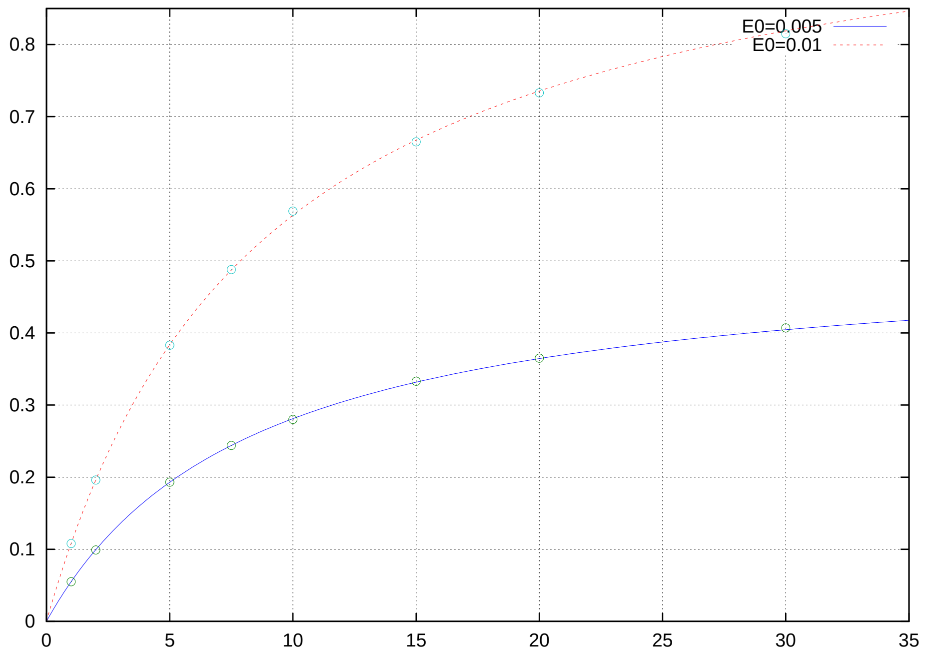
com

$$y = \frac{[E]_0}{r},\ x = \frac{1}{[S]},\ b_0 = \frac{1}{k_2},\ b_1 = \frac{K_m}{k_2}.$$

Denotemos os dados x e y por x_i e y_i, com $i = 1, \ldots, 8$. Este sistema de equações lineares tem a representação matricial

$$A \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} = y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_8 \end{bmatrix}$$

com $A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_8 & 1 \end{bmatrix}$. A única solução de $A^T A \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} = y$ indica-nos a solução no sentido dos mínimos quadrados da equação matricial, e daqui obtemos os valores de k_2 e de K_m.

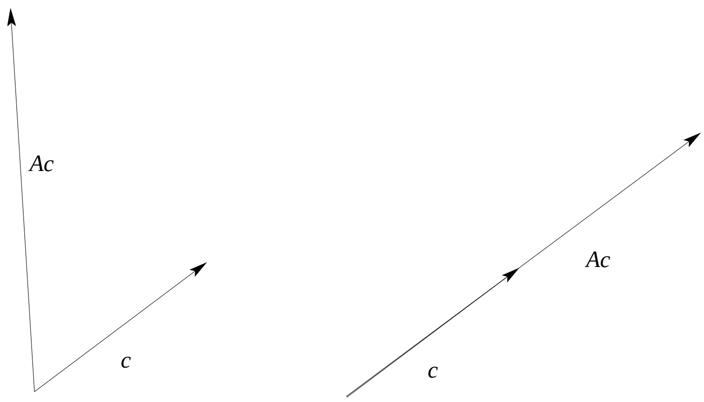

E0=0.005
E0=0.01
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
5
10
15
20
25
30
35

Capítulo 5

Valores e vectores próprios

5.1 Motivação e definições

Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$. Para $b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, obtemos $Ab = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Mas se tomarmos $c = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, temos que $Ac = 2c$. Ou seja, Ac é um múltiplo de c.



Dada uma matriz complexa A quadrada, $n \times n$, um vector $x \in \mathbb{C}^n$ não nulo diz-se um *vector próprio* de A se $Ax = \lambda x$, para algum $\lambda \in \mathbb{C}$. O complexo λ é denominado *valor próprio*, e dizemos que x é vector próprio associado a λ. O conjunto dos valores próprios de A é denotado por $\sigma(A)$ e é chamado de espectro de A.

No exemplo apresentado atrás, temos que $2 \in \sigma(A)$ e que $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio 2.

Uma questão que colocamos desde já é:

Como encontrar $\sigma(A)$?

Ora, sendo A uma matriz complexa $n \times n$ e se λ é valor próprio de A então existe $x \in \mathbb{C}^n \setminus \{0\}$ para o qual $Ax = \lambda x$. Ou seja, $\lambda I_n x - Ax = \lambda x - Ax = 0$, o que equivale a $(\lambda I_n - A)x = 0$. Como $x \neq 0$, tal significa que a equação $(\lambda I_n - A)x = 0$ é consistente e que tem solução não nula. Isto é, a matriz quadrada $\lambda I_n - A$ tem característica estritamente inferior ao número de colunas, o que acontece se e só se não é invertível, ou de forma equivalente, o seu determinante é nulo. Os valores próprios de A são os escalares λ que tornam $\lambda I_n - A$ uma matriz singular, ou seja, que satisfazem $|\lambda I_n - A| = 0$. Ora $|\lambda I_n - A|$ é um polinómio em λ, usando o teorema de Laplace, denominado *polinómio característico* de A, e denotado por Δ_A. Os valores próprios de A são as raizes do polinómio característico Δ_A, ou seja, as soluções da equação $\Delta_A(\lambda) = 0$. Esta equação é chamada a equação característica de A.

Determinar os valores próprios de uma matriz equivalente a determinar as raizes do seu polinómio característico. Usando o teorema de Laplace, este polinómio tem grau igual à ordem da matriz A, que assumimos $n \times n$, e é mónico: o coeficiente de λ^n de $\Delta_A(\lambda)$ é 1. Pelo Teorema Fundamental da Álgebra, sendo o grau de Δ_A igual a n este tem n raizes (contando as suas multiplicidades) sobre $\mathbb{C}$. Ou seja, a matriz A do tipo $n \times n$ tem então n valores próprios (contando com as suas multiplicidades). Sabendo que se $z \in \mathbb{C}$ é raiz de Δ_A então o conjugado $\bar{z}$ de z é raiz de Δ_A, segue que se $\lambda \in \sigma(A)$ então $\bar{\lambda} \in \sigma(A)$. Em particular, se A tem um número ímpar de valores próprios (contado as suas multiplicidades) então tem pelo menos um valor próprio real. Isto é, $\sigma(A) \cap \mathbb{R} \neq \emptyset$. A *multiplicidade algébrica* de um valor próprio λ é a multiplicidade da raiz λ de Δ_A.

Vimos no que se discutiu acima uma forma de determinar os valores próprios de uma matriz. Dado um valor próprio λ,

Como determinar os vectores próprios associados a $\lambda \in \sigma(A)$?

Recorde que os vectores próprios associados a $\lambda \in \sigma(A)$ são as soluções *não-nulas* de $Ax = \lambda x$, ou seja, as soluções não nulas de $(\lambda I_n - A)x = 0$. Isto é, os vectores próprios de A associados a λ são os elementos não nulos de $N(\lambda I_n - A)$. Recorde que o núcleo de qualquer matriz é um espaço vectorial, e portanto $N(\lambda I_n - A)$ é o espaço vectorial dos vectores próprios de A associados a λ juntamente com o vector nulo, e denomina-se *espaço próprio de A associado a λ*. A *multiplicidade geométrica* de λ é a dimensão do espaço próprio associado a λ, isto é, $\dim N(\lambda I_n - A)$.

