Contents

Chapter 1

Matrizes e sistemas de equações lineares

1.1 CORPOS

Um *corpo* é uma estrutura algébrica do tipo $\mathbb{K} = (K, \oplus, \odot, \mathbb{0}, \mathbb{1})$ onde K é um conjunto não vazio $\mathbb{0}$ e $\mathbb{1}$ são elementos fixos de K e, $\oplus$, $\odot$ são duas operações binárias em K i.e., fazem corresponder a cada par (a, b) de elementos de K um elemento de K que, no primeiro caso se denota por $a \oplus b$ e no segundo $a \odot b$. Além disso devem ser verdadeiras as seguintes propriedades (axiomas de corpo):

1. $a \oplus b = b \oplus a$, quaisquer que sejam $a, b \in K$;
2. $(a \oplus b) \oplus c = a \oplus (b \oplus c)$, quaisquer que sejam $a, b, c \in K$;
3. $a \oplus \mathbb{0} = a$, para qualquer $a \in K$;
4. para cada $a \in K$ existe um único elemento que denotamos a^- que satisfaz $a \oplus a^- = \mathbb{0}$;
5. $a \odot b = b \odot a$, para quaisquer $a, b \in K$;
6. $(a \odot b) \odot c = a \odot (n \odot c)$, para quaisquer $a, b, c \in K$;
7. $a \odot \mathbb{1} = a$, para qualquer $a \in K$;
8. para cada $a \in K \setminus \{\mathbb{0}\}$ existe um único elemento que denotamos a^{-1} que satisfaz $a \odot a^{-1} = \mathbb{1}$;
9. $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$, para quaisquer $a, b, c \in K$.

Exemplos que corpos são $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, os conjuntos dos números racionais, reais e complexos, respectivamente, equipados com as as operações de adição e produto usuais, e considerando $\mathbb{0} := 0$ e $\mathbb{1} := 1$.

Se $n \in \mathbb{N} \setminus \{0\}$, então a expressão $n\mathbb{1}$ denota a *soma*

$$\underbrace{\mathbb{1} \oplus \mathbb{1} \oplus \cdots \oplus \mathbb{1}}_{n \text{ vezes}}.$$

Se num corpo $\mathbb{K}$, para cada $n \in \mathbb{N} \setminus \{0\}$, se tem $n\mathbb{1} \neq \mathbb{0}$ dizemos que o corpo tem *característica* 0. É o caso dos corpos $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$ que mencionámos anteriormente.

A *característica* de um corpo $\mathbb{K}$ ou é zero, nas circunstâncias que acabámos de descrever, ou então, se existir algum $n \in \mathbb{N} \setminus \{0\}$ tal que $n\mathbb{1} = \mathbb{0}$ é o menor natural não nulo nessas circunstâncias. É possível demonstrar que a característica de um corpo, ou é zero, ou então é um número primo.

Um facto interessante é que existem corpos de todas as características. Se p é um número primo então existem corpos de característica p, um exemplo é o corpo que se denota por $\mathbb{Z}_p$ e é denominado *corpo dos inteiros módulo p*. Tem-se que

$$\mathbb{Z}_p = (\{0, 1, \dots, p-1\}, \oplus_p, \odot_p, 0, 1),$$

onde as operações $\oplus_p$ e $\odot_p$ se definem de acordo com o seguinte $m \oplus_p n$ é o resto da divisão inteira de $m+n$ por p e $m \odot_p n$ é o resto da divisão inteira de mn por p. Por exemplo, $\mathbb{Z}_2 = (\{0, 1\}, \oplus_2, \odot_2, 0, 1)$ e as operações são:[1]

$\oplus_2$	0	1
0	0	1
1	1	0

$\odot_2$	0	1
0	0	0
1	0	1

[1] Será instrutivo para o leitor verificar que as leis 1.–9., acima são válidas nesta estrutura.

Em álgebra linear estudamos uma variedade de estruturas e objectos matemáticos e.g., espaços lineares e matrizes, que utilizam como objectos básicos os elementos de um determinado corpo. *Ao longo deste curso consideraremos essencialmente os corpos* $\mathbb{R}$ *e* $\mathbb{C}$. Desta forma recorremos à letra $\mathbb{K}$ para indicar um corpo que deve ser entendido como $\mathbb{R}$ ou $\mathbb{C}$. Em casos onde um corpo diferente destes deva ser considerado, essa excepção será devidamente assinalada. No contexto da álgebra linear os elementos de $\mathbb{K}$ designam-se de *escalares*.

1.2 MATRIZES E ÁLGEBRA DE MATRIZES

Uma matriz é uma função $A : \{1, \dots, m\} \times \{1, \dots, n\} \to \mathbb{K}$ onde consideramos $m, n \in \mathbb{N} \setminus \{0\}$. Sendo assim do ponto de vista mais formal, é no entanto mais conveniente descrever as matrizes como *quadros* compostos de linhas e colunas.[2] No caso de uma matriz como acima, trata-se de um quadro com m linhas e n-colunas, onde na posição determinada pela linha i e pela coluna j se dispõe o elemento $A(i, j)$ que se denomina

[2] As matrizes são objectos finitos e este tipo de apresentação permite visualizar facilmente a totalidade da informação contida numa matriz.

de *entrada-i,j* da matriz A. Ou seja:

$$A = \begin{bmatrix} A(1,1) & A(1,2) & \cdots & A(1,n) \\ A(2,1) & A(2,2) & \cdots & A(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ A(m,1) & A(m,2) & \cdots & A(m,n) \end{bmatrix}.$$

Em geral, em lugar de escrevermos $A(i,j)$ escreveremos A_{ij} (ou $A_{i,j}$). Uma matriz A com m linhas e n colunas diz-se que é do tipo $m \times n$ e o conjunto matrizes $m \times n$ cujas entradas são elementos de um corpo $\mathbb{K}$ denota-se por $\mathbb{K}^{m \times n}$.

As matrizes representam-se em geral por letras maiúsculas com a excepção das matrizes dos tipos $1 \times n$ e $n \times 1$ que, por razões mais ou menos claras se denominam de *vectores coluna* e *vectores linha*, respectivamente, que denotaremos por letras minúsculas, e.g. A, b, c, … , u, v, x, y, ….

Observe-se que duas funções f, g são iguais se $\mathrm{dom}(f) = \mathrm{dom}(g)$ e se, para cada $x \in \mathrm{dom}(f) = \mathrm{dom}(g)$ se tem $f(x) = g(x)$. Traduzindo este facto para o contexto das matrizes temos o seguinte critério de igualdade de matrizes:

CRITÉRIO DE IGUALDADE DE MATRIZES.—*Sejam , B matrizes. Tem-se $A = B$ se e só se, A e B são de um mesmo tipo $m \times n$ e se, para cada $(i,j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$ se tem $A_{ij} = B_{ij}$, ou seja, se e só se as entradas correspondentes forem iguais.*

Em certo sentido as matrizes constituem uma generalização dos números. Podemos considerá-las uma espécie de números mais complicados. Ora, o números seriam pouco úteis se a eles não se encontrasse associada uma álgebra. A mesma ideia se aplica às matrizes pelo que iremos agora descrever a respectiva álgebra.

1.2.1 ADIÇÃO DE MATRIZES

Podem adicionar-se matrizes do mesmo tipo. Ou seja, existe uma operação de adição, $+ : \mathbb{K}^{m \times n} \times \mathbb{K}^{m \times n} \to \mathbb{K}^{m \times n}$ que associa a cada par de matrizes do tipo $m \times n$ uma outra matriz do mesmo tipo que se designa de *soma das duas primeiras*.

Mais precisamente, dadas matrizes $A, B \in \mathbb{K}^{m \times n}$ a respectiva adição, que se denota $A + B$, é a matriz $C \in \mathbb{K}^{m \times n}$ que é definida por:

$$C_{ij} = A_{ij} + B_{ij}. \tag{1.1}$$

Tem-se assim que:

$$\begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix} + \begin{bmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,n} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m,1} & B_{m,2} & \cdots & B_{m,n} \end{bmatrix} = \begin{bmatrix} A_{1,1}+B_{1,1} & A_{1,2}+B_{1,2} & \cdots & A_{1,n}+B_{1,n} \\ A_{2,1}+B_{2,1} & A_{2,2}+B_{2,2} & \cdots & A_{2,n}+B_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1}+B_{m,1} & A_{m,2}+B_{m,2} & \cdots & A_{m,n}+B_{m,n} \end{bmatrix}$$

Lema 1.1.— *Consideremos matrizes $A, B, C \in \mathbb{K}^{m\times n}$. A operação de adição de matrizes possui as seguintes propriedades:*

(1) $(A+B)+C = A+(B+C)$ *i.e., a adição de matrizes é uma operação* associativa.

(2) $A+B = B+A$ *i.e., a adição de matrizes é uma operação* comutativa.

(3) *Sendo $\mathbb{0}_{m\times n} \in \mathbb{K}^{m\times n}$ a matriz cujas entradas são todas nulas (esta matriz designa-se de* matriz nula*) tem-se $A + \mathbb{0}_{m\times n} = A$.*

(4) *Dada uma matriz A, a matriz simétrica de A é a matriz $-A = [-A_{i,j}]$. Tem-se $A + (-A) = \mathbb{0}$.*

Demonstração.— (1) Tem-se que $((A+B)+C)_{i,j} = (A_{i,j}+B_{i,j})+C_{i,j}$ mas, como a adição de escalares é associativa tem-se que

$$(A_{i,j}+B_{i,j})+C_{i,j} = A_{i,j}+(B_{i,j}+C_{i,j}) = (A+(B+C))_{i,j}.$$

Como isto é verdade para quaisquer i, j, o resultado segue.

(2) Análogo ao anterior considerando que a adição de escalares é comutativa.

(3) Tem-se que $(A+\mathbb{0})_{i,j} = A_{i,j} + \mathbb{0}_{i,j} = A_{i,j} + 0 = A_{i,j}$.

(4) Análogo ao anterior. ◪

1.2.2 multiplicação de um escalar por uma matriz

Dados, um escalar $\alpha \in \mathbb{K}$, e uma matriz $A \in \mathbb{K}^{m\times n}$ o produto do escalar α pela matriz A e é a matriz $C \in \mathbb{K}^{m\times n}$ definida por:

$$C_{ij} = \alpha A_{ij} \tag{1.2}$$

ou seja,

$$\alpha \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix} = \begin{bmatrix} \alpha A_{1,1} & \alpha A_{1,2} & \cdots & \alpha A_{1,n} \\ \alpha A_{2,1} & \alpha A_{2,2} & \cdots & \alpha A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha A_{m,1} & \alpha A_{m,2} & \cdots & \alpha A_{m,n} \end{bmatrix}.$$

O resultado seguinte resume as propriedades básicas da operação de multiplicação de uma matriz por um escalar.

Lema 1.2.— *Consideremos $A, B \in \mathbb{K}^{m\times n}$ e $\alpha, \beta \in \mathbb{K}$. Tem-se:*

(1) $\alpha(A + B) = \alpha A + \alpha B$ *(distributividade do produto por escalar em relação à adição de matrizes).*

(2) $(\alpha + \beta)A = \alpha A + \beta A$ *(distributividade do produto por escalar em relação à adição de escalares).*

(3) $1A = A$, $0A = \mathbb{0}$ *e* $(-1)A = -A$.

Demonstração.— A demonstração é trivial, seguindo a estratégia da demonstração do lema 1.1. ◪

1.2.3 multiplicação de matrizes

O produto de matrizes é uma operação $\cdot : \mathbb{K}^{m\times p} \times \mathbb{K}^{p\times n} \to \mathbb{K}^{m\times n}$. Tal como no caso da adição de matrizes observa-se uma restrição nos tipos das matrizes que se podem multiplicar. No caso do produto de matrizes A e B, para que o produto de A por B, que se denota AB, se encontre definido, é preciso que *o número de colunas do factor esquerdo iguale o número de linhas do factor direito*. O resultado da operação é neste caso uma matriz com o mesmo número de linhas do factor esquerdo e o mesmo número de colunas do factor direito.

É conveniente começar por descrever o produto de matrizes num caso básico, mais precisamente quando A é um vector linha (do tipo $1 \times p$) e B é um vector coluna (do tipo $p \times 1$, para que o produto se encontre definido). Neste caso particular o produto (que é uma matriz 1×1) define-se:

$$AB = \begin{bmatrix} A_{1,1}B_{1,1} + A_{1,2}B_{2,1} + \cdots + A_{1,p}B_{p,1} \end{bmatrix}.$$

De uma maneira geral iremos considerar como essencialmente iguais uma matriz 1×1 e o escalar que constitui a sua única entrada. Optando por uma interpretação ou por outra consoante o contexto.[3]

[3] Se o contexto requer uma matriz, interpretamos $[a]$ estritamente como sendo a matriz $[a]$, se o contexto requer um número, interpretamos $[a]$ como sendo a.

Dada uma matriz A a linha i de A denota-se por $A_{i,*}$ e é um vector linha, enquanto que a coluna j de A se denota por $A_{*,j}$ e é um vector coluna. Ou seja,

$$A_{i,*} = \begin{bmatrix} A_{i,1} & \cdots & A_{i,j} & \cdots & A_{i,n} \end{bmatrix}, \qquad A_{*,j} = \begin{bmatrix} A_{1,j} \\ \vdots \\ A_{i,j} \\ \vdots \\ A_{m,j} \end{bmatrix}$$

Podemos agora definir o produto de matrizes no caso geral. Assim, se

$A \in \mathbb{K}^{m\times p}$ e $B \in \mathbb{K}^{p\times n}$ então AB é a matriz $C \in \mathbb{K}^{m\times n}$ onde:

$$C_{ij} = A_{i,*}B_{*,j}. \tag{1.3}$$

(Em (1.3) estamos a identificar $A_{i,*}B_{*,j}$ com um escalar e não com uma matriz 1×1.) Assim, o elemento da posição i, j da matriz produto resulta de multiplicar a linha i do factor esquerdo, pela coluna j do factor da direita.[4]

Definição 1.3 (combinação linear de vectores).— *Uma soma do tipo $\alpha_1 \mathrm{u}_1 + \cdots + \alpha_k \mathrm{u}_k$ onde os α_i são escalares e os u_i são vectores (linha ou coluna) diz-se uma* combinação linear dos u_i. *Os escalares $\alpha_1, \ldots, \alpha_k$ dizem-se os* coeficientes *da combinação linear.*

A seguinte observação (importante) resulta imediatamente da definição de produto de duas matrizes. Se $A \in \mathbb{K}^{m\times p}$, $B \in \mathbb{K}^{p\times n}$ e $AB = C$ então,

$$C_{i,*} = A_{i,*}B = A_{i,1}B_{1,*} + A_{i,2}B_{2,*} + \cdots + A_{i,p}B_{p,*}, \tag{1.4}$$

ou seja, a linha i de AB resulta de multiplicar a linha i de A por B e é a combinação linear das linhas de B obtida considerando como coeficientes os elementos da linha i de A. Por outro lado,

$$C_{*,j} = AB_{*,j} = B_{1,j}A_{*,1} + B_{2,j}A_{*,2} + \cdots + B_{p,j}A_{*,p}, \tag{1.5}$$

ou seja a coluna j de AB resulta de multiplicar A pela coluna j de B e é a combinação linear das colunas de A que se obtém considerando os elementos da coluna j de B como coeficientes.

As relações (1.4) e (1.5) podem estabelecer-se da seguinte forma:

$$\begin{aligned} C_{i,*} = \; & \begin{bmatrix} A_{i,1}B_{1,1} + \cdots + A_{i,p}B_{p,1} & A_{i,1}B_{1,2} + \cdots + A_{i,p}B_{p,2} & \cdots & A_{i,1}B_{1,n} + \cdots + A_{i,p}B_{p,n} \end{bmatrix} = \\ & \begin{bmatrix} A_{i,1}B_{1,1} & A_{i,1}B_{1,2} & \cdots & A_{i,1}B_{1,n} \end{bmatrix} + \cdots + \begin{bmatrix} A_{i,p}B_{p,1} & A_{i,p}B_{p,2} & \cdots & A_{i,p}B_{p,n} \end{bmatrix} = \\ & A_{1,1}B_{1,*} + \cdots + A_{i,p}B_{p.*}. \end{aligned}$$

o que estabelece (1.4). Quanto a (1.5), pode estabelecer-se de forma totalmente análoga.

O lema que se segue enumera mais algumas propriedades básicas da álgebra de matrizes, envolvendo a multiplicação.

Lema 1.4.— *Sempre que os produtos das matrizes forem possíveis tem-se,*

(1) $\mathbb{0}A = \mathbb{0}$ *e* $A\mathbb{0} = \mathbb{0}$.

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

(3) $A(BC) = (AB)C$ *(a multiplicação é associativa)*

(4) $A(B + C) = (AB) + (AC)$;

(5) $(A + B)C = (AC) + (BC)$;

[4] Uma primeira observação é oportuna desde já: o produto de matrizes não é, em geral, comutativo. De facto pode até acontecer que AB esteja definido e BA, não. Por outro lado AB e BA podem estar definidos mas serem matrizes de tipos diferentes pelo que, necessariamente, $AB \neq BA$. Mas, mesmo quando AB e BA são do mesmo tipo, pode acontecer que $AB \neq BA$.
Se considerarmos as matrizes

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

tem-se que

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

enquanto que

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

(6) *Se $A \in \mathbb{K}^{m\times n}$ e denotarmos por $\mathbb{1}_k$ é a matriz do tipo $k \times k$ em que as entradas nas posições (i,i) são iguais a 1 e as restantes são nulas (a* matriz identidade de ordem *k) tem-se: $\mathbb{1}_m A = A$ e $A\mathbb{1}_n = A$.*[5]

[5] Em geral escrevemos simplesmente $\mathbb{1}$ em lugar de $\mathbb{1}_n$ (omitindo a ordem da matriz) porque quando escrevemos e.g., $A + \mathbb{1}$ ou $A\mathbb{1}$, a ordem da matriz fica inequivocamente determinada pelo facto de a operação ter que fazer sentido. Assim, se escrevemos $\mathbb{1}A$ e $A\mathbb{1}$ e se $A \in \mathbb{R}^{2\times 3}$, no primeiro caso $\mathbb{1}$ tem ordem 2 enquanto que no segundo tem ordem 3.

Demonstração.— (1) É trivial e fica a cargo do leitor.
(2) Suponhamos que $A \in \mathbb{K}^{m\times p}$ e $B \in \mathbb{K}^{p\times n}$. Tem-se:

$$\begin{aligned}
(\alpha(AB))_{i,j} &= \alpha \sum_{r=1}^{p} A_{i,r}B_{r,j} = \sum_{r=1}^{p}(\alpha A_{i,r})B_{r,j} = \sum_{r=1}^{p}(\alpha A)_{i,r}B_{r,j} = \\
&= ((\alpha A)B)_{i,j} = \\
&= \sum_{r=1}^{p} A_{i,r}(\alpha B_{r,j}) = \sum_{r=1}^{p} A_{i,r}(\alpha B)_{r,j} = \\
&= (A(\alpha B))_{i,j}.
\end{aligned}$$

(3) Suponhamos que $A \in \mathbb{K}^{m\times p}$, $A \in \mathbb{K}^{p\times n}$ e $C \in \mathbb{K}^{n\times k}$. Tem-se:

$$\begin{aligned}
(A(BC))_{i,j} &= A_{i,*}(BC)_{*,j} = A_{i,*}(C_{1,j}B_{*,1} + \cdots + C_{n,j}B_{*,n}) = \\
&= C_{1,j}A_{i,*}B_{*,1} + \cdots + C_{n,j}A_{i,*}B_{*,n} = \\
&= (AB)_{i,1}C_{1,j} + \cdots + (AB)_{i,n}C_{n,j} = \\
&= (AB)_{i,*}C_{*,j} = ((AB)C)_{i,j}.
\end{aligned}$$

(4) e (5) são análogos ao anterior. (6) resume-se a uma simples verificação. ◪

A razão para que se enunciem duas leis distributivas, designadamente $A(B + C) = AB + AC$ e $(A + B)C = AC + BC$ resulta do facto de a multiplicação de matrizes não ser em geral comutativa.

