Álgebra Linear

M. Esmeralda Sousa Dias

Capítulo 1

Sistemas de Equações Lineares e Matrizes

A resolução de sistemas de equações lineares é uma tarefa comum a diversas áreas de aplicação da matemática, sendo frequente a ocorrência de sistemas lineares com um grande número de equações e de incógnitas em física, química, astronomia, engenharia, economia, etc. Do ponto de vista teórico, a resolução de um sistema de equações lineares não é difícil. Em particular, sistemas com um número reduzido de equações podem ser resolvidos sem recurso a um qualquer método específico. Porém, quando o número de equações é significativo, é natural utilizar-se um procedimento sistemático que possa ser facilmente programável.

Na primeira década do século XIX, Carl Friedrich Gauss[1] apresentou um método geral para a resolução de sistemas de equações lineares designado modernamente por *método de eliminação de Gauss*. Trata-se de um algoritmo bastante simples, o qual permanece como um dos marcos fundamentais da álgebra linear, dada a sua relevância quer do ponto de vista teórico quer do ponto de vista computacional. Neste capítulo, esse método é utilizado não apenas para resolver sistemas lineares, mas também como suporte à introdução de conceitos fundamentais, servindo de motivação para o estudo de matrizes e do espaço $\mathbb{R}^n$.

Neste capítulo são definidas as operações básicas com matrizes à custa de operações com vectores coluna. As operações de adição de vectores e de multiplicação de um vector por um número real são introduzidas como generalizações das operações bem conhecidas com vectores do plano e do espaço. Em $\mathbb{R}^n$ trata-

[1]Johann Carl Friedrich Gauss (ou Gauß), (1777-1855), matemático alemão com inúmeras contribuições em diversas áreas da matemática e física.

se com algum detalhe o conceito de combinação linear de vectores e conjunto gerador, dando especial ênfase aos aspectos geométricos no caso de vectores do plano e do espaço. Estes conceitos são fundamentais, não só na introdução de certas operações com matrizes mas também em estudos subsequentes como, por exemplo, no estudo de espaços lineares (Capítulo 3).

Na Secção 1.4 estudam-se matrizes invertíveis e descreve-se um método para determinar a inversa de uma matriz quadrada, conhecido pela designação de método de eliminação de Gauss-Jordan[2]. Estabelecem-se ainda relações entre matrizes invertíveis e a sua característica, bem como entre matrizes invertíveis e soluções de sistemas lineares.

Finalmente, tendo em vista utilizações posteriores, abordamos outros tópicos, como seja a factorização LU de uma matriz, bem como alguns aspectos relacionados com matrizes particionadas em blocos.

1.1 Sistemas de equações lineares

No contexto deste livro, quando nos referimos a números ou a *escalares* estamos a subentender tratar-se de números reais ou complexos. Comecemos por definir o que se entende por equação linear e por solução.

Definição 1.1. Uma *equação linear* nas *variáveis* (ou incógnitas) $x_1, x_2, \dots, x_n$ é uma equação do tipo

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = d, \tag{1.1}$$

onde $a_1, \dots, a_n$ e d são constantes (reais ou complexas).

Os números $a_1, \dots, a_n$ são designados por *coeficientes* da equação (1.1), e d é designado por *segundo membro* ou *termo independente* dessa equação.

Exemplo 1.1.

$$\begin{array}{rl}
2x - 3y + 5z - 1 = 0 & \text{é uma equação linear nas variáveis } x, y, z. \\
2x - y^2 + 5z - 1 = 0 & \text{não é uma equação linear.} \\
\sqrt{2}x + y = 1 & \text{é uma equação linear nas variáveis } x, y. \\
\sqrt{2x} + y = 1 & \text{não é uma equação linear.} \\
2x - \operatorname{sen} y = 0 & \text{não é uma equação linear.}
\end{array}$$

[2] Wilhelm Jordan, (1842-1899), geodesista alemão.

♦

Uma solução da equação (1.1) é uma sequência de n números $s_1, s_2, \ldots, s_n$ que satisfaz a equação quando se substitui $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$. Nomeadamente, quando se verifica a igualdade

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = d.$$

Esta solução é habitualmente designada pelo n-uplo $(s_1, s_2, \ldots, s_n)$.

Um sistema de equações lineares

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = d_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = d_2 \\ \qquad\qquad\qquad\qquad\qquad \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n = d_p \end{cases} \tag{1.2}$$

tem solução $(s_1, s_2, \ldots, s_n)$ se este n-uplo é solução de todas as equações do sistema.

Definição 1.2. O n-uplo $(s_1, s_2, \ldots, s_n)$ é solução da equação linear (1.1) se

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = d.$$

O n-uplo $(s_1, s_2, \ldots, s_n)$ é solução do sistema de equações lineares (1.2) se é solução de todas as equações do sistema.

Ao conjunto de todas as soluções de um sistema chamamos *solução geral*.

Um sistema de equações lineares também é designado por *sistema linear*.

Exemplo 1.2. O par $(x_1, x_2) = (1, 2)$ é solução do sistema

$$\begin{cases} x_1 - x_2 = -1 \\ -x_1 + 2x_2 = 3. \end{cases} \tag{1.3}$$

Porém, o par $(2, 1)$ não é uma solução (verifique).

♦

Relembremos que uma recta do plano é definida por uma equação linear em duas variáveis (isto é, $ax + by = d$). Por exemplo, as equações do sistema linear (1.3) definem duas rectas que designamos respectivamente por l_1 e l_2. Uma

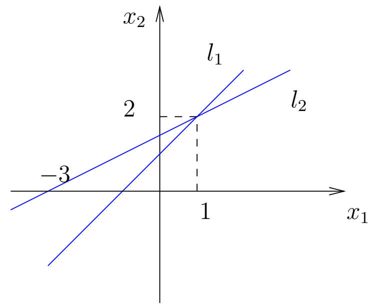


Figura 1.1: O sistema (1.3) tem solução única $(1, 2)$.

solução deste sistema é um ponto comum a l_1 e l_2. Na Figura 1.1 encontram-se representadas estas rectas bem como a (única) solução do sistema.

Geometricamente a solução geral de um sistema de duas equações a duas incógnitas é o conjunto dos pontos de intersecção de duas rectas do plano. Assim, a solução geral de sistemas de duas equações a duas incógnitas é de um dos seguintes três tipos: (i) solução única (as duas rectas intersectam-se num único ponto) e portanto a solução geral é um conjunto com um único elemento; (ii) não existe solução (as rectas são paralelas), pelo que a solução geral é o conjunto vazio; (iii) existem infinitas soluções (as rectas são coincidentes) pelo que a solução geral tem infinitos elementos.

No caso de um sistema linear com um número qualquer de equações e de incógnitas, distinguimos os casos que se seguem.

Um sistema de equações lineares diz-se:

1. *Possível e determinado* se tem uma única solução;
2. *Impossível* se não tem soluções;
3. *Indeterminado* se tem infinitas soluções.

Nota 1. *Como se mostra na Proposição 1.4 (pág. 43), um sistema linear com mais do que uma solução possui uma infinidade de soluções. Conclui-se portanto que os três casos acima esgotam todas as possibilidades.*

1.2 Método de eliminação de Gauss

Para resolver sistemas com várias equações e incógnitas é conveniente usar um algoritmo, ou seja, um procedimento sistemático. O método de eliminação de Gauss é um algoritmo de eleição e baseia-se na substituição de um dado sistema por outro (de mais fácil resolução) que possui exactamente as mesmas soluções. Este novo sistema é obtido aplicando três tipos de operações que eliminam de forma sistemática as variáveis do sistema. As operações permitidas, e que passamos a designar por *operações elementares*, são:

Operações elementares sobre as equações de um sistema

1. Trocar a ordem das equações.
2. Multiplicar uma equação por uma constante não nula.
3. Substituir uma equação pela sua soma com um múltiplo de outra equação.

Note-se que a aplicação de uma operação elementar a um sistema produz um sistema com as mesmas soluções, isto é, um sistema equivalente.

A escolha das operações elementares a aplicar num sistema tem como objectivo eliminar incógnitas das equações por forma a ser possível determinar a solução geral do sistema por substituição regressiva (isto é, determinar as soluções da última equação, substituir essas soluções na penúltima equação e determinar as suas soluções, continuar este processo até à primeira equação). No exemplo seguinte ilustramos quais os critérios habitualmente utilizados na escolha das operações a aplicar.

Exemplo 1.3. Apliquemos operações elementares ao sistema

$$\begin{cases} \phantom{-x_1 + {}} x_2 - 3x_3 = 1 \\ x_1 - 5x_2 + x_3 = 4 \\ -x_1 + 4x_2 - 2x_3 = 1, \end{cases}$$

por forma a obter um sistema com as seguintes características: a terceira equação é uma equação somente na variável x_3, a segunda uma equação nas variáveis x_2 e x_3, e a primeira equação uma equação nas variáveis x_1, x_2 e x_3.

Como a primeira equação não possui a incógnita x_1, vamos trocar a segunda equação com a primeira equação

$$\begin{cases} x_2 - 3x_3 = 1 \\ x_1 - 5x_2 + x_3 = 4 \\ -x_1 + 4x_2 - 2x_3 = 1 \end{cases} \xrightarrow{\text{trocar equação 1 com a 2}} \begin{cases} x_1 - 5x_2 + x_3 = 4 \\ x_2 - 3x_3 = 1 \\ -x_1 + 4x_2 - 2x_3 = 1 \end{cases}$$

A segunda equação já é uma equação nas variáveis x_2 e x_3, como pretendíamos. Teremos agora de eliminar x_1 e x_2 da terceira equação. Para eliminar x_1 da terceira equação, vamos substituir a terceira equação pela sua soma com a primeira

$$\begin{cases} x_1 - 5x_2 + x_3 = 4 \\ x_2 - 3x_3 = 1 \\ -x_1 + 4x_2 - 2x_3 = 1 \end{cases} \xrightarrow{\text{substituir equação 3 pela sua soma com a equação 1}} \begin{cases} x_1 - 5x_2 + x_3 = 4 \\ x_2 - 3x_3 = 1 \\ -x_2 - x_3 = 5 \end{cases}$$

Para eliminar x_2 da terceira equação sem introduzir x_1, temos que usar a segunda equação. Nomeadamente, substituindo a terceira equação pela sua soma com a segunda obtemos o seguinte sistema

$$\begin{cases} x_1 - 5x_2 + x_3 = 4 \\ x_2 - 3x_3 = 1 \\ -x_2 - x_3 = 5 \end{cases} \xrightarrow{\text{substituir equação 3 pela sua soma com a equação 2}} \begin{cases} x_1 - 5x_2 + x_3 = 4 \\ x_2 - 3x_3 = 1 \\ -4x_3 = 6 \end{cases}$$

Podemos agora determinar a solução do sistema por substituição regressiva. Resolvendo a última equação temos $x_3 = \frac{-3}{2}$. Substituindo este valor na segunda equação obtemos $x_2 = \frac{-7}{2}$ e, finalmente, substituindo x_3 e x_2 na primeira equação obtemos $x_1 = 4 + 5 \times \left(\frac{-7}{2}\right) + \frac{3}{2} = -12$.

O sistema tem solução única:

$$(x_1, x_2, x_3) = \left(-12, \frac{-7}{2}, \frac{-3}{2}\right).$$

♦

É claro do exemplo anterior que a informação essencial de um sistema se encontra nos seus coeficientes e nos termos do segundo membro de cada equação. Esta informação pode ser descrita de forma sucinta construindo um quadro de números dispostos em linhas e colunas onde os coeficientes das variáveis de uma dada equação se situam numa linha, e os coeficientes de uma dada variável do sistema se encontram numa coluna. Este quadro, que se designa por *matriz dos coeficientes* do sistema, permite identificar os coeficientes de cada incógnita em qualquer das equações do sistema. Por exemplo, para o sistema do último exemplo temos

$$\begin{cases} x_2 - 3x_3 = 1 \\ x_1 - 5x_2 + x_3 = 4 \\ -x_1 + 4x_2 - 2x_3 = 1 \end{cases} \rightsquigarrow \begin{bmatrix} 0 & 1 & -3 \\ 1 & -5 & 1 \\ -1 & 4 & -2 \end{bmatrix}$$

(**Matriz dos coeficientes do sistema**).

Na matriz dos coeficientes deste sistema, os coeficientes das variáveis da 1ª equação situam-se na 1ª linha, da 2ª equação na segunda linha e da 3ª equação na terceira linha. Os coeficientes da variável x_1 situam-se na 1ª coluna, da variável x_2 na segunda coluna, e da variável x_3 na terceira coluna.

Como as operações elementares sobre as equações de um sistema também afectam o segundo membro de cada equação, é conveniente aumentar em uma coluna (correspondente ao segundo membro das equações) a matriz dos coeficientes do sistema. A esta matriz chamamos *matriz aumentada* do sistema. Por exemplo,

$$\begin{cases} x_2 - 3x_3 = 1 \\ x_1 - 5x_2 + x_3 = 4 \\ -x_1 + 4x_2 - 2x_3 = 1 \end{cases} \rightsquigarrow \left[\begin{array}{ccc|c} 0 & 1 & -3 & 1 \\ 1 & -5 & 1 & 4 \\ -1 & 4 & -2 & 1 \end{array}\right]$$

(**Matriz aumentada do sistema**).

Note-se que a matriz aumentada de um sistema determina completamente o sistema no sentido em que dada uma tal matriz podemos escrever um sistema correspondente.

Passamos a designar por *matriz* qualquer quadro de números dispostos por linhas e colunas, e por *entradas* os números que aparecem neste quadro.

Exemplo 1.4. O sistema linear nas incógnitas x, y, z e w correspondente à matriz aumentada

$$\left[\begin{array}{cccc|c} 1 & 2 & -5 & 1 & 3 \\ 3 & -1 & 1 & -2 & 5 \\ 4 & -2 & 7 & -3 & 6 \end{array}\right] \quad \text{é} \quad \begin{cases} x + 2y - 5z + w = 3 \\ 3x - y + z - 2w = 5 \\ 4x - 2y + 7z - 3w = 6. \end{cases}$$

♦

As operações elementares sobre as equações de um sistema traduzem-se imediatamente em operações sobre as linhas da matriz aumentada do sistema.

Operações elementares sobre as linhas de uma matriz
1. Trocar linhas. 2. Multiplicar uma linha por uma constante não nula. 3. Substituir uma linha pela sua soma com outra linha multiplicada por uma constante.

Exemplo 1.5. Apliquemos operações elementares sobre as linhas da matriz aumentada do sistema do Exemplo 1.3, a fim de reproduzirmos as operações ali efectuadas.

$$\begin{cases} x_2 - 3x_3 = 1 \\ x_1 - 5x_2 + x_3 = 4 \\ -x_1 + 4x_2 - 2x_3 = 1 \end{cases} \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 0 & 1 & -3 & 1 \\ 1 & -5 & 1 & 4 \\ -1 & 4 & -2 & 1 \end{array}\right].$$

Abreviaremos as operações elementares usando a seguinte notação: L_i designa a linha i da matriz; $L_i \leftrightarrow L_j$ designa a troca da linha i com a linha j; αL_i designa a substituição da linha i por essa linha multiplicada por $\alpha \neq 0$; $L_i + \alpha L_j$ designa a substituição da linha i pela sua soma com a linha j multiplicada pelo escalar α. Note-se que nesta notação a linha que é substituída aparece em primeiro lugar na operação indicada, pelo que o significado de $L_i + \alpha L_j$ é distinto de $\alpha L_j + L_i$.

$$\left[\begin{array}{ccc|c} 0 & 1 & -3 & 1 \\ 1 & -5 & 1 & 4 \\ -1 & 4 & -2 & 1 \end{array}\right] \xrightarrow{L_1 \leftrightarrow L_2} \left[\begin{array}{ccc|c} 1 & -5 & 1 & 4 \\ 0 & 1 & -3 & 1 \\ -1 & 4 & -2 & 1 \end{array}\right] \xrightarrow{L_3 + L_1} \left[\begin{array}{ccc|c} 1 & -5 & 1 & 4 \\ 0 & 1 & -3 & 1 \\ 0 & -1 & -1 & 5 \end{array}\right]$$

$$\xrightarrow{L_3 + L_2} \left[\begin{array}{ccc|c} 1 & -5 & 1 & 4 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & -4 & 6 \end{array}\right]$$

Escrevendo o sistema correspondente à matriz (aumentada) resultante, obtemos exactamente o sistema final do Exemplo 1.3. Como operações elementares não alteram as soluções de um sistema, o sistema obtido possui o mesmo conjunto de soluções que o sistema de partida. ♦

Notação:

$L_i \leftrightarrow L_j$	trocar a linha i com a linha j.
αL_i	multiplicar a linha i por $\alpha \neq 0$.
$L_i + \alpha L_j$	substituir a linha i pela soma da linha i com a linha j multiplicada pelo escalar α.

Nota 2. *Para indicar a passagem de uma matriz a outra obtida por meio de uma certa operação elementar sobre as linhas, usou-se uma seta e não o sinal de igualdade. Como veremos adiante, a igualdade de matrizes tem outro significado.*

A questão que naturalmente se coloca é a de saber quando devemos parar de aplicar operações elementares, ou seja, qual deve ser a forma da matriz (aumentada) final. A definição que se segue responde a essa questão.

Definição 1.3. Matriz em escada por linhas

Uma matriz diz-se que está em *escada por linhas*, ou simplesmente em *escada*, se satisfaz as seguintes propriedades:

a) Não possui linhas nulas seguidas de linhas não nulas.

b) A primeira entrada não nula de cada linha encontra-se numa coluna à direita da coluna a que pertence a primeira entrada não nula da linha imediatamente anterior.

Numa matriz em escada por linhas chamamos *pivô* à primeira entrada não nula de cada linha.

Uma matriz diz-se em escada por linhas na *forma reduzida* se está em escada por linhas e além disso satisfaz as condições:

- todos os pivôs são iguais a 1;
- cada pivô é a única entrada não nula da coluna respectiva.