O resultado seguinte resume o que foi afirmado na discussão anterior.

Teorema 5.1.1. *Sejam A uma matriz $n \times n$ e $\lambda \in \mathbb{C}$. As afirmações seguintes são equivalentes:*

1. $\lambda \in \sigma(A)$;
2. $(\lambda I_n - A)x = 0$ *é uma equação possível indeterminada;*
3. $\exists_{x \in \mathbb{C}^n \setminus \{0\}}\, Ax = \lambda x$;
4. λ *é solução de* $|\tilde{\lambda} I_n - A| = 0$.

Para a matriz considerada acima, $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$, o seu polinómio característico é $\Delta_A(\lambda) = \begin{vmatrix} \lambda - 1 & -2 \\ -2 & \lambda + 2 \end{vmatrix} = \lambda^2 + \lambda - 6$, cujas raizes são $-3, 2$. Portanto, $\sigma(A) = \{-3, 2\}$, e cada valor próprio de A tem multiplicidade algébrica igual a 1.

Teorema 5.1.2. *Sejam A uma matriz quadrada e $\lambda \in \sigma(A)$ com multiplicidade algébrica ν_λ e multiplicidade geométrica η_λ. Então*

$$\nu_\lambda \geq \eta_\lambda.$$

Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$. O polinómio característico de A iguala $\Delta_A(\lambda) = \lambda^2 + \lambda - 6$. As raizes de Δ_A são os elementos de $\sigma(A) = \{-3, 2\}$. A multiplicidade algébrica de cada um deles é 1.

Resta-nos determinar vectores próprios associados a cada um destes valores próprios. Recorde que os vectores próprios associados a -3 [resp. 2] são os elementos não nulos de $N(-3I_2 - A)$ [resp. $N(2I_2 - A)$], pelo que nos basta calcular uma base para cada espaço próprio. Como a multiplicidade geométrica (ou seja, a dimensão do espaço próprio) não pode ser superior à algébrica, e cada uma delas vale 1, segue que as multiplicidades algébrica e geométrica de cada um dos valores próprios são iguais.

5.2 Propriedades

Nos resultados que se seguem descrevemos algumas propriedades dos valores própios.

Teorema 5.2.1. *Dada uma matriz quadrada A,*

$$\sigma(A) = \sigma(A^T).$$

Demonstração. Recorde que $|\lambda I - A| = |(\lambda I - A)^T| = |\lambda I - A^T|$. □

Teorema 5.2.2. *Os valores próprios de uma matriz triangular (inferior ou superior) são os seus elementos diagonais.*

Demonstração. Seja $A = [a_{ij}]$ triangular superior, $n \times n$. Ora $\sigma(A)$ é o conjunto das soluções de $|\lambda I_n - A|$. Mas $\lambda I_n - A$ é de novo uma matriz triangular superior já que λI_n é diagonal. Portanto $|\lambda I_n - A|$ é o produto dos seus elementos diagonais, ou seja, $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$, que tem como raizes $a_{11}, a_{22}, \ldots, a_{nn}$. □

Teorema 5.2.3. *Uma matriz A, quadrada, é invertível se e só se $0 \notin \sigma(A)$.*

Demonstração. Sejam A uma matriz quadrada de ordem n e $\Delta_A(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$ o polinómio característico de A. Ora $0 \in \sigma(A)$ se e só se 0 é raiz de Δ_A, ou de forma equivalente, $c_n = 0$.

Por definição, $\Delta_A(\lambda) = |\lambda I_n - A|$. Tomando $\lambda = 0$ obtemos $(-1)^n|A| = |-A| = c_n$. tal implica que $|A| = 0$ se e só se $c_n = 0$. Portanto A não é invertível se e só se $c_n = 0$ o que por sua vez vimos ser equivalente a $0 \in \sigma(A)$. □

Teorema 5.2.4. *Sejam A uma matriz quadrada e $k \in \mathbb{N}$. Se $\lambda \in \sigma(A)$ e x é vector próprio associado a λ então $\lambda^k \in \sigma(A^k)$ e x é vector próprio de A^k associado a λ^k.*

Demonstração. Se $\lambda \in \sigma(A)$ e x é vector próprio associado a λ então $Ax = \lambda x$. Desta igualdade segue que, para qualquer $k \in \mathbb{N}$, se tem

$$A^k x = A^{k-1}Ax = A^{k-1}\lambda x = \lambda A^{k-1}x = \cdots = \lambda^k x$$

e portanto $\lambda \in \sigma(A^k)$ e x é vector próprio de A^k associado a λ^k. □

Recordamos que uma matriz N, $n \times n$, se diz nilpotente se existir um natural k para o qual $N^k = 0_{n\times n}$.

Alertamos ainda para o facto de $\sigma(0_{n\times n}) = \{0\}$; isto é, a matriz nula só tem um valor próprio: o zero.

Corolário 5.2.5. *Se N é uma matriz nilpotente então $\sigma(N) = \{0\}$.*

Demonstração. Suponha que k é tal que $N^k = 0_{n\times n}$. Seja $\lambda \in \sigma(N)$. Então λ^k é valor próprio de $N^k = 0_{n\times n}$; portanto, $\lambda^k = 0$, do que segue que $\lambda = 0$. □

Terminamos esta secção com duas observações, omitindo a sua prova:

(i) O determinante de uma matriz iguala o produto dos seus valores próprios.

(ii) O traço de uma matriz (ou seja, a soma dos elementos diagonais de uma matriz) iguala a soma dos seus valores próprios.

5.3 Matrizes diagonalizáveis

Nesta secção, vamo-nos debruçar sobre dois problemas, que aliás, e como veremos, estão relacionados. Assume-se que A é uma matriz $n \times n$ sobre $\mathbb{C}$. Essas questões são:

1. Existe uma base de $\mathbb{C}^n$ constituída por vectores próprios de A?

2. Existe uma matriz U invertível para a qual $U^{-1}AU$ é uma matriz diagonal?

Recordamos a noção de semelhança entre matrizes. As matriz A e B dizem-se *semelhantes*, e denota-se por $A \approx B$, se existir uma matriz invertível U para a qual $B = U^{-1}AU$. Repare que as matrizes A, B são necessariamente quadradas.

É óbvio que se $A \approx B$ então $B \approx A$; de facto, se $B = U^{-1}AU$ então $UBU^{-1} = A$.

Definição 5.3.1. *Uma matriz quadrada A diz-se* diagonalizável *se existir uma matriz diagonal D tal que $A \approx D$. Isto é, $A = UDU^{-1}$, para alguma matriz U invertível. À matriz U chamamos matriz diagonalizante.*

É óbvio que uma matriz diagonal é diagonalizável, bastando tomar a matriz identidade como matriz diagonalizante.

O resultado seguinte não só nos caracteriza as matrizes diagonalizáveis, mas também, à custa da sua prova, obtemos um algoritmo para encontrar a matriz diagonal e a a respectiva matriz diagonalizante.

Teorema 5.3.2. *Uma matriz $n \times n$ é diagonalizável se e só se tiver n vectores próprios linearmente independentes.*

Demonstração. Em primeiro lugar, assumimos que A é diagonalizável; ou seja, existe uma matriz $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$ invertível tal que $U^{-1}AU = D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$.