Definição 1.5.— *Uma matriz $A \in \mathbb{K}^{n\times n}$ diz-se* quadrada *e a respectiva* ordem *é n.*

Definição 1.6.— *Seja $A \in \mathbb{K}^{n\times n}$ uma matriz quadrada de ordem n. Dizemos que $B \in \mathbb{K}^{n\times n}$ é uma inversa de A se $AB = BA = \mathbb{1}$.*[6]

[6] Observe-se entretanto que para se ter $AB = BA$ ou seja para que ambos os produtos estejam definidos e sejam iguais é necessário que as matrizes A e B sejam quadradas da mesma ordem.

Veremos adiante (secção 1.2.5) como saber se uma matriz A tem inversa (ou como também se diz *é invertível*) e como calcular essa inversa se ela existir.

Lema 1.7.— *Se a matriz $A \in \mathbb{K}^{n\times n}$ tem inversa então, tem uma única inversa (que, nesse caso, se denota por A^{-1}).*

Demonstração.— Suponhamos que B e $\bar{B}$ são inversas de A. Tem-se então que $B = B\mathbb{1} = B(A\bar{B}) = (BA)\bar{B} = \mathbb{1}\bar{B} = \bar{B}$, como se pretendia. ◪

Enquanto que as propriedades da adição de matrizes são análogas às da adição de escalares, o mesmo não sucede com o produto de matrizes, onde desde logo falha a comutatividade. Mas as diferenças não se resumem à comutatividade.

Lema 1.8.— *As leis seguintes* não são*, de uma maneira geral, válidas no caso do produto de matrizes:*

(1) $AB = \mathbb{0} \Leftrightarrow A = \mathbb{0} \vee B = \mathbb{0}$.

(2) $AB = AC \Rightarrow B = C;$

(3) $BA = CA \Rightarrow B = C;$

(4) $(A + B)^2 = A^2 + 2AB + B^2;$

(5) $(A - B)(A + B) = A^2 - B^2$.

Demonstração.— O falhanço destas propriedades, em geral, pode ser evidenciado através de contra-exemplos. Uma matriz $A \neq \mathbb{0}$ tal que $A^2 = \mathbb{0}$ servirá de contra-exemplo (considerando $B = A$ e $C = \mathbb{0}$) para as três primeiras propriedades. Ora,

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \mathbb{0}.$$

No caso das propriedades (4) e (5) elas dependem da comutatividade do produto que, como vimos não é uma propriedade geral. Por exemplo, considerando (4), tem-se:

$$(A+B)^2 = (A+B)(A+B) = A(A+B)+B(A+B) = A^2+AB+BA+B^2.$$

Ora $A^2 + AB + BA + B^2 = A^2 + 2AB + B^2$ sse $AB + BA = 2AB$, o que sucede sse $AB = BA$, ou seja se as matrizes A e B comutam e, como se sabe, isso nem sempre sucede. Analogamente para (5). ⍁

Apesar do resultado anterior tem-se o seguinte:

Lema 1.9.— *Suponhamos que A é invertível. Tem-se:*

(1) $AB = AC \Rightarrow B = C;$

(2) $BA = CA \Rightarrow B = C$.

Além disso, se A é invertível ou B é invertível e $AB = \mathbb{0}$ então $A = \mathbb{0}$ ou $B = \mathbb{0}$. Finalmente se A e B comutam então:

(1) $(A + B)^2 = A^2 + 2AB + B^2;$

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

Demonstração.— Se $AB = AC$ e A é invertível então, multiplicando ambos os membros à esquerda por A^{-1} obtemos $A^{-1}AB = A^{-1}AC$ ou seja $B = C$, como se pretendia. O outro caso é análogo.

Suponhamos agora que $AB = \mathbb{0}$ e, digamos, B é invertível. Multiplicando à direita por B^{-1} obtemos $ABB^{-1} = \mathbb{0}B^{-1}$, ou seja $A = \mathbb{0}$, como se pretendia. Se for A invertível, multiplicando à esquerda por A^{-1} obtém-se $B = \mathbb{0}$.

A última parte do resultado é evidente a partir das considerações feitas na parte final da demonstração do lema precedente. ◪

Definição 1.10 (Inversas laterais).— *Seja $A \in \mathbb{K}^{n\times m}$ uma matriz. Uma* inversa esquerda *de A é uma matriz $B \in \mathbb{K}^{m\times n}$ tal que $BA = \mathbb{1}$. Analogamente, uma* inversa direita *de A é uma matriz $C \in \mathbb{K}^{m\times n}$ tal que $BC = \mathbb{1}$.*

Veremos que se A é quadrada e possui uma inversa direita (resp. esquerda) então ela também é inversa esquerda (resp. direita). Mas, no caso de uma matriz não ser quadrada, perde-se a unicidade da inversa lateral, i.e. uma matriz pode ter muitas inversas laterais. (Ver secção 1.2.5.)

1.2.4 transposição de matrizes

A *transposição de matrizes* associa a cada matriz $A \in \mathbb{K}^{m\times n}$ a sua transposta $A^\top \in \mathbb{K}^{n\times m}$. A definição é simples:

$$A^\top_{ij} = A_{ji}. \tag{1.6}$$

Essencialmente, as linhas de $A^\top$ são as colunas de A (e as colunas de $A^\top$ são as linhas da matriz A).

Se para uma matriz A se tem $A = A^\top$ a matriz diz-se *simétrica*.[7]

O resultado seguinte enumera as propriedades básicas da operação de *transposição*.

[7] Uma matriz simétrica é necessariamente uma matriz quadrada.

Lema 1.11.— *Sejam $A, B \in \mathbb{K}^{m\times n}$ e $\alpha \in \mathbb{K}$ tem-se:*

(1) $(A + B)^\top = A^\top + B^\top$.

(2) $(AB)^\top = B^\top A^\top$.

(3) *Se $A^\top$ é invertível então,* $(A^\top)^{-1} = (A^{-1})^\top$.

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

(5) $(A^\top)^\top = A$.

Demonstração.—
(1) $((A + B)^\top)_{i,j} = (A + B)_{j,i} = A_{j,i} + B_{j,i} = A^\top_{i,j} + B^\top_{i,j}$.
(2) $(AB)^\top_{i,j} = (AB)_{j,i} = A_{j,*}B_{*,i} = B^\top_{i,*}A^\top_{*,j} = (B^\top A^\top)_{i,j}$.
(3) Tem-se $(A^{-1})^\top A^\top = (AA^{-1})^\top = \mathbb{1}^\top = \mathbb{1} = (A^{-1}A)^\top = A^\top(A^{-1})^\top$.

(4) $((\alpha A)^\top)_{i,j} = \alpha A_{j,i} = \alpha A^\top_{i,j}$.
(5) Óbvio. ◪

1.3 SISTEMAS DE EQUAÇÕES LINEARES

Uma equação linear nas incógnitas (ou variáveis) $x_1, \dots, x_n$ com coeficientes no corpo $\mathbb{K}$ é uma equação da forma:

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \delta, \tag{1.7}$$

onde $\alpha_i \in \mathbb{K}$ é o coeficiente de x_i (para $i = 1, \dots, n$) e $\delta \in \mathbb{K}$ é habitualmente designado de *termo independente*.

Uma *solução* da equação (1.7) é um n-úplo $(\beta_1, \dots, \beta_n) \in \mathrm{K}^n$ para o qual a igualdade $\alpha_1\beta_1 + \cdots + \alpha_n\beta_n = \delta$ é verdadeira.[8]

Um sistema de equações lineares nas incógnitas $x_1, \dots, x_n$ é um conjunto de equações lineares, normalmente apresentado na forma:

$$\Sigma = \begin{cases} \alpha_{1,1}x_1 + \alpha_{1,2}x_2 + \cdots + \alpha_{1,n}x_n = \delta_1 \\ \alpha_{2,1}x_1 + \alpha_{2,2}x_2 + \cdots + \alpha_{2,n}x_n = \delta_2 \\ \qquad\qquad\vdots \\ \alpha_{m,1}x_1 + \alpha_{m,2}x_2 + \cdots + \alpha_{m,n}x_n = \delta_m \end{cases} \tag{1.8}$$

Uma solução do sistema Σ, é um n-úplo $(\beta_1, \dots, \beta_n)$ que é solução de *todas* as equações lineares que compõem o sistema. O *conjunto solução* ou *conjunto de soluções* do sistema Σ é o conjunto cujos elementos são todas as soluções do sistema. O conjunto de soluções de um sistema Σ denota-se por $S(\Sigma)$.

Relativamente à existência de soluções, um sistema de equações lineares, Σ pode ser *impossível*, i.e. não admitir soluções (conjunto solução é vazio) e.g. se o sistema Σ é composto pela única equação $0 = 1$ então $S(\Sigma) = \emptyset$;[9] pode ser *possível e determinado*, i.e. possuir uma única solução (o conjunto solução tem um único elemento) e.g. se Σ é o sistema nas variáveis x_1, x_2 consistindo nas equações $x_+x_2 = 1$ e $x_1 = 0$ então $S(\Sigma) = \{(0,1)\}$; ou ainda *possível e indeterminado*, i.e. possui mais que uma solução (o conjunto solução tem vários elementos) e.g. se Σ é um sistema nas variáveis x_1, x_2 e consiste na única equação $x_1 - x_2 = 0$ então $S(\Sigma) = \{(\alpha, \alpha) \mid \alpha \in \mathbb{K}\}$ que é um conjunto infinito.

DEFINIÇÃO 1.12.— *Dois sistemas de equações lineares dizem-se* equivalentes *se tiverem as mesmas soluções i.e., se os conjuntos solução forem iguais.*

[8] Observe-se que quando dizemos que $(\beta_1, \dots, \beta_n)$ é uma solução da equação (1.7), a afirmação pressupõe sempre uma ordenação das incógnitas neste caso, $(\beta_1, \dots, \beta_n)$ corresponde a $(x_1, \dots, x_n)$ indicando que β_i se deve substituir no lugar de x_i $(i = 1, \dots, n)$ na equação. A ordenação das variáveis tem que se pressupor mas não tem que ser nenhuma em particular.

[9] A equação $0 = 1$, ou mais geralmente uma equação da forma $0 = \delta$, pode ser vista como uma equação linear nas variáveis $x_1, \dots, x_n$ pois equivale a $0x_1 + \cdots + 0x_n = \delta$.

1.3.1 REPRESENTAÇÃO MATRICIAL DE SISTEMAS

Dado um sistema como (1.8) associamos-lhe duas matrizes. A *matriz dos coeficientes do sistema*:

$$\begin{array}{cc} & \begin{array}{cccc} x_1 & x_2 & \cdots & x_n \end{array} \\ A = & \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} \end{bmatrix} \begin{array}{l} \text{eq. } 1 \\ \text{eq. } 2 \\ \\ \text{eq. } m \end{array} \end{array}$$

que é constituída precisamente pelos coeficientes das variáveis nas equações do sistema – cada coluna diz respeito aos coeficientes de uma mesma variável e, cada linha guarda informação de uma única equação.

A segunda matriz que associamos ao sistema é a denominada *coluna dos termos independentes* que é o vector coluna: $\mathrm{b} = [\delta_1\ \delta_2\ \cdots\ \delta_m]^\top$. Tal como o nome indica, esta matriz guarda os termos independentes das diferentes equações, consideradas pela mesma ordem que na matriz dos coeficientes. O sistema pode assim ser descrito recorrendo à denominada *matriz aumentada do sistema* que é a matriz:

$$[A|\mathrm{b}] = \left[\begin{array}{cccc|c} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} & \delta_1 \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} & \delta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} & \delta_m \end{array}\right]$$

(O traço vertical não é essencial e é apenas uma forma conveniente de separar os coeficientes dos termos independentes.)

As descrições das matrizes A e b também permitem descrever um sistema de equações como uma equação matricial. De facto considerando x o vector coluna das variáveis i.e., $\mathrm{x} = [x_1\ x_2\ \cdots\ x_n]^\top$, as soluções do sistema são exactamente os x que tornam verdadeira a igualdade $A\mathrm{x} = \mathrm{b}$.[10]

[10] Para constatar este facto basta efectuar o produto do lado esquerdo do sinal de igualdade e utilizar o critério de igualdade de duas matrizes.

Relacionando os dois tipos de representação, a i-ésima linha do sistema representado por $[A|\mathrm{b}]$ descreve uma equação linear e, uma coluna u é solução da equação representada pela i-ésima linha de $[A|\mathrm{b}]$ se e só se $A_{i,*}\mathrm{u} = \delta_i$.

Resumindo, podemos associar a um sistema de equações lineares como (1.8) dois tipos de representação matricial: uma na forma $[A|\mathrm{b}]$ onde A é a matriz dos coeficientes e b a coluna dos termos independentes.; ou então, uma equação matricial $A\mathrm{x} = \mathrm{b}$, onde A, b são como anteriormente e x é o vector coluna das variáveis. Recorremos a uma interpretação ou a outra consoante a nossa conveniência: iremos descrever um método geral de resolução de sistemas de equações lineares que envolve representações do tipo $[A|\mathrm{b}]$; por outro lado se A tem inversa então, considerando o sistema na forma $Ax = b$ obtemos imediatamente a solução:

$$Ax = b \Leftrightarrow A^{-1}(Ax) = A^{-1}b \Leftrightarrow x = A^{-1}b,$$

ou seja o vector $A^{-1}b$ é a única solução do sistema.[11]

Um sistema de equações da forma $Ax = \mathbb{0}$ i.e., onde os termos independentes são todos nulos diz-se *homogéneo*. Como é evidente $A\mathbb{0} = \mathbb{0}$ pelo que os sistemas homogéneos têm sempre soluções ou seja, são sempre possíveis (em particular possuem a solução nula que corresponde a interpretar todas as variáveis por zero). No entanto um sistema homogéneo pode ser indeterminado, i.e. podem existir n-úplos de escalares $(\xi_1, \xi_2, \dots, \xi_n)$, não todos nulos que permitam obter $\mathbb{0}$ como combinação linear das colunas de A, i.e.

$$\xi_1 A_{*,1} + \xi_2 A_{*,2} + \cdots + \xi_n A_{*,n} = \mathbb{0} \text{ e } (\exists i \in \{1, \dots, n\})\xi_i \neq 0. \quad (1.9)$$

[11] Esta observação também revela que se a matriz dos coeficientes é invertível então o sistema é *possível* i.e., admite soluções, e é *determinado* i.e., possui uma única solução.
Por outro lado, tendo em conta as propriedades da multiplicação de matrizes, concluímos imediatamente que o sistema $Ax = b$ é possível se e só se a coluna dos termos independentes b é uma combinação linear das colunas de A (uma solução u de $Ax = b$ fornece precisamente os coeficientes de uma tal combinação linear).

Definição 1.13 (dependência e independência linear).— *Dados vectores* $u_1, \dots, u_n$, *dizemos que são* linearmente dependentes *se existirem escalares não todos nulos* $\xi_1, \dots, \xi_n$ *tais que*

$$\xi_1 u_1 + \cdots + \xi_n u_n = \mathbb{0}.$$

Caso contrário i.e., se $\xi_1 u_1 + \cdots + \xi_n u_n = \mathbb{0}$ *implica que* $\xi_1 = \cdots = \xi_n = 0$, *os vectores dizem-se* linearmente independentes.

Definição 1.14 (núcleo de uma matriz).— *Seja* $A \in \mathbb{K}^{m\times n}$. *O núcleo de* A, *que se denota* $\mathrm{Nuc}(A)$ *é o conjunto de soluções do sistema homogéneo* $Ax = \mathbb{0}$.

Das definições anteriores resulta imediatamente o seguinte:

Lema 1.15.— *Seja* A *uma matriz. As colunas de* A *são linearmente dependentes se e só se* $\mathrm{Nuc}(A) \neq \{\mathbb{0}\}$ *(ou seja, são linearmente independentes se e só se* $\mathrm{Nuc}(A) = \{\mathbb{0}\}$*).*

Lema 1.16.— *Tem-se o seguinte:*

(a) $\mathrm{Nuc}(B) \subset \mathrm{Nuc}(AB)$;

(b) *se* A *é invertível então* $\mathrm{Nuc}(AB) = \mathrm{Nuc}(B)$.

Demonstração.—
(a) Se $x \in \mathrm{Nuc}(B)$, tem-se $Bx = \mathbb{0}$. Neste caso $ABx = A\mathbb{0} = \mathbb{0}$.
(b) Suponhamos que a matriz A é invertível. Tem-se que $ABx = \mathbb{0}$ se e só se $A^{-1}ABx = A^{-1}\mathbb{0}$ ou seja se e só se $Bx = \mathbb{0}$. Por outras palavras $x \in \mathrm{Nuc}(AB)$ se e só se $x \in \mathrm{Nuc}(B)$. ◻

Definição 1.17.— *Dado um sistema* $A\mathrm{x} = \mathrm{b}$ *o* sistema homogéneo correspondente *é o sistema* $A\mathrm{x} = \mathbb{0}$.

Lema 1.18.— *Consideremos o sistema* $A\mathrm{x} = \mathrm{b}$. *Se* u_0 *é uma solução do sistema então, qualquer solução do sistema é da forma* $\mathrm{u}_0 + \mathrm{v}$ *onde* v *é uma solução do sistema homogéneo correspondente.*

Demonstração.— Tem-se que $A\mathrm{u}_0 = \mathrm{b}$. Suponhamos que u_1 é uma outra solução do sistema, caso em que se tem $A\mathrm{u}_1 = \mathrm{b}$. Tem-se então que:

$$A(\mathrm{u}_1 - \mathrm{u}_0) = A\mathrm{u}_1 - A\mathrm{u}_0 = \mathrm{b} - \mathrm{b} = \mathbb{0},$$

ou seja $\mathrm{v} = \mathrm{u}_1 - \mathrm{u}_0$ é uma solução do sistema homogéneo que corresponde ao sistema original. Tem-se assim que $\mathrm{u}_1 = \mathrm{u}_0 + \mathrm{v}$.

Por outro lado, se v é solução do sistema homogéneo, tem-se para $\mathrm{u}_1 = \mathrm{u}_0 + \mathrm{v}$ que,

$$A\mathrm{u}_1 = A(\mathrm{u}_0 + \mathrm{v}) = A\mathrm{u}_0 + A\mathrm{v} = \mathrm{b} + \mathbb{0} = \mathrm{b},$$

pelo que v é solução de $A\mathrm{x} = \mathrm{b}$. ◪

Não obstante estas considerações, o método mais geral de resolução de sistemas de equações lineares recorre à representação do sistema na forma $[A|\mathrm{b}]$. A ideia consiste em recorrer a um algoritmo através do qual se obtém um sistema equivalente ao sistema dado mas, relativamente ao qual, a determinação das soluções é trivial. Esse algoritmo conhecido como *método de eliminação de Gauss* procede a essa transformação de matrizes recorrendo a certas operações que actuam sobre as linhas da matriz e que se denominam de *operações elementares*.

1.3.2 operações elementares, matrizes elementares e o método de eliminação de gauss

Dada uma matriz $A \in \mathbb{K}^{m\times n}$ designamos de *operações elementares sobre linhas* a uma qualquer operação dos seguintes três tipos:

(1) troca de duas linhas de A;

(2) multiplicação de uma linha de A por um escalar $\alpha \neq 0$;

(3) adição a uma linha, de uma outra previamente multiplicada por um escalar α.

Os efeitos destas operações numa matriz podem ser obtidos algebricamente, multiplicando à esquerda a matriz A por certas matrizes que designamos de *matrizes elementares* e que passamos desde já a descrever.

Denotamos por $E_{ij}^m \in \mathbb{K}^{m\times m}$ a matriz, que é como a matriz identidade excepto que com as linhas i e j da identidade trocadas entre si. Denotamos por $E_{ij}^m(\alpha) \in \mathbb{K}^{m\times m}$ a matriz que é como a matriz identidade excepto que a entrada i, j é α. Finalmente, denotamos por $E_i^m(\alpha) \in \mathbb{K}^{m\times m}$ (para $\alpha \neq 0$) a matriz que é como a identidade excepto que na i-ésima posição da diagonal está o escalar α. Tem-se então o seguinte:

(1) $E_{ij}^m A = B$, onde B é a matriz que resulta de A trocando as linhas i e j entre si.

(2) $E_i^m(\alpha)A = B$, onde B é a matriz que resulta de A substituindo a linha i de A por $\alpha A_{i,*}$, i.e. multiplicando a linha i de A pelo escalar $\alpha \neq 0$.

(3) $E_{ij}^m(\alpha)A = B$, onde B é a matriz que resulta de A substituindo a linha i de A por $A_{i,*} + \alpha A_{j,*}$, i.e. adicionando à linha i de A a linha j multiplicada por α.

