Escola Superior Náutica Infante D. Henrique

Álgebra Linear

Apontamentos de apoio à disciplina

Luís Cruz-Filipe & Patrícia Engrácia

Setembro de 2011

Conteúdo

Introdução

Estes apontamentos foram escritos para servir de apoio à disciplina de Álgebra Linear dos cursos de licenciatura da Escola Superior Náutica Infante D. Henrique.

Iniciamos estes apontamentos com um tema familiar e já conhecido dos alunos, introduzindo técnicas eficientes e fundamentais para os restantes temas abordados. Aproveitamos assim para rever alguns conceitos do Ensino Secundário, relevantes para a disciplina em questão e de extrema utilidade para outras disciplinas.

A exposição dos conteúdos é acompanhada de explicações intuitivas e muitos exemplos, de forma a tornar o texto acessível. Houve ainda o cuidado de colocar a seguir a cada conjunto de exemplos alguns exercícios simples de aplicação directa dos assuntos estudados, permitindo uma avaliação imediata dos conhecimentos adquiridos.

Paço d'Arcos, Setembro de 2010

Luís Cruz-Filipe e Patrícia Engrácia

Capítulo 1

Matrizes e aritmética matricial

Sem dúvida a ferramenta mais útil da Álgebra Linear é o cálculo matricial. Este capítulo apresenta o conceito de matriz e um conjunto de operações fundamentais sobre estes objectos, que serão da maior utilidade para os outros temas discutidos em capítulos posteriores.

Em vez de começarmos a tratar de matrizes em abstracto, vamos começar por estudar um problema concreto em que este conceito surge naturalmente: a resolução de sistemas de equações lineares. Teremos oportunidade de ver que o uso de matrizes simplifica substancialmente a organização da informação e o próprio processo de resolução dos sistemas; em contrapartida, as operações sobre matrizes que surgem neste contexto revelar-se-ão de extrema importância em aplicações completamente distintas.

1.1 Resolução de sistemas de equações lineares

1.1.1 Equações lineares

Uma *equação* é uma relação de igualdade envolvendo variáveis de valor desconhecido (incógnitas). *Resolver* uma equação é encontrar valores dessas incógnitas que tornam essa igualdade verdadeira. Por exemplo: $x = 2$ é uma solução de $x^2 - 2x = 0$, pois se nessa equação substituirmos x por 2 obtemos a expressão $2^2 - 2 \times 2 = 0$, que é verdadeira.

Duas equações dizem-se *equivalentes* se admitem o mesmo conjunto de soluções. As técnicas de resolução de equações passam sempre por ir transformando as equações noutras equivalentes até chegar a expressões que indicam directamente o valor das incógnitas (por exemplo, $x = 2$).

Uma *equação linear* é uma equação em que as incógnitas só aparecem multiplicadas por números (constantes) ou em somas. Assim, as equações

$$2x = 0 \qquad 3y + 2 = -3 \qquad 2x + 3z = 4w - 5 \qquad 5z + 2 = -1$$

são equações lineares, enquanto que

$$x^2 + 3x = 0 \qquad \sin(x + y) = 1 \qquad e^x = 0 \qquad \sqrt{x} = 3z + 2$$

não o são.

As equações lineares são extremamente simples de resolver, já que é preciso conhecer apenas em dois princípios fundamentais.

1. Somar quantidades iguais a ambos os membros duma equação produz uma equação equivalente à original.

2. Multiplicar ambos os membros duma equação pelo mesmo valor (diferente de 0) produz uma equação equivalente à original.

Vejamos como podemos aplicar estas técnicas para resolver equações lineares. Considere-se a equação

$$3x + 6 = -3\,.$$

O primeiro passo é isolar os termos em x dum dos lados da equação e os termos independentes do outro. Para tal, vamos somar -6 a ambos os membros e simplificar, obtendo uma equação equivalente.

$$3x + 6 = -3 \iff 3x + 6 - \mathbf{6} = -3 - \mathbf{6} \iff 3x = -9$$

Esta operação é vulgarmente designada por "passar o 6 para o outro lado trocando-lhe o sinal". De facto, ao somarmos -6 a ambos os membros da equação, o termo independente desapareceu do primeiro membro, surgindo o seu simétrico do lado direito. Esquematicamente, podemos representar esta operação da seguinte forma.

$$3\mathrm{x}+6 \qquad = \qquad -3\mathbf{-6}$$

O próximo passo é eliminar o coeficiente do x. Para tal, vamos dividir ambos os lados da equação por 3.

$$3x = -9 \iff \frac{3x}{\mathbf{3}} = \frac{-9}{\mathbf{3}} \iff x = -3$$

Esta operação é vulgarmente designada por "passar o 3 para o outro lado a dividir", podendo representar-se da forma seguinte.

$$3\,\mathrm{x} \qquad = \qquad \frac{-9}{\mathbf{3}}$$

Exemplo. Para resolver $2x + 3 = -1$, aplicamos o mesmo raciocínio:

$$2x + 3 = -1 \iff 2x + 3 - \mathbf{3} = -1 - \mathbf{3} \iff 2x = -4 \iff \frac{2x}{\mathbf{2}} = \frac{-4}{\mathbf{2}} \iff x = -2$$

obtendo-se a solução $x = -2$. Normalmente não se indicam os passos intermédios, pelo que esta resolução se apresentaria como

$$2x + 3 = -1 \iff 2x = -4 \iff x = -2\,.$$

Nos casos em que a divisão não é exacta, podem surgir fracções. Convém por isso ter presentes as regras de multiplicação e divisão por números fraccionários.

Exemplo. Para resolver $3x - 1 = 1$ procedemos como atrás, obtendo

$$3x - 1 = 1 \iff 3x = 2 \iff x = \frac{2}{3}\,.$$

Exercício 1. Resolva as seguintes equações lineares.

(a) $3x + 2 = -1$	(c) $-3x + 4 = 0$	(e) $-x + 1 = 5$	(g) $-4x - 3 = 1$
(b) $2x + 1 = 2$	(d) $x - 2 = -4$	(f) $2x - 2 = -2$	(h) $3x - 1 = -1$

Quando ocorrem várias variáveis, não encontramos uma solução mas uma relação entre as diferentes variáveis. Por exemplo, a equação $2x + 3y = 5$ pode ser simplificada por forma a exprimir x em função de y (resolver em ordem a x) ou y em função de x (resolver em ordem a y).

Resolvendo esta equação em ordem a x, obtemos

$$2x + 3y = 5 \Longleftrightarrow 2x = 5 - 3y \Longleftrightarrow x = \frac{5-3y}{2}\,.$$

Isto significa que para cada valor que atribuirmos a y somos capazes de encontrar um valor de x que torna a equação válida — precisamente o valor $\frac{5-3y}{2}$. Por exemplo, para $y = 1$ obtemos $x = \frac{5-3\times 1}{2} = 1$; se substituirmos na equação original x por 1 e y por 1, obtemos de facto $2 \times 1 + 3 \times 1 = 5$, que é uma equação verdadeira.

Mas podemos escolher outros valores para y. Tomando $y = 0$, obtemos $x = \frac{5-3\times 0}{2} = \frac{5}{2}$, que é outra solução da equação:

$$2 \times \frac{5}{2} + 3 \times 0 = 5\,.$$

Para $y = 2$ obtemos $x = -\frac{1}{2}$, e assim sucessivamente.

Resolvendo a mesma equação em ordem a y, obtemos

$$2x + 3y = 5 \Longleftrightarrow 3y = 5 - 2x \Longleftrightarrow y = \frac{5-2x}{3}\,.$$

Podemos agora escolher valores para x e determinar o valor de y correspondente que torna a equação verdadeira. Por exemplo, para $x = 0$ obtemos $y = \frac{5}{3}$ (e tem-se de facto $2\times 0+3\times\frac{5}{3} = 5$), enquanto para $x = 4$ se obtém $y = -1$ (e pode-se verificar que $2 \times 4 + 3 \times (-1) = 5$). Já tomando $x = 1$ obtemos novamente $y = 1$, o que era de esperar, pois todas estas equações são equivalentes.

Se fizermos um gráfico com as soluções desta equação, marcando os pontos (x, y) que satisfazem a condição prescrita pela equação, obtemos uma linha recta (ver Figura 1.1, à esquerda); é por este motivo que estas equações são chamadas equações lineares.

Quando o número de variáveis aumenta, o tipo de conjunto de soluções ganha mais dimensões. Por exemplo, consideremos a equação $x + 2y + 3z = 5$; para obter um valor para x é necessário fixar não apenas y, mas também z. Graficamente, o conjunto das soluções é agora um plano em $\mathbb{R}^3$ (mesma figura, à direita).

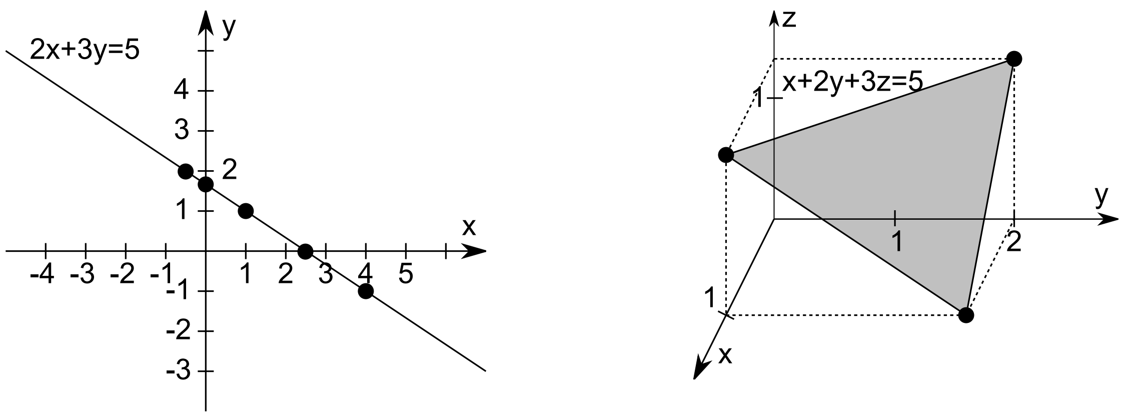


Figura 1.1: Conjunto de soluções de equações lineares a duas e três variáveis.

Exercício 2. Resolva as seguintes equações lineares em ordem a todas as variáveis que nelas ocorrem.

(a) $3x + 2y = -1$

(b) $2z + y - 1 = 2$

(c) $-3y + 4x - 2 = 0$

(d) $x - 2 = -4t$

(e) $-y + 2z + w = 5x - 4$

(f) $2t - 2 = -2x$

(g) $-4w - 3 = -4x - 3$

(h) $3x - 1 = y - 1$

(i) $x + y + z = 0$

Definição. Uma equação diz-se uma *equação linear* nas n variáveis $x_1, x_2, \ldots, x_n$ se puder ser escrita na forma

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

onde $a_1, a_2, \ldots, a_n, b$ são números reais não simultaneamente nulos.

Uma equação linear diz-se *homogénea* quando $b = 0$, ou seja, quando tem a forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

com $a_1, a_2, \ldots, a_n$ números reais.

Exercício 3. Quais das seguintes equações são lineares? Quais são homogéneas?

(a) $x - 3y = 2$

(b) $x^2 - 3y = 2$

(c) $5x + xy - z = 0$

(d) $2\sqrt{\pi}x = \frac{3}{4}y$

(e) $x + 7^2y = 1$

(f) $(x + y)^2 = z^2$

(g) $xyz = 1$

(h) $7x - 5y + 3 = 3$

As equações homogéneas vão ter uma importância especial em toda a Álgebra Linear. Nesta fase, podemos fazer desde já uma observação: uma equação homogénea admite sempre a solução $x_1 = x_2 = \cdots = x_n = 0$ em que todas as incógnitas têm o valor 0.

Exemplo. A equação $x + 2y = 0$ é uma equação homogénea. Podemos verificar que $x = y = 0$ é de facto uma solução: $0 + 2 \times 0 = 0$.

Observe-se que uma equação pode ser homogénea sem esse facto ser imediato: a equação $5(x+1) - 2(y+3) = -1$ não parece ser homogénea — pode nem ser óbvio que é linear — mas simplificando obtém-se

$$\begin{aligned} 5(x+1) - 2(y+3) = -1 &\iff 5x + 5 - 2y - 6 = -1 \iff 5x - 2y - 1 = -1 \\ &\iff 5x - 2y = 0\,, \end{aligned}$$

que é claramente uma equação homogénea. Outra forma de verificar seria substituir x e y por 0 na equação original e verificar que se obtém uma identidade válida, neste caso $5 \times (0+1) - 2 \times (0+3) = -1$.

O conjunto das soluções duma equação homogénea a duas variáveis é portanto uma recta que passa pela origem (ver Figura 1.2). A três variáveis obtemos um plano que passa pela origem (mesma figura, à direita).

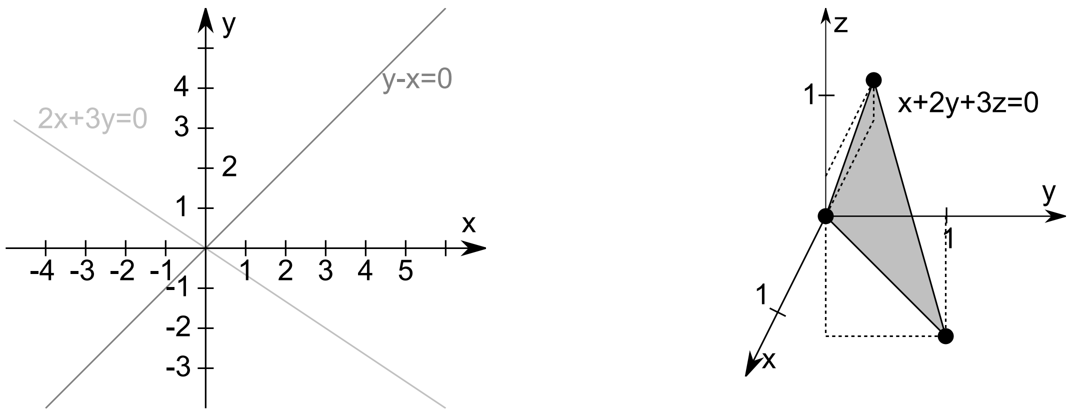


Figura 1.2: Conjunto de soluções de equações lineares homogéneas a duas e três variáveis.

Exercício 4. Quais das equações dos Exercícios 1 e 2 eram homogéneas? Verifique que de facto todas elas admitem uma solução com todas as incógnitas iguais a 0.

1.1.2 Sistemas de equações lineares

Muito frequentemente, as incógnitas estão relacionadas de várias formas diferentes, obtendo-se um conjunto de equações que queremos satisfazer *em simultâneo*. Um tal conjunto de equações chama-se um *sistema de equações*.

Definição. Um *sistema de equações lineares* é um conjunto de equações lineares em n variáveis $x_1, x_2, \ldots, x_n$, que se escreve habitualmente como

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

onde os números a_{ij} são reais e não simultaneamente zero e os números b_i são também reais.

Se todas as equações do sistema forem homogéneas (ou seja, se $b_1 = b_2 = \cdots = b_m = 0$), o sistema de equações lineares diz-se *homogéneo*.

Uma *solução particular* do sistema é um conjunto de números reais $(s_1, s_2, \ldots, s_n)$ que satisfazem simultaneamente todas as equações do sistema. A *solução geral* ou *conjunto solução* é o conjunto de todas as soluções particulares.

Exemplo.

1. Considere-se o sistema seguinte.

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

Resolvendo a segunda equação, obtemos

$$2x = -2 \iff x = -1\,;$$

substituindo este valor na primeira equação (uma vez que as queremos satisfazer em simultâneo), obtemos

$$3 \times (-1) + 2y = 1 \iff -3 + 2y = 1 \iff 2y = 4 \iff y = 2\,.$$

A solução geral do sistema é então $(x = -1, y = 2)$, que é também uma solução particular (a única).

2. Tomando agora o sistema

$$\begin{cases} 3x + 2y - 4z = 1 \\ 2x = -2 \end{cases}$$

podemos começar a resolvê-lo como atrás obtendo $x = -1$. Substituindo na primeira equação e resolvendo em ordem a y, obtemos

$$3 \times (-1) + 2y - 4z = 1 \iff -3 + 2y - 4z = 1 \iff 2y = 4 + 4z \iff y = 2 + 2z\,.$$

A solução geral do sistema é agora $(x = -1, y = 2 + 2z)$ (com z arbitrário). Soluções particulares são $(x = -1, y = 2, z = 0)$, obtida tomando $z = 0$; $(x = -1, y = 4, z = 1)$, obtida com $z = 1$; ou $(x = -1, y = 0, z = -1)$, obtida fazendo $z = -1$.

Normalmente convenciona-se apresentar as soluções por uma ordem fixa das variáveis (a ordem em que ocorrem no sistema). Assim, poderíamos dizer que este sistema tem solução geral $(-1, 2+t, t)$, onde escolhemos um novo parâmetro t para representar o valor arbitrário da variável z, e soluções particulares $(-1, 2, 0)$, $(-1, 4, 2)$ e $(-1, 0, -1)$, entre outras.

Exercício 5. Encontre soluções dos seguintes sistemas de equações lineares.

(a) $\begin{cases} 2x + 3y = 5 \\ x - 3y = -2 \end{cases}$ (b) $\begin{cases} 2x + y = 0 \\ 2x - y = 4 \end{cases}$ (c) $\begin{cases} x - 2y = 0 \\ x + 2y = 4 \end{cases}$ (d) $\begin{cases} 2x + 3y = 6 \\ 3x + 2y = 9 \end{cases}$

Definição. Um sistema de equações lineares diz-se:

- *possível determinado* se tem solução e essa solução é única;
- *possível indeterminado* se o sistema tem várias soluções diferentes;
- *impossível* se não tem solução.

Vimos no exemplo anterior que o sistema

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

é possível determinado, enquanto

$$\begin{cases} 3x + 2y - 4z = 1 \\ 2x = -2 \end{cases}$$

é possível indeterminado. Um exemplo de um sistema impossível seria o seguinte.

$$\begin{cases} x + y = 2 \\ x + y = 3 \end{cases}$$

De facto, independentemente dos valores atribuídos a x e a y, a sua soma ou é 2 ou é 3 (ou é outro valor qualquer), mas nunca poderá ser simultaneamente 2 e 3. Aquelas equações são por isso ditas *incompatíveis* e o sistema por elas formado é impossível.

Num sistema indeterminado, há incógnitas cujo valor pode ser escolhido arbitrariamente e outras que ficam expressas em função dessas. As incógnitas cujo valor pode ser escolhido dizem-se *variáveis livres* ou *independentes*; as incógnitas que são determinadas em função das variáveis livres dizem-se *variáveis dependentes*.

Note-se que o conceito de variável livre e dependente depende da forma de resolução da equação: no exemplo que demos atrás da equação $2x + 3y = 5$, quando a pomos na forma

$$x = \frac{5 - 3y}{2}$$

a variável y é livre e x dependente; mas se a resolvermos como

$$y = \frac{5 - 2x}{3}$$

é a variável x que é livre e a variável y que é dependente.

Uma outra situação que ocorre na prática é um sistema de equações lineares depender de parâmetros. Nesse caso, a sua caracterização dependerá do valor dos parâmetros. Por exemplo, o sistema

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

depende dos parâmetros α e β. Quando escolhemos os valores 0 e 1 para α e β, respectivamente, obtemos o sistema

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

que é claramente impossível. Para $\alpha = \beta = 0$, temos o sistema

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

com apenas duas equações para três incógnitas, que é possível indeterminado. Para os valores $\alpha = 1$ e $\beta = 2$, obtemos

$$\begin{cases} 2x + y + z = 2 \\ x + y = 2 \\ z = 2 \end{cases},$$

que é possível determinado:

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

Nas próximas secções, veremos como caracterizar de forma sistemática um sistema que dependa de parâmetros.

Exercício 6. Determine a solução dos seguintes sistemas de equações lineares em função dos parâmetros a, b e c e classifique-os em função destes parâmetros.

(a) $\begin{cases} x+y=a \\ x-y=b \end{cases}$ (b) $\begin{cases} 2x+y=a \\ 3x+6y=b \end{cases}$ (c) $\begin{cases} x+y+z=a \\ 2x+2z=b \\ 3y+3z=c \end{cases}$

Finalmente, observemos que, da mesma forma que uma equação homogénea tem sempre uma solução com todas as variáveis nulas, também um sistema homogéneo tem sempre pelo menos essa solução. Logo, um sistema de equações lineares homogéneo é possível determinado ou possível indeterminado, mas nunca impossível.

Geometricamente, podemos interpretar estas três situações tendo em conta o que atrás foi dito sobre os conjuntos de soluções. Para simplificar, consideremos um sistema de duas equações lineares envolvendo duas incógnitas x e y. Sabemos que o conjunto de soluções de cada equação é uma recta no plano; em geral, estas rectas intersectam-se num ponto (a solução), caso em que o sistema é possível determinado (Figura 1.3 (a)); mas pode suceder que as rectas sejam paralelas e o sistema se torne impossível (Figura 1.3 (b)); a terceira alternativa corresponde a ambas as equações descreverem a mesma recta, dando um sistema indeterminado (Figura 1.3 (c)).

Se o sistema for homogéneo, a segunda situação é impossível uma vez que ambas as rectas passam necessariamente pela origem (Figura 1.4).

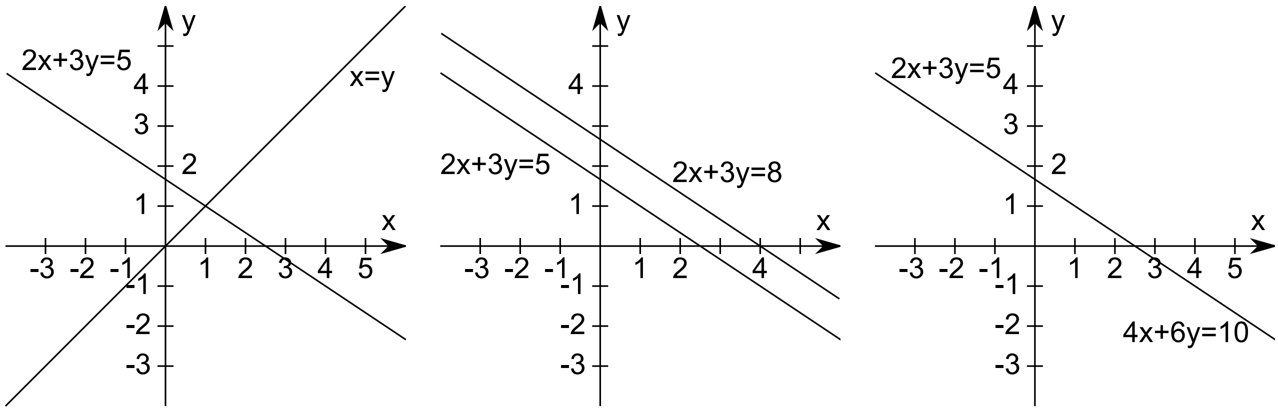
Com três equações a duas variáveis, podemos ter os mesmos três tipos de sistema, embora haja situações novas (Figura 1.5).

Em dimensões superiores, os comportamentos possíveis são generalizações destes.

Exercício 7. Para cada um dos seguintes sistemas de equações lineares a duas incógnitas, trace as rectas correspondentes aos conjuntos de soluções de cada equação e classifique o sistema como possível determinado, possível indeterminado ou impossível.

(a) $\begin{cases} -x+3y=-1 \\ 2x+y=2 \end{cases}$ (c) $\begin{cases} 2x+4y=2 \\ -x-2y=-1 \end{cases}$ (e) $\begin{cases} -2x+y=0 \\ -3x+2y=0 \end{cases}$

(b) $\begin{cases} 3x+2y=-1 \\ x+y=2 \end{cases}$ (d) $\begin{cases} x+y=1 \\ 2x+2y=-1 \end{cases}$ (f) $\begin{cases} 2x-y=1 \\ 4x-2y=0 \end{cases}$



(a) (b) (c)

Figura 1.3: Possíveis comportamentos dum sistema de duas equações lineares a duas incógnitas.

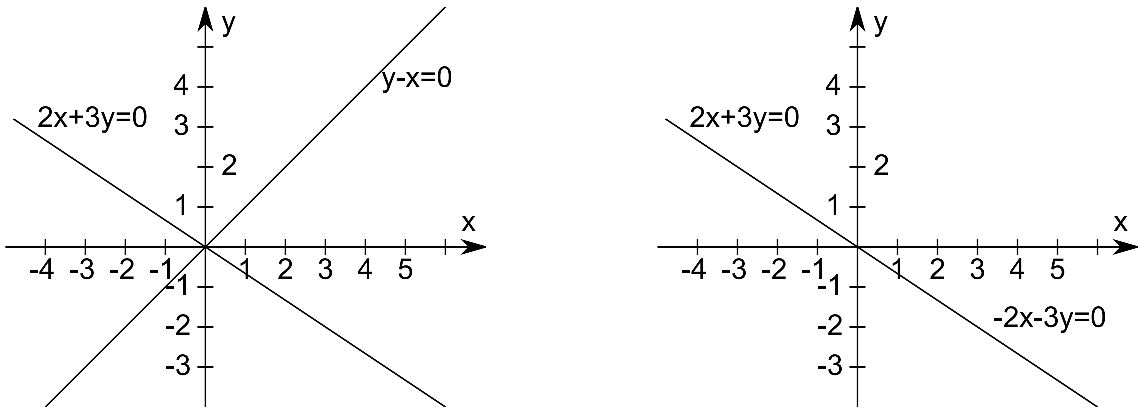


Figura 1.4: Possíveis comportamentos dum sistema de duas equações lineares homogéneo a duas incógnitas.

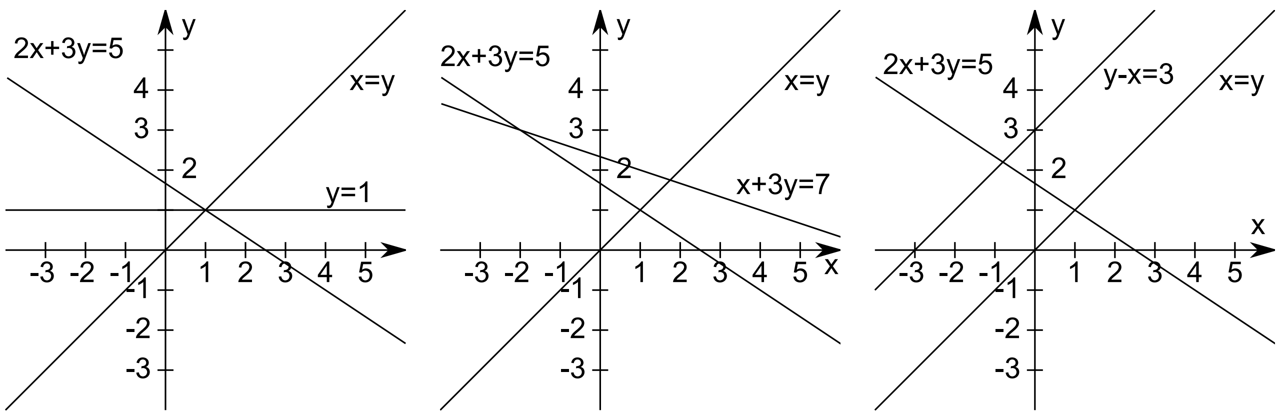


Figura 1.5: Alguns possíveis comportamentos dum sistema de três equações lineares a duas incógnitas.

1.1.3 Método de eliminação de Gauss

O método gráfico que discutimos na secção anterior não é muito prático para resolver sistemas de equações lineares em geral, nem se aplica facilmente em casos em que haja mais de duas variáveis. Nesta secção vamos estudar um método sistemático para resolver sistemas de equações lineares: o método de eliminação de Gauss. Este método, para além de ser simples de usar, vai ser de grande utilidade em muitos outros contextos que estudaremos mais adiante.

Tal como o método que apresentámos para resolver equações lineares, o método de eliminação de Gauss baseia-se em duas observações muito simples.

- Multiplicar ambos os membros duma equação por uma constante produz uma equação equivalente à original.
- Substituir uma equação pela sua soma com um múltiplo de outra produz um sistema de equações equivalente ao original.

Esta segunda observação é a novidade relativamente ao caso anterior. Consideremos o seguinte sistema muito simples de equações lineares.

$$\begin{cases} x+y=2 \\ x-y=-4 \end{cases}$$

Se somarmos as duas equações membro a membro, obtemos a nova equação $2x=-2$, donde se retira imediatamente o valor de x. Por outro lado, se tivessemos somado a primeira com (-1) vezes a segunda, obteríamos

$$(x+y)+(-1)(x-y)=2+(-1)(-4) \Longleftrightarrow 2y=6\,,$$

donde se retira o valor de y. Observe-se que os valores $x=-1$ e $y=3$ são de facto a solução do sistema original!

Definição. As *operações elementares* sobre um sistema de equações lineares são as seguintes.

- Troca de linhas: trocar as equações nas linhas i e j, denotada por $l_i \leftrightarrow l_j$.
- Multiplicação de uma linha por uma constante não nula: substituir a equação na linha i pelo seu produto por $\alpha \neq 0$, denotada por $l_i \to \alpha l_i$ ou simplesmente por αl_i.
- Soma duma equação com um múltiplo de outra: substituir a equação na linha i pela sua soma com α vezes a equação na linha j, denotada por $l_i \to l_i + \alpha l_j$ ou simplesmente por $l_i + \alpha l_j$.

Por vezes usaremos uma operação que é combinação das duas últimas, nomeadamente a substituição da linha i pela combinação $\alpha l_i + \beta l_j$.

Definição. Um coeficiente $c_{ij} \neq 0$ da variável x_j na equação na linha i diz-se um *pivot* se satisfizer as duas condições seguintes.

1. Os coeficientes de $x_1, x_2, \ldots, x_{j-1}$ na equação da linha i são todos nulos.
2. Para cada linha acima da linha i, os coeficientes de $x_1, x_2, \ldots, x_{j-1}$ não são todos nulos.

Por exemplo, no sistema

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

os coeficientes assinalados a negrito são pivots. O primeiro coeficiente da primeira equação satisfaz trivialmente a definição de pivot (não há equações antes nem coeficientes antes na mesma equação). O 2 da segunda linha também é um pivot, já que o coeficiente de x na segunda equação é 0 e o coeficiente de x na primeira equação não é 0.

Definição. O *método de eliminação de Gauss* consiste na aplicação sucessiva de operações elementares a um sistema de equações por forma a eliminar todos os coeficientes por baixo dos pivots.

Antes de vermos exemplos, convém fazer algumas observações que serão relevantes.

1. Por definição, os pivots não podem ser zero. Se não existir nenhum pivot por tratar, aplica-se uma troca de linhas por forma a obter um pivot na primeira linha que ainda não foi processada.

2. A eliminação faz-se sempre da esquerda para a direita: só depois de eliminar todos os termos da primeira variável abaixo do pivot é que se passa para a segunda variável, e assim sucessivamente.

3. Depois de se ter eliminado a coluna abaixo de um dado pivot, não é necessário voltar a alterar a equação que contém o pivot. Assim, a eliminação de Gauss também anda sempre de cima para baixo.

Vamos aplicar o método de eliminação de Gauss para resolver o seguinte sistema de equações lineares.

$$\begin{cases} x + y + 2z = 9 \\ 2x + 4y - 3z = 1 \\ 3x + 6y - 5z = 0 \end{cases}$$

Em cada passo, vamos eliminar as parcelas abaixo das parcelas a negrito. Indicamos a seguir a cada equação qual a operação elementar que lhe vamos aplicar.

$$\begin{cases} \mathbf{x} + y + 2z = 9 \\ 2x + 4y - 3z = 1 & (l_2 - 2l_1) \\ 3x + 6y - 5z = 0 & (l_3 - 3l_1) \end{cases} \longrightarrow \begin{cases} \mathbf{x} + y + 2z = 9 \\ \qquad \mathbf{2y} - 7z = -17 \\ \qquad 3y - 11z = -27 & (2l_3 - 3l_2) \end{cases}$$

$$\longrightarrow \begin{cases} \mathbf{x} + y + 2z = 9 \\ \qquad -\mathbf{2y} + 7z = 17 \\ \qquad\qquad -\mathbf{z} = -3 \end{cases} \longrightarrow \begin{cases} x + y + 2z = 9 \\ -2y + 7z = 17 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x + y + 6 = 9 \\ -2y + 21 = 17 \\ z = 3 \end{cases}$$

$$\longrightarrow \begin{cases} x + y = 3 \\ y = 2 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x + 2 = 3 \\ y = 2 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x = 1 \\ y = 2 \\ z = 3 \end{cases}$$

Observe-se que depois de o método de eliminação de Gauss terminar, obtemos um sistema de equações que se pode resolver simplesmente substituindo os valores das incógnitas de baixo para cima.

Este sistema tem uma única solução, logo é um sistema possível determinado.

Exemplo.

1. Vamos usar o método de eliminação de Gauss para resolver o seguinte sistema.

$$\begin{cases} - 2y + 3z = 1 \\ 3x + 6y - 3z = 18 \\ 6x + 6y = 30 \end{cases}$$

Uma vez que a primeira linha não tem pivot (não aparece nenhuma parcela contendo a variável x), vamos começar por fazer uma troca de linhas.

$$\begin{cases} \mathbf{3x} + 6y - 3z = 18 \\ - 2y + 3z = 1 \\ 6x + 6y = 30 \end{cases} \quad (l_3 - 2l_1) \longrightarrow \begin{cases} \mathbf{3x} + 6y - 3z = 18 \\ \mathbf{-2y} + 3z = 1 \\ - 6y + 6z = -6 \end{cases} \quad (l_3 - 3l_2)$$

$$\longrightarrow \begin{cases} \mathbf{3x} + 6y - 3z = 18 \\ \mathbf{-2y} + 3z = 1 \\ \mathbf{-3z} = -9 \end{cases} \longrightarrow \begin{cases} 3x + 6y - 3z = 18 \\ -2y + 3z = 1 \\ z = 3 \end{cases} \longrightarrow \begin{cases} 3x + 6y - 9 = 18 \\ -2y + 9 = 1 \\ z = 3 \end{cases}$$

$$\longrightarrow \begin{cases} 3x + 6y = 27 \\ y = 4 \\ z = 3 \end{cases} \longrightarrow \begin{cases} 3x + 24 = 27 \\ y = 4 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x = 1 \\ y = 4 \\ z = 3 \end{cases}$$

Novamente este sistema tem solução única, logo é possível determinado.