Como é óbvio, de $U^{-1}AU = D$ segue que $AU = UD$. Portanto,

$$\begin{aligned} \begin{bmatrix} Au_1 & Au_2 & \cdots & Au_n \end{bmatrix} &= AU = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n u_n \end{bmatrix} \end{aligned}$$

e portanto

$$\begin{cases} Au_1 = \lambda_1 u_1 \\ Au_2 = \lambda_2 u_2 \\ \quad \vdots \qquad \vdots \\ Au_n = \lambda_n u_n \end{cases}.$$

Como U é invertível, então não pode ter colunas nulas, pelo que $u_i \neq 0$. Portanto, $\lambda_1, \lambda_2, \ldots, \lambda_n$ são valores próprios de A e $u_1, u_2, \ldots, u_n$ são respectivos vectores próprios. Sendo U invertível, as suas colunas são linearmente independentes, e portanto A tem n vectores próprios linearmente independentes.

Reciprocamente, suponha que A tem n vectores próprios linearmente independentes. Sejam eles os vectores $u_1, u_2, \ldots, u_n$, associados aos valores próprios (não necessariamente distintos) $\lambda_1, \lambda_2, \ldots, \lambda_n$. Seja U a matriz cujas colunas são os vectores próprios considerados acima. Ou seja, $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$. Ora esta matriz quadrada $n \times n$ tem característica igual a n, pelo que é invertível. De

$$\begin{cases} Au_1 = \lambda_1 u_1 \\ Au_2 = \lambda_2 u_2 \\ \quad \vdots \qquad \vdots \\ Au_n = \lambda_n u_n \end{cases}$$

segue que $\begin{bmatrix} Au_1 & Au_2 & \cdots & Au_n \end{bmatrix} = \begin{bmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n u_n \end{bmatrix}$ e portanto

$$A \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

Multiplicando ambas as equações, à esquerda, por $\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^{-1}$, obtemos

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^{-1} A \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

$\square$

Realçamos o facto da demonstração do teorema nos apresentar um algoritmo de diagonalização de uma matriz $n \times n$ com n vectores linearmente independentes. De facto, de $\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^{-1} A \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ obtemos

$$A = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^{-1}.$$

Uma matriz diagonalizante é a matriz cujas colunas são os vectores próprios linearmente independentes dados, e a matriz diagonal correspondente é a matriz cuja entrada (i, i) é o valor próprio λ_i correspondente à coluna i (e portanto ao $i-$ésimo vector próprio) da matriz diagonalizante.

Para a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$, vimos atrás que $\sigma(A) = \{-3, 2\}$. Será A diagonalizável? Um vector próprio associado ao valor próprio -3 é um elemento não nulo de $N(-3I_2 - A)$. Encontrar um vector próprio associado a -3 é equivalente a encontrar uma solução não nula de $(-3I_2 - A)x = 0$. Fica ao cargo do leitor verificar que $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ é vector próprio associado ao valor próprio -3, e fazendo o mesmo raciocínio, que $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio 2. Ora estes dois vectores são linearmente independentes, visto car $\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix} = 2$. Portanto, a matriz A é diagonalizável, sendo a matriz diagonalizante $U = \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$ e a matriz diagonal $\begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$.

Considere agora a matriz $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Esta matriz é nilpotente, pelo que $\sigma(B) = \{0\}$. O espaço próprio associado a 0 é $N(-B) = N(B)$. Ora $\text{car}(B) = 1$, pelo que $\text{nul}(B) = 1$, e portanto a multiplicidade geométrica do valor próprio 0 é 1 (repare que a multiplicidade algébrica do valor próprio 0 é 2). Ou seja, não é possível encontrar 2 vectores próprios linearmente independentes.

A matriz $C = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, sendo triangular superior, tem como valores próprios os elementos diagonais da matriz. Isto é, $\sigma(C) = \{2\}$. Repare que a multiplicidade algébrica do valor próprio 2 é 2. Repare que $\operatorname{car}(2I_2 - C) = \operatorname{car} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 1$, pelo que $\operatorname{nul}(2I_2 - C) = 1$. Logo, não é possível encontrar 2 vectores próprios de C linearmente independentes, e portanto C não é diagonalizável.

A matriz $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -2 & 2 \\ 0 & 0 & -3 \end{bmatrix}$. tem como espectro $\sigma(A) = \{-3, 2\}$, sendo as multiplicidades algébricas de -3 e 2, respectivamente, 2 e 1. Como $\operatorname{car}(-3I_3 - A) = 2$, temos que $\operatorname{nul}(-3I_3 - A) = 1$, e portanto a multiplicidade geométrica do valor próprio -3 é 1. Portanto, a matriz não é diagonalizável pois não é possível encontrar 3 vectores próprios linearmente independentes.

O que se pode dizer em relação à independência linear de um vector próprio associado a -3 e um vector próprio associado a 2?

Teorema 5.3.3. *Sejam $v_1, v_2, \ldots, v_k$ vectores próprios associados a valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_k$ distintos entre si. Então $\{v_1, v_2, \ldots, v_k\}$ é um conjunto linearmente independente.*

Demonstração. Suponhamos que $\{v_1, v_2, \ldots, v_k\}$ é um conjunto linearmente dependente, sendo $v_1, v_2, \ldots, v_k$ vectores próprios associados a valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_k$ distintos entre si. Pretendemos, desta forma, concluir um absurdo.

Seja r o menor inteiro para o qual o conjunto $\{v_1, v_2, \ldots, v_r\}$ é linearmente independente. Ora $r \geq 1$ já que $v_1 \neq 0$ (pois é v_1 é vector próprio) e $r < k$ já que o conjunto dos vectores próprios é linearmente dependente. Sendo o conjunto $\{v_1, v_2, \ldots, v_{r+1}\}$ linearmente dependente, existem escalares $\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}$ não todos nulos para os quais

$$\sum_{i=1}^{r+1} \alpha_i v_i = 0$$

o que implica que $A \sum_{i=1}^{r+1} \alpha_i v_i = \sum_{i=1}^{r+1} \alpha_i A v_i = 0$, e portanto

$$\sum_{i=1}^{r+1} \alpha_i \lambda_i v_i = 0.$$

Por outro lado, $\sum_{i=1}^{r+1} \alpha_i v_i = 0$ implica que $\lambda_{r+1} \sum_{i=1}^{r+1} \alpha_i v_i = 0$ e portanto

$$\sum_{i=1}^{r+1} \alpha_i \lambda_{r+1} v_i = 0.$$

Fazendo a diferença das duas equações, obtemos $\sum_{i=1}^{r+1} \alpha_i(\lambda_i - \lambda_{r+1})v_i = 0$, e portanto $\sum_{i=1}^{r} \alpha_i(\lambda_i - \lambda_{r+1})v_i = 0$. Como $\{v_1, v_2, \ldots, v_r\}$ é linearmente independente, segue que $\alpha_i(\lambda_i - \lambda_{r+1}) = 0$, o que implica, e visto $\lambda_i - \lambda_{r+1} \neq 0$ já que os valores próprios são distintos, que $\alpha_i = 0$, com $i = 1 \ldots, r$. Mas $\sum_{i=1}^{r+1} \alpha_i v_i = 0$, o que juntamente com as igualdades $\alpha_i = 0$, com $i = 1 \ldots, r$, leva a que $\alpha_{r+1}v_{r+1} = 0$. Como $v_{r+1} \neq 0$ já que é vector próprio, segue que $\alpha_{r+1} = 0$. Tal contradiz o facto de existirem escalares $\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}$ não todos nulos para os quais $\sum_{i=1}^{r+1} \alpha_i v_i = 0$. □

Alertamos para o facto do recíproco do teorema ser *falso*. Repare que a matriz identidade I_n tem 1 como único valor próprio, e a dimensão de $N(I_n - I_n)$ ser n, e portanto há n vectores próprios linearmente independentes associados a 1.