As matrizes elementares são todas invertíveis e não é difícil constatar que $(E_{ij}^m)^{-1} = E_{ij}^m$; $(E_i^m(\alpha))^{-1} = E_i^m(\alpha^{-1})$; e $(E_{ij}^m(\alpha))^{-1} = E_{ij}^m(-\alpha)$. Em particular, *as inversas das matrizes elementares são matrizes elementares do mesmo tipo*.

É interessante notar que, se em vez de multiplicarmos matrizes elementares à esquerda, efectuarmos as multiplicações à direita, obtemos efeitos semelhantes, mas desta vez sobre as colunas. Mais precisamente:

(1) $AE_{ij}^m = B$, onde B é a matriz que resulta de A trocando as colunas i e j entre si.

(2) $AE_i^m(\alpha) = B$, onde B é a matriz que resulta de A substituindo a coluna i de A por $\alpha A_{*,i}$, i.e. multiplicando a coluna i de A pelo escalar $\alpha \neq 0$.

(3) $AE_{ij}^m(\alpha) = B$, onde B é a matriz que resulta de A substituindo a coluna j de A por $A_{j,*} + \alpha A_{i,*}$, i.e. adicionando à coluna j de A a coluna i multiplicada por α.

As operações elementares sobre linhas podem ser usadas estrategicamente para transformar uma matriz A numa matriz que possui uma forma característica e se diz estar *em escada por linhas*.

Definição 1.19 (matriz em escada por linhas).— *Seja $A \in \mathbb{K}^{m\times n}$ uma matriz. Dizemos que A se encontra em* escada por linhas *se, designado por* pivô *o primeiro elemento não nulo em cada linha e acumulando-se as (eventuais) linhas nulas da matriz nas últimas linhas, percorrendo os pivôs de forma descendente, os índices de coluna das respectivas posições formam uma* *sequência estritamente crescente, i.e se* $(a_{1,j_1}, a_{2,j_2}, \dots, a_{k,j_k})$ *for a sequência dos pivôs da matriz se tem* $j_1 < j_2 < \cdots < j_k$.

Definição 1.20.— *Duas matrizes* A, B *dizem-se* Gauss-equivalentes, *escrevemos* $A \equiv_G B$, *se existe uma sequência,* $(E_1, E_2, \dots, E_k)$, *de matrizes elementares tal que*

$$B = E_1 E_2 \cdots E_k A.$$

Teorema 1.21.— *A relação* $\equiv_G$ *possui as seguintes propriedades:*

(1) $A \equiv_G A$;

(2) $A \equiv_G B$ *se e só se* $B \equiv_G A$;

(3) *se* $A \equiv_G B$ *e* $B \equiv_G C$ *então* $A \equiv_G C$.

Demonstração.— (1) $A = \mathbb{1}A$ e $\mathbb{1}$ é uma matriz elementar.

(2) Tem-se $B = E_1 E_2 \cdots E_k A$ se e só se $A = E_k^{-1} \cdots E_2^{-1} E_1^{-1} A$ e, pelo que vimos anteriormente $(E_k^{-1}, \dots, E_2^{-1}, E_1^{-1})$ é uma sequência de matrizes elementares.

(3) Se $B = E_1 E_2 \cdots E_k A$ e $C = \bar{E}_1 \bar{E}_2 \cdots \bar{E}_s B$ então,

$$C = \bar{E}_1 \bar{E}_2 \cdots \bar{E}_s E_1 E_2 \cdots E_k A,$$

onde, $(\bar{E}_1, \bar{E}_2, \dots, \bar{E}_s, E_1, E_2, \dots, E_k)$ é uma sequência de operações elementares. $\boxslash$

Teorema 1.22.— *Seja* A *uma matriz arbitrária. Existe uma matriz em escada por linhas,* K, *que é Gauss-equivalente a* A, *i.e.* $A \equiv_G K$.

Demonstração.— Uma matriz como K obtém-se através do *algoritmo de eliminação de Gauss*. Qualquer matriz $A \in \mathbb{K}^{m\times n}$ pode ser decomposta na forma:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \cdots & a_{k,n} \\ a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} A^{\underline{k}} \\ \\ A^{\overline{k}} \end{bmatrix} = [A^{\underline{k}} \; A^{\overline{k}}]^\top$$

onde $k = 0, \dots m$, $A^{\underline{k}}$ consiste nas primeiras k linhas de A e $A^{\overline{k}}$ consiste nas linhas de A a partir da linha $k+1$ (inclusive). (Tem-se assim que $A^{\underline{0}}$ é a matriz vazia e $A^{\overline{0}} = A$, enquanto que $A^{\underline{m}} = A$ e $A^{\overline{m}}$ é a matriz vazia.)

Dada uma matriz A a matriz A^G é a matriz que resulta de A depois de efectuadas as seguintes transformações, trocando linhas (se for necessário) passamos uma das linhas que tem o pivô o mais à esquerda possível para primeira linha. Depois, usando a operação de adicionar a uma linha, uma outra previamente multiplicada por um escalar α, podemos anular todas as posições abaixo do pivô na primeira linha. Findo este processo, obtemos A^G.

O método de eliminação de Gauss aplicado a A produz uma sequência de matrizes:

$$A = A_0 \to A_1 \to A_2 \to \cdots \to A_s,$$

onde $A_{i+1} = [A_i^{\underline{i}}\, (A_i^{\overline{i}})^G]^\top$, para $i < s$. Além disso, $s = m$ ou então $s < m$ e s é mínimo tal que $A_s^{\overline{s}} = \mathbb{0}$.

Por indução é fácil mostrar que $A_i^{\underline{i}}$ é uma matriz em escada por linhas e daqui conclui-se imediatamente que A_s é uma matriz em escada por linhas. (Ou seja tomamos $K = A_s$.) ⧄

Neste ponto é conveniente considerar um exemplo. Apliquemos então o algoritmo de eliminação de Gauss à matriz

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

Na primeira etapa trocando linhas, se necessário, asseguramos que a primeira linha seja uma das linhas da matriz com o pivô o mais à esquerda possível. Assim, trocaremos a primeira linha com a terceira (a troca da primeira com a segunda seria igualmente admissível). Obtemos então

$$A_0 = A = \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \underset{L_1+L_3\to L_3}{\longrightarrow} \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = A_1$$

Obtemos assim, ao fim da primeira etapa a matriz A_1. A partir deste momento a primeira linha da matriz nunca mais se modifica. A etapa 2, irá agora incidir sobre a parte da matriz que corresponde às linhas a partir da linha 2, repetido neste bloco o procedimento anterior. A primeira linha deste bloco já é (neste bloco) uma das que tem o pivô o mais à esquerda possível pelo que a manteremos como primeira linha. Procedemos então para anular todas as posições abaixo do pivô:

$$A_1 = \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \\ \\ -2L_2+L_3 \to L_3 \\ \\ \\ \end{matrix} \to \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = A_2$$

Chegados ao final da segunda etapa, as duas primeiras linhas da matriz vão, a partir de agora, permanecer alteradas e vamos repetir o processo, considerando o bloco que consiste nas três últimas linhas. A primeira linha deste bloco já é uma das que neste bloco tem o pivô o mais à esquerda possível e podemos usá-la. No entanto as constas simplificam-se muito quando o pivô é 1 ou -1 pelo que, ainda assim, iremos trocar as linhas 3 e 5, obtendo-se:

$$\begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ 2L_3+L_5 \to L_5 \end{matrix} \to \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix} = A_3$$

Repetindo o processo uma última vez obtemos

$$K = A_4 = \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

que é uma matriz em escada por linhas.

A matriz K no lema precedente não é única, e.g se K está em escada de linhas então $K' = E_1(2)K$ também está em escada de linhas e é Gauss-equivalente a K. Existe no entanto uma forma em escada por linhas para o qual se pode obter um resultado de unicidade.

Definição 1.23 (matriz em escada por linhas reduzida).— *Uma matriz K, em escada de linhas, encontra-se em* escada por linhas reduzida *se todos os pivôs são 1 e acima de cada pivô só existem zeros.*

Corolário 1.23.1.— *Seja A uma matriz arbitrária. Existe uma matriz, $\bar{K}$, em escada por linhas reduzida tal que $A \equiv_G \bar{K}$.*

Demonstração.— O teorema 1.22 garante a existência de uma matriz, K, em escada de linhas, tal que $A \equiv_G K$. Para, a partir de K se obter $\bar{K}$, usando operações elementares procedemos da seguinte forma. Multiplicamos cada linha pelo inverso do pivô, obtendo desta forma uma matriz, B, cujos pivôs são todos iguais a 1. Podemos agora eliminar todas as posições acima de cada pivô. Se uma coluna contém um pivô numa posição que não seja a primeira (o 1 aparecendo na linha $r > 1$), então

$$\begin{bmatrix} & B_{1,j} & \\ * & \vdots & * \\ & B_{r-1,j} & \\ 0 \cdots\cdots 0 & 1 & * \\ 0 \cdots\cdots 0 & 0 & \\ \ddots & \vdots & * \\ 0 \cdots\cdots 0 & 0 & \end{bmatrix} \xrightarrow{E_{r-1},\ldots,E_1} \begin{bmatrix} & 0 & \\ * & \vdots & * \\ & 0 & \\ 0 \cdots\cdots 0 & 1 & * \\ 0 \cdots\cdots 0 & 0 & \\ \ddots & \vdots & * \\ 0 \cdots\cdots 0 & 0 & \end{bmatrix}$$

onde, para $i = 1, \ldots, r-1$ se tem que

$$E_i = E_{r,i}(-B_{i,j}^{-1}).$$

Repetindo o procedimento com todos os pivôs, obtemos uma matriz $\bar{K}$ que se encontra em escada de linhas reduzida e é tal que $A \equiv_G \bar{K}$. ☒

1.3.3 UNICIDADE DA MATRIZ EM ESCADA DE LINHAS REDUZIDA

Já observámos antes que toda a matris A é Gauss equivalente a uma matriz K em escada de linhas reduzida. Muito embora A seja Gauss equivalente a uma infinidade de matrizes em escada de linhas, ela existe uma única matriz em escada de linhas reduzida nestas circunstâncias. Tem-se então o seguinte:

TEOREMA 1.24.— *Sejam $A, K_1, K_2 \in \mathbb{K}^{m\times n}$ uma matriz onde K_1 e K_2 são matrizes em escada de linhas reduzida. Se $A \equiv_G K_1$ e $A \equiv_G K_2$ então $K_1 = K_2$. Além disso se $K_1 = E_p \cdots E_1 A$ e $K_2 = F_s \cdots F_1 A$ onde as matrizes $E_1, \ldots, E_p, F_1, \ldots, F_s$ são elementares então $E_p \cdots E_1 = F_s \cdots F_1$.*

DEMONSTRAÇÃO.— Fixemos a matriz $C = E_p \cdots E_1 F_1^{-1} \cdots F_s^{-1}$. Temos que:

$$K_1 = CK_2 \text{ e } K_2 = C^{-1}K_1.$$

Provaremos que $K_1 = K_2$ e $E_p \cdots E_1 = F_s \cdots F_1$ (provando que $C = \mathbb{1}$) por indução em $1 \leq k \leq m$.

Denotamos por e_j^s a j-ésima coluna da matriz identidade de ordem s e consideramos $\mathrm{a}_j = A\mathrm{e}_j^n$, $\mathrm{k}_j^1 = K_1\mathrm{e}_j^n$, $\mathrm{k}_j^2 = K_2\mathrm{e}_j^n$, e $\mathrm{c}_j = C\mathrm{e}_j^m$ (ou seja os vectores $\mathrm{a}_j, \mathrm{k}_j^1, \mathrm{k}_j^2$ e c_j correspondem à j-ésima coluna das matrizes A, K_1, K_2 e C, respectivamente).

A primeira observação interessante é que se tem $\mathrm{k}_j^1 = \mathbb{0}$ se e só se $\mathrm{k}_j^2 = \mathbb{0}$ se e só se $\mathrm{a}_j = \mathbb{0}$ (e desta forma podemos supor que nem A nem K_1 nem K_2 possuem colunas nulas). Isto constata-se da seguinte forma: se $\mathrm{k}_j^2 = \mathbb{0}$ então $\mathbb{0} = C\mathrm{k}_j^2 = CK_2\mathrm{e}_j^n = K_1\mathrm{e}_j^n = \mathrm{k}_j^1$. A implicação contrária obtém-se considerando a matriz C^{-1}. Por outro lado, como $U = E_p \cdots E_1 A$ e $E_p \cdots E_1$ é uma matriz invertível (é um produto de matrizes invertíveis) tem-se que $\mathrm{Nuc}(U) = \mathrm{Nuc}(A)$ e assim $\mathrm{e}_1^n \in \mathrm{Nuc}(U)$ se e só se $\mathrm{e}_1^n \in \mathrm{Nuc}(A)$.

Agora que nos concentramos em matrizes que não possuem colunas de zeros, tem-se que $\mathrm{k}_1^1 = \mathrm{k}_1^2 = \mathrm{e}_1^m$.

Afirmação 1.—Para $1 \leq k \leq n$ tem-se que e_k^m é a j_k-ésima coluna de K_1 se e só se é a j_k-ésima coluna de K_2. Além disso, se e_k^m é a j_k-ésima coluna de K_1 então:

(1) As primeiras j_k colunas de K_1 e K_2 coincidem.

(2) As colunas com índice superior a j_k, que são nulas nas posições correspondentes às linhas de índice superior a k são comuns a K_1 e K_2.

(3) As primeiras k colunas de C coincidem com as primeiras k colunas da matriz identidade de ordem m.

Consideremos então indução em k (com $1 \leq k \leq m$). O primeiro caso é o caso $k = 1$ e já vimos que a primeira coluna de K_1 é e_1^m sse e só se esta é também a primeira coluna de K_2. Assim

$$C\mathrm{e}_1^m = CK_2\mathrm{e}_1^n = K_1\mathrm{e}_1^n = \mathrm{e}_1^m.$$

Por outro lado se uma coluna j de K_2 é da forma $\lambda\mathrm{e}_1$, ou seja, se $K_2\mathrm{e}_j^n = \lambda\mathrm{e}_1^m$ então:

$$K_1\mathrm{e}_1^n = CK_2\mathrm{e}^n = C\mathrm{e}_1^m = \mathrm{e}_1^m,$$

pelo anterior. A implicação contrária estabelece-se usando C^{-1} (observe-se que também se tem $C^{-1}\mathrm{e}_1^m = \mathrm{e}_1^m$).

O caso $k = 1$ fica assim estabelecido. Admitimos agora que o resultado é válido para um dado $k < m$ e provaremos que permanece válido para $k + 1$.

Só temos que nos preocupar com o caso em que e_{k+1} ocorre em K_1 pois caso contrário, a hipótese de indução já garante a igualdade das matrizes K_1 e K_2. Suponhamos assim que $\mathrm{e}_{k+1}^m \in K_1$. Pela cláusula (2) da hipótese de indução tem-se que $\mathrm{e}_{k+1}^m \in K_2$ ocorrendo necessariamente na mesma coluna j_{k+1} que ocorre em K_1. Tem-se então que:

$$c_{k+1} = C\mathrm{e}_{k+1}^m = CK_2\mathrm{e}_{j_{k+1}}^n = K_1\mathrm{e}_{j_{k+1}}^n = \mathrm{e}_{k+1}^m.$$

Este resultado e a hipótese de indução revelam que as primeira $k + 1$ colunas de C coincidem com as colunas de índice correspondente na matriz identidade de ordem m.

Consideremos agora as colunas com índices superiores a j_{k+1} em K_2 e cujas posições abaixo da linha $k+1$ são todas nulas. Uma tal coluna k_j^2 (com $j > j_{k+1}$) é uma combinação linear das colunas $j_1, \dots, j_{k+1}$. Neste caso, tendo em conta que C é como a identidade até à coluna j_{k+1} concluímos que

$$\mathrm{k}_j^1 = C\mathrm{k}_j^2 = \mathrm{k}_j^2.$$

Usando C^{-1} podemos trocar os papeis de K_1 e K_2 e concluir que esta matrizes partilham exactamente as mesmas colunas com índice superior a j_{k+1} que são nulas nas posições correspondentes às linhas de índice superior a $k+1$. Completamos assim a demonstração do caso $k+1$.

Temos então necessariamente que $K_1 = K_2$. Por outro lado, como $K_1 = CK_2$ resulta que $C = \mathbb{1}$, ou seja,

$$E_p \cdots E_1 F_1^{-1} \cdots F_s^{-1} = E_p \cdots E_1 (F_s \cdots F_1)^{-1} = \mathbb{1}.$$

Mas isto significa que $E_p \cdots E_1 = F_s \cdots F_1$, concluindo-se assim a demonstração. ◪

1.3.4 RESOLUÇÃO DE SISTEMAS DE EQUAÇÕES PELO MÉTODO DE ELIMINAÇÃO DE GAUSS

LEMA 1.25.— *Suponhamos que* $[A|\mathrm{b}]$ *é a matriz aumentada de um sistema* S *e que* $[A'|\mathrm{b}']$ *é uma matriz em escada de linhas que é Gauss-equivalente a* $[A|\mathrm{b}]$. *Nestas condições, o sistema que é representado por* $[A'|\mathrm{b}']$ *é equivalente a* S.[12]

[12] De facto o resultado é ainda mais geral: se $[A|b] \equiv_G [A'|b']$, então os sistemas representados por $[A|b]$ e $[A'|b']$ são equivalentes.

DEMONSTRAÇÃO.— Basta-nos mostrar que se $[A'|\mathrm{b}']$ se obtém da matriz $[A|\mathrm{b}]$ efectuando uma das operações elementares então, os sistemas representados por estas duas matrizes são equivalentes.

CASO 1,.—(Troca de linhas.) Trocar linhas numa matriz que representa um sistema corresponde a trocar a ordem das equações e isso não altera o conjunto de soluções do sistema. (Se solução do sistema corresponde a ser solução de todas as suas equações, independentemente pela ordem pela qual as consideramos.)

CASO 2.—(Multiplicação de uma linha por um escalar não nulo.) Uma linha na matriz do sistema representa uma equação linear que é uma relação do tipo $u = v$. Ora, considerando $\alpha \neq 0$ tem-se $u = v$ se e só se $\alpha u = \alpha v$, pelo que as equações são equivalente e, consequentemente, os sistemas são equivalentes.

CASO 3.—(Adicionar a uma linha, outra previamente multiplicada por um escalar.) Começamos por observar que se um sistema de equações

$\sigma = \{\mathscr{E}_1, \ldots, \mathscr{E}_m\}$ consiste de m equações $\mathscr{E}_1, \ldots, \mathscr{E}_m$ então, o conjunto solução desse sistema, que denotamos por $\mathrm{Sol}(\sigma)$ é:

$$\mathrm{Sol}(\sigma) = \mathrm{Sol}(\mathscr{E}_1) \cap \cdots \cap \mathrm{Sol}(\mathscr{E}_m),$$

onde, para $i = 1, \ldots, m$, $\mathrm{Sol}(\mathscr{E}_i)$ denota o conjunto de soluções da equação $\mathscr{E}_i$. Mas, como a intersecção de conjuntos é associativa e, para qualquer X se tem $X \cap X = X$, resulta que se considerarmos subconjuntos $I, J \subset \{1, \ldots, m\}$ tais que $I \cup J = \{1, \ldots, m\}$ se tem também que:

$$\mathrm{Sol}(\sigma) = \bigcap_{i \in I} \mathrm{Sol}(\mathscr{E}_i) \cap \bigcap_{j \in J} \mathrm{Sol}(\mathscr{E}_l).$$

Os sistemas $\sigma_I = \{\mathscr{E}_i \mid i \in I\}$ e $\sigma_J = \{\mathscr{E}_j \mid j \in J\}$ dizem-se *subsistemas* do sistema σ.[13] Podemos assim considerar o sistema original, $[A|\mathrm{b}]$ decomposto em dois subsistemas: o sistema $\sigma_1 = \{\mathscr{E}_s \mid s \neq i, j\}$ e o sistema $\sigma_2 = \{\mathscr{E}_i, \mathscr{E}_j\}$. O sistema correspondente a $[A'|\mathrm{b}']$ pode ser analogamente decomposto em dois subsistemas σ_1' e σ_2' onde $\sigma_1' = \sigma_1$ e $\sigma_2' = \{\mathscr{E}_i, \mathscr{E}_j^*\}$. Assim, o nosso resultado fica demonstrado se mostrarmos que os sistemas σ_2 e σ_2' são equivalentes.