Nota 3. *A definição de matriz em escada por linhas implica que uma tal matriz tenha todas as entradas nulas por baixo de cada pivô.*

Eis um exemplo esquemático de uma matriz em escada por linhas:

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{bmatrix},$$

onde $\blacksquare$ indica um número diferente de zero, 0 indica um zero e $*$ pode ser qualquer número. A correspondente matriz em escada por linhas na forma reduzida é

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

Exemplo 1.6. Ilustremos a noção de matriz em escada com alguns exemplos.

- $\begin{bmatrix} 1 & 5 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ não está em escada por linhas, visto que tem uma linha nula seguida de uma linha não nula.
- $\begin{bmatrix} 1 & 5 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ não está em escada por linhas, uma vez que a primeira entrada não nula da terceira linha se encontra numa coluna à esquerda da coluna a que pertence a primeira entrada não nula da segunda linha.
- $\begin{bmatrix} 2 & 0 & 4 & 6 & 4 \\ 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ está em escada por linhas.
- $\begin{bmatrix} 2 & 5 & 4 & 6 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ não está em escada por linhas, já que por baixo da primeira entrada não nula da segunda linha há entradas não nulas.

♦

Aplicando operações elementares a uma matriz não nula podemos sempre obter uma matriz em escada na forma reduzida. Se uma matriz está em escada por linhas, a sua forma reduzida obtém-se dividindo cada linha pelo pivô dessa linha, e anulando as entradas acima de cada pivô usando operações elementares. Por exemplo,

$$\underbrace{\begin{bmatrix} 2 & 4 & 6 & 0 & 2 \\ 0 & 1 & 1 & 2 & -2 \\ 0 & 0 & 0 & 2 & 4 \end{bmatrix}}_{\text{escada por linhas}} \underset{L_1/2}{\xrightarrow{L_3/2}} \begin{bmatrix} 1 & 2 & 3 & 0 & 1 \\ 0 & 1 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{L_2-2L_3} \begin{bmatrix} 1 & 2 & 3 & 0 & 1 \\ 0 & 1 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \longrightarrow$$

$$\xrightarrow{L_1-2L_2} \underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 & 13 \\ 0 & 1 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}}_{\text{escada por linhas na forma reduzida}}.$$

Verifica-se facilmente que não é única a matriz em escada resultante da aplicação de operações elementares a uma dada matriz. Por exemplo, se alterarmos a sequência das operações elementares efectuadas na matriz do Exemplo 1.5 obtemos uma matriz em escada distinta:

$$\left[\begin{array}{ccc|c} 0 & 1 & -3 & 1 \\ 1 & -5 & 1 & 4 \\ -1 & 4 & -2 & 1 \end{array}\right] \xrightarrow{L_1 \leftrightarrow L_3} \left[\begin{array}{ccc|c} -1 & 4 & -2 & 1 \\ 1 & -5 & 1 & 4 \\ 0 & 1 & -3 & 1 \end{array}\right] \xrightarrow{L_2+L_1} \left[\begin{array}{ccc|c} -1 & 4 & -2 & 1 \\ 0 & -1 & -1 & 5 \\ 0 & 1 & -3 & 1 \end{array}\right]$$

$$\xrightarrow{L_3+L_2} \left[\begin{array}{ccc|c} -1 & 4 & -2 & 1 \\ 0 & -1 & -1 & 5 \\ 0 & 0 & -4 & 6 \end{array}\right].$$

Para a matriz do exemplo referido obtivemos as seguintes matrizes em escada (distintas)

$$\left[\begin{array}{ccc|c} 1 & -5 & 1 & 4 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & -4 & 6 \end{array}\right] \quad \text{e} \quad \left[\begin{array}{ccc|c} -1 & 4 & -2 & 1 \\ 0 & -1 & -1 & 5 \\ 0 & 0 & -4 & 6 \end{array}\right].$$

No entanto, a matriz em escada na forma reduzida correspondente a estas duas matrizes é a matriz

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 12 \\ 0 & 1 & 0 & -7/2 \\ 0 & 0 & 1 & -3/2 \end{array}\right].$$

No Teorema 1.2 (pág. 43), mostra-se que dada uma matriz a respectiva forma reduzida em escada por linhas *é única*. A unicidade da matriz em escada na forma reduzida é um resultado com consequências importantes. Nomeadamente, uma vez que a forma reduzida de uma matriz em escada R tem o mesmo número de pivôs que R, a unicidade da matriz em escada na forma reduzida permite concluir que o número de pivôs de uma matriz em escada obtida de uma matriz A é independente do número e da sequência de operações elementares utilizadas. Definimos a *característica* da matriz A como sendo o número de pivôs de uma qualquer matriz em escada obtida de A por meio de operações elementares sobre as suas linhas.

Definição 1.4. Característica

Chama-se *característica* de uma matriz A ao número de linhas não nulas de uma qualquer matriz em escada obtida de A por meio de operações elementares sobre as suas linhas. Designamos a característica de A por:

$$\operatorname{car} A.$$

A terminologia anglo-saxónica para a característica de uma matriz é “rank”.

Exemplo 1.7. **(a)** Consideremos a matriz aumentada $\left[A \mid b\right]$ do Exemplo 1.5, e a respectiva matriz em escada obtida nesse exemplo:

$$\left[A \mid b\right] = \left[\begin{array}{ccc|c} 0 & 1 & -3 & 1 \\ 1 & -5 & 1 & 4 \\ -1 & 4 & -2 & 1 \end{array}\right] \xrightarrow{\text{Operações elementares}} \left[\begin{array}{ccc|c} 1 & -5 & 1 & 4 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & -4 & 6 \end{array}\right].$$

Como a matriz em escada por linhas possui 3 pivôs, têm-se que $\operatorname{car}\left[A \mid b\right] = 3$. Podemos também concluir que, neste caso, a característica da matriz aumentada $\left[A \mid b\right]$ é igual à característica da matriz A dos coeficientes do sistema, isto é

$$\operatorname{car}\left[A \mid b\right] = 3 \quad \text{e} \quad \operatorname{car} A = 3.$$

Como observámos no Exemplo 1.3, o sistema correspondente é possível e determinado. De facto, como veremos já no próximo exemplo, é necessário que $\operatorname{car}\left[A \mid b\right] = \operatorname{car} A$ para que o sistema seja possível.

(b) Consideremos agora a matriz aumentada $\left[A \mid b\right] = \left[\begin{array}{cc|c} 1 & -2 & 5 \\ -2 & 4 & 4 \end{array}\right]$. Calculemos a sua característica.

$$\left[A \mid b\right] = \left[\begin{array}{cc|c} 1 & -2 & 5 \\ -2 & 4 & 4 \end{array}\right] \xrightarrow{L_2+2L_1} \left[\begin{array}{cc|c} 1 & -2 & 5 \\ 0 & 0 & 14 \end{array}\right].$$

Donde se conclui que

$$\operatorname{car}\left[A \mid b\right] = 2 \quad \text{e} \quad \operatorname{car} A = 1.$$

Neste caso, a característica da matriz aumentada $[A|b]$ é diferente da característica de A. Como o sistema inicial tem as mesmas soluções que o sistema correspondente à matriz em escada obtida, ou seja, que o sistema

$$\begin{cases} x - 2y = \;\; 5 \\ \qquad\;\; 0 = 14, \end{cases}$$

o sistema original é impossível. ♦

Os exemplos (a) e (b) fazem prever que a existência ou não de solução para um determinado sistema linear está dependente do facto de ser ou não verdadeira a igualdade $\operatorname{car} A = \operatorname{car}[A|b]$. Ver adiante a Proposição 1.1.

Algoritmo para redução de uma matriz a uma matriz em escada

Tendo em vista a construção de um algoritmo facilmente programável, é conveniente efectuar as operações elementares segundo uma determinada ordem, a fim de obter uma matriz em escada. Tal algoritmo recebe usualmente a designação de *método de eliminação de Gauss*. Ilustramos esse algoritmo através do cálculo de uma matriz em escada por linhas para a matriz

$$\begin{bmatrix} 0 & 2 & -6 & 4 & 1 & 2 \\ 3 & 9 & 5 & 1 & 0 & 1 \\ 3 & 7 & 2 & -1 & 2 & 1 \end{bmatrix}.$$

Passo 1:

Selecionar a primeira coluna não nula da matriz. Se o primeiro elemento dessa coluna for não nulo, esse elemento será designado por pivô[a]. Caso contrário, efectua-se troca de linhas colocando um elemento não nulo na primeira linha dessa coluna.

[a] Embora tenhamos definido pivô apenas para matrizes em escada, neste algoritmo usamos a mesma designação de pivô para os elementos que vão ser precisamente os pivôs da matriz em escada final.

$$\begin{bmatrix} 0 & 2 & -6 & 4 & 1 & 2 \\ 3 & 9 & 5 & 1 & 0 & 1 \\ 3 & 7 & 2 & -1 & 2 & 1 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 3 & 9 & 5 & 1 & 0 & 1 \\ 0 & 2 & -6 & 4 & 1 & 2 \\ 3 & 7 & 2 & -1 & 2 & 1 \end{bmatrix}$$

> **Passo 2:**
> Usar operações elementares para obter zeros abaixo do pivô selecionado no passo anterior.

$$\begin{bmatrix} 3 & 9 & 5 & 1 & 0 & 1 \\ 0 & 2 & -6 & 4 & 1 & 2 \\ 3 & 7 & 2 & -1 & 2 & 1 \end{bmatrix} \xrightarrow{L_3 - L_1} \begin{bmatrix} 3 & 9 & 5 & 1 & 0 & 1 \\ 0 & 2 & -6 & 4 & 1 & 2 \\ 0 & -2 & -3 & -2 & 2 & 0 \end{bmatrix}$$

> **Passo 3:**
> Repetir os passos 1 e 2 para a matriz que se obtém suprimindo a linha e coluna que contêm o pivô.

$$\begin{bmatrix} 3 & 9 & 5 & 1 & 0 & 1 \\ 0 & 2 & -6 & 4 & 1 & 2 \\ 0 & -2 & -3 & -2 & 2 & 0 \end{bmatrix} \xrightarrow{L_3 + L_2} \begin{bmatrix} 3 & 9 & 5 & 1 & 0 & 1 \\ 0 & 2 & -6 & 4 & 1 & 2 \\ 0 & 0 & -9 & 2 & 3 & 2 \end{bmatrix}$$

pivôs

A matriz final é uma matriz em escada por linhas.

A aplicação do método de eliminação de Gauss para a resolução de sistemas é a seguir resumido.

> **Método de Eliminação de Gauss**
> A resolução de um sistema de equações lineares pelo método de eliminação de Gauss consiste em:
>
> 1. Escrever a matriz aumentada do sistema.
> 2. Por meio de operações elementares sobre as linhas da matriz aumentada, obter uma matriz em escada de acordo com o algoritmo anteriormente descrito.
> 3. Escrever o sistema correspondente à matriz (aumentada) em escada. Resolver este sistema por substituição (regressiva) começando pela última equação.

Ilustremos o método de eliminação de Gauss através de alguns exemplos.

Exemplo 1.8. Usemos o método de eliminação de Gauss para resolver o sistema seguinte.

$$\begin{cases} x_1 - 3x_2 - x_3 \phantom{{}- 4x_4} = -2 \\ x_2 + x_3 - 4x_4 = 1 \\ \phantom{x_1 - 3x_2 + {}} x_3 + 9x_4 = 4. \end{cases}$$

A matriz aumentada do sistema

$$\left[\begin{array}{cccc|c} 1 & -3 & -1 & 0 & -2 \\ 0 & 1 & 1 & -4 & 1 \\ 0 & 0 & 1 & 9 & 4 \end{array}\right] \tag{1.4}$$

já está em escada por linhas. Logo, resolvendo a última equação em ordem a x_3, vem $x_3 = 4 - 9x_4$. Resolvendo a segunda equação em ordem a x_2 e substituindo x_3, obtém-se $x_2 = 1 - x_3 + 4x_4 = -3 + 13x_4$. Substituindo x_2 e x_3 na primeira equação, resulta

$$x_1 = -2 + 3\,(-3 + 13x_4) + (4 - 9x_4) = -7 + 30x_4.$$

O sistema possui infinitas soluções e a sua solução geral é o conjunto

$$\{(x_1, x_2, x_3, x_4) :\ x_1 = -7 + 30x_4,\ x_2 = -3 + 13x_4,\ x_3 = 4 - 9x_4 \text{ e } x_4 \in \mathbb{R}\} =$$

$$= \{(-7 + 30x_4, -3 + 13x_4, 4 - 9x_4, x_4) : x_4 \in \mathbb{R}\}.$$

Uma vez que podemos atribuir livremente valores a x_4, esta variável parametriza o conjunto das soluções. Ou seja, a cada valor do parâmetro x_4 corresponde uma solução do sistema. ♦

Exemplo 1.9. Usemos o método de eliminação de Gauss para resolver o sistema

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

Reduza-se a matriz aumentada do sistema a uma matriz em escada.

$$\left[\begin{array}{ccc|c} 1 & -1 & 0 & 2 \\ 1 & 0 & 1 & 1 \\ -2 & 2 & 0 & 2 \end{array}\right] \xrightarrow{L_2 - L_1} \left[\begin{array}{ccc|c} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ -2 & 2 & 0 & 2 \end{array}\right] \xrightarrow{L_3 + 2L_1} \left[\begin{array}{ccc|c} 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 6 \end{array}\right].$$

O sistema correspondente a esta matriz em escada é

$$\begin{cases} x - y = 2 \\ y + z = -1 \\ 0 = 6. \end{cases}$$

Da última equação concluímos que o sistema é impossível. Saliente-se que a característica da matriz aumentada do sistema é 3 enquanto a característica da matriz dos coeficientes do sistema é 2. ♦

1.2.1 A solução geral de um sistema

Como vimos, um sistema linear ou é possível ou é impossível. No caso de ser possível ou possui solução única e diz-se *possível e determinado*, ou tem um número infinito de soluções sendo por isso *indeterminado* (ver Proposição 1.4 na página 43). Uma condição necessária e suficiente[3] para um sistema ser possível é que a matriz aumentada do sistema e a matriz dos coeficientes do sistema tenham a mesma característica. Enunciamos este resultado na proposição a seguir.

Proposição 1.1. Seja $[A|b]$ a matriz aumentada de um sistema de equações lineares. O sistema é possível se e só se

$$\operatorname{car}[A|b] = \operatorname{car} A.$$

Demonstração. [4] Comecemos por mostrar que se $\operatorname{car}[A|b] > \operatorname{car}(A)$ o sistema é impossível. Se $\operatorname{car}[A|b] > \operatorname{car}(A)$, o método de eliminação de Gauss aplicado a $[A|b]$ dá origem a uma matriz em escada que tem uma linha com todas as entradas nulas excepto a entrada mais à direita. O sistema correspondente à matriz em escada obtida possui uma equação do tipo $0 = k$ com $k \neq 0$, por conseguinte o respectivo sistema é impossível. Como o sistema original e o sistema correspondente à matriz em escada por linhas têm as mesmas soluções, o sistema original é impossível.

[3]Uma proposição P diz-se uma condição necessária e suficiente para a proposição Q se são satisfeitas as implicações: $P \Longrightarrow Q$ e $Q \Longrightarrow P$. Quando se verifica esta dupla implicação, ou seja, "se P, então Q" e "se Q, então P", resume-se este facto numa única proposição, a saber, "P se e só se Q". A expressão "P se e só se Q" (ou abreviadamente "P sse Q") pode também escrever-se na forma $P \Longleftrightarrow Q$, traduzindo que P é equivalente a Q.

[4]Assume-se implicitamente nesta demonstração que a característica de uma matriz é um número bem definido, resultado que segue do Teorema 1.2. Não há contudo qualquer erro lógico em demonstrar esta proposição antes desse teorema, uma vez que a prova do teorema é independente desta demonstração.

Para a implicação recíproca, atenda-se a que a característica da matriz aumentada é sempre maior ou igual do que a característica da matriz dos coeficientes. Como por hipótese o sistema é possível, a característica da matriz aumentada não pode ser maior que a característica da matriz dos coeficientes. Logo, se o sistema é possível, então $\mathrm{car}[A|b] = \mathrm{car}(A)$.

□

A solução geral de um sistema é o conjunto de todas as soluções do sistema. Assim, um sistema impossível (sem soluções) tem para solução geral o conjunto vazio, enquanto que um sistema possível e determinado tem para solução geral um conjunto com um único elemento. Quando um sistema é indeterminado, isto é com infinitas soluções, a descrição do conjunto das soluções é mais complexa.

No Exemplo 1.8 o sistema é indeterminado e optámos por resolver a última equação em ordem a x_3 (isto é, $x_3 = 4 - 9x_4$), obtendo x_3 em função de x_4. As soluções do sistema obtêm-se atribuindo um valor arbitrário a x_4 e substituindo esse valor nas igualdades obtidas para as restantes variáveis. Dizemos então que x_4 é uma variável *livre* e que as restantes são *dependentes*. No entanto, poderíamos ter optado por resolver a última equação em ordem a x_4, obtendo-se x_4 dependente de x_3, nomeadamente $x_4 = 4/9 - x_3/9$. Neste caso, o conjunto solução geral seria descrito em termos de x_3, ou seja, x_3 poderia ser qualquer número real, e as restantes variáveis dependeriam de x_3. Este exemplo mostra que há várias maneiras de descrever a solução geral, conforme as variáveis que escolhamos para descrever (ou *parametrizar*) os elementos do conjunto das soluções.

É habitual introduzir-se alguma nomenclatura que corresponde a uma escolha criteriosa das variáveis usadas na parametrização do conjunto das soluções de um sistema. Relembremos que na notação matricial, cada coluna da matriz dos coeficientes de um sistema está associada a uma variável. No Exemplo 1.8 a solução geral foi descrita em termos da variável x_4, nomeadamente a solução geral é o conjunto

$$\{(-7 + 30x_4, -3 + 13x_4, 4 - 9x_4, x_4) : x_4 \in \mathbb{R}\}\,.$$

Observando a matriz em escada em (1.4), verifica-se que a variável x_4 é a variável associada à coluna dessa matriz que não contém qualquer pivô e que x_1, x_2 e x_3 são variáveis associadas a colunas com um pivô.

Definição 1.5. Numa matriz em escada por linhas obtida por eliminação de Gauss da matriz dos coeficientes de um sistema possível, designamos por:

i) variáveis *livres* ou *independentes*, as variáveis correspondentes a colunas sem pivô;

ii) variáveis *dependentes*, as variáveis correspondentes a colunas com pivô.