2. Vamos agora resolver o sistema seguinte.

$$\begin{cases} -2y + 3z = 1 \\ 3x + 6y - 3z = -2 \\ 6x + 6y + 3z = 5 \end{cases}$$

Começamos novamente por trocar linhas, pois a primeira equação não tem parcela correspondente à variável x.

$$\begin{cases} \mathbf{3x} + 6y - 3z = -2 \\ - 2y + 3z = 1 \\ 6x + 6y + 3z = 5 \end{cases} \quad (l_3 - 2l_1) \longrightarrow \begin{cases} \mathbf{3x} + 6y - 3z = -2 \\ \mathbf{-2y} + 3z = 1 \\ - 6y + 9z = 9 \end{cases} \quad (l_3 - 3l_2)$$

$$\longrightarrow \begin{cases} \mathbf{3x} + 6y - 3z = -2 \\ \mathbf{-2y} + 3z = 1 \\ 0 = 6 \end{cases}$$

A última equação deste sistema é impossível. Uma vez que todas as operações elementares produzem sistemas equivalentes, concluímos que o sistema original não tem soluções, sendo portanto um sistema impossível.

3. Vejamos agora outro sistema de equações.

$$\begin{cases} \mathbf{2x}+2y+2z=0 \\ -2x+5y+2z=1 & (l_2+l_1) \\ 8x+y+4z=-1 & (l_3-4l_1) \end{cases} \longrightarrow \begin{cases} \mathbf{2x}+2y+2z=0 \\ \quad \mathbf{7y}+4z=1 \\ \quad -7y-4z=-1 & (l_3+l_2) \end{cases}$$

$$\longrightarrow \begin{cases} \mathbf{2x}+2y+2z=0 \\ \quad \mathbf{7y}+4z=1 \\ \quad 0=0 \end{cases} \longrightarrow \begin{cases} 2x+2y+2z=0 \\ 7y+4z=1 \end{cases} \longrightarrow \begin{cases} 2x+2y+2z=0 \\ y=\frac{1-4z}{7} \end{cases}$$

$$\longrightarrow \begin{cases} 2x+2\frac{1-4z}{7}+2z=0 \\ y=\frac{1-4z}{7} \end{cases} \longrightarrow \begin{cases} 14x+6z=-2 \\ y=\frac{1-4z}{7} \end{cases} \longrightarrow \begin{cases} x=\frac{-3z-1}{7} \\ y=\frac{1-4z}{7} \end{cases}$$

O sistema tem infinitas soluções (uma para cada valor que se escolha para a variável livre z), logo é um sistema possível indeterminado.

A sua solução geral é

$$\left\{(x,y,z)\in\mathbb{R}^3 \mid x=\frac{-3z-1}{7} \text{ e } y=\frac{1-4z}{7}\right\} \text{ ou } \left\{\left(\frac{-3z-1}{7},\frac{1-4z}{7},z\right) \mid z\in\mathbb{R}\right\}.$$

4. Vejamos agora um sistema com quatro incógnitas. O leitor deve procurar reproduzir o raciocínio a partir das indicações apresentadas.

$$\begin{cases} 10y-4z+w=1 \\ x+4y-z+w=2 & (l_1 \leftrightarrow l_2) \\ -2x-8y+2z-2w=-4 \\ 3x+2y+z+w=5 \end{cases}$$

$$\longrightarrow \begin{cases} \mathbf{x}+4y-z+w=2 \\ \quad 10y-4z+w=1 \\ -2x-8y+2z-2w=-4 & (l_3+2l_1) \\ 3x+2y+z+w=5 & (l_4-3l_1) \end{cases} \longrightarrow \begin{cases} \mathbf{x}+4y-z+w=2 \\ \quad \mathbf{10y}-4z+w=1 \\ \quad 0=0 \\ \quad -10y+4z-2w=-1 \end{cases}$$

$$\longrightarrow \begin{cases} \mathbf{x}+4y-z+w=2 \\ \quad \mathbf{10y}-4z+w=1 \\ \quad -10y+4z-2w=-1 & (l_3+l_2) \end{cases} \longrightarrow \begin{cases} \mathbf{x}+4y-z+w=2 \\ \quad \mathbf{10y}-4z+w=1 \\ \quad -\mathbf{w}=0 \end{cases}$$

$$\longrightarrow \begin{cases} x+4y-z+w=2 \\ 10y-4z+w=1 \\ w=0 \end{cases} \longrightarrow \begin{cases} x+4y-z=2 \\ 10y-4z=1 \\ w=0 \end{cases} \longrightarrow \begin{cases} x+4y-z=2 \\ y=\frac{4z+1}{10} \\ w=0 \end{cases}$$

$$\longrightarrow \begin{cases} x=\frac{8-3z}{5} \\ y=\frac{4z+1}{10} \\ w=0 \end{cases}$$

O sistema tem novamente infinitas soluções (uma para cada valor atribuído à variável livre z), logo é possível indeterminado e a sua solução geral é a seguinte.

$$\left\{(x,y,z)\in\mathbb{R}^3 \mid x=\frac{8-3z}{5} \text{ e } y=\frac{4z+1}{10} \wedge w=0\right\} = \left\{\left(\frac{8-3z}{5},\frac{4z+1}{10},z,0\right) : z\in\mathbb{R}\right\}$$

Exercício 8. Recorrendo ao método de eliminação de Gauss, encontre soluções dos seguintes sistemas de equações lineares. Indique em cada passo o pivot e os multiplicadores usados.

(a) $\begin{cases} 2x+3y=5 \\ x-3y=-2 \end{cases}$

(b) $\begin{cases} x+y+z=0 \\ 7x+3y+2z=3 \\ 5x+y+5z=8 \end{cases}$

(c) $\begin{cases} 3x-2y=0 \\ x+2z=0 \\ y+3z=0 \end{cases}$

(d) $\begin{cases} x-2y=0 \\ x+2y=4 \end{cases}$

(e) $\begin{cases} 2x+3y=6 \\ 3x+2y=9 \end{cases}$

(f) $\begin{cases} 2x+y-2z=-1 \\ x+2y+z=5 \\ 3x-z=1 \end{cases}$

1.1.4 Sistemas homogéneos e solução geral de sistemas de equações lineares

A cada sistema de equações lineares

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases}$$

corresponde um sistema homogéneo associado

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases}$$

que é obtido do sistema original eliminando os termos independentes.

A solução geral S de um qualquer sistema de equações lineares pode ser determinada usando a solução S_h do sistema homogéneo associado, se conhecermos uma solução particular S_p do sistema inicial. De facto, verifica-se que

$$S = S_p + S_h\,.$$

Exemplo. Consideremos o sistema seguinte.

$$\begin{cases} 4x+y-z=4 \\ 2x+3y+2z=7 \end{cases}$$

Vamos começar por resolver este sistema pelo método de eliminação de Gauss.

$$\begin{cases} \mathbf{4x}+y-z=4 \\ 2x+3y+2z=7 \end{cases} \quad (2l_2-l_1) \quad \longrightarrow \begin{cases} \mathbf{4x}+y-z=4 \\ \quad\mathbf{5y}+5z=10 \end{cases}$$

$$\longrightarrow \begin{cases} 4x+y-z=4 \\ y=2-z \end{cases} \quad \longrightarrow \begin{cases} x=\frac{1}{2}+\frac{1}{2}z \\ y=2-z \end{cases}$$

Logo, a solução geral deste sistema é

$$S = \left\{(x,y,z) \in \mathbb{R}^3 \mid x = \frac{1}{2} + \frac{1}{2}z \text{ e } y = 2 - z\right\} = \left\{\left(\frac{1}{2} + \frac{1}{2}z, 2 - z, z\right) \mid z \in \mathbb{R}\right\}.$$

Vamos agora verificar que obtemos a mesma solução do sistema usando uma sua solução particular e o conjunto solução do sistema homogéneo associado.

Para determinarmos uma solução particular, fixemos $z = 0$. Obtemos o sistema seguinte.

$$\begin{cases} \mathbf{4x} + y = 4 \\ 2x + 3y = 7 \end{cases} \quad (2l_2 - l_1) \quad \longrightarrow \begin{cases} 4x + y = 4 \\ 5y = 10 \end{cases} \longrightarrow \begin{cases} 4x + y = 4 \\ y = 2 \end{cases}$$

$$\longrightarrow \begin{cases} 4x = 2 \\ y = 2 \end{cases} \longrightarrow \begin{cases} x = \frac{1}{2} \\ y = 2 \end{cases}$$

Obtemos a solução particular $S_p = \left(\frac{1}{2}, 2, 0\right)$. Vamos agora resolver o sistema homogéneo associado.

$$\begin{cases} \mathbf{4x} + y - z = 0 \\ 2x + 3y + 2z = 0 \end{cases} \quad (2l_2 - l_1) \quad \longrightarrow \begin{cases} \mathbf{4x} + y - z = 0 \\ \quad \mathbf{5y} + 5z = 0 \end{cases} \longrightarrow \begin{cases} 4x + y - z = 0 \\ y = -z \end{cases}$$

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

A solução geral deste sistema homogéneo é então

$$S_h = \left\{(x,y,z) \in \mathbb{R}^3 \mid x = \frac{1}{2}z \text{ e } y = -z\right\} = \left\{\left(\frac{1}{2}z, -z, z\right) \mid z \in \mathbb{R}\right\}.$$

Temos que

$$S_p + S_h = \left\{\left(\frac{1}{2}, 2, 0\right)\right\} + \left\{\left(\frac{1}{2}z, -z, z\right) \mid z \in \mathbb{R}\right\} = \left\{\left(\frac{1}{2} + \frac{1}{2}z, 2 - z, z\right) \mid z \in \mathbb{R}\right\},$$

tendo-se a relação $S = S_p + S_h$.

Vale a pena analisar um pouco em detalhe este exemplo para perceber porque é que se tem a relação $S = S_p + S_h$. Seja (x^*, y^*, z^*) uma solução do sistema homogéneo; então necessariamente tem-se

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

por definição de solução. Ora sabemos que $\left(\frac{1}{2}, 2, 0\right)$ é uma solução particular do sistema original ou, de forma equivalente, que

$$\begin{cases} 4 \times \frac{1}{2} + 2 - 0 = 4 \\ 2 \times \frac{1}{2} + 3 \times 2 + 2 \times 0 = 7 \end{cases}$$

são equações verdadeiras. Substituindo x por $\frac{1}{2} + x^*$, y por $2 + y^*$ e z por z^*, obtemos o sistema

$$\begin{cases} 4\left(\frac{1}{2}+x^*\right)+(2+y^*)-z^*=4 \\ 2\left(\frac{1}{2}+x^*\right)+3\,(2+y^*)+2z^*=7 \end{cases} \iff \begin{cases} 4\times 12+4x*+2+y^*-z^*=4 \\ 2\times\frac{1}{2}+2x^*+3\times 2+3y^*+2z^*=7 \end{cases}$$

$$\iff \begin{cases} \underbrace{\left(4\times\frac{1}{2}+2-0\right)}_{4}+\underbrace{(4x^*+y^*-z^*)}_{0}=4 \\ \underbrace{\left(2\times\frac{1}{2}+3\times 2+2\times 0\right)}_{7}+\underbrace{(2x^*+3y^*+2z^*)}_{0}=7 \end{cases}$$

que é verdadeiro devido à escolha de x^*, y^* e z^*. O raciocínio no caso geral é análogo.

Exercício 9. Escreva o sistema homogéneo correspondente a cada um dos seguintes sistemas de equações e resolva-o. Usando a solução particular fornecida, escreva a expressão da solução geral do sistema de equações em questão.

(a) $x-y=2$

Solução particular: $(x,y)=(3,1)$

(b) $\begin{cases} x-y=3 \\ 2x-2y=6 \end{cases}$

Solução particular: $(x,y)=(2,-1)$

(c) $x+y+z=2$

Solução particular: $(x,y,z)=(1,1,0)$

(d) $\begin{cases} 2x+y+z=1 \\ 3x+2y-z=-8 \end{cases}$

Solução particular: $(x,y,z)=(-2,1,4)$

(e) $\begin{cases} 3x-2y+z+w=6 \\ -x+y+2z=3 \\ y+w=1 \end{cases}$

Solução particular: $(x,y,z,w)=(1,0,2,1)$

1.1.5 Representação matricial

Vamos agora ver uma forma mais expedita de aplicar o método de eliminação de Gauss a sistemas de equações. Muito do que escrevemos nos exemplos das secções anteriores era na realidade desnecessário; uma vez que as incógnitas aparecem sempre pela mesma ordem, tudo o que precisamos para resolver os sistemas é de saber realizar as operações elementares sobre os coeficientes.

A representação matricial é uma forma mais compacta de representar sistemas de equações. Em vez de as escrevermos por extenso, representamos um sistema por uma tabela, com uma coluna correspondendo a cada variável; em cada linha, escrevemos o coeficiente de cada variável que ocorre na equação na coluna correspondente, tendo o cuidado de inserir um 0 nas colunas correspondentes a variáveis que não ocorrem nessa equação. A última coluna contém os termos independentes.

Definição. Dado um sistema de equações lineares

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases}$$

a sua *matriz de coeficientes* é a tabela

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

e a sua *representação em matriz aumentada* é a tabela

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

Desde que se saiba a ordem das incógnitas, é possível reescrever o sistema original a partir da matriz aumentada. Daqui em diante, convencionaremos que as variáveis surgem sempre pela ordem x, y, z, w, t. O termo *matriz aumentada* salienta que a matriz tem mais uma coluna do que o número de incógnitas (a matriz não aumentada é a matriz dos coeficientes, que teremos oportunidade de estudar mais adiante).

O método de eliminação de Gauss pode ser aplicado à matriz aumentada exactamente da mesma forma que era usado em sistemas de equações; podemos somar linhas da matriz e multiplicá-las por constantes realizando essas operações sobre as entradas da mesma coluna.

Vejamos um exemplo. Vamos resolver o sistema

$$\begin{cases} \mathbf{x} + y + 2z = 9 \\ 2x + 4y - 3z = 1 \\ 3x + 6y - 5z = 0 \end{cases}$$

por eliminação de Gauss nas duas formas: usando a representação em sistema (à esquerda) e em matriz aumentada (à direita).

$$\begin{cases} \mathbf{x} + y + 2z = 9 \\ 2x + 4y - 3z = 1 & (l_2 - 2l_1) \\ 3x + 6y - 5z = 0 & (l_3 - 3l_1) \end{cases} \qquad \left[\begin{array}{ccc|c} \mathbf{1} & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right] \begin{array}{l} \\ l_2 - 2l_1 \\ l_3 - 3l_1 \end{array}$$

$$\longrightarrow \begin{cases} \mathbf{x} + y + 2z = 9 \\ \quad \mathbf{2y} - 7z = -17 \\ \quad 3y - 11z = -27 & (2l_3 + 3l_2) \end{cases} \qquad \longrightarrow \left[\begin{array}{ccc|c} \mathbf{1} & 1 & 2 & 9 \\ 0 & \mathbf{2} & -7 & -17 \\ 0 & 3 & -11 & -27 \end{array}\right] \begin{array}{l} \\ \\ 2l_3 + 3l_2 \end{array}$$

$$\longrightarrow \begin{cases} \mathbf{x} + y + 2z = 9 \\ \quad -\mathbf{2y} + 7z = 17 \\ \qquad -\mathbf{z} = -3 \end{cases} \qquad \longrightarrow \left[\begin{array}{ccc|c} \mathbf{1} & 1 & 2 & 9 \\ 0 & -\mathbf{2} & 7 & 17 \\ 0 & 0 & -\mathbf{1} & -3 \end{array}\right]$$

Neste ponto concluimos a aplicação do método de eliminação de Gauss: já não há elementos para eliminar abaixo dos pivots. Assim, podemos voltar a escrever o sistema correspondente à última matriz aumentada e resolvê-lo.

$$\begin{cases} x + y + 2z = 9 \\ -2y + 7z = 17 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x + y = 3 \\ -2y = -4 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x + y = 3 \\ y = 2 \\ z = 3 \end{cases} \longrightarrow \begin{cases} x = 1 \\ y = 2 \\ z = 3 \end{cases}$$

Exercício 10. Escreva cada um dos sistemas do Exercício 5 na forma matricial e resolva-os aplicando o método de eliminação de Gauss à matriz obtida.

A matriz que resulta do método de eliminação de Gauss tem um formato especial: os valores diferentes de zero foram uma "escada", em que cada linha começa mais à direita do que a linha anterior. Estas matrizes chamam-se matrizes em escada de linhas.

Definição. Uma matriz diz-se *em escada de linhas* se satisfaz as duas condições seguintes.

1. Em duas linhas consecutivas da matriz, o primeiro elemento não nulo da linha inferior encontra-se estritamente à direita do primeiro elemento não nulo da linha superior.
2. Todas as linhas nulas (linhas em que todas as entradas são zero) estão agrupadas na base da matriz.

Por analogia com os sistemas de equações lineares, à primeira entrada não nula de cada linha chamamos *pivot*. Ao método de eliminação de Gauss aplicado a matrizes também se chama *método de redução*, sendo a matriz em escada de linhas que se obtém no final da aplicação do método chamada *matriz reduzida*.

Definição. A *característica* duma matriz em escada de linhas é o número de pivots que contém.

Exemplo. As matrizes

$$\begin{bmatrix} \mathbf{2} & -1 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & \mathbf{1} & 2 & -1 \\ 0 & 0 & -\mathbf{3} & 1 \end{bmatrix} \qquad \text{e} \qquad \begin{bmatrix} \mathbf{1} & -1 & 2 & 0 & 0 & 2 \\ 0 & 0 & -\mathbf{3} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

são matrizes em escadas de linhas de características, respectivamente, 2, 2 e 3. Os elementos assinalados a negrito são os pivots.

As matrizes seguintes não são matrizes em escada de linhas.

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

Exercício 11. Quais das seguintes matrizes são em escada de linhas?

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

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

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

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

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

(f) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

Os sistemas de equações podem ser caracterizados analisando a sua matriz reduzida, que tem uma das seguintes três formas.

1. Se a matriz contém uma linha em que apenas a última entrada não é nula, ou seja, a matriz tem o aspecto
$$\left[\begin{array}{ccccc|c} \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & k \\ \vdots & \vdots & \vdots & & \vdots & \vdots \end{array}\right]$$
em que $k \neq 0$, então essa linha corresponde à equação $0 = k$, que é impossível. Neste caso, o sistema de equações lineares é impossível.

2. Se a matriz não contém nenhuma linha daquela forma, então o sistema é possível. Neste caso, se a sua característica for igual ao número de incógnitas, o sistema é possível determinado.

3. Se a matriz não contém nenhuma linha daquela forma e a sua característica é inferior ao número de incógnitas, então o sistema é possível indeterminado.

Exercício 12. Classifique cada um dos sistemas de equações seguintes como determinado, indeterminado ou impossível. Nos primeiros dois casos, encontre a(s) solução(ões) do sistema.

(a) $\begin{cases} 3x + 2y = 7 \\ 2x - 2y = -2 \end{cases}$ (b) $\begin{cases} x + 2y - z = -3 \\ 3x - y + 2z = 8 \\ 2x + y - 4z = -7 \end{cases}$ (c) $\begin{cases} x + 2z - 2w = 1 \\ 2x + y = 2 \\ -x + 3y - z + w = -1 \end{cases}$

Esta classificação é especialmente útil quando estamos a tratar de sistemas que dependem de parâmetros.

Exemplo. Consideremos o sistema de equações lineares seguinte.

$$\begin{cases} x - 4z = -3 \\ 2x + \alpha y - 3z = -2 \\ x + y + \alpha z = 1 \end{cases}$$

A caracterização deste sistema depende do valor atribuído ao parâmetro α. Para encontrar os valores de α que tornam o sistema possível determinado, possível indeterminado e impossível, começamos por aplicar o método de eliminação de Gauss à matriz aumentada associada ao sistema.

$$\left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 2 & \alpha & -3 & -2 \\ 1 & 1 & \alpha & 1 \end{array}\right] \begin{array}{l} \\ l_2 - 2l_1 \\ l_3 - l_1 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 0 & \alpha & 5 & 4 \\ 0 & 1 & 4+\alpha & 4 \end{array}\right]$$

Com $\alpha = 0$, temos

$$\left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 0 & 0 & 5 & 4 \\ 0 & 1 & 4 & 4 \end{array}\right] \; l_2 \leftrightarrow l_3 \; \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 0 & 1 & 4 & 4 \\ 0 & 0 & 5 & 4 \end{array}\right]$$

e obtemos um sistema com característica igual ao número de incógnitas e sem equações impossíveis. Então este sistema é possível determinado para $\alpha = 0$.

Se $\alpha \neq 0$, então

$$\left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 0 & \alpha & 5 & 4 \\ 0 & 1 & 4+\alpha & 4 \end{array}\right] \begin{array}{c} \\ \\ \alpha l_3 - l_2 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 0 & \alpha & 5 & 4 \\ 0 & 0 & \alpha^2+4\alpha-5 & 4\alpha-4 \end{array}\right]$$

e o sistema é possível determinado quando esta matriz tem três pivots, o que sucede se $\alpha^2 + 4\alpha - 5 \neq 0$. Ora

$$\alpha^2 + 4\alpha - 5 = 0 \Longleftrightarrow \alpha = -2 \pm \sqrt{4+5} \Longleftrightarrow \alpha = -2 \pm 3 \Longleftrightarrow \alpha = 1 \text{ ou } \alpha = -5$$

donde o sistema é possível determinado se $\alpha \neq 1$ e $\alpha \neq -5$.

Nos casos $\alpha = 1$ e $\alpha = -5$, o lado esquerdo da última equação vale 0. No primeiro caso, $\alpha = 1$ e $4\alpha - 4 = 0$, donde a última equação é $0 = 0$ e o sistema é possível indeterminado; no segundo caso, $\alpha = -5$ e $4\alpha - 4 = -24$, donde a última equação é $0 = -24$ e o sistema é impossível.

Em suma: o sistema é impossível para $\alpha = -5$, possível indeterminado para $\alpha = 1$ e possível determinado para todos os outros valores de α.

Exemplo. Consideremos agora o sistema

$$\begin{cases} 2x + 7y - z = 26 \\ x - z = 3 \\ \alpha y + 6z = \beta \end{cases}$$

com parâmetros reais α e β. Para caracterizar o sistema, começamos por aplicar eliminação de Gauss à sua representação matricial.

$$\left[\begin{array}{ccc|c} 2 & 7 & -1 & 26 \\ 1 & 0 & -1 & 3 \\ 0 & \alpha & 6 & \beta \end{array}\right] \begin{array}{c} \\ 2l_2 - l_1 \\ \\ \end{array} \longrightarrow \left[\begin{array}{ccc|c} 2 & 7 & -1 & 26 \\ 0 & -7 & -1 & -20 \\ 0 & \alpha & 6 & \beta \end{array}\right] \begin{array}{c} \\ \\ 7l_3 + \alpha l_2 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} 2 & 7 & -1 & 26 \\ 0 & -7 & -1 & -20 \\ 0 & 0 & 42-\alpha & 7\beta - 20\alpha \end{array}\right]$$

Novamente, o sistema é possível determinado desde que tenha três pivots, o que sucede se $42 - \alpha \neq 0$, ou seja, se $\alpha \neq 42$.

Se $\alpha = 42$, então o sistema é impossível ou possível indeterminado consoante o valor de $7\beta - 20\alpha$. Se $7\beta - 20\alpha = 0$, o que corresponde a $\beta = 120$, a última equação é $0 = 0$ e o sistema é indeterminado; qualquer outro valor de β conduz a um sistema impossível.

Então o sistema é possível determinado se $\alpha \neq 42$, possível indeterminado se $\alpha = 42$ e $\beta = 120$ e impossível se $\alpha = 42$ e $\beta \neq 120$.

Exercício 13. Classifique cada um dos seguintes sistemas de equações lineares como possível determinado, indeterminado ou impossível em função dos parâmetros α e β.

(a) $\begin{cases} x+y+2z=2 \\ 2x-y+3z=2 \\ 5x-y+\alpha z=6 \end{cases}$ (b) $\begin{cases} x+2y=3 \\ \alpha y=\beta \end{cases}$ (c) $\begin{cases} (\alpha-3)x+y=0 \\ x+(\alpha-3)y=0 \end{cases}$

1.2 Matrizes e operações elementares

Na secção anterior recorremos a matrizes para simplificar a representação dos sistemas de equações lineares. Contudo, as matrizes têm uma aplicabilidade muito mais vasta do que a esse problema. Nesta secção, vamos dar início ao estudo destes objectos e das suas propriedades, que serão ferramentas essenciais em toda a Álgebra Linear.

1.2.1 Conceitos fundamentais

Uma matriz não é mais do que uma tabela rectangular, bidimensional, contendo informação. Tipicamente, na Álgebra Linear apenas estamos interessados em matrizes numéricas (cujas entradas são números reais), mas noutros contextos utilizam-se matrizes doutros tipos.

Definição. Uma *matriz real* é uma tabela rectangular contendo um número real em cada posição.

$$\begin{bmatrix} \text{linha } 1 \\ \text{linha } 2 \\ \vdots \\ \text{linha } n \end{bmatrix} \qquad \begin{bmatrix} \text{c} & \text{c} & & \text{c} \\ \text{o} & \text{o} & & \text{o} \\ \text{l} & \text{l} & & \text{l} \\ \text{u} & \text{u} & \cdots & \text{u} \\ \text{n} & \text{n} & & \text{n} \\ \text{a} & \text{a} & & \text{a} \\ & & & \\ 1 & 2 & & m \end{bmatrix}$$

Diz-se que uma matriz A com n linhas e m colunas tem *dimensão* $n \times n$, o que se denota por $\dim(A) = n \times m$. Também é frequente escrever $A_{n\times m}$ para indicar que se está a falar duma matriz A com n linhas e m colunas.

Ao valor na intersecção da linha i com a coluna j chama-se *entrada* (i,j) da matriz. A entrada (i,j) duma matriz A denota-se habitualmente por a_{ij} ou, por vezes, por $(A)_{ij}$.

Exemplo.

1. A matriz $\begin{bmatrix} 2 & 1 & 2 \\ 0 & 1 & -1 \end{bmatrix}$ é uma matriz de dimensão 2×3. A entrada $(2,1)$ desta matriz é 0, enquanto as entradas $(1,1)$ e $(1,3)$ são ambas 2. Denotando a matriz por A, podemos resumir esta informação dizendo que é uma matriz $A_{2\times 3}$, que $a_{21}=0$ e que $a_{11}=a_{13}=2$.

2. A matriz $B = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 6 & 0 \end{bmatrix}$ é uma matriz de dimensão 3×2. Comparando com a matriz A

acima, observa-se que a ordem dos valores na dimensão é importante: a matriz A tem 2 linhas e 3 colunas, a matriz B tem 3 linhas e apenas 2 colunas.

Recorrendo às notações acima, podemos afirmar por exemplo que $b_{22} = -1$ e $b_{31} = 6$.

3. Considere-se uma matriz $C_{2\times 2}$ tal que $c_{11} = c_{22} = 1$ e $c_{12} = 2$. Então a matriz C tem a forma
$$\begin{bmatrix} 1 & 2 \\ * & 1 \end{bmatrix}$$
onde $*$ é um valor desconhecido.

4. Considere-se a seguinte matriz D.
$$D = \begin{bmatrix} 1 & -2 & 3 & 0 \\ \pi & 0 & 1 & 1 \\ 0 & -4 & 2 & 1 \end{bmatrix}.$$
A matriz D tem dimensão 3×4 e satisfaz $d_{14} = d_{22} = d_{31} = 0$.

Alguns tipos de matrizes têm nomes especiais, devido à importância que têm em determinadas aplicações, que é importante conhecer. Relativamente à dimensão, há duas famílias importantes: as matrizes quadradas e os vectores. A razão de ser do nome destes últimos tornar-se-á mais clara nos capítulos seguintes.

Definição. Uma matriz com igual número de linhas e de colunas diz-se uma *matriz quadrada*.

Definição. Uma matriz com apenas uma coluna diz-se um *vector coluna*; uma matriz com apenas uma linha diz-se um *vector linha*.

Exemplo.

1. A matriz C do exemplo anterior é uma matriz quadrada de dimensão 2×2. Uma vez que o número de linhas e colunas coincide sempre para matrizes quadradas, também podemos afirmar simplesmente que C é uma matriz quadrada de dimensão 2.

2. A matriz $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ é uma matriz quadrada de dimensão 3.

3. A matriz $\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$ é uma matriz de dimensão 1×4; é portanto um vector linha (só tem uma linha), sendo também usual dizer-se que é um vector linha de dimensão 4: o número de linhas fica determinado ao dizermos que se trata de um vector linha.

4. De forma análoga, a matriz $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ é um vector coluna de dimensão 3.

5. A matriz $\begin{bmatrix} 3 \end{bmatrix}$ é simultaneamente uma matriz quadrada, um vector linha e um vector coluna (de dimensão 1, em qualquer dos casos). É típico identificar qualquer matriz desta forma com a sua única entrada; esta matriz corresponde, portanto, ao número 3.

6. As matrizes A e B do exemplo anterior não são nem quadradas, nem vectores. Por vezes, diz-se que são *matrizes rectangulares* para chamar a atenção para esse facto.

Exercício 14. Para cada uma das seguintes matrizes, indique a sua dimensão e descreva algumas das suas entradas. Quais das matrizes são quadradas? Quais são vectores linha? Quais são vectores coluna?

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

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

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

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

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

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

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

(h) $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

(i) $\begin{bmatrix} 1 \end{bmatrix}$

1.2.2 Matrizes quadradas

As matrizes quadradas desempenham um papel especial em muitos contextos, conforme teremos ocasião de apreciar ainda neste capítulo. Por esse motivo, há terminologia específica para estas matrizes que é importante conhecer.

Definição. A *diagonal principal* duma matriz quadrada $A_{n\times n}$ é composta pelas entradas $a_{11}, a_{22}, \ldots, a_{nn}$, assinaladas a negrito na matriz abaixo.

$$\begin{bmatrix} \mathbf{a_{11}} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & \mathbf{a_{22}} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & \mathbf{a_{33}} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & \mathbf{a_{nn}} \end{bmatrix}$$

Um caso particular muito importante são as matrizes que têm o valor 0 em todas as entradas fora da diagonal principal.

Definição. Uma matriz quadrada diz-se uma *matriz diagonal* se todas as suas entradas forem nulas, exceptuando eventualmente os valores na diagonal principal.

Definição. Uma matriz quadrada de dimensão n que tem 1 em todas as entradas sobre a diagonal principal e 0 em todas as restantes é uma matriz diagonal que se diz *matriz identidade* de dimensão n. Esta matriz é normalmente denotada por $\mathbf{I_n}$.

Observe-se que a matriz identidade duma dada dimensão é única. Por exemplo, as matrizes identidade de dimensão 2 e 3 são, respectivamente,

$$\mathbf{I_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{e} \qquad \mathbf{I_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercício 15. Escreva as matrizes identidade de dimensão 4 e 5.

Exercício 16. Quais das seguintes matrizes são diagonais?

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

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

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

(d) $\begin{bmatrix} 1 \end{bmatrix}$

(e) $\begin{bmatrix} 0 \end{bmatrix}$

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

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

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

(i) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & -3 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

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

Observe-se que uma matriz diagonal fica completamente definida se indicarmos os valores das entradas sobre a diagonal. Por exemplo: a única matriz diagonal com diagonal principal $\begin{bmatrix} 1 & -1 & 2 \end{bmatrix}$ é a matriz

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

Esta correspondência entre vectores linha e matrizes diagonais é um exemplo de *transformação linear*. As transformações lineares serão o objecto de estudo do Capítulo 3.

Exercício 17. Escreva as matrizes diagonais correspondentes a cada um dos seguintes vectores linha.

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

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

(c) $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & -1 \end{bmatrix}$

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

(f) $\begin{bmatrix} 0 & 2 \end{bmatrix}$

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

(h) $\begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}$

(i) $\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$

Um outro conceito que está intimamente ligado com o de diagonal principal (e que portanto só faz sentido para matrizes quadradas) é o conceito de traço. Embora neste contexto não tenhamos oportunidade de aprofundar o seu estudo, o traço duma matriz é uma ferramenta de grande utilidade em diversas aplicações e será referido em disciplinas mais avançadas.

Definição. Seja A uma matriz quadrada de dimensão n. O *traço* de A é a soma dos elementos da diagonal principal de A.

Simbolicamente:

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{k=1}^{n} a_{kk}\,.$$

Exemplo. O traço da matriz

$$A = \begin{bmatrix} 3 & 4 & 8 & 3 \\ 8 & 7 & 5 & 8 \\ 6 & 7 & 4 & 8 \\ 5 & 9 & 9 & 9 \end{bmatrix}$$

é dado por $\operatorname{tr}(A) = 3 + 7 + 4 + 9 = 23$.

O traço duma matriz não muda se lhe alterarmos entradas fora da diagonal principal.

Exercício 18. Qual é o traço da matriz identidade?

1.2.3 Outras formas de definir matrizes

Até agora, temos sempre apresentado explicitamente as matrizes com que trabalhámos, com uma excepção. Num dos primeiros exemplos, definimos uma matriz C duma forma diferente das outras, indicando duma forma descritiva os valores as suas entradas. Esta forma de especificar matrizes é especialmente útil quando os valores as entradas são fáceis de escrever como função dos índices i e j de linha e coluna.

Por exemplo, a matriz identidade de dimensão 2 é uma matriz $A_{2\times 2}$ que tem o valor 1 nas entradas sobre a diagonal principal (ou seja, nas entradas em que os índices de linha i e de coluna j coincidem) e 0 nas restantes. Podemos então dizer que esta matriz A satisfaz

$$a_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

Como é que se constrói a matriz A a partir desta especificação? Percorrem-se as suas entradas e avalia-se a expressão de a_{ij} para cada valor de i e j, escrevendo o resultado na entrada correspondente. Por exemplo, para a_{11} tem-se que $i = 1 = j$, pelo que $a_{11} = 1$; para a_{12} temos $1 = i \neq j = 2$, donde $a_{12} = 0$. Prosseguindo de forma semelhante, concluiríamos que a matriz era a matriz $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, que é de facto a matriz identidade de dimensão 2.

Exemplo. Consideremos uma matriz $B_{2\times 3}$ tal que $b_{ij} = i + j$. Para construir a matriz B vamos dando valores a i (entre 1 e 2) e a j (entre 1 e 3) e calculando o valor da entrada correspondente: $b_{11} = 1 + 1 = 2$; $b_{12} = 1 + 2 = 3$; $b_{13} = 1 + 3 = 4$; $b_{21} = 2 + 1 = 3$; e assim por diante. Obtemos então a matriz

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

Exercício 19. Escreva as seguintes matrizes.