[13] Podemos evidentemente subdividir o sistema σ num maior número de subsistemas, considerando

$$I_1 \cup \cdots \cup I_n = \{1, \ldots, m\}.$$

Os dois subsistemas σ_2 e σ_2' podem escrever-se nas formas:

$$\begin{bmatrix} a_{i,1} \cdots\cdots a_{i,n} \\ a_{j,1} \cdots\cdots a_{j,n} \end{bmatrix} \mathrm{x} = \begin{bmatrix} b_i \\ b_j \end{bmatrix}, \qquad \begin{bmatrix} a_{i,1} \cdots\cdots\cdots\cdots a_{i,n} \\ a_{j,1} + \alpha a_{i,1} \cdots\cdots a_{j,n} + \alpha a_{i,n} \end{bmatrix} \mathrm{x} = \begin{bmatrix} b_i \\ b_j + \alpha b_i \end{bmatrix}$$

Podemos agora verificar que qualquer solução de um dos sistemas é também solução do outro. Verificamos uma das direcção deixando a outra ao cuidado do leitor. Suponhamos então que $(u_1, \ldots, u_n)$ é solução do segundo sistema. Tem-se então que:

$$\begin{aligned} &a_{i,1}u_1 + \cdots + a_{i,n}u_n = b_i \\ &(a_{j,1} + \alpha a_{i,1})u_1 + \cdots + (a_{j,n} + \alpha a_{i,n})u_n = b_j + \alpha b_i. \end{aligned}$$

Observe-se que se tivermos uma igualdade $\eta = \zeta$ então, para qualquer α tem-se $\alpha\eta = \alpha\zeta$ e, desta forma, para quaisquer ϕ, ψ tem-se:

$$\phi = \psi \Leftrightarrow \phi + \alpha\eta = \psi + \alpha\zeta.$$

Fazendo uso destas considerações podemos somar à segunda igualdade a primeira multiplicada por $-\alpha$ sabendo que o resultado será uma igualdade equivalente. Assim, aquelas duas equações são equivalentes à duas seguintes:

$$\begin{aligned} a_{i,1}u_1 + \cdots + a_{i,n}u_n = b_i \\ a_{j,1}u_1 + \cdots + a_{j,n}u_n = b_j. \end{aligned}$$

mostrando que $(u_1, \ldots, u_n)$ é solução de σ_2. A outra direcção é análoga e a demonstração fica assim completa. ◪

A conveniência de ter um sistema descrito por uma matriz em escada de linhas é que a determinação das suas soluções é simples. Ilustramos este facto recorrendo a dois exemplos.

Exemplo 1.— Consideremos o sistema nas variáveis x, y, z representado pela matriz:

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

Para escrever o sistema que corresponde a esta matriz ainda é necessária uma informação adicional: a que variável corresponde cada coluna da matriz (à esquerda do traço vertical). De facto as variáveis poderiam ser consideradas por qualquer ordem, continuando a representar sistemas equivalentes (a adição é comutativa). Não se dizendo nada considera-se a ordem alfabética, i.e.

$$\begin{array}{c} \begin{array}{cccc} x & y & z & \end{array} \\ \left[\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 3 & 2 \end{array}\right] \end{array} \longrightarrow \begin{cases} x + 2y + z = -1 \\ -y + z = 1 \\ 3z = 2 \end{cases} .$$

Observe-se agora que as variáveis se podem determinar começando pela última equação: $z = 2/3$; substituindo na segunda equação obtemos $y = z - 1 = 2/3 - 1 = -1/3$; finalmente, na primeira equação temos $x = -1 - 2y - z = -1 - 2/3 - 2/3 = -7/3$.

Exemplo 2.— Consideremos o sistema nas variáveis x_1, x_2, x_3, x_4, x_5 descrito pela matriz:

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{array}\right]$$

Mais uma vez não se indicando uma ordem para as variáveis admite-se que se consideram pela ordem natural:

$$\begin{array}{c} \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & \end{array} \\ \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{array}\right] \end{array} \longrightarrow \begin{cases} x_1 + x_4 + 2x_5 = -1 \\ x_3 + x_4 = 0 \\ -x_5 = 2 \end{cases} .$$

A resolução procede como no exemplo anterior, começando pela última equação concluímos que $x_5 = -2$; chegando à segunda equação, o que podemos fazer é escrever uma das variáveis à custa das restantes e embora isso não seja necessário, podemos por exemplo, escolher a variável correspondente à coluna com o pivô para ser expressa em função das restantes. Neste caso $x_3 = -x_4$; chegados à primeira equação e podemos escrever $x_1 = -1 - x_4 - x_5 = -1 - x_4 + 2 = 1 - x_4$. Neste caso o conjunto das soluções do sistema é o conjunto:

$$\{(1 - x_4, x_2, -x_4, -2) \mid x_2, x_4 \in \mathbb{R}\}.$$

O sistema é portanto possível e indeterminado. Observe-se que três das variáveis ficam expressas em função das duas restantes. Estas variáveis que se podem exprimir em função das restantes dizem-se *dependentes* em qualquer outro caso, tal como neste, são numa quantidade que iguala a quantidade de pivôs na matriz do sistema quando esta se encontra em escada de linhas. As restantes variáveis dizem-se *livres*. A sua quantidade é assim a diferença entre o número de variáveis e o número de pivôs. Esta diferença designa-se de *grau de indeterminação do sistema*.

Um sistema $A\mathrm{x} = \mathrm{b}$ que seja possível e indeterminado é equivalente a um sistema $K(x) = \mathrm{b}^*$, onde K é uma matriz em escada de linhas reduzida. Podemos resolver o sistema exprimindo as variáveis correspondentes às colunas com pivôs em função das que não possuem essa propriedade, nesta forma de exprimir a solução, as variáveis que correspondem às colunas com pivôs dizem-se *variáveis dependentes* as restantes dizem-se *livres*. O número de variáveis livres corresponde àquilo que se designa de *grau de indeterminação do sistema*.

Lema 1.26.— *O grau de indeterminação de um sistema, $A\mathrm{x} = \mathrm{b}$, possível, é uma característica do sistema e corresponde à diferença entre o número de colunas de A e o número de pivôs de uma qualquer matriz em escada de linhas, K, tal que $A \equiv_G K$.*

Demonstração.— O método que descrevemos na demonstração do corolário 1.23.1 permite-nos passar de uma matriz em escada de linhas para uma matriz em escada de linhas reduzida sem alterar nem o número de pivôs nem as respectivas localizações. ◪

1.3.5 O PROBLEMA DA EXISTÊNCIA DA INVERSA E A SUA DETERMINAÇÃO.

Retomamos neste ponto a questão de decidir se uma matriz é ou não invertível e, em caso afirmativo de determinar essa inversa (que, já se sabe, é única).

Consideremos então uma matriz quadrada $A \in \mathbb{K}^{n \times n}$. Sabemos que existe uma matriz invertível E (que é um produto de matrizes elementares) tal que $EA = K$ onde K é uma matriz em escada de linhas reduzida. Ora, se K não é a matriz indentidade então tem pelo menos uma linha de zeros e, neste caso, $\mathrm{Nuc}(K)$ não é trivial e, consequentemente não pode ser invertível porque o núcleo de uma matriz invertível é composto unicamente pelo vector nulo. Ora, sendo o produto EA não invertível e sendo E invertível, temos que concluir que A não é invertível.

Ou seja, tem-se o seguinte resultado,

LEMA 1.27.— *Uma matriz $A \in \mathbb{K}^{n\times n}$ é invertível se e só se o número de pivôs de uma qualquer matriz em escada de linhas, Gauss-equivalente a A, é n.*

De acordo com o lema precedente, no caso de A ser invertível, todos os sistemas da forma $A\mathrm{x} = \mathrm{b}$ são possíveis e determinados. Desta forma, existem soluções únicas para os sistemas

$$A\mathrm{x} = \mathbb{1}_{*,j} \qquad (1 \leq j \leq n).$$

Dispondo as soluções destes sistemas como colunas de uma matriz B, tendo em conta a definição de produto de duas matrizes, resulta que $AB = \mathbb{1}$. Ora esta matriz B é de facto a inversa de A pois, de acordo com o lema seguinte também se verifica $BA = \mathbb{1}$.

LEMA 1.28.— *Seja $A \in \mathbb{K}^{n\times n}$. São equivalentes:*

(1) *A é invertível;*

(2) *Existe B tal que $AB = \mathbb{1}$;*

(3) *Existe B tal que $BA = \mathbb{1}$.*

DEMONSTRAÇÃO.— A única coisa a demonstrar é que se $AB = \mathbb{1}$ então também se tem $BA = \mathbb{1}$. Suponhamos então que $AB = \mathbb{1}$. Neste caso também se tem que $B^\top A^\top = \mathbb{1}$ e então, forçosamente o sistema $A^\top \mathrm{x} = \mathbb{0}$ só pode ter como solução o vector nulo e assim, qualquer sistema da forma $A^\top \mathrm{x} = \mathrm{b}$ é possível e determinado. Reproduzindo o argumento acima podemos concluir a existência de uma matriz C tal que $A^\top C = \mathbb{1}$. Transpondo obtém-se $C^\top A = \mathbb{1}$. Considerando $D = C^\top$ temos assim que:

$$DA = \mathbb{1} \quad \text{e} \qquad AB = \mathbb{1}.$$

Mas agora tem-se:

$$D = D\mathbb{1} = D(AB) = (DA)B = \mathbb{1}B = B,$$

o que nos permite concluir a demonstração. ◻

Chapter 2

Espaços lineares

2.1 Definições e exemplos

Seja $\mathbb{K}$ é um corpo. Um *espaço linear* ou *espaço vectorial* sobre $\mathbb{K}$ é um conjunto V, não vazio, cujos elementos se designam genericamente de *vectores* (independentemente da sua natureza), equipado com uma operação de *adição de vectores* que, essencialmente, é uma função $V \times V \to V$ que associa a cada par (u, v) de elementos de V um elemento de V que se designa de *soma de* u *com* v e que, em geral, se denota por $\mathrm{u} +_V \mathrm{v}$ ou simplesmente $\mathrm{u} + \mathrm{v}$ quando V for claro. Além disso, a cada elemento $\alpha \in \mathbb{K}$ corresponde uma operação $\hat{\alpha}_V : V \to V$ que a cada vector $\mathrm{v} \in V$ faz corresponder um vector de V que se denota $\hat{\alpha}_V(\mathrm{v})$ ou mais simplesmente $\alpha\mathrm{v}$ (no contexto da teoria dos espaços vectoriais, os elementos do corpo $\mathbb{K}$ designam-se habitualmente de *escalares*). Finalmente, estas operações devem verificar o conjunto de propriedades (os *axiomas de espaço linear*) que se enunciam a seguir:

(1) para quaisquer $\mathrm{u}, \mathrm{v}, \mathrm{w} \in V$ tem-se $\mathrm{u} + (\mathrm{v} + \mathrm{w}) = (\mathrm{u} + \mathrm{v}) + \mathrm{w}$ (ou seja, a adição de vectores é associativa).

(2) para quaisquer $\mathrm{u}, \mathrm{v} \in V$ tem-se $\mathrm{u} + \mathrm{v} = \mathrm{v} + \mathrm{u}$ (ou seja, a adição de vectores é comutativa);

(3) existe um *elemento neutro para a adição*, ou seja, um elemento $\mathrm{e} \in V$ com a seguinte propriedade: para qualquer $\mathrm{u} \in V$ tem-se que $\mathrm{e} + \mathrm{u} = \mathrm{u}$.

De facto, não pode existir mais que um elemento neutro nestas circunstâncias, pois se $\mathrm{e}, \mathrm{f} \in V$ são elementos neutros para a soma de vectores então: $\mathrm{e} = \mathrm{e} + \mathrm{f}$ porque f é elemento neutro, enquanto que $\mathrm{f} = \mathrm{e} + \mathrm{f}$ porque e é elemento neutro. Mas então tem-se que $\mathrm{e} = \mathrm{e} + \mathrm{f} = \mathrm{f}$ ou seja, o elemento neutro para a adição de vectores é *único*. Tendo em conta esta unicidade, reservam-se, em geral uma notação e uma designação particulares para este elemento. Ele designa-se de *vector nulo de V* e denota-se por $\mathbb{0}_V$ ou simplesmente $\mathbb{0}$ quando não existe perigo de confusão.

Continuando a enunciar os axiomas de espaço linear:

(4) para cada $\mathrm{u} \in V$ existe um vector $\mathrm{w} \in V$ com a propriedade: $\mathrm{u} + \mathrm{w} = \mathbb{0}_V$.

Mais uma vez, para cada $\mathrm{u} \in V$ existe um único w nas condições de (4). Com efeito: se w e w′ satisfazem (4) relativamente a u então:

$$\mathrm{w} = \mathrm{w} + \mathbb{0}_V = \mathrm{w} + (\mathrm{u} + \mathrm{w}') = ((\mathrm{w} + \mathrm{u})'\mathrm{w}' = \mathbb{0}_V + \mathrm{w}' = \mathrm{w}'.$$

Perante esta unicidade, mais uma vez reservamos uma designação e uma notação especiais para w como em (4), ele designa-se de *simétrico de* u e denota-se $-\mathrm{u}$.

Os restantes axiomas são:

(5) para quaisquer $\alpha \in \mathbb{K}$ e $\mathrm{u}, \mathrm{v} \in V$ tem-se $\alpha(\mathrm{u} + \mathrm{v}) = (\alpha\mathrm{u}) + (\alpha\mathrm{v})$;

(6) para quaisquer $\alpha, \beta \in \mathbb{K}$ e $\mathrm{u} \in V$ tem-se $(\alpha + \beta)\mathrm{u} = (\alpha\mathrm{u}) + (\beta\mathrm{u})$;

(7) para quaisquer $\alpha, \beta \in \mathbb{K}$ e $\mathrm{u} \in V$ tem-se $(\alpha\beta)\mathrm{u} = \alpha(\beta\mathrm{u})$;

(8) $1\mathrm{u} = \mathrm{u}$.

Se $\mathbb{K} = \mathbb{C}$ o espaço linear diz-se um *espaço linear complexo*, se $\mathbb{K} = \mathbb{R}$ o espaço V diz-se um *espaço linear real*.

Observamos que convencionalmente a operação de multiplicação por escalar tem precedência sobre a adição de vectores, assim uma expressão como $\alpha\mathrm{x} + \beta\mathrm{y}$ significa o mesmo que $(\alpha\mathrm{x}) + (\beta\mathrm{y})$. Somas deste tipo, envolvendo um número finito de parcelas têm um papel fundamental no estudo destas estruturas pelo que lhes dedicamos a próxima definição:

Definição 2.1.— *Sejam V um espaço linear sobre $\mathbb{K}$, $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ e $x_1, x_2, \dots, x_n \in V$. Uma soma da forma:*

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \sum_{i=1}^{n} \alpha_i x_i, \tag{2.1}$$

onde $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ diz-se uma combinação linear *de $x_1, x_2, \dots, x_n$.*[1]

[1] Note-se que esta não é mais que a generalização aos vectores enquanto elementos de um espaço linear, da correspondente noção envolvendo vectores linha e coluna no caso das matrizes.

Considerando todas as possíveis variações de escalares em (2.1) obtemos todas as combinações lineares dos vectores $\mathrm{x}_1, \dots, \mathrm{x}_n$ cujo respectivo conjunto se denota por $L_V(\{\mathrm{x}_1, \dots, \mathrm{x}_n\})$ e designa de *expansão linear dos vectores* $\mathrm{x}_1, \dots, \mathrm{x}_n$ *em V*. Uma vez que produtos de um escalar por um vector e somas de vectores são elementos de V tem-se claramente que

$$L_V(\{\mathrm{x}_1, \dots, \mathrm{x}_n\}) \subset V.$$

Não existe nenhum motivo para considerar apenas expansões lineares de conjuntos finitos de vectores (embora as combinações lineares sejam sempre somas finitas). Se A é um conjunto infinito de vectores então a respectiva expansão linear, que se denota por $L_V(A)$, define-se como:

$$L_V(A) = \left\{ \sum_{i=0}^{n} \alpha_i \mathrm{x}_i \;\middle|\; n \in \mathbb{N}, x_1, \dots, x_n \in A, \alpha_1, \dots, \alpha_n \in \mathbb{K} \right\},$$

onde $\mathbb{N}$ é o conjunto dos números naturais (recorde-se que adoptamos a convenção segundo a qual $0 \in \mathbb{N}$ i.e., $\mathbb{N} = \{0, 1, 2, 3, \dots\}$.[2]

Exibimos agora alguns exemplos de espaços lineares.

[2] Note-se que apesar de termos fornecido duas definições de $L_V(X)$, uma para o caso em X é finito e outra para o caso em que X é infinito, a segunda definição poderia adoptar-se para os dois casos. Isto acontece porque se $X = \{\mathrm{x}_1, \dots, \mathrm{x}_n\}$ se tem que qualquer soma de elementos de X é equivalente a uma soma da forma $\alpha_1 \mathrm{x}_1 + \cdots + \mathrm{x}_n$.

Exemplo 3.— O conjunto ${}^{\mathbb{K}}\mathbb{K}$ que consiste nas funções $f : \mathbb{K} \to \mathbb{K}$, onde se definem as operações $(f, g) \mapsto f + g$ e $(\alpha, f) \mapsto \alpha f$ através de:

1. $(f + g)(k) \coloneqq f(k) + g(k)$, para $k \in \mathbb{K}$;
2. $(\alpha f)(k) \coloneqq \alpha \cdot f(k)$, para $k \in \mathbb{K}$.

Observe-se o *duplo sentido* do símbolo «+» acima: em $f + g$ ele descreve a operação de adição de funções em ${}^{\mathbb{K}}\mathbb{K}$, enquanto que em $f(k) + g(k)$ descreve a adição no corpo $\mathbb{K}$ (note-se que os valores $f(k)$ e $g(k)$ são elementos de $\mathbb{K}$. Por outro lado $\alpha \cdot f(k)$ denota também o produto em $\mathbb{K}$. Habitualmente este tipo de distinções é claro no contexto e abstraímo-nos de as produzir explicitamente.

Exemplo 4.— A análise do exemplo precedente mostra que a estrutura de espaço linear em ${}^{\mathbb{K}}\mathbb{K}$ provém da estrutura de $\mathbb{K}$ aplicada às imagens $f(k)$. Este facto sugere a generalização daquele através das estruturas ${}^{X}\mathbb{K}$ das funções $f : X \to \mathbb{K}$ onde, $X \neq \emptyset$ é agora um conjunto arbitrário. Efectivamente, definindo as operações de adição de vectores e de multiplicação por escalar como anteriormente, i.e.

1. $(f + g)(x) \coloneqq f(x) + g(x)$, para $x \in X$;
2. $(\alpha f)(x) \coloneqq \alpha \cdot f(x)$, para $x \in X$.

continuamos a obter um espaço linear sobre $\mathbb{K}$.

Muitos exemplos de espaços vectoriais são essencialmente da forma descrita no exemplo anterior, muito embora geralmente isso não seja explícito pois é mais conveniente descrever os vectores noutras formas mais convenientes à sua manipulação.

Exemplo 5.— Os espaços $\mathbb{K}^n$ consistem nos n-úplos ordenados de elementos de $\mathbb{K}$, i.e. elementos da forma $(\alpha_1, \dots, \alpha_n)$, onde $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ considerando a álgebra definida por:

1. $(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) \coloneqq (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$;
2. $k(\alpha_1, \dots, \alpha_n) \coloneqq (k\alpha_1, \dots, k\alpha_n)$.

Vale a pena mencionar a propriedade fundamental dos n-úplos ordenados relacionada com o critério de igualdade de dois objectos deste tipo: tem-se que $(\alpha_1, \dots, \alpha_n) = (\beta_1, \dots, \beta_n)$ se e só se

$$\alpha_1 = \beta_1 \wedge \alpha_2 = \beta_2 \wedge \cdots \wedge \alpha_n = \beta_n,$$

i.e. dois n-úplos são iguais se e só se os elementos nas posições correspondentes são iguais, e.g. $(1, 1, 2) \neq (1, 2, 1)$. Ora, aderindo à ideia segundo a qual os objectos em Matemática são essencialmente determinados pela informação que contêm, facilmente concluímos que os vectores em $\mathbb{K}^n$ podem ser identificados com as funções $f \in {}^n\mathbb{K}$,[3] ou seja, identificando o n-úplo $(\alpha_1, \dots, \alpha_n)$ com a função $f \in {}^n\mathbb{K}$ dada por $f(0) = \alpha_1, \dots, f(n-1) = \alpha_n$. Desta forma, tem-se que $\mathbb{K}^n$ é essencialmente ${}^n\mathbb{K}$.