Chamamos *grau de indeterminação* de um sistema (possível) ao número de variáveis livres.

Note-se que o grau de indeterminação de um sistema possível é igual ao número de variáveis do conjunto das soluções que podem tomar qualquer valor, ou seja, o número de *parâmetros* necessários para descrever a solução geral.

Como o número de colunas de uma matriz em escada é igual à soma do número de colunas sem pivô com número de colunas com pivô, da definição anterior e da definição de característica de uma matriz (Definição 1.4), seguem os resultados que resumimos na próxima proposição.

Proposição 1.2. Seja A a matriz dos coeficientes de um sistema possível com n incógnitas. Então,

1) $\mathrm{car}(A) =$ número de variáveis dependentes.

2) $n - \mathrm{car}(A)$ = número de variáveis livres = grau de indeterminação do sistema.

Demonstração. Os itens 1) e 2) são imediatos usando as definições de característica e de variáveis livres. □

Exemplo 1.10. No Exemplo 1.8 temos $\mathrm{car}[A|b] = \mathrm{car}[A]$ e o grau de indeterminação do sistema é 1 (a matriz A só tem uma coluna sem pivô, ou seja, o sistema correspondente tem uma variável livre).

No Exemplo 1.9, como $\mathrm{car}[A|b] > \mathrm{car}(A)$, o sistema é impossível.

No Exemplo 1.2 verifica-se que uma matriz (aumentada) em escada é

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

Logo, $\mathrm{car}[A|b] = \mathrm{car}[A] = 2$ e não existem variáveis livres, uma vez que o número de variáveis do sistema é igual ao número de pivôs. Por conseguinte, o sistema é possível e determinado. ⧫

1.3 Álgebra de matrizes

Vimos que o uso de matrizes na resolução de sistemas representa uma economia de notação considerável. No entanto, a utilização de matrizes é útil num contexto muito mais vasto. Em particular, a manipulação algébrica de matrizes permite demonstrar facilmente resultados importantes, e tem aplicações que vão muito além da resolução de sistemas. Nesta secção definimos as operações usuais com matrizes e estabelecemos certas identificações envolvendo matrizes e vectores.

Comecemos por precisar alguns conceitos já usados.

Definição 1.6. Uma *matriz* do tipo $p \times n$ é um quadro de números com p linhas e n colunas. Cada número na matriz é designado por *entrada*.

Se $p \neq n$ a matriz diz-se *rectangular* e se $p = n$ a matriz diz-se *quadrada* e de *ordem* n.

As matrizes são habitualmente designadas por letras maiúsculas, por exemplo A, e cada entrada pela correspondente letra minúscula indexada por dois índices que representam a posição da entrada na matriz. Assim, a_{ij} é o número que se encontra na linha i e na coluna j da matriz A. Por exemplo, a matriz

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

é uma matriz com duas linhas e três colunas e portanto é do tipo 2×3. A entrada situada na primeira linha e na segunda coluna é $a_{12} = 3$, enquanto que a entrada na segunda linha e na primeira coluna é $a_{21} = 1$. O primeiro índice em a_{ij} designa a linha e o segundo a coluna.

Uma matriz A do tipo $p \times n$ (com p linhas e n colunas), de entradas a_{ij} é representada por

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix} = \begin{bmatrix} a_{ij} \end{bmatrix}_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}}$$

Exemplo 1.11. A matriz $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{\substack{i=1,2,3 \\ j=1,2}}$ definida por

$$a_{ij} = \begin{cases} i-1 & \text{se } i > j \\ 0 & \text{se } i = j \\ j+2 & \text{se } i < j, \end{cases}$$

é a seguinte matriz 3×2:

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

♦

Numa matriz quadrada A do tipo $n \times n$, chamamos *diagonal principal* de A à diagonal constituída pelas entradas $a_{11}, a_{22}, \ldots, a_{nn}$.

Definição 1.7. A uma matriz com uma única coluna (isto é, do tipo $p \times 1$) chamamos *vector coluna*. As entradas de um vector coluna são designadas por *componentes* ou *coordenadas*.

Os vectores coluna são denotados por letras minúsculas a cheio, e por vezes com uma seta em cima.

Até ao início da próxima secção, onde se clarificam as relações entre vectores (geométricos) e vectores coluna, abreviamos vector coluna para vector.

Eis um exemplo de um vector,

$$\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \text{ ou } \quad \vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \quad \text{é um vector cuja segunda componente é } 3.$$

Soluções de equações lineares podem ser identificadas com vectores. Nomeadamente, uma solução $(s_1, \ldots, s_n)$ da equação linear $a_1x_1 + a_2x_2 + \cdots + a_nx_n = d$ pode escrever-se como o vector

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

que satisfaz a igualdade $a_1s_1 + a_2s_2 + \cdots + a_ns_n = d$.

Sempre que conveniente, também escrevemos $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ para designar o vector coluna

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

Nota 4. *a) Os parênteses habitualmente utilizados para delimitar as entradas de uma matriz são parênteses rectos ou curvos. As chavetas e os traços verticais não se usam para este efeito pois estão reservados para denotar os conceitos de conjunto e de determinante (a estudar no próximo capítulo).*

b) Por vezes usaremos a designação de tamanho quando nos referimos ao tipo da matriz.

c) Noutros livros o termo vector é também usado para designar matrizes com uma única linha mas aqui segue-se outra convenção.

Se encararmos uma matriz como um conjunto (ordenado) de vectores coluna é natural definirmos operações com matrizes baseando-as em operações com vectores coluna. Esta é a estratégia que seguiremos na definição de operações com matrizes, sendo por isso necessário começar por estudar operações com vectores colunas.

Como veremos de seguida, um vector coluna de duas ou três componentes reais representa um vector geométrico, respectivamente, no plano e no espaço. Um vector com $n > 3$ componentes (isto é, um vector de $\mathbb{R}^n$) generaliza os vectores geométricos. As operações com vectores coluna serão definidas como generalizações das correspondentes operações com vectores do plano ou do espaço.

1.3.1 Vectores de $\mathbb{R}^n$

Na representação de certas grandezas físicas tais como a velocidade, a aceleração ou a força, é necessário ter em conta não só a intensidade mas também uma direcção e sentido segundo os quais essa grandeza se manifesta. Este tipo de grandezas físicas são grandezas representadas por vectores.

Nesta secção introduzimos geometricamente a noção de vector, estabelecemos relações entre vectores, pontos, e vectores em sistemas de coordenadas. As noções estudadas no plano e no espaço são depois generalizadas para vectores de $\mathbb{R}^n$.

Vectores e geometria

Dois pontos (distintos) P e Q, definem um segmento de recta. Tomando P como ponto inicial e Q como ponto final, o segmento de recta fica orientado no sentido de P para Q. A notação $\overrightarrow{PQ}$ designa esse segmento orientado.

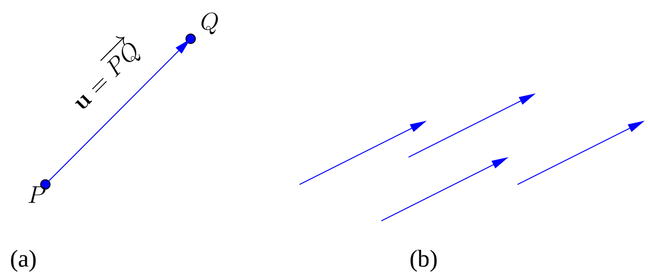


Figura 1.2: (a) O vector $\mathbf{u}$ é representado pelo segmento orientado $\vec{PQ}$. (b) Vectores equivalentes.

Segmentos orientados paralelos, com o mesmo sentido e o mesmo comprimento dizem-se *segmentos orientados equivalentes*. Isto é, dois segmentos orientados são equivalentes se coincidem quando se desloca um deles, paralelamente a si próprio, por forma a que os pontos iniciais dos dois segmentos orientados coincidam (ver Figura 1.2-(b)).

Frequentemente interessa considerar uma grandeza vectorial, independente do ponto do espaço em que a grandeza se manifesta, sendo apenas relevantes a sua intensidade e sentido. Uma maneira de formalizar este conceito consiste em definir um *vector*, ou *vector livre*, como o conjunto de *todos* os segmentos orientados equivalentes a um dado segmento orientado. O *vector nulo* é definido como o conjunto dos segmentos degenerados (segmentos com ponto inicial e final coincidentes).

Quando se representa um vector por um segmento orientado particular (trata-se de representar um conjunto por um dos seus elementos) diz-se que o vector se encontra aplicado no ponto inicial do segmento orientado que o representa (ver Figura 1.2-(a)).

A seguir definimos a soma de dois vectores bem como a multiplicação de um vector por um número real.

Definição 1.8. A soma $\mathbf{u}+\mathbf{v}$ de dois vectores $\mathbf{u}$ e $\mathbf{v}$ é o vector representado pelo segmento orientado obtido da seguinte forma: considerem-se representantes dos vectores $\mathbf{u}$ e $\mathbf{v}$ tais que o ponto inicial do segmento orientado que representa o vector $\mathbf{v}$ coincide com ponto final do segmento orientado que representa $\mathbf{u}$, o vector $\mathbf{u}+\mathbf{v}$ é o vector representado pelo segmento orientado cujo ponto inicial é o ponto inicial do representante de $\mathbf{u}$, e o ponto final é o ponto final do representante de $\mathbf{v}$ (ver Figura 1.3).

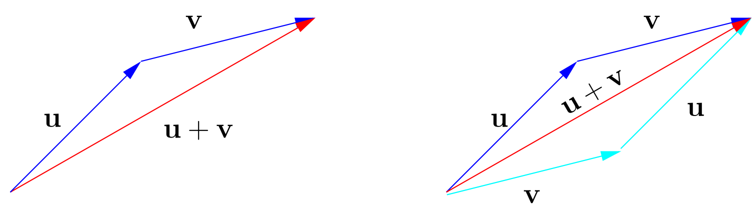


Figura 1.3: A soma $\mathbf{u}+\mathbf{v}$ e $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

Na Figura 1.3 verificamos geometricamente que a adição de vectores é comutativa, isto é,

$$\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}.$$

Definição 1.9. Se $\mathbf{u}$ é um vector não nulo e k um número real, então $k\mathbf{u}$ é o vector representado pelo segmento orientado cujo comprimento é $|k|$ vezes o comprimento do segmento orientado que representa $\mathbf{u}$, e o sentido é o desse segmento se $k>0$, e é o oposto se $k<0$. Definimos $k\mathbf{u}$ como sendo o vector nulo (ou zero) $\mathbf{0}$ se $k=0$ ou $\mathbf{u}$ é o vector nulo.

Na Figura 1.4 é ilustrada a multiplicação de um vector por um número real.

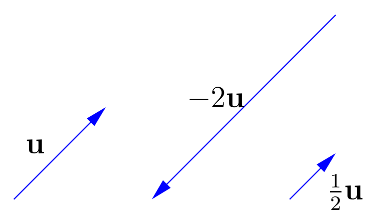


Figura 1.4: O produto de um número real por um vector.

Como consequência da definição de soma de vectores e das propriedades da multiplicação por um escalar, o simétrico $(-\mathbf{v})$ de um vector $\mathbf{v}$ é o vector que se obtém multiplicando $\mathbf{v}$ pelo escalar (-1).

Podemos definir diferença de dois vectores como sendo a seguinte soma:

$$\mathbf{u}-\mathbf{v}=\mathbf{u}+(-\mathbf{v}).$$

Como se ilustra na Figura 1.5, a construção geométrica de $\mathbf{u}-\mathbf{v}$ pode efectuar-se do seguinte modo: considerem-se representantes de $\mathbf{u}$ e de $\mathbf{v}$ com o mesmo ponto inicial, o vector $\mathbf{u}-\mathbf{v}$ é representado pelo segmento orientado cujo ponto inicial

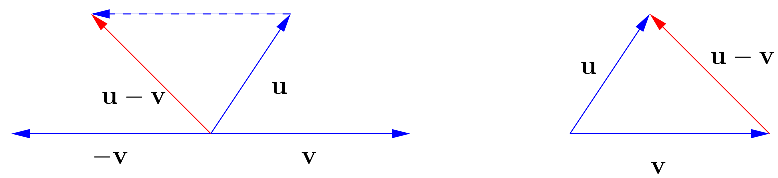


Figura 1.5: O vector diferença $\mathbf{u} - \mathbf{v}$.

é o ponto terminal do representante de $\mathbf{v}$ e cujo ponto final é o ponto final do segmento orientado que representa $\mathbf{u}$ (ver Figura 1.5).

Fixando um ponto O no espaço, um qualquer vector livre $\mathbf{u}$ pode sempre representar-se por um (e um só) segmento orientado cujo ponto inicial é o ponto O. Em particular, o vector nulo é representado pelo ponto O. Existe portanto uma correspondência biunívoca entre vectores (livres) e segmentos orientados com ponto inicial O. Além disso, uma vez fixado um ponto O do espaço, pode estabelecer-se uma correspondência biunívoca entre vectores (livres) e pontos do espaço. Esta correspondência é estabelecida do seguinte modo: ao vector $\mathbf{u}$ representado pelo segmento orientado $\overrightarrow{OP}$ corresponde o ponto P; reciprocamente, ao ponto P corresponde o conjunto de segmentos orientados equivalentes a $\overrightarrow{OP}$. Assim, fixado um ponto do espaço, temos a correspondência biunívoca

$$\text{Vectores} \longleftrightarrow \text{Pontos}.$$

Vectores em sistemas de coordenadas

Num plano podemos fixar um referencial, considerando para tal um ponto que se toma como origem e duas rectas não paralelas (eixos coordenados) que se intersectam na origem. Estas rectas são munidas duma unidade de comprimento e de um sentido.

Um ponto do plano fica completamente caracterizado se forem especificadas as suas coordenadas, isto é, se for dado o par ordenado (a, b) que indica que o ponto se projecta (paralelamente aos eixos coordenados) nos eixos coordenados em pontos que se situam, respectivamente, a a e b unidades (positivas ou negativas) da origem.

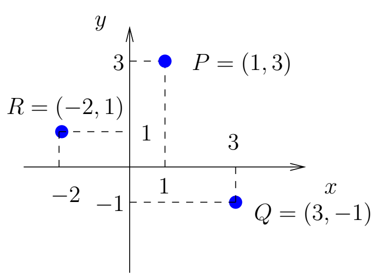
Cada vector livre do plano pode representar-se por um único segmento orientado cujo ponto inicial é a origem O do referencial fixado.

A correspondência biunívoca entre pontos do plano e vectores, permite identificar cada ponto do plano de coordenadas (a, b) com o segmento orientado cujo ponto inicial é a origem do referencial e cujo ponto final é (a, b). Ou seja, cada ponto pode identificar-se com um vector aplicado na origem do referencial. Este vector é designado por *vector de posição* do ponto (Figura 1.6).

Denotamos pontos por letras maiúsculas e os respectivos vectores de posição pela correspondente letra minúscula (a cheio). Isto é, para o ponto P o vector de posição respectivo é denotado por $\mathbf{p}$.

Existe portanto uma identificação natural entre o conjunto de vectores do plano e o conjunto dos vectores coluna de duas componentes reais, o qual se designa por $\mathbb{R}^2$. Isto é,

$$\mathbb{R}^2 = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 \in \mathbb{R} \text{ e } x_2 \in \mathbb{R} \right\} = \{(x_1, x_2) : x_1 \in \mathbb{R} \text{ e } x_2 \in \mathbb{R}\}.$$



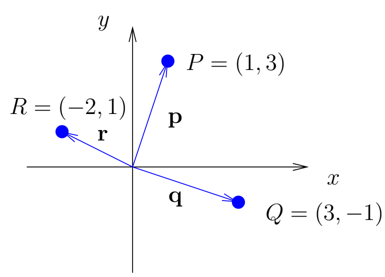


Figura 1.6: Representação de pontos e respectivos vectores de posição no referencial de eixos x e y.

Mediante a identificação descrita acima entre vectores do plano e vectores coluna, a soma de vectores do plano corresponde à soma componente a componente de vectores coluna. Ou seja,

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}, \qquad \mathbf{u} \in \mathbb{R}^2, \ \mathbf{v} \in \mathbb{R}^2.$$

Geometricamente, o vector soma $\mathbf{u} + \mathbf{v}$ é o vector de posição de um ponto que é o vértice de um paralelogramo cujos outros vértices são a origem e os pontos correspondentes aos vectores de posição $\mathbf{u}$ e $\mathbf{v}$ (ver Figura 1.7).

A multiplicação de um vector $\mathbf{u}$ de $\mathbb{R}^2$ por um número real k corresponde à multiplicação de cada componente do vector coluna por k. Isto é,

$$k\mathbf{u} = k \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} ku_1 \\ ku_2 \end{bmatrix}, \ k \in \mathbb{R}.$$

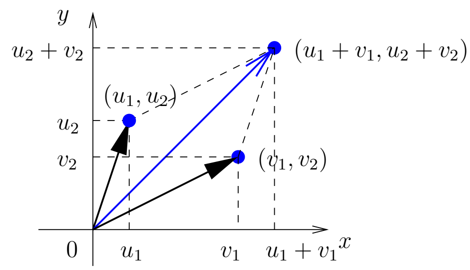


Figura 1.7: Adição de vectores de $\mathbb{R}^2$.

Na Figura 1.8 ilustra-se a representação da multiplicação de um vector de $\mathbb{R}^2$ por um real.

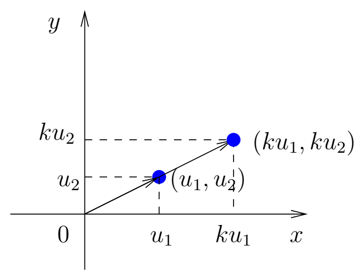


Figura 1.8: Multiplicação de vectores de $\mathbb{R}^2$ pelo escalar $k > 0$.

Vectores de $\mathbb{R}^3$

Pontos do espaço podem ser representados por ternos ordenados de números reais, desde que esteja fixado um referencial. Um referencial no espaço é obtido fixando um ponto do espaço que se toma como origem, e três rectas (eixos coordenados) distintas que se intersectam na origem. Em cada eixo coordenado escolhe-se ainda um sentido e uma unidade de comprimento.