(a) $A_{2\times 3}$ tal que $a_{ij} = i \times j$

(b) $B_{3\times 2}$ tal que $b_{ij} = i \times j$

(c) $A_{3\times 3}$ tal que $a_{ij} = i - j$

(d) $B_{3\times 3}$ tal que $b_{ij} = \begin{cases} 0 & \text{se } i = j \\ i & \text{se } i \neq j \end{cases}$

(e) $A_{4\times 4}$ tal que $a_{ij} = |i - j|$

(f) $C_{3\times 3}$ tal que $c_{ij} = \begin{cases} i - j & \text{se } i \geq j \\ i + j & \text{se } i < j \end{cases}$

(g) $C_{2\times 4}$ tal que $c_{ij} = \begin{cases} |i - j| & \text{se } |i - j| \leq 1 \\ 0 & \text{se } |i - j| \geq 2 \end{cases}$

Para cada dimensão $n \times m$, a matriz que contém o valor zero em todas as entradas diz-se a *matriz nula* dessa dimensão, e denota-se habitualmente como $\mathbf{0_{n\times m}}$.

Exemplo. As matrizes nulas de dimensão 2×2, 2×3 e 3×3 são, respectivamente,

$$\mathbf{0_{2\times 2}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{0_{2\times 3}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{0_{3\times 3}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

1.2.4 Operações lineares sobre matrizes

Tal como acontece com os números reais, é possível comparar matrizes e efectuar operações entre matrizes; contrariamente ao que se passa com os números reais, estas operações estão sujeitas a restrições sobre as dimensões das matrizes.

Para compreender estas restrições, vamos começar por estudar a igualdade de matrizes. Intuitivamente, duas matrizes são iguais se contêm os mesmos valores nas mesmas entradas.

Definição. Duas matrizes A e B dizem-se *iguais*, denotado por $A = B$, se tiverem a mesma dimensão e todas as suas entradas forem iguais. Equivalentemente, $A = B$ se $\dim(A) = \dim(B)$ e $a_{ij} = b_{ij}$ para quaisquer valores de i e j.

A condição relativamente à dimensão é essencial: as duas matrizes $A = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 2 & 3 \\ -2 & -1 & 0 \end{bmatrix}$ claramente não são iguais, embora se tenha $a_{ij} = b_{ij}$ para todos os valores de i e j que dão sentido a esta expressão.

Duas matrizes da mesma dimensão podem ser somadas componente a componente, obtendo-se uma nova matriz com a mesma dimensão que as matrizes originais.

Definição. Sejam A e B duas matrizes com a mesma dimensão $n \times m$. A *soma* de A e B é a matriz $A + B$, também de dimensão $n \times m$, cujas entradas são a soma das entradas respectivas de A e de B: $(A + B)_{ij} = (A)_{ij} + (B)_{ij}$.

Explicitamente:

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

Exemplo. Sejam A e B as seguintes matrizes

$$A = \begin{bmatrix} 3 & 8 & 2 \\ 9 & 3 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & \pi & \frac{1}{2} \\ 1 & 1 & -2 \end{bmatrix}.$$

A sua soma é dada por

$$A + B = \begin{bmatrix} 3 & 8+\pi & \frac{5}{2} \\ 10 & 4 & -1 \end{bmatrix}.$$

Novamente, podemos observar a importância da restrição sobre as dimensões: se quiséssemos somar as matrizes

$$A = \begin{bmatrix} 3 & 8 & 2 \\ 9 & 3 & 1 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix}$$

deparar-nos-íamos com um problema: a matriz resultado deve ter duas colunas ou três? E no caso de serem três, que valores devemos incluir nessa coluna?

Exercício 20. Calcule as seguintes somas matriciais.

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

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

(c) $\begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix} + \mathbf{0_{2\times 2}}$

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

(e) $\mathbf{0_{2\times 3}} + \begin{bmatrix} -1 & -1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$

(f) $\begin{bmatrix} 7 & 1 & -3 \\ 1 & 2 & -1 \\ -9 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 0 \\ -1 & -1 & -1 \end{bmatrix}$

(g) $\begin{bmatrix} 2 & 1 & -1 \\ -2 & 0 & -2 \\ 9 & 6 & 3 \end{bmatrix} + \begin{bmatrix} -2 & -1 & 1 \\ 2 & 0 & 2 \\ -9 & -6 & -3 \end{bmatrix}$

Visto que a soma de matrizes é definida directamente usando a soma dos reais, é intuitivo que todas as propriedades da soma de reais se continuem a verificar.

Proposição 1. Sejam A e B duas matrizes de dimensão $n \times m$. Então verificam-se as seguintes igualdades.

1. Comutatividade: $A + B = B + A$
2. Associatividade: $A + (B + C) = (A + B) + C$
3. Existência de elemento neutro: $A + \mathbf{0_{n\times m}} = \mathbf{0_{n\times m}} + A = A$
4. Existência de simétrico: $A + (-A) = (-A) + A = \mathbf{0_{n\times m}}$, onde $-A$ é a matriz cujas entradas são as simétricas das entradas de A

A partir da última propriedade, podemos definir *subtracção* de matrizes como habitualmente, tomando $A - B = A + (-B)$. Alternativamente, poderíamos ter definido esta operação directamente: a matriz $A-B$ é a matriz que satisfaz $(A-B)_{ij} = a_{ij} - b_{ij}$. Ambas as alternativas conduzem ao mesmo resultado.

Exemplo. As propriedades acima permitem muitas vezes simplificar expressões sem efectuar explicitamente os cálculos. Por exemplo, consideremos duas matrizes A e B dadas por

$$A = \begin{bmatrix} 2 & 6 & -3 \\ 1 & 5 & -1 \\ 3 & 1 & 0 \end{bmatrix} \text{ e } B = \begin{bmatrix} 1 & 2 & 5 \\ -2 & 1 & -9 \\ 0 & 2 & 1 \end{bmatrix}$$

e suponhamos que pretendíamos calcular o valor da expressão $(A + B) - A$. Usando as propriedades acima, podemos simplificar esta expressão:

$$(A + B) - A = (B + A) - A = B + (A - A) = B + \mathbf{0_{3\times 3}} = B$$

e chegamos ao resultado sem fazer quaisquer cálculos.

Outra operação que ocorre frequentemente é a multiplicação dum número real por uma matriz. Neste caso, o resultado é a matriz que contém o produto de cada entrada da matriz pelo real em causa.

Definição. Sejam α um número real e A uma matriz de dimensão $n \times m$. A matriz αA, produto de α por A, é uma matriz de dimensão $n \times m$ tal que $(\alpha A)_{ij} = \alpha(A)_{ij}$. Equivalentemente,

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

A esta operação também é costume chamar *produto da matriz A pelo escalar α*. O termo *escalar* surge no contexto dos espaços lineares; a justificação para este nome será discutida no Capítulo 2.

Exemplo. Seja A a matriz

$$A = \begin{bmatrix} 2 & 9 & 8 & 4 \\ 0 & 1 & 8 & 3 \\ 2 & 0 & 8 & 4 \end{bmatrix}.$$

A matriz $3A$ obtém-se multiplicando todas as entradas de A por 3:

$$3A = \begin{bmatrix} 6 & 27 & 24 & 12 \\ 0 & 3 & 24 & 9 \\ 6 & 0 & 24 & 12 \end{bmatrix}.$$

Tal como a soma, esta operação deriva da multiplicação de reais. Assim, verificam-se mais uma vez um conjunto de propriedades que vêm das propriedades das operações sobre os reais.

Proposição 2. Sejam A e B duas matrizes de dimensão $n \times m$ e α e β dois números reais. Então verificam-se as seguintes identidades.

1. Distributividade (na matriz): $\alpha(A + B) = \alpha A + \alpha B$

2. Distributividade (no escalar): $(\alpha + \beta)A = \alpha A + \beta A$

3. Associatividade: $\alpha(\beta A) = (\alpha\beta)A$

4. Compatibilidade: $(-1)A = -A$ e $0A = \mathbf{0_{n\times m}}$

A última propriedade diz que multiplicando uma matriz A pelo escalar -1 se obtém a matriz simétrica de A e que multiplicando qualquer matriz por 0 se obtém a matriz nula, justificando assim as notações que temos usado para matrizes nulas e matrizes simétricas.

Estas propriedades permitem-nos trabalhar com somas de matrizes e produtos de matrizes por escalares de forma muito semelhante à forma como trabalhamos com números reais. Por exemplo, podemos agrupar os termos em que ocorre a mesma matriz somando os coeficientes. A simplificação de expressões envolvendo matrizes é portanto muito mais eficiente usando estas propriedades.

Exemplo. Considerando as mesmas matrizes A e B do exemplo da página 27, podemos simplificar a expressão $(A+B)-5(B-A)+3B-2(A-B)$ começando por agrupar todos os termos em A e em B e somando os coeficientes; obtemos sucessivamente

$$\begin{aligned}(A+B)-5(B-A)+3B-2(A-B) &= (A+5A-2A)+(B-5B+3B+2B)\\ &= 4A+B = 4\begin{bmatrix} 2 & 6 & -3\\ 1 & 5 & -1\\ 3 & 1 & 0\end{bmatrix} + \begin{bmatrix} 1 & 2 & 5\\ -2 & 1 & -9\\ 0 & 2 & 1\end{bmatrix}\\ &= \begin{bmatrix} 8 & 24 & -12\\ 4 & 20 & -4\\ 12 & 4 & 0\end{bmatrix} + \begin{bmatrix} 1 & 2 & 5\\ -2 & 1 & -9\\ 0 & 2 & 1\end{bmatrix} = \begin{bmatrix} 9 & 26 & -7\\ 2 & 21 & -13\\ 12 & 6 & 1\end{bmatrix},\end{aligned}$$

expressão que é muito mais simples de obter do que efectuando directamente os cálculos.

Exercício 21. Considere as matrizes X, Y e Z seguintes.

$$X = \begin{bmatrix} 2 & -1 & 2\\ 1 & 0 & 0\end{bmatrix} \qquad Y = \begin{bmatrix} -3 & 2 & 5\\ 0 & 4 & 9\end{bmatrix} \qquad Z = \begin{bmatrix} -1 & 1 & 1\\ 3 & 2 & -2\end{bmatrix}$$

Calcule as matrizes denotadas pelas expressões seguintes, começando por as simplificar recorrendo às propriedades acima.

(a) $(X-Y)+(Y-Z)$

(b) $(X+Y)+(X-Y)$

(c) $(X-Z)+(Z-Y)+(Y-X)$

(d) $(3X+Y)-3(X+Y)-2Z$

(e) $2(X+Y+Z)-3(X+Y+Z)$

(f) $3X-2Y+3(X+Y+Z)-6(Z+X)$

No caso das matrizes quadradas, o traço também se relaciona de forma particularmente simples com estas operações. As propriedades seguintes são muito simples de verificar.

Proposição 3. Sejam A e B matrizes quadradas de dimensão n. Então verificam-se as seguintes identidades.

1. $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

2. $\mathrm{tr}(\alpha A) = \alpha\, \mathrm{tr}(A)$

Demonstração. Sejam A e B genericamente dadas por

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix} \text{ e } B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ b_{n1} & b_{n2} & \cdots & b_{nn}\end{bmatrix}.$$

Temos então as seguintes relações.

$$\begin{aligned}
\operatorname{tr}(A+B) &= \operatorname{tr}\left(\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \cdots & a_{nn}+b_{nn} \end{bmatrix}\right) \\
&= a_{11}+b_{11}+a_{22}+b_{22}+\cdots+a_{nn}+b_{nn} \\
&= (a_{11}+a_{22}+\cdots+a_{nn})+(b_{11}+b_{22}+\cdots+b_{nn}) \\
&= \operatorname{tr}\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\right) + \operatorname{tr}\left(\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}\right) \\
&= \operatorname{tr}(A)+\operatorname{tr}(B)
\end{aligned}$$

O raciocínio para o produto por um escalar é semelhante.

$$\begin{aligned}
\operatorname{tr}(\alpha\ A) &= \operatorname{tr}\left(\begin{bmatrix} \alpha\ a_{11} & \alpha\ a_{12} & \cdots & \alpha\ a_{1n} \\ \alpha\ a_{21} & \alpha\ a_{22} & \cdots & \alpha\ a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha\ a_{n1} & \alpha\ a_{n2} & \cdots & \alpha\ a_{nn} \end{bmatrix}\right) \\
&= \alpha\ a_{11}+\alpha\ a_{22}+\cdots+\alpha\ a_{nn} \\
&= \alpha\ \ (a_{11}+a_{22}+\cdots+a_{nn}) \\
&= \alpha\ \operatorname{tr}\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\right) \\
&= \alpha\ \operatorname{tr}(A)
\end{aligned}$$

$\square$

A última operação que vamos discutir nesta secção é uma operação dum tipo diferente, uma vez que altera a dimensão da matriz à qual é aplicada.

Definição. Seja A uma matriz de dimensão $n \times m$. A *matriz transposta* de A, designada por A^T, é uma matriz de dimensão $m \times n$ que se obtém escrevendo as linhas de A como colunas da nova matriz.

Se A for a matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

então a sua transposta é dada por

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \dots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}$$

Exemplo. Considere-se a seguinte matriz

$$A = \begin{bmatrix} 4 & 8 & 0 & 1 & 8 \\ 5 & 6 & 7 & 5 & 4 \end{bmatrix}.$$

A transposta de A é a matriz

$$A^T = \begin{bmatrix} 4 & 5 \\ 8 & 6 \\ 0 & 7 \\ 1 & 5 \\ 8 & 4 \end{bmatrix}.$$

Exercício 22. Calcule as transpostas das seguintes matrizes.

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

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

(c) $\begin{bmatrix} 1 & 6 \\ 8 & -1 \end{bmatrix}$

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

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

(f) $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

Exercício 23. Qual é a transposta da matriz identidade de dimensão n, $\mathbf{I_n}$? E qual é a transposta da matriz nula $\mathbf{0_{n\times m}}$?

Exercício 24. Quais são as matrizes que têm a mesma dimensão que as suas transpostas?

Tal como o traço, a transposição de matrizes comuta com as operações algébricas anteriormente definidas. A verificação das seguintes propriedades segue o mesmo raciocínio das anteriores e deixa-se como exercício.

Proposição 4. Sejam A e B matrizes de dimensão $n \times m$ e α um número real. Verificam-se as seguintes identidades.

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(\alpha A)^T = \alpha\, A^T$
4. $\mathrm{tr}\left(A^T\right) = \mathrm{tr}(A)$

Todas estas operações sobre matrizes (traço, soma, produto por um escalar e transposição) são operações bastante simples e que se comportam de forma semelhante às operações correspondentes sobre números reais. Na próxima secção discutiremos outras operações um pouco mais complexas e cujo comportamento já apresenta algumas características inesperadas.

Exercício 25. Para cada um dos seguintes pares de matrizes A e B, diga se é possível somá-las e, em caso afirmativo, calcule a matriz $A+B$.

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

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

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

Exercício 26. Sejam A, B e C as seguintes matrizes.

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

Verifique se as seguintes expressões estão correctas e, em caso afirmativo, calcule o seu valor.

(a) $\mathrm{tr}(A)+\mathrm{tr}(C)$ (b) $\mathrm{tr}(A+B)$ (c) $(A+B)^T$ (d) C^T+A (e) $2A^T+B$

1.3 Produto de matrizes

Nesta secção vamos discutir uma operação fundamental sobre matrizes: o produto de matrizes. Teremos oportunidade de verificar que este produto goza de propriedades bastante diferentes das do produto de números reais, mas o decorrer da exposição justificará esta definição. A partir do produto de matrizes, definem-se ainda outras operações que serão fundamentais nos capítulos posteriores.

Comecemos por recordar a definição do produto interno de dois vectores de números reais, que deve ser já conhecida do Ensino Secundário. Esta operação está definida para dois vectores (sequências de números reais) do mesmo tamanho; se $u = (u_1, u_2, \ldots, u_k)$ e $v = (v_1, v_2, \ldots, v_k)$ forem dois vectores de comprimento k, então o seu produto interno é um número real dado por

$$u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_kv_k = \sum_{i=1}^{k} u_iv_i \,.$$

Exercício 27. Calcule o produto interno dos seguintes pares de vectores.

(a) $(2,1)$ e $(1,-2)$

(b) $(2,0,1)$ e $(1,2,0)$

(c) $(1,1,1)$ e $(-2,-2,-1)$

(d) $(2,1,1,2)$ e $(1,1,1,-2)$

(e) $(2,1,2,1)$ e $(0,0,0,0)$

(f) $(1,1,0,1)$ e $(2,-2,1,1)$

Recordemos que convencionámos chamar vector a matrizes com uma linha e uma coluna. Escrevendo o vector u como vector linha e o vector v como vector coluna, podemos *definir* o produto de duas matrizes, sendo a primeira de dimensão $1 \times k$ e a segunda de dimensão $k \times 1$, como

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_k \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} = u \cdot v\,.$$

Para obtermos uma matriz como resultado, recordemos que convencionámos identificar matrizes de dimensão 1×1 com números reais; vendo o produto interno de u por v como uma matriz, podemos então escrever

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_k \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^k u_i v_i \end{bmatrix}\,.$$

Observe-se o que se passa em termos de dimensão.

$$U_{1\times\mathbf{k}} V_{\mathbf{k}\times 1} \rightsquigarrow (UV)_{1\times 1}$$

Suponhamos agora que em vez do vector linha u temos uma matriz arbitrária A, com n linhas e k colunas. Podemos ver esta matriz como uma lista de n vectores linha (um em cada linha da matriz), pelo que podemos formar n produtos internos, um de cada uma destas linhas com o vector v. Obtemos n números reais, que podemos dispor em coluna numa matriz resultado.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^k a_{1i} v_i \\ \sum_{i=1}^k a_{2i} v_i \\ \vdots \\ \sum_{i=1}^k a_{2i} v_i \end{bmatrix} = \begin{bmatrix} A_{1*} \cdot v \\ A_{2*} \cdot v \\ \vdots \\ A_{n*} \cdot v \end{bmatrix},$$

convencionando designar por A_{i*} a linha i da matriz A.

Antes de considerar o caso geral, há duas observações importantes a fazer. Em primeiro lugar, observe-se o que se passa em termos de dimensões:

$$A_{n\times\mathbf{k}} V_{\mathbf{k}\times 1} \rightsquigarrow (AV)_{n\times 1}\,.$$

Para o produto estar definido, é necessário que o número de colunas da matriz A corresponda à dimensão do vector v; o resultado terá tantas linhas quantas a matriz A.

Em segundo lugar, podemos desde já (antes sequer de considerar o caso geral) justificar a notação matricial utilizada para a resolução de sistemas de equações. Recorde-se que o sistema

$$\begin{cases} 3x + 2y - z = 1 \\ 2x - y + 5z = 0 \\ -2x + y = -2 \end{cases}$$

era representado matricialmente como

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

Podemos ler esta matriz aumentada como uma abreviatura para a relação matricial

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

De facto, de acordo com a definição de produto que demos acima, tem-se

$$\begin{bmatrix} 3 & 2 & -1 \\ 2 & -1 & 5 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + 2y - z \\ 2x - y + 5z \\ -2x + y \end{bmatrix}$$

e impor que este resultado seja igual à matriz da expressão anterior corresponde precisamente, atendendo à definição de igualdade entre matrizes, a exigir que as equações constantes do sistema inicial se verifiquem.

Também é frequente, denotando a matriz dos coeficientes por A,

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

e o vector coluna resultado por $\underline{b}$,

$$\underline{b} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix},$$

escrever o sistema como $A\underline{x} = \underline{b}$, onde os sublinhados chamam a atenção para o facto de se tratarem de vectores coluna.

Exercício 28. Escreva o sistema

$$\begin{cases} 3x + 2y - z + 4w = 1 \\ 2x - 5z + 2w = -1 \\ 3y + 2z - 3w = 9 \end{cases}$$

em forma de matriz aumentada. Expandindo esta matriz conforme exemplificado acima, verifique que se obtêm precisamente as três equações originais.

Exercício 29. Calcule os seguintes produtos de matrizes por vectores.

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

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

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

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

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

(f) $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}$

O passo seguinte de generalização deverá agora ser claro. Se o segundo factor não for um vector coluna mas sim uma matriz B, podemos interpretá-la como um conjunto de vectores coluna (um vector em cada coluna de B), efectuar os vários produtos da matriz A por cada uma dessas colunas, e escrever cada produto numa coluna do resultado. Obtemos uma matriz com tantas colunas como B.

Por exemplo: sejam $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & -1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}$. O resultado do produto de A pela primeira coluna de B é

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

e o produto de A pela segunda coluna de B é

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

donde o produto de A por B é

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

Continua a ser preciso que o número de colunas de A seja igual ao número de linhas de B.

Definição. Sejam $A_{n\times k}$ e $B_{k\times m}$ duas matrizes com as dimensões indicadas. O *produto matricial* de A e B é uma matriz de dimensão $n \times m$ que tem na entrada (i,j) o produto interno da linha i de A pela coluna j de B.

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \cdots & b_{1j} & \cdots \\ \cdots & b_{2j} & \cdots \\ & \vdots & \\ \cdots & b_{kj} & \cdots \end{bmatrix} = \begin{bmatrix} \ddots & \vdots & \iddots \\ \cdots & (AB)_{ij} & \cdots \\ \iddots & \vdots & \ddots \end{bmatrix}$$

Designando como atrás a linha i de A por A_{i*} e escrevendo a coluna j de B como B_{*j}, podemos também escrever

$$AB = \begin{bmatrix} A_{1*}B_{*1} & A_{1*}B_{*2} & \cdots & A_{1*}B_{*m} \\ A_{2*}B_{*1} & A_{2*}B_{*2} & \cdots & A_{2*}B_{*m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n*}B_{*1} & a_{n*}B_{*2} & \cdots & a_{n*}B_{*m} \end{bmatrix}$$

Exemplo. Sejam A e B as seguintes matrizes.

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

O produto AB é a matriz 2×2 seguinte.

$$\begin{aligned} AB &= \begin{bmatrix} 1 & -2 & 0 \\ 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 3 & -1 \end{bmatrix} \\ &= \begin{bmatrix} (1,-2,0)\cdot(0,1,3) & (1,-2,0)\cdot(2,1,-1) \\ (4,3,1)\cdot(0,1,3) & (4,3,1)\cdot(2,1,-1) \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 6 & 10 \end{bmatrix} \end{aligned}$$

Mais uma vez, mantém-se a relação fundamental entre as dimensões das matrizes envolvidas.

$$A_{n\times k} B_{k\times m} \rightsquigarrow (AB)_{n\times m}$$

Exercício 30. Calcule os seguintes produtos de matrizes.

(a) $\begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} -3 & 6 \\ 4 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 2 & 6 & 6 \\ 2 & 7 & 6 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 3 & 1 \end{bmatrix}$

Embora o produto de matrizes possa ser visto como uma generalização do produto de números reais, é uma operação substancialmente diferente e que não goza de muitas das propriedades a que estamos habituados. Para começar, não podemos multiplicar duas matrizes arbitrárias devido à restrição sobre as dimensões.

Exemplo. Não é possível calcular o produto de $\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$ e $\begin{bmatrix} 2 & 1 \\ -2 & 1 \\ 1 & -1 \end{bmatrix}$, uma vez que a primeira matriz tem duas colunas e a segunda tem três linhas.

Outra grande diferença diz respeito à comutatividade. O produto de números reais é comutativo (a ordem pela qual escrevemos os factores não afecta o resultado), mas com matrizes a situação é substancialmente diferente. Para começar, há situações em que o produto AB está definido mas o produto BA não está (ou vice-versa); em segundo lugar, mesmo quando ambos os produtos estão definidos, as dimensões dos resultados podem ser diferentes; finalmente, mesmo quando ambos os produtos estão definidos e têm a mesma dimensão, não há qualquer razão para serem iguais.

Exemplo. Considerem-se as seguintes matrizes.

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

1. Vimos num exemplo anterior que $AB = \begin{bmatrix} -2 & 0 \\ 6 & 10 \end{bmatrix}$. Também é fácil verificar que

$$BA = \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 4 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 6 & 2 \\ 5 & 1 & 1 \\ -1 & -9 & -1 \end{bmatrix}$$

e estamos portanto perante uma situação em que os produtos AB e BA não têm a mesma dimensão.

2. Se calcularmos o produto CA, concluímos que se tem

$$CA = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 4 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 14 & 5 & 3 \end{bmatrix}$$

mas o produto AC não está definido (A tem três colunas, mas C tem apenas duas linhas). Estamos perante uma situação em que um dos produtos faz sentido mas o outro não.

3. Considerando as matrizes C e D, podemos calcular ambos os produtos CD e DC, mas tem-se

$$CD = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 7 & 7 \end{bmatrix} \neq \begin{bmatrix} 0 & -3 \\ 7 & 9 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = DC$$

Exercício 31. Para cada um dos seguintes pares de matrizes A e B, diga se é possível multiplicá-las e, em caso afirmativo, calcule a matriz AB.

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

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

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

Exercício 32. Sejam $A_{2\times 2}$, $B_{2\times 3}$, $C_{3\times 2}$ e $D_{2\times 4}$ matrizes com as dimensões indicadas. Indique quais dos seguintes produtos de matrizes estão definidos e qual a dimensão do resultado.

(a) AB (c) AC (e) AD (g) $A(BC)$ (i) $\mathbf{I_2}A$

(b) BA (d) CB (f) $(AB)C$ (h) DA (j) $\mathbf{I_3}\,(C\mathbf{I_2})$

Contudo, a ausência de comutatividade não é tão grave como à partida poderia parecer; requer apenas algum cuidado com a ordem pela qual as operações são realizadas. Em contrapartida, o produto de matrizes goza dum conjunto de propriedades bastante mais importantes: é associativo e comporta-se "bem" com a soma de matrizes, o produto por escalares e a transposição.

Proposição 5. Sejam $A_{n\times m}$, B e C matrizes com as dimensões adequadas para as expressões apresentadas fazerem sentido. Então verificam-se as seguintes identidades.

1. Associatividade: $A(BC) = (AB)C$

2. Existência de elemento neutro: $A\mathbf{I_m} = A$ e $\mathbf{I_n}A = A$

3. Existência de elemento absorvente: $A\mathbf{0_{m\times k}} = \mathbf{0_{n\times k}}$ e $\mathbf{0_{k\times n}}A = \mathbf{0_{k\times m}}$

4. Distributividade à direita: $A(B+C) = AB + AC$
5. Distributividade à esquerda: $(A+B)C = AC + BC$
6. Associatividade com o produto por escalares: $\alpha(AB) = (\alpha A)B$
7. Relação com o traço: $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
8. Relação com a transposta: $(AB)^T = B^T A^T$

As provas destas propriedades são em geral simples e deixam-se como exercício. Apresentam-se algumas a título de exemplo do estilo de raciocínio envolvido.

Demonstração.

2. A entrada (i,j) do produto $A\mathbf{I_m}$ é calculada como o produto interno da linha i de A pela coluna j de $\mathbf{I_m}$. Ora a coluna j de $\mathbf{I_m}$ é um vector que tem 0 em todas as entradas excepto na entrada j; o seu produto interno com a linha i de A vai corresponder precisamente ao elemento na posição j desse vector — que é a_{ij}.

$$\begin{aligned}(A\mathbf{I_n})_{ij} &= (a_{i1}, a_{i2}, \dots, a_{im}) \cdot (0, \dots, 0, \underbrace{1}_{j}, 0, \dots, 0) \\ &= a_{i1} \times 0 + a_{i2} \times 0 + \dots + a_{ij} \times 1 + \dots + a_{im} \times 0 \\ &= a_{ij}\end{aligned}$$

Como isto se verifica para todos os valores de i e j, a matriz $A\mathbf{I_m}$ é precisamente a matriz A.

Para o produto $\mathbf{I_n}A$ a situação é análoga: a sua entrada (i,j) é o produto interno da linha i de $\mathbf{I_n}$ (um vector que tem 0 em todas as entradas excepto na i-ésima) pela coluna j de A. Por um raciocínio semelhante, o valor deste produto interno é precisamente a_{ij}, pelo que mais uma vez o resultado do produto é a matriz A.

5. A entrada (i,j) da matriz $(A+B)C$ é dada pelo produto interno entre a linhha i de $(A+B)$ e a coluna j de C. Tendo em conta a definição de soma de matrizes e as propriedades das operações com números reais, podemos facilmente verificar que

$$\begin{aligned}((A+B)C)_{ij} &= (A+B)_{i1}c_{1j} + (A+B)_{i2}c_{2j} + \dots + (A+B)_{im}c_{mj} \\ &= (a_{i1}+b_{i1})\,c_{1j} + (a_{i2}+b_{i2})\,c_{2j} + \dots + (a_{im}+b_{im})\,c_{mj} \\ &= a_{i1}c_{1j} + b_{i1}c_{1j} + a_{i2}c_{2j} + b_{i2}c_{2j} + \dots + a_{im}c_{mj} + b_{im}c_{mj} \\ &= (a_{i1}c_{1j} + a_{i2}c_{2j} + \dots + a_{im}c_{mj}) + (b_{i1}c_{1j} + b_{i2}c_{2j} + \dots + b_{im}c_{mj}) \\ &= (AC)_{ij} + (BC)_{ij}\end{aligned}$$

tendo em conta que as somas na penúltima expressão correspondem precisamente aos produtos internos da linha i de A com a coluna j de C e da linha i de B com a coluna j de C.

□

Tal como se passava com as operações discutidas na secção anterior, o principal interesse destas propriedades é permitir simplificar expressões envolvendo matrizes antes de efectuar os cálculos.

Exercício 33. Considere as seguintes matrizes.

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

Calcule o resultado das seguintes expressões, simplificando-as se possível antes de efectuar os cálculos.

(a) $AC + BC$

(b) $CD - A$

(c) $3(A + B)C - 2AC + B(-\mathbf{I_2}C)$

(d) $\left((CD)^T - D^TC^T\right)(AB - BA)$

(e) $C^TA^T + C^TB^T$

(f) $(\mathbf{I_2 0_{2\times 3} I_3} + D)(A + B + CD)$

Uma particularidade interessante do produto de matrizes é a existência de *divisores de zero*. Nos números reais, verifica-se a lei do anulamento do produto: se o produto de dois números é zero, então pelo menos um deles é zero.

$$xy = 0 \Longrightarrow x = 0 \text{ ou } y = 0$$

No cálculo com matrizes, a situação é radicalmente diferente: existem matrizes não nulas cujo produto é zero; um exemplo bastante simples é o produto

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{0_{2\times 2}},$$

mas existem matrizes com todas as entradas diferentes de zero cujo produto é a matriz nula.

Exemplo. Considere-se as matrizes $\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}$ e $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$. É fácil verificar que o seu produto é a matriz $\mathbf{0_{2\times 2}}$.

Exercício 34. Verifique que o produto das matrizes $\begin{bmatrix} 1 & -1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$ e $\begin{bmatrix} -9 & 18 \\ -1 & 2 \\ 4 & -8 \end{bmatrix}$ também é a matriz nula.

A partir do produto de matrizes, é natural definir potências de forma semelhante às potências dos números reais: por iteração da multiplicação. Por forma a que uma matriz possa ser multiplicada por si própria, tem de ter o mesmo número de linhas e colunas; ou seja, a potência de matrizes só está definida para matrizes quadradas.

Definição. Seja A uma matriz quadrada de dimensão n e seja p um número natural. A potência A^p de A define-se como

$$A^p = \underbrace{A \cdots A}_{p \text{ vezes}}$$

ou, em alternativa, de forma indutiva por

$$A^0 = \mathbf{I_n}$$
$$A^{p+1} = A\ A^p$$

Exemplo. Seja A a matriz

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

Para calcular A^4, temos de começar por calcular A^3, que requer calcular antes A^2.

$$A^2 = A\ A = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} \qquad A^3 = A\ A^2 = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix} \qquad A^4 = A\ A^3 = \begin{bmatrix} 1 & 8 \\ 0 & 1 \end{bmatrix}$$

Exercício 35. Calcule as potências das seguintes matrizes.

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

Exercício 36. Seja A uma matriz quadrada de dimensão 1. Quanto vale a única entrada de A^n?

A operação de potenciação de matrizes satisfaz muitas das propriedades habituais das potências de números reais.

Proposição 6. Seja A uma matriz quadrada de dimensão n. Então verificam-se as seguintes igualdades.

1. $A^n A^m = A^{n+m}$
2. $(A^n)^m = A^{nm}$
3. $(A^n)^T = \left(A^T\right)^n$

Contudo, é importante salientar que muitas outras propriedades *não* se verificam, devido à não comutatividade do produto de matrizes. Por exemplo, $(AB)^2 = (AB)(AB) \neq A^2B^2$.

Exercício 37. Sejam $A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$. Verifique que:

(a) $(AB)^2 \neq A^2B^2$ (b) $(A+B)^2 \neq A^2 + 2AB + B^2$

1.4 Inversão de matrizes

Sejam A e B as duas matrizes seguintes.

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

Ao calcular os produtos AB e BA, verifica-se que ambos são iguais à matriz identidade $\mathbf{I_2}$.

Definição. Sejam A e B duas matrizes quadradas de dimensão n. Se $AB = BA = \mathbf{I_n}$, então as matrizes A e B dizem-se *inversas*. Nesta situação, é frequente escrever $B = A^{-1}$.

O conceito de matriz inversa é extremamente útil e essencial ao trabalho com matrizes. Contudo, e mais uma vez ao contrário do que se passa com os números reais, existem matrizes quadradas que não têm inversa, como as matrizes $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ e $\begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}$.

Definição. Uma matriz quadrada A de dimensão n diz-se *invertível* se existir uma matriz B tal que A e B são matrizes inversas.

Se não existir nenhuma matriz B nestas condições, a matriz A diz-se *não invertível* ou *singular*.

No seguimento veremos como determinar se uma matriz é invertível e daremos exemplos de matrizes singulares de dimensões superiores a 2.

Exemplo. A matriz $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ é invertível, porque tomando $B = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$ tem-se $AB = BA = \mathbf{I_2}$.

1.4.1 Propriedades

A notação A^{-1} sugere que se fale *na* matriz inversa de A. De facto, uma matriz não pode ter duas inversas distintas.

Proposição 7. Se uma matriz A é invertível, então a sua inversa é única.