[3] Esta notação merece um esclarecimento dado que descrevemos ${}^XY$ como sendo o conjunto das funções de X em Y pressupondo-se assim que X e Y sejam conjuntos. O que acontece é que os números naturais podem ser vistos como conjuntos (é de resto assim que, como qualquer objecto matemático, são definidos no seio da teoria de conjuntos) mais precisamente; $0 = \varnothing$ e nos restantes casos $n = \{0, \dots, n-1\}$ i.e., cada natural é o conjunto dos naturais que o precedem.

Exemplo 6.— Os espaços $\mathbb{K}^{m\times n}$ das matrizes do tipo $m \times n$ com entradas em $\mathbb{K}$. Com as operações usuais de adição de matrizes e de multiplicação de um escalar por uma matriz tem-se que $\mathbb{K}^{m\times n}$ é um espaço linear.

Estamos uma vez mais perante um exemplo que corresponde na sua essência a uma construção do tipo da descrita no exemplo 2. De facto, uma matriz $A \in \mathbb{K}^{m\times n}$ corresponde a uma função $f \in {}^{\{1,\dots,m\}\times\{1,\dots,n\}}\mathbb{K}$ definida por $f(i, j) = A_{i,j}$. Este facto revela que o espaço $\mathbb{K}^{m\times n}$ é essencialmente o espaço ${}^{\{1,\dots,m\}\times\{1,\dots,n\}}\mathbb{K}$.[4]

[4] De facto, neste tipo de identificações, é preciso identificar mais que os elementos dos dois conjuntos. Feita essa correspondência as *estruturas*, ou seja, as operações algébricas, devem também corresponder-se i.e., o resultado de operar objectos numa das estruturas deve corresponder ao resultado de operar, na outra estrutura os objectos correspondentes. Neste exemplo concreto, denotando por $\bar{A}$ a função que corresponde à matriz A tem-se: $\overline{A+B} = \bar{A} + \bar{B}$ e $\overline{\alpha A} = \alpha\bar{A}$, pelo que as duas estruturas são de facto equivalentes.

Exemplo 7.— Uma sucessão de elementos de $\mathbb{K}$ é uma sequência infinita da forma:

$$(x_0, x_1, x_2, \dots, x_n, \dots) \qquad n \in \mathbb{N}.$$

Usualmente, denotamos uma sucessão como acima recorrendo à notação (x_n). Uma sucessão é assim uma generalização da noção de sequência finita. De resto é válido um critério de igualdade semelhante i.e. tem-se

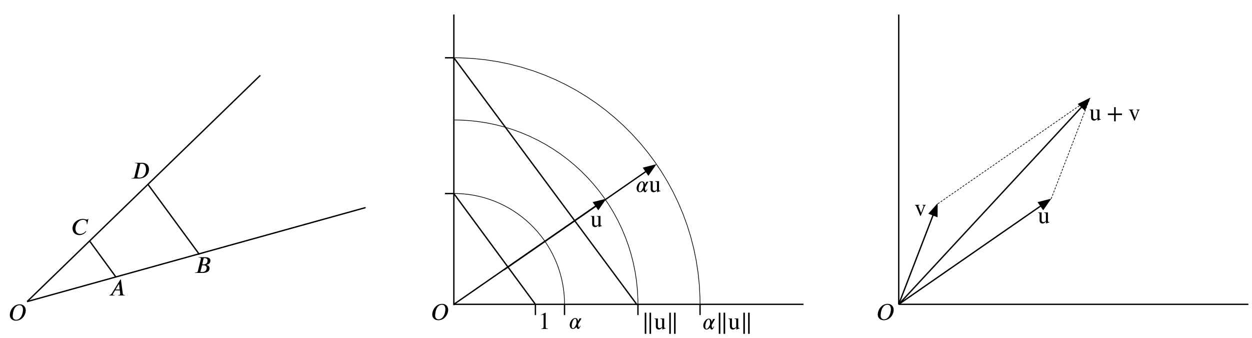
$$(x_n) = (y_n) \Leftrightarrow (\forall i \in \mathbb{N}) x_i = y_i.$$

Mais uma vez, as sucessões podem ser identificadas com funções, neste caso com funções $f : \mathbb{N} \to \mathbb{K}$. Por essa razão denotaremos o conjunto das sucessões em $\mathbb{K}$ por ${}^{\mathbb{N}}\mathbb{K}$. As operações são definidas como no caso ${}^X\mathbb{K}$ e, com estas operações, ${}^{\mathbb{N}}\mathbb{K}$ é um espaço linear (ver exemplo 2).

Consideramos um último exemplo. Trata-se de um exemplo particularmente interessante pois trata-se de um exemplo descrevendo uma estrutura muito próxima da intuição geométrica. Apesar de se tratar de um exemplo muito simples, constitui aquilo que poríamos considerar como uma boa *estrutura de teste* i.e., uma estrutura à qual podemos recorrer para testar conjecturas acerca de espaços lineares.

Exemplo 8.— O exemplo que vamos considerar consiste na estrutura dos *vectores fixos do plano*. Depois de fixarmos um ponto O no plano (*a origem*) um vector fixo do plano é um segmento orientado com origem em O e extremidade num ponto A desse mesmo plano, simbolicamente representamos esse vector por $\overrightarrow{OA}$. Em particular é possível considerar o vector $\mathbb{0} = \overrightarrow{OO}$. Graficamente o vector $\overrightarrow{OA}$ é representado através de uma seta com origem em O e extremidade em B. (Claro que no caso do vector $\mathbb{0}$ essa seta é degenerada e a origem coincide com a extremidade.) Para induzir neste conjunto de vectores uma estrutura de álgebra linear, consideramos operações que se definem geometricamente como se indica na figura 2.1.

Figure 2.1:



(A) (B) (C)

Em (A) encontra-se descrita a forma como se podem multiplicar geometricamente dois comprimentos. Considerando que o comprimento de OA é 1 e DB é paralelo a CA então, como se sabe, dado que os triângulos OAC e OBD são semelhantes, tem-se que:

$$\frac{\overline{OA}}{\overline{OC}} = \frac{\overline{OB}}{\overline{OD}},$$

ou seja, $\overline{OD} = \overline{OA} \cdot \overline{OB}$ (pois $\overline{OC} = 1$). Em (B) descreve-se a multiplicação por escalar (usando a construção em (A). Essencialmente, se u é um vector e $\alpha \in \mathbb{R}$, então αu é o vector que tem como suporte a mesma recta que u, comprimento $\alpha\|u\|$ onde $\|u\|$ denota o comprimento de u.

Além disso αu tem o mesmo sentido que u (está na mesma semi-recta) se $\alpha > 0$ enquanto que tem sentido contrário ao de u se $\alpha < 0$. Finalmente (C) descreve a denominada *regra do paralelogramo* que permite descrever a adição de vectores através da construção de um paralelogramo, tal como se descreve na figura.

EXEMPLO 9.— Dado um conjunto $X \neq \emptyset$ o *espaço linear livre sobre* X que denotamos por $C_{\mathbb{K}}(X)$ é o subespaço de ${}^{X}\mathbb{K}$ que consiste nas funções $f : X \to \mathbb{K}$ com *suporte* finito, onde o suporte de f, que se denota supp(f) é

$$\operatorname{supp}(f) := \{x \in X \mid f(x) \neq 0\}.$$

As funções da forma

$$c_a(x) := \begin{cases} 1 & (\text{se } x = a) \\ 0 & (\text{caso contrário}) \end{cases}$$

desempenham um papel importante. De facto, o conjunto $\{c_a \mid a \in X\}$ constitui um *conjunto gerador* de $C_{\mathbb{K}}(X)$ i.e., cada elemento $f \in C_{\mathbb{K}}(X)$ é da forma,

$$f(x) = \xi_1 c_{a_1}(x) + \cdots + \xi_n c_{a_n}(x),$$

onde $a_1, \ldots, a_n \in X$ e $\xi_1, \ldots, \xi_n \in \mathbb{K}$. Mais precisamente, se o suporte de f é $S = \{a_1, \ldots, a_n\}$ então

$$f(x) = f(a_1)c_{a_1}(x) + \cdots + f(a_n)c_{a_n}(x).$$

Além disso, este conjunto de geradores é *minimal* ou seja, removendo algum dos seus elementos deixa de ser possível gerar todo o $C_{\mathbb{K}}(X)$. Mais precisamente, removendo c_a deixa de ser possível geral qualquer função f cujo suporte contém a como elemento.

As propriedades seguintes são úteis apesar de simples consequências dos axiomas de espaço linear (e por isso são válidas em qualquer estrutura deste tipo).

LEMA 2.2.— *Seja V um espaço linear sobre $\mathbb{K}$. Tem-se:*

1. *Se* $\mathrm{x} \in V$ *então* $0\mathrm{x} = \mathbb{0}$.
2. *Se* $\mathrm{x} \in V$ *então* $(-1)\mathrm{x} = -\mathrm{x}$.
3. *Se* $\mathrm{x} \in V$, $\alpha, \beta \in \mathbb{K}$ *e* $\mathrm{x} \neq \mathbb{0}$ *então,* $\alpha\mathrm{x} = \beta\mathrm{x}$ *se e só se* $\alpha = \beta$.
4. *Se* $\mathrm{x}, \mathrm{y} \in V$, $\alpha \in \mathbb{K} \setminus \{0\}$ *então,* $\alpha\mathrm{x} = \alpha\mathrm{y}$ *se e só se* $\mathrm{x} = \mathrm{y}$.

DEMONSTRAÇÃO.— 1. Tem-se que $0\mathrm{x} = (0 + 0)\mathrm{x} = 0\mathrm{x} + 0\mathrm{x}$. Adicionando $-(0\mathrm{x})$ a ambos os membros da igualdade anterior obtém-se imediatamente que $\mathbb{0} = 0\mathrm{x}$.

2. Tem-se que $(x)+(-1)\mathrm{x} = 1\mathrm{x}+(-1)\mathrm{x} = (1-1)\mathrm{x} = 0\mathrm{x} = \mathbb{0}$. Ora, como $-\mathrm{x}$ é único para a propriedade $\mathrm{y} + \mathrm{x} = \mathbb{0}$, concluímos que $(-1)\mathrm{x} = -\mathrm{x}$.

3. Consideremos $\mathrm{x} \neq \mathbb{0}$ e $\alpha\mathrm{x} = \beta\mathrm{x}$ com $\alpha \neq \beta$. Ten-se então:

$$\alpha\mathrm{x} = \beta\mathrm{x} \Leftrightarrow \alpha\mathrm{x} - \beta\mathrm{x} = \mathbb{0} \Leftrightarrow (\alpha - \beta)\mathrm{x} = \mathbb{0}.$$

Como estamos a supor que $\alpha \neq \beta$ temos $\alpha-\beta \neq 0$ e podemos considerar $(\alpha - \beta)^{-1}$. Multiplicando a igualdade acima por este valor concluímos:

$$(\alpha - \beta)^{-1}(\alpha - \beta)\mathrm{x} = \mathbb{0},$$

ou seja $\mathrm{x} = \mathbb{0}$, o que contradiz a suposição original. Temos então que ter $\alpha = \beta$.

4. Temos que

$$\alpha\mathrm{x} = \alpha\mathrm{y} \Leftrightarrow \alpha\mathrm{x} - \alpha\mathrm{y} = \mathbb{0} \Leftrightarrow \alpha(\mathrm{x} - \mathrm{y}) = \mathbb{0}.$$

Como $\alpha \neq 0$ podemos multiplicar ambos os membros da última igualdade por α^{-1} obtemos $\mathrm{x} - \mathrm{y} = \mathbb{0}$ ou seja $\mathrm{x} = \mathrm{y}$. ⧄

2.1.1 SUBESPAÇOS DE UM ESPAÇO LINEAR

Suponhamos que V é um espaço linear sobre $\mathbb{K}$. Suponhamos que a operação de adição de vectores é $+_V$ e que as operações de multiplicação por escalar são $\hat{\alpha}_V$ (para $\alpha \in \mathbb{K}$). (Neste ponto estamos a usar uma notação mais complicada para descrever as operações no espaço V porque será conveniente distinguir claramente as operações em diferentes espaços.) Consideremos agora um subconjunto não-vazio $X \subset V$. A *restrição da operação de adição de vectores em V, ao conjunto X*, denota-se por $+_{V|X}$ e é a função $+_{V|X} : X \times X \to V$ definida por:

$$\mathrm{x} +_{V|X} \mathrm{y} \coloneqq \mathrm{x} +_V \mathrm{y},$$

ou seja, os resultados da restrição $+_{V|X}$ são calculados da mesma maneira que recorrendo à adição em V, simplesmente enquanto função, o domínio de $+_{V|X}$ envolve apenas vectores em X. Analogamente, para as multiplicações por escalar: a restrição de $\hat{\alpha}_V$ a X denota-se $\hat{\alpha}_{V|X}$ e para cada $\mathrm{x} \in X$ tem-se $\hat{\alpha}_{V|X}(\mathrm{x}) = \hat{\alpha}_V(\mathrm{x})$.

Resumindo, para calcular o resultado de uma restrição aplicada a elementos de X, *esquecemos* que eles são elementos de X e calculamos o resultado através da correspondente operação em V, considerando os vectores enquanto elementos de V. Em geral, não distinguiremos simbolicamente uma operação e a sua restrição até porque, os resultados são calculados da mesma forma.

Pode perfeitamente acontecer que aplicando uma restrição a X de operações em V a elementos de X não resulte num elemento de X. E.g. se $X = \{\mathrm{x}\}$ (com $\mathrm{x} \neq \mathbb{0}$) e se $\mathbb{K} = \mathbb{R}$ então $2\mathrm{x}$ não é elemento de X (porque, como $\mathrm{x} \neq \mathbb{0}$ tem-se que $2\mathrm{x} \neq \mathrm{x}$). Quando esta situação ocorre para alguma das operações, dizemos que X *não é fechado para as operações de* V. Caso contrário diremos que X *é fechado para as operações de* V.

DEFINIÇÃO 2.3.— *Sejam,* V *um espaço linear sobre* $\mathbb{K}$ *e* $U \subset V$ *tal que* $U \neq \emptyset$. *Diremos que* U *é um* subespaço *de* V *se* U, *equipado com as restrições das operações de* V *a* U, *é um espaço linear. Se* U *é um subespaço de* V *escrevemos* $U \leqslant V$.

O resultado seguinte é de grande utilidade prática, pois permite verificar se um subconjunto de um espaço linear é um subespaço sem que tenhamos que fazer uma verificação de todos os axiomas de espaço linear.

TEOREMA 2.4.— *Sejam,* V *um espaço linear sobre* $\mathbb{K}$ *e* $U \subset V$ *com* $U \neq \emptyset$. *Temos que* U *é um subespaço de* V *se e só se* U *é fechado para as operações de* V, *i.e. se e só se,*

1. $\alpha\mathrm{x} \in U$, *para quaisquer* $\alpha \in \mathbb{K}$ *e* $\mathrm{x} \in U$;
2. $\mathrm{x} + \mathrm{y} \in U$, *para quaisquer* $\mathrm{x}, \mathrm{y} \in U$.

Ou ainda, se e só se

$$\alpha\mathrm{x} + \beta\mathrm{y} \in U, \text{ quaisquer que sejam } \alpha, \beta \in \mathbb{K} \text{ e } \mathrm{x}, \mathrm{y} \in U.$$

DEMONSTRAÇÃO.— Temos que verificar que nestas condições todos os axiomas de espaço linear são verdadeiros em U com as restrições das operações em V. Um dos axiomas que temos que estabelecer é a comutatividade da adição de vectores. A propriedade em causa estabelece que quaisquer que sejam x e y se tem $\mathrm{x} + \mathrm{y} = \mathrm{y} + \mathrm{x}$. A expressão *quaisquer que sejam* corresponde ao uso de quantificadores universais e, uma asserção que envolve apenas este tipo de quantificador diz-se *universal*. A asserção da comutatividade é assim uma asserção universal. Por que é que ela é verdadeira em U. Pela simples razão que dados $\mathrm{x}, \mathrm{y} \in U$ eles se somam em U da mesma forma que se somam em V, e como a adição comuta para quaisquer vectores em V ela comuta também para quaisquer vectores em U.

O argumento anterior pode usar-se para qualquer para qualquer outra asserção universal (dados que, insistimos, as operações em U produzem os mesmos resultados que em V).

Analizando os axiomas de espaço linear constatamos que, com poucas excepções, elas são asserções universais. As únicas que não são deste tipo são, a existência de simétrico, e a existência do vector nulo. Mas, tendo em conta que $U \neq \varnothing$ é fechado para as operações tem-se que $0\mathrm{x} = \mathbb{0} \in U$ e $(-1)\mathrm{x} = -\mathrm{x} \in U$ (onde x é um qualquer elemento de U). A condição é assim suficiente.

A condição também é necessária porque num espaço linear o resultado de qualquer operação envolvendo vectores tem que ser um elemento desse espaço (por definição). ⍁

Corolário 2.4.1.— *Tem-se que* $\{\mathbb{0}\} \leqslant V$ *e* $V \leqslant V$.

Exemplos típicos de subespaços são as denominadas *expansões lineares* de conjuntos de vectores que já definimos anteriormente.

Teorema 2.5.— *Sejam,* V *um espaço linear sobre* $\mathbb{K}$ *e* $X \subset V$. *Tem-se que* $L_V(X)$ *é um subespaço de* V.

Demonstração.— O vector nulo pode sempre ser escrito como uma combinação linear de quaisquer vectores, usando os coeficientes todos nulos. Tem-se assim que $\mathbb{0} \in L_V(X)$, pelo que $L_V(X) \neq \varnothing$.[5] Suponhamos agora que $\mathrm{x}_1, \mathrm{x}_2 \in L_V(X)$. Neste caso podemos considerar vectores $\mathrm{x}_{1,1} \dots, \mathrm{x}_{1,r}, \mathrm{x}_{2,1}, \dots, \mathrm{x}_{2,s} \in X$ tais que

$$\mathrm{x}_1 = \sum_{i=1}^{r} \xi_i \mathrm{x}_{1,i}, \quad \mathrm{x}_2 = \sum_{i=1}^{s} \zeta_i \mathrm{x}_{2,i}.$$

Sendo assim tem-se que:

$$\mathrm{x}_1 + \mathrm{x}_2 = \sum_{j=1}^{r+s} \alpha_j \mathrm{w}_j,$$

onde, para $1 \leq i \leq r$ se tem $\alpha_i = \xi_i$ e $\mathrm{w}_i = \mathrm{x}_{1,i}$ e para $1 \leq i \leq s$ se tem $\alpha_{r+i} = \zeta_i$ e $\mathrm{w}_{r+i} = \mathrm{x}_{2,i}$. Concluímos assim que $\mathrm{x}_1 + \mathrm{x}_2$ é uma combinação linear de elementos de X ou seja, $\mathrm{x}_1 + \mathrm{x}_2 \in L_V(X)$.

Consideremos agora $\mathrm{x} \in L_V(X)$ e $\alpha \in \mathbb{K}$. Como anteriormente, podemos considerar $\mathrm{x}_1 \dots, \mathrm{x}_r \in X$ e escalares $\xi_1, \dots, \xi_r$ tais que

$$\mathrm{x} = \sum_{i=1}^{r} \xi_i \mathrm{x}_i.$$

[5] Uma das condições para verificar que U é subespaço de V é verificar que $U \neq \varnothing$. Uma boa forma de proceder a este teste é verificar que $\mathbb{0} \in U$. Com efeito se $\mathbb{0} \in U$ conclui-se imediatamente que $U \neq \varnothing$ e, se $\mathbb{0} \notin U$ então podemos imediatamente concluir que U não é subespaço já que o vector nulo tem que pertencer a qualquer subespaço.

Neste caso,

$$\alpha \mathrm{x} = \alpha\left(\sum_{i=1}^{r} \xi_i \mathrm{x}_i\right) = \sum_{i=1}^{r} (\alpha\xi_i)\mathrm{x}_i \in L_V(X).$$

Concluímos assim que $L_V(X)$ é um subespaço de V. ◻

Exemplo 10.— Uma questão natural é a de saber se um vector $\mathrm{x} \in V$ é um elemento do espaço gerado por um conjunto finito, X, de vectores de V i.e. a questão de saber para um dado $\mathrm{x} \in V$ se tem $\mathrm{x} \in L_V(X)$.