Um terno ordenado (u_1, u_2, u_3) de números reais representa um ponto do espaço cujas coordenadas no referencial em causa valem, respectivamente, u_1 unidades do primeiro eixo, u_2 unidades do segundo eixo e u_3 unidades do terceiro. Um ponto P do espaço de coordenadas (u_1, u_2, u_3) pode ser identificado com o vector de posição $\mathbf{p} = (u_1, u_2, u_3)$, representado pelo segmento orientado cujo ponto inicial é a origem e cujo ponto final é o ponto P (ver Figura 1.9).

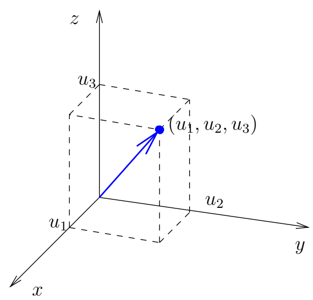


Figura 1.9: Representação de um vector de $\mathbb{R}^3$ no referencial de eixos x, y, z.

O conjunto de todos os vectores coluna de três componentes reais designa-se por $\mathbb{R}^3$, isto é,

$$\begin{aligned}\mathbb{R}^3 &= \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : \; x_1 \in \mathbb{R}, x_2 \in \mathbb{R} \text{ e } x_3 \in \mathbb{R} \right\} \\ &= \{(x_1, x_2, x_3) : \; x_1, x_2, x_3 \in \mathbb{R}\}\,.\end{aligned}$$

De forma inteiramente análoga ao caso de vectores de $\mathbb{R}^2$, não é difícil verificar que as definições de igualdade de vectores, adição de vectores e multiplicação de um vector por um número real, se traduzem da forma que se segue. Para quaisquer vectores $\mathbf{u} = (u_1, u_2, u_3)$ e $\mathbf{v} = (v_1, v_2, v_3)$ de $\mathbb{R}^3$ tem-se:

- $\mathbf{u} = \mathbf{v}$ se e só se $u_1 = v_1$, $u_2 = v_2$ e $u_3 = v_3$.
- $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$.
- $k\mathbf{u} = (ku_1, ku_2, ku_3)$ para $k \in \mathbb{R}$.

Por vezes é necessário determinar as coordenadas de um vector $\overrightarrow{PQ}$ conhecendo-se as coordenadas dos pontos P e de Q. Por exemplo, sabendo que as coordenadas dos pontos P e Q são respectivamente (p_1, p_2, p_3) e (q_1, q_2, q_3), pretende-se as coordenadas do vector $\overrightarrow{PQ}$ representado na Figura 1.10. Designando por O a origem do referencial, tem-se

$$\overrightarrow{OP} + \overrightarrow{PQ} = \overrightarrow{OQ}.$$

Os vectores $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ são os vectores de posição $\mathbf{p}$ e $\mathbf{q}$, respectivamente, dos pontos P e Q. Logo,

$$\begin{aligned}\overrightarrow{PQ} &= \overrightarrow{OQ} - \overrightarrow{OP} = \mathbf{q} - \mathbf{p}\\ &= (q_1, q_2, q_3) - (p_1, p_2, p_3) = (q_1 - p_1, q_2 - p_2, q_3 - p_3).\end{aligned}$$

As coordenadas do vector pretendido são portanto as coordenadas do vector $\mathbf{q}-\mathbf{p}$.

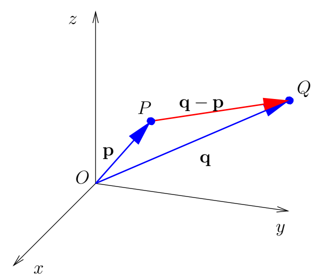


Figura 1.10: $\vec{PQ} = \mathbf{q} - \mathbf{p}$.

Vectores de $\mathbb{R}^n$

Por analogia com os casos $n = 2$ e $n = 3$, definimos $\mathbb{R}^n$ (n inteiro positivo) como o conjunto dos vectores coluna de n componentes reais. Nomeadamente

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : \; x_i \in \mathbb{R}, \text{ para } i = 1, \ldots, n \right\}$$

$$= \left\{ (x_1, x_2, \ldots, x_n) : \; x_i \in \mathbb{R} \text{ para } i = 1, \ldots, n \right\}.$$

O vector de $\mathbb{R}^n$ com todas as componentes iguais a zero chama-se o *vector nulo* (ou *vector zero*), que designamos por $\mathbf{0}$. A adição de vectores de $\mathbb{R}^n$ e a multiplicação destes vectores por números reais definem-se de modo análogo à definição apresentada para $n = 2$ e $n = 3$. Nomeadamente, para

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n,$$

definimos

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{e} \quad k\mathbf{x} = \begin{bmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{bmatrix} \text{ para } k \in \mathbb{R}. \tag{1.5}$$

Com base nas propriedades da adição e multiplicação de números reais, é fácil verificar as propriedades que a seguir enunciamos.

Propriedades algébricas de $\mathbb{R}^n$

Para quaisquer vectores $\mathbf{u}, \mathbf{v}, \mathbf{w}$ de $\mathbb{R}^n$ e números reais k e l, tem-se:

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

(iii) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u}$

(iv) $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$
onde $-\mathbf{u}$ designa $(-1)\mathbf{u}$

(v) $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

(vi) $(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$

(vii) $k(l\mathbf{u}) = (kl)\mathbf{u}$

(viii) $1\mathbf{u} = \mathbf{u}$

1.3.2 Combinação linear de vectores

Os conceitos de combinação linear de vectores e de conjunto gerado por um conjunto de vectores desempenham um papel fundamental em toda a álgebra linear. Em particular, o produto de uma matriz A por um vector (coluna) $\mathbf{x}$ será definido como uma combinação linear dos vectores coluna de A. Nesta secção introduzimos estas noções para vectores de $\mathbb{R}^n$, enfatizando os seus aspectos geométricos no caso de vectores do plano ou do espaço.

Definição 1.10. Diz-se que o vector $\mathbf{y} \in \mathbb{R}^n$ é uma *combinação linear* dos vectores $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_p$ de $\mathbb{R}^n$, se existem p números reais $c_1, c_2, \ldots, c_p$ tal que

$$\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p.$$

Aos escalares $c_1, c_2, \ldots, c_p$ chamamos *coeficientes* da combinação linear.

Designamos o conjunto de todas as combinações lineares de $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_p$ por

$$\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_p\} = \{c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p : \; c_1, \ldots, c_p \in \mathbb{R}\}.$$

A este subconjunto de $\mathbb{R}^n$ chamamos *conjunto gerado*[5]por, ou *expansão linear* de, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$.

Note-se que o conjunto de todas as combinações lineares de um vector $\mathbf{v}$, é o conjunto de todos os vectores da forma $\alpha\mathbf{v}$, onde α é um escalar real. Dois vectores $\mathbf{u}$ e $\mathbf{v}$ tais que $\mathbf{u} = \alpha\mathbf{v}$, com α um escalar real, recebem a designação de vectores *colineares*.

Exemplifiquemos agora, com vectores de $\mathbb{R}^2$ e $\mathbb{R}^3$, as noções que acabámos de definir.

Exemplo 1.12. Na Figura 1.11 ilustram-se algumas combinações lineares dos vectores $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ e $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Os conjuntos $\mathrm{Span}\{\mathbf{u}_1\}$ e $\mathrm{Span}\{\mathbf{u}_2\}$ são, respectivamente, os conjuntos

$$\mathrm{Span}\{\mathbf{u}_1\} = \{\alpha\mathbf{u}_1 : \; \alpha \in \mathbb{R}\}, \quad \mathrm{Span}\{\mathbf{u}_2\} = \{\alpha\mathbf{u}_2 : \alpha \in \mathbb{R}\}.$$

Geometricamente, estes conjuntos são rectas que passam pela origem e têm, respectivamente, a direcção de $\mathbf{u}_1$ e de $\mathbf{u}_2$.

O conjunto gerado por dois vectores é o conjunto das somas de todos os múltiplos de um vector com todos os múltiplos de outro vector. Para os vectores $\mathbf{u}_1$ e $\mathbf{u}_2$ o conjunto $\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ é o plano. Assim, de acordo com a Figura 1.11, tem-se:

$$\begin{aligned} &\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\} = \mathbb{R}^2, && \mathrm{Span}\{\mathbf{u}_1\} = l_1, && \mathrm{Span}\{\mathbf{u}_2\} = l_2 \\ &\mathbf{w} \in \mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\}, && \mathbf{w} \notin \mathrm{Span}\{\mathbf{u}_1\}, && \mathbf{w} \notin \mathrm{Span}\{\mathbf{u}_2\}. \end{aligned}$$

♦

[5] A designação “Span” para conjunto gerado é a designação habitualmente utilizada em língua inglesa.

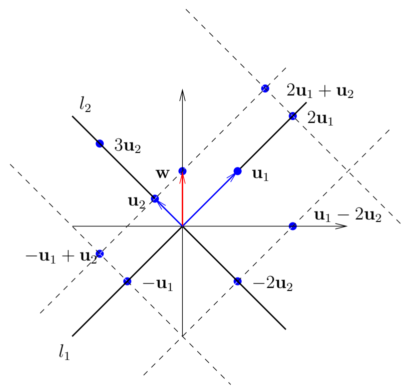


Figura 1.11: Combinação linear de vectores de $\mathbb{R}^2$. O vector $\mathbf{w}$ é $\mathbf{w} = \frac{1}{2}\mathbf{u}_1 + \mathbf{u}_2$.

Exemplo 1.13. Consideremos os vectores

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix} \quad \text{e} \quad \mathbf{w} = \begin{bmatrix} -4 \\ 18 \\ -1 \end{bmatrix}.$$

Pretendemos saber se $\mathbf{w}$ é ou não combinação linear de $\mathbf{u}_1$ e $\mathbf{u}_2$. Ou seja, se existem números reais c_1 e c_2 tais que

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \iff \begin{bmatrix} -4 \\ 18 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}.$$

Isto é,

$$\begin{bmatrix} -4 \\ 18 \\ -1 \end{bmatrix} = \begin{bmatrix} c_1 \\ 3c_1 \\ -2c_1 \end{bmatrix} + \begin{bmatrix} -2c_2 \\ 4c_2 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 - 2c_2 \\ 3c_1 + 4c_2 \\ -2c_1 + c_2 \end{bmatrix}.$$

Como dois vectores são iguais se as componentes correspondentes são iguais, a igualdade acima é equivalente ao sistema

$$\begin{cases} c_1 - 2c_2 = -4 \\ 3c_1 + 4c_2 = 18 \\ -2c_1 + c_2 = -1. \end{cases}$$

Usando o método de eliminação de Gauss para resolver este sistema, temos

$$\left[\begin{array}{cc|c} 1 & -2 & -4 \\ 3 & 4 & 18 \\ -2 & 1 & -1 \end{array}\right] \xrightarrow{L_2-3L_1} \left[\begin{array}{cc|c} 1 & -2 & -4 \\ 0 & 10 & 30 \\ -2 & 1 & -1 \end{array}\right]$$

$$\xrightarrow{L_3+2L_1} \left[\begin{array}{cc|c} 1 & -2 & -4 \\ 0 & 10 & 30 \\ 0 & -3 & -9 \end{array}\right] \xrightarrow{L_3+\frac{3}{10}L_2} \left[\begin{array}{cc|c} 1 & -2 & -4 \\ 0 & 10 & 30 \\ 0 & 0 & 0 \end{array}\right].$$

Como o sistema

$$\begin{cases} c_1 - 2c_2 = -4 \\ 10c_2 = 30 \end{cases}$$

tem solução $c_1 = 2$ e $c_2 = 3$, podemos dizer que $\mathbf{w}$ é combinação linear de $\mathbf{u}_1$ e $\mathbf{u}_2$. De facto,

$$\begin{bmatrix} -4 \\ 18 \\ -1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + 3\begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix} \Longleftrightarrow \mathbf{w} = 2\mathbf{u}_1 + 3\mathbf{u}_2.$$

Ou seja, o vector $\mathbf{w}$ pertence ao conjunto gerado por $\mathbf{u}_1$ e $\mathbf{u}_2$, isto é, $\mathbf{w} \in \mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

Note-se que o conjunto gerado por um único vector não nulo de $\mathbb{R}^3$ é uma recta que passa pela origem e tem a direcção do vector, enquanto que o conjunto gerado pelos vectores $\mathbf{u}_1$ e $\mathbf{u}_2$ é um plano que contém a origem. As rectas que passam pela origem e têm as direcções de $\mathbf{u}_1$ e de $\mathbf{u}_2$ pertencem a este plano, bem como o vector $\mathbf{w}$ (ver Figura 1.12).

♦

Exemplo 1.14. Consideremos os vectores

$$\mathbf{u}_1 = (1,1), \quad \mathbf{u}_2 = (-2,-2) \quad \text{e} \quad \mathbf{u}_3 = (1,2).$$

Os vectores $\mathbf{u}_1$ e $\mathbf{u}_2$ são colineares ($\mathbf{u}_1$ é proporcional a $\mathbf{u}_2$). Por conseguinte, o conjunto gerado por $\mathbf{u}_1$ e $\mathbf{u}_2$ é a recta que passa pela origem e tem a direcção destes vectores.

Os vectores $\mathbf{u}_1$ e $\mathbf{u}_3$ não são colineares (nem os vectores $\mathbf{u}_2$ e $\mathbf{u}_3$), logo o conjunto gerado por estes vectores é $\mathbb{R}^2$. Isto é,

$$\mathrm{Span}\{\mathbf{u}_1, \mathbf{u}_3\} = \mathbb{R}^2, \text{ bem como } \mathrm{Span}\{\mathbf{u}_2, \mathbf{u}_3\} = \mathbb{R}^2.$$

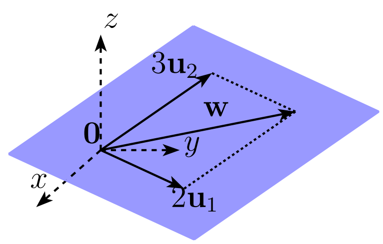


Figura 1.12: Os vectores $\mathbf{u}_1$ e $\mathbf{u}_2$ geram um plano e o vector $\mathbf{w}$ pertence a esse plano.

O conjunto gerado pelos três vectores $\mathbf{u}_1, \mathbf{u}_2$ e $\mathbf{u}_3$ é igualmente todo o $\mathbb{R}^2$. Ou seja,

$$\text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \mathbb{R}^2.$$

♦

Dos exemplos e definições apresentadas anteriormente podemos tirar as conclusões que se seguem. Para vectores não nulos, temos:

- Um vector de $\mathbb{R}^2$ (ou de $\mathbb{R}^3$) gera uma recta que passa pela origem e tem a direcção do vector.
- Dois vectores colineares de $\mathbb{R}^2$ (ou de $\mathbb{R}^3$) geram uma recta que passa pela origem e tem a direcção dos vectores.
- Dois vectores não colineares de $\mathbb{R}^2$ geram o plano $\mathbb{R}^2$. Se forem vectores de $\mathbb{R}^3$ geram um plano que passa pela origem e que contém as rectas que passam pela origem e cujas direcções são as dos vectores dados.
- Três ou mais vectores de $\mathbb{R}^2$, em que pelo menos dois deles não são colineares geram $\mathbb{R}^2$.
- Três ou mais vectores de $\mathbb{R}^3$, em que pelo menos três deles não são colineares dois a dois, geram $\mathbb{R}^3$.

Exercício 1.1. Mostrar que para quaisquer vectores $\mathbf{u}$ e $\mathbf{v}$ de $\mathbb{R}^n$ se verifica

$$\mathrm{Span}\{\mathbf{u},\mathbf{v}\} = \mathrm{Span}\{\mathbf{u}, \alpha\mathbf{u}+\beta\mathbf{v}\}, \text{ para todos os } \alpha \in \mathbb{R} \text{ e } \beta \in \mathbb{R}\setminus\{0\}.$$

▲

1.3.3 A equação matricial $A\mathbf{x} = \mathbf{b}$

Nesta secção mostra-se que um sistema de p equações lineares a n incógnitas reais pode escrever-se na forma matricial como $A\mathbf{x} = \mathbf{b}$, onde A é uma matriz $p\times n$, $\mathbf{x}$ é um vector de $\mathbb{R}^n$ e $\mathbf{b}$ é um vector de $\mathbb{R}^p$. Para tal necessitamos de definir o produto $A\mathbf{x}$. Iremos definir este produto como uma combinação linear das colunas de A.

Definição 1.11. Seja A a matriz $A = [a_{ij}]_{\substack{i=1,\ldots,p\\ j=1,\ldots,n}}$ e $\mathbf{x}$ um vector de n componentes.

O *produto de A por $\mathbf{x}$* é a combinação linear dos vectores coluna de A que tem para coeficientes as componentes de $\mathbf{x}$. Ou seja,

$$\begin{aligned} A\mathbf{x} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \underbrace{x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{pn} \end{bmatrix}}_{\text{combinação linear das colunas de } A}. \end{aligned}$$

Nota 5.

1. *O produto $A\mathbf{x}$ só está definido se o número de colunas de A for igual ao número de componentes de $\mathbf{x}$ (isto é, o número de colunas de A é igual ao número de linhas da matriz coluna $\mathbf{x}$).*

2. *De acordo com a definição acima, se A é do tipo $p \times n$ e $\mathbf{x}$ é do tipo $n \times 1$, então $A\mathbf{x}$ é um vector do tipo $p \times 1$.*

Da definição anterior, e usando as definições de produto de um escalar por um vector e de adição de vectores, temos

$$\begin{aligned} A\mathbf{x} &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{p1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{p2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{pn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pn}x_n \end{bmatrix}. \end{aligned} \tag{1.6}$$

Note-se que a entrada da linha i do vector $A\mathbf{x}$ é igual à soma dos produtos das entradas da linha i de A pelas componentes correspondentes de $\mathbf{x}$.