Demonstração. Seja A uma matriz quadrada de dimensão n invertível e suponha-se que B e C são duas matrizes inversas de A. Por um lado, se B é inversa de A, então tem-se

$$AB = BA = \mathbf{I_n}\,.$$

Por outro lado, como C também é inversa de A, tem-se

$$AC = CA = \mathbf{I_n}\,.$$

Consideremos o produto BAC. Usando a associatividade do produto, podemos efectuar primeiro quer a multiplicação da esquerda, quer a da direita. No primeiro caso, obtemos $(BA)C = \mathbf{I_n}C = C$; no segundo, $B(AC) = B\mathbf{I_n} = B$. Mas estas expressões são iguais, pelo que necessariamente $B = C$. □

Antes de nos debruçarmos sobre o cálculo de matrizes inversas, vamos ver algumas propriedades simples da relação da inversão de matrizes com as operações anteriormente estudadas.

Proposição 8. Sejam A e B duas matrizes invertíveis de dimensão n, p um número natural e $\alpha \neq 0$ um número real. Então:

1. AB é uma matriz invertível e $(AB)^{-1} = B^{-1}A^{-1}$;
2. A^{-1} é invertível e $(A^{-1})^{-1} = A$;
3. A^p é invertível e $(A^p)^{-1} = (A^{-1})^p$, sendo usual denotar esta matriz por A^{-p};
4. αA é invertível e $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$;
5. A^T é invertível e $\left(A^T\right)^{-1} = (A^{-1})^T$.

Demonstração. Uma vez que A e B são matrizes invertíveis, existem A^{-1} e B^{-1} tais que $AA^{-1} = A^{-1}A = \mathbf{I_n}$ e $BB^{-1} = B^{-1}B = \mathbf{I_n}$.

1. Tome-se $X = B^{-1}A^{-1}$. Então

$$\begin{aligned}(AB)X &= (AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = A\mathbf{I_n}A^{-1} = AA^{-1} = \mathbf{I_n}\\ X(AB) &= \left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}\mathbf{I_n}B = B^{-1}B = \mathbf{I_n}\end{aligned}$$

   donde X é inversa de AB. Logo AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.

2. Da relação $AA^{-1} = A^{-1}A = \mathbf{I_n}$ conclui-se que A é a inversa de A^{-1}, ou seja, $(A^{-1})^{-1} = A$.

3. Intuitivamente, sabemos que

$$A^p = \underbrace{A\cdots A}_{p \text{ vezes}}.$$

   Multiplicando p vezes por A^{-1}, obtemos produtos que sabemos corresponderem a matrizes identidade:

$$\begin{aligned}A^p\left(A^{-1}\right)^p &= \underbrace{A\cdots A}_{p \text{ vezes}}\underbrace{A^{-1}\cdots A^{-1}}_{p \text{ vezes}} = \underbrace{A\cdots A}_{p-1 \text{ vezes}}\;\mathbf{I_n}\underbrace{A^{-1}\cdots A^{-1}}_{p-1 \text{ vezes}}\\ &= \underbrace{A\cdots A}_{p-1 \text{ vezes}}\;\underbrace{A^{-1}\cdots A^{-1}}_{p-1 \text{ vezes}} = \cdots = \mathbf{I_n}\end{aligned}$$

   Uma alternativa um pouco mais rigorosa desta propriedade é recorrer ao método de indução.

   Para $p = 1$, obtemos simplesmente $\left(A^1\right)^{-1} = A^{-1} = \left(A^{-1}\right)^1$.

   Supondo que o resultado é válido para p, vamos ver que continua válido para $p+1$. Pela definição de potência, temos que $A^{p+1} = AA^p$ e, pela primeira propriedade, sabemos que $\left(AA^p\right)^{-1} = \left(A^p\right)^{-1}A^{-1}$. Ora estamos a assumir que $\left(A^p\right)^{-1} = (A^{-1})^p$, donde

$$\left(A^{p+1}\right)^{-1} = \left(A^p\right)^{-1}A^{-1} = \left(A^{-1}\right)^pA^{-1} = \left(A^{-1}\right)^{p+1}.$$

4. Pelas propriedades do produto de escalares por matrizes, temos que

$$(\alpha A)\left(\frac{1}{\alpha}A^{-1}\right) = \left(\alpha\frac{1}{\alpha}\right)\left(AA^{-1}\right) = 1\mathbf{I_n} = \mathbf{I_n}.$$

   Analogamente, verifica-se que também $\left(\frac{1}{\alpha}A^{-1}\right)(\alpha A) = \mathbf{I_n}$, donde αA é invertível e a sua inversa é dada por $\frac{1}{\alpha}A^{-1}$.

5. Pelas propriedades da transposição de matrizes, $A^T\left(A^{-1}\right)^T = \left(A^{-1}A\right)^T = \mathbf{I_n}^T = \mathbf{I_n}$.
   Analogamente, $\left(A^{-1}\right)^T A^T = \mathbf{I_n}$. Tal como atrás, concluímos que A^T é invertível e a sua inversa é $\left(A^{-1}\right)^T$.

$\square$

Exercício 38. Nos exemplos acima, encontrámos já duas matrizes invertíveis:

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

Usando as propriedades da inversão de matrizes, calcule as inversas das seguintes matrizes.

(a) AB
(b) BA^2
(c) $(BA)^2$
(d) $A^{-1}BA$
(e) $(AB)B^{-1}$
(f) $\left(A^2\right)^{-1}A$
(g) $3AA^T$
(h) $-B^2$

1.4.2 Inversão de matrizes elementares

Na resolução de sistemas de equações lineares, definimos operações elementares sobre matrizes: troca de linhas, produto duma linha por um escalar e soma duma linha com um múltiplo de outra. Estas operações dão origem a uma classe de matrizes a que chamamos matrizes elementares, que são particularmente úteis no contexto da inversão de matrizes porque são muito simples de inverter.

Definição. Uma matriz quadrada de dimensão n diz-se uma *matriz elementar* se pode ser obtida a partir da matriz identidade $\mathbf{I_n}$ por aplicação de uma única operação elementar.

Sendo $*$ uma operação elementar, existe uma única matriz elementar E obtida por aplicação de $*$ à matriz $\mathbf{I_n}$. Esta matriz diz-se a matriz elementar *associada* a $*$.

Por exemplo: considerando as operações elementares

$$*_1 = l_1 \leftrightarrow l_2 \qquad *_2 = 4l_3 \qquad *_3 = l_1 + 5l_2$$

e designando as matrizes elementares correspondentes a $*_1$, $*_2$ e $*_3$ por E_1, E_2 e E_3, respectivamente, temos que

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercício 39. Escreva as matrizes elementares de dimensão 3 associadas a cada uma das seguintes operações elementares.

(a) $l_1 \leftrightarrow l_3$
(b) $l_3 \to 2l_3$
(c) $l_3 \to l_3 + 3l_1$
(d) $l_3 \to l_3 - 2l_2$
(e) $l_2 \to -2l_2$
(f) $l_2 \leftrightarrow l_3$
(g) $l_2 \to l_2 - 2l_3$
(h) $l_1 \to 5l_1$

Exercício 40. Quais das seguintes matrizes são elementares? Quais as operações elementares a que correspondem?

(a) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ (i) $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ (m) $\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ (j) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (n) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (g) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (k) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ (o) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (h) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (l) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (p) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

A importância das matrizes elementares vem de permitirem representar as operações elementares de forma algébrica. É esse o sentido da seguinte propriedade, cuja verificação é simples.

Proposição 9. Sejam A uma matriz e $*$ uma operação elementar com matriz elementar associada E_*. Então a matriz que se obtém de A por aplicação de $*$ é precisamente E_*A.

Esta propriedade é fundamental para muitas aplicações e constitui também um dos primeiros exemplos de aplicação do produto de matrizes.

Uma outra propriedade importante das operações elementares é que toda a operação elementar tem uma operação elementar inversa:

- a inversa de $l_i \leftrightarrow l_j$ é $l_i \leftrightarrow l_j$;
- a inversa de $l_i \to \alpha l_i$ é $l_i \to \frac{1}{\alpha} l_i$;
- a inversa de $l_i \to l_i + \alpha l_j$ é $l_i \to l_i - \alpha l_j$.

Dito doutra forma:

- se trocarmos linhas duma matriz, para recuperar a matriz original temos de voltar a trocar as mesmas linhas;
- se multiplicarmos uma linha duma matriz por uma constante, para recuperar a matriz original temos de dividir a mesma linha pela mesma constante;
- se a uma dada linha duma matriz adicionamos um múltiplo de outra linha, para recuperar a matriz original temos de subtrair o mesmo multiplo da linha que lhe tinhamos adicionado.

Exercício 41. Indique as operações elementares inversas das operações do Exercicio 39.

É fácil verificar que se $*$ for uma operação elementar associada à matriz elementar E, então a matriz elementar associada à inversa da operação $*$ é precisamente a matriz inversa E^{-1} de E. Dito de outra forma, a inversa de uma matriz elementar associada a uma operação elementar $*$ é a matriz elementar associada à operação inversa de $*$.

Exemplo. Consideremos a operação elementar $l_2 \to l_2 \to 2{+}3l_3$. A matriz elementar associada a esta operação é

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

A operação elementar inversa é $l_2 - 3l_3$, à qual está associada a matriz elementar

$$E^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

e é simples verificar que de facto se tem

$$EE^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I_4}$$

e

$$E^{-1}E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I_4}\,.$$

Ou seja, E e E^{-1} são de facto inversas uma da outra.

Estas observações podem ser sintetizadas no seguinte resultado.

Proposição 10. Toda a matriz elementar é invertível, sendo a sua matriz inversa uma matriz elementar.

Exemplo. Consideremos a matriz

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

Para determinar a inversa de A, podemos começar por observar que A é a matriz elementar associada à operação $l_2 + 2l_3$. Sendo uma matriz elementar, o resultado anterior garante que

A é invertível, sendo a sua inversa a matriz elementar associada à operação inversa de $l_2 + 2l_3$, que é $l_2 - 2l_3$. Então

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

Exercício 42. Verifique que as matrizes do exemplo anterior são de facto inversas uma da outra.

Exercício 43. Calcule as matrizes inversas das matrizes determinadas no Exercicio 39.

Exercício 44. Calcule as inversas das seguintes matrizes.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$ (g) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (h) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

1.4.3 Método de inversão de Gauss–Jordan

O estudo efectuado para a inversão de matrizes elementares vai-nos permitir resolver o caso geral. Para inverter uma matriz quadrada arbitrária, vamos considerar uma extensão do método de eliminação de Gauss, conhecida como o *método de Gauss–Jordan*. Enquanto o método de eliminação de Gauss transforma uma matriz arbitrária numa matriz em escada de linhas por aplicação de operações elementares, o método de eliminação de Gauss–Jordan transforma uma matriz arbitrária numa matriz em escada de linhas reduzida — e permite decidir rapidamente se a matriz é ou não invertível e, caso o seja, encontrar a sua inversa.

Recordemos os conceitos de matriz em escada de linhas e de pivot, introduzidos na Secção 1.1. Numa matriz em escada de linhas, qualquer pivot é simultaneamente o *último* elemento não nulo da coluna em que se encontra. Por isso, é também comum falar-se do pivot duma coluna.

Definição. Uma matriz A diz-se uma *matriz em escada de linhas reduzida* se for uma matriz em escada de linhas satisfazendo as duas condições seguintes.

- Todos os pivots são iguais a 1.

- As colunas com pivots têm todas as restantes entradas nulas.

Exemplo. A matriz

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

é uma matriz em escada de linhas que não é reduzida: a segunda e a quarta colunas contêm elementos diferentes de zero para além dos pivots. Já a matriz

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

é uma matriz em escada de linhas reduzida: a única coluna com mais do que um elemento não nulo é a terceira, que não contém nenhum pivot.

Exercício 45. Quais das seguintes matrizes são em escada de linhas reduzida?

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

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

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

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

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

(f) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

Vamos agora ver como transformar uma matriz numa matriz em escada de linhas reduzida. Recorde-se que, no final do método de eliminação de Gauss, se obtinha uma matriz em escada de linhas. Nessa matriz, partindo do último pivot, é possível eliminar todos os outros elementos nessa coluna usando apenas operações elementares. De seguida, pode-se proceder de forma semelhante para o penúltimo pivot: a linha que o contém vai necessariamente ter um 0 na coluna correspondente ao último pivot. Prosseguindo para a esquerda, em cada passo eliminamos todos os outros elementos da coluna contendo um pivot. Finalmente, multiplicando cada linha pelo valor adequado consegue-se garantir que todos os pivots ficam iguais a 1.

Tal como sucedia com o método de eliminação de Gauss, o método de Gauss–Jordan usa apenas operações elementares sobre matrizes.

Vejamos um exemplo. Consideremos a matriz

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

Aplicando o método de eliminação de Gauss, chegamos à seguinte matriz em escada de linhas.

$$\begin{bmatrix} 2 & 1 & -2 & 0 \\ 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{array}{c} \\ 2l_2 - l_1 \\ \\ \end{array} \longrightarrow \begin{bmatrix} 2 & 1 & -2 & 0 \\ 0 & -1 & 0 & 4 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{array}{c} \\ \\ l_2 + l_3 \end{array} \longrightarrow \begin{bmatrix} 2 & 1 & -2 & 0 \\ 0 & -1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

O último pivot é o 1 da terceira coluna. Para eliminar o -2 da primeira linha é necessário substituir l_1 por $l_1 + 2l_3$.

$$\begin{bmatrix} 2 & 1 & -2 & 0 \\ 0 & -1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix} \begin{array}{c} l_1 + 2l_3 \\ \longrightarrow \\ \end{array} \begin{bmatrix} 2 & 1 & 0 & 10 \\ 0 & -1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

O penúltimo pivot é o -1 da segunda linha; observe-se que esta linha já tem um 0 na terceira coluna. Para eliminar o 1 da primeira linha, vamos substituí-la por $l_1 + l_2$.

$$\begin{bmatrix} 2 & 1 & 0 & 10 \\ 0 & -1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix} \begin{array}{c} l_1 + l_2 \\ \longrightarrow \\ \end{array} \begin{bmatrix} 2 & 0 & 0 & 14 \\ 0 & -1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

Finalmente, dividimos a primeira linha por dois e trocamos o sinal da segunda para obter uma matriz em escada de linhas reduzida.

$$\begin{bmatrix} 2 & 0 & 0 & 14 \\ 0 & -1 & 0 & 4 \\ 0 & 0 & 1 & 5 \end{bmatrix} \begin{array}{c} \frac{1}{2}l_1 \\ -l_2 \longrightarrow \\ \end{array} \begin{bmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

Exercício 46. Aplique o método de Gauss–Jordan às seguintes matrizes.

(a) $\begin{bmatrix} 2 & 1 \\ 1 & -2 \\ 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 6 \\ 8 & -1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 1 & -4 & 3 \\ 0 & 2 & -1 & 2 \end{bmatrix}$

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

O interesse principal deste método é permitir caracterizar rapidamente as matrizes invertíveis.

Teorema 1. Seja A uma matriz quadrada de dimensão n. Então as seguintes afirmações são equivalentes:

i) A é invertível;

ii) o sistema de equações lineares $A\underline{x} = \mathbf{0_{n\times 1}}$ (em que $\underline{x}$ é um vector coluna com n entradas) tem solução única e trivial $\underline{x} = \mathbf{0_{n\times 1}}$;

iii) a matriz em escada de linhas reduzida de A é $\mathbf{I_n}$.

A terceira condição equivale a dizer que, no final da aplicação do método de eliminação de Gauss–Jordan a uma matriz invertível, se obtém a matriz identidade.

Exemplo. Seja A a matriz

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

Para ver se a matriz A é invertível, vamos aplicar o método de eliminação de Gauss–Jordan para transformar A numa matriz em escada de linhas reduzida, verificando no final se esta matriz é a identidade.

Vimos já que o primeiro passo é aplicar o método de eliminação de Gauss.

$$\begin{bmatrix} \mathbf{1} & 2 & 3 \\ 2 & 3 & 7 \\ 1 & 0 & 8 \end{bmatrix} \begin{array}{l} \\ l_2 - 2l_1 \\ l_3 - l_1 \end{array} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & \mathbf{-1} & 1 \\ 0 & -2 & 5 \end{bmatrix} \begin{array}{l} \\ \\ l_3 - 2l_2 \end{array} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 0 & 0 & 3 \end{bmatrix}$$

Agora vamos eliminar os elementos acima dos pivots, começando pela coluna da direita.

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 0 & 0 & \mathbf{3} \end{bmatrix} \begin{array}{l} l_1 - l_3 \\ 3l_2 - l_3 \\ \\ \end{array} \longrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & \mathbf{-3} & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{array}{l} 3l_1 + 2l_2 \\ \\ \\ \end{array} \longrightarrow \begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Finalmente, dividimos cada linha pelo seu pivot.

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{array}{l} \frac{1}{3}l_1 \\ -\frac{1}{3}l_2 \\ \frac{1}{3}l_3 \end{array} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Como a matriz em escada de linhas reduzida obtida é a identidade $\mathbf{I_3}$, concluimos que A é invertível.

Exercício 47. Quais das matrizes do Exercicio 46 são invertíveis?

Pelo teorema anterior, se A é uma matriz invertível então existe uma sequência de operações elementares que, quando aplicada à matriz A, a transforma na matriz identidade.

$$A \xrightarrow{*_1} A_1 \xrightarrow{*_2} A_2 \longrightarrow \cdots \xrightarrow{*_k} \mathbf{I_n}$$

Ora vimos na subsecção anterior que aplicar uma operação elementar à matriz A corresponde a multiplicar a matriz elementar correspondente a essa operação pela matriz A. Estamos portanto a dizer que existem matrizes elementares E_1, E_2, …, E_k tais que

$$E_k \cdots E_2 E_1 A = \mathbf{I_n} \,,$$

donde A é invertível e tem inversa $E_k \cdots E_2 E_1$.

Por sua vez, esta matriz pode ser calculada aplicando as mesmas transformações elementares à matriz identidade $\mathbf{I_n}$. Obtemos assim um algoritmo baseado no método de Gauss–Jordan para calcular a matriz inversa de A: aplicamos o método de eliminação de Gauss–Jordan à matriz aumentada $[A|\mathbf{I_n}]$.

O método termina assim que a matriz A fica reduzida à sua matriz em escada de linhas reduzida. Mas se A for invertível sabemos que se reduz à identidade, pelo que a matriz aumentada terá a forma $[\mathbf{I_n}|B]$; neste caso, a matriz B é a inversa de A: $B = A^{-1}$. Se a matriz A não for invertível, haverá uma linha só com 0s na parte esquerda da matriz aumentada.

Exemplo. Continuando o exemplo anterior, já vimos que a matriz A aí apresentada é invertível. Vamos agora calcular a sua inversa.

$$\left[\begin{array}{ccc|ccc} \mathbf{1} & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 7 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ l_2 - 2l_1 \\ l_3 - l_1 \end{array} \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & \mathbf{-1} & 1 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \\ l_3 - 2l_2 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 1 & -2 & 1 & 0 \\ 0 & 0 & \mathbf{3} & 3 & -2 & 1 \end{array}\right] \begin{array}{l} l_1 - l_3 \\ 3l_2 - l_3 \\ \\ \end{array} \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 2 & -1 \\ 0 & \mathbf{-3} & 0 & -9 & 5 & -1 \\ 0 & 0 & 3 & 3 & -2 & 1 \end{array}\right] \begin{array}{l} 3l_1 + 2l_2 \\ \\ \\ \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 3 & 0 & 0 & -24 & 16 & -5 \\ 0 & -3 & 0 & -9 & 5 & -1 \\ 0 & 0 & 3 & 3 & -2 & 1 \end{array}\right] \begin{array}{l} \frac{1}{3}l_1 \\ -\frac{1}{3}l_2 \\ \frac{1}{3}l_3 \end{array} \longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -8 & \frac{16}{3} & -\frac{5}{3} \\ 0 & 1 & 0 & 3 & -\frac{5}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 1 & -\frac{2}{3} & \frac{1}{3} \end{array}\right]$$

Logo

$$A^{-1} = \begin{bmatrix} -8 & \frac{16}{3} & -\frac{5}{3} \\ 3 & -\frac{5}{3} & \frac{1}{3} \\ 1 & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

Exemplo. Suponhamos agora que queremos inverter a matriz

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

Aplicando o método de eliminação de Gauss–Jordan à matriz aumentada $[B|\mathbf{I_3}]$ obtemos sucessivamente

$$\left[\begin{array}{ccc|ccc} \mathbf{2} & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ 2l_2 + l_1 \\ \\ \end{array} \longrightarrow \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 1 & 2 & 0 \\ 0 & 0 & \mathbf{3} & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ 3l_2 - 2l_3 \\ \\ \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & \mathbf{9} & 0 & 3 & 6 & -2 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} 9l_1 - l_2 \\ \\ \\ \end{array} \longrightarrow \left[\begin{array}{ccc|ccc} 18 & 0 & 0 & 6 & -6 & 2 \\ 0 & 9 & 0 & 3 & 6 & -2 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \frac{1}{18}l_1 \\ \frac{1}{9}l_2 \\ \frac{1}{3}l_3 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{9} \\ 0 & 1 & 0 & \frac{1}{3} & \frac{2}{3} & -\frac{2}{9} \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{3} \end{array}\right]$$

donde

$$B^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{9} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{9} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

Exercício 48. Aplique o método de Gauss–Jordan para determinar quais das seguintes matrizes são invertíveis, indicando as inversas das que o forem.

(a) $\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 2 & 6 \\ -1 & -3 \end{bmatrix}$

(e) $\begin{bmatrix} 2 & 6 \\ 1 & -3 \end{bmatrix}$
(g) $\begin{bmatrix} 1 & -3 & 4 \\ 0 & 0 & -2 \\ 1 & 2 & 5 \end{bmatrix}$
(i) $\begin{bmatrix} -1 & 0 & 0 & 1 \\ 2 & 0 & 0 & -2 \\ 0 & 1 & 2 & -1 \\ -1 & 1 & 1 & 0 \end{bmatrix}$
(k) $\begin{bmatrix} 2 & 1 & 5 \\ 1 & 1 & -1 \\ -2 & 0 & 0 \end{bmatrix}$

(f) $\begin{bmatrix} 2 & 3 & 1 \\ 3 & 0 & -2 \\ -2 & 1 & 2 \end{bmatrix}$
(h) $\begin{bmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 1 & 1 & 1 & -2 \end{bmatrix}$
(j) $\begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & -5 \\ 3 & 0 & 2 \end{bmatrix}$
(l) $\begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & -5 \\ -3 & 0 & -3 \end{bmatrix}$

1.5 Determinantes

O conceito de determinante duma matriz é central a diversas aplicações da Álgebra Linear, surgindo com frequência em variadíssimos contextos. Assim, mais do que conhecer fórmulas de cálculo de determinantes, importa ter alguma intuição sobre o seu significado e um razoável à vontade com as suas propriedades.

1.5.1 Definição e propriedades

Há várias formas de introduzir o conceito de determinante duma matriz. Nesta exposição, vamos apresentá-lo através da sua interpretação geométrica, que será relevante em particular nas aplicações à Análise Matemática, por forma a obter intuição sobre as suas propriedades.

Definição. O *determinante* duma matriz quadrada A de dimensão $n \times n$ é um número real cujo módulo corresponde ao volume do sólido n-dimensional que tem as linhas de A por arestas.

Exemplo. Considerem-se as matrizes seguintes.

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

Então $|\det(A)|$ é a área do paralelogramo à esquerda na Figura 1.6, enquanto $|\det(B)|$ é o volume do paralelipípedo à direita.

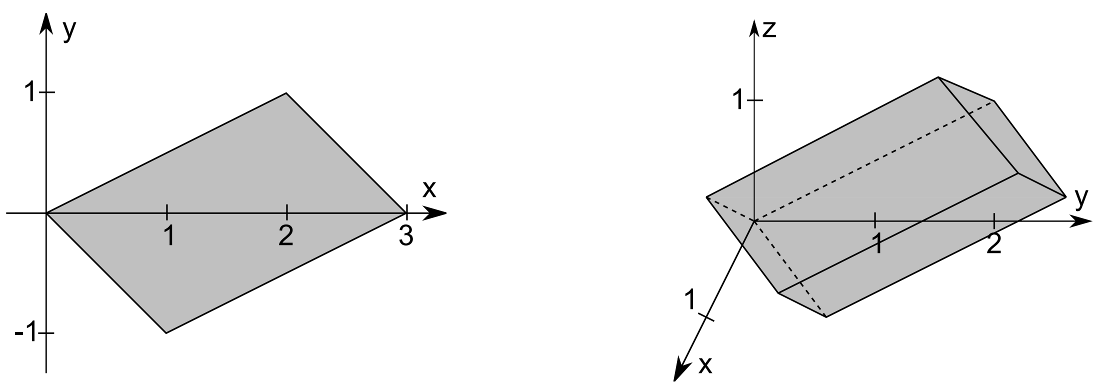


Figura 1.6: Interpretação geométrica do determinante.

Em dimensões superiores a 3, esta interpretação geométrica mantém-se válida, embora não seja possível visualizá-la.

Para começarmos a calcular determinantes de matrizes, comecemos por considerar um caso simples. Se A for uma matriz diagonal, então a figura geométrica definida pelas suas linhas tem arestas paralelas aos eixos coordenados, pelo que a sua área é simplesmente o produto dos valores na diagonal da matriz. A Figura 1.7 exemplifica esta situação para as matrizes C e D seguintes.

$$C = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{7}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

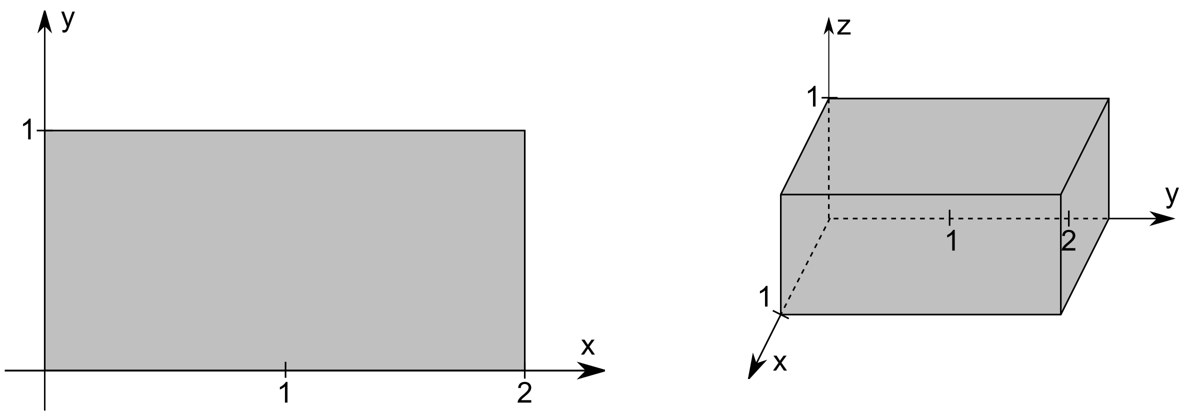


Figura 1.7: Figuras geométricas correspondendo às matrizes C e D.

Proposição 11. O determinante duma matriz diagonal é o produto dos elementos na diagonal da matriz.

Exemplo.

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

$$\det \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} = 2 \times 1 \times \frac{1}{2} = 1$$

$$\det \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 2 \times 1 \times 3 \times 1 = 6$$

O último exemplo acima mostra como podemos generalizar as intuições obtidas para os casos de matrizes 2×2 e 3×3 a dimensões superiores.

Exercício 49. Indique o valor do determinante das seguintes matrizes.

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

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

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

(d) $\begin{bmatrix} 6 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$

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

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

(g) $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$

(h) $\begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{3}{5} \end{bmatrix}$

(i) $\begin{bmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{3}{2} & 0 \\ 0 & 0 & -2 \end{bmatrix}$

(j) $\begin{bmatrix} \pi & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \frac{7}{4} \end{bmatrix}$

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

(l) $\begin{bmatrix} -3 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & \frac{2}{5} & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$

(m) $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$

Por outro lado, a definição apresentada permite ter determinantes negativos; por exemplo:

$$\det \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = -1\,.$$

Como se pode interpretar este valor?

Em primeiro lugar, na definição de determinante foi dito que o seu *módulo* correspondia à área da figura geométrica definida pelas linhas da matriz. Claramente isto aplica-se neste caso: a figura geométrica correspondente àquela matriz é um quadrado de lado 1 e $|-1| = 1$.

O que se passa, porém, é que o valor do determinante corresponde a um *volume orientado*, situação que ocorre também na correspondência entre áreas e integrais definidos. No caso de $\mathbb{R}^2$, diz-se que um ângulo tem orientação positiva se corresponde ao percurso entre as semi-rectas que o delimitam, no sentido contrário ao dos ponteiros do relógio. O rectângulo definido pelos vectores $(1,0)$ e $(0,-1)$ tem orientação negativa: para ir do primeiro lado para o segundo, segue-se no sentido dos ponteiros do relógio. Assim, a área orientada do rectângulo tem sinal negativo.

Em dimensões superiores, o conceito de orientação também existe, embora seja um pouco mais complicado de definir. Por convenção, os eixos estão sempre ordenados da forma que corresponde à orientação positiva — justificando assim a fórmula acima apresentada para matrizes diagonais. Em $\mathbb{R}^3$, a orientação é definida pela "regra da mão direita", muito usada na Física.

Por outro lado, do que acima foi dito não surpreende a propriedade seguinte.

Proposição 12. Sejam A e B matrizes tais que B se obtém de A por troca de duas linhas. Então $\det(B) = -\det(A)$.

No caso de $\mathbb{R}^2$, a troca de linhas corresponde a trocar a ordem dos lados do rectângulo, mudando a sua orientação. É possível verificar que o mesmo se passa em $\mathbb{R}^3$, justificando por analogia o resultado para dimensões superiores.

Exemplo.

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

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

Note-se que este resultado permite que se estabeleçam relações entre determinantes de matrizes, mesmo que não se consiga (para já) calculá-los.

Exemplo.

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

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

Exercício 50. Sabendo que

$$\det\begin{bmatrix}2 & 1 & -2\\ 0 & 1 & 2\\ 1 & -1 & 0\end{bmatrix} = 8\,,$$

indique o valor do determinante das seguintes matrizes.

(a) $\begin{bmatrix}1 & -1 & 0\\ 2 & 1 & -2\\ 0 & 1 & 2\end{bmatrix}$ (b) $\begin{bmatrix}2 & 1 & -2\\ 1 & -1 & 0\\ 0 & 1 & 2\end{bmatrix}$ (c) $\begin{bmatrix}0 & 1 & 2\\ 1 & -1 & 0\\ 2 & 1 & -2\end{bmatrix}$ (d) $\begin{bmatrix}0 & 1 & 2\\ 2 & 1 & -2\\ 1 & -1 & 0\end{bmatrix}$

Outra propriedade importante é a linearidade: o que acontece ao determinante duma matriz se uma das suas linhas for multiplicada por uma constante ou somada com outro vector?

É simples responder à primeira questão. Multiplicar uma linha duma matriz A por um número real corresponde a esticar ou encolher a figura geométrica correspondente ao longo duma das suas direcções (Figura 1.8).

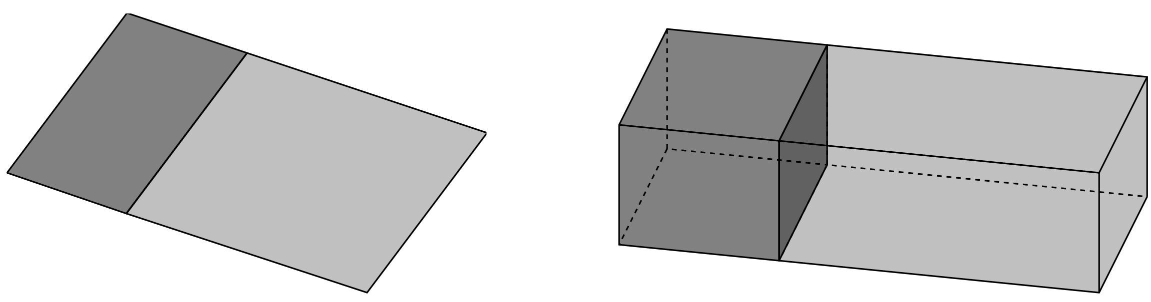


Figura 1.8: Relação entre o determinante e a multiplicação duma linha por um escalar.

Tendo em conta que as áreas e volumes são directamente proporcionais ao comprimento das arestas, obtém-se o seguinte resultado.

Proposição 13. Sejam A e B matrizes tais que B se obtém de A por multiplicação da linha i pelo número real c. Então $\det(B) = c \times \det(A)$.

Em particular, uma matriz que tenha uma linha só com zeros tem determinante nulo — facto que podia ser justificado directamente com um argumento geométrico.

Exercício 51. Sabendo que

$$\det \begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} = 8\,,$$

indique o valor do determinante das seguintes matrizes.

(a) $\begin{bmatrix} 2 & 1 & -2 \\ 0 & -2 & -4 \\ 1 & -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 & -2 \\ 0 & 3 & 6 \\ 2 & -2 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} \frac{2}{5} & \frac{1}{5} & -\frac{2}{5} \\ 1 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 2 & 4 \\ 3 & -3 & 0 \\ 2 & 1 & -2 \end{bmatrix}$

Considere-se agora o outro caso, em que se adiciona um vector a uma linha da matriz. A Figura 1.9 ilustra esta situação para o caso de $\mathbb{R}^2$: obtém-se um paralelogramo cuja área é igual à soma das áreas de dois paralelogramos — o original, e aquele definido substituindo a linha em causa pelo vector adicionado. No exemplo, os paralelogramos da esquerda correspondem às matrizes

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$$

e o da direita à matriz

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

cuja primeira linha é a soma das primeiras linhas das matrizes anteriores.

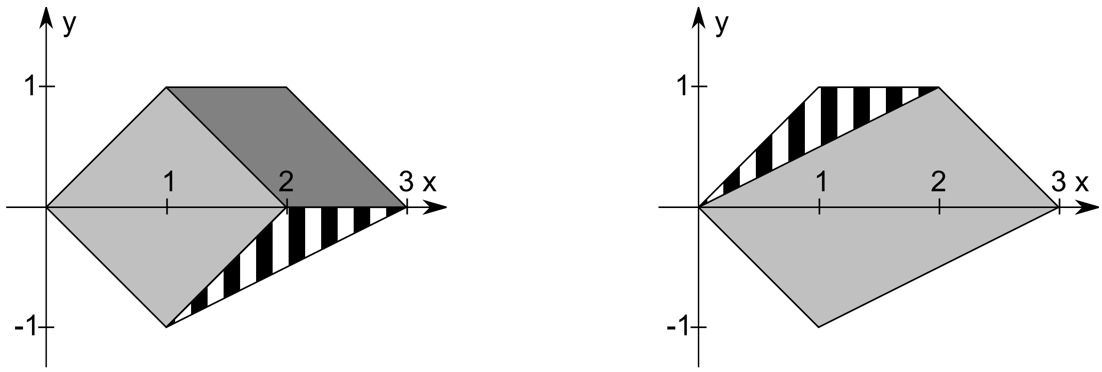


Figura 1.9: Soma de um vector a uma linha. Observe-se que os triângulos assinalados com riscas verticais são idênticos, pelo que a área total sombreada em ambas as figuras é idêntica.