Se uma matriz $n \times n$ tem os seus n valores próprios distintos então, pelo teorema, tem n vectores próprios linearmente independentes, o que é equivalente a afirmar que a matriz é diagonalizável.

Corolário 5.3.4. *Uma matriz com os seus valores próprios distintos é diagonalizável.*

Mais uma vez alertamos para o facto do recíproco do corolário ser *falso*. Isto é, há matrizes diagonalizáveis que têm valores próprios com multiplicidade algébrica *superior* a 1.

Considere a matriz $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$. Esta matriz tem dois valores próprios distintos, e $\sigma(A) = \{1, 2\}$. Repare que o valor próprio 2 tem multiplicidade algébrica igual a 2, enquanto que a multiplicidade algébrica do valor próprio 1 é 1. Pelo teorema anterior, um vector próprio associado a 2 e um vector próprio associado a 1 são linearmente independentes. Repare que a multiplicidade geométrica de 2 é também 2, já que $\text{car}(2I_3 - A) = 1$ implica que $\text{nul}(2I_3 - A) = 3 - \text{car}(2I_3 - A) = 2$. Ou seja, $\dim N(2I_3 - A) = 2$, e portanto existem dois vectores próprios linearmente independentes associados a 2. Como exercício, determine uma base do espaço próprio associado a 2. Esses dois vectores, juntamente com um vector próprio associado ao valor próprio 1, formam um conjunto linearmente independente, pois vectores próprios associados a valor próprios distintos são linearmente independentes. Ou seja, há 3 vectores próprios linearmente independentes, donde segue que a matriz A é diagonalizável.

Exercícios

1. Considere a matriz $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

   (a) Calcule o polinómio característico de A.

   (b) Calcule os valores próprios de A.

   (c) Compare o determinante de A com o produto dos seus valores próprios.

(d) Compare o traço[1] de A com a soma dos seus valores próprios.

(e) Calcule os valores próprios de A^2 e compare-os com os quadrados dos valores próprios de A.

(f) Indique uma base para cada um dos espaços próprios associados aos valores próprios de A.

(g) Seja U a matriz cujas colunas são os vectores das bases que obteve na alínea anterior. Mostre que U é invertível.

(h) Calcule $U^{-1}AU$.

(i) Troque duas colunas da matriz U e efectue, de novo, o produto $U^{-1}AU$. Comente o resultado obtido.

(j) Use as alíneas anteriores para calcular A^5.

(k) Calcule B por forma a que $B^2 = A$.

2. Considere a matriz $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

   (a) Calcule o polinómio característico e mostre que B não tem valores próprios reais.

   (b) Compare o determinante de B com o produto dos seus valores próprios.

   (c) Compare o traço de B com a soma dos seus valores próprios.

   (d) Calcule os valores próprios de B^2 compare-os com os quadrados dos valores próprios de B.

   (e) Seja U a matriz cujas colunas são os vectores próprios de B linearmente independentes. Calcule $U^{-1}BU$.

   (f) Troque duas colunas da matriz U descrita na alínea anterior e efectue, de novo, o produto $U^{-1}BU$. Comente o resultado obtido.

3. Considere a matriz $C = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

   (a) Calcule o polinómio característico e os valores próprios (e a sua multiplicidade algébrica).

   (b) Calcule a dimensão dos respectivos espaços próprios.

   (c) Compare as multiplicidades algébrica e geométrica.

   (d) Mostre que a matriz não é diagonalizável.

4. Mostre que

   (a) se λ é valor próprio de A então λ^k é valor próprio de A^k;

   (b) uma matriz nilpotente não tem valores próprios não nulos.

[1] O traço de uma matriz é a soma dos seus elementos diagonais.

5. Mostre que duas matrizes semelhantes têm o mesmo espectro.

6. Para cada uma das seguintes matrizes, calcule os valores próprios e os respectivos espaços próprios (indicando uma base para os espaços próprios).

   (a) $\begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

   (d) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

   (g) $\begin{bmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{bmatrix}$ (h) $\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$

7. Calcule $\begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix}^9$.

8. Considere a matriz $A = \begin{bmatrix} 8 & 1 & 3 \\ 0 & -3 & -6 \\ 0 & 1 & 2 \end{bmatrix}$.

   (a) Determine os valores próprios de A.

   (b) Determine os vectores próprios de A e diagonalize A.

   (c) Usando o resultado da alínea (b), determine uma matriz B tal que $B^3 = A$.

9. Considere a matriz $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$.

   (a) Verifique que $5, -1$ são os valores próprios de A.

   (b) Verifique se $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio -1 e se $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio 5.

   (c) Diga, justificando, se a matriz $\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$ é invertível, e caso afirmativo calcule a sua inversa.

   (d) Diga, justificando, se a matriz A é diagonalizável, e caso afirmativo, diagonalize-a.

10. Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

   (a) Verifique que $3, -1$ são os valores próprios de A.

   (b) Verifique se $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio 3 e se $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio -1.

(c) Diga, justificando, se a matriz $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ é invertível, e caso afirmativo calcule a sua inversa.

(d) Diga, justificando, se a matriz A é diagonalizável, e caso afirmativo, diagonalize-a.

11. Considere a matriz $A = \begin{bmatrix} 2 & -2 \\ 2 & -3 \end{bmatrix}$.

    (a) Verifique que $1, -2$ são os valores próprios de A.

    (b) Verifique se $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio 1 e se $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ é vector próprio associado ao valor próprio -2.

    (c) Diga, justificando, se a matriz $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ é invertível, e caso afirmativo calcule a sua inversa.

    (d) Diga, justificando, se a matriz A é diagonalizável, e caso afirmativo, diagonalize-a.

12. Considere a matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

    (a) Verifique que $2, -3$ são os valores próprios de A.

    (b) Verifique se $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ é vector próprio associado ao valor próprio -3 e se $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ é vector próprio associado ao valor próprio 2.

    (c) Diga, justificando, se a matriz $\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$ é invertível, e caso afirmativo calcule a sua inversa.

    (d) Diga, justificando, se a matriz A é diagonalizável, e caso afirmativo, diagonalize-a.

13. Dada a matriz real $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$,

    (a) calcule os valores próprios e respectivos espaços próprios;

    (b) verifique que a matriz dada não é diagonalizável.

Capítulo 6

Transformações lineares

6.1 Definição e exemplos

Definição 6.1.1. *Sejam V, W espaços vectoriais sobre $\mathbb{K}$. Uma* transformação linear *ou aplicação linear de V em W é uma função $T : V \to W$ que satisfaz, para $u, v \in V, \alpha \in \mathbb{K}$,*

1. *$T(u + v) = T(u) + T(v)$;*

2. *$T(\alpha u) = \alpha T(u)$.*

Para $F, G : \mathbb{R}^2 \to \mathbb{R}^4$ definidas por

$$F(x, y) = (x - y, 2x + y, 0, y)$$

e

$$G(x, y) = (x^2 + y^2, 1, |x|, y),$$

tem-se que F é linear enquanto G não o é. De facto, para $u_1 = (x_1, y_1), u_2 = (x_2, y_2)$ e $\alpha \in \mathbb{K}$, temos $F(u_1 + u_2) = F(x_1 + x_2, y_1 + y_2) = (x_1 + x_2 - y_1 - y_2, 2x_1 + 2x_2 + y_1 + y_2, 0, y_1 + y_2) = (x_1 - y_1, 2x_1 + y_1, 0, y_1) + (x_2 - y_2, 2x_2 + y_2, 0, y_2) = F(u_1) + F(u_2)$, e $F(\alpha u_1) = F(\alpha x_1, \alpha y_1) = (\alpha x_1 - \alpha y_1, 2\alpha x_1 + \alpha y_1, 0, \alpha y_1) = \alpha(x_1 - y_1, 2x_1 + y_1, 0, y_1) = \alpha F(u_1)$, enquanto que $G(-(1, 1)) = G(-1, -1) = ((-1)^2 + (-1)^2, 1, |-1|, -1) = (2, 1, 1, -1) \neq -(2, 1, 1, 1) = -G(1, 1)$

Apresentamos alguns exemplos clássicos de transformações lineares:

1. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ e $T_A : \mathbb{K}^n \to \mathbb{K}^m$ definida por $T_A(x) = Ax$. A aplicação T_A é uma transformação linear. Ou seja, dada uma matriz, existe uma transformação linear associada a ela. No entanto, formalmente são entidades distintas. Mais adiante, iremos ver que qualquer transformação linear está associada a uma matriz.