Exemplo 1.15. (a)

$$\underbrace{\begin{bmatrix} 1 & 3 & 4 & 2 \\ 2 & 4 & 5 & 1 \end{bmatrix}}_{2\times 4} \underbrace{\begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix}}_{4\times 1} = 1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} + 3 \begin{bmatrix} 4 \\ 5 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1\times 1 - 2\times 3 + 3\times 4 + 4\times 2 \\ 1\times 2 - 2\times 4 + 3\times 5 + 4\times 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 15 \\ 13 \end{bmatrix}}_{2\times 1}.$$

(b) $$\underbrace{\begin{bmatrix} 1 & 3 \\ 5 & 1 \\ 0 & 2 \\ 7 & 4 \end{bmatrix}}_{4\times 2} \underbrace{\begin{bmatrix} 1 \\ -2 \end{bmatrix}}_{2\times 1} = 1 \begin{bmatrix} 1 \\ 5 \\ 0 \\ 7 \end{bmatrix} - 2 \begin{bmatrix} 3 \\ 1 \\ 2 \\ 4 \end{bmatrix} = \underbrace{\begin{bmatrix} -5 \\ 3 \\ -4 \\ -1 \end{bmatrix}}_{4\times 1}.$$

♦

Exemplo 1.16. Considere-se o sistema de equações lineares

$$\begin{cases} x_1 + 3x_2 + 4x_3 + 2x_4 = 5 \\ 2x_1 + 4x_2 + 5x_3 + x_4 = 3. \end{cases}$$

Este sistema pode ser escrito como uma igualdade entre dois vectores coluna de $\mathbb{R}^2$, nomeadamente

$$\begin{bmatrix} x_1 + 3x_2 + 4x_3 + 2x_4 \\ 2x_1 + 4x_2 + 5x_3 + x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \mathbf{b}.$$

Porém, o lado esquerdo da equação anterior pode ser escrito como uma combinação linear de vectores de $\mathbb{R}^2$, isto é

$$\begin{aligned} \begin{bmatrix} x_1 + 3x_2 + 4x_3 + 2x_4 \\ 2x_1 + 4x_2 + 5x_3 + x_4 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 5 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 & 4 & 2 \\ 2 & 4 & 5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = A\mathbf{x}. \end{aligned}$$

Conclui-se que o sistema pode ser escrito na forma matricial $A\mathbf{x} = \mathbf{b}$, onde A é a matriz dos coeficientes do sistema e $\mathbf{b}$ é o vector coluna dos termos independentes. ♦

A igualdade (1.6) permite escrever qualquer sistema de equações lineares na forma $A\mathbf{x} = \mathbf{b}$, onde A é a matriz dos coeficientes do sistema, $\mathbf{x}$ é o vector cujas componentes são as variáveis do sistema, e $\mathbf{b}$ é o vector coluna cujas componentes são o segundo membro das equações do sistema. Por conseguinte, dizer que um sistema de equações lineares é possível é dizer que existe um vector $\mathbf{x}$ que satisfaz a igualdade $A\mathbf{x} = \mathbf{b}$. Ou seja, o sistema $A\mathbf{x} = \mathbf{b}$ é possível se e só se o vector $\mathbf{b}$ é uma combinação linear das colunas de A. Na proposição seguinte enunciamos este resultado para referência futura.

Proposição 1.3. Um sistema de equações lineares $A\mathbf{x} = \mathbf{b}$ é possível se e só se $\mathbf{b}$ é uma combinação linear das colunas de A.

Quando no sistema $A\mathbf{x} = \mathbf{b}$ o vector $\mathbf{b}$ é o vector nulo, dizemos que o sistema é um *sistema homogéneo*. Um sistema homogéneo é sempre possível já que, admite pelo menos a solução nula.

Exemplo 1.17. Façamos a discussão do sistema homogéneo $A_\alpha \mathbf{x} = \mathbf{0}$ em termos do parâmetro real α, onde

$$A_\alpha = \begin{bmatrix} 1 & 1 & -1 \\ 4 & \alpha^2 & -4 \\ 2 & 2 & \alpha \end{bmatrix}.$$

Apliquemos o método de eliminação de Gauss à matriz dos coeficientes do sistema.

$$\begin{bmatrix} 1 & 1 & -1 \\ 4 & \alpha^2 & -4 \\ 2 & 2 & \alpha \end{bmatrix} \xrightarrow{L_2-4L_1} \begin{bmatrix} 1 & 1 & -1 \\ 0 & \alpha^2-4 & 0 \\ 2 & 2 & \alpha \end{bmatrix} \xrightarrow{L_3-2L_1} \begin{bmatrix} 1 & 1 & -1 \\ 0 & \alpha^2-4 & 0 \\ 0 & 0 & \alpha+2 \end{bmatrix}.$$

A característica de A_α é

$$\operatorname{car}(A_\alpha) = \begin{cases} 1 & \text{se } \alpha = -2 \\ 2 & \text{se } \alpha = 2 \\ 3 & \text{se } \alpha \in \mathbb{R} \setminus \{-2, 2\}. \end{cases}$$

De acordo com a Proposição 1.2 (pág. 18), concluimos que para $\alpha \in \mathbb{R} \setminus \{-2, 2\}$ o sistema só possui a solução nula (não existem incógnitas livres), enquanto que para $\alpha = 2$ o sistema tem grau de indeterminação 1, e para $\alpha = -2$ tem grau de indeterminação 2.

O sistema $A_\alpha \mathbf{x} = \mathbf{0}$ é equivalente ao sistema

$$\begin{cases} x & +y & -z & = 0 \\ & (\alpha^2-4)y & & = 0 \\ & & (\alpha+2)z & = 0. \end{cases} \tag{1.7}$$

Quando $\alpha = 2$ o sistema reduz-se à primeira e terceira equações. Resolvendo esse sistema obtém-se $z = 0$ e $x = -y$. Consequentemente, para $\alpha = 2$ a solução geral é

$$\begin{aligned} \left\{(x, y, z) \in \mathbb{R}^3 : x = -y \text{ e } z = 0\right\} &= \{(-y, y, 0) : y \in \mathbb{R}\} \\ &= \{y(-1, 1, 0) : y \in \mathbb{R}\} = \operatorname{Span}\{(-1, 1, 0)\}. \end{aligned}$$

Para $\alpha = -2$ o sistema (1.7) reduz-se à primeira equação, donde $x = -y + z$. Neste caso a solução geral é

$$\begin{aligned} \left\{(x, y, z) \in \mathbb{R}^3 : x = -y + z\right\} &= \{(-y + z, y, z) : y, z \in \mathbb{R}\} \\ &= \{y(-1, 1, 0) + z(1, 0, 1) : y, z \in \mathbb{R}\} \\ &= \operatorname{Span}\{(-1, 1, 0), (1, 0, 1)\}. \end{aligned}$$

♦

1.3.4 Operações com matrizes

Já definimos as operações de adição de vectores coluna, multiplicação de um escalar por um vector e multiplicação de uma matriz por um vector coluna. Baseados nestas operações, definimos agora operações sobre matrizes com várias colunas encarando para tal uma matriz como um conjunto ordenado de vectores coluna. Posteriormente, estudamos as propriedades das operações que definimos a seguir.

Sejam A e B duas matrizes do mesmo tipo, por exemplo $p \times n$,

$$A = [a_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}} \quad \text{e} \quad B = [b_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}} .$$

- A e B são *iguais* se as entradas respectivas são iguais, isto é, $a_{ij} = b_{ij}$ para todo $i = 1, \ldots, p$ e $j = 1, \ldots, n$.

- A soma $A + B$ é a matriz $p \times n$ cujas entradas são as somas das entradas de A com as respectivas entradas de B. Tal significa que a entrada na linha i e na coluna j de $A + B$ é $a_{ij} + b_{ij}$, para todo $i = 1, \ldots, p$ e $j = 1, \ldots, n$.

- A multiplicação de uma matriz A, por um escalar α é a matriz cujas entradas são as entradas de A multiplicadas por α.

Resumindo:

Sejam $A = [a_{ij}]$ e $B = [b_{ij}]$ matrizes do tipo $p \times n$ e α um escalar. Então,

$$A + B = [a_{ij} + b_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}} \tag{1.8}$$

$$\alpha A = [\alpha a_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}} . \tag{1.9}$$

Nota 6. *A adição de matrizes só está definida quando as matrizes são do mesmo tipo.*

Seguidamente definimos multiplicação de matrizes. Se considerarmos uma matriz B como um conjunto ordenado de vectores coluna, é natural definir o produto AB como o conjunto das colunas que são o produto de A pelas colunas de B, respeitando a ordem de B.

Definição 1.12. Seja A uma matriz $p \times n$ e B uma matriz $n \times k$ com colunas $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$.

O produto AB é a matriz $p \times k$ cujas colunas são $A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_k$. Isto é,

$$AB = A \begin{bmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_k \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A\mathbf{b}_1 & A\mathbf{b}_2 & \ldots & A\mathbf{b}_k \\ | & | & & | \end{bmatrix}.$$

Note-se que o produto AB só está definido se o número de colunas de A é igual ao número de linhas de B. Além disso, a matriz AB é uma matriz com o mesmo número de linhas de A e o mesmo número de colunas de B.

$$\begin{array}{cccc} A & B & = & AB. \\ p \times n & n \times k & & p{\times}k \end{array}$$

No caso particular da matriz B ser $n \times 1$ (ou seja, um vector coluna) a definição do produto AB coincide com a definição anteriormente apresentada (como não podia deixar de ser).

Por razões práticas convém deduzir como se calcula uma entrada específica do produto de duas matrizes. Para tal comecemos com dois exemplos.

Exemplo 1.18. Calculemos

$$A\mathbf{u} = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}.$$

Por definição de produto de uma matriz por um vector, o produto $A\mathbf{u}$ é a matriz 2×1 dada por

$$A\mathbf{u} = A \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = 5 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 6 \begin{bmatrix} 3 \\ 3 \end{bmatrix} + 7 \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \times 5 + 3 \times 6 + 4 \times 7 \\ 1 \times 5 + 3 \times 6 - 2 \times 7 \end{bmatrix} = \begin{bmatrix} 56 \\ 9 \end{bmatrix}.$$

♦

O exemplo anterior ilustra como podemos obter as entradas de uma certa coluna do produto de duas matrizes. Nomeadamente, se $C = AB$, a entrada c_{ij} de C obtém-se multiplicando as entradas da linha i de A pelas entradas correspondentes da coluna j de B e somando os produtos obtidos.

Exemplo 1.19. Considerem-se as matrizes

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

Como A é uma matriz 2×3 e B é 3×4, o produto AB está definido e AB é uma matriz do tipo 2×4. Para determinar a entrada na linha 2 e na coluna 3 da matriz AB escolhemos a linha 2 de A e a coluna 3 de B, multiplicamos as entradas da referida linha pelas entradas respectivas da coluna selecionada, e somamos. Ou seja,

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 1 \\ -1 & 0 & 2 & 3 \\ 5 & 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \underset{\uparrow}{\boxed{26}} & \square \end{bmatrix} = AB.$$

$$\boxed{0 \times 3 + 4 \times 2 + 6 \times 3 = 26}$$

Por exemplo, a entrada na linha 1 e na coluna 4 de AB é

$$\boxed{2 \times 1 + 3 \times 3 + 1 \times 1 = 12}$$

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 1 \\ -1 & 0 & 2 & 3 \\ 5 & 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \overset{\downarrow}{\boxed{12}} \\ \square & \square & \square & \square \end{bmatrix}.$$

Calculando as restantes entradas de AB, obtemos

$$AB = \begin{bmatrix} 4 & 6 & 15 & 12 \\ 26 & 12 & 26 & 18 \end{bmatrix}.$$

♦

Para $A = [a_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}}$, $B = [b_{ij}]_{\substack{i=1,\ldots,n \\ j=1,\ldots,k}}$ e $AB = C = [c_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,k}}$, isto é

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1k} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots & \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nk} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1k} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ik} \\ \vdots & & \vdots & & \vdots \\ c_{p1} & \cdots & c_{pj} & \cdots & c_{pk} \end{bmatrix},$$

tem-se $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{r=1}^{n} a_{ir}b_{rj}$. Resumindo:

Se $A = [a_{ij}]$ é do tipo $p \times n$ e $B = [b_{ij}]$ do tipo $n \times k$, a entrada na linha i e na coluna j da matriz $AB = C = [c_{ij}]$ é dada por

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{r=1}^{n} a_{ir}b_{rj}. \tag{1.10}$$

Propriedades das operações com matrizes

A adição de matrizes reais (resp. complexas) goza das propriedades comutativa e associativa como facilmente se deduz usando as propriedades comutativa e associativa dos números reais (resp. complexos) e a definição (1.8) de adição de matrizes. Contudo, a multiplicação de matrizes **não goza** em geral da propriedade comutativa, ou seja, pode ter-se $AB \neq BA$ mesmo no caso em que ambos os produtos AB e BA estão definidos. Por exemplo, para

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

tem-se

$$AB = \begin{bmatrix} 0 & 14 \\ 5 & 17 \end{bmatrix} \quad \text{e} \quad BA = \begin{bmatrix} 14 & 8 \\ 14 & 3 \end{bmatrix}.$$

Logo, $AB \neq BA$. Quando $AB = BA$ dizemos que A e B *comutam*.

Passamos a enunciar algumas propriedades das operações com matrizes.

Teorema 1.1. Sejam A, B e C matrizes para as quais as operações abaixo indicadas estão definidas, e α, β escalares. São válidas as seguintes propriedades:

(a) $A + B = B + A$	**Comutatividade da adição**
(b) $A + (B + C) = (A + B) + C$	**Associatividade da adição**
(c) $A(BC) = (AB)C$	**Associatividade da multiplicação**
(d) $A(B + C) = AB + AC$	**Distributividade à esquerda**
(e) $(B + C)A = BA + CA$	**Distributividade à direita**
(f) $\alpha(A + B) = \alpha A + \alpha B$	
(g) $(\alpha + \beta)A = \alpha A + \beta A$	
(h) $\alpha(AB) = (\alpha A)B = A(\alpha B)$	

As alíneas (a), (b), (f), (g) e (h) são imediatas usando propriedades das operações com números reais ou complexos. A seguir apresentamos apenas a demonstração da alínea (d), demonstração esta que, no entanto, é bem ilustrativa dos procedimentos a seguir para as outras alíneas.

Demonstração. (d): Para que $A(B + C)$ seja igual a $AB + AC$, as matrizes B e C deverão ter o mesmo tamanho e o número de colunas de A deverá ser igual ao número de linhas de B e de C. Suponha-se que $A = [a_{ij}]$ tem n colunas e que as matrizes $B = [b_{ij}]$ e $C = [c_{ij}]$ têm n linhas. De modo a facilitar a compreensão da demonstração, usaremos simultaneamente a notação d_{ij} e $[D]_{ij}$ para designar a entrada ij da matriz D.

Pretendemos mostrar que as entradas homólogas de $A(B+C)$ e de $AB+AC$ são iguais, ou seja, que é satisfeita a igualdade

$$[A(B + C)]_{ij} = [AB]_{ij} + [AC]_{ij},$$

para quaisquer valores de i e de j. Usando (1.8) e (1.10) obtemos

$$\begin{aligned}
[A(B + C)]_{ij} &= a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \cdots + a_{in}(b_{nj} + c_{nj}) \\
&= (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{in}c_{nj}) \\
&= [AB]_{ij} + [AC]_{ij}.
\end{aligned}$$

□

Nota 7. *Para quaisquer números reais a e b é válida a propriedade $ab = 0$ se e só se $a = 0$ ou $b = 0$. Este resultado não é em geral válido para o produto de matrizes, como se verifica no exemplo que apresentamos a seguir.*

Exemplo 1.20. Considere-se $A = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}$ e $B = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}$. O produto AB é igual à matriz nula

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

porém, nem A nem B são matrizes nulas. ♦

As propriedades algébricas das matrizes são essenciais na demonstração de um grande número de resultados. Por exemplo, as propriedades das operações com matrizes habilitam-nos agora a apresentar a demonstração de duas proposições anteriormente referidas. Sem estas propriedades a prova destas proposições seria com certeza muito mais extensa.

A proposição seguinte diz-nos que um sistema possível ou tem solução única ou uma infinidade de soluções.

Proposição 1.4. Se um sistema de equações lineares admite mais do que uma solução, então possui uma infinidade de soluções.

Demonstração. Sejam $\mathbf{u}$ e $\mathbf{v}$ soluções distintas do sistema $A\mathbf{x} = \mathbf{b}$. O vector $\mathbf{y} = \mathbf{u} - \mathbf{v}$ é uma solução não nula do sistema homogéneo $A\mathbf{x} = \mathbf{0}$, já que

$$A\mathbf{y} = A(\mathbf{u} - \mathbf{v}) = A\mathbf{u} - A\mathbf{v} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Para qualquer escalar α, o vector $\mathbf{w} = \mathbf{u} + \alpha\mathbf{y}$ é solução do sistema $A\mathbf{x} = \mathbf{b}$ já que

$$A\mathbf{w} = A(\mathbf{u} + \alpha\mathbf{y}) = A\mathbf{u} + \alpha A\mathbf{y} = \mathbf{b} + \alpha\mathbf{0} = \mathbf{b}.$$

Logo, se $\mathbf{u}$ e $\mathbf{v}$ são soluções distintas do sistema $A\mathbf{x} = \mathbf{b}$, então qualquer vector da forma $\mathbf{w} = \mathbf{u} + \alpha\mathbf{y}$, com $\alpha \in \mathbb{R}$ e $\mathbf{y} = \mathbf{u} - \mathbf{v}$, é solução do sistema. Ou seja, há uma infinidade de soluções $\mathbf{w} = \mathbf{u} + \alpha\mathbf{y}$ (uma para cada α fixado em $\mathbb{R}$). □

Relembremos que na génese da noção de característica de uma matriz, apresentada na Secção 1.2 (pág. 12), encontra-se a *unicidade da forma reduzida em escada por linhas de uma matriz*. O teorema que enunciamos a seguir estabelece este resultado.

Teorema 1.2. A forma reduzida em escada por linhas de uma matriz é única.

Demonstração. [6] Seja A uma matriz do tipo $p \times n$. Usemos indução[7] sobre n.

Para $n = 1$, a matriz A tem uma única coluna e portanto a correspondente forma reduzida em escada por linhas é obviamente única. Considere-se como hipótese de indução que para qualquer matriz do tipo $p \times (n-1)$ a respectiva forma reduzida em escada é única.