Proposição 14. Sejam A e B matrizes cujas linhas são todas iguais excepto a linha i. Então $\det(A) + \det(B) = \det(C)$, onde C é a matriz com todas as linhas iguais às de A e B excepto a linha i, que é a soma das linhas i de A e de B.

Exercício 52. Sabendo que

$$\det\begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} = 8 \qquad \det\begin{bmatrix} 2 & 1 & -2 \\ 2 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} = -4 \qquad \det\begin{bmatrix} 2 & 1 & -2 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} = 3\,,$$

indique o valor do determinante das seguintes matrizes.

(a) $\begin{bmatrix} 2 & 1 & -2 \\ 2 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 3 \\ 1 & -1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 1 & -2 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$

Esta propriedade é fundamental para a obtenção duma fórmula de cálculo de determinantes. Observe-se que *qualquer* vector pode ser escrito como a soma de vectores em que todas as componentes são nulas, excepto uma. Por exemplo, para a matriz anterior, $(2,1) = (2,0)+(0,1)$ e $(1,-1) = (1,0)+(0,1)$. Então tem-se sucessivamente, aplicando a propriedade da soma primeiro à linha de cima e depois à linha de baixo:

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

Nesta última expressão, há duas matrizes cujos determinantes já se sabe calcular (a segunda e a terceira). Para calcular o determinante das outras duas, basta observar que definem paralelogramos de área nula, uma vez que os dois vectores que os definem são paralelos. Este facto é geral, conforme se verifica facilmente para dimensão 3.

Proposição 15. Seja A uma matriz em que uma linha é múltipla doutra. Então o determinante de A é 0.

Este resultado permite concluir os cálculos acima iniciados.

$$\begin{aligned}\det\begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} &= \underbrace{\det\begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}}_{0} + \det\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} + \det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \underbrace{\det\begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}}_{0} \\ &= \det\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} - \det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -2-1 = -3\end{aligned}$$

É fácil verificar que o paralelogramo correspondente a esta matriz (Figura 1.6) tem de facto área 3, que é o valor absoluto deste determinante.

Aplicando o mesmo raciocínio a uma matriz 2×2 genérica, obtém-se uma fórmula explícita para o cálculo de determinantes.

$$\begin{aligned}
\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \det\begin{bmatrix} a & 0 \\ c & d \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \\
&= \det\begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} + \det\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix} \\
&= 0 + \det\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} - \det\begin{bmatrix} b & 0 \\ 0 & c \end{bmatrix} + 0 = ad - bc
\end{aligned}$$

Exercício 53. Indique o valor do determinante das seguintes matrizes.

(a) $\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} -3 & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} \frac{1}{3} & -2 \\ -1 & 3 \end{bmatrix}$ (d) $\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$

Não faz sentido procurar escrever fórmulas explícitas para matrizes de dimensão superior; o determinante duma matriz 3×3, por exemplo, é uma soma de seis produtos de três entradas da matriz. Ao invés de procurar estas fórmulas, apresentar-se-ão de seguida processos de cálculo para determinantes de matrizes de dimensões superiores — que permitirão, com algum cuidado, reduzir ao mínimo o número de operações aritmética necessárias.

Antes de prosseguir, convém resumir as propriedades do determinante já referidas e deduzir mais algumas.

Proposição 16. Seja A uma matriz $n\times n$.

(i) Se A é uma matriz diagonal, então o seu determinante é o produto dos elementos na diagonal.

(ii) Seja B obtida a partir de A por troca de duas linhas. Então $\det(B) = -\det(A)$.

(iii) Seja B uma matriz que se obtém de A por multiplicação da linha i pelo número real c. Então $\det(B) = c\times\det(A)$.

(iv) Seja B uma matriz cujas linhas são todas iguais às de A com excepção da linha i. Então $\det(A)+\det(B) = \det(C)$, onde C é a matriz que se obtém de A substituindo a linha i pela soma das linhas i de A e de B.

(v) Seja A uma matriz em que uma linha é múltipla doutra. Então $\det(A) = 0$.

(vi) Seja B uma matriz obtida somando a uma linha de A um múltiplo de outra linha de A. Então $\det(B) = \det(A)$.

Demonstração. As propriedades (i) a (v) já foram discutidas anteriormente. A propriedade (vi) é consequência de (iv) e (v). Se a linha j de B se obtém pela operação $l_j + kl_i$ a partir de A, então a matriz B pode ser escrita como a soma de A com a matriz obtida de A substituindo a linha j por kl_i (que tem determinante nulo), donde $\det(B) = \det(A)$. □

Por outro lado, todo o raciocínio feito até aqui poderia ter sido realizado sem qualquer alteração considerando a figura geométrica definida pelos vectores nas *colunas* duma matriz;

conclui-se que as propriedades anteriores são também válidas para colunas, ou seja, tem-se o seguinte resultado.

Proposição 17. Seja A uma matriz $n \times n$.

(vii) Seja B obtida a partir de A por troca de duas colunas. Então $\det(B) = -\det(A)$.

(viii) Seja B uma matriz que se obtém de A por multiplicação da coluna i pelo número real c. Então $\det(B) = c \times \det(A)$.

(ix) Seja B uma matriz cujas colunas são todas iguais às de A com excepção da coluna i. Então $\det(A) + \det(B) = \det(C)$, onde C é a matriz que se obtém de A substituindo a coluna i pela soma das colunas i de A e de B.

(x) Seja A uma matriz em que uma coluna é múltipla doutra. Então $\det(A) = 0$.

(xi) Seja B uma matriz obtida somando a uma coluna de A um múltiplo de outra coluna de A. Então $\det(B) = \det(A)$.

Para além destas, o determinante goza ainda das seguintes propriedades.

Proposição 18. Sejam A e B duas matrizes de dimensão $n \times n$.

(xii) $\det\left(A^T\right) = \det(A)$

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

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

Demonstração.

(xii) A figura geométrica definida pelas linhas de A é a mesma figura definida pelas colunas de A^T, donde os determinantes destas duas matrizes coincidem.

(xiii) Geometricamente, o sólido definido pelas linhas de AB obtém-se substituindo cada aresta da figura definida pelas linhas de A pelo seu produto por B. Esta transformação é uma operação linear — o volume da figura resultante varia linearmente com o comprimento de cada aresta —, pelo que basta estudar o que acontece no caso de a matriz A ser a identidade. Ora $\det(\mathbf{I_n}B) = \det(B)$, pelo que a constante de proporcionalidade é $\det(B)$. Então o volume da figura correspondente a AB é $\det(A)\det(B)$.

(xiv) Consequência da propriedade anterior, tendo em conta que $AA^{-1} = \mathbf{I_n}$ e que $\det(\mathbf{I_n}) = 1$.

□

1.5.2 Regra de Laplace

Consideremos agora então uma matriz genérica 3×3. Tal como exemplificado acima para matrizes 2×2, pode-se ver a primeira linha desta matriz como uma soma e expandir o determinante da forma seguinte.

$$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \det\begin{bmatrix} a & 0 & 0 \\ d & e & f \\ g & h & i \end{bmatrix} + \det\begin{bmatrix} 0 & b & 0 \\ d & e & f \\ g & h & i \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Por troca de colunas, esta expressão é ainda igual a

$$\det \begin{bmatrix} a & 0 & 0 \\ d & e & f \\ g & h & i \end{bmatrix} - \det \begin{bmatrix} b & 0 & 0 \\ e & d & f \\ h & g & i \end{bmatrix} + \det \begin{bmatrix} c & 0 & 0 \\ f & d & e \\ i & g & h \end{bmatrix},$$

observando que na terceira matriz se efectuaram duas trocas de colunas (a primeira com a segunda e a segunda com a terceira), correspondendo a duas trocas de sinal.

Tal como no método de eliminação de Gauss, somando em cada uma das matrizes os múltiplos adequados da primeira linha às outras linhas pode-se anular os restantes termos da primeira coluna. Uma vez que esta operação não altera o determinante da matriz (propriedade (vi)) e que as outras entradas da primeira linha são nulas, conclui-se que a última expressão é ainda igual a

$$\det \begin{bmatrix} a & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{bmatrix} - \det \begin{bmatrix} b & 0 & 0 \\ 0 & d & f \\ 0 & g & i \end{bmatrix} + \det \begin{bmatrix} c & 0 & 0 \\ 0 & d & e \\ 0 & g & h \end{bmatrix}.$$

Em qualquer dos casos, o paralelipípedo correspondente à matriz tem uma aresta sobre o eixo dos xx e as duas outras no plano yOz. Então, uma vez que se trata de um sólido com uma aresta perpendicular à base, o seu volume é dado pelo comprimento da aresta multiplicado pela área da base; mas a área da base é o determinante da submatriz que se obtém ignorando a primeira linha e a primeira coluna. Ou seja:

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a \times \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \times \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \times \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}.$$

A técnica usada pode ser generalizada de três formas diferentes. Em primeiro lugar, não há necessidade de se escolher a primeira linha — poder-se-ia ter escolhido qualquer linha da matriz, apenas tendo de se efectuar também trocas de linhas para colocar a linha com os zeros no topo. Em segundo lugar, se se tivesse usado a expansão duma coluna em vez de uma linha obter-se-ia uma expansão análoga. Finalmente, o mesmo processo pode ser usado para matrizes de dimensão superior, permitindo escrever o seu determinante à custa de matrizes mais pequenas.

Um pouco de reflexão permite concluir que o número de trocas necessárias para colocar o elemento a_{ij} na primeira posição é igual a $(i-1)+(j-1)$. Este valor tem a mesma paridade que $i+j$; uma vez que é a paridade do número de trocas que determina o sinal do determinante (um número par de trocas não o altera, um número ímpar altera-o), pode-se escrever o método descrito acima simbolicamente nas seguintes forma, conhecida como *regra de Laplace*:

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

onde A_{ij}, o *menor-i,j* de A, é a matriz que se obtém de A eliminando a linha i e a coluna j. A primeira fórmula corresponde a seleccionar a linha i; a segunda corresponde a seleccionar a coluna j.

Exemplo. Considere-se a matriz B correspondente ao paralelipípedo da Figura 1.6. Aplicando a regra de Laplace à primeira linha e usando a fórmula acima estabelecida para o determinante de matrizes 2×2 é fácil calcular $\det(B)$.

$$\begin{aligned}\det\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} &= 1 \times \det\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} - 0 \times \det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 1 \times \det\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \\ &= 1 \times -1 + 1 \times -1 = -2\end{aligned}$$

Neste exemplo, o número de determinantes de menores a calcular foi reduzido devido à presença de um 0 na linha escolhida. Em geral, o cálculo de matrizes de dimensões maiores pode ser bastante simplificado escolhendo uma linha ou coluna com vários zeros.

Exemplo. Considere-se a matriz

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

e observe-se em primeiro lugar que na segunda coluna apenas existe um elemento diferente de 0, na entrada $(4,2)$. Uma vez que $4+2=6$ é par, conclui-se que

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

Agora é a quarta coluna que só tem uma entrada não nula, na posição $(1,4)$; uma vez que $1+4=5$ é ímpar, nova aplicação da regra de Laplace permite concluir que

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

Finalmente, aplicando novamente a regra de Laplace à terceira linha desta matriz (a aplicação à terceira coluna daria um resultado idêntico) permite terminar os cálculos.

$$\begin{aligned}3 \times (-1) \times \det\begin{bmatrix} 2 & 1 & 2 \\ -1 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix} &= 3 \times (-1) \times \left(1 \times \det\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} - 2 \times \det\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\right) \\ &= 3 \times (-1) \times (1 \times (-5) - 2 \times 4) = 39\end{aligned}$$

Exercício 54. Para cada uma das seguintes matrizes 3×3, determine todos os seus menores e calcule o seu determinante aplicando a regra de Laplace a cada linha e coluna. Verifique que o valor obtido para cada matriz é sempre o mesmo.

(a) $\begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix}$ (b) $\begin{bmatrix} -1 & 3 & 4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 2 & 6 & 6 \\ 2 & 7 & 6 \\ 2 & 7 & 7 \end{bmatrix}$

Antes de prosseguir, apresenta-se uma regra explícita para determinantes de matrizes 3×3, a regra de Sarros, que é consequência directa da regra de Laplace. Conforme foi visto acima,

$$\begin{aligned} \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= a \times \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \times \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \times \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} \\ &= a \times (ei - fh) - b \times (di - fg) + c \times (dh - eg)\,. \end{aligned}$$

Acrescentando uma cópia das duas primeiras colunas à matriz original, obtém-se

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

e é fácil observar que os termos de sinal positivo na expressão anterior (aei, bfg e cdh) ocorrem nas três diagonais descendentes para a direita, enquanto que os termos de sinal negativo (ceg, bdi e afh) ocorrem nas três diagonais descendentes para a esquerda, conforme assinalado no esquema seguinte.

$$\begin{bmatrix} \mathbf{a} & \underline{b} & \boxed{c} & a & b \\ d & \mathbf{e} & \underline{f} & \boxed{d} & e \\ g & h & \mathbf{i} & \underline{g} & \boxed{h} \end{bmatrix} \qquad \begin{bmatrix} a & b & \mathbf{c} & \underline{a} & \boxed{b} \\ d & \mathbf{e} & \underline{f} & \boxed{d} & e \\ \mathbf{g} & \underline{h} & \boxed{i} & g & h \end{bmatrix}$$

Exemplo. Vamos usar a regra de Sarros para calcular o determinante da matriz

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

Começamos por aumentar esta matriz, obtendo

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

e a regra diz-nos que o seu determinante é

$$\begin{aligned} \det(A) &= (2 \times 3 \times 0 + 1 \times 1 \times 1 + 2 \times (-1) \times 2) - (2 \times 3 \times 1 + 2 \times 1 \times 2 + 1 \times (-1) \times 0) \\ &= (0 + 1 - 4) - (6 + 4 + 0) = -3 - 10 = -13\,. \end{aligned}$$

Observe-se, contudo, que embora seja útil conhecer esta fórmula, na maioria dos casos é muito mais simples calcular determinantes de matrizes 3×3 recorrendo à regra de Laplace.

Exercício 55. Calcule o determinante das matrizes do Exercício 54 recorrendo à regra de Sarros.

1.5.3 Eliminação de Gauss

Outra forma prática de calcular determinantes, e que é a mais eficiente para matrizes de dimensões maiores, recorre ao método de eliminação de Gauss. Recorde-se que este método recorre apenas a operações elementares:

- troca de linhas;
- multiplicação duma linha por uma constante;
- substituição de uma linha pela sua soma com um múltiplo doutra.

Em termos de determinante da matriz, a primeira operação troca o seu sinal, a segunda multiplica-o pela mesma constante, e a terceira não o altera. Assim, é fácil analisar a evolução do determinante ao longo da aplicação do método. Se B é obtida de A através de:

- $l_i \leftrightarrow l_j$, então $\det(B) = -\det(A)$;
- αl_i, então $\det(B) = \alpha \det(A)$;
- $l_i \to l_i + \alpha l_j$, então $\det(B) = \det(A)$.

Finalmente, no caso de B ser obtida de A através da operação (não elementar) $l_i \to \alpha l_i + \beta l_j$, verifica-se ainda que $\det(B) = \alpha \det(A)$.

Por outro lado, o método de eliminação de Gauss produz sempre uma matriz em escada de linhas. É fácil ver que o determinante duma tal matriz é simplesmente o produto dos elementos na sua diagonal (este facto verifica-se facilmente recorrendo, por exemplo, à regra de Laplace).

Proposição 19. O determinante duma matriz em escada de linhas é o produto dos elementos na sua diagonal.

Demonstração. Por aplicação sucessiva da regra de Laplace, seleccionando sempre a primeira coluna de cada matriz — que só tem um elemento não-nulo na entrada $(1,1)$, correspondendo ao pivot dessa coluna. $\square$

Exemplo. Considere-se a seguinte matriz e a sequência de passos no método de eliminação de Gauss até obter uma matriz em escada de linhas.

$$\begin{bmatrix} 2 & 1 & -1 & 2 \\ 1 & 2 & -1 & 2 \\ 4 & 2 & 1 & -2 \\ 1 & -1 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \mathbf{2}l_2 - l_1 \\ l_3 - 2l_1 \\ \mathbf{2}l_4 - l_1 \end{matrix} \longrightarrow \begin{bmatrix} 2 & 1 & -1 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 3 & -6 \\ 0 & -3 & 1 & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ l_4 + l_2 \end{matrix} \longrightarrow \begin{bmatrix} 2 & 1 & -1 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 3 & -6 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

No primeiro passo, o determinante da matriz foi multiplicado por 4 (a segunda linha foi multiplicada por 2, a quarta também); no segundo passo, o determinante não se alterou (só foram somadas duas linhas). A última matriz é uma matriz em escada de linhas, pelo que o seu determinante é $2 \times 3 \times 3 \times 2 = 36$; logo o determinante da primeira é 9, pois $9 \times 4 = 36$.

O exemplo seguinte ilustra como ambos os métodos podem ser combinados.

Exemplo. Considere-se a seguinte matriz.

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

Uma vez que a quarta coluna só tem um elemento não-nulo (a entrada $(4,4)$), da regra de Laplace deduz-se que o determinante desta matriz é igual a

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

Aplicando o método de eliminação de Gauss para eliminar os elementos da primeira coluna, substitui-se a segunda linha por $l_2 - 2l_1$, a terceira por $l_3 + l_1$ e a quarta por $l_4 - l_1$. Uma vez que nenhuma destas operações altera o determinante da matriz, este vale

$$-3 \times \det \begin{bmatrix} 1 & 2 & -1 & 1 \\ 0 & 2 & 1 & 2 \\ 0 & -1 & 2 & 1 \\ 0 & 4 & 2 & 4 \end{bmatrix}.$$

Nesta matriz a última linha é o dobro da segunda, pelo que o seu determinante é zero. Conclui-se então que o determinante da matriz original é 0.

Exercício 56. Calcule o determinante das seguintes matrizes recorrendo ao método de eliminação de Gauss.

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

1.5.4 Aplicações

Para terminar esta exposição, apresentam-se duas aplicações dos determinantes, uma à inversão de matrizes e outra à resolução de sistemas de equações. Embora em ambos os casos o método de eliminação de Gauss (ou Gauss–Jordan) seja por norma mais eficiente, nalguns casos particulares pode ser prático conhecer estas fórmulas de resolução.

A regra de Laplace apresenta uma fórmula para o cálculo do determinante duma matriz A que recorre ao determinante dos menores de A. Ao determinante do menor-i,j de A, com o sinal correspondente à troca de linhas e colunas, chama-se *cofactor-i,j* de A: $a_{ij}(-1)^{i+j}\det(A_{ij})$.

A título de exemplo, considere-se a matriz

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

e os seus menores.

$$A_{11} = \begin{bmatrix} 3 & 2 \\ -2 & -3 \end{bmatrix} \qquad A_{12} = \begin{bmatrix} -1 & 2 \\ 4 & -3 \end{bmatrix} \qquad A_{13} = \begin{bmatrix} -1 & 3 \\ 4 & -2 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} 1 & 5 \\ -2 & -3 \end{bmatrix} \qquad A_{22} = \begin{bmatrix} 2 & 5 \\ 4 & -3 \end{bmatrix} \qquad A_{23} = \begin{bmatrix} 2 & 1 \\ 4 & -2 \end{bmatrix}$$

$$A_{31} = \begin{bmatrix} 1 & 5 \\ 3 & 2 \end{bmatrix} \qquad A_{32} = \begin{bmatrix} 2 & 5 \\ -1 & 2 \end{bmatrix} \qquad A_{33} = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$$

Tendo em conta os índices das linhas e colunas correspondentes a cada uma destas matrizes, pode-se construir uma nova matriz, a *matriz dos cofactores* de A, colocando na entrada i, j o cofactor-i, j de A. Determina-se assim a seguinte matriz.

$$\operatorname{cof}(A) = \begin{bmatrix} -5 & 5 & -10 \\ -7 & -26 & 8 \\ -13 & -9 & 7 \end{bmatrix}$$

Observe-se que o produto inteiro duma linha (ou coluna) de A pela linha (ou coluna) correspondente de $\operatorname{cof}(A)$ é igual a $\det(A)$ — é precisamente o que a regra de Laplace afirma. Por exemplo:

- o produto interno da primeira linha de A pela primeira linha de $\operatorname{cof}(A)$ é

$$2 \times (-5) + 1 \times 5 + 5 \times (-10) = -55$$

- o produto interno da segunda coluna de A pela segunda coluna de $\operatorname{cof}(A)$ é

$$1 \times 5 + 3 \times (-26) + (-2) \times (-9) = -55$$

e analogamente para os restantes casos.

Por outro lado, o produto interno duma linha (ou coluna) de A por uma linha (ou coluna) de $\operatorname{cof}(A)$ que não a correspondente é igual a 0. De facto, esse produto interno corresponde, pela regra de Laplace, à formula de cálculo do determinante duma matriz com duas linhas (ou colunas) iguais, que se sabe já valer 0. Por exemplo:

- o produto interno da segunda linha de A pela terceira linha de $\operatorname{cof}(A)$ é

$$(-1) \times (-13) + 3 \times (-9) + 2 \times (-7) = 0$$

- o produto interno da primeira coluna de A pela segunda coluna de $\operatorname{cof}(A)$ é

$$2 \times 5 + (-1) \times (-26) + 4 \times (-9) = 0\,.$$

Transpondo a matriz dos cofactores, estes resultados podem ser sintetizados na relação $A \cdot (\operatorname{cof}(A))^T = (\operatorname{cof}(A))^T \cdot A = \det(A) \cdot \mathbf{I_n}$. Daqui obtêm-se duas propriedades fundamentais.

Proposição 20. Seja A uma matriz quadrada de dimensão n. Então:

- A é invertível sse $\det(A) \neq 0$;

- se A é invertível, então a sua inversa é $\frac{1}{\det(A)}\mathrm{cof}(A)^T$.

Demonstração. Se $\det(A) \neq 0$, então

$$A \cdot \left(\frac{1}{\det(A)}\mathrm{cof}(A)^T\right) = \frac{1}{\det(A)}\left(A \cdot \mathrm{cof}(A)^T\right) = \frac{1}{\det(A)}\det(A) \cdot \mathbf{I_n} = \mathbf{I_n}$$

donde A é invertível e a sua inversa é $\mathrm{cof}(A)^T$.

Por outro lado, se $\det(A) = 0$ então reduzindo A a uma matriz em escada de linhas obtém-se necessariamente uma matriz cuja última linha é nula, donde A não pode ser invertível. □

À matriz $\mathrm{cof}(A)^T$ também se chama *matriz adjunta* de A, denotada por $\mathrm{adj}(A)$.

Exercício 57. Escreva a matriz dos cofactores e a matriz adjunta de cada uma das matrizes 3×3 do Exercício 54. Decida quais delas são invertíveis e calcule a sua inversa usando a matriz adjunta. Compare o resultado com o que obteria aplicando o método de Gauss–Jordan.

A proposição acima também é útil para determinar, dada uma matriz com parâmetros, quais os valores dos parâmetros que a tornam matriz invertível. Consideremos por exemplo a matriz

$$A_\alpha = \begin{bmatrix} \alpha & 1 & 5 \\ -1 & 3 & 2 \\ 2\alpha & -2 & -3 \end{bmatrix}$$

com parâmetro $\alpha \in \mathbb{R}$.

Para determinar que valores do parâmetro tornam a matriz invertível, é necessário apenas calcular o seu determinante e encontrar os valores de α para os quais este não é nulo. Aplicando a regra de Laplace, obtemos

$$\begin{aligned} \det A_\alpha &= \det \begin{bmatrix} \alpha & 1 & 5 \\ -1 & 3 & 2 \\ 2\alpha & -2 & -3 \end{bmatrix} \\ &= \alpha \det \begin{bmatrix} 3 & 2 \\ -2 & -3 \end{bmatrix} - \det \begin{bmatrix} -1 & 2 \\ 2\alpha & -3 \end{bmatrix} + 5 \det \begin{bmatrix} -1 & 3 \\ 2\alpha & -2 \end{bmatrix} \\ &= -5\alpha - (3 - 4\alpha) + 5(2 - 6\alpha) = -5\alpha - 3 + 4\alpha + 10 - 30\alpha \\ &= 7 - 31\alpha\,. \end{aligned}$$

Então a matriz A_α só é singular quando $\alpha = \frac{7}{31}$. Em particular, quando $\alpha = 2$, obtemos a matriz do exemplo anterior com determinante -55 e inversa dada por

$$A_2^{-1} = \frac{1}{\det(A_2)}\mathrm{adj}(A_2) = -\frac{1}{55}\begin{bmatrix} -5 & -7 & -13 \\ 5 & -26 & -9 \\ -10 & 8 & 7 \end{bmatrix}.$$

Exemplo. Seja A a matriz dada por

$$\begin{bmatrix} \alpha & 2 & 1 \\ 0 & \alpha & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Para encontrar os valores do parâmetro α que tornam a matriz A invertível temos de resolver a equação $\det A = 0$. Usando a fórmula de Laplace aplicada à primeira coluna, obtemos

$$\det A = \alpha \det \begin{bmatrix} \alpha & 1 \\ 1 & 1 \end{bmatrix} = \alpha(\alpha - 1)$$

e, portanto,

$$\det A = 0 \iff \alpha(\alpha - 1) = 0 \iff \alpha = 0 \text{ ou } \alpha = 1\,.$$

Assim, a matriz A é invertível para todos os reais diferentes de 0 e de 1.

Exercício 58. Determine o(s) valor(es) de α que tornam as matrizes seguintes singulares.

(a) $\begin{bmatrix} \alpha - 2 & 1 \\ \alpha - 1 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} \alpha - 1 & \alpha \\ \alpha & \alpha + 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 2 & 4 \\ 3 & 1 & 6 \\ \alpha & 3 & 2 \end{bmatrix}$

Outra aplicação dos determinantes é a chamada *regra de Cramer* para a resolução de sistemas de equações. De uma forma geral, esta regra diz que, num sistema $Ax = b$ de n equações a n incógnitas que seja possivel e determinado (e portanto em que a matriz A dos coeficientes tem determinante diferente de 0), se tem

$$x_i = \frac{\det(A_i)}{\det(A)}$$

onde A_i é a matriz obtida de A substituindo a coluna i pelo vector resultado b.

Exemplo. Considere-se o seguinte sistema de equações.

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

Este sistema pode ser escrito na forma matricial como

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

Para obter o valor de x, substitui-se a primeira coluna de A pelo vector resultado e calcula-se a expressão acima; para encontrar o valor de y procede-se de forma análoga, substituindo agora o vector resultado na segunda coluna de A. Ou seja:

$$x = \frac{\det \begin{bmatrix} 2 & 2 \\ -1 & -3 \end{bmatrix}}{\det \begin{bmatrix} 3 & 2 \\ 1 & -3 \end{bmatrix}} = \frac{-4}{-11} = \frac{4}{11} \qquad y = \frac{\det \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}}{\det \begin{bmatrix} 3 & 2 \\ 1 & -3 \end{bmatrix}} = \frac{-5}{-11} = \frac{5}{11}$$

Esta solução pode ser verificada.

$$\begin{aligned}3x + 2y &= 3 \times \frac{4}{11} + 2 \times \frac{5}{11} = \frac{12}{11} + \frac{10}{11} = \frac{22}{11} = 2 \\ x - 3y &= \frac{4}{11} - 3 \times \frac{5}{11} = \frac{4}{11} - \frac{15}{11} = \frac{-11}{11} = -1\end{aligned}$$

Observe-se que em geral o método de eliminação de Gauss é muito mais eficiente do que a regra de Cramer para resolver sistemas de equações. Porém, em determinados casos — nomeadamente, se só interessa conhecer o valor de uma variável e se obtém um determinante simples de calcular — pode ser útil aplicar esta regra.

Exemplo. Considere-se o sistema de equações seguinte.

$$\begin{cases} 2x + 3y - 2z + w = -1 \\ 3x - 2y + 4z - 2w = 2 \\ x - 2y + 2z + 3w = 1 \\ x + y + w = 0 \end{cases}$$

Pela regra de Cramer, o valor de z pode ser calculado como se segue.

$$z = \frac{\det \begin{bmatrix} 2 & 3 & -1 & 1 \\ 3 & -2 & 2 & -2 \\ 1 & -2 & 1 & 3 \\ 1 & 1 & 0 & 1 \end{bmatrix}}{\det \begin{bmatrix} 2 & 3 & -2 & 1 \\ 3 & -2 & 4 & -2 \\ 1 & -2 & 2 & 3 \\ 1 & 1 & 0 & 1 \end{bmatrix}} = \frac{\det \begin{bmatrix} 2 & 3 & -1 & 1 \\ 3 & -2 & 2 & -2 \\ 1 & -2 & 1 & 3 \\ 1 & 1 & 0 & 1 \end{bmatrix}}{2\det \begin{bmatrix} 2 & 3 & -1 & 1 \\ 3 & -2 & 2 & -2 \\ 1 & -2 & 1 & 3 \\ 1 & 1 & 0 & 1 \end{bmatrix}} = \frac{1}{2}$$

No segundo passo, recorreu-se às propriedades do determinante para simplificar a expressão, escrevendo a terceira coluna da matriz inferior como o dobro de outro vector.

É fácil justificar a validade da regra de Cramer usando apenas as propriedades já vistas dos determinantes. Apresenta-se de seguida esta justificação para o caso dum sistema de 3 equações a 3 incógnitas, por forma a simplificar a notação; o caso geral é análogo.

Teorema 2. Considere-se o sistema de n equações a n incógnitas $Ax = b$, escrito na forma matricial, com $\det(A) \neq 0$. Então as soluções deste sistema são dadas por

$$x_i = \frac{\det(A_i)}{\det(A)},$$

onde A_i é a matriz obtida de A substituindo a coluna i pelo vector resultado b.

Demonstração. Vamos ver o que se passa no caso $n = 3$, em que o sistema tem a forma

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}.$$

Segundo a regra de Cramer,

$$x = \frac{\det \begin{bmatrix} p & b & c \\ q & e & f \\ r & h & i \end{bmatrix}}{\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}}$$

e o sistema de equações diz que $p = ax + by + cz$, e analogamente para q e r. Substituindo estes valores na expressão anterior obtém-se

$$\frac{\det \begin{bmatrix} ax+by+cz & b & c \\ dx+ey+fz & e & f \\ gx+hy+iz & h & i \end{bmatrix}}{\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}}$$

e a primeira coluna do determinante da matriz no numerador é uma soma de três vectores. Aplicando a distributividade da soma (propriedade (ix)), a expressão anterior transforma-se em

$$\frac{\det \begin{bmatrix} ax & b & c \\ dx & e & f \\ gx & h & i \end{bmatrix} + \det \begin{bmatrix} by & b & c \\ ey & e & f \\ hy & h & i \end{bmatrix} + \det \begin{bmatrix} cz & b & c \\ fz & e & f \\ iz & h & i \end{bmatrix}}{\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}}.$$

Analisando novamente o numerador desta fracção, há duas propriedades do determinante que se podem aplicar. Na primeira matriz, a primeira coluna é resultado da multiplicação de x pelo vector (a, b, c), podendo a expressão ser simplificada com base na propriedade (viii). Na segunda matriz, a primeira coluna é múltipla da segunda, donde o seu determinante é 0 (propriedade (x)), enquanto que na terceira matriz a primeira coluna é múltipla da terceira. A expressão anterior simplifica-se então a

$$\frac{x \times \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}}{\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}}$$

que tem precisamente o valor x. $\square$

Exercício 59. Resolva os seguintes sistemas de equações recorrendo à regra de Cramer.

(a) $\begin{cases} 3x + 2y = 1 \\ 2x + 5y = -2 \end{cases}$ (b) $\begin{cases} x - 2y = 3 \\ 2x + y = 3 \end{cases}$ (c) $\begin{cases} 3x - y = -2 \\ x + y = 1 \end{cases}$ (d) $\begin{cases} 2x + y = 0 \\ x - 2y = 3 \end{cases}$

1.6 Exercícios e problemas

Resolução de sistemas de equações lineares

60. Indique, justificando, quais das seguintes equações são lineares.

(a) $x = -\pi y + \frac{2}{3}w - \sqrt{3}z$ (b) $x^{\frac{2}{5}} + 8y - 5z = \sqrt{7}$ (c) $3x+2y+z = 3z+2y+x$

61. Recorrendo ao método de eliminação de Gauss, encontre soluções dos seguintes sistemas de equações lineares.

(a) $\begin{cases} 2x - 2y = -2 \\ x + y = -1 \end{cases}$ (b) $\begin{cases} 3x - 5y = 1 \\ 5x - 3y = 7 \end{cases}$ (c) $\begin{cases} x - y = 0 \\ 3x - 2y = 2 \end{cases}$ (d) $\begin{cases} 2x + 3y = 0 \\ -x - 3y = 0 \end{cases}$

62. Recorrendo ao método de eliminação de Gauss, encontre soluções dos seguintes sistemas de equações lineares. Indique em cada passo o pivot e os multiplicadores usados.

(a) $\begin{cases} -3x + y + 3z = -1 \\ 2x - 2y + z = -5 \\ 3x + y - 2z = 4 \end{cases}$ (c) $\begin{cases} 3x - 4y + z = 1 \\ 6x - 2y - z = 1 \\ z = 0 \end{cases}$ (e) $\begin{cases} 2x + 4y - 3z = 1 \\ -x + 2y + 6z = 1 \\ x + y + 3z = 2 \end{cases}$

(b) $\begin{cases} 2x - 2z = 0 \\ x - 2y + 2z = 4 \\ 3x - 2y - z = 2 \end{cases}$ (d) $\begin{cases} x - y - z = -4 \\ 2x + 2y + 3z = 6 \\ -x + y - 2z = -2 \end{cases}$ (f) $\begin{cases} 4x - y + 2z = -2 \\ -8x + 2y + z = 4 \\ 2x - y + 3z = -3 \end{cases}$

63. Escreva o sistema homogéneo correspondente a cada um dos seguintes sistemas de equações e resolva-o. Usando a solução particular fornecida, escreva a expressão da solução geral do sistema de equações em questão.