Nos espaços $\mathbb{K}^n$ existe um procedimento efectivo que permite responder a esta questão. Na génese desse procedimento encontra-se uma observação básica: se $A \in \mathbb{K}^{m\times n}$ é uma matriz e $\mathrm{x} \in \mathbb{K}^{n\times 1}$ é o vector coluna $[\alpha_1 \ \cdots \ \alpha_n]^T$ então,

$$A\mathrm{x} = \alpha_1 A_{1,*} + \cdots + \alpha_n A_{n,*},$$

ou seja $A\mathrm{x}$ é uma combinação linear das colunas de A, os escalares que determinam essa combinação linear são as entradas do vector coluna x. Ou seja um vector b é uma combinação linear das colunas de A se e só se existe um vector x tal que $A\mathrm{x} = \mathrm{b}$, ou seja, se e só se o sistema $A\mathrm{x} = \mathrm{b}$ *é possível*.

Exemplo 11.— Suponhamos que $V = \mathbb{R}^3$ e consideremos os vectores $(1,0,1), (2,2,3) \in \mathbb{R}^3$. Será que o vector $(1,2,3)$ é uma combinação linear daqueles dois vectores?

Tendo em conta as observações precedentes, para responder a esta questão temos apenas que ver os vectores como colunas de uma matriz A e ver se a coluna $\mathrm{b} = [1\ 2\ 3]^T$ é combinação linear das colunas de A, ou seja, verificar se o sistema $A\mathrm{x} = \mathrm{b}$ é possível. Neste caso obtemos o sistema:

$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \text{ou} \qquad \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 2 & 2 \\ 1 & 3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{cc|r} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{array}\right]$$

pelo que o sistema é *impossível*. Concluímos assim que o vector $(1,2,3)$ não é elemento da expansão linear $L_V(\{(1,0,1),(2,2,3)\})$.

Podemos até considerar o problema genericamente e caracterizar exactamente os vectores $(x,y,z) \in \mathbb{R}^3$ que são elementos daquela expansão linear, ou seja, que são combinações lineares daqueles dois vectores. A questão é semelhante, queremos determinar em que condições, o sistema:

$$\left[\begin{array}{cc|c} 1 & 2 & x \\ 0 & 2 & y \\ 1 & 3 & z \end{array}\right]$$

é possível. Ora, recorrendo ao método de eliminação de Gauss, obtemos:

$$\left[\begin{array}{cc|c}1&2&x\\0&2&y\\1&3&z\end{array}\right]\to\left[\begin{array}{cc|c}1&2&x\\0&2&y\\0&1&z-x\end{array}\right]\to\left[\begin{array}{cc|l}1&2&x\\0&1&z-x\\0&2&y\end{array}\right]\to$$
$$\to\left[\begin{array}{cc|l}1&2&x\\0&1&z-x\\0&0&2x+y-2z\end{array}\right]$$

O que mostra que o sistema é possível se e só se $2x+y-z=0$. Assim, concluímos que

$$L_V(\{(1,0,1),(2,2,3)\}) = \{(x,y,z)\in\mathbb{R}^3 \mid 2x+y-z=0\}.$$

O exemplo precedente ilustra o seguinte resultado:

Lema 2.6.— *Dados vectores* $\mathrm{u}_1,\ldots,\mathrm{u}_k\in\mathbb{K}^n$ *um vector* w *é combinação linear dos vectores* $\mathrm{u}_1,\ldots,\mathrm{u}_k\in\mathbb{K}^n$ *i.e.,* $\mathrm{w}\in L_{\mathbb{K}}(\{\mathrm{u}_1,\ldots,\mathrm{u}_k\})$ *se e só se o sistema* $A\mathrm{x}=\mathrm{w}$, *onde* A *é a matriz que tem como colunas os vectores* $\mathrm{u}_1,\ldots,\mathrm{u}_k$, *é possível.*

Definição 2.7.— *Um espaço linear* V *sobre* $\mathbb{K}$ *diz-se de* dimensão finita *ou* finitamente gerado *se existe* $X\subset V$, *finito, tal que* $V=L_V(X)$ *i.e., qualquer vector em* V *é uma combinação linear dos vectores em* X. *Se, pelo contrário, para qualquer* $X\subset V$, *finito, se tem* $L_V(X)\neq V$ *então, dizemos que* V *é um espaço de* dimensão infinita.[6]

Como veremos adiante o lema 2.6 tem um âmbito mais geral do que aparenta.[7]

Por definição, um espaço linear de dimensão finita, V, é um espaço da forma $V=L_V(\{\mathrm{u}_1,\ldots,\mathrm{u}_k\})$. Teremos a ocasião de estabelecer mais adiante que se $U\leq V$ então U também é finitamente gerado. Assim, todo o subespaço de um espaço linear de dimensão finita, V, é da forma $L_V(\{\mathrm{e}_1,\ldots,\mathrm{e}_s\})$.

Retomando o exemplo 10, fomos capazes de caracterizar os elementos do subespaço de $\mathbb{R}^3$ aí considerado através de equações lineares homogéneas i.e., descrevendo-o como o núcleo de uma matriz.

A situação pode generalizar-se.

Lema 2.8.— *Seja* $A\in\mathbb{K}^{m\times n}$. *Tem-se que* $\mathrm{Nuc}(A)$ *é um subespaço de* $\mathbb{K}^n$.

Demonstração.— Tem-se que $A\mathbb{0}=\mathbb{0}$ assim $\mathbb{0}\in\mathrm{Nuc}(A)$. Por outro lado, se $\mathrm{u},\mathrm{v}\in\mathrm{Nuc}(A)$ então, $A(\alpha\mathrm{u}+\beta\mathrm{v})=\alpha A\mathrm{u}+\beta A\mathrm{v}=\mathbb{0}+\mathbb{0}=\mathbb{0}$, pelo que $\alpha\mathrm{u}+\beta\mathrm{v}\in\mathrm{Nuc}(A)$. Desta forma $\mathrm{Nuc}(A)$ é fechado para combinações lineares. Conclui-se que é um subespaço de $\mathbb{K}^n$. ⍁

[6] Iremos ocupar-nos essencialmente com os espaços de dimensão finita. Desta forma, ao longo deste texto, o termo *espaço linear* deve ser entendido como sinónimo de *espaço de dimensão finita*.

[7] Essencialmente, um espaço linear V sobre $\mathbb{K}$ é essencialmente equivalente a um espaço $\mathbb{K}^n$ para algum n e, qualquer problema (de álgebra linear) envolvendo V, pode ser reduzido a um problema do mesmo tipo em $\mathbb{K}^n$. Em particular, o problema de saber se $\mathrm{w}\in L_V\mathrm{u}_1,\ldots,\mathrm{u}_k)$ (em V) pode ser resolvido em $\mathbb{K}^n$ resolvendo um problema do mesmo tipo $\mathrm{w}^*\in L_V\mathrm{u}_1^*,\ldots,\mathrm{u}_k^*)$ onde $\mathrm{w}^*,\mathrm{u}_1^*,\ldots,\mathrm{u}_k^*$ são vectores de coordenadas de $\mathrm{w},\mathrm{u}_1,\ldots,\mathrm{u}_k$, respectivamente.

2.2 Bases e dimensão

Recordamos que se $V = L_V(X)$ então dizemos que X é um *conjunto gerador* de V. Como $L_V(\varnothing) = \{\mathbb{0}_V\}$ o subespaço $\{\mathbb{0}_V\}$ é gerado pelo conjunto vazio. Além disso, o conjunto vazio é o seu único gerador pois se $X \neq \varnothing$ então $L_V(X) \neq \{\mathbb{0}_V\}$. Este último facto é uma consequência directa de uma propriedade geral: quaisquer que sejam $X \subset V$ e $x \in X$ tem-se sempre que $1x \in \mathscr{L}(X)$ o que implica $X \subset L_V(X)$. Além disso, todo o espaço linear V possui trivialmente um conjunto gerador dado que $L_V(V) = V$.

Claro que o facto $V = L_V(V)$ tem pouco interesse prático. Seria interessante poder aceder à estrutura de V a partir de um conjunto mínimo de informação ou, pelo menos, a partir de um conjunto relativamente pequeno de informação. Num caso típico tem-se $V = L_V(X)$ com X finito. Quando isto acontece, o espaço V diz-se um *espaço de dimensão finita*.

2.2.1 Espaços de dimensão finita

Esta secção é acerca do problema de determinar conjuntos geradores óptimos no contexto dos espaços lineares de dimensão finita. Entendemos aqui o termo *óptimos* no sentido de *o mais económicos possível*. Para ilustrar a questão: se $V = L_V(\{\mathrm{x}, 2\mathrm{x}\})$ (onde V é um e.l. sobre $\mathbb{R}$) então o conjunto $\{\mathrm{x}, 2\mathrm{x}\}$ não é óptimo: de facto se $\mathrm{y} \in V$ é uma combinação linear dos vectores x e 2x, i.e. se se tem $y = \alpha\mathrm{x} + \beta(2\mathrm{x})$ para certos $\alpha, \beta \in \mathbb{R}$ então:

$$\mathrm{y} = \alpha\mathrm{x} + \beta(2\mathrm{x}) = (\alpha + 2\beta)\mathrm{x},$$

o que, visto y poder ser qualquer, revela que $L_V(\{\mathrm{x}, 2\mathrm{x}\}) = L_V(\{\mathrm{x}\})$. Aquilo que impede que $\{\mathrm{x}, 2\mathrm{x}\}$ seja um conjunto gerador óptimo é o facto de um dos seus elementos, e.g. 2x ser já uma combinação linear dos restantes (neste caso apenas x), podendo ser eliminado sem restringir a respectiva expansão linear. Estas observações sugerem a definição seguinte.

Definição 2.9 (conjunto linearmente dependente). —
Um conjunto de vectores $\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}$ *diz-se* linearmente dependente *se algum dos vectores* x_i *for combinação linear dos restantes, ou seja se para algum* $i \in \{1, \ldots n\}$ *se tem* $\mathrm{x}_i \in L_V(X \setminus \{\mathrm{x}_i\})$.[8]

[8] A propósito desta notação, recorde-se que se X, Y são conjuntos então, $X \setminus Y$ (diz-se X *excepto* Y) denota o conjunto que contém os elementos de X que não são também elementos de Y.

Definição 2.10 (conjunto linearmente independente). —
Um conjunto de vectores $\{x_1, \ldots, x_n\}$ *diz-se* linearmente independente *se*

não for linearmente dependente, i.e. se nenhum dos vectores no conjunto for uma combinação linear dos restantes.

Observe-se que $\mathbb{0} \in L_V(X \setminus \{\mathbb{0}\})$ pelo que, *se um conjunto X contém o vector nulo ele é necessariamente linearmente dependente.*

O resultado seguinte fornece critérios alternativos para a verificação da independência linear de um conjunto de vectores.

TEOREMA 2.11.— *Sejam, V um espaço linear sobre $\mathbb{K}$ e $X = \{x_1, \dots, x_n\}$ um subconjunto de V. Nestas condições são equivalentes:*

(a) *X é linearmente independente.*

(b) *Se $\mathbb{0} = \alpha_1 x_1 + \cdots + \alpha_n x_n$ então $\alpha_1 = \cdots = \alpha_n = 0$, ou seja, a única forma de obter o vector nulo como combinação linear de $x_1, \dots, x_n$ é considerando a combinação linear em que os coeficientes são todos nulos.*

(c) *Se $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta_1 x_1 + \cdots + \beta_n x_n$ então, $\alpha_1 = \beta_1, \dots, \alpha_n = \beta_n$ ou seja, se um vector é uma combinação linear de $x_1, \dots, x_n$ essa combinação linear é única.*

DEMONSTRAÇÃO.—
(a) $\Rightarrow$ (b): Mostramos que $\neg$(a) $\Rightarrow$ $\neg$(b). Suponhamos então que se tem $\mathbb{0} = \alpha_1 x_1 + \cdots + \alpha_n x_n$ e que os escalares $\alpha_1, \dots, \alpha_n$ não são todos nulos. Suponhamos ainda que $i \in \{1, \dots, n\}$ é tal que $\alpha_i \neq 0$. Tem-se:

$$x_i = \frac{1}{\alpha_i} \sum_{\substack{1 \le j \le n \\ j \neq i}} -\alpha_j x_j \in L_V(\{x_1, \dots, x_n\} \setminus \{x_i\}),$$

contradizendo o facto de $\{x_1, \dots, x_n\}$ ser linearmente independente.

(b) $\Rightarrow$ (c): Supondo que $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta_1 x_1 + \cdots + \beta_n x_n$ então,

$$(\alpha_1 - \beta_1)x_1 + \cdots + (\alpha_n - \beta_n)x_n = \mathbb{0}.$$

Usando (b) concluímos que se tem $\alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0$, ou seja, $\alpha_1 = \beta_1, \dots, \alpha_n = \beta_n$, como se pretendia.

(c) $\Rightarrow$ (a): Mais uma vez optamos por mostrar $\neg$(a) $\Rightarrow$ $\neg$(c). Suponhamos que $\{x_1, \dots, x_n\}$ é linearmente dependente. Então um dos vectores, digamos que x_i, é uma combinação linear dos restantes:

$$x_i = \sum_{\substack{1 \le j \le n \\ j \neq i}} \alpha_j x_j.$$

Nestas condições: $\mathbb{0} = \alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} - x_i + \alpha_{i+1} x_{i+1} + \cdots + \alpha_n x_n$. Ou seja, não existe unicidade na representação do vector nulo, contrariando a condição (c). $\boxslash$

Lema 2.12.— *Sejam V um espaço linear sobre $\mathbb{K}$, $X = \{u_1, \dots, u_n\} \subset V$ linearmente independente e $x \notin X \subset V$. Tem-se que $X \cup \{x\}$ é linearmente independente se e só se $x \notin L_V(X)$.*

Demonstração.— Consideremos X e x nas condições do enunciado e suponhamos que que $X \cup \{\mathrm{x}\}$ é linearmente dependente. Então, existem escalares (não todos nulos) tais que $\alpha_1 \mathrm{u}_1 + \cdots + \alpha_n \mathrm{u}_n + \beta \mathrm{x} = \mathbb{0}$. Não se pode ter $\beta = 0$ pois nesse caso ter-se-ia $\alpha_1 \mathrm{u}_1 + \cdots + \alpha_n \mathrm{u}_n = \mathbb{0}$ contrariando o facto de X ser um conjunto linearmente independente. Mas, tendo-se $\beta \neq 0$, também se tem que:

$$\mathrm{x} = -\frac{\alpha_1}{\beta}\mathrm{x}_1 - \cdots - \frac{\alpha_n}{\beta}\mathrm{x}_n \in L_V(X),$$

contrariando as hipóteses do enunciado. Tem-se assim que $X \cup \{\mathrm{x}\}$ é linearmente independente. ⊠

Os conjuntos geradores que são linearmente independentes são óptimos (no sentido rigoroso descrito no resultado seguinte).

Teorema 2.13.— *Sejam V um espaço linear sobre $\mathbb{K}$ e $X = \{x_1, \dots, x_n\}$ um subconjunto de V. São equivalentes:*

(a) *$V = L_V(X)$ e X é um conjunto linearmente independente.*

(b) *Todo o elemento de V pode ser escrito de modo único como combinação linear dos vectores em X.*

(c) *$V = L_V(X)$ e se $Y \subsetneq X$ então $L_V(Y) \neq V$ (i.e. X é um conjunto gerador de V que é minimal).*

(d) *X é linearmente independente e se $X \subsetneq Y \subset V$ então Y é linearmente dependente (i.e. X é em V um conjunto linearmente independente maximal).*

Demonstração.—
(a)⇒(b): É uma consequência do teorema anterior. Como $V = L_V(X)$ todo o elemento de V é combinação linear dos elementos de X. Como X é linearmente independente, pelo teorema precedente, essa combinação linear é única.

(b)⇒(c): Admitindo que existe $Y \subsetneq X$ tal que $L_V(Y) = L_V(X)$, consideremos um $x_i \in X$ tal que $x_i \notin Y$. Por hipótese temos que

$$x_i = \sum_{x \in Y} \alpha_x x.$$

Como $x_i = 1x_i$, temos que $1x_i$ e $\sum_{x \in Y} \alpha_x x$ são duas combinações lineares distintas de elementos de X que permitem obter o vector x_i. Esta circunstância contraria, no entanto, (b).

(c)⇒(d): Se Y não fosse dependente então seria um conjunto gerador de V linearmente independente, mas não seria minimal, contrariando (c).

(c)⇒(a): Se não se tivesse $\mathscr{L}_V(X) = V$ então poderíamos considerar $y \notin \mathscr{L}_V(X)$ e então ter-se-ia $X \subsetneq X \cup \{y\}$ com $X \cup \{y\}$ linearmente independente, contrariando assim (d). ◪

De acordo com o resultado precedente um conjunto gerador de um espaço que seja linearmente independente é um conjunto minimal de vectores que gera esse espaço, i.e. nenhuma das suas partes próprias pode gerar todo o espaço. Também é maximal no seguinte sentido, se lhe acrescentarmos outros vectores, embora continue a gerar o espaço, deixa de ser linearmente independente.

Definição 2.14 (base de um espaço linear).— *Um conjunto $X \subset V$ que satisfaça uma qualquer das condições do teorema precedente diz-se uma* base de V.

Teorema 2.15 (existência de bases).— *Suponhamos que V é um espaço linear sobre $\mathbb{K}$, finitamente gerado. Nestas condições V tem uma base.*[9]

[9] De facto, se um espaço linear possui uma base então possui uma infinidade de bases.

Demonstração.— O caso $V = \{\mathbb{0}_V\}$ possui trivialmente uma base que é o conjunto vazio. Consideremos então o caso em que $V \neq \{\mathbb{0}_V\}$. Uma vez que V é finitamente gerado, existe um conjunto (não vazio) $X = \{u_1, \dots, u_n\}$ tal que $V = L_V(X)$. Entre as partes de X é sempre possível encontrar uma, digamos Y, que constitua um conjunto linearmente independente, maximal.[10] Neste caso $L_V(Y) = V$. Para constatar este facto basta observar que para cada $1 \leq i \leq n$ se tem que $Y \cup \{x_i\}$ ou é Y ou é um conjunto linearmente independente. Assim, $\{x_1, \dots, x_n\} \subset L_V(Y)$, ou seja, $V = L_V(X) \subset L_V(Y) \subset V$. Concluímos desta forma que Y é uma base de V. ◪

[10] I.e., um subconjunto $Y \subset X$ tal que Y é linearmente independente e dado qualquer $Z \subset X$ tal que $Y \subset Z$ se tem que Z é linearmente dependente. (Um tal Y pode ser obtido da seguinte forma: definimos a sequência $Y_1 \subset Y_2 \subset \cdots \subset Y_n$ de subconjuntos de X da seguinte forma $Y_1 = \{u_1\}$ se $u_1 \neq \mathbb{0}$ e $Y_1 = \emptyset$ caso contrário. Se $i < n$ e Y_i está definido então, $Y_{i+1} = Y_i \cup \{x_{i+1}\}$ se este conjunto é linearmente independente, caso contrário $Y_{i+1} = Y_i$.

Teorema 2.16.— *Um conjunto $\{x_1, \dots, x_n\}$ é uma base de V se e só se*

$$V = L_V(\{x_1\}) \oplus \cdots \oplus L_V(\{x_n\}).$$

2.2.2 dimensão

Um facto interessante é que se $V = L_V(X)$ então podemos isolar um subconjunto $Y \subset X$ que é uma base de V. (Desta forma qualquer espaço de dimensão finita tem uma base.) Por outro lado, se X e Y são bases de um espaço V então a cardinalidade de X (i.e. a quantidade de elementos de X) é igual à cardinalidade de Y. (Se X é um conjunto finito, denotamos por $|X|$ a respectiva cardinalidade.)