Seja A' a matriz que se obtém de A suprimindo a última coluna. Note-se que qualquer sequência de operações elementares sobre A que reduza A a uma matriz em escada na forma reduzida também reduz A' a uma matriz em escada por linhas na forma reduzida. Assim, pela hipótese de indução, se B e C são formas reduzidas em escada por linhas da matriz A, estas matrizes diferem quando muito na coluna número n.

Suponha-se que B e C são matrizes em escada na forma reduzida, obtidas de A, e que (por absurdo) $B \neq C$. Pretendemos chegar a uma contradição. Sendo $B \neq C$, existe um inteiro k tal que a linha k de B é distinta da linha k de C. Seja $\mathbf{u}$ uma qualquer solução do sistema $B\mathbf{x} = \mathbf{0}$. Como B e C são matrizes obtidas de A por meio de operações elementares sobre as suas linhas, os sistemas $B\mathbf{x} = \mathbf{0}$ e $C\mathbf{x} = \mathbf{0}$ têm as mesmas soluções. Dado que $B\mathbf{u} = \mathbf{0} = C\mathbf{u}$, da propriedade distributiva segue $(B - C)\mathbf{u} = \mathbf{0}$. Como B e C têm as primeiras $(n-1)$ colunas iguais (por hipótese de indução) a matriz $(B-C)$ tem as primeiras $(n-1)$ colunas nulas, e portanto a componente k do vector $(B-C)\mathbf{u}$ é $(b_{kn} - c_{kn})u_n = 0$. Esta equação possui como única solução $u_n = 0$ já que, por hipótese, $b_{kn} \neq c_{kn}$. Logo, qualquer solução $\mathbf{u}$ dos sistemas $B\mathbf{x} = \mathbf{0}$ e $C\mathbf{x} = \mathbf{0}$ tem a $n^{\text{ésima}}$ componente nula. Assim sendo, a coluna número n de B e de C tem um pivô (igual a 1), caso contrário x_n seria uma variável livre dos sistemas $B\mathbf{x} = \mathbf{0}$ e $C\mathbf{x} = \mathbf{0}$. Como as primeiras $(n-1)$ colunas de B e de C são iguais, o pivô da coluna n de B e de C aparece na mesma linha (a última). Logo, por definição de matriz em escada por linhas na forma reduzida, tem-se $B = C$, o que é uma contradição. □

Matriz Identidade

A matriz identidade de ordem n é uma matriz quadrada $n \times n$ cujas entradas da diagonal principal são todas iguais a 1 e as restantes entradas são nulas. Por

[6] Esta demonstração é apresentada por T. Yuster em [13].

[7] O método de indução matemática é um método de demonstração usado frequentemente quando se pretende provar que uma certa proposição é verdadeira para todos os números naturais n. Uma versão deste método é a seguinte: (i) mostrar que a proposição é válida para o primeiro valor de n; (ii) supor que a proposição é válida para $n = r$ (esta hipótese é conhecida como hipótese de indução); (iii) deduzir de (ii) que a proposição é válida para $n = r + 1$.

exemplo,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

são respectivamente a matriz identidade de ordem 2, 3, e 4.

Definição 1.13. A *matriz identidade de ordem* n é uma matriz quadrada $n \times n$ cujas entradas na diagonal principal são todas iguais a 1 e as outras entradas iguais a zero.

A matriz identidade de ordem n será designada I_n, ou simplesmente por I quando no contexto for claro qual a ordem da matriz. Isto é,

$$I_n = [\delta_{ij}]_{i,j=1,\dots,n} \quad \text{com} \quad \delta_{ij} = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}$$

As entradas da matriz identidade são tradicionalmente denotadas pelo símbolo δ_{ij}. Este símbolo é conhecido por *delta de Kronecker*[8] e tem o significado especificado na definição anterior.

A partir da definição de multiplicação de matrizes, é fácil mostrar (basta usar (1.10)) que o produto de uma matriz A por uma matriz identidade é sempre igual a A. Ilustremos este facto com o exemplo seguinte.

Exemplo 1.21. Seja $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$.

Apenas podemos multiplicar A à esquerda pela matriz identidade de ordem 3, e à direita pela de ordem 4.

$$I_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = A$$

e

$$AI_4 = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = A.$$

[8] Leopold Kronecker (1823-1891), matemático alemão.

♦

A matriz identidade é um exemplo de uma matriz *diagonal*. Isto é, uma matriz quadrada com todas as entradas iguais a zero excepto possivelmente as entradas da diagonal principal. Por exemplo, a matriz

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

é uma matriz diagonal.

Transposta de uma matriz

A *transposta* de uma matriz A do tipo $p \times n$, é a matriz $n \times p$ cujas linhas são as colunas de A, pela mesma ordem em que aparecem em A. Designamos por A^T a matriz transposta de A. Por exemplo, para

$$A = \begin{bmatrix} 1 & 2 & 4 & 5 \\ 2 & 3 & 5 & 6 \end{bmatrix} \quad \text{tem-se} \quad A^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 5 \\ 5 & 6 \end{bmatrix}.$$

Como as linhas de A^T são as colunas de A, pela mesma ordem, tem-se:

$$\text{se } A = [a_{ij}]_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}} \text{ e } A^T = [a'_{ij}]_{\substack{i=1,\ldots,n \\ j=1,\ldots,p}}, \text{ então } a'_{ij} = a_{ji}.$$

Enunciamos a seguir algumas propriedades da transposta de uma matriz.

Propriedades da matriz transposta

Sejam A e B matrizes para as quais as operações abaixo indicadas estão definidas. Verificam-se as igualdades:

(a) $(A^T)^T = A$;

(b) $(A + B)^T = A^T + B^T$;

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

Demonstração. Deixamos como exercício a prova de (a) e (b) e mostramos somente (c). Para tal, considere-se

$$A = [a_{ij}]_{\substack{i=1,\dots p \\ j=1,\dots n}} \quad \text{e} \quad B = [b_{ij}]_{\substack{i=1,\dots n \\ j=1,\dots k}}.$$

A matriz AB é do tipo $p \times k$ e portanto $(AB)^T$ é do tipo $k \times p$. Da mesma forma, $B^T A^T$ é do tipo $k \times p$. Designemos por $\left[(AB)^T\right]_{ij}$ a entrada ij de $(AB)^T$ e por $[B^T A^T]_{ij}$ a entrada ij de $B^T A^T$. Assim, como as linhas da transposta de uma matriz são as colunas da matriz, tem-se

$$\left[(AB)^T\right]_{ij} = [AB]_{ji} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni}, \tag{1.11}$$

onde na última igualdade usámos a igualdade (1.10), da definição de produto de matrizes.

Para calcular $[B^T A^T]_{ij}$ designemos por b'_{ij} a entrada ij de B^T e por a'_{ij} a de A^T. É claro que

$$b'_{ij} = b_{ji} \quad \text{e} \quad a'_{ij} = a_{ji}.$$

Usando de novo (1.10) temos

$$\left[B^T A^T\right]_{ij} = b'_{i1}a'_{1j} + b'_{i2}a'_{2j} + \cdots + b'_{in}a'_{nj} = b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ni}a_{jn}.$$

Comparando a última igualdade com (1.11) fica mostrado o resultado enunciado na alínea (c).

□

Definição 1.14. Uma matriz A diz-se *simétrica* se $A = A^T$ e *anti-simétrica* se $A = -A^T$.

É de realçar que a definição anterior implica que matrizes simétricas e anti-simétricas sejam matrizes quadradas. De facto, a igualdade $A = \pm A^T$ só é possível se A é quadrada. Além disso, como as entradas da diagonal principal de A e A^T são as mesmas, uma matriz anti-simétrica deverá ter estas entradas iguais a zero, uma vez que satisfaz a igualdade $A = -A^T$. A designação de matriz simétrica está relacionada com a simetria que este tipo de matrizes possui relativamente à diagonal principal, conforme se verifica nos exemplos apresentados a seguir.

Exemplo 1.22. No sentido de clarificar a noção de matriz simétrica e anti-simétrica apresentamos alguns exemplos.

$$\text{A matriz } \begin{bmatrix} 2 & 5 & 3 \\ 5 & 1 & 0 \\ 3 & 0 & 2 \end{bmatrix} \qquad \text{é simétrica.}$$

$$\text{A matriz } \begin{bmatrix} 0 & 5 & -3 \\ -5 & 0 & 1 \\ 3 & -1 & 0 \end{bmatrix} \qquad \text{é anti-simétrica.}$$

$$\text{A matriz } \begin{bmatrix} 1 & 5 & -3 \\ -5 & 0 & 1 \\ 3 & -1 & 0 \end{bmatrix} \qquad \text{não é anti-simétrica (nem simétrica).}$$

♦

Exemplo 1.23. Dada uma matriz A do tipo $p \times n$, as matrizes A^TA e AA^T são simétricas, respectivamente, do tipo $n\times n$ e $p\times p$. De facto, usando as propriedades da matriz transposta atrás enunciadas, resultam as seguintes igualdades:

$$(A^TA)^T \overset{(c)}{=} A^T(A^T)^T \overset{(a)}{=} A^TA \quad \text{e} \quad (AA^T)^T \overset{(c)}{=} (A^T)^TA^T \overset{(a)}{=} AA^T.$$

♦

Exercício 1.2. Seja A uma matriz quadrada.

a) Mostrar que $(A + A^T)$ é simétrica e $(A - A^T)$ é anti-simétrica.
b) Usar a alínea anterior para escrever A como a soma de uma matriz simétrica com uma anti-simétrica.

1.4 Inversa de uma matriz

Nesta secção apresentamos a definição de matriz inversa de uma matriz (quadrada) e estudamos as suas propriedades, nomeadamente a relação entre a existência de inversa e a característica da matriz. O cálculo da inversa de uma matriz é efectuado recorrendo ao método de eliminação de Gauss-Jordan, método este que pode ser descrito em termos de multiplicação sucessiva por certas matrizes ditas elementares. Mostra-se ainda que a matriz dos coeficientes de um sistema, com igual número de equações e de incógnitas, é invertível se e só se o sistema tem uma única solução.

Definição 1.15. Uma matriz quadrada A do tipo $n \times n$, diz-se *invertível*, ou *não singular*, se existe uma matriz B do tipo $n \times n$, tal que

$$AB = I_n = BA. \tag{1.12}$$

Se B é uma matriz que satisfaz a igualdade anterior dizemos que B é uma *inversa* de A.

Uma matriz que não seja invertível diz-se uma *matriz singular*.

Exemplo 1.24. A matriz

$$A = \begin{bmatrix} 1 & 10 & 3 \\ 0 & 0 & 0 \\ 1 & 3 & 2 \end{bmatrix}$$

não é invertível. Para verificar este facto considere-se uma matriz B do tipo 3×3 de entradas b_{ij} e efectue-se o produto AB. Como a segunda linha de AB é nula, este produto nunca pode ser igual à matriz identidade de terceira ordem.

♦

Nota 8. *O exemplo anterior ilustra o facto mais geral: qualquer matriz quadrada que possua uma linha nula é uma matriz singular. Este resultado pode ser provado usando o mesmo procedimento do exemplo anterior e a expressão* (1.10).

A unicidade da matriz inversa é estabelecida no próximo teorema.

Teorema 1.3. A inversa de uma matriz, quando existe, é única.

Demonstração. Suponha-se que B e C são inversas de A. Isto é,

$$AB = BA = I \quad \text{e} \quad AC = CA = I.$$

Uma vez que $BA = I = AC$, tem-se

$$BAC = B(AC) = BI = B,$$

e

$$BAC = (BA)C = IC = C.$$

Logo, $B = C$. □

Passamos a designar a matriz inversa de uma matriz invertível A por:

$$A^{-1}$$

Proposição 1.5. Se $ad - bc \neq 0$, a inversa da matriz

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

é dada por

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Demonstração. Deixamos como exercício verificar que $AA^{-1} = I_2$ e $A^{-1}A = I_2$. □

Exemplo 1.25. A matriz $\begin{bmatrix} 2 & 1 \\ 4 & 4 \end{bmatrix}$ é invertível e a sua inversa é

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

♦

Verifiquemos agora que o produto de matrizes invertíveis ainda é invertível.

Proposição 1.6. Se A e B são matrizes invertíveis da mesma ordem, então AB é invertível e a sua inversa é $B^{-1}A^{-1}$. Ou seja,

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{1.13}$$

Demonstração. Como a inversa de uma matriz é única, basta-nos mostrar que $B^{-1}A^{-1}$ é a inversa de AB.

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} && \text{(associatividade da multiplicação)} \\ &= AIA^{-1} && \text{(já que } B^{-1} \text{ é a inversa de } B\text{)} \\ &= AA^{-1} \\ &= I. \end{aligned}$$

Da mesma forma,

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

□

Potências de uma matriz

Para qualquer inteiro não negativo k e qualquer matriz (quadrada) A, definimos A^k como o produto de k factores todos iguais a A

$$A^k = \underbrace{AA\cdots A}_{k \text{ factores}}$$

sendo a expressão A^0 interpretada como a matriz identidade.

Se A é uma matriz invertível, definimos potências negativas de A em termos de potências positivas de A^{-1}. Isto é,

$$A^{-k} = (A^{-1})^k = \underbrace{A^{-1}A^{-1}\cdots A^{-1}}_{k \text{ factores}}$$

A proposição seguinte sumariza alguns resultados sobre matrizes invertíveis e potências destas matrizes.

Proposição 1.7. Seja A uma matriz invertível. São válidas as proposições seguintes.

1) A^{-1} é invertível e $\left(A^{-1}\right)^{-1} = A$.

2) A^k é invertível e $\left(A^k\right)^{-1} = \left(A^{-1}\right)^k = A^{-k}$.

3) Para qualquer escalar não nulo α, a matriz αA é invertível e $(\alpha A)^{-1} = \dfrac{1}{\alpha}A^{-1}$.

Demonstração. As provas das alíneas 1) e 2) são deixadas como exercício. Façamos a prova de 3). Como

$$(\alpha A)\left(\frac{1}{\alpha}A^{-1}\right) = \alpha\frac{1}{\alpha}AA^{-1} = I,$$

e $\left(\dfrac{1}{\alpha}A^{-1}\right)(\alpha A) = \alpha\dfrac{1}{\alpha}A^{-1}A = I$, segue o resultado. $\square$

Como já referimos, o produto de matrizes não é em geral comutativo. Assim, a potência

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

é em geral diferente de $A^2+2AB+B^2$ (excepto se as matrizes A e B comutarem). Sugerimos a procura de um contra-exemplo.

1.4.1 Matrizes elementares e cálculo da inversa

Nesta secção apresentamos um algoritmo para calcular a inversa de uma matriz. Subjacente a este algoritmo está o facto de ser possível obter a matriz identidade aplicando operações elementares à matriz invertível dada. Cada operação elementar (ver definição na página 8) aplicada a uma matriz vai traduzir-se na multiplicação dessa matriz por uma matriz dita elementar.

Definição 1.16. Um matriz *elementar* é uma matriz quadrada $n \times n$ que é obtida da identidade I_n por meio de uma única operação elementar.

Exemplo 1.26. Exemplos de matrizes elementares:

(a) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Troca da segunda linha de I_3 com a primeira linha.

(b) $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$ A primeira linha de I_2 foi substituída pela sua soma com a 2ª linha multiplicada por 3. ♦

(c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ A segunda linha de I_4 foi multiplicada por 2.

Quando uma matriz A é multiplicada à *esquerda* por uma matriz elementar, a matriz produto é a matriz que se obtém de A efectuando a mesma operação elementar efectuada em I para obter a matriz elementar em causa. A proposição seguinte enuncia este facto. A sua demonstração é deixada como exercício.

Proposição 1.8. Se E é uma matriz elementar, então EA é a matriz que se obtém de A efectuando a mesma operação elementar que permite obter E a partir da identidade.

Ilustremos esta proposição através de um exemplo.

Exemplo 1.27. Consideremos a matriz

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 & 2 \\ 5 & 3 & 1 & 5 & 8 \\ 10 & 2 & 7 & 2 & 6 \end{bmatrix},$$

e a matriz elementar

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

obtida de I_2 por:

$$I_2 \overset{L_2+2L_1}{\longrightarrow} E.$$

A matriz E resulta de I_2 substituindo a segunda linha pela soma desta linha com a primeira linha multiplicada por 2 . Calculando EA, verificamos que

$$A = \begin{bmatrix} 1 & 3 & 5 & 7 & 2 \\ 5 & 3 & 1 & 5 & 8 \\ 10 & 2 & 7 & 2 & 6 \end{bmatrix} \overset{L_2+2L_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 5 & 7 & 2 \\ 7 & 9 & 11 & 19 & 12 \\ 10 & 2 & 7 & 2 & 6 \end{bmatrix} = EA.$$

♦

Nota 9. *A Proposição 1.8 só é válida para multiplicação à esquerda. De facto, se uma matriz A é multiplicada à direita por uma matriz elementar E, a matriz AE é a matriz que se obtém de A efectuando a operação elementar correspondente a E sobre as colunas de A. Por exemplo, para $\alpha \neq 0$ e*

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix},\ E = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \textit{ tem-se } EA = \begin{bmatrix} 1 & 3 \\ 2\alpha & 4\alpha \end{bmatrix} \textit{ e } AE = \begin{bmatrix} 1 & 3\alpha \\ 2 & 4\alpha \end{bmatrix}.$$

Como veremos, qualquer matriz elementar E é invertível já que, se E é obtida da identidade por meio de uma operação elementar, então efectuando sobre E a operação elementar "inversa" obtemos a identidade. Ou seja, existe sempre uma matriz elementar (correspondente à operação elementar "inversa") que quando multiplicada (à esquerda) por E dá a identidade (e esta matriz é a inversa de E). No exemplo seguinte ilustramos este facto bem como o que se entende por operação elementar "inversa".

Exemplo 1.28. Calculemos a inversa das matrizes elementares E seguintes.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \overset{3L_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \overset{\frac{1}{3}L_2}{\longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{e} \quad E^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}.$$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{L_2+2L_1} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{L_2-2L_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

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

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{L_3 \leftrightarrow L_1} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \xrightarrow{L_3 \leftrightarrow L_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{e} \quad E^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Para cada uma das alíneas anteriores, calculando os produtos EE^{-1} e $E^{-1}E$ confirma-se que é obtida a matriz identidade. ♦

No quadro seguinte resumimos o que entendemos por operação elementar inversa.