(a) $\begin{cases} x - y + z = 0 \\ 2x + y + 2z = 6 \end{cases}$
Solução particular:
$(x, y, z) = (1, 2, 1)$

(c) $\begin{cases} 2x + 3y + 2z = 10 \\ x - 2y + 2z = 7 \end{cases}$
Solução particular:
$(x, y, z) = (3, 0, 2)$

(e) $\begin{cases} 3x - y + 2z = 0 \\ 4x - y + 2z = 1 \end{cases}$
Solução particular:
$(x, y, z) = (1, 3, 0)$

(b) $\begin{cases} x + 2y + 3z = 6 \\ 2x - y + 2z = 3 \\ 3x + y + 5z = 9 \end{cases}$
Solução particular:
$(x, y, z) = (1, 1, 1)$

(d) $\begin{cases} 3x + 2z - w = 3 \\ x - y + z - w = -2 \\ 2x + 3z + w = 7 \end{cases}$
Solução particular:
$(x, y, z, w) = (2, 1, 0, 3)$

(f) $\begin{cases} 2x + 3y + 4z = 9 \\ 4x + 3y + 2z = 9 \end{cases}$
Solução particular:
$(x, y, z, w) = (1, 1, 1, -1)$

(g) $\begin{cases} x + z = 0 \\ y + w = 0 \\ x - y + z = 1 \end{cases}$
Solução particular:
$(x, y, z, w) = (-1, -1, 1, 1)$

(h) $\begin{cases} x + y + w = 0 \\ z = 4 \end{cases}$
Solução particular:
$(x, y, z, w) = (1, 2, 4, -3)$

64. Classifique cada um dos sistemas de equações seguintes como determinado, indeterminado ou impossível. Nos primeiros dois casos, encontre a(s) solução(ões) do sistema.

(a) $\begin{cases} x+y=0 \\ 3x+3y=2 \end{cases}$

(b) $\begin{cases} x+y=2 \\ 3x-y=2 \\ 2x-2y=0 \end{cases}$

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

(d) $\begin{cases} 3x-2y=1 \\ -6x+4y=2 \end{cases}$

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

(f) $\begin{cases} 2x-6y+4z=6 \\ x+2y-4z=3 \\ x-8y+8z=3 \end{cases}$

(g) $\begin{cases} x-z=0 \\ x+z=4 \\ 2y-z=0 \end{cases}$

(h) $\begin{cases} 2x+3y-z+2w=2 \\ -x+2y+z-2w=1 \\ 2x+10y=-3 \end{cases}$

(i) $\begin{cases} 3x-2y+z=5 \\ 2x+y+w=3 \\ z+3w=5 \end{cases}$

(j) $\begin{cases} x+y+z+w=1 \\ 2x-z+w=-2 \\ -2x+y+3z=3 \\ x+2y+z+w=3 \end{cases}$

65. Para cada um dos seguintes sistemas de equações lineares, determine para que valor(es) dos parâmetros α, β e γ eles são possíveis determinados, possíveis indeterminados e impossíveis.

(a) $\begin{cases} x+y+2z=\alpha \\ x+z=\beta \\ 2x+y+3z=\gamma \end{cases}$

(b) $\begin{cases} x+y+z=2 \\ x-y+z=2 \\ \alpha x-z=-2 \\ 3x+y+\beta z=6 \end{cases}$

(c) $\begin{cases} x+y-2z=\alpha \\ x+3z=\beta \\ 2x+y+3z=\gamma \end{cases}$

(d) $\begin{cases} x+2y-3z=0 \\ 3x-y+5z=0 \\ 4x+y+(\alpha^2-14)\,z=0 \end{cases}$

(e) $\begin{cases} x+2y-3z=4 \\ 3x-y+5z=2 \\ 4x+y+(\alpha^2-14)\,z=\alpha+2 \end{cases}$

Aritmética matricial

66. Escreva cada um dos sistemas do Exercício 62 na forma matricial e resolva-os aplicando o método de eliminação de Gauss à matriz obtida.

67. Repita o exercício anterior para os sistemas de equações do Exercício 64.

68. Escreva o sistema de equações correspondente a cada uma das seguintes matrizes aumentadas. Resolva-o pelo método de eliminação de Gauss (aplicado ao sistema ou à matriz).

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

(b) $\left[\begin{array}{ccc|c} 3 & 0 & -2 & 5 \\ 7 & 1 & 4 & -3 \\ 0 & -2 & 1 & 7 \end{array}\right]$

(c) $\left[\begin{array}{cccc|c} 7 & 2 & 1 & -3 & 5 \\ 1 & 2 & 4 & 0 & 1 \end{array}\right]$

(d) $\left[\begin{array}{cccc|c} 3 & 1 & 1 & 1 & 0 \\ 5 & -1 & 1 & -1 & 0 \end{array}\right]$

(e) $\left[\begin{array}{ccc|c} 5 & -2 & 6 & 0 \\ -2 & 1 & 3 & 1 \end{array}\right]$

(f) $\left[\begin{array}{ccc|c} 1 & 1 & 2 & 8 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & 10 \end{array}\right]$

(g) $\left[\begin{array}{cc|c} 4 & -8 & 12 \\ 3 & -6 & 9 \\ -2 & 4 & -6 \end{array}\right]$

(h) $\left[\begin{array}{ccc|c} 2 & 1 & 3 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$ (j) $\left[\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ -2 & 5 & 2 & 1 \\ 8 & 1 & 4 & -1 \end{array}\right]$ (l) $\left[\begin{array}{cc|c} 2 & -3 & -2 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \end{array}\right]$

(i) $\left[\begin{array}{ccc|c} 2 & -1 & -3 & 0 \\ -1 & 2 & -3 & 0 \\ 1 & 1 & 4 & 0 \end{array}\right]$ (k) $\left[\begin{array}{ccc|c} 0 & -2 & 3 & 1 \\ 3 & 6 & -3 & -2 \\ 6 & 6 & 3 & 5 \end{array}\right]$ (m) $\left[\begin{array}{ccc|c} 3 & 2 & -1 & -15 \\ 5 & 3 & 2 & 0 \\ 3 & 1 & 3 & 11 \\ -6 & -4 & 2 & 30 \end{array}\right]$

(n) $\left[\begin{array}{cccc|c} 0 & 10 & -4 & 1 & 1 \\ 1 & 4 & -1 & 1 & 2 \\ 3 & 2 & 1 & 2 & 5 \\ -2 & -8 & 2 & -2 & -4 \\ 1 & -6 & 3 & 0 & -1 \end{array}\right]$ (p) $\left[\begin{array}{cccc|c} 0 & 2 & 2 & 4 & 0 \\ 1 & 0 & -1 & -3 & 0 \\ 2 & 3 & 1 & 1 & 0 \\ -2 & 1 & 3 & -2 & 0 \end{array}\right]$

(o) $\left[\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 1 \\ 1 & 3 & 7 & 2 & 2 \\ 1 & 12 & -11 & -16 & 5 \end{array}\right]$ (q) $\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{array}\right]$

69. Aplique o método de eliminação de Gauss–Jordan para transformar as seguintes matrizes na sua forma em escada de linhas reduzida.

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

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

70. Indique a dimensão de cada uma das seguintes matrizes.

(a) $\begin{bmatrix} 1 & 2 \\ 1 & 7 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 5 & 3 & -7 \end{bmatrix}$ (k) $\begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}$

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

(c) $\begin{bmatrix} 3 & -2 \end{bmatrix}$ (h) $\begin{bmatrix} 3 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$ (m) $\begin{bmatrix} \frac{1}{2} & -5 \\ 0 & -1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 3 & 1 & 2 \\ 1 & 2 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -3 & 2 & 0 \end{bmatrix}$ (i) $\begin{bmatrix} 2 & 1 & 3 & 0 & -2 \\ 0 & -2 & 3 & 4 & 5 \end{bmatrix}$ (n) $\begin{bmatrix} 1 & 3 & 1 & 2 \\ 2 & -3 & 4 & 1 \\ 1 & 3 & 1 & 2 \\ -1 & 2 & 2 & 1 \\ 0 & 5 & 1 & 0 \\ 0 & -3 & 2 & 0 \end{bmatrix}$

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

71. Para cada um dos seguintes pares de matrizes A e B, diga se é possível somá-las e, em caso afirmativo, calcule a matriz $A+B$.

(a) $A = \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 5 \\ 0 & 4 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 1 & 5 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 9 & 1 \\ 3 & -2 \\ 5 & 0 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$ $B = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 5 & 3 \end{bmatrix}$

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

(e) $A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & -1 & 2 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 0 & -2 \\ 1 & 1 \\ 2 & 1 \\ -1 & 9 \end{bmatrix}$

(f) $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $B = \begin{bmatrix} 2 & 1 \end{bmatrix}$

72. Para cada um dos pares de matrizes A e B do exercício anterior, diga se é possível multiplicá-las e, em caso afirmativo, calcule a matriz AB.

73. Sejam A, B, C, D e E matrizes com as dimensões seguintes.

A	B	C	D	E
4×5	4×5	5×2	4×2	5×4

Para cada uma das seguintes operações, diga se estão definidas e, em caso afirmativo, indique a dimensão da matriz resultado.

(a) AB
(b) BA
(c) BC
(d) CB
(e) $AC+D$
(f) $C+AD$
(g) $AE+B$
(h) $AE+B^T$
(i) $AB+B$
(j) AB^T
(k) AB^T+B
(l) $E(A+B)$
(m) $(A+B)C$
(n) $E(AC)$
(o) $(EA)C$
(p) EA
(q) E^TA
(r) $(A+B)^T+E$
(s) $(A^T+E)D$
(t) A^T+B^T+E

74. Considere as seguintes matrizes.

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

$$D = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix} \quad E = \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}$$

Verifique se as seguintes operações estão bem definidas e, em caso afirmativo, calcule o resultado.

(a) $D+E$
(b) $5A$
(c) $2B-C$
(d) $-3(D+2E)$
(e) $\mathrm{tr}(D-3E)$
(f) $2A^T+C$
(g) D^T-E^T
(h) B^T+5C^T
(i) $B-B^T$
(j) $\left(2E^T-3D^T\right)^T$
(k) $\mathrm{tr}(4E^T-D)$
(l) AB

(m) BA (p) CC^T (s) $\operatorname{tr}\left((AC)^T + 2E^T\right)$ (v) $(-AC)^T + 5D^T$

(n) $(AB)C$ (q) $(DA)^T$ (t) $\left(2D^T - E\right)A$ (w) $\left(BA^T - 2C\right)^T$

(o) $A(BC)$ (r) $\left(C^T B\right) A^T$ (u) $(4B)C + 2B$ (x) $B^T\left(CC^T - A^T A\right)$

75. Escreva as matrizes descritas em cada uma das seguintes alíneas. Use letras para assinalar as entradas cujo valor não está especificado.

    (a) Uma matriz 4×4 tal que $a_{ij} = 1$ se $i = j$ e $a_{ij} = 0$ se $i \neq j$.
    (b) Uma matriz 4×4 tal que $a_{ij} = 0$ se $i \neq j$.
    (c) Uma matriz 3×4 tal que $a_{ij} = i$ se $i \leq j$.
    (d) Uma matriz 5×3 tal que $a_{ij} = 0$ se $i > j$.
    (e) Uma matriz 5×5 tal que $a_{ij} = 0$ se $|i - j| > 1$.
    (f) Uma matriz 3×5 tal que $a_{ij} = i + j$.
    (g) Uma matriz 4×4 tal que $a_{ij} = (-1)^{i+j}$.
    (h) Uma matriz 3×4 tal que $a_{ij} = (3 - i)(j - 2)$.
    (i) Uma matriz 5×4 tal que $a_{ij} = i^2/j$.

76. Considere as matrizes A, B e C abaixo e tome $a = 4$, $b = -7$.

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & 5 \\ -2 & 1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 8 & -3 & 5 \\ 0 & 1 & 2 \\ 4 & -7 & 6 \end{bmatrix} \quad C = \begin{bmatrix} 0 & -2 & 3 \\ 1 & 7 & 4 \\ 3 & 5 & 9 \end{bmatrix}$$

Confirme que as seguintes igualdades se verificam, efectuando os cálculos de ambos os lados de cada relação.

(a) $(A + B) + C = A + (B + C)$
(b) $(AB)C = A(BC)$
(c) $(a + b)C = aB + aC$
(d) $a(B - C) = aB - aC$
(e) $a(BC) = (aB)C = B(aC)$
(f) $A(B - C) = AB - AC$
(g) $(B + C)A = BA + CA$
(h) $a(bC) = (ab)C$
(i) $\left(A^T\right)^T = A$
(j) $(A + B)^T = A^T + B^T$
(k) $(aC)^T = aC^T$
(l) $(AB)^T = B^T A^T$

77. Resolva a equação matricial

$$\begin{bmatrix} a - b & b + c \\ 3d + c & 2a - 4d \end{bmatrix} = \begin{bmatrix} 8 & 1 \\ 7 & 6 \end{bmatrix}$$

78. Verifique se as seguintes matrizes são invertíveis e, em caso afirmativo, calcule a sua inversa.

(a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 6 & 6 \\ 2 & 7 & 6 \\ 2 & 7 & 7 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & 3 & 4 \\ 2 & 4 & 1 \\ -4 & 2 & -9 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix}$

(f) $\begin{bmatrix} -3 & 6 \\ 4 & 5 \end{bmatrix}$

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

(h) $\begin{bmatrix} \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{10} \\ \frac{1}{5} & -\frac{4}{5} & \frac{1}{10} \end{bmatrix}$

(i) $\begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{bmatrix}$

(j) $\begin{bmatrix} -8 & 17 & 2 & \frac{1}{3} \\ 4 & 0 & \frac{2}{5} & -9 \\ 0 & 0 & 0 & 0 \\ -1 & 13 & 4 & 2 \end{bmatrix}$

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

(l) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 3 & 5 & 0 \\ 1 & 3 & 5 & 7 \end{bmatrix}$

(m) $\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix}$

(n) $\begin{bmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{bmatrix}$

(o) $\begin{bmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{bmatrix}$

Determinantes

79. Sem efectuar os cálculos, indique quais das seguintes matrizes têm determinante zero.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

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

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

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

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

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

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

(h) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

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

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

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

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

(m) $\begin{bmatrix} 1 & 0 & 2 \\ -1 & 2 & 1 \\ 1 & 2 & 5 \end{bmatrix}$

80. Considere matrizes quadradas A, B e C de dimensão n tais que $\det(A) = 2$, $\det(B) = -1$ e $\det(C) = 1$. Para cada uma das seguintes expressões, indique, justificando, se é possível calcular o seu determinante e, em caso afirmativo, qual o seu valor.

(a) A^T (b) B^{-1} (c) AB

(d) BA^{-1} (e) C^TB (f) $(AA^T)C^{-1}$

(g) $C + B^T$ (h) ABC (i) $A\left(B - C^T\right)$

(j) $B^2A^T + C^TB$ (k) $(AB)^{-1}$ (l) $B^2\left(C^{-1}A\right)$

(m) $A^TB + \mathbf{0_{n\times n}}(C + B)$ (n) $AC^T - \left(CA^T\right)^T$ (o) $\mathbf{I_n}B^TC^{-1}$

81. Calcule directamente o determinante de cada uma das seguintes matrizes.

(a) $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & -2 \\ \frac{1}{2} & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -1 \\ \frac{3}{4} & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 2 & \frac{1}{2} \\ -2 & 3 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

(f) $\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ 1 & 2 & 1 \end{bmatrix}$ (g) $\begin{bmatrix} 1 & 1 & -1 \\ 2 & \frac{1}{3} & 1 \\ -\frac{1}{2} & 2 & 1 \end{bmatrix}$ (h) $\begin{bmatrix} 2 & -1 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$ (i) $\begin{bmatrix} 1 & 2 & 3 \\ -3 & -2 & -1 \\ 1 & 1 & -1 \end{bmatrix}$

82. Para cada uma das seguintes matrizes 3×3, determine todos os seus menores. Usando a fórmula para determinantes de matrizes 2×2, calcule as suas matrizes dos cofactores e adjunta. Determine se a matriz é invertível e, em caso afirmativo, escreva a sua inversa. Compare os resultados com os obtidos no Exercício 78.

(a) $\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{10} \\ \frac{1}{5} & -\frac{4}{5} & \frac{1}{10} \end{bmatrix}$ (c) $\begin{bmatrix} 1 & -2 & 3 \\ 6 & 7 & -1 \\ -3 & 1 & 4 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{bmatrix}$

83. Calcule o determinante de cada uma das seguintes matrizes recorrendo eventualmente à fórmula de Laplace.

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

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

(h) $\begin{bmatrix} 5 & 1 & -2 \\ 2 & 0 & -1 \\ 0 & 3 & 0 \end{bmatrix}$ (i) $\begin{bmatrix} 2 & 0 & -1 \\ 2 & 6 & -2 \\ -1 & 0 & 2 \end{bmatrix}$ (j) $\begin{bmatrix} \frac{1}{2} & 1 & -1 \\ 1 & \frac{1}{2} & 1 \\ -1 & 1 & -2 \end{bmatrix}$

(k) $\begin{bmatrix} 1 & 5 & 2 & -9 \\ 0 & 2 & -5 & 1 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (l) $\begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 3 & 1 & -1 & 2 \\ 0 & 2 & 0 & 1 \end{bmatrix}$

(m) $\begin{bmatrix} -3 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 \\ 7 & 9 & -3 & 0 \\ 5 & -2 & 5 & -2 \end{bmatrix}$ (n) $\begin{bmatrix} -2 & 1 & 3 & -1 \\ 3 & -1 & 2 & 2 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 2 & 1 \end{bmatrix}$ (o) $\begin{bmatrix} 2 & 0 & -4 & 1 \\ 4 & 3 & 3 & 1 \\ 1 & 0 & 0 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}$

(p) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 5 & 1 & 4 & 2 & 1 \end{bmatrix}$ (q) $\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

(r) $\begin{bmatrix} 2 & 1 & 3 & 1 & 3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & -2 & 4 & 0 \\ 0 & 2 & 0 & -1 & 0 \\ -1 & 4 & 1 & 2 & -2 \end{bmatrix}$ (s) $\begin{bmatrix} 6 & 1 & 4 & 0 & 0 & 2 & 2 & 1 \\ 4 & 0 & 0 & 3 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 3 & 1 & 0 & 4 & 0 & 1 \\ 5 & 0 & 0 & 6 & 0 & 2 & 0 & 0 \\ 8 & 2 & 1 & 0 & 1 & 3 & 2 & 4 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 5 & 1 & 0 & 3 & 0 & 1 & 0 & 6 \end{bmatrix}$

84. Sabendo que

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = -6$$

calcule o determinante das seguintes matrizes.

(a) $\begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}$ (b) $\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$ (c) $\begin{bmatrix} d & e & f \\ g & h & i \\ a & b & c \end{bmatrix}$ (d) $\begin{bmatrix} a & a & c \\ d & d & f \\ g & g & i \end{bmatrix}$

(e) $\begin{bmatrix} a & g & d \\ b & h & e \\ c & i & f \end{bmatrix}$ (g) $\begin{bmatrix} 3a & -b & -2c \\ 3d & -e & -2f \\ 3g & -h & -2i \end{bmatrix}$ (i) $\begin{bmatrix} -3a & -b & -c \\ 6d & 2e & 2f \\ 3g & h & i \end{bmatrix}$

(f) $\begin{bmatrix} 2a & 2b & 2c \\ -d & -e & -f \\ 3g & 3h & 3i \end{bmatrix}$ (h) $\begin{bmatrix} 2a & 2b & 2c \\ 2d & 2e & 2f \\ 2g & 2h & 2i \end{bmatrix}$ (j) $\begin{bmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{bmatrix}$

(k) $\begin{bmatrix} a+g & b+h & c+i \\ a+d & b+e & c+f \\ g & h & i \end{bmatrix}$ (l) $\begin{bmatrix} 2a+3g & 2b+3h & 2c+3i \\ -d & -e & -f \\ 3a+2g & 3b+2h & 3c+2i \end{bmatrix}$

85. Calcule o determinante das seguintes matrizes recorrendo ao método de eliminação de Gauss.

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

86. Calcule o determinante de cada uma das seguintes matrizes pelo método mais adequado.

(a) $\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & -1 & -1 \end{bmatrix}$ (h) $\begin{bmatrix} -2 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 0 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

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

(j) $\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (k) $\begin{bmatrix} 0 & 2 & 0 & 0 \\ 3 & 1 & -1 & 3 \\ 0 & 3 & 0 & 1 \\ 2 & -1 & 2 & 1 \end{bmatrix}$ (l) $\begin{bmatrix} 1 & -3 & 2 & 1 \\ 3 & 2 & -1 & 3 \\ -2 & 6 & -4 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

87. Determine o(s) valor(es) de α que torna(m) as matrizes seguintes singulares.

(a) $\begin{bmatrix} 1 & 2 \\ -1 & \alpha \end{bmatrix}$ (c) $\begin{bmatrix} \alpha-3 & -2 \\ -2 & \alpha-2 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & -2 & 1 \\ 2 & -1 & 0 \\ \alpha & -\alpha & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 0 \\ \alpha & 4 & -2 \end{bmatrix}$

(b) $\begin{bmatrix} \alpha & 1 \\ 1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} \alpha & 1 \\ 1 & \alpha \end{bmatrix}$

88. Quais das matrizes dos exercícios 81, 83 e 86 são invertíveis? Quais são não-singulares?

89. Resolva cada um dos seguintes sistemas de equações recorrendo à regra de Cramer.

(a) $\begin{cases} x-2y+3z=3 \\ 3x-y+2z=2 \\ 2x+2y-z=-1 \end{cases}$ (b) $\begin{cases} x+2y=1 \\ y-2z=2 \\ x-z=2 \end{cases}$ (c) $\begin{cases} 4x-y+2z=-2 \\ -8x+2y+z=4 \\ 2x-y+3z=-3 \end{cases}$

(d) $\begin{cases} -3x + y + 3z = -1 \\ 2x - 2y + z = -5 \\ 3x + y - 2z = 4 \end{cases}$ (e) $\begin{cases} 2x + y - 2z = -1 \\ x + 2y + z = 5 \\ 3x - z = 1 \end{cases}$ (f) $\begin{cases} 2x + y - 2z = 1 \\ x + 2y + z = 4 \\ 3x - z = 2 \end{cases}$

Problemas

90. Considere a seguinte matriz.

$$\begin{cases} 4x + 2y + 4z = 2 \\ 2x + (2 - \alpha)z = -1 \\ 2x + 2z = \beta - 1 \end{cases}$$

(a) Determine para que parâmetros α e β o sistema é possível determinado, possível indeterminado e impossível.

(b) Encontre a solução geral do sistema com $\alpha = 1$ e $\beta = 0$.

Seja A a matriz de coeficientes associada ao sistema com $\alpha = 1$ e $\beta = 0$.

(c) Calcule o determinante da matriz A.

(d) Use a regra de Cramer para confirmar o valor de x.

(e) Verifique se a matriz A é invertível e, em caso afirmativo, calcule a inversa de A.

(f) Calcule o volume do paralelipípedo definido pelos vectores $(2, 0, 2), (2, 0, 1), (8, 4, 8)$.

91. Considere a seguinte matriz A.

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

(a) Calcule o determinante de A.

(b) Indique se A é invertível e, em caso afirmativo, encontre A^{-1}.

(c) Diga quais as soluções do sistema seguinte.

$$A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

(d) Encontre a matriz X que satisfaz a equação seguinte.

$$AX = \begin{bmatrix} 8 & 6 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$$

(e) Qual é o volume do paralelipípedo definido pelos vectores $(-2, 2, -1)$, $(1, 2, 1)$ e $(2, 2, 1)$?

92. Considere o seguinte sistema de equações lineares.

$$\begin{cases} 2x + 3y = 4 \\ -x + \alpha y = -2 \end{cases}$$

(a) Para que valores de α é que este sistema é indeterminado?

(b) Para que valores de α é que a matriz dos coeficientes do sistema apresentado é invertível?

93. Considere o seguinte sistema de equações lineares.

$$\begin{cases} x - y + 2w = 2 \\ 3x - 5y - 3z + 10w = 2 \\ 2x + 6z + 5w = 14 \\ -x + y + 6z + 3w = 10 \end{cases}$$

Seja A a matriz de coeficientes associada a este sistema.

(a) Encontre as soluções do sistema.

(b) Encontre as soluções do sistema homogéneo correspondente.

(c) Calcule o determinante da matriz A.

(d) Calcule o determinante da matriz seguinte.

$$\begin{bmatrix} 2 & 0 & 6 & 5 \\ 3 & -5 & -3 & 10 \\ 1 & -1 & 0 & 2 \\ -1 & 1 & 6 & 3 \end{bmatrix}$$

(e) Recorrendo à regra de Cramer, determine o valor da incógnita z no sistema seguinte.

$$A \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ -12 \\ -12 \end{bmatrix}$$

94. Considere o sistema

$$\begin{cases} x - 4z = -3 \\ 2x + \alpha y - 3z = -2 \\ x + y + \alpha z = 1 \end{cases}$$

e seja A a matriz de coeficientes associada a este sistema.

(a) Determine para que valores de α o sistema é possível determinado, possível indeterminado e impossível.

(b) Determine a solução geral do sistema para $\alpha = 1$.

(c) Calcule o determinante da matriz A, com $\alpha = 1$.

(d) Verifique se o sistema

$$\begin{cases} x - 4z = 6 \\ 2x + \alpha y - 3z = 1 \end{cases}$$

é possível ou impossível.

Capítulo 2

Espaços lineares

Toda a Álgebra Linear assenta no conceito de linearidade — a relação de operações com a soma e a multiplicação por escalares. No capítulo anterior, tivemos oportunidade de estudar diversas operações sobre matrizes que se comportavam bem com estas operações. Neste capítulo, vamos estudar as estruturas mais gerais em que faz sentido falar em linearidade: os espaços lineares ou vectoriais.

O tema dos espaços lineares é vastíssimo, pelo que nesta apresentação apenas se focarão alguns aspectos com interesse nas aplicações que mais tarde consideraremos. Começaremos por discutir alguns exemplos já conhecidos doutros contextos, para ganhar alguma intuição sobre a noção abstracta de espaço vectorial. O resto do capítulo apresenta um conjunto de operações e conceitos que fazem sentido em qualquer espaço vectorial, mostrando que de alguma forma os exemplos apresentados de início fornecem uma boa intuição para o conceito geral.

2.1 Motivação

Começamos por apresentar alguns exemplos já conhecidos de espaços vectoriais, por forma a dar alguma intuição para as propriedades que serão interessantes mais adiante.

2.1.1 Espaços de vectores

O estudo de certos fenómenos por vezes requer grandezas físicas que são completamente caracterizadas por um valor real. Por exemplo, o estudo da evolução da temperatura ao longo do tempo num determinado lugar é completamente determinado por um valor (a temperatura) em função do tempo. As grandezas caracterizadas por um único valor real dizem-se *grandezas escalares.*

Por outro lado, a velocidade de uma partícula que se desloca no espaço ao longo do tempo não pode ser completamente caracterizada por um único valor: é necessário conhecer não apenas a intensidade da velocidade, mas também a sua direcção e sentido. Pense-se num exemplo concreto: se soubermos que um automóvel partiu de Lisboa a uma velocidade constante de 100 km/h, conseguimos concluir que ao fim de três horas ele se encontra no máximo a 300 km de Lisboa, mas não conseguimos determinar a sua posição. O automóvel pode estar no Porto, em Faro, em Elvas ou mesmo em Lisboa (por exemplo, se tiver percorrido 150 km numa direcção e 150 km da direcção oposta).

Este tipo de grandezas que são caracterizadas por uma intensidade, uma direcção e um sentido dizem-se *grandezas vectoriais* ou simplesmente *vectores.*

Geometricamente, é usual representar um vector como uma seta entre um ponto inicial A e um ponto final B, denotado por $\overrightarrow{AB}$. É também habitual representar vectores genéricos por letras minúsculas com uma seta por cima: $\overrightarrow{u}$, $\overrightarrow{v}$, $\overrightarrow{w}$, $\overrightarrow{u_1}$, $\overrightarrow{u_2}$, etc. Uma vez que um vector geometricamente é um segmento de recta (orientado), é comum chamar *comprimento* à sua intensidade.

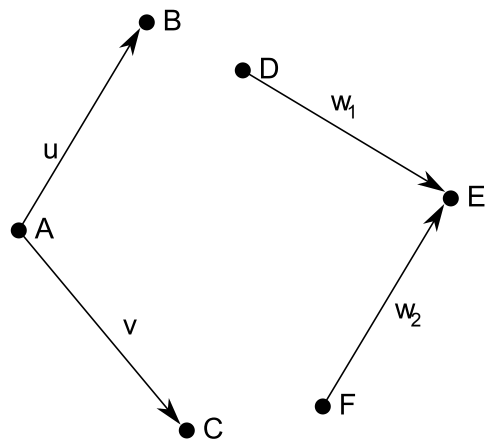


Figura 2.1: Vectores no plano. Os vectores $\overrightarrow{u} = \overrightarrow{AB}$ e $\overrightarrow{v} = \overrightarrow{AC}$ têm a mesma origem; os vectores $\overrightarrow{w_1}$ e $\overrightarrow{w_2}$ têm o mesmo destino; os vectores $\overrightarrow{u}$ e $\overrightarrow{w_1}$ têm o mesmo comprimento mas direcções diferentes; e os vectores $\overrightarrow{u}$ e $\overrightarrow{w_2}$ têm o mesmo comprimento, direcção e sentido.

Para caracterizar completamente um vector, é necessário indicar a sua direcção, o seu sentido, o seu comprimento e o seu ponto de aplicação. Porém, dois vectores que só difiram no ponto de aplicação (por exemplo, os vectores $\overrightarrow{u}$ e $\overrightarrow{w_2}$ da Figura 2.1) são para todos os efeitos práticos iguais.

Definição. Dois vectores com a mesma direcção, comprimento e sentido dizem-se *vectores equivalentes*.

Em particular, qualquer vector é equivalente a um vector aplicado na origem. Assim, quando estamos a raciocinar a um nível mais abstracto sobre vectores (ignorando o seu ponto de aplicação) podemos sempre supor que estes estão aplicados na origem. Desta forma, qualquer vector fica caracterizado pelo seu destino. Num plano, escolhendo um sistema de eixos, podemos portanto caracterizar um vector como um par de números reais; no espaço, podemos caracterizá-lo como três números reais. (Esta representação foi aliás já referida no capítulo anterior a propósito da definição de determinante.) A estes números chama-se *coordenadas* do vector.

Definição. O conjunto dos vectores com n componentes reais denota-se por $\mathbb{R}^n$.

Relativamente à Figura 2.2, o vector da esquerda é o vector $(3, 2)$ de $\mathbb{R}^2$ e o da direita é o vector $\left(\frac{1}{2}, 2, 1\right)$ de $\mathbb{R}^3$.

Em qualquer destes espaços, existe um vector com todas as entradas nulas (correspondendo a um vector com comprimento 0, ou um vector da origem para a origem).

Definição. O vector de $\mathbb{R}^n$ que tem todas as entradas nulas chama-se *vector zero* ou *vector nulo* (de $\mathbb{R}^n$) e denota-se por $\overrightarrow{0}$.

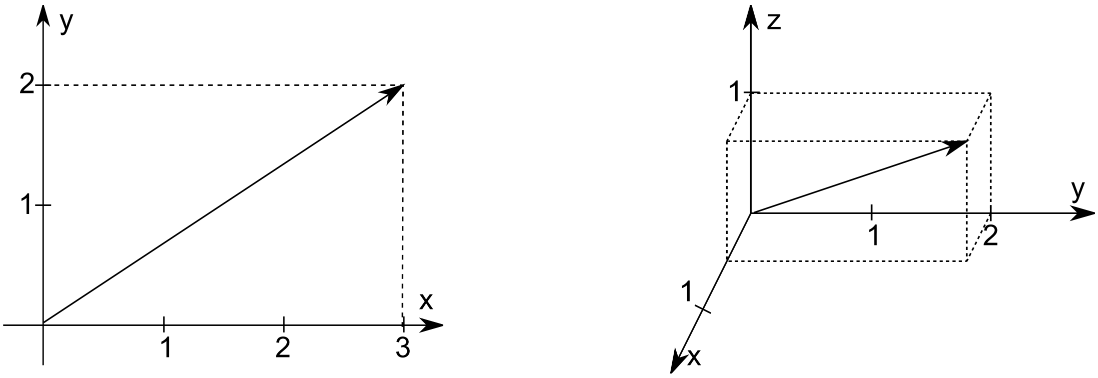


Figura 2.2: Determinação das coordenadas de um vector com ponto de aplicação na origem no plano (esquerda) ou no espaço (direita).

Note-se que os vectores nulos de espaços diferentes são vectores diferentes: o vector nulo de $\mathbb{R}^2$ é $(0,0)$, mas o vector nulo de $\mathbb{R}^3$ é $(0,0,0)$. Embora denotemos ambos os vectores por $\overrightarrow{0}$, o contexto permitir-nos-á sempre saber a qual deles nos estamos a referir.

Existem duas operações principais que se podem fazer com vectores, ambas com uma interpretação geométrica muito clara.

Definição. Se $\overrightarrow{u}$ e $\overrightarrow{v}$ são vectores de $\mathbb{R}^n$, então a sua soma $\overrightarrow{u}+\overrightarrow{v}$ é o vector de $\mathbb{R}^n$ cujas componentes são a soma das componentes correspondentes de $\overrightarrow{u}$ e $\overrightarrow{v}$.

$$\overrightarrow{u}+\overrightarrow{v}=(u_1,u_2,\ldots,u_n)+(v_1,v_2,\ldots,v_n)=(u_1+v_1,u_2+v_2,\ldots,u_n+v_n)$$

Geometricamente, somar dois vectores corresponde a aplicar o segundo deles no ponto de chegada do primeiro. A Figura 2.3 ilustra esta construção.

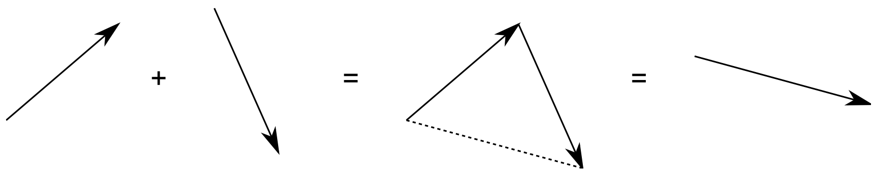


Figura 2.3: Soma de dois vectores.

A segunda operação é a multiplicação dum vector por um número real α.