LEMA 2.17.— *Seja V um e.l. sobre $\mathbb{K}$. Suponhamos que $X, Y \subset V$ e $Y \subset L_V(X)$. Nestas condições, $L_V(Y) \subset L_V(X)$.*

DEMONSTRAÇÃO.— A conclusão decorre de uma simples aplicação do facto de um subespaço ser fechado para combinações lineares i.e., contém todas as combinações lineares dos seus elementos. ⧄

LEMA 2.18.— *Seja V um espaço linear sobre $\mathbb{K}$. Suponhamos que $X \subset V$ e $\mathbb{0} \neq y \in L_V(X)$. Nestas condições, existe um $x \in X$ tal que*

$$L_V(X) = L_V((X \setminus \{x\}) \cup \{y\}).$$

DEMONSTRAÇÃO.— Suponhamos que $y = \alpha_1 x_1 + \cdots + \alpha_n x_n$. Como $y \neq \mathbb{0}$ pelo menos um dos escalares α_i naquela combinação linear não é nulo. Desta forma obtemos:

$$-\alpha_i x_i = -y + \sum_{j \neq i} \alpha_j x_j.$$

ou seja,

$$x_i = \frac{1}{\alpha_i} y - \sum_{j \neq i} \frac{\alpha_j}{\alpha_i} x_j.$$

Tem-se assim que $L_V(X) \subset L_V((X \setminus \{x_i\}) \cup \{y\}) \subset L_V(X)$, ou seja, $L_V(X) = L_V((X \setminus \{x_i\}) \cup \{y\})$. ⧄

TEOREMA 2.19.— *Se $V = L_V(X)$ e $Y \subset V$ é linearmente independente então $|Y| \leq |X|$.*

DEMONSTRAÇÃO.— Suponhamos que, contrariando a conclusão do teorema, se tem que $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_k\}$ e $k = n + r$, para algum $r \geq 1$. Consideremos os dois conjuntos lado a lado:

$$\{x_1, x_2, \ldots, x_n\} \quad \{y_1, \ldots y_n, y_{n+1}, \ldots, y_{n+r}\}$$

Mudando o vector y_1 para o primeiro conjunto obtemos:

$$\{y_1, x_1, x_2, \ldots, x_n\} \quad \{y_2, \ldots y_n, y_{n+1}, \ldots, y_{n+r}\}.$$

Uma vez que y_1 é combinação linear dos x_i, recorrendo ao resultado precedente podemos eliminar um dos x_i e continuar a gerar V. Sem perda de generalidade, podemos supor que podemos eliminar x_1. Concluímos então que os conjuntos

$$\{y_1, x_2, \ldots, x_n\} \quad \{y_2, \ldots y_n, y_{n+1}, \ldots, y_{n+r}\}.$$

são respectivamente um conjunto gerador de V e um conjunto linearmente independente de vectores. O procedimento anterior pode assim ser repetido até ao momento em que obtemos conjuntos:

$$\{y_1, y_2, \ldots, y_n\} \quad \{y_{n+1}, \ldots, y_{n+r}\}.$$

o da esquerda, um conjunto gerador de V, o da direita um conjunto linearmente independente.

Temos assim que $y_{n+1} \in L_V(\{y_1, y_2, \ldots, y_n\})$ contrariando a independência linear de Y. Não se pode pois ter $k > n$ e o teorema fica estabelecido. ⧄

Corolário 2.19.1.— *Se $B_1, B_2 \subset V$ são bases de V então $|B_1| = |B_2|$.*

Demonstração.— Pelo teorema precedente B_1 e B_2 são conjuntos geradores, linearmente independentes, de V. Assim, como B_1 é gerador e B_2 é linearmente independente tem-se que $|B_2| \leq |B_1|$. Trocando os papéis de B_1 e B_2 concluímos que $|B_1| \leq |B_2|$ tendo-se finalmente que $|B_1| = |B_2|$, como pretendido. ⧄

Definição 2.20 (dimensão).— *Seja V um espaço linear sobre $\mathbb{K}$. A dimensão de V, que se denota* $\dim(V)$ *é a cardinalidade de uma (e neste caso de qualquer) base de V.*

Teorema 2.21.— *Seja V um espaço linear sobre $\mathbb{K}$.*

1. *Seja B é uma base de V tal que $B = B_1 \cup B_2$ onde $B_1 \cap B_2 = \emptyset$ então, $V = L_V(B_1) \oplus L_V(B_2)$.*
2. *Se $V = U \oplus W$, B_1 é uma base de U e B_2 é uma base de W então $B_1 \cap B_2 = \emptyset$ e $B_1 \cup B_2$ é uma base de V.*

Teorema 2.22.— *Suponhamos que $U, W \in S(V)$. Tem-se que:*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Em particular:

$$\dim(U \oplus W) = \dim(U) + \dim(W).$$

2.2.3 espaços das linhas e das colunas de uma matriz. característica

Suponhamos que $A \in \mathbb{K}^{m \times n}$. As linhas de A podem ser vistas como elementos de $\mathbb{K}^n$ enquanto que as colunas de A podem ser vistas como elementos de $\mathbb{K}^m$. Feita esta observação, definimos o *espaço das colunas* de A, que denotamos por, $\mathrm{Col}(A)$, é o subespaço de $\mathbb{K}^m$ gerado pelas colunas de A. Analogamente, o espaço das linhas de A, que se denota por $\mathrm{Lin}(A)$, é o subespaço de $\mathbb{K}^n$ gerado pelas linhas de A.

Teorema 2.23.— *Suponhamos que $A \in \mathbb{K}^{m\times n}$. As operações do método de eliminação de Gauss não alteram o espaço das linhas de A. Além disso não alteram a dimensão do espaço das colunas.*

Teorema 2.24.— *Suponhamos que $A \in \mathbb{K}^{m\times n}$. Tem-se que:*

$$\dim(\mathrm{Lin}(A)) = \dim(\mathrm{Col}(A)).$$

Demonstração.— Seja $A \in \mathbb{K}^{m\times n}$ e suponhamos que a dimensão de $\mathrm{Col}(A)$ é r. Fixemos uma base $\mathrm{c}_1, \dots, \mathrm{c}_r$ de $\mathrm{Col}(A)$. Consideremos a matriz $C \in \mathbb{K}^{m\times r}$ cujas colunas são os vectores c_r, i.e. tem-se $C_{*,j} = \mathrm{c}_j$, para $1 \leq j \leq r$. As restantes $n-r$ colunas de A são combinações lineares de $\mathrm{c}_1, \dots, \mathrm{c}_r$. Isto significa que existe uma matriz $R \in \mathbb{K}^{r\times n}$ tal que $A = CR$. (R é a matriz onde $R_{*,j}$ é o vector de coordenadas da j-ésima coluna de A na base $\mathrm{c}_1, \dots, \mathrm{c}_r$.

Agora, cada linha de A é uma combinação linear das r linhas de R e, desta forma, a dimensão do espaço das linhas não pode exceder r.

Uma vez que esta conclusão vale para qualquer matriz, vale para a matriz $A^\top$ onde o papel das linhas e colunas é trocado. Assim concluímos que as duas dimensões devem coincidir. ◪

Definição 2.25.— *Tendo em causa o resultado anterior, a* característica *de uma matriz A é*

$$\mathrm{car}(A) \coloneqq \dim(\mathrm{Lin}(A)) = \dim(\mathrm{Col}(A)).$$

Observação 1.— Tendo em conta as considerações anteriores, se A^* é uma matriz em escada de linhas, que se obtém de A usando o método de eliminação de Gauss então

$$\mathrm{car}(A) = \mathrm{car}(A^*) = \text{número de pivôs de } A^*.$$

2.2.4 Bases ordenadas e coordenadas

Definição 2.26.— *Seja V um espaço linear de dimensão n. Uma base ordenada de V é um n-úplo ordenado de vectores de V, $(\mathrm{e}_1, \dots, \mathrm{e}_n)$, tal que $\{\mathrm{e}_1, \dots, \mathrm{e}_n\}$ é uma base de V.*

Observe-se que qualquer conjunto $\{\mathrm{v}_1, \dots, \mathrm{v}_n\}$ origina $n!$ n-úplos distintos, desta forma qualquer base B de V, com $|B| = n$, origina $n!$ bases ordenadas distintas.

Suponhamos que $B = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ é uma base ordenada de V. Dado $\mathrm{x} \in V$ existe uma única sequência de escalares $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ tal que:

$$\mathrm{x} = \alpha_1 \mathrm{e}_1 + \cdots + \alpha_n \mathrm{e}_n.$$

O vector $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$, contendo os coeficientes daquela combinação linear diz-se o *vector de coordenadas de* x *na base* B e denota-se x_B. Reciprocamente, um qualquer vector $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ determina, relativamente à base B um único vector de V, respectivamente:

$$\mathrm{x} = (\alpha_1, \ldots, \alpha_n)^B = \alpha_1 \mathrm{e}_1 + \cdots + \alpha_n \mathrm{e}_n.$$

Podemos até definir uma função $\phi_B : V \to \mathbb{K}^n$ que associa a cada vector de V o respectivo vector de coordenadas na base B, i.e. $\phi_B(\mathrm{x}) = \mathrm{x}_B$. Esta função possui propriedades interessantes:

(1) ϕ_B tem uma inversa $\phi_B^{-1} : \mathbb{K}^n \to V$ que é definida por:

$$\phi_B^{-1}(\alpha_1, \ldots, \alpha_n) = (\alpha_1, \ldots, \alpha_n)^B.$$

São válidas as seguintes igualdades:

$$(\mathrm{x}_B)^B = \mathrm{x};$$
$$((\alpha_1, \ldots, \alpha_n)^B)_B = (\alpha_1, \ldots, \alpha_n).$$

(2) ϕ_B preserva combinações lineares, i.e.

$$(\alpha_1 \mathrm{x}_1 + \cdots + \alpha_r \mathrm{x}_r)_B = \alpha_1 (\mathrm{x}_1)_B + \cdots + \alpha_r (\mathrm{x}_r)_B.$$

ou seja, o vector de coordenadas de uma combinação linear de vectores é a (mesma) combinação linear dos respectivos vectores de coordenadas.

(3) ϕ_B^{-1} também preserva combinações lineares no sentido da alínea precedente.[11]

Se $X \subset V$ e $W \subset \mathbb{K}^n$ então, definimos

$$\phi_B[X] := \{\mathrm{x}_B \mid \mathrm{x} \in X\};$$
$$\phi_B^{-1}[W] = \{(\xi_1, \ldots, \xi_n)^B \mid (\xi_1, \ldots, \xi_n) \in W\}.$$

(4) $X \leqslant V$ sse $\phi_B[X] \leqslant \mathbb{K}^n$;

(5) $W \leqslant \mathbb{K}^n$ sse $\phi_B^{-1}[W] \leqslant V$;

(6) $X \subset V$ é linearmente independente (resp. dependente) sse $\phi_B[X]$ é linearmente independente (resp. dependente);

(7) $W \subset \mathbb{K}^n$ é linearmente independente (resp. dependente) sse $\phi_B^{-1}[W]$ é linearmente independente (resp. dependente);

(8) $X = L_V(\mathrm{x}_1, \ldots, \mathrm{x}_k)$ se e só se $\phi_B[X] = L_{\mathbb{K}^n}((\mathrm{x}_1)_B, \ldots, (\mathrm{x}_k)_B)$;

(9) $W = L_{\mathbb{K}^n}(\mathrm{w}_1, \ldots, \mathrm{w}_k)$ se e só se $\phi_B^{-1}[W] = L_V((\mathrm{w}_1)^B, \ldots, (\mathrm{w}_k)^B)$.

[11] Uma função que preserva combinações lineares diz-se uma *transformação linear*. Assim a função ϕ_B é uma transformação linear. Por outro lado a inversa de uma transformação linear é sempre uma transformação linear (ver o próximo capítulo).

Exemplo 12.— Consideremos o espaço $\mathbb{R}_3[t]$ dos polinómios reais de grau ≤ 3. Neste espaço consideremos a base ordenada

$$B = (1, 1 + t, t^2, t^2 + t^3).$$

1. Determinar $(t + t^3)_B$;
2. Determinar $(0, 1, -1, 1)^B$;

Para calcular $(t + t^3)_B$, ou seja, o vector de coordenadas de $t + t^3$ na base B, temos que descobrir os coeficientes da combinação linear dos vectores de B que coincide com $t + t^3$, i.e.

$$(t + t^3)_B = (\xi_1, \xi_2, \xi_3, \xi_4) \Leftrightarrow t + t^3 = \xi_1 + \xi_2(1 + t) + \xi_3 t^2 + \xi_4(t^2 + t^3).$$

A igualdade do lado direito do símbolo de equivalência ocorre sse

$$t + t^3 = (\xi_1 + \xi_2) + \xi_2 t + (\xi_3 + \xi_4)t^2 + \xi_4 t^3.$$

Recordando que dois polinómios são iguais sse os coeficientes correspondentes são iguais, a última igualdade verifica-se sse

$$\begin{cases} \xi_1 + \xi_2 = 0 \\ \xi_2 = 1 \\ \xi 3 + \xi_4 = 0 \\ \xi_4 = 1 \end{cases}$$

ou seja, se e só se $\xi_1 = -1, \xi_2 = 1, \xi_3 = -1, \xi_4 = 1$. Conclui-se então que $(t + t^3)_B = (-1, 1, -1, 1) \in \mathbb{R}^4$.

Quanto à segunda questão, por definição:

$$(0, 1, -1, 1)^B = 0 \cdot 1 + 1(1 + t) + (-1)t^2 + 1(t^2 + t^3)$$

ou seja, $(0, 1, -1, 1)^B = 1 + t + t^3$.

Exemplo 13.— Se $(\alpha_1, \dots, \alpha_n)$ e B_c é a base canónica de $\mathbb{K}^n$ então:

$$(\alpha_1, \dots, \alpha_n)_{B_c} = (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n,$$

ou seja fixando a base canónica em $\mathbb{K}^n$ o vector de coordenadas de um qualquer vector coincide com o próprio vector. Da mesma forma, se estiver fixa a base canónica em $\mathbb{K}^n$ então:

$$(\alpha_1, \dots, \alpha_n)^{B_c} = (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n.$$

Exemplo 14.— Consideremos o espaço $V = \mathbb{R}^{2\times 2}$. Será que:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \in L_V\left(\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\}\right)?$$

Tendo em conta as propriedades da função ϕ_B e da sua inversa. Podemos responder a esta questão *colocando a mesma questão acerca das coordenadas dos objectos envolvidos (numa qualquer base de* $\mathbb{R}^{2\times 2}$*).* As bases canónicas dos diferentes espaços são em geral as mais convenientes para este propósito, pois os vectores de coordenadas são fáceis de calcular. Se fixarmos em $\mathbb{R}^{2\times 2}$ a base canónica i.e., a base $B = (E^{(1,1)}, E^{(1,2)}, E^{(2,1)}, E^{(2,2)})$ então, tem-se que

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}_B = (a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}).$$

O problema inicial, em termos de coordenadas na base B é então:

$$(1,2,3,4) \in L_{\mathbb{R}^4}(\{(1,0,0,1),(1,0,0,-1),(0,1,1,0)\})?$$

Ora, sabemos que a resposta a esta questão é positiva se e só se o sistema

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

é possível. Feita a eliminação de Gauss:

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

ou seja a resposta é negativa, tanto para a questão envolvendo as coordenadas como para a questão que envolve os vectores originais.

Este último exemplo a importância dos vectores de coordenadas. Através desta noção qualquer espaço de dimensão finita é essencialmente uma versão de um $\mathbb{K}^n$. No entanto, os vectores de coordenadas são calculados relativamente a uma base e, por diversas razões certas bases são mais interessantes que outras (ou seja, nem sempre as bases canónicas são as mais interessantes). Vamos agora descrever um esquema que nos permite de uma forma sistemática converter coordenadas numa base em coordenadas noutra.

Dadas duas bases ordenadas $B = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ e $\bar{B} = (\bar{\mathrm{e}}_1, \ldots, \bar{\mathrm{e}}_n)$ de V a *matriz de mudança da base $\bar{B}$ para a base B*, que se denota $M_{B\leftarrow\bar{B}}$ é a matriz que satisfaz a seguinte condição:

$$(M_{B\leftarrow\bar{B}})_{*,i} = (\bar{e}_i)_B, \qquad i = 1, \ldots, n. \tag{2.2}$$

ou seja a i-ésima coluna de $M_{B\leftarrow\bar{B}}$ é o vector de coordenadas do i-ésimo vector da base $\bar{B}$ na base B.

Lema 2.27.— *Sejam B e $\bar{B}$ bases ordenadas de um espaço V, Se A é uma matriz tal que $A\mathrm{x}_{\bar{B}} = \mathrm{x}_B$, para todo o $\mathrm{x} \in V$ então $A = M_{B\leftarrow\bar{B}}$.*

Teorema 2.28.— *Sejam B e $\bar{B}$ duas bases ordenadas de V. Tem-se:*

(1) $\mathrm{x}_B = M_{B\leftarrow\bar{B}}\mathrm{x}_{\bar{B}}$, *para qualquer* $\mathrm{x} \in V$*;*

(2) $M_{\bar{B}\leftarrow B} = M^{-1}_{B\leftarrow\bar{B}}$*;*

(3) *se* $\ddot{B}$ *é uma terceira base ordenada de* V *então,*

$$M_{\ddot{B}\leftarrow\bar{B}} = M_{\ddot{B}\leftarrow B}M_{B\leftarrow\bar{B}}.$$

Exemplo 15.— Considerando as bases ordenadas

$$B = ((0,1,1),(1,0,1),(1,1,0)) \text{ e } \bar{B} = ((1,0,0),(1,1,0),(1,1,1))$$

de $\mathbb{R}^3$, podemos obter a matriz $M_{B\leftarrow\bar{B}}$ de duas formas. Primeiro recorrendo à definição: esta matriz tem nas suas colunas os vectores de coordenadas $(1,0,0)_B, (1,1,0)_B$ e $(1,1,1)_B$.

Ora o problema de determinar as coordenadas de um vector (x, y, z) em B é simples e equivale a resolver o sistema (digamos que nas variáveis α, β, γ):

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & x \\ 1 & 0 & 1 & y \\ 1 & 1 & 0 & z \end{array}\right]$$

Resolvido o sistema, a solução geral é da forma:

$$\left(\frac{z+y-x}{2}, \frac{z-y+x}{2}, \frac{x+y-z}{2}\right)$$

ou seja,

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

pelo que:

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

Usando este facto concluímos que:

$$(1,0,0)_B = (-1/2, 1/2, 1/2),$$

$$(1,1,0)_B = (0,0,1)\ ;$$

$$(1,1,1)_B = (1/2, 1/2, 1/2).$$

Ou seja,

$$M_{B\leftarrow\bar{B}} = \begin{bmatrix} -1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1 & 1/2 \end{bmatrix}.$$

Podemos contornar a definição de matriz de mudança de base, explorando as propriedades algébricas deste tipo de matrizes e o facto de as matrizes de mudança de uma base para a base canónica se poderem escrever directamente (recorde-se que o vector de coordenadas na base canónica de um vector de $\mathbb{K}^n$ coincide com o próprio vector). Denotando a base canónica de $\mathbb{R}^3$ por $\boldsymbol{B}_c$ tem-se então que:

$$M_{B_c \leftarrow B} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad M_{B_c \leftarrow \bar{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Tem-se então que:

$$M_{B \leftarrow \bar{B}} = M_{B \leftarrow B_c} M_{B_c \leftarrow \bar{B}} = M_{B_c \leftarrow B}^{-1} M_{B_c \leftarrow \bar{B}}$$

ou seja,

$$M_{B \leftarrow \bar{B}} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

2.3 O RETICULADO DOS SUBESPAÇOS DE UM ESPAÇO LINEAR

DEFINIÇÃO 2.29.— *Denotamos por* $\mathrm{Sub}(V)$ *o conjunto de todos os subespaços de um espaço linear* V.

Uma relação binária $\leq$ num conjunto X é uma ordem parcial se for *reflexiva*, *anti-simétrica* e *transitiva* i.e.,

1. $x \leq x$, para qualquer $x \in X$;
2. $x \leq y \wedge y \leq x \Rightarrow x = y$, para quaisquer $x, y \in X$;
3. $x \leq y \wedge y \leq z \Rightarrow x \leq z$, para quaisquer $x, y, z \in X$.