Operação elementar	Operação elementar inversa
$L_i \leftrightarrow L_j$	$L_i \leftrightarrow L_j$
$\alpha L_i \ \ (\alpha \neq 0)$	$\frac{1}{\alpha} L_i \ \ (\alpha \neq 0)$
$L_i + \alpha L_j$	$L_i - \alpha L_j$

Uma matriz elementar que é obtida da matriz identidade por troca de duas das suas linhas é um exemplo de uma matriz de *permutação*[9]. Como se deduz facilmente, a inversa de uma matriz elementar de permutação coincide com a própria matriz (ver Exemplo1.28-(c)).

Proposição 1.9. Qualquer matriz elementar é invertível e a sua inversa é uma matriz elementar.

[9]Uma matriz de permutação é uma matriz que se obtém da identidade efectuando um número finito de trocas de linhas.

Demonstração. Sendo E uma matriz elementar, esta matriz é obtida da identidade efectuando uma operação elementar. Designe-se por E_1 a matriz que se obtém da identidade efectuando a operação inversa da que permitiu obter E. Então, pela Proposição 1.8, temos

$$E_1 E = I.$$

De igual modo, como E é uma matriz elementar, também se verifica $EE_1 = I$. Portanto, a matriz elementar E_1 é a inversa de E. $\square$

Como veremos, dada uma matriz invertível A, podemos realizar um número finito de operações elementares sobre as linhas de A até obter a matriz identidade. Isto significa que a matriz em escada na forma reduzida de uma matriz invertível é a matriz identidade. Para reduzir uma matriz invertível à matriz identidade por meio de operações elementares aplicamos o método conhecido por *método de eliminação de Gauss-Jordan*. Este método consiste na aplicação do método de eliminação de Gauss até obter uma matriz em escada com todos os pivôs iguais a 1, prosseguindo-se com a eliminação das entradas acima da diagonal principal usando os mesmos procedimentos do algoritmo de eliminação de Gauss, mas começando este algoritmo com a última coluna da matriz. Ilustramos no próximo exemplo os diversos passos do método de eliminação de Gauss- Jordan.

Exemplo 1.29. Vejamos como obter a matriz identidade por meio de operações elementares para a seguinte matriz invertível A.

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix} \underset{L_3+L_1}{\overset{L_2-2L_1}{\longrightarrow}} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 4 \end{bmatrix} \underset{1/2L_2}{\overset{1/4L_3}{\longrightarrow}} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & \frac{-3}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

$$\overset{L_2+3/2L_3}{\longrightarrow} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{L_1-2L_3}{\longrightarrow} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{L_1+L_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

As operações realizadas na primeira linha da expressão anterior correspondem ao método de eliminação de Gauss, e as operações na segunda linha eliminam as entradas acima da diagonal principal começando com a última coluna da matriz em escada por linhas já obtida.

Realizámos 7 operações elementares. A cada operação elementar corresponde uma matriz elementar. Designemos por E_i a matriz elementar correspondente à operação elementar número i[10]. Para as operações elementares efectuadas, as matrizes E_i são:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$$

$$E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix} \qquad E_6 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_7 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Da Proposição 1.8, concluimos

$$E_7E_6E_5E_4E_3E_2E_1A = I. \tag{1.14}$$

Como qualquer matriz elementar é invertível (Proposição 1.9), e o produto de matrizes invertíveis é invertível (Proposição 1.6), a matriz $(E_7E_6E_5E_4E_3E_2E_1)$ é invertível. Logo, multiplicando a igualdade (1.14) à esquerda por $(E_7E_6E_5E_4E_3E_2E_1)^{-1}$ obtemos

$$A = (E_7E_6E_5E_4E_3E_2E_1)^{-1}I = (E_7E_6E_5E_4E_3E_2E_1)^{-1}. \tag{1.15}$$

Assim, pelo primeiro item da Proposição 1.7, a matriz A é invertível e

$$A^{-1} = E_7E_6E_5E_4E_3E_2E_1 = \begin{bmatrix} \frac{-1}{8} & \frac{1}{2} & \frac{-1}{8} \\ \frac{-5}{8} & \frac{1}{2} & \frac{3}{8} \\ \frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix}.$$

Pode confirmar-se que a matriz anterior é de facto a inversa de A calculando o produto de A por essa matriz.

[10] As operações que aparecem sobrepostas na primeira linha da aplicação do método podem ser realizadas por qualquer ordem. Em particular, isso significa que as matrizes elementares correspondentes comutam.

A igualdade (1.15) é equivalente a

$$A = (E_7E_6E_5E_4E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1}E_6^{-1}E_7^{-1},$$

onde na última igualdade aplicámos (1.13). Como a inversa de uma matriz elementar é ainda uma matriz elementar, a expressão anterior dá a matriz invertível A como um produto de matrizes elementares. ♦

Algumas propriedades que vimos no exemplo anterior caracterizam as matrizes invertíveis. Esta afirmação é parte do conteúdo do teorema seguinte.

Teorema 1.4. Sendo A é uma matriz quadrada, são equivalentes as afirmações:

(i) A é invertível.

(ii) O sistema $A\mathbf{x} = \mathbf{b}$ tem solução única.

(iii) É possível aplicar operações elementares sobre as linhas de A e obter a matriz identidade.

(iv) A matriz A pode exprimir-se como um produto de matrizes elementares.

Demonstração. Provemos a seguinte sequência de implicações: $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$: Se A é invertível, o sistema $A\mathbf{x} = \mathbf{b}$ tem solução única uma vez que, multiplicando (à esquerda) a igualdade $A\mathbf{x} = \mathbf{b}$ por A^{-1} se obtém $\mathbf{x} = A^{-1}\mathbf{b}$.

$(ii) \Rightarrow (iii)$: Seja $\mathbf{x} = (c_1, \ldots, c_n)$ a única solução de $A\mathbf{x} = \mathbf{b}$. Assim, a matriz aumentada deste sistema pode ser reduzida (por meio de operações elementares sobre as suas linhas) à matriz

$$\left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_n \end{array}\right].$$

Isto é, existe uma sequência de operações elementares que reduzem A à matriz identidade.

$(iii) \Rightarrow (iv)$: Se é possível aplicar operações elementares sobre as linhas de A e obter a matriz identidade, então existe um número finito de matrizes elementares, sejam $E_1, \ldots, E_k$, tal que

$$E_k \cdots E_2E_1A = I.$$

Como qualquer matriz elementar é invertível (Proposição 1.9) e o produto de matrizes invertíveis é invertível (Proposição 1.6), podemos multiplicar a igualdade anterior, à esquerda, por $(E_k \cdots E_2 E_1)^{-1}$, obtendo-se

$$A = (E_k \cdots E_2 E_1)^{-1} I = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

Uma vez que a inversa de uma matriz elementar ainda é uma matriz elementar (Proposição 1.9), a igualdade anterior exprime A como um produto de matrizes elementares.

$(iv) \Rightarrow (i)$: Se A é igual ao produto de matrizes elementares, a matriz A é invertível visto que cada matriz elementar é invertível, e o produto de matrizes invertíveis ainda é invertível. $\square$

Método de eliminação de Gauss-Jordan e cálculo da inversa

O Teorema 1.4 diz-nos que uma matriz invertível pode reduzir-se à matriz identidade por meio de operações elementares sobre as suas linhas. Assim, se A é invertível existem matrizes elementares $E_1, \ldots, E_k$ tais que $E_k \cdots E_2 E_1 A = I$. Esta igualdade é equivalente a

$$A^{-1} = E_k \cdots E_2 E_1 I. \tag{1.16}$$

Como o produto EA é a matriz que se obtém de A efectuando a operação elementar que permitiu obter E da identidade, a igualdade (1.16) diz-nos que a inversa de A se pode obter realizando sobre a matriz identidade as mesmas operações elementares (e pela mesma ordem) que reduzem A à matriz identidade. Esquematicamente,

$$\begin{bmatrix} A & | & I \end{bmatrix} \overset{\text{Operações elementares}}{\longrightarrow} \begin{bmatrix} I & | & A^{-1} \end{bmatrix}.$$

Exemplo 1.30. Determinemos a inversa de

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

Coloquemos a matriz identidade de terceira ordem a par da matriz A, e efectuemos operações elementares sobre $[A \,|\, I]$ até obtermos a matriz $[I \,|\, A^{-1}]$.

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow[L_3+L_1]{L_2-L_1} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{1/2L_3} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right] \xrightarrow{L_2-L_3} \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{-3}{2} & 1 & \frac{-1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right]$$

$$\xrightarrow{L_1+L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-1}{2} & 1 & \frac{-1}{2} \\ 0 & 1 & 0 & \frac{-3}{2} & 1 & \frac{-1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right].$$

Logo, a matriz inversa de A é

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

♦

Algumas consequências da invertibilidade de uma matriz

Um sistema de equações lineares com igual número de equações e incógnitas possui matriz dos coeficientes quadrada, e portanto podemos falar da invertibilidade desta matriz. O Teorema 1.4 diz-nos que um tal sistema é possível e determinado se e só se a matriz dos coeficientes do sistema é invertível. Além disso, a demonstração desse teorema estabelece que a solução (única) se obtém multiplicando o vector correspondente ao termo independente do sistema pela inversa da matriz dos coeficientes do sistema. Para referência futura enunciemos este resultado.

Proposição 1.10. Seja $\mathbf{b}$ um vector qualquer. Um sistema $A\mathbf{x} = \mathbf{b}$ cuja matriz dos coeficientes é quadrada, tem solução única se e só se A é invertível. Além disso, sendo A invertível a solução do sistema $A\mathbf{x} = \mathbf{b}$ é

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Em particular, um sistema homogéneo $A\mathbf{x} = \mathbf{0}$ tem solução única $\mathbf{x} = \mathbf{0}$ se e só se a matriz A é invertível.

Demonstração. Pelo Teorema 1.4, um sistema $A\mathbf{x} = \mathbf{b}$ cuja matriz dos coeficientes é quadrada, tem solução única se e só se A é invertível. Se A^{-1} existe, podemos multiplicar $A\mathbf{x} = \mathbf{b}$, à esquerda, por A^{-1} e obtemos

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \Longleftrightarrow \mathbf{x} = A^{-1}\mathbf{b}.$$

Se $\mathbf{b} = \mathbf{0}$, como o produto de uma qualquer matriz pelo vector nulo é o vector nulo, a solução anterior é $\mathbf{x} = \mathbf{0}$. □

Nota 10. *Na demonstração anterior não faz sentido multiplicar $A\mathbf{x} = \mathbf{b}$ à direita por A^{-1}, já que tal multiplicação não está definida (o vector $\mathbf{b}$ é do tipo $n \times 1$ e A^{-1} é do tipo $n \times n$).*

Exemplo 1.31. Consideremos o sistema

$$\begin{cases} 2x_1 + x_2 = 6 \\ 4x_1 + 4x_2 = 8. \end{cases}$$

A matriz dos coeficientes do sistema é exactamente a matriz A do Exemplo 1.25, na página 50. Nesse exemplo calculámos a inversa de A, e portanto o sistema tem solução única dada por

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A^{-1} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 & \frac{-1}{4} \\ -1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}.$$

♦

De acordo com a definição de matriz invertível, uma matriz A é invertível se existir uma matriz B que satisfaz as condições:

$$AB = I \quad \text{e} \quad BA = I.$$

Na proposição seguinte mostramos que é suficiente que uma das condições se verifique para a outra ser automaticamente válida (e portanto A ser invertível).

Proposição 1.11. Seja A uma matriz quadrada.

a) Se B é uma matriz que satisfaz $BA = I$ então B é a inversa de A.

b) Se B é uma matriz que satisfaz $AB = I$ então B é a inversa de A.

Demonstração. Mostremos o item a). Comecemos por provar que se $BA = I$ então A é invertível. Para tal, basta mostar que a única solução do sistema homogéneo $A\mathbf{x} = \mathbf{0}$ é a solução nula (conforme Proposição 1.10). Seja $\mathbf{u}$ uma qualquer solução de $A\mathbf{x} = \mathbf{0}$. Multiplicando ambos os membros da igualdade $A\mathbf{u} = \mathbf{0}$ por B, obtém-se

$$BA\mathbf{u} = B\mathbf{0} \iff I\mathbf{u} = \mathbf{0} \iff \mathbf{u} = \mathbf{0}.$$

Como $\mathbf{u}$ é uma qualquer solução de $A\mathbf{x} = \mathbf{0}$, este sistema apenas admite a solução nula, e portanto a Proposição 1.10 garante que a matriz A é invertível.

Sendo A invertível e $BA = I$, multiplicando (à direita) ambos os membros desta igualdade por A^{-1}, tem-se

$$BAA^{-1} = A^{-1} \iff BI = A^{-1} \iff B = A^{-1}.$$

Para demonstrar o item b) usamos o item a) trocando os papéis de A e B. Assim, se $AB = I$ o item a) diz-nos que A é a inversa de B, e portanto

$$A = B^{-1} \iff A^{-1} = B.$$

□

A Proposição 1.6 diz que se A e B são matrizes (do mesmo tipo) invertíveis, então o produto AB é invertível. No teorema a seguir mostramos que é condição necessária e suficiente para a invertibilidade do produto AB que ambas as matrizes A e B sejam invertíveis.

Teorema 1.5. Sejam A e B matrizes quadradas da mesma ordem. A matriz produto AB é invertível se e só se as matrizes A e B são invertíveis. Além disso, a inversa do produto é dada por

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

Demonstração. Tendo em conta a Proposição 1.6, resta mostrar que se AB é uma matriz invertível, então as matrizes A e B são invertíveis.

Sendo a matriz AB invertível, por definição de inversa existe uma matriz D tal que

$$D(AB) = I \quad \text{e} \quad (AB)D = I.$$

Como o produto de matrizes é associativo, as igualdades anteriores são equivalentes a

$$(DA)B = I \quad \text{e} \quad A(BD) = I.$$

Usando, respectivamente, os itens a) e b) da Proposição 1.11, tem-se

$$(DA)B = I \Longrightarrow B \text{ é invertível}$$
$$A(BD) = I \Longrightarrow A \text{ é invertível.}$$

□

Finalizamos esta secção enunciando algumas características de matrizes reais invertíveis.

Proposição 1.12. Seja A uma matriz real do tipo $n \times n$. São equivalentes as afirmações seguintes.

a) A é invertível.

b) A tem característica n.

c) Os vectores coluna de A geram $\mathbb{R}^n$.

Demonstração. Esta prova será realizada mostrando as seguintes relações: $a) \iff b)$, $a) \Rightarrow c)$ e $c) \Rightarrow a)$.

$a) \iff b)$: Uma matriz é invertível se e só se a sua forma reduzida em escada por linhas é a identidade. Consequentemente, uma matriz do tipo $n \times n$ é invertível se e só se tem característica igual a n.

$a) \Rightarrow c)$: Como A é invertível, o sistema $A\mathbf{x} = \mathbf{b}$ é possível, qualquer que seja o vector $\mathbf{b}$ de $\mathbb{R}^n$. Assim, pela Proposição 1.3, qualquer vector de $\mathbb{R}^n$ é uma combinação linear dos vectores colunas de A. Isto significa que qualquer vector de $\mathbb{R}^n$ pertence ao conjunto gerado pelas colunas de A.

$c) \Rightarrow a)$: Como qualquer vector de $\mathbb{R}^n$ é uma combinação linear das colunas de A, os vectores

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$$

são combinações lineares das colunas de A. Isto significa, que existem vectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ tais que

$$A\mathbf{x}_1 = \mathbf{e}_1, \ A\mathbf{x}_2 = \mathbf{e}_2, \ \ldots, \ A\mathbf{x}_n = \mathbf{e}_n.$$

Seja B a matriz cujas colunas são os vectores $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, por esta ordem. Usando a Definição 1.12 de produto de matrizes, tem-se

$$\begin{aligned} AB &= A\begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x_2} & \ldots & \mathbf{x}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \ldots & A\mathbf{x}_n \\ | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \\ | & | & & | \end{bmatrix} = I_n. \end{aligned}$$

Da Proposição 1.11, segue que B é a inversa de A.

□

1.5 Matrizes triangulares

Matrizes ditas triangulares surgem nos mais diversos contextos. Em particular, a aplicação do método de eliminação de Gauss a uma matriz quadrada produz uma matriz em escada que é triangular superior.

Definição 1.17. Uma matriz quadrada com todas as entradas abaixo das entradas da diagonal principal iguais a zero é designada por matriz *triangular superior*.

Uma matriz quadrada com todas as entradas acima das entradas da diagonal principal iguais a zero diz-se *triangular inferior*.

Uma matriz A do tipo $n \times n$ que seja triangular superior e triangular inferior, isto é, tal que todas as entradas não nulas estão na diagonal principal, diz-se uma matriz *diagonal* e denota-se

$$A = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}).$$

Um exemplo de uma matriz diagonal é a matriz identidade $I_n = \operatorname{diag}(1, 1, \ldots, 1)$.

A definição de matriz triangular exprime-se em termos das entradas da matriz da forma que se segue.

Seja $A = [a_{ij}]_{i,j=1,\ldots,n}$.

i) A é triangular superior $\Longleftrightarrow a_{ij} = 0$ para todo $i > j$.

ii) A é triangular inferior $\Longleftrightarrow a_{ij} = 0$ para todo $i < j$.

iii) A é diagonal $\Longleftrightarrow a_{ij} = 0$ para todo $i \neq j$.

Exemplo 1.32. A matriz

$$\begin{bmatrix} 1 & 7 & 5 & 2 \\ 0 & 8 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

é triangular superior e

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

é triangular inferior. ♦

A matriz transposta de uma matriz triangular superior é uma matriz triangular inferior e vice-versa.

Uma matriz triangular é invertível se e só se a sua diagonal principal não tem entradas nulas. De facto, como a característica de uma matriz triangular é igual ao número de entradas não nulas da sua diagonal principal, pela Proposição 1.12 (página 62) uma matriz triangular é invertível se e só se todas as entradas da diagonal principal são distintas de zero. No Exemplo 1.32 a primeira matriz não é invertível enquanto que a segunda matriz é.