2. Seja V um espaço vectorial arbitrário sobre $\mathbb{K}$. As aplicações $I, O : V \to V$ definidas por $I(v) = v$ e $O(v) = 0$ são transformações lineares. Denominam-se, respectivamente, por transformação identidade e transformação nula.

Definição 6.1.2. *Seja T uma transformação linear do espaço vectorial V para o espaço vectorial W.*

1. *Se $V = W$, diz-se que T é um* endomorfismo *de V.*
2. *A um homomorfismo injectivo de V sobre W chama-se* monomorfismo *de V sobre W; a um homomorfismo sobrejectivo de V sobre W chama-se* epimorfismo *de V sobre W; a um homomorfismo bijectivo de V sobre W chama-se* isomorfismo *de V sobre W; a um endomorfismo bijectivo de V chama-se* automorfismo *de V.*
3. *V e W são ditos* isomorfos*, e representa-se por $V \cong W$, se existir uma transformação linear de V em W que seja um isomorfismo.*

6.2 Propriedades das transformações lineares

Proposição 6.2.1. *Sejam V, W espaços vectoriais sobre $\mathbb{K}$ e $T\colon V \to W$ uma transformação linear. Então*

1. *$T(0_v) = 0_w$ para $0_v \in V,\ 0_w \in W$;*
2. *$T(-v) = -T(v), \forall v \in V$;*
3. *$T(\sum_{i=0}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i T(v_i),\ v_i \in V,\ \alpha_i \in \mathbb{K}$;*
4. *Se $v_1, v_2, ..., v_n$ são vectores de V linearmente dependentes, então $T(v_1), T(v_2), \ldots, T(v_n)$ são vectores de W linearmente dependentes.*

Demonstração. As afirmações 1–3 seguem da definição de transformação linear. Mostremos (4).

Se $v_1, v_2, ..., v_n$ são vectores de V linearmente dependentes então um deles, digamos v_k, escreve-se como combinação linear dos restantes:

$$v_k = \sum_{i=0, i\neq k}^{n} \alpha_i v_i.$$

Aplicando T a ambos os membros da equação,

$$T(v_k) = T\left(\sum_{i=0, i\neq k}^{n} \alpha_i v_i\right) = \sum_{i=1, i\neq k}^{n} \alpha_i T(v_i),$$

e portanto $T(v_k)$ escreve-se como combinação linear de $T(v_1), T(v_2), T(v_{k-1}), \ldots, T(v_{k+1}), \ldots, T(v_n)$. Segue que $T(v_1), T(v_2), \ldots, T(v_n)$ são vectores de W linearmente dependentes. □

Em geral, uma transformação **não** preserva a independência linear. Por exemplo, a transformação linear

$$\begin{array}{rccc} & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ T: & (x, y) & \longrightarrow & (0, y). \end{array}$$

As imagens da base canónica de $\mathbb{R}^2$ não são linearmente independentes.

Recordamos que, apesar de indicarmos uma base como um conjunto de vectores, é importante a ordem pela qual estes são apresentados. Ou seja, uma base é um n-uplo de vectores. Por forma a não ser confundida por um n-uplo com entradas reais, optámos por indicar uma base como um conjunto. É preciso enfatizar esta incorrecção (propositadamente) cometida.

Teorema 6.2.2. *Sejam V, W espaços vectoriais com dimensão finita,* $\dim V = n = \dim W$, $\{v_1, \ldots, v_n\}$ *uma base de V e $w_1, \ldots, w_n \in W$. Então existe uma única transformação linear $T : V \to W$ tal que*

$$T(v_1) = w_1, T(v_2) = w_2, \ldots, T(v_n) = w_n.$$

Demonstração. Se $\{v_1, \ldots, v_n\}$ é uma base de V, então todo o elemento de v escreve-se de forma única como combinação linear de $v_1, \ldots, v_n$. Isto é, para qualquer $v \in V$, existem $\alpha_i \in \mathbb{K}$ tais que

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Seja $T : V \to W$ definida por

$$T\left(\sum_i \alpha_i v_i\right) = \sum_i \alpha_i w_i.$$

Obviamente, $T(v_i) = w_i$. Observe-se que T é de facto uma aplicação pela unicidade dos coeficientes da combinação linear relativamente à base. Mostre-se que T assim definida é linear. Para $\alpha \in \mathbb{K}$, $u = \sum_i \beta_i v_i$ e $w = \sum_i \gamma_i v_i$,

$$\begin{aligned}
T(u+w) &= T\left(\sum_i \beta_i v_i + \sum_i \gamma_i v_i\right) \\
&= T\left(\sum_i (\beta_i + \gamma_i) v_i\right) \\
&= \sum_i (\beta_i + \gamma_i) w_i \\
&= \sum_i \beta_i w_i + \sum_i \gamma_i w_i \\
&= T(u) + T(w)
\end{aligned}$$

e

$$\begin{aligned}
T(\alpha u) &= T(\alpha \sum_i \beta_i v_i) \\
&= T(\sum_i \alpha \beta_i v_i) \\
&= \sum_i \alpha \beta_i w_i \\
&= \alpha \sum_i \beta_i w_i = \alpha T(u).
\end{aligned}$$

Portanto, T assim definida é linear.

Mostre-se, agora, a unicidade. Suponhamos que T' é uma aplicação linear que satisfaz $T'(v_i) = w_i$, para todo o i no conjunto dos índices. Seja $v \in V$, com $v = \sum_i \alpha_i v_i$. Então

$$\begin{aligned} T'(v) &= T'\left(\sum_i \alpha_i v_i\right) \\ &= \sum_i T'(v_i) \\ &= \sum_i \alpha_i w_i \\ &= \sum_i \alpha_i T(v_i) \\ &= T\left(\sum_i \alpha_i v_i\right) = T(v). \end{aligned}$$

Portanto, $T = T'$. □

Teorema 6.2.3. *Todo o espaço vectorial de dimensão n sobre o corpo $\mathbb{K}$ é isomorfo a $\mathbb{K}^n$.*

Demonstração. Seja $\{v_1, v_2, ..., v_n\}$ uma base de V e v um vector qualquer de V. Então $v = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n,\ \alpha_i \in \mathbb{K}$. Vamos definir uma transformação T,

$$T: \begin{array}{ccc} V & \longrightarrow & \mathbb{K}^n \\ v & \longrightarrow & (\alpha_1, \alpha_2, ..., \alpha_n) \end{array}.$$

Pretendemos mostrar que esta aplicação é um isomorfismo de espaços vectoriais.