Definição. Sejam α um número real e $\overrightarrow{u}$ um vector de $\mathbb{R}^n$. Então o produto de α por $\overrightarrow{u}$ é o vector cujas coordenadas são o produto de α pela coordenada correspondente de $\overrightarrow{u}$.

$$\alpha\overrightarrow{u}=\alpha\,(u_1,u_2,\ldots,u_n)=(\alpha u_1,\alpha u_2,\ldots,\alpha u_n)$$

Geometricamente, esta operação corresponde a manter a direcção do vector, mas multiplicar o seu comprimento por $|\alpha|$, mantendo o sentido se $\alpha>0$ e invertendo-o se $\alpha<0$, conforme ilustrado na Figura 2.4. Observe-se que o vector aumenta de comprimento se $|\alpha|>1$ e diminui se $|\alpha|<1$.

Se $\alpha=-1$, o vector $(-1)\overrightarrow{u}$ escreve-se simplesmente $-\overrightarrow{u}$ e chama-se o *simétrico* de $\overrightarrow{u}$. Este vector obtém-se de $\overrightarrow{u}$ mantendo o seu comprimento e direcção, mas trocando-lhe o sentido.

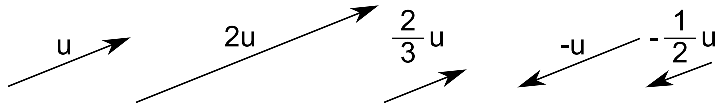


Figura 2.4: Produto de um vector por diferentes escalares.

Embora a interpretação geométrica seja extremamente útil para raciocinar sobre vectores, as propriedades algébricas da soma e produto por escalares são muito mais simples de perceber se pensarmos nestas operações em termos das coordenadas dos vectores.

Proposição 21. Sejam $\overrightarrow{u}$, $\overrightarrow{v}$ e $\overrightarrow{w}$ vectores de $\mathbb{R}^n$ e α e β números reais arbitrários. Então verificam-se as seguintes igualdades.

1. Associatividade da soma: $\overrightarrow{u} + (\overrightarrow{v} + \overrightarrow{w}) = (\overrightarrow{u} + \overrightarrow{v}) + \overrightarrow{w}$.
2. Comutatividade da soma: $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{u}$.
3. Existência de elemento neutro para a soma: $\overrightarrow{u} + \overrightarrow{0} = \overrightarrow{u}$.
4. Existência de simétrico: $\overrightarrow{u} + (-\overrightarrow{u}) = \overrightarrow{0}$.
5. Associatividade do produto por escalares: $\alpha(\beta\overrightarrow{u}) = (\alpha\beta)\overrightarrow{u}$.
6. Existência de elemento neutro para o produto: $1\overrightarrow{u} = \overrightarrow{u}$.
7. Distributividade do produto pela soma de vectores: $\alpha(\overrightarrow{u} + \overrightarrow{v}) = \alpha\overrightarrow{u} + \alpha\overrightarrow{v}$.
8. Distributividade do produto pela soma de escalares: $(\alpha + \beta)\overrightarrow{u} = \alpha\overrightarrow{u} + \beta\overrightarrow{u}$.

Demonstração. Todas as propriedades são consequência das propriedades correspondentes da soma e produto de números reais. A soma de reais é associativa, comutativa, tem elemento neutro e todo o número real tem simétrico; o produto é associativo e tem elemento neutro; e a soma é distributiva em relação ao produto.

Uma vez que um vector é uma lista de reais e que a soma e produto por escalar são efectuadas coordenada a coordenada, as propriedades das operações sobre vectores resultam imediatamente do facto de as igualdades indicadas se verificarem em cada coordenada. □

Veremos adiante que os espaços $\mathbb{R}^n$ (para qualquer n) são exemplos de espaços vectoriais; num sentido que precisaremos mais adiante, são mesmo os espaços vectoriais mais importantes.

2.1.2 Espaços de matrizes

Vamos agora ver que se tomarmos conjuntos de matrizes com dimensão fixa podemos construir espaços com propriedades muito semelhantes aos espaços $\mathbb{R}^n$ discutidos na secção anterior.

Definição. O conjunto $M_{n\times m}$ contém todas as matrizes de entradas reais de dimensão $n \times m$.

Por exemplo, o conjunto $M_{2\times 2}$ contém todas as matrizes quadradas de dimensão 2; o conjunto $M_{3\times 1}$ contém todos os vectores coluna de dimensão 3; o conjunto $M_{2\times 3}$ contém todas as matrizes rectangulares com duas linhas e três colunas; e assim sucessivamente.

Vimos já que duas matrizes com a mesma dimensão podem ser somadas, sendo a sua soma novamente uma matriz com a dimensão das parcelas. Também sabemos que podemos multiplicar uma matriz por um escalar arbitrário, obtendo outra matriz com a mesma dimensão. Ou seja, tal como acontecia em $\mathbb{R}^n$, também em $M_{n\times m}$ temos definidas uma soma e um produto por escalares.

Fixados n e m, existe uma única matriz em $M_{n\times m}$ contendo zeros em todas as suas entradas — a matriz nula $\mathbf{0_{n\times m}}$. Por outro lado, vimos também como definir a matriz simétrica $-A$ duma matriz arbitrária $A_{n\times m}$. Entre as propriedades destas operações vistas no capítulo anterior, contam-se as seguintes.

Proposição 22. Sejam A, B e C matrizes em $M_{n\times m}$ (ou seja, matrizes de dimensão $n \times m$) e α e β números reais arbitrários. Então verificam-se as seguintes igualdades.

1. Associatividade da soma: $A + (B + C) = (A + B) + C$.

2. Comutatividade da soma: $A + B = B + A$.

3. Existência de elemento neutro para a soma: $A + \mathbf{0_{n\times m}} = A$.

4. Existência de simétrico: $A + (-A) = \mathbf{0_{n\times m}}$.

5. Associatividade do produto por escalares: $\alpha(\beta A) = (\alpha\beta)A$.

6. Existência de elemento neutro para o produto: $1A = A$.

7. Distributividade do produto pela soma de vectores: $\alpha(A + B) = \alpha A + \alpha B$.

8. Distributividade do produto pela soma de escalares: $(\alpha + \beta)A = \alpha A + \beta A$.

Repare-se que estas propriedades são precisamente as mesmas propriedades satisfeitas pelos vectores de $\mathbb{R}^n$. Dito doutra forma, substituindo $\overrightarrow{u}$ por A, $\overrightarrow{v}$ por B, $\overrightarrow{w}$ por C e $\overrightarrow{0}$ por $\mathbf{0_{n\times m}}$ na Proposição 21, obtém-se precisamente o enunciado da Proposição 22. Esta observação sugere que há uma estrutura comum subjacente aos dois tipos de espaços — espaços de vectores $\mathbb{R}^n$ e espaços de matrizes $M_{n\times m}$ — que pode ser generalizada.

É importante observar que a bidimensionalidade das matrizes não intervém de forma significativa nem na soma nem no produto por escalares; é apenas quando falamos de traço, matriz transposta, determinante, produto, inversa ou característica que é relevante a estrutura da matriz com linhas e colunas. Assim, enquanto nos cingimos a somas e produtos de matrizes por escalares, podemos de facto pensar numa matriz de dimensão $n \times m$ como um vector contendo todas as entradas numa única lista (ou seja, como um vector de $\mathbb{R}^{n\times m}$). Esta identificação explica de certa forma que todas as propriedades listadas na proposição anterior se mantenham válidas — para além de mostrar que, no contexto destas duas operações, de facto não há diferenças substanciais entre trabalhar com vectores ou trabalhar com matrizes.

2.1.3 Espaços de sucessões

Vamos agora ver um exemplo que, apesar de parecer bastante diferente dos anteriores, também se comporta de forma semelhante.

Recorde-se que uma sucessão de números reais é uma sequência infinita de números reais. Sendo u uma sucessão, é comum chamar u_n ao elemento que ocupa a posição n nessa lista.

Podemos então pensar numa sucessão como uma função dos números naturais para os reais — a função que transforma o índice n no valor u_n da sucessão.

Por exemplo, a sucessão $u = 1, 2, 3, 4, 5, 6, 7, \ldots$ tem o número natural n na posição n, podendo ser escrita de forma abreviada como $u_n = n$. Já a sucessão dos quadrados perfeitos $v = 1, 4, 9, 16, 25, 36, 49, \ldots$ tem na posição n o valor n^2, pelo que se pode escrever como $v_n = n^2$.

Definição. O conjunto $\mathbb{R}^{\mathbb{N}}$ é o conjunto de todas as sucessões de números reais.

Tal como fizemos nos exemplos anteriores, podemos definir soma de duas sucessões e produto de uma sucessão por um escalar.

Definição. Sejam u e v duas sucessões de números reais. A sua *soma* é a sucessão $(u + v)$ tal que $(u + v)_n = u_n + v_n$ (soma pontual das duas sucessões).

Por exemplo, para as duas sucessões u e v apresentadas acima, tem-se

$$u + v = 2, 6, 12, 20, 30, \ldots \;;$$

se quisermos encontrar a expressão funcional desta sucessão, podemos usar a definição:

$$(u + v)_n = u_n + v_n = n + n^2 \,.$$

Exercício 1. Considere as sucessões

$$a = 0, 1, 0, 1, 0, 1, \ldots \qquad b = 1, 2, 4, 8, 16, 32, \ldots \qquad c = 4, 2, 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots \,.$$

Determine os primeiros termos de $a + b$, $a + c$ e $b + c$. Qual a expressão geral de $(b + c)_n$?

Definição. Sejam u uma sucessão de números reais e α um número real arbitrário. O *produto* de α por u é a sucessão (αu) tal que $(\alpha u)_n = \alpha u_n$ (produto de cada termo de u por α).

Por exemplo, para as duas sucessões u e v que usámos atrás, temos

$$2u = 2, 4, 6, 8, 10, \ldots$$

e em geral $(2u)_n = 2u_n = 2n$, enquanto que

$$-\frac{1}{3}v = -\frac{1}{3}, -\frac{4}{3}, -3, -\frac{16}{3}, -\frac{25}{3}, \ldots$$

e, em geral, $\left(-\frac{1}{3}v\right)_n = -\frac{1}{3}v_n = -\frac{n^2}{3}$.

Exercício 2. Considerando as mesmas sucessões a, b e c do exercício anterior, indique os primeiros termos de $-2a$, $\frac{1}{2}b$ e $0c$.

É habitual designar por 0 a sucessão constante com todos os termos iguais a zero (ou seja, $0_n = 0$; 0 designa a sucessão $0, 0, 0, 0, 0, \ldots$) e por $-u$ a sucessão $(-1)u$ que tem os simétricos dos elementos de u. Com estas notações, estas duas operações (soma e produto por escalares) gozam novamente de várias propriedades já vistas atrás.

Proposição 23. Sejam u, v e w sucessões de números reais e α e β números reais arbitrários. Então verificam-se as seguintes igualdades.

1. Associatividade da soma: $u + (v + w) = (u + v) + w$.

2. Comutatividade da soma: $u + v = v + u$.

3. Existência de elemento neutro para a soma: $u + 0 = u$.

4. Existência de simétrico: $u + (-u) = 0$.

5. Associatividade do produto por escalares: $\alpha(\beta u) = (\alpha\beta)u$.

6. Existência de elemento neutro para o produto: $1u = u$.

7. Distributividade do produto pela soma de vectores: $\alpha(u + v) = \alpha u + \alpha v$.

8. Distributividade do produto pela soma de escalares: $(\alpha + \beta)u = \alpha u + \beta u$.

Demonstração. A verificação destas propriedades é muito semelhante à prova para o caso de $\mathbb{R}^n$, dando alguma intuição sobre as semelhanças existentes entre estes dois espaços. De facto, se virmos as sucessões de reais como listas *infinitas* de números reais, podemos pensar nelas como "vectores" com um número infinito de coordenadas. Novamente, as operações de soma e produto por um escalar são definidas coordenada a coordenada à custa das operações análogas sobre números reais, pelo que estas propriedades são consequência imediata das propriedades idênticas das operações de soma e produto de números reais. □

Repare-se que este raciocínio corresponde a ver $\mathbb{R}^\mathbb{N}$ como um "limite" de $\mathbb{R}^n$ cujos vectores já têm infinitas coordenadas. Esta semelhança estrutural explica mais uma vez a coincidência das propriedades das operações definidas para os diferentes espaços.

2.1.4 Espaços de funções

Na secção anterior observámos que uma sucessão de números reais pode ser vista como uma regra de transformação de números naturais em números reais, correspondendo portanto a uma função de $\mathbb{N}$ em $\mathbb{R}$. De facto, esta visão pode ser aplicada a todos os exemplos das secções anteriores: um vector $\overrightarrow{u}$ em $\mathbb{R}^n$ não é mais do que uma função de $\{1, 2, \ldots, n\}$ em $\mathbb{R}$ — concretamente, a função que a cada valor i entre 1 e n associa a componente u_i de u. Já uma matriz $A_{n\times m}$ pode ser vista como uma função de duas variáveis, que a cada par (i, j) com $1 \leq i \leq n$ e $1 \leq j \leq m$ associa a entrada a_{ij} da matriz.

Reciprocamente, se fixarmos um conjunto X e pensarmos no conjunto de todas as funções de X para $\mathbb{R}$, podemos definir uma soma e um produto por escalares com propriedades semelhantes às dos espaços anteriores (que são casos particulares deste). Nesta secção vamos discutir um caso especialmente importante para muitas aplicações: o das funções reais de variável real.

Definição. O espaço de todas as funções reais de variável real designa-se por $\mathcal{F}(\mathbb{R})$.

Assim, são elementos de $\mathcal{F}(\mathbb{R})$ as funções $f(x) = x + 1$, $g(x) = x^2 + 3x - 1$, $h(x) = \sin(x)$, entre outras.

Conforme é sabido, duas funções podem ser somadas ponto a ponto.

Definição. Sejam f e g duas funções reais de variável real. A sua soma $f+g$ é a função definida por $(f+g)(x) = f(x)+g(x)$.

Também podemos multiplicar uma função por uma constante real arbitrária.

Definição. Sejam f uma função reais de variável real e α um número real. O produto de α por f é a função αf definida por $(\alpha f)(x) = \alpha f(x)$.

Designando por $-f$ a função tal que $(-f)(x) = -f(x)$ e por 0 a função identicamente igual a zero (ou seja, tal que $0(x) = 0$ para qualquer real x), verificam-se as seguintes propriedades.

Proposição 24. Sejam f, g e h funções reais de variável real e α e β números reais arbitrários. Então verificam-se as seguintes igualdades.

1. Associatividade da soma: $f+(g+h) = (f+g)+h$.
2. Comutatividade da soma: $f+g = g+f$.
3. Existência de elemento neutro para a soma: $f+0 = f$.
4. Existência de simétrico: $f+(-f) = 0$.
5. Associatividade do produto por escalares: $\alpha(\beta f) = (\alpha\beta)f$.
6. Existência de elemento neutro para o produto: $1f = f$.
7. Distributividade do produto pela soma de vectores: $\alpha(f+g) = \alpha f + \alpha g$.
8. Distributividade do produto pela soma de escalares: $(\alpha+\beta)f = \alpha f + \beta f$.

A prova é idêntica às anteriores.

2.2 Espaços e subespaços lineares

Neste momento, deve ser claro que há várias semelhanças entre os vários exemplos da secção anterior. Em todos os casos ($\mathbb{R}^n$, $M_{n\times m}$, $\mathbb{R}^{\mathbb{N}}$ e $\mathcal{F}(\mathbb{R})$) estamos perante um conjunto de objectos com duas operações definidas: uma soma (que permite criar um novo objecto do mesmo tipo a partir de dois outros) e um produto por escalares (que transforma um objecto num novo objecto do mesmo tipo, dado um número real) que satisfazem o mesmo conjunto de propriedades. Isto significa que podemos abstrair dessas operações e das suas propriedades numa estrutura geral. É a esta estrutura geral que chamamos *espaço linear* ou *espaço vectorial*.

2.2.1 Axiomática dos espaços lineares

Definição. Um *espaço linear* ou *espaço vectorial* é um conjunto V onde estão definidas uma operação de *soma* $+ : V \times V \to V$ e um *produto por escalares* $\cdot : \mathbb{R} \times V \to V$ e contendo um elemento 0_V tais que, para quaisquer $u, v, w \in V$ e $\alpha, \beta \in \mathbb{R}$, são satisfeitas as seguintes propriedades.

Axioma 1. $u+(v+w) = (u+v)+w$ (associatividade da soma)

Axioma 2. $u+v = v+u$ (comutatividade da soma)

Axioma 3. $u + 0_V = u$ (existência de elemento neutro da soma)

Axioma 4. $u + (-u) = 0_V$ para algum elemento $(-u) \in V$ (existência de simétrico)

Axioma 5. $\alpha(\beta u) = (\alpha\beta)u$ (associatividade do produto)

Axioma 6. $1u = u$ (existência de identidade)

Axioma 7. $(\alpha + \beta)u = \alpha u + \beta u$ (distributividade à esquerda)

Axioma 8. $\alpha(u + v) = \alpha u + \alpha v$ (distributividade à direita)

A vantagem de definir este conceito é permitir-nos determinar propriedades que são válidas em *todos* os espaços lineares — propriedades que dependem apenas da existência destas operações com estas propriedades, e não da estrutura particular do espaço em si. Por exemplo, em todos os espaços da secção anterior se verifica a identidade $0v = 0_V$ para qualquer v; dois casos particulares são, em $\mathbb{R}^3$, a identidade $0(x, y, z) = (0, 0, 0)$, enquanto em $\mathbb{R}^\mathbb{N}$ se tem que $0u$ é a sucessão nula para qualquer u. Na realidade, esta propriedade é consequência dos axiomas listados acima, pelo que se verifica em qualquer espaço linear e não carece de verificação.

Os exemplos da secção anterior podem ser sintetizados no seguinte resultado.

Proposição 25.

- Os espaços de vectores $\mathbb{R}^n$ são espaços lineares, para qualquer n.
- Os espaços de matrizes $M_{n\times m}$ são espaços lineares, para quaisquer n e m.
- O espaço das sucessões de números reais $\mathbb{R}^\mathbb{N}$ é um espaço linear.
- O espaço $\mathcal{F}(\mathbb{R})$ de todas as funções reais de variável real é um espaço linear.

Ao longo deste capítulo teremos ocasião de verificar que os espaços $\mathbb{R}^n$ são de alguma forma os espaços lineares "típicos", no sentido de que trabalhar num outro espaço linear pode ser em geral reduzido àquele caso. É por essa razão que surge o termo *espaço vectorial*: quando falamos em espaços lineares, estamos a chamar a atenção para a existência duma soma e produto por escalar que estão associadas ao conceito de lineariedade; ao falar em espaços vectoriais, estamos a salientar a ligação aos espaços $\mathbb{R}^n$. Neste texto usaremos os dois termos indistintamente; um elemento dum espaço vectorial genérico será também comummente designado por *vector*, de acordo com a terminologia habitual. Assim, neste sentido, um "vector" de $M_{n\times m}$ é uma matriz, um "vector" de $\mathbb{R}^\mathbb{N}$ é uma sucessão de números reais e um "vector" de $\mathcal{F}(\mathbb{R})$ é uma função real de variável real. Embora a princípio esta terminologia possa ser estranha, o seu uso é essencial para a interiorização dos conceitos mais abstractos com que estaremos a trabalhar a partir deste momento.

Vejamos um exemplo dum espaço vectorial um pouco diferente.

Exemplo. Consideremos o conjunto $V = \{(x, 0) \mid x \in \mathbb{R}\}$ com a soma e produto usuais em $\mathbb{R}^2$. Por outras palavras, V contém os pontos do eixo horizontal de $\mathbb{R}^2$ e estes são somados e multiplicados por escalares como normalmente em $\mathbb{R}^2$.

Para ver que V é de facto um espaço vectorial temos de verificar que as operações fazem sentido e que os axiomas se verificam. O primeiro passo é simples: os elementos de V são vectores cuja segunda coordenada é nula; é necessário verificar que ao somarmos dois vectores destes ou ao multiplicarmos um vector destes por um escalar obtemos como resultado um vector que ainda está em V.

De facto, é simples ver que

$$(x,0)+(x',0)=(x+x',0) \text{ e } \alpha(x,0)=(\alpha x,0)\,,$$

donde estas operações estão bem definidas.

De seguida, temos de verificar as identidades expressas nos axiomas. Para tal, vamos escolher vectores genéricos $u=(x,0)$, $v=(y,0)$ e $w=(z,0)$ e verificar que as igualdades consideradas se verificam, escolhendo os simétricos e o zero usuais de $\mathbb{R}^2$.

1. $(x,0)+((y,0)+(z,0))=(x+y+z,0)=((x,0)+(y,0))+(z,0)$
2. $(x,0)+(y,0)=(x+y,0)=(y,0)+(x,0)$
3. $(x,0)+(0,0)=(x,0)$
4. $(x,0)+(-x,0)=(0,0)$
5. $\alpha(\beta(x,0))=(\alpha\beta x,0)=(\alpha\beta)(x,0)$
6. $1(x,0)=(x,0)$
7. $(\alpha+\beta)(x,0)=(\alpha x+\beta x,0)=\alpha(x,0)+\beta(x,0)$
8. $\alpha((x,0)+(y,0))=(\alpha x+\alpha y,0)=\alpha(x,0)+\alpha(y,0)$

Concluímos portanto que V, com estas operações, é um espaço vectorial.

Mais adiante veremos uma forma mais expedita de resolver problemas semelhantes a este; porém, há casos em que a única forma de verificar que um conjunto é espaço vectorial é mesmo esta — testar directamente a satisfação dos axiomas.

Exercício 3. Quais das seguintes estruturas são espaços vectoriais reais?

(a) $\{(x,y)\mid x\geq 0, y\in\mathbb{R}\}$ com a soma e produto usuais em $\mathbb{R}^2$

(b) $\{(x,\dots,x)\mid x\in\mathbb{R}\}$ com a soma e produto usuais em $\mathbb{R}^n$

Embora usualmente nos refiramos ao espaço vectorial V sem mais, é importante salientar que um espaço vectorial fica definido não apenas pelo conjunto dos seus elementos, mas também pelas próprias operações. De facto, é possível (e tem utilidade em determinados contextos) definir estruturas diferentes de espaço linear sobre o mesmo conjunto de suporte.

Por outro lado, é preciso ter cuidado com a definição das operações: não é fácil garantir que todos os axiomas são satisfeitos.

Exemplo. Considere-se o espaço $V=\mathbb{R}^2$ com a soma de vectores definida atrás e o produto por escalares dado por $\alpha(x,y)=(\alpha x,0)$.

Vamos ver que $\mathbb{R}^2$ com estas operações não é espaço linear: o Axioma 6 não é satisfeito por todos os vectores. De facto, tomando-se $\overrightarrow{u}=(1,1)$, tem-se $1(1,1)=(1,0)\neq(1,1)$.

Exercício 4. Quais das seguintes estruturas são espaços vectoriais reais?

(a) $\mathbb{R}^3$ com a soma usual e o produto definido por $\alpha(x,y,z) = (\alpha x, y, z)$

(b) $\{(1,y) \mid y \in \mathbb{R}\}$ com as operações definidas por $(1,y)+(1,y') = (1,y+y')$ e $\alpha(1,y) = (1,\alpha y)$

(c) $\left\{ \begin{bmatrix} a & 0 \\ 1 & b \end{bmatrix} \mid a,b \in \mathbb{R} \right\}$ com a soma e produto definidos por

$$\begin{bmatrix} a & 0 \\ 1 & b \end{bmatrix} + \begin{bmatrix} a' & 0 \\ 1 & b' \end{bmatrix} = \begin{bmatrix} a+a' & 0 \\ 1 & b+b' \end{bmatrix} \qquad \alpha \begin{bmatrix} a & 0 \\ 1 & b \end{bmatrix} = \begin{bmatrix} \alpha a & 0 \\ 1 & \alpha b \end{bmatrix}$$

Vamos agora ilustrar a utilidade da noção abstracta de espaço linear dando exemplos de algumas propriedades que se verificam em todos os espaços vectoriais.

Proposição 26. Sejam V um espaço linear com vector zero 0_V, u, v e w vectores arbitrários de V e $\alpha \in \mathbb{R}$. Então verificam-se as seguintes propriedades.

1. $0u = 0_V$
2. $\alpha 0_V = 0_V$
3. Se $\alpha u = 0_V$, então $\alpha = 0$ ou $u = 0_V$.
4. $(-1)u = -u$
5. Se $u + v = u + w$, então $v = w$ (lei do corte).

Demonstração.

1. Como $0 = 0 + 0$, pelo Axioma 7 temos que

$$0u = (0+0)u = 0u + 0u\,.$$

Pelo Axioma 4, o elemento $0u$ tem um simétrico $(-0u)$. Somando este elemento a ambos os membros da igualdade anterior, obtemos

$$\begin{array}{rcl} & -(0u) + (0u + 0u) = -(0u) + 0u & \\ \iff & (-(0u) + 0u) + 0u = -(0u) + 0u & \text{(Axioma 1)} \\ \iff & (0u + (-0u)) + 0u = 0u + (-0u) & \text{(Axioma 2)} \\ \iff & 0_V + 0u = 0_V & \text{(Axioma 4)} \\ \iff & 0u = 0_V & \text{(Axioma 3)} \end{array}$$

consoante afirmámos.

2. Pelos Axiomas 8 e 3, temos que

$$\alpha 0_V + \alpha 0_V = \alpha\,(0_V + 0_V) = \alpha 0_V\,.$$

Seja $-(\alpha 0_V)$ o simétrico de $\alpha 0_V$. Então

$$\begin{array}{rcr} & \alpha 0_V + \alpha 0_V = \alpha 0_V & \\ \Longrightarrow & (\alpha 0_V + \alpha 0_V) + (-(\alpha 0_V)) = \alpha 0_V + (-(\alpha 0_V)) & \\ \Longleftrightarrow & \alpha 0_V + (\alpha 0_V + (-(\alpha 0_V))) = \alpha 0_V + (-(\alpha 0_V)) & \text{(Axioma 1)} \\ \Longleftrightarrow & \alpha 0_V = 0_V & \text{(Axioma 4)} \end{array}$$

o que termina a prova.

3. Suponhamos que $\alpha u = 0_V$. Se $\alpha = 0$, não temos mais nada a mostrar. Assuma-se então que $\alpha \neq 0$; temos que

$$u \overset{\text{Ax6}}{=} 1u = \left(\alpha^{-1}\alpha\right) u \overset{\text{Ax5}}{=} \alpha^{-1} \underbrace{(\alpha u)}_{0_V} = \alpha^{-1} 0_V = 0_V$$

recorrendo à Propriedade 2.

4. Para mostrar que $(-1)u$ é o simétrico de u, basta verificar que $u + (-1)u = 0_V$. Ora

$$u + (-1)u \overset{\text{Ax6}}{=} 1u + (-1)u \overset{\text{Ax1}}{=} (1 + (-1))u = 0u = 0_V$$

usando a Propriedade 1.

5. Assumamos que $u + v = u + w$. Tem-se então

$$\begin{array}{rcr} & u + v = u + w & \\ \Longrightarrow & (-u) + (u + v) = (-u) + (u + w) & \\ \Longleftrightarrow & ((-u) + u) + v = ((-u) + u) + w & \text{(Axioma 1)} \\ \Longleftrightarrow & (u + (-u)) + v = (u + (-u)) + w & \text{(Axioma 2)} \\ \Longleftrightarrow & 0_V + v = 0_V + w & \text{(Axioma 4)} \\ \Longleftrightarrow & v = w & \text{(Axioma 3)} \end{array}$$

conforme tínhamos afirmado. □

O interesse de provar estas propriedades a partir dos axiomas é, como foi dito atrás, garantir que se verificam em todos os espaços lineares. Daqui em diante, se precisarmos por exemplo de calcular o produto $0v$ num espaço vectorial qualquer V, sabemos já que garantidamente este vale 0, sem ter de olhar para a definição do produto por escalares.

A consequência fundamental destes resulados é podermos trabalhar com as operações de soma e produto por escalares da forma a que estamos habituados a trabalhar com os números reais.

Exercício 5. Verifique directamente que as propriedades demonstradas na proposição acima são válidas em todos os espaços vectoriais descritos na Secção 2.1.

2.2.2 Subespaços lineares

Nos exercícios da secção anterior encontrámos alguns exemplos de espaços vectoriais que eram definidos como subconjuntos de outros espaços maiores, preservando as operações (em particular, no Exercício 3). Esta situação é muito recorrente na prática e, como veremos nesta secção, é de facto muito mais simples de tratar do que o caso geral de mostrar que um conjunto com determinadas operações forma um espaço vectorial.

Definição. Sejam V um espaço linear e $W \subseteq V$ um subconjunto de V. Se W, conjuntamente com a soma, produto por escalares, e vector zero definidos para V for um espaço linear, então diz-se que W é um *subespaço linear* ou *subespaço vectorial* de V.

À partida, sendo V um espaço linear, para mostrar que $W \subseteq V$ também é espaço linear com as mesmas operações, teríamos de mostrar que todas as propriedades dos espaços lineares se verificam. Em particular, teríamos de mostrar que as operações de soma e produto por escalares estão bem definidas em W (ou seja, que a soma de elementos de W ainda é um elemento de W e que o produto de um elemento de W por um real ainda está um W) e que satisfazem os axiomas acima apresentados.

Contudo, a realidade é outra. Os axiomas do espaço linear dizem respeito a propriedades das *operações* e não dos elementos de V; verificando-se estes para todos os elementos de V, também se verificarão necessariamente para todos os elementos de W, uma vez que $W \subseteq V$. Assim, a única questão que é preciso verificar é que W é *fechado* para as operações de soma e produto por escalar (no sentido descrito no parágrafo anterior), que W contém o vector zero e que todos os elementos de W têm o seu simétrico em W. Na realidade, os dois últimos factos são consequência de W ser fechado para a soma e produto por escalares.

Teorema 3. Sejam V um espaço linear e $W \subseteq V$. Então $W \subseteq V$ é um subespaço linear de V se e só se W satisfaz as duas propriedades seguintes.

- Fecho para a soma: se $u, v \in W$, então $u + v \in W$.

- Fecho para o produto por escalares: se $\alpha \in \mathbb{R}$ e $u \in W$, então $\alpha u \in W$.

Demonstração. Claramente todo o subespaço linear de V satisfaz os axiomas de fecho. Reciprocamente, seja $W \subseteq V$ um conjunto que satisfaz os axiomas de fecho.

Pela Propriedade 1, $0_V \in W$, visto que W é fechado para o produto por escalares e $0_V = 0u$ para $u \in W$. Da mesma forma, se $u \in W$ então $-u = (-1)u$ também está em W pela Propriedade 4. Logo W contém o elemento neutro da soma e simétrico de todos os seus elementos, necessários para os Axiomas 3 e 4.

A verificação dos axiomas é automática, já que todos os elementos de W são elementos de V e V é um espaço linear. Logo W também é um espaço linear, sendo portanto um subespaço linear de V. □

Qualquer espaço vectorial tem dois subespaços muito simples.

Exemplo. Seja V um espaço linear qualquer. Os conjuntos $\{0_V\}$ (*subespaço trivial*) e V são subespaços lineares de V.

No caso do espaço trivial, uma vez que $0_V + 0_V = 0_V$ e que $\alpha 0_V = 0_V$, este conjunto é fechado para a soma e produto por escalares, constituindo portanto um subespaço vectorial de V.

No segundo caso, a verificação é directa pela definição: como V é um espaço linear e $V \subseteq V$, tem-se que V é subespaço linear de V.

Obviamente que estes dois exemplos são pouco interessantes, pelo que é comum falar de *subespaços próprios* dum espaço vectorial V para designar subespaços de V que não sejam nem o próprio V nem o espaço nulo; por outras palavras, um subespaço próprio de V é um subespaço linear W de V tal que $0 \neq W \neq V$.

Vejamos alguns exemplos de subespaços de $\mathbb{R}^2$.

Exemplo.

1. Consideremos o espaço vectorial $\mathbb{R}^2$ com as operações usuais de soma e produto por escalares. Geometricamente, este espaço corresponde a um plano; vamos ver que o eixo dos xx é um subespaço linear de $\mathbb{R}^2$.

   O eixo dos xx corresponde ao conjunto dos pontos tais que $y = 0$; ou seja, é o subconjunto
   $$W = \{(x, y) \in \mathbb{R}^2 \mid y = 0\} = \{(x, 0) \mid x \in \mathbb{R}\}$$
   de $\mathbb{R}^2$.

   Para verificar que W é um subespaço vectorial de $\mathbb{R}^2$, é necessário provar que W é fechado para a soma o produto por escalares de $\mathbb{R}^2$; ou seja, é necessário verificar que a soma de dois vectores sobre o eixo dos xx é ainda um vector sobre esse eixo e que o produto dum vector sobre o eixo dos xx por um escalar é ainda um vector sobre o eixo dos xx.

   Poderíamos fazer esta verificação geometricamente — o que é aliás bastante simples — mas vamos proceder por via algébrica, uma vez que será a técnica mais útil em geral. No próximo exemplo apresentaremos uma prova geométrica duma situação semelhante.

   Fecho para soma. Sejam $(x, 0)$ e $(y, 0)$ dois elementos escolhidos arbitrariamente de W. Então $(x, 0) + (y, 0) = (x + y, 0) \in W$, donde W é fechado para a soma.

   Fecho para o produto por escalares. Sejam α um real e $(x, 0)$ um elemento de W arbitrários. Então $\alpha(x, 0) = (\alpha x, 0) \in W$, donde W é fechado para o produto por escalares.

   Assim, o eixo dos xx é subespaço linear de $\mathbb{R}^2$.

2. Consideremos uma recta R de $\mathbb{R}^2$ que passa pela origem. Vamos ver que R também é subespaço linear de $\mathbb{R}^2$. Note-se que o exemplo anterior é um caso particular deste.

   Em vez de proceder analiticamente, vamos recorrer a um argumento geométrico. Escolhendo dois vectores arbitrários ao longo de R, vamos calcular a sua soma e verificar que esta ainda é um vector sobre R; escolhendo um vector sobre R e um real α, vamos verificar que o seu produto é ainda um vector sobre R. As construções encontram-se na Figura 2.5. Para garantir que estamos a tratar todos os casos, convém verificar, na soma, o caso em que os vectores têm sentidos iguais (a) e o caso em que têm sentidos opostos (b); no produto, o caso em que $\alpha > 0$ (c) e o caso em que $\alpha < 0$ (d).

Exercício 6. Faça a verificação de que R é um subespaço vectorial de $\mathbb{R}^2$ por via algébrica. Recorde que a equação geral duma recta passando pela origem é $y = mx$ (recta não vertical) ou $x = 0$ (recta vertical), pelo que R será o conjunto dos pontos $\{(x, mx) \mid x \in \mathbb{R}\}$ (recta não vertical) ou $\{(0, y) \mid y \in \mathbb{R}\}$ (recta vertical).