Nota 11. *Se uma matriz não é triangular podem existir zeros na diagonal principal e a matriz ser invertível. Por exemplo, a matriz*

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

é invertível (é uma matriz de permutação).

O algoritmo apresentado para o cálculo da inversa (método de eliminação de Gauss-Jordan) permite concluir que a inversa de uma matriz triangular inferior é triangular inferior, e a inversa de uma matriz triangular superior é triangular superior.

Deixamos como exercício mostrar que o produto de matrizes triangulares superiores (resp. inferiores) é ainda uma matriz triangular superior (resp. inferior). Para tal, basta usar a expressão (1.10) para o cálculo da entrada ij da matriz produto.

Resumimos no quadro seguinte os resultados que acabámos de referir.

Propriedades das matrizes triangulares

(a) A transposta de uma matriz triangular superior (resp. inferior) é triangular inferior (resp. superior).

(b) O produto de matrizes triangulares superiores (resp. inferiores) é triangular superior (resp. inferior).

(c) Uma matriz triangular é invertível se e só se as entradas da diagonal principal são todas diferentes de zero.

(d) A inversa de uma matriz triangular superior (resp. inferior) é uma matriz triangular superior (resp. inferior).

1.5.1 Factorização LU

Uma factorização de uma matriz A é uma igualdade que exprime A como um produto de uma ou mais matrizes. A factorização $A = LU$ exprime a matriz quadrada A como o produto de uma matriz triangular inferior L (do inglês “lower triangular”) por uma matriz triangular superior U (do inglês “upper triangular”). Quando as matrizes triangulares L e U gozam das propriedades adicionais que especificaremos a seguir, dizemos que $A = LU$ é uma factorização LU de A.

A factorização LU desempenha um papel fundamental na classificação de matrizes simétricas (estudo efectuado no Capítulo 7) e está na base de algoritmos vantajosos do ponto de vista computacional para a resolução de sistemas lineares. Por exemplo, suponha-se que é dada uma factorização $A = LU$, com L e U invertíveis, e que pretendemos resolver um sistema linear $A\mathbf{x} = \mathbf{b}$. A solução deste sistema pode ser determinada da seguinte forma:

- O sistema $A\mathbf{x} = \mathbf{b}$ é reescrito na forma
$$LU\mathbf{x} = \mathbf{b}. \tag{1.17}$$
- Defina-se o vector
$$\mathbf{y} = U\mathbf{x}. \tag{1.18}$$
- Usando (1.18), o sistema (1.17) é equivalente ao sistema triangular $L\mathbf{y} = \mathbf{b}$. Resolva-se este sistema para obter $\mathbf{y}$.
- Substituindo $\mathbf{y}$ em (1.18) e resolvendo o sistema triangular correspondente, obtém-se a solução do sistema inicial $A\mathbf{x} = \mathbf{b}$.

O processo acabado de descrever substitui o sistema $A\mathbf{x} = \mathbf{b}$ por dois sistemas ($L\mathbf{y} = \mathbf{b}$ e $U\mathbf{x} = \mathbf{y}$) cujas matrizes dos coeficientes são triangulares. Tal representa uma economia computacional considerável, em particular quando se pretende resolver vários sistemas com a mesma matriz dos coeficientes A e termos independentes distintos.

No exemplo seguinte ilustramos a aplicação deste método.

Exemplo 1.33. Consideremos a seguinte factorização LU de A

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

Podemos obter a solução do sistema

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

resolvendo os sistemas

$$L\mathbf{y} = \mathbf{b} \iff \begin{cases} y_1 = 3 \\ \frac{1}{2}y_1 + y_2 = 1 \end{cases} \quad \text{e} \quad U\mathbf{x} = \mathbf{y} \iff \begin{cases} 4x_1 + \phantom{\frac{1}{2}} x_2 = y_1 \\ \frac{1}{2}x_2 = y_2. \end{cases}$$

A solução do sistema $L\mathbf{y} = \mathbf{b}$ é $y_1 = 3$ e $y_2 = -\frac{1}{2}$. Substituindo o vector $\mathbf{y}$ obtido em $U\mathbf{x} = \mathbf{y}$, a solução do sistema é $x_1 = 1$ e $x_2 = -1$. ♦

Apresentamos a seguir o que entendemos por uma factorização LU.

Factorização LU

Diz-se que a matriz A (invertível) admite uma factorização LU se $A = LU$, em que

- L é uma matriz triangular inferior com todas as entradas da diagonal principal iguais a 1.
- U é uma matriz triangular superior com todas as entradas da diagonal principal diferentes de zero.

Vejamos como calcular uma factorização LU para uma matriz invertível A no caso em que é possível aplicar o método de eliminação de Gauss sem troca de linhas. Como A é quadrada, o método de eliminação de Gauss produz uma matriz em escada U que é triangular superior e que possui na diagonal principal os pivôs.

Como estamos supondo que não efectuamos troca de linhas, não podemos encontrar durante o processo nenhuma entrada nula na diagonal principal, caso contrário seria necessário usar troca de linhas para obter U. Logo, U tem entradas na diagonal principal diferentes de zero.

Operações elementares sobre as linhas de A traduzem-se na multiplicação sucessiva (à esquerda) por matrizes elementares. Assim, existe um número finito de matrizes elementares $E_1, E_2, \ldots, E_k$, tais que

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

Como não foram efectuadas troca de linhas, as matrizes $E_1, \ldots, E_k$ são triangulares inferiores com todas as entradas na diagonal principal iguais a 1. Logo, o produto $E_k \cdots E_2 E_1$ é uma matriz triangular inferior com 1's na diagonal principal. Por outro lado, como matrizes elementares são invertíveis, o seu produto é invertível, tendo-se portanto

$$E_k \cdots E_2 E_1 A = U \iff A = (E_k \cdots E_2 E_1)^{-1} U. \tag{1.19}$$

A matriz $(E_k \cdots E_2 E_1)^{-1}$ é a inversa de uma matriz triangular inferior pelo que também é triangular inferior. Podemos assim tomar para L a matriz

$$L = (E_k \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

A matriz L é igual a um produto de matrizes (elementares) triangulares inferiores com 1's na diagonal principal. Consequentemente, L tem todas as entradas na diagonal principal iguais a 1. Ou seja, a equação (1.19) exprime A na forma LU.

Acabámos de mostrar o resultado que enunciamos a seguir.

Teorema 1.6. Se a aplicação do método de eliminação de Gauss, sem troca de linhas, a uma matriz quadrada A produz uma matriz triangular superior U com todas as entradas na diagonal principal não nulas, então A admite a factorização $A = LU$, onde L é uma matriz triangular inferior com todas as entradas da diagonal principal iguais a 1.

Nota 12. *Para que uma matriz admita uma factorização LU é necessário que a matriz seja invertível, uma vez que L e U são invertíveis. A condição de invertibilidade da matriz não é contudo uma condição suficiente. No Exercício 7.13 (pág. 398) indica-se uma matriz invertível que não admite uma factorização LU.*

Refira-se ainda que a Proposição 7.9 na página 398, fornece uma condição necessária e suficiente para a existência de uma factorização LU.

Exemplo 1.34. Verifiquemos que a matriz

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

admite uma factorização LU e calculemos os respectivos factores L e U.

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 0 & 3 \\ 4 & 1 & 1 \end{bmatrix} \xrightarrow{L_2-3L_1} \begin{bmatrix} 2 & 2 & 2 \\ 0 & -6 & -3 \\ 4 & 1 & 1 \end{bmatrix} \quad \rightsquigarrow \quad E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 & 2 \\ 0 & -6 & -3 \\ 4 & 1 & 1 \end{bmatrix} \xrightarrow{L_3-2L_1} \begin{bmatrix} 2 & 2 & 2 \\ 0 & -6 & -3 \\ 0 & -3 & -3 \end{bmatrix} \quad \rightsquigarrow \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 2 & 2 & 2 \\ 0 & -6 & -3 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{L_3-1/2L_2} \begin{bmatrix} 2 & 2 & 2 \\ 0 & -6 & -3 \\ 0 & 0 & \frac{-3}{2} \end{bmatrix} = U \quad \rightsquigarrow \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{-1}{2} & 1 \end{bmatrix}$$

Logo, $E_3E_2E_1A = U$ e $L = (E_3E_2E_1)^{-1}$. Efectuando o produto $L = E_1^{-1}E_2^{-1}E_3^{-1}$, tem-se

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & \frac{1}{2} & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 2 & 2 & 2 \\ 0 & -6 & -3 \\ 0 & 0 & \frac{-3}{2} \end{bmatrix}.$$

♦

A matriz L obtida no exemplo anterior é uma matriz triangular inferior cujas entradas da diagonal principal são iguais a 1 e as entradas abaixo da diagonal principal são os simétricos dos valores (ou *multiplicadores*) usados no método de eliminação de Gauss para anular essa entrada. Esta propriedade da matriz L é verificada para qualquer factorização LU. De facto, atendendo à forma particular da matriz inversa de uma matriz elementar (que não seja de permutação), conclui-se facilmente que o produto $L = (E_k \cdots E_2E_1)^{-1}$ tem sempre a forma observada no exemplo anterior. Resumindo,

Se A admite uma factorização LU a matriz L é uma matriz triangular inferior com todas as entradas da diagonal principal iguais a 1, e cada entrada abaixo da diagonal principal é igual ao simétrico do escalar usado no método de eliminação de Gauss para anular essa entrada. Estes escalares designam-se por *multiplicadores*.

Deixamos como exercício mostrar que se uma matriz é factorizável na forma LU, a factorização é única.

Exercício 1.3. Mostre que a factorização $A = LU$ é única.

Sugestão: Supor que L_1U_1 e L_2U_2 são duas factorizações LU de A. Usar o facto dos factores L e U serem matrizes triangulares invertíveis para verificar que $L_2^{-1}L_1 = I = U_2U_1^{-1}$.

▲

Para finalizar esta secção, refira-se que se A é invertível é possível que durante o processo de eliminação de Gauss se encontre uma entrada nula na diagonal principal que pode ser removida por troca de linhas. Mostra-se neste caso que existe uma matriz P obtida da matriz identidade I efectuando uma sucessão de trocas de linhas tal que $PA = LU$. O procedimento para obter a matriz (de permutação) P sai do âmbito deste livro. Remetemos o leitor interessado nesta questão, ou na eficiência computacional dos algoritmos aqui estudados para a resolução de sistemas lineares, para um livro de álgebra linear numérica como por exemplo [12].

1.6 Partição de matrizes em blocos

Muitas vezes é conveniente encarar uma matriz como uma matriz cujas entradas são matrizes mais pequenas chamadas blocos. Matrizes em blocos aparecem em várias áreas de aplicação da álgebra linear, sendo que a partição em blocos está quase sempre relacionada com a estrutura do modelo em estudo, o qual é frequentemente traduzido através de um sistema linear. Neste texto, em particular nalgumas demonstrações, usamos operações com matrizes em blocos. Por isso, nesta secção fazemos uma breve referência a estas operações. Comecemos por dar um exemplo de uma partição em blocos de uma matriz do tipo 3×6. Seja

$$A = \left[\begin{array}{ccc|cc|c} 2 & 1 & 3 & 4 & 1 & 2 \\ 5 & 3 & 2 & 7 & 8 & 1 \\ \hline -1 & 3 & -2 & 4 & 5 & 1 \end{array}\right]. \tag{1.20}$$

A matriz A pode ser vista como uma matriz em blocos do tipo 2×3

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix},$$

onde os blocos são respectivamente as matrizes

$$A_{11} = \begin{bmatrix} 2 & 1 & 3 \\ 5 & 3 & 2 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 4 & 1 \\ 7 & 8 \end{bmatrix}, \qquad A_{13} = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} -1 & 3 & -2 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 4 & 5 \end{bmatrix}, \quad \text{e} \;\; A_{23} = \begin{bmatrix} 1 \end{bmatrix}.$$

Todas as operações com matrizes em blocos são definidas como se as submatrizes (blocos) fossem as entradas de uma matriz. Por exemplo, se A e B são do mesmo tipo e estão particionadas por forma a que os blocos correspondentes sejam do mesmo tipo, a soma $A+B$ é ainda uma matriz em blocos cujos blocos são a soma dos blocos de A com os blocos homólogos de B. A multiplicação de uma matriz em blocos por um escalar pode também ser efectuada bloco a bloco.

Para obter o produto AB, onde A e B são matrizes em blocos, é necessário que a partição de A por colunas coincida com a partição de B por linhas. Por exemplo, a partição da matriz A em (1.20) foi realizada usando 3, 2 e 1 colunas. Logo, o produto AB só está definido se a partição de B for realizada em 3, 2 e 1 linhas, como é o caso da matriz

$$B = \left[\begin{array}{cc} 1 & 7 \\ 2 & 3 \\ 4 & 2 \\ \hline 2 & 3 \\ 1 & 4 \\ \hline 8 & 1 \end{array}\right], \quad \text{isto é,} \quad B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}.$$

O produto AB é efectuado aplicando a fórmula utilizada para definir o produto de matrizes usando nessa fórmula os blocos em vez das entradas. Mais precisamente, para a matriz A dada em (1.20), tem-se

$$AB = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 + A_{13}B_3 \\ A_{21}B_1 + A_{22}B_2 + A_{23}B_3 \end{bmatrix},$$

com

$$A_{11}B_1 + A_{12}B_2 + A_{13}B_3 = \begin{bmatrix} 2 & 1 & 3 \\ 5 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 7 \\ 2 & 3 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 1 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 8 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 16 & 23 \\ 19 & 48 \end{bmatrix} + \begin{bmatrix} 9 & 16 \\ 22 & 53 \end{bmatrix} + \begin{bmatrix} 16 & 2 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 41 & 41 \\ 49 & 102 \end{bmatrix},$$

$$\begin{aligned} A_{21}B_1 + A_{22}B_2 + A_{23}B_3 &= \begin{bmatrix} -1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 1 & 7 \\ 2 & 3 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 5 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} 8 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -3 & -2 \end{bmatrix} + \begin{bmatrix} 13 & 32 \end{bmatrix} + \begin{bmatrix} 8 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 31 \end{bmatrix}. \end{aligned}$$

Assim, a matriz em blocos AB é

$$AB = \left[\begin{array}{cc} 41 & 41 \\ 49 & 102 \\ \hline 18 & 31 \end{array}\right].$$

Sugere-se que verifique que a matriz obtida coincide com a matriz produto AB que se obtém aplicando a fórmula (1.10) na página 41.

Uma matriz quadrada em blocos, da forma

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ \mathbf{0} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & A_{kr} \end{bmatrix},$$

onde os blocos A_{ii} são matrizes quadradas e $\mathbf{0}$ denota matrizes nulas, diz-se uma *matriz triangular superior por blocos*. De modo análogo se define *matriz triangular inferior por blocos* como uma matriz em que os blocos acima da diagonal são nulos. Uma matriz quadrada A diz-se uma *matriz diagonal por blocos* se os únicos blocos não nulos são os blocos (necessariamente quadrados) na diagonal.

Uma matriz triangular por blocos não é necessariamente uma matriz triangular, como se pode observar pelos exemplos que se seguem.

Exemplo 1.35. Exemplos de matrizes triangulares por blocos.

a) $\left[\begin{array}{cc|c} 2 & 1 & 0 \\ 3 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$ é uma matriz diagonal por blocos.

b) $B = \left[\begin{array}{cc|ccc} 2 & 1 & 5 & 0 & 2 \\ 3 & 2 & 3 & 0 & 3 \\ \hline 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 1 \end{array}\right]$ é uma matriz triangular superior por blocos.

c) $C = \left[\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ \hline 1 & 2 & 1 & 2 \end{array}\right]$ não é uma matriz triangular por blocos uma vez que o bloco C_{22} não é uma matriz quadrada (a matriz C também não é quadrada).

♦

Consideremos agora a matriz triangular por blocos $A = \begin{bmatrix} A_{11} & A_{12} \\ \mathbf{0} & A_{22} \end{bmatrix}$, em que as submatrizes A_{11} e A_{22} são quadradas respectivamente $p \times p$ e $q \times q$. Pretende-se saber em que condições esta matriz é invertível, e nesse caso, calcular a sua inversa.

Usando a definição de inversa, a matriz A^{-1} pode ser vista como uma matriz em blocos B tal que $AB = I$. Isto é,

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ \mathbf{0} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & \mathbf{0} \\ \mathbf{0} & I_q \end{bmatrix},$$

onde I_p e I_q são as matrizes identidade de ordens p e q respectivamente. Efectuando este produto, obtém-se

$$\begin{aligned} A_{11}B_{11} + A_{12}B_{21} &= I_p \\ A_{11}B_{12} + A_{12}B_{22} &= \mathbf{0} \\ A_{22}B_{21} &= \mathbf{0} \\ A_{22}B_{22} &= I_q. \end{aligned} \tag{1.21}$$

A (última) igualdade $A_{22}B_{22} = I_q$ é válida se e só se A_{22} for invertível (Proposição 1.11, pág. 60). Logo, $B_{22} = A_{22}^{-1}$. Multiplicando, à esquerda, a penúltima equação por A_{22}^{-1} obtém-se a matriz nula $B_{21} = \mathbf{0}$. Substituindo $B_{21} = \mathbf{0}$ na primeira equação vem $A_{11}B_{11} = I_p$. Esta equação é satisfeita se e só se A_{11} for invertível. Assim, se as matrizes A_{11} e A_{22} admitem inversa, a segunda equação de (1.21) é equivalente a

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \iff B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}.$$

Por conseguinte, A é invertível se e só se as matrizes A_{11} e A_{22} são invertíveis, e a inversa de A é

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ \mathbf{0} & A_{22}^{-1} \end{bmatrix}. \tag{1.22}$$

Não é difícil verificar que a prova realizada para uma matriz triangular com apenas 2 blocos na diagonal se generaliza a qualquer matriz triangular por blocos com um número superior de blocos. Concluindo,

Uma matriz triangular por blocos é invertível se e só se os blocos na diagonal (principal) são matrizes invertíveis.

Exercício 1.4. Calcular a inversa da matriz B do Exemplo 1.35 usando a expressão (1.22). Confirmar o resultado aplicando o método de eliminação de Gauss-Jordan a B.

♦