(a) A aplicação T é bijectiva.

Primeiro, verificamos que T é injectiva, i.e., que

$$T(u) = T(v) \Longrightarrow u = v,\ \forall\ u, v \in V.$$

Ora,

$$\begin{aligned} T(u) = T(v) &\iff T(\sum_{i=1}^n \alpha_i v_i) = T(\sum_{i=1}^n \beta_i v_i) \\ &\iff (\alpha_1, \alpha_2, ..., \alpha_n) = (\beta_1, \beta_2, ..., \beta_n) \\ &\iff \alpha_i = \beta_i \\ &\iff \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \beta_i v_i \\ &\iff u = v. \end{aligned}$$

Mostramos, agora, que T é sobrejectiva, i.e., que

$$\forall x \in \mathbb{K}^n, \exists w \in V : f(w) = x.$$

Temos sucessivamente,

f é sobrejectiva $\iff \forall x \in \mathbb{K}^n, \exists w \in V : f(w) = x \iff \forall(\delta_1, \delta_2, ..., \delta_n) \in \mathbb{K}^n, \exists w = \delta_1 v_1 + \delta_2 v_2 + ... + \delta_n v_n \in V : T(\delta_1 v_1 + \delta_2 v_2 + ... + \delta_n v_n) = (\delta_1, \delta_2, ..., \delta_n)$.

(b) A aplicação T é linear.

$$\begin{aligned}
T(u+v) &= T(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n + \beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n) \\
&= T[(\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + ... + (\alpha_n + \beta_n)v_n] \\
&= (\alpha_1 + \beta_1, \alpha_2 + \beta_2, ..., \alpha_n + \beta_n) \\
&= (\alpha_1, \alpha_2, ..., \alpha_n) + (\beta_1, \beta_2, ..., \beta_n) \\
&= T(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) + T(\beta_1 v_1 + \beta_2 v_2 + ... + \beta_n v_n) \\
&= T(u) + T(v)
\end{aligned}$$

e

$$\begin{aligned}
T(\alpha u) &= T\left(\alpha(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n)\right) \\
&= T\left((\alpha\alpha_1)v_1 + (\alpha\alpha_2)v_2 + ... + (\alpha\alpha_n)v_n\right) \\
&= (\alpha\alpha_1, \alpha\alpha_2, ..., \alpha\alpha_n) \\
&= \alpha(\alpha_1, \alpha_2, ..., \alpha_n) \\
&= \alpha T(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) \\
&= \alpha T(u)
\end{aligned}$$

□

Por exemplo, o subespaço $U = \langle(1,1,0),(0,1,1)\rangle$ de $\mathbb{R}^3$ é isomorfo a $\mathbb{R}^2$. De facto, $\dim U = 2$ e portanto, como U é espaço vectorial sobre $\mathbb{R}$, tem-se U isomorfo a $\mathbb{R}^2$

Corolário 6.2.4. *Sejam U e V dois espaços vectoriais sobre mesmo corpo $\mathbb{K}$. Se U e V têm a mesma dimensão, então U e V são isomorfos.*

O subespaço U apresentado atrás é isomorfo a $W = \langle(0,1,0,1),(1,0,0,1)\rangle$ já que $\dim U = \dim W$.

Terminamos esta secção com uma observação que será útil na forma como se define uma transformação linear. É bem conhecido que o conhecimento da imagem de um certo número de objectos por uma função arbitrária não é suficiente para se conhecer a imagem de *todos* os objectos. As transformações lineares são funções especiais que satisfazem a aditividade e homogeneidade. Estas duas condições permitem-nos conhecer a imagem de todos os objectos conhecendo, à partida, apenas a de alguns. De facto, se V e W são espaços vectoriais de dimensão finita, e $v_1, \ldots, v_n$ uma base de V, então $T : V \to W$ fica bem definida se se souber os valores de $T(v_1), T(v_2), \ldots, T(v_n)$. Se $v \in V$ então existem $\alpha_i \in \mathbb{K}$, únicos, para os quais $v = \sum \alpha_i v_i$. Sabendo que a imagem por T de uma soma é a soma das imagens, e que a imagem de um múltiplo de um vector é o múltiplo da imagem do vector, obtemos

$$T(v) = T(\sum \alpha_i v_i) = \sum T(\alpha_i v_i) = \sum \alpha_i T(v_i).$$

Usando o exposto acima, a transformação linear

$$T : U = \langle (1,1,0), (0,1,1) \rangle \to W = \langle (0,1,0,1), (1,0,0,1) \rangle$$

definida por

$$T(1,1,0) = (0,1,0,1) \text{ e } T(0,1,1) = (1,0,0,1)$$

é um isomorfismo entre U e W. Para sabermos a imagem por T de $(1,3,2)$ basta escrever $(1,3,2) = 1(1,1,0) + 2(0,1,1)$ e notar que

$$T(1,3,2) = 1T(1,1,0) + 2T(0,1,1) = (0,1,0,1) + 2(1,0,0,1) = (2,1,0,3).$$

6.3 Matriz associada a uma transformação linear

Iremos concluir que todas as transformações lineares de $\mathbb{K}^n \to \mathbb{K}^m$ podem ser representadas por matrizes do tipo $m \times n$. Como motivação, consideramos alguns exemplos.

Sejam e_1, e_2, e_3 elementos da base canónica B_1 de $\mathbb{R}^3$ e e_1^*, e_2^* os elementos da base canónica B_1^* de $\mathbb{R}^2$. Seja ainda $T : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$\begin{aligned} T(e_1) &= a_{11}e_1^* + a_{21}e_2^* \\ T(e_2) &= a_{12}e_1^* + a_{22}e_2^* \\ T(e_3) &= a_{13}e_1^* + a_{23}e_2^*. \end{aligned}$$

Recorde que a transformação linear está bem definida à custa das imagens dos vectores de uma base.

Se $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, então

$$\begin{aligned} T(x) &= T(x_1e_1 + x_2e_2 + x_3e_3) = x_1T(e_1) + x_2T(e_2) + x_3T(e_3) \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = Ax \end{aligned}$$

Por outras palavras, a transformação linear definida atrás pode ser representado à custa de uma matriz $A \in \mathcal{M}_{2\times 3}(\mathbb{R})$, que tem como colunas as coordenadas em relação a B_1^* das imagens dos vectores $e_i \in \mathbb{R}^3, i = 1, 2, 3$ por T. Desta forma, dizemos que nas condições do exemplo anterior, a matriz A é a representação matricial de T relativamente às bases canónicas de $\mathbb{R}^2$ e $\mathbb{R}^3$.

Por exemplo, considere a aplicação linear T definida por

$$\begin{aligned} T(1,0,0) &= (4,-1) = 4(1,0) - 1(0,1) \\ T(0,1,0) &= (-2,5) = -2(1,0) + 5(0,1) \\ T(0,0,1) &= (3,-2) = 3(1,0) - 2(0,1) \end{aligned}$$

A matriz que representa T em relação às bases canónicas de $\mathbb{R}^3$ e $\mathbb{R}^2$ é $A = \begin{bmatrix} 4 & -2 & 3 \\ -1 & 5 & -2 \end{bmatrix}$. Para todo $v \in \mathbb{R}^3$, $T(v) = Av$.