Se $(X, \leq)$ é uma ordem parcial, e $x, y \in X$ o *supremo* de x e y, que se denota $x \vee y$, é (se existir) o elemento $z \in X$ que satisfaz:

$$x, y \leq z \wedge (\forall w \in X)[x, y \leq w \Rightarrow z \leq w].$$

Analogamente, o *ínfimo* de x e y, que se denota $x \wedge y$, é (se existir) o elemento $z \in X$ que satisfaz:

$$z \leq x, y \wedge (\forall w \in X)[w \leq x, y \Rightarrow w \leq z].$$

Podemos generalizar as noções de supremo e ínfimo a qualquer conjunto (não-vazio) de elementos de X. Assim, se $\varnothing \neq A \subset X$ o supremo de A, que se denota $\sup A$ é o $z \in X$ (se existir) que satisfaz:

$$(\forall x \in A) x \leq z \wedge (\forall w \in X)[(\forall x \in A) x \leq w \Rightarrow z \leq w].$$

Analogamente, o ínfimo de A, que se denota $\inf A$ é o $z \in X$ (se existir) que satisfaz:

$$(\forall x \in A) z \leq x \wedge (\forall w \in X)[(\forall x \in A) w \leq x \Rightarrow w \leq z].$$

Observe-se que $x \vee y = \sup\{x, y\}$ e $x \wedge y = \inf\{x, y\}$.

O conjunto $\mathrm{Sub}(V)$ é parcialmente ordenado pela relação de inclusão de conjuntos. Como $V \in \mathrm{Sub}(V)$ tem-se que $\mathrm{Sub}(V)$ tem um elemento máximo, pois para qualquer $U \in \mathrm{Sub}(V)$ tem-se que $U \subset V$. Por outro lado, $\{\mathbb{0}\} \in \mathrm{Sub}(V)$ pelo que $\mathrm{Sub}(V)$ também tem um elemento mínimo, precisamente o subespaço $\{\mathbb{0}\}$.

A estrutura $(\mathrm{Sub}(V), \subset)$ para além de ser uma ordem parcial possui uma estrutura rica do ponto de vista algébrico, uma estrutura vulgarmente designada de *reticulado*.

2.3.1 RETICULADOS

DEFINIÇÃO 2.30 (RETICULADO).— *Uma ordem parcial* $(L, \leq)$ *é um* reticulado *se dados quaisquer* $x, y \in L$ *existirem sempre* $x \vee y$ *e* $x \wedge y$*. Se um reticulado tem um elemento máximo ele denota-se por* $\mathbb{1}$ *e se tem um mínimo ele denota-se por* $\mathbb{0}$*. Se existirem o ínfimo e o supremo de qualquer subconjunto de* L *então o reticulado diz-se um* reticulado completo.

Como é claro, se $(L, \leq)$ é um reticulado então o ínfimo e o supremo de qualquer conjunto finito existem sempre, e.g.

$$\sup\{x_1, x_2 \ldots, x_n\} = x_1 \vee (\sup\{x_2, \ldots, x_n\})$$

e, analogamente

$$\inf\{x_1, x_2 \ldots, x_n\} = x_1 \wedge (\inf\{x_2, \ldots, x_n\}).$$

Tem-se o seguinte resultado:

Teorema 2.31.— *Sejam, V um espaço linear sobre $\mathbb{K}$ e $U, W \in \mathrm{Sub}(V)$. Nestas condições,* $\sup\{U, W\}$ *i.e. o menor subespaço de V que contém U e W, e* $\inf\{U, W\}$ *i.e., o maior subespaço de V que está contido em U e em W, existem.*[12] *Além disso tem-se que:*

1. $\sup\{U, W\} = U + W := \{\mathrm{u} + \mathrm{w} \mid \mathrm{u} \in U \wedge \mathrm{w} \in W\}$.
2. $\inf\{U, W\} = U \cap W$.

O teorema anterior generaliza-se sem grande dificuldade. Além disso como este teorema é um caso particular do próximo não faremos a demonstração deste.

Teorema 2.32.— *Sejam V um espaço linear sobre $\mathbb{K}$ e $X \subset \mathrm{Sub}(V)$. Nestas condições,* $\sup X$ *e* $\inf X$ *existem. Além disso tem-se que:*

1. $\sup X = \sum X$ *onde* $\sum X$ *consiste nos* $\mathrm{x} = \mathrm{u}_0 + \cdots + \mathrm{u}_n$ *em que* $n \in \mathbb{N}$, $Z = \{U_0, \ldots, U_n\} \subset X$ *e* $\mathrm{u}_0 \in U_0, \ldots, \mathrm{u}_n \in U_n$. *Dizemos que Z é um* suporte *de* x.[13]
2. $\inf X = \cap X = \{\mathrm{u} \mid (\forall U \in X)\mathrm{u} \in U\}$.

Demonstração.— Sejam $S = \sum X$ e $I = \cap X$.

1. Como qualquer $\mathrm{u} \in U \in X$ pode ser escrito na forma $1\mathrm{u}$ tem-se que se $U \in X$ então $U \subset S$. (Em particular $\mathbb{0} \in S$.) Vejamos agora que S é fechado para as operações. Consideremos então $\mathrm{x}, \mathrm{y} \in S$. Sem perda de generalidade podemos supor que x e y têm um suporte comum i.e., existem $\mathrm{u}_1, \mathrm{w}_1 \in U_1 \in X, \ldots, \mathrm{u}_n, \mathrm{w}_n \in U_n \in X$ tais que:

$$\mathrm{x} = \sum_{i=1}^{n} \xi_i \mathrm{u}_i, \quad \mathrm{y} = \sum_{i=1}^{n} \zeta_i \mathrm{w}_i.$$

Tem-se então que:

$$\begin{aligned} \alpha\mathrm{x} + \beta\mathrm{y} \quad &= \alpha\left(\sum_{i=1}^{n} \xi_i \mathrm{u}_i\right) + \beta\left(\sum_{i=1}^{n} \zeta_i \mathrm{w}_i\right) \\ &= \sum_{i=1}^{n} (\alpha\xi_i \mathrm{u}_i + \beta\zeta_i \mathrm{w}_i). \end{aligned}$$

Como cada $U_i \leqslant V$ tem-se que $\alpha\xi_i\mathrm{u}_i + \beta\zeta_i\mathrm{w}_i \in U_i$ e assim concluímos que $\mathrm{x}+\mathrm{y} \in S$. Consequentemente S é um subespaço que contém todos os subespaços em X. Supondo agora que $W \leqslant V$ contém todos os subespaços em X então, como W é fechado para combinações lineares, resulta imediatamente que $S \subset X$. Isto significa precisamente que se

[12] Tomamos aqui os termos *maior* e *menor* no sentido da inclusão de conjuntos i.e. um conjunto X é o maior relativamente a uma dada propriedade se satisfaz a propriedade e não está propriamente contido em nenhum outro conjunto que satisfaça essa propriedade. Analogamente, será o menor conjunto com uma dada propriedade se a possuir e não contiver propriamente nenhum outro conjunto que satisfaça essa propriedade.

[13] Se $\mathrm{x} \in \sum X$, Z é um suporte de x, $Z \subset W \subset \mathrm{Sub}(V)$ e W é finito então W também é um suporte de x. De facto, se $Z = \{U_1, \ldots, U_n\}$, se $\mathrm{u}_1 \in U_1, \ldots, \mathrm{u}_n \in U_n$ e se $\mathrm{x} = \mathrm{u}_1 + \cdots + \mathrm{u}_n$ então, se $W = \{U_1, \ldots, U_n, U_{n+1}, \ldots, U_{n+k}\}$ tem-se também que:

$$\mathrm{x} = \mathrm{u}_1 + \cdots + \mathrm{u}_n + \mathbb{0} + \cdots + \mathbb{0}.$$

Deste modo, quando lidamos com um número finito de vectores em $\sum X$ podemos supor que eles têm um suporte comum.

tem $S = \sup X$.

2. Como $\cap X$ é o maior conjunto contido em todos os membros de X, basta demonstrar que $\cap X$ é um subespaço de V. Suponhamos então que $\mathrm{x}, \mathrm{y} \in I$. Isto significa que $(\forall U \in X)\mathrm{x}, \mathrm{y} \in U$. Mas cada um destes U é um subespaço de V pelo que, dados quaisquer escalares $\xi, \zeta \in \mathbb{K}$ se tem $\xi\mathrm{x} + \zeta\mathrm{y} \in U$ i.e., $(\forall U \in X)\xi\mathrm{x} + \zeta\mathrm{y} \in U$. Mas isto significa precisamente que $\xi\mathrm{x} + \zeta\mathrm{y} \in I$. Tem-se então que $I \leqslant V$. ⧄

Teorema 2.33.— *Suponhamos que V é um espaço linear sobre $\mathbb{K}$ onde $\mathbb{K}$ é um corpo infinito. Se $n \in \mathbb{N}$ e $U_1, \ldots, U_n \in \mathrm{Sub}(V) \setminus \{V\}$. Tem-se que $U_1 \cup \cdots \cup U_n \neq V$.*

Este teorema estabelece que se V é um espaço linear sobre um corpo infinito (é o caso de $\mathbb{R}$ ou $\mathbb{C}$) então não pode ser uma união finita de subespaços próprios ($\neq V$).

Demonstração.— Suponhamos que $U_1, \ldots, U_n \in \mathrm{Sub}(V) \setminus \{V\}$ são tais que $U_1 \cup \cdots \cup U_n = V$. Neste caso, sem perda de generalidade, podemos supor que para algum dos U_i se tem:

$$U_i \not\subset \bigcup_{\substack{1 \leq j \leq n \\ j \neq i}} U_j.$$

Depois de uma eventual re-indexação podemos sempre supor que se tem $U_1 \not\subset U_2 \cup \cdots \cup U_n$. Fixemos vectores $\mathrm{w} \in S_1 \setminus (U_2 \cup \cdots \cup U_n)$ e $\mathrm{v} \notin U_1$. Veremos agora que a recta $A = \{\xi\mathrm{w} + \mathrm{v} \mid \xi \in \mathbb{K}\}$ (que contém um número infinito de vectores pois $\mathbb{K}$ é infinito)[14] tem, no máximo, um vector comum com cada U_i (ver figura 2.2) e, neste caso, $A \not\subset U_1 \cup \cdots \cup U_n$, contradizendo o facto de se ter $U_1 \cup \cdots \cup U_n = V$.

Supondo então que $\mathrm{x} = \xi\mathrm{w}+\mathrm{v} \in U_1$ então também seria verdade que $\mathrm{x}-\mathrm{v} = \xi\mathrm{w} \in U_1$. Mas então, para $\xi \neq 0$ ter-se-ia, $\xi^{-1}(\mathrm{x}-\mathrm{v}) = \xi^{-1}\xi\mathrm{w} = \mathrm{w} \in U_1$ o que é absurdo. Tem-se assim que $A \cap U_1 = \{\mathrm{v}\}$. Para $i \geq 2$ usamos um argumento diferente. Suponhamos que para certos $\xi \neq \zeta$ se tem $\xi\mathrm{w} + \mathrm{v}, \zeta\mathrm{w} + \mathrm{v} \in U_i$. Neste caso,

$$\xi\mathrm{w} + \mathrm{v} + \zeta\mathrm{w} + \mathrm{v} \in U_i = (\xi + \zeta)\mathrm{w} + 2\mathrm{v} \in U_i,$$

significando que

$$\mathrm{v} = -\frac{\xi + \zeta}{2}\mathrm{w} \in U_2 \cup \cdots \cup U_n,$$

contrariando a escolha de v. ⧄

[14] Note-se que $\xi\mathrm{w} + \mathrm{v} = \zeta\mathrm{w} + \mathrm{v}$ se e só se $(\xi - \zeta)\mathrm{w} = \mathbb{0}$ o que, como $\mathrm{w} \neq \mathbb{0}$ (porquê?) implica que $\xi = \zeta$.

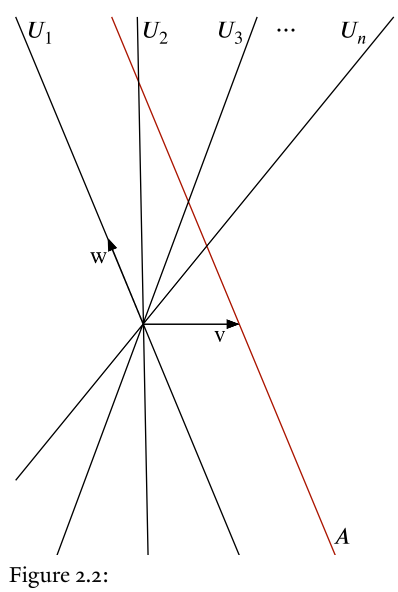


Figure 2.2:

2.3.2 SOMAS DIRECTAS

Já referimos na secção precedente a noção de *soma de dois subespaços*. Recorde-se que dados $U, W \in \mathrm{Sub}(V)$ a respectiva soma é o subespaço de V que denotamos por $U + V$ e é

$$U + V = \{\mathrm{u} + \mathrm{v} \mid \mathrm{u} \in U \wedge \mathrm{v} \in V\}.$$

Quando $U \cap V = \{\mathbb{0}\}$ a soma de U com W designa-se de *soma directa* e denota-se $U \oplus W$. Note-se que $U \oplus W$ coincide com $U + W$ a nova notação serve apenas para indicar que a intersecção de U com W consiste apenas do vector nulo. Apesar disso existem consequências importantes do facto de uma soma ser directa. Todo o elemento de uma soma de dois subespaços U e W (directa ou não) é uma soma de um elemento $\mathrm{u} \in U$ com um elemento $\mathrm{w} \in W$. No caso da soma não ser directa e a intersecção conter um vector não nulo, digamos x, tem-se que $\mathrm{u} + \mathrm{w} = (\mathrm{u} - \mathrm{x}) + (\mathrm{w} + \mathrm{x})$ (sendo que $\mathrm{u} - \mathrm{x} \in U$ e $\mathrm{w} + \mathrm{x} \in W$), ou seja, um elemento em $U + W$ não é neste caso obtido univocamente como a soma de um elemento de U com um elemento de W. No caso de uma soma directa a situação é a oposta:

TEOREMA 2.34.— *Se V é um e.l.,*[15] $U, W \in \mathrm{Sub}(V)$ *e* $\mathrm{x} \in U \oplus W$ *então, existem* $\mathrm{u} \in U$ *e* $\mathrm{w} \in W$ únicos *tais que* $\mathrm{x} = \mathrm{u} + \mathrm{w}$.

[15] muitas vezes escrevemos «e.l.» em lugar de escrever «espaço linear».

DEMONSTRAÇÃO.— Que existem u e w nas condições indicadas resulta imediatamente do facto de uma soma directa ser uma soma. Quanto à unicidade: suponhamos que não temos unicidade e que é possível considerar $\mathrm{x} \in U \oplus W$, bem como $\mathrm{u}_1, \mathrm{u}_2 \in U$ e $\mathrm{w}_1, \mathrm{w}_2 \in W$ tais que:

$$\mathrm{u}_1 + \mathrm{w}_1 = \mathrm{x} = \mathrm{u}_2 + \mathrm{w}_2.$$

Ora daqui conclui-se que $\mathrm{u}_1 - \mathrm{u}_2 = \mathrm{w}_2 - \mathrm{w}_1$ o que, não se tendo $\mathrm{u}_1 = \mathrm{u}_2$ e $\mathrm{w}_1 = \mathrm{w}_2$ implica que $U \cap W$ não é trivial, o que contradiz o facto de a soma ser directa. Tem assim que se ter a unicidade. ◪

O tipo de soma directa que acabámos de descrever designa-se de *soma directa interna* já que envolve dois subespaços de um mesmo espaço linear V.

2.3.3 SOMA E SOMA DIRECTA DE ESPAÇOS—GENERALIZAÇÕES

A noção de soma de dois subespaços generaliza-se sem dificuldade a um qualquer número finito de subespaços de V.

DEFINIÇÃO 2.35.— *Sejam, V um e.l. sobre $\mathbb{K}$ e* $U_1, \ldots, U_n \in \mathrm{Sub}(V)$.

Definimos:

$$\begin{aligned}\textstyle\sum_{i=1}^{n} U_i &= U_1 + \cdots + U_n \\ &= \{\mathrm{u}_1 + \cdots + \mathrm{u}_n \mid \mathrm{u}_1 \in U_1 \wedge \cdots \wedge \mathrm{u}_n \in U_n\}.\end{aligned}$$

Teorema 2.36.— *Sejam, V um e.l. sobre $\mathbb{K}$ e $U_1, \dots, U_n \in \mathrm{Sub}(V)$. Nestas condições, $U_1 + \cdots + U_n \in \mathrm{Sub}(V)$.*

Uma soma $U_1 + \cdots + U_n$ diz-se *directa* se para cada $i = 1, \dots, n$ se tem que:

$$U_i \cap \sum_{j \neq i} U_j = \{\mathbb{0}\}.$$

Se $U_1 + \cdots + U_n$ é uma soma directa (interna) escrevemos $U_1 \oplus \cdots \oplus U_n$ ou $\bigoplus_{i=1}^{n} U_i$ para exprimir claramente esse facto.

Como anteriormente tem-se o seguinte

Teorema 2.37.— *Se $U_1, \dots, U_n \in \mathrm{Sub}(V)$ e $\mathrm{x} \in U_1 \oplus \cdots \oplus U_n$ então existem vectores $\mathrm{u}_1 \in U_1, \dots, \mathrm{u}_n \in U_n$ (únicos), tais que $\mathrm{x} = \mathrm{u}_1 + \cdots + \mathrm{u}_n$.*

Definição 2.38.— *Se $V = U \oplus W$ dizemos que U é um complemento de W em V (ou que W é um complemento de U em V).*

Teorema 2.39.— *Se $U \in \mathrm{Sub}(V)$ então existe $W \in \mathrm{Sub}(V)$ tal que $V = U \oplus W$. (Este W não é único.)*

Concluímos esta secção com o seguinte resultado:

Teorema 2.40.— *Consideremos uma família $\{U_i \mid i = 1, \dots, n\}$ onde $n \in \mathbb{N}$. São equivalentes:*

1. *A família é uma* família independente de subespaços, *i.e. para cada $i = 1, \dots, n$,*
$$S_i \cap \sum_{j \neq i} U_i = \{\mathbb{0}\}.$$
2. Unicidade de representação do vector nulo *i.e. a única maneira de obter $\mathbb{0}$ como soma de vectores dos $U_1, \dots, U_n$ é somando vectores nulos.*
3. *Cada vector $\mathrm{x} \in U_1 + \cdots + U_n$ pode ser obtido de modo único (excepto a ordem das parcelas) como uma soma $\mathrm{u}_1 + \cdots + \mathrm{u}_n$ onde para $i = 1, \dots, n$ se tem $\mathrm{u}_i \in U_i$.*

Assim, uma soma é directa se e só se uma destas condições se verifica.

Um outro tipo de soma directa consiste nas denominadas *somas directas externas* que envolvem espaços vectoriais que não são necessariamente subespaços de um mesmo espaço.

2.3.4 SOMA DIRECTA EXTERNA

Consideremos espaços vectoriais $V_1, \dots, V_n$. A soma directa externa dos V_i, onde $i = 1, \dots, n$ denota-se por $V_1 \boxplus \cdots \boxplus V_n$ e define-se:

$$V_1 \boxplus \cdots \boxplus V_n = \{(\mathrm{v}_1, \dots, \mathrm{v}_n) \mid \mathrm{v}_1 \in V_1, \dots, \mathrm{v}_n \in V_n\}.$$

e as operações definidas por:

1. $(\mathrm{v}_1, \dots, \mathrm{v}_n) + (\mathrm{u}_1, \dots, \mathrm{u}_n) := (\mathrm{v}_1 + \mathrm{u}_1, \dots, \mathrm{v}_n + \mathrm{u}_n)$;
2. $\alpha(\mathrm{v}_1, \dots, \mathrm{v}_n) := (\alpha\mathrm{v}_1, \dots, \alpha\mathrm{v}_n)$.

De facto não existe qualquer razão para nos restringirmos ao caso finito. Recorde-se que se $A = \{X_i \mid i \in I\}$ é uma família (eventualmente infinita) de conjuntos então, o produto cartesiano é definido da seguinte forma:

$$\prod A = \prod_{i \in I} X_i := \{f \in {}^I(\cup A) \mid (\forall i \in I) f(i) \in X_i\},$$

ou seja, o produto cartesiano dos membros de A consiste nas funções $f : I \to \cup_{i \in I} X_i$ que, para cada $i \in I$ satisfazem $f(i) \in X_i$.

A soma directa externa de uma família $\{V_i \mid i \in I\}$ de espaços vectoriais, que denotamos por $\boxplus_{i \in I} V_i$, pode ser agora ser definida como sendo o conjunto $\prod_{i \in I} V_i$ equipado com as operações:

1. $(f + g)(i) := f(i) +_{V_i} g(i)$;
2. $(\alpha f)(i) := \hat{\alpha}_{V_i}(f(i))$, para $\alpha \in \mathbb{K}$.