Repare que os cálculos envolvidos foram simples de efectuar já que usámos as bases canónicas dos espaços vectoriais. Tal não será, certamente, o caso se usarmos outras bases que não as canónicas. Neste caso, teremos que encontrar as coordenadas das imagens dos elementos da base do primeiro espaço vectorial em relação à base fixada previamente do segundo espaço vectorial. Vejamos o exemplo seguinte:

Sejam $\{u_1, u_2, u_3\}$ base B_2 de $\mathbb{R}^3$ e $\{v_1, v_2\}$ base B_2^* de $\mathbb{R}^2$. Se $x \in \mathbb{R}^3$, então $x = \xi_1 u_1 + \xi_2 u_2 + \xi_3 u_3$, e consequentemente

$$T(x) = \xi_1 T(u_1) + \xi_2 T(u_2) + \xi_3 T(u_3).$$

Por outro lado, $T(u_i) \in \mathbb{R}^2, i = 1, 2, 3$, logo, podemos escrever estes vectores como combinação linear de v_1 e v_2. Assim,

$$\begin{array}{rcl} T(u_1) & = & b_{11}v_1 + b_{21}v_2 \\ T(u_2) & = & b_{12}v_1 + b_{22}v_2 \\ T(u_3) & = & b_{13}v_1 + b_{23}v_2. \end{array}$$

Verificamos, então, que,

$$\begin{array}{rcl} T(x) & = & \xi_1(b_{11}v_1 + b_{21}v_2) + \xi_2(b_{12}v_1 + b_{22}v_2) + \xi_3(b_{13}v_1 + b_{23}v_2) \\ & = & (\xi_1 b_{11} + \xi_2 b_{12} + \xi_3 b_{13} v_1) + (\xi_1 b_{21} + \xi_2 b_{22} + \xi_3 b_{23} v_2) \\ & = & \alpha_1 v_1 + \alpha_2 v_2 \end{array}$$

onde

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

Dizemos, agora, que $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$ é a matriz de T relativamente às bases B_2 e B_2^* de $\mathbb{R}^3$ e $\mathbb{R}^2$, respectivamente.

Passamos de seguida a expôr o caso geral.

Sejam $B_1 = \{u_1, u_2, ..., u_n\}$ uma base de U, $B_2 = \{v_1, v_2, ..., v_m\}$ uma base de V, e

$$T \ : \ U \to V$$

$$x \to T(x) = \sum_{j=1}^{n} x_j T(u_j)$$

uma transformação linear, sabendo que as coordenadas de x na base B_1 é $(x_1, x_2, \ldots, x_n)$.

O vector $T(u_j)$ pode ser escrito – de modo único – como combinação linear dos vectores $v_1, v_2, ..., v_m$. Assim

$$T(u_j) = \sum_{i=1}^{m} a_{ij} \cdot v_i, \quad j = 1, ..., n.$$

Logo

$$T(x) = \sum_{j=1}^{n} x_j T(u_j) = \sum_{j=1}^{n} [\sum_{i=1}^{m} a_{ij} v_i] = \sum_{j=1}^{n} [a_{1j}x_j]v_1 + \sum_{j=1}^{n} [a_{2j}x_j]v_2 + ... + \sum_{j=1}^{n} [a_{mj}x_j]v_m =$$

$\sum_{i=1}^{n} \varphi_i v_i$, com $\varphi_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, ..., m.$

Verificamos, assim, que existe entre as coordenadas $(x_1, x_2, ..., x_n)$ de x (relativa à base B_1), em U, e as coordenadas $(\varphi_1, \varphi_2, ..., \varphi_m)$ de $T(x)$ (relativa à base B_2) em V. Tal ligação exprime-se pelas seguintes equações

$$\varphi_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, ..., m.$$

O que se pode ser escrito como a equação matricial seguinte:

$$\begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & b_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & b_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_m \end{bmatrix}$$

Assim, concluímos:

Teorema 6.3.1. *Se fixamos uma base de U e uma base de V, a aplicação linear $T : U \longrightarrow V$ fica perfeitamente definida por $m \times n$ escalares. Ou seja, a aplicação linear $T : U \longrightarrow V$ fica perfeitamente definida por uma matriz do tipo $m \times n$*

$$M_{B_1,B_2}(T)$$

cujas colunas são as coordenadas dos transformados dos vectores da base de U, em relação à base de V.

Vimos, então, que dada uma transformação linear $G : \mathbb{K}^n \to \mathbb{K}^m$, existe uma matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$ tal que $G = T_A$. Mais, se $\{e_1, \ldots, e_n\}$ e $\{f_1, \ldots, f_m\}$ são as bases canónicas, respectivamente, de $\mathbb{K}^n$ e $\mathbb{K}^m$, então a matriz A é tal que a coluna i de A são as coordenadas de $G(e_i)$ em relação à base $\{f_1, \ldots, f_m\}$. No entanto, se se considerarem bases que não as canónicas, então é preciso ter um pouco mais de trabalho.

Por exemplo, considere[1] a base B_1 de $\mathbb{R}^3$ constituída pelos vectores $(0, 1, 1), (1, 1, 0), (1, 0, 1)$, e a base B_2 de $\mathbb{R}^2$ constituída pelos vectores $(2, 1), (1, 2)$. Vamos calcular a matriz G que representa $T : \mathbb{R}^3 \to \mathbb{R}^2$, com $T(x, y, z) = (x - y, x + y + z)$, nas bases apresentadas. Em primeiro lugar, calculamos as imagens dos elementos da base escolhida:

$$\begin{aligned} T(0, 1, 1) &= (-1, 2) = v_1 \\ T(1, 1, 0) &= (0, 2) = v_2 \\ T(1, 0, 1) &= (1, 2) = v_3 \end{aligned}$$

[1] Verifique que de facto formam uma base!

Agora, encontramos as coordenadas de v_1, v_2, v_3 relativamente à base de $\mathbb{R}^2$ que fixámos. Ou seja, encontramos as soluções dos sistemas possíveis determinados[2]

$$Ax = v_1,\ Ax = v_2,\ Ax = v_3,$$

onde $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. A matriz que representa T em relação às bases apresentadas é $G = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$, onde u_1 é a única solução de $Ax = v_1$, u_2 é a única solução de $Ax = v_2$ e u_3 é a única solução de $Ax = v_3$.

Fixadas as bases dos espaços vectoriais envolvidos, a matriz associada à transformação linear G será, doravante, denotada por $[G]$.

Antes de passarmos ao resultado seguinte, consideremos as transformações lineares

$$\begin{array}{rccc} H: & \mathbb{R}^2 & \to & \mathbb{R}^3 \\ & (x,y) & \mapsto & (x-y, y, 0) \end{array}, \qquad \begin{array}{rccc} G: & \mathbb{R}^3 & \to & \mathbb{R} \\ & (r,s,t) & \mapsto & 2r - s + t \end{array}$$

Obtemos, então,

$$\begin{aligned} G \circ H(x,y) &= 2(x-y) - 1 \cdot y + 1 \cdot 0 \\ &= \begin{bmatrix} 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x-y \\ y \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= [G][H] \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

Portanto, $[G \circ H] = [G][H]$.

Vejamos o que podemos afirmar em geral:

Teorema 6.3.2. *Sejam U, V, W espaços vectoriais sobre $\mathbb{K}$ e $H : U \to V$, $G : V \to W$ duas transformações lineares. Então*

1. *$G \circ H$ é uma transformação linear;*
2. *$G \circ H = T_{[G][H]}$ e $[G \circ H] = [G]\,[H]$.*

Demonstração. A demonstração de (1) fica como exercício. Para mostrar (2), observe-se que, para qualquer $u \in U$,

$$G \circ H(u) = G(H(u)) = G([H]u) = [G][H]u = T_{[G][H]}u.$$

□

[2] Consegue explicar por que razão os sistemas são possíveis determinados?

Terminamos, assim, como iniciámos: a algebrização do conjunto das matrizes. As matrizes não são mais do que representantes de um certo tipo de funções (as transformações lineares) entre conjuntos muitos especiais (espaços vectoriais). Se a soma de matrizes corresponde à soma de transformações lineares (em que a soma de funções está definida como a função definida pela soma das imagens), o produto de matrizes foi apresentado como uma operação bem mais complicada de efectuar. No entanto, a forma como o produto matricial foi definido corresponde à *composição* das transformações lineares definidas pelas matrizes.

Exercícios

1. Seja $T : V \to W$ uma transformação linear entre os espaços vectoriais V e W. Mostre que

   (a) $T(0) = 0$.

   (b) Para $v_1, ..., v_n \in V$, $\alpha_1, ..., \alpha_n \in \mathbb{K}$,
   $T(\alpha_1 v_1 + ... + \alpha_n v_n) = \alpha_1 T(v_1) + ... + \alpha_n(v_n)$.

2. Diga quais das aplicações seguintes, entre espaços vectoriais reais, são transformações lineares:

   (a) $F : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $F(x, y) = (2x + y, x, y - x)$, $\forall (x, y) \in \mathbb{R}^2$.

   (b) $G : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $G(x, y, z) = (y^2, y)$, $\forall (x, y, z) \in \mathbb{R}^3$.

   (c) $T : \mathbb{R}^2 \to \mathbb{R}$ definida por $T(a, b) = 5a - 2b$, $\forall\ (a, b) \in \mathbb{R}^2$.

3. Diga, justificando, se existe

   (a) uma transformação linear $F : \mathbb{R}^3 \to \mathbb{R}^3$ tal que
   $$F(1, 0, 0) = (0, 0, 1),\ F(0, 0, 1) = (1, 0, 0),\ F(7, 0, 14) = (0, 0, 7).$$

   (b) uma transformação linear $G : \mathbb{R}^2 \to \mathbb{R}^4$ tal que
   $$G(-1, 2) = (0, 1, 2, 3),\ G(2, -1) = (0, -1, -2, -3).$$

4. Considere as transformações lineares

   $F : \mathbb{R}^4 \to \mathbb{R}^3$ definida por $F(x, y, z, w) = (x - y, x + w, y + z)$, $\forall (x, y, z, w) \in \mathbb{R}^4$;

   $G : \mathbb{R}^4 \to \mathbb{R}^4$ definida por $G(x, y, z, w) = (x, x + z, -w, 2y + z)$, $\forall (x, y, z, w) \in \mathbb{R}^4$;

   $H : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $H(1, 1, 1) = (2, 0, 1)$, $H(1, 1, 0) = (1, 0, -1)$, $H(1, 0, 0) = (0, 0, 2)$;

   $T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por $T(x, y, z) = (x - y, 0, 0, x + y + z)$, $\forall (x, y, z) \in \mathbb{R}^3$.

   Determine, em relação às bases canónicas:
   i) $[F]$. ii) $[G]$. iii) $[H]$. iv) $[T]$. v) $[H \circ F]$.

5. Considere a transformação linear $F : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

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

Determine a matriz que representa F nas bases $B_1 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ e na canónica.

6. Seja $G : \mathbb{R}^3 \to \mathbb{R}^4$ a transformação linear definida por

$$G(1, 0, 0) = (1, 0, 1, 0),\ G(0, 1, 0) = (0, 1, -2, 0),\ G(0, 0, 1) = (1, 1, 0, 0).$$

(a) Determine i) $G(2, 3, 1)$. ii) $G(-1, 2, 0)$.

(b) Para as bases

$B_1 = \{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}$ e $B_2 = \{(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)\}$,

determine a matriz que representa G nessas bases.

7. Considere, no espaço vectorial real $\mathbb{R}^3$, a base $\{(1, 0, 1), (1, 1, 0), (0, 1, 1)\}$ e o endomorfismo F definido por $F(1, 0, 1) = (1, -1, 1)$, $F(1, 1, 0) = (2, 1, 1)$, $F(0, 1, 1) = (1, 0, 0)$.

(a) Determine $[F]$.

(b) Determine $F(a, b, c)$, $\forall (a, b, c) \in \mathbb{R}^3$.

8. Considere as transformações lineares

$F : \mathbb{R}^4 \to \mathbb{R}^3$ definida por $F(x, y, z, w) = (x - y, x + w, y + z)$, $\forall (x, y, z, w) \in \mathbb{R}^4$;

$G : \mathbb{R}^4 \to \mathbb{R}^4$ definida por $G(x, y, z, w) = (x, x + z, -w, 2y + z)$, $\forall (x, y, z, w) \in \mathbb{R}^4$;

$H : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $H(1, 1, 1) = (2, 0, 1)$, $H(1, 1, 0) = (1, 0, -1)$, $H(1, 0, 0) = (0, 0, 2)$;

$T : \mathbb{R}^3 \to \mathbb{R}^4$ definida por $T(x, y, z) = (x - y, 0, 0, x + y + z)$, $\forall (x, y, z) \in R^3$;

e as bases $B_1 = \{(0, 0, 0, 1), (0, 0, 1, 1), (0, 1, 1, 1), (1, 1, 1, 1)\}$ e

$B_2 = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$ de $\mathbb{R}^4$;

$B'_1 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ e $B'_2 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ de $\mathbb{R}^3$.

Determine

a) $[F]_{B_1,B'_1}$. b) $[F]_{B_1,B'_2}$. c) $[G]_{B_2,B_1}$. d) $[G]_{B_1,B_2}$

e) $[H]_{B'_1,B'_2}$. f) $[H]_{B'_1,B'_1}$ g) $[T]_{B'_2,B_2}$. h) $[T]_{B'_1,B_1}$.

Bibliografia

[1] F. R. Dias Agudo, *Introdução à álgebra linear e geometria analítica*, Escolar Editora, 1996.

[2] Howard Anton, Chris Rorres, *Elementary linear algebra: applications version*, John Wiley & Sons, 1994.

[3] Kenneth J. Beers, *Numerical Methods for Chemical Engineering, Applications in Matlab®*, Cambridge University Press, 2007.

[4] I. S. Duff, A. M. Erisman, J. K. Reid, *Direct methods for sparse matrices*, Oxford University Press, 1989.

[5] Bruce A. Finlayson, *Introduction to Chemical Engineering Computing*, Wiley, 2006.

[6] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence, *Linear Algebra (2nd edition)*, Prentice-Hall International, Inc., 1989.

[7] Emília Giraldes, Vitor Hugo Fernandes, M. Paula Marques Smith, *Curso de álgebra linear e geometria analítica*, McGraw-Hill, 1995.

[8] David R. Hill, David E. Zitarelli, *Linear algebra labs with MATLAB*, Prentice Hall, 1996

[9] Roger Horn, Charles Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[10] Peter Lancaster, Miron Tismenetsky, *The Theory of Matrices*, second edition with applications, Academic Press, 1985.

[11] Luís T. Magalhães, *Álgebra Linear como introdução à matemática aplicada*, IST, 1987.

[12] J. M. Powers, *Method of least squares*, University of Notre Dame, 2003, `http://www.nd.edu/~powers/ame.332/leastsquare/leastsquare.pdf`

[13] Ana Paula Santana, João Filipe Queiró, *Álgebra Linear e Geometria Analítica*, Departamento de Matemática, Universidade de Coimbra, 2003.

[14] Gilbert Strang, *Linear algebra and its applications*, Academic Press, 1976.

[15] Maria Raquel Valença, *Métodos numéricos*, INIC, 1990.