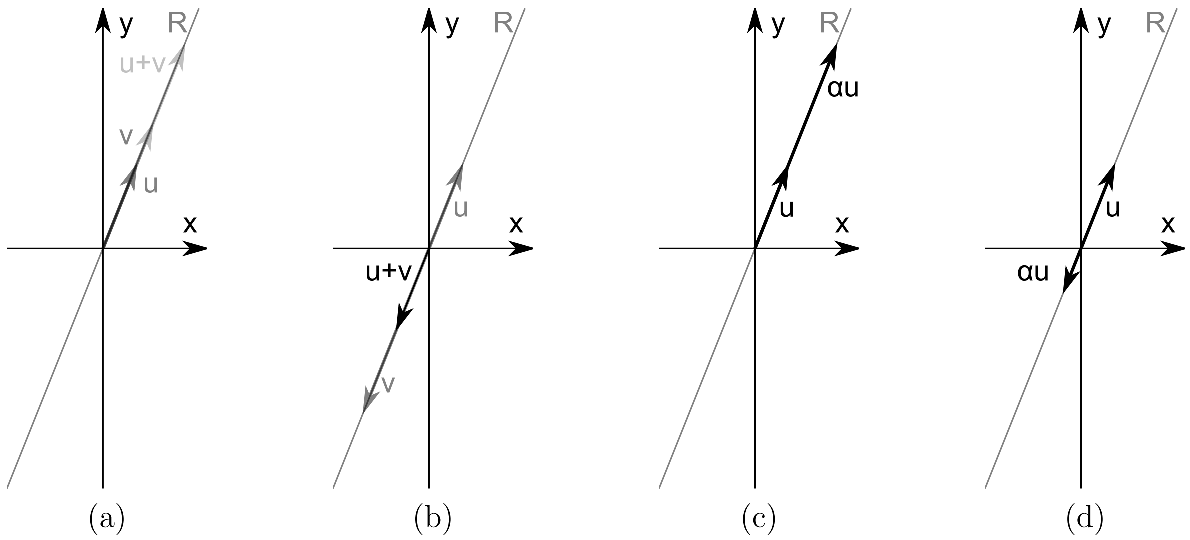


Figura 2.5: Uma recta R de $\mathbb{R}^2$ é fechada para a soma e produto por escalares.

Exemplo.

3. Vejamos agora que uma recta que não passa pela origem não é subespaço linear de $\mathbb{R}^2$. Seja
$$W = \{(x, y) \in \mathbb{R}^2 : y = mx + b, b \neq 0\}$$
uma tal recta.

A primeira observação a fazer é que W não passa na origem: se $x = 0$, então $y = b \neq 0$. Automaticamente, concluímos que W não pode ser um espaço vectorial.

Outra alternativa seria verificar que este conjunto não é fechado para a soma. Escolhendo por exemplo $u = (1, m + b)$, temos que $u + u = (1, m + b) + (1, m + b) = (2, 2m + 2b)$.

Ora $(2, 2m+2b)$ está em W se e só se $2m+2b = 2m+b$, ou seja, se $b = 2b$, o que implicaria $b = 0$; uma vez que isto não se passa, concluímos que $u + u \notin W$. Logo W não é fechado para a soma, donde não é um subespaço linear de $\mathbb{R}^2$.

Exercício 7. Verifique que as rectas e planos que passam pela origem são subespaços lineares de $\mathbb{R}^3$, enquanto as rectas e planos que não passam pela origem não o são. Recorde que a equação geral dum plano em $\mathbb{R}^3$ é $Ax + By + Cz = D$ (e o plano passa pela origem se e só se $D = 0$) enquanto uma recta é definida pelas duas equações $y = m_1x + b_1$ e $z = m_2x + b_2$, passando pela origem se e só se $b_1 = b_2 = 0$.

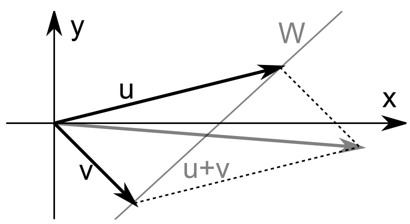


Figura 2.6: Uma recta de $\mathbb{R}^2$ que não passa pela origem não é um subespaço linear de $\mathbb{R}^2$.

Exercício 8. Verifique que os seguintes conjuntos *não* são subespaços vectoriais de $\mathbb{R}^2$.

(a) A parábola de equação $y = x^2 + 3x$.

(b) A união das duas rectas $y = 2x$ e $y = -x$.

Em geral, qualquer recta ou plano em $\mathbb{R}^n$ que passe pela origem é subespaço linear de $\mathbb{R}^n$, enquanto que qualquer recta ou plano de $\mathbb{R}^n$ que não passe na origem não o é. A segunda afirmação é simples de verificar: nenhum desses conjuntos contém a origem, que é precisamente o vector zero em $\mathbb{R}^n$. Mais adiante veremos formas simples de demonstrar a primeira afirmação.

Vamos agora ver exemplos de subespaços lineares de outros espaços.

Exemplo.

4. Consideremos o espaço linear $M_{2\times 2}$. As matrizes de dimensão 2×2 cujo determinante é igual a 1 não formam um subespaço linear de $M_{2\times 2}$: o vector 0 em $M_{2\times 2}$ é a matriz nula $\mathbf{0_{2\times 2}}$, que tem determinante 0, logo este conjunto não pode ser subespaço linear de $M_{2\times 2}$.

5. O conjunto S das matrizes singulares de dimensão 2×2 também não é um subespaço linear de $M_{2\times 2}$. Observe-se que o argumento que usámos atrás não nos é útil neste caso: a matriz $\mathbf{0_{2\times 2}}$ tem determinante 0, pelo que está em S.

   Para S ser um espaço linear, S teria de ser fechado para a soma e para o produto por escalares. Vamos ver que de facto não o é. Escolhendo as duas matrizes

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

   temos que $\det(A) = 0$ e $\det(B) = 0$, pelo que A e B estão ambas em S, mas $A + B = \mathbf{I_2}$, cujo determinante é 1 e portanto não está em S. Então S não é fechado para a soma, pelo que não é um espaço linear.

6. Considere-se o conjunto Z das sucessões reais cujo limite é 0 e vejamos que Z é um subespaço linear de $\mathbb{R}^{\mathbb{N}}$. Se u e v são duas sucessões com limite 0, então das propriedades operatórias dos limites sabe-se que a sua soma $u+v$ também tem limite 0. Analogamente, se α é um real arbitrário, então $\lim \alpha u = 0$. Logo Z é fechado para a soma e produto por escalares, pelo que é um subespaço linear de $\mathbb{R}^{\mathbb{N}}$.

7. Considere-se agora o espaço $\mathcal{F}(\mathbb{R})$ de todas as funções reais e consideremos o conjunto $\mathcal{C}^1(\mathbb{R})$ das funções diferenciáveis em todo o $\mathbb{R}$.

   Dos resultados da Análise, sabe-se que a soma de funções diferenciáveis é ainda uma função diferenciável, tal como o produto de uma função diferenciável por uma constante real. Então $\mathcal{C}^1(\mathbb{R})$ é um subconjunto de $\mathcal{F}(\mathbb{R})$ que é fechado para a soma e produto por escalares, sendo portanto um subespaço vectorial de $\mathcal{F}(\mathbb{R})$.

8. Escolha-se um número natural n e tome-se o conjunto P_n dos polinómios de grau menor ou igual a n. Por exemplo, P_1 é o conjunto de todas as funções da forma $p(x) = ax + b$; P_2 é o conjunto das funções da forma $p(x) = ax^2 + bx + c$; P_3 contém as funções da forma $p(x) = ax^3 + bx^2 + cx + d$; e em geral P_n contém todas as funções da forma

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \,.$$

Claramente $P_n \subseteq \mathcal{F}(\mathbb{R})$, uma vez que os polinómios são funções reais de variável real definidas em toda a recta real. Mais, tem-se ainda $P_n \subseteq \mathcal{C}^1(\mathbb{R})$. Vamos verificar que P_n é um subespaço linear de $\mathcal{F}(\mathbb{R})$ e de $\mathcal{C}^1(\mathbb{R})$.

Fecho para a soma. Se p e q são dois polinómios de grau menor ou igual a n, então a sua soma ainda é um polinómio de grau menor ou igual a n, uma vez que $(p+q)(x)$ pode ser calculado como

$$(p+q)(x) = (a_n + b_n)\, x^n + \cdots + (a_1 + b_1)\, x + (a_0 + b_0)$$

se assumirmos que

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \text{ e } q(x) = b_n x^n + \cdots + b_1 x + b_0 \,.$$

Fecho para o produto por escalares. Tomando p como acima e escolhendo um real α, a função αp continua a ser um polinómio de grau menor ou igual a n, já que $(\alpha p)(x)$ é uma função com os coeficientes de p multiplicados por α:

$$(\alpha p)(x) = \alpha a_n x^n + \cdots + \alpha a_1 x + \alpha a_0 \,.$$

Então P_n é um espaço linear, pelo que é subespaço linear de $\mathcal{C}^1(\mathbb{R})$ e de $\mathcal{F}(\mathbb{R})$.

9. Considere-se finalmente o espaço P dos polinómios de grau menor ou igual a 2 tais que $p'(2) = 0$ e vejamos que P é um subespaço linear de P_2.

Fecho para a soma. Se p e q são dois polinómios de grau menor ou igual a 2 tais que $p'(2) = 0$ e $q'(2) = 0$, então a sua soma é um polinómio diferenciável tal que

$$(p+q)'(2) = p'(2) + q'(2) = 0 + 0 = 0 \,.$$

Fecho para o produto por escalares. Tomando p como acima e escolhendo um real α, o polinímio αp é diferenciável e

$$(\alpha p)'(2) = \alpha p'(2) = \alpha \times 0 = 0 \,.$$

O conjunto P é portanto um subespaço linear de P_2 (e por consequência também de $\mathcal{C}^1(\mathbb{R})$ e $\mathcal{F}(\mathbb{R})$).

Exercício 9. Quais dos seguintes espaços de matrizes são subespaços de $M_{n\times m}$?

(a) $\{A_{3\times 3} \mid \det(A) \neq 0\}$

(b) $\{A_{2\times 2} \mid A \text{ é uma matriz diagonal}\}$

(c) $\{A_{3\times 3} \mid \mathrm{tr} A = 0\}$

(d) $\{A_{2\times 3} \mid a_{ij} > 0\}$

Exercício 10. Quais dos seguintes conjuntos de polinómios são subespaços vectoriais de P_n?

(a) $\{ax^2 + bx + c \mid b^2 - ac < 0\}$

(b) $\{ax^3 + bx^2 + cx \mid a, b, c \in \mathbb{R}\}$

(c) $\{ax^2 + bx + c \mid b^2 - ac \geq 0\}$

(d) $\{p(x) \in P_n \mid p(-1) = p(1)\}$

Exercício 11. Quais dos seguintes conjuntos de funções são subespaços vectoriais de $\mathcal{F}(\mathbb{R})$?

(a) $\{f \in \mathcal{F}(\mathbb{R}) \mid x^2 + f''(x) = 0\}$

(b) $\{f \in \mathcal{F}(\mathbb{R}) \mid f(x) \leq 0\}$

(c) $\{f \in \mathcal{F}(\mathbb{R}) \mid f \text{ é constante}\}$

(d) $\{f \in \mathcal{F}(\mathbb{R}) \mid f \text{ é duas vezes diferenciável}\}$

2.3 Representações de vectores

O facto de qualquer espaço linear dispor de operações de soma e produto por escalares permite--nos representar vectores à custa de outros. Nesta secção estudamos o conceito de combinação linear e as suas implicações.

2.3.1 Combinações lineares

Num espaço vectorial, usando as suas operações (soma e produto por escalares) podemos construir novos vectores a partir de outros. A uma expressão envolvendo somas e produtos por escalares chamamos combinação linear.

Se a designação de "espaço vectorial" dá ênfase à ligação entre estes espaços e os espaços $\mathbb{R}^n$, já o termo "espaço linear" frisa a importância das combinações lineares — que são uma das ferramentas básicas da Álgebra Linear.

Definição. Sejam V um espaço linear e S um subconjunto não vazio de elementos de V. Diz-se que $u \in V$ é *combinação linear* dos elementos de S se existirem elementos $s_1, s_2, \ldots, s_n \in S$ e números reais $c_1, c_2, \ldots, c_n$ tais que

$$u = c_1 s_1 + c_2 s_2 + \cdots + c_n s_n \,.$$

Por outras palavras, uma combinação linear de elementos de S é um vector que pode ser escrito a partir dum número finito de elementos de S usando apenas somas e multiplicações por escalares.

Este conceito está longe de ser novidade. Logo na Secção 1.1, quando introduzimos o método de eliminação de Gauss para a resolução de sistemas de equações lineares, definimos uma das operações elementares como "substituir uma linha pela sua soma com um múltiplo de outra" (operação $l_i + \alpha l_j$). De seguida, generalizámos esta operação para a substituição de l_i por $\alpha l_i + \beta l_j$ — que não é mais do que uma combinação linear das linhas i e j da matriz. Ou seja: o método de eliminação de Gauss assenta na troca de linhas e na substituição de uma linha por uma combinação linear dessa linha com outra.

Consideremos um exemplo simples: seja S o subconjunto $\{(1,0),(0,1)\}$ de $\mathbb{R}^2$. Então o vector $(2,1)$ é combinação linear dos elementos de S, já que $(2,1) = 2(1,0) + 1(0,1)$.

Na realidade, qualquer vector de $\mathbb{R}^2$ pode ser escrito como combinação linear de elementos de S: para um vector arbitrário (x,y), tem-se a relação $(x,y) = x(1,0) + y(0,1)$. Porém, nem todos os subconjuntos de $\mathbb{R}^2$ têm esta propriedade.

Exemplo. Considere-se o conjunto $S = \{(1,1),(-1,-1),(2,2)\}$. Se u é combinação linear de elementos de S, então existem reais c_1, c_2, c_3 tais que

$$u = c_1(1,1) + c_2(-1,-1) + c_3(2,2) = (c_1 - c_2 + 2c_3, c_1 - c_2 + 2c_3) \ ,$$

pelo que as coordenadas de u são iguais. Então o vector $(2,1)$ não é combinação linear de elementos de S.

Como é que podemos determinar se um vector é combinação linear de elementos dum conjunto de vectores duma forma sistemática? Em geral, não há um algoritmo para o fazer, já que a definição concreta do espaço linear V pode tornar a questão bastante complexa. Porém, em muitos exemplos (e em particular nas situações que consideraremos nas secções seguintes) conseguimos reduzir o problema à resolução dum sistema de equações lineares.

Para ilustrar este método, imaginemos que queríamos decidir se $(1,2,3)$ é combinação linear dos elementos de $S = \{(1,1,1),(0,2,1),(-1,0,1)\}$, subconjunto de $\mathbb{R}^3$. Para tal, teriam de existir números reais c_1, c_2 e c_3 tais que

$$\begin{aligned}(1,2,3) &= c_1(1,1,1) + c_2(0,2,1) + c_3(-1,0,1)\\ &= (c_1,c_1,c_1) + (0,2c_2,c_2) + (-c_3,0,c_3) = (c_1 - c_3, c_1 + 2c_2, c_1 + c_2 + c_3) \ .\end{aligned}$$

Ora dois vectores de $\mathbb{R}^3$ são iguais se todas as suas coordenadas o forem. Obtemos então o seguinte sistema de equações lineares.

$$\begin{cases} c_1 - c_3 = 1 \\ c_1 + 2c_2 = 2 \\ c_1 + c_2 + c_3 = 3 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 1 & 2 & 0 & 2 \\ 1 & 1 & 1 & 3 \end{array}\right] \begin{array}{l} \\ l_2 - l_1 \\ l_3 - l_1 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 2 \end{array}\right] \begin{array}{l} \\ \\ 2l_3 - l_2 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{array}\right] \longrightarrow \begin{cases} c_1 - c_3 = 1 \\ 2c_2 + c_3 = 1 \\ 3c_3 = 3 \end{cases} \longrightarrow \begin{cases} c_1 - c_3 = 1 \\ 2c_2 + c_3 = 1 \\ c_3 = 1 \end{cases} \longrightarrow \begin{cases} c_1 = 2 \\ c_2 = 0 \\ c_3 = 1 \end{cases}$$

Ou seja: $(1,2,3) = 2(1,1,1) + (-1,0,1)$, pelo que é combinação linear dos elementos de S.

Vejamos alguns exemplos que ilustram esta técnica noutros espaços vectoriais.

Exemplo.

1. Considere-se o espaço P_2 de todos os polinómios de grau menor ou igual a 2. Vamos ver se $4 - x + 8x^2$ é combinação linear dos elementos de

   $$S = \left\{1 + 2x - x^2, 6 + 4x + 2x^2\right\} \ .$$

   O polinómio $4 - x + 8x^2$ é combinação linear dos elementos de S se existirem reais c_1 e c_2 tais que

   $$4 - x + 8x^2 = c_1\left(1 + 2x - x^2\right) + c_2\left(6 + 4x + 2x^2\right) \ .$$

Simplificando a equação acima, obtemos

$$\begin{aligned}4 - x + 8x^2 &= c_1\left(1 + 2x - x^2\right) + c_2\left(6 + 4x + 2x^2\right)\\ &= c_1 + 2c_1x - c_1x^2 + 6c_2 + 4c_2x + 2c_2x^2\\ &= (c_1 + 6c_2) + (2c_1 + 4c_2)\,x + (-c_1 + 2c_2)\,x^2\end{aligned}$$

Ora é um facto conhecido que dois polinómios são iguais precisamente quando os coeficientes de todas as potências de x coincidem. A condição acima reduz-se portanto ao seguinte sistema de equações lineares.

$$\begin{cases}c_1 + 6c_2 = 4\\ (2c_1 + 4c_2) = -x\\ (-c_1 + 2c_2)\,x^2 = 8x^2\end{cases} \longrightarrow \begin{cases}c_1 + 6c_2 = 4\\ 2c_1 + 4c_2 = -1\\ -c_1 + 2c_2 = 8\end{cases} \longrightarrow \left[\begin{array}{cc|c}1 & 6 & 4\\ 2 & 4 & -1\\ -1 & 2 & 8\end{array}\right] \begin{array}{l}\\ l_2 - 2l_1\\ l_3 + l_1\end{array}$$

$$\longrightarrow \left[\begin{array}{cc|c}1 & 6 & 4\\ 0 & -8 & -9\\ 0 & 8 & 12\end{array}\right] \begin{array}{l}\\ \\ l_3 + l_2\end{array} \longrightarrow \left[\begin{array}{cc|c}1 & 6 & 4\\ 0 & -8 & -9\\ 0 & 0 & 3\end{array}\right] \longrightarrow \begin{cases}c_1 + c_2 = 4\\ -8c_2 = -9\\ 0 = 3\end{cases}$$

Como o último sistema é impossível, não existem constantes c_1 e c_2 nas condições pretendidas. Então $4 - x + 8x^2$ não é combinação linear dos elementos de S.

2. Seja S o subconjunto

$$S = \left\{\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix}\right\}$$

de $M_{2\times 2}$ e seja A a matriz $\begin{bmatrix}3 & 5\\ 3 & 0\end{bmatrix}$. Para determinar se A é combinação linear de elementos de S, temos de procurar reais c_1, c_2 e c_3 tais que

$$\begin{aligned}\begin{bmatrix}3 & 5\\ 3 & 0\end{bmatrix} &= c_1\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix} + c_2\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} + c_3\begin{bmatrix}1 & 2\\ 1 & 0\end{bmatrix}\\ &= \begin{bmatrix}2c_1 & c_1\\ c_1 & -c_1\end{bmatrix} + \begin{bmatrix}c_2 & 0\\ 0 & -c_2\end{bmatrix} + \begin{bmatrix}c_3 & 2c_3\\ c_3 & 0\end{bmatrix}\\ &= \begin{bmatrix}2c_1 + c_2 + c_3 & c_1 + 2c_3\\ c_1 + c_3 & -c_1 - c_2\end{bmatrix}.\end{aligned}$$

Tendo em conta que duas matrizes são iguais quando todas as suas entradas são iguais, obtemos um sistema de quatro equações a três incógnitas.

$$\begin{cases}2c_1 + c_2 + c_3 = 3\\ c_1 + 2c_3 = 5\\ c_1 + c_3 = 3\\ -c_1 - c_2 = 0\end{cases} \longrightarrow \left[\begin{array}{ccc|c}2 & 1 & 1 & 3\\ 1 & 0 & 2 & 5\\ 1 & 0 & 1 & 3\\ -1 & -1 & 0 & 0\end{array}\right] \begin{array}{l}\\ 2l_2 - l_1\\ 2l_3 - l_1\\ 2l_4 + l_1\end{array} \longrightarrow \left[\begin{array}{ccc|c}2 & 1 & 1 & 3\\ 0 & -1 & 3 & 7\\ 0 & -1 & 1 & 3\\ 0 & -1 & 1 & 3\end{array}\right] \begin{array}{l}\\ \\ l_3 - l_2\\ l_4 - l_2\end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c}2 & 1 & 1 & 3\\ 0 & -1 & 3 & 7\\ 0 & 0 & -2 & -4\\ 0 & 0 & -2 & -4\end{array}\right] \begin{array}{l}\\ \\ \\ l_4 - l_3\end{array} \longrightarrow \left[\begin{array}{ccc|c}2 & 1 & 1 & 3\\ 0 & -1 & 3 & 7\\ 0 & 0 & -2 & -4\\ 0 & 0 & 0 & 0\end{array}\right]$$

$$\longrightarrow \begin{cases}2c_1 + c_2 + c_3 = 3\\ -c_2 + 3c_3 = 7\\ -2c_3 = -4\end{cases} \longrightarrow \begin{cases}c_1 = 1\\ c_2 = -1\\ c_3 = 2\end{cases}$$

donde

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

sendo portanto combinação linear daquelas matrizes.

Exercício 12. Escreva os seguintes vectores como combinação linear de elementos do conjunto $S = \{(0,-2,2),(1,3,-1)\}$.

(a) $(2,2,2)$ (c) $(2,8,-4)$ (e) $(-2,-2,-2)$ (g) $(3,3,3)$

(b) $(1,-1,3)$ (d) $(0,2,-2)$ (f) $(2,6,-2)$ (h) $(0,0,0)$

Exercício 13. Seja $S = \{2+x+4x^2, 1-x+3x^2, 3+2x+5x^2\}$. Escreva os seguintes vectores como combinação linear de elementos de S.

(a) $-9-7x-15x^2$ (d) $7+8x+9x^2$ (g) $-x+10x^2$ (j) $6+2x+6x^2$

(b) $6+11x+6x^2$ (e) $2+2x+4x^2$ (h) $15+3x+9x^2$ (k) 1

(c) 0 (f) $2-3x+6x^2$ (i) $1-2x+x^2$ (l) x^2

Exercício 14. Seja S o conjunto

$$S = \left\{ \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \right\}.$$

Quais dos seguintes vectores são combinação linear de elementos de S?

(a) $\begin{bmatrix} 6 & -2 \\ 2 & 8 \end{bmatrix}$ (b) $\begin{bmatrix} 12 & 0 \\ 0 & 15 \end{bmatrix}$ (c) $\begin{bmatrix} 12 & 0 \\ 0 & 14 \end{bmatrix}$ (d) $\begin{bmatrix} 4 & 2 \\ -1 & 2 \end{bmatrix}$

2.3.2 Espaço gerado por um conjunto de vectores

Partindo dum conjunto S de vectores, podemos agora pensar em *todos* os vectores que se podem construir a partir desses por formação de combinações lineares.

Definição. Sejam V um espaço linear e S um subconjunto não vazio de V. A *expansão linear* de S ou *espaço gerado por* S é o conjunto $L(S)$ de todas as combinações lineares de elementos de S. Diz-se ainda que S é o *conjunto gerador* de V, ou que S *gera* V, e aos elementos de S chama-se *geradores*.

O grande interesse deste conceito é permitir representar elementos dum conjunto tipicamente grande ($L(S)$ é quase sempre infinito) à custa de poucos elementos. Vejamos alguns exemplos.

Exemplo.

1. Em $\mathbb{R}^2$, considere-se o conjunto $S = \{(1,1)\}$. O espaço gerado por S é o conjunto de todas as combinações lineares envolvendo o vector $(1,1)$, ou seja, o conjunto de todos os pontos da forma $c(1,1) = (c,c)$. Portanto $L(S)$ é a bissectriz dos quadrantes ímpares.

2. Em $\mathbb{R}^3$, considere-se o conjunto $S = \{(1,0,0),(1,0,1),(0,0,1)\}$. O espaço gerado por S é o conjunto de todos os vectores $\overrightarrow{u}$ de $\mathbb{R}^3$ da forma
$$\overrightarrow{u} = c_1(1,0,0) + c_2(1,0,1) + c_3(0,0,1) = (c_1 + c_2, 0, c_2 + c_3)\ .$$
É fácil ver que qualquer vector (x,y,z) satisfazendo a condição $y = 0$ é desta forma (pode ser escrito tomando $c_1 = x$, $c_3 = z$ e $c_2 = 0$); então $L(S)$ é o plano $y = 0$.

3. Em $\mathbb{R}^{\mathbb{N}}$, considere-se o conjunto $S = \{u, v\}$ com $u_n = 1$ para todo o n e $v_n = -1$ para todo o n. O espaço gerado por S é o conjunto de todas as sucessões da forma $c_1u + c_2v$, ou seja, das sucessões s tais que
$$s(n) = c_1 + c_2 \text{ para todo o } n.$$
Então $L(S)$ é o conjunto de todas as sucessões constantes.

4. Não é necessário que S seja um conjunto finito, embora cada combinação linear só possa usar um número finito de elementos de S. Na teoria da aproximação, usam-se muitas vezes espaços gerados por conjuntos infinitos de funções. Por exemplo, os espaços P_n são gerados por $\{1, x, x^2, \ldots, x^n\}$, enquanto o espaço $\mathcal{P}$ de todos os polinómios é gerado por todas as potências de x (ou seja, $\mathcal{P} = L\left(\{1, x, x^2, \ldots, x^n, \ldots\}\right)$), que já não é um conjunto finito.

É importante observar que em todos estes casos o conjunto $L(S)$ é um espaço vectorial.

Exercício 15. Descreva os espaços $L(S)$ para os conjuntos S seguintes.

(a) $S = \{(1,0),(0,1)\} \subseteq \mathbb{R}^2$

(b) $S = \{(1,0,0),(0,1,0)\} \subseteq \mathbb{R}^3$

(c) $S = \{(1,0,1),(0,1,-1)\} \subseteq \mathbb{R}^3$

(d) $S = \left\{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\} \subseteq M_{2\times 2}$

(e) $S = \left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right\} \subseteq M_{2\times 2}$

Exercício 16. Verifique que $L(S)$ é um espaço vectorial para todos os conjuntos S no exercício anterior.

O conceito de combinação linear está intimamente relacionado quer com o de espaço gerado por um conjunto, quer com o de espaço linear. Não surpreende portanto que haja fortes relações entre estes dois conceitos.

Teorema 4. Sejam V um espaço linear e S um subconjunto não vazio de V. Então $L(S)$ é subespaço linear de V.

Demonstração. Para verificar que $L(S)$ é subespaço linear de V, basta mostrar que é fechado para as operações de soma e produto por escalares. Dito por outras palavras, temos de mostrar que a soma de duas combinações lineares de elementos de S é ainda um elemento de S e que o produto duma combinação linear de elementos de S por um escalar continua a ser uma combinação linear de elementos de S.

Fecho para a soma. Sejam u e v elementos de $L(S)$. Então u e v são combinações lineares de elementos de S, pelo que existem elementos $s_1, s_2, \ldots, s_n$ e $r_1, r_2, \ldots, r_m$ de S e reais $c_1, c_2, \ldots, c_n$ e $d_1, d_2, \ldots, d_m$ tais que

$$u = c_1s_1 + c_2s_2 + \cdots + c_ns_n \qquad v = d_1r_1 + d_2r_2 + \cdots + d_mr_m\,.$$

Então

$$u + v = c_1s_1 + c_2s_2 + \cdots + c_ns_n + d_1r_1 + d_2r_2 + \cdots + d_mr_m$$

é combinação linear de elementos de S. Logo $u + v \in L(S)$, ou seja, $L(S)$ é fechado para a soma.

Fecho para o produto por escalares. Sejam u como atrás e $\alpha \in \mathbb{R}$. Então

$$\alpha u = \alpha\,(c_1s_1 + c_2s_2 + \cdots + c_ns_n) = \alpha c_1s_1 + \alpha c_2s_2 + \cdots + \alpha c_ns_n\,,$$

pelo que $\alpha u \in L(S)$ e, portanto, $L(S)$ também é fechado para o produto por escalares.

Conclui-se assim que $L(S)$ é subespaço linear de V. □

O conceito de combinação linear também permite caracterizar subespaços lineares de V.

Proposição 27. Sejam V um espaço linear e $W \subseteq V$. Então W é subespaço linear de V se e só se qualquer combinação linear de elementos de W está em W.

Demonstração. Se qualquer combinação linear de elementos de W estiver em W, então W é fechado para a soma e para o produto por escalares, já que estes são casos particulares de combinações lineares; logo W é subespaço linear de V.

Reciprocamente, suponha-se W é fechado para a soma e para o produto por escalares e seja $u = c_1w_1 + c_2w_2 + \cdots + c_nw_n$ uma combinação linear de elementos de W. Como W é fechado para o produto por escalares, os vectores $c_1w_1, c_2w_2, \ldots, c_nw_n$ são todos elementos de W; como W é fechado para a soma, a sua soma também pertence a W. Logo $u \in W$. □

Dito doutra forma: um subconjunto W dum espaço vectorial V é um subespaço linear de V se e só se $L(W) = W$. Este resultado dá-nos ainda outra forma de provar que um conjunto é subespaço vectorial doutro: basta mostrar que é fechado para combinações lineares. Na realidade, basta até mostrar que é fechado para combinações lineares de dois elementos.

2.3.3 Dependência e independência linear

Vimos na secção anterior que pode haver conjuntos relativamente pequenos que geram todos os elementos dum espaço vectorial V. Um exemplo foi o do conjunto $S = \{(1,0),(0,1)\}$, que gera $\mathbb{R}^2$. Nesta secção vamo-nos preocupar com uma questão de certa forma complementar: quando é que podemos *retirar* elementos de um conjunto de geradores mantendo o espaço por eles gerado?

Definição. Sejam V um espaço linear com elemento zero 0_V e S um subconjunto não vazio de V. O conjunto S é *linearmente independente* se para quaisquer elementos $s_1, s_2, \ldots, s_n \in S$ e $c_1, c_2, \ldots, c_n \in \mathbb{R}$, a equação

$$c_1 s_1 + c_2 s_2 + \cdots + c_n s_n = 0_V$$

tem uma única solução $c_1 = c_2 = \cdots = c_n = 0$. Caso contrário, S diz-se *linearmente dependente*.

Por outras palavras: S é linearmente independente se a única combinação linear de elementos de S que representa o vector nulo é a combinação linear com todos os coeficientes nulos.

O conceito de independência linear é fundamental para o que se segue, pelo que é importante discuti-lo com algum cuidado. Vamos começar por analisar alguns exemplos para perceber como é que se verifica se um conjunto é linearmente independente.

Exemplo.

1. Tome-se o subconjunto

$$S = \{(1,2,-1), (0,2,-3), (2,0,1)\}$$

de $\mathbb{R}^3$ e vamos ver se S é linearmente independente. Para tal, temos de estudar as soluções da equação

$$c_1(1,2,-1) + c_2(0,2,-3) + c_3(2,0,1) = (0,0,0)\,.$$

Ora

$$\begin{aligned} c_1(1,2,-1) + c_2(0,2,-3) + c_3(2,0,1) &= (c_1, 2c_1, -c_1) + (0, 2c_2, -3c_2) + (2c_3, 0, c_3) \\ &= (c_1 + 2c_3, 2c_1 + 2c_2, -c_1 - 3c_2 + c_3)\ . \end{aligned}$$

Daqui obtém-se o seguinte sistema de equações.

$$\begin{cases} c_1 + 2c_3 = 0 \\ 2c_1 + 2c_2 = 0 \\ -c_1 - 3c_2 + c_3 = 0 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 2 & 2 & 0 & 0 \\ -1 & -3 & 1 & 0 \end{array}\right] \begin{array}{l} \\ l_2 - 2l_1 \\ l_3 + l_1 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 0 \\ 0 & -3 & 3 & 0 \end{array}\right] \begin{array}{l} \\ \\ 2l_3 + 3l_2 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 2 & -4 & 0 \\ 0 & 0 & -6 & 0 \end{array}\right]$$

Este sistema é homogéneo, possível e determinado, logo a sua única solução é a solução $c_1 = c_2 = c_3 = 0$. Assim, conclui-se que S é linearmente independente.

2. Considere-se agora o espaço linear P_2 e tome-se

$$S = \left\{1 + x, -1 + 4x - x^2, -2 + 3x - x^2\right\}\ .$$

O conjunto S é linearmente independente se a equação

$$c_1(1 + x) + c_2\left(-1 + 4x - x^2\right) + c_3\left(-2 + 3x - x^2\right) = 0$$

tiver como única solução $c_1 = c_2 = c_3 = 0$.

Ora

$$\begin{aligned} & c_1(1+x)+c_2\left(-1+4x-x^2\right)+c_3\left(-2+3x-x^2\right) \\ & \quad = c_1+c_1x-c_2+4c_2x-c_2x^2-2c_3+3c_3x-c_3x^2 \\ & \quad = (c_1-c_2-2c_3)+(c_1+4c_2+3c_3)\,x+(-c_2-c_3)\,x^2 \end{aligned}$$

e este polinómio é o polinómio nulo se todos os seus coeficientes forem 0. Obtemos então o sistema homogéneo de equações

$$\begin{cases} c_1-c_2-2c_3=0 \\ c_1+4c_2+3c_3=0 \\ -c_2-c_3=0 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 1 & 4 & 3 & 0 \\ 0 & -1 & -1 & 0 \end{array}\right] l_2-l_1 \longrightarrow \left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & -1 & -1 & 0 \end{array}\right] l_2+5l_3$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & -5 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

que tem dois pivots para três incógnitas, sendo portanto possível indeterminado. Assim, há vários valores possíveis para c_1, c_2 e c_3, pelo que S é linearmente dependente.

3. Considere-se o subconjunto

$$S=\left\{\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right],\left[\begin{array}{cc} 1 & -1 \\ 1 & 0 \end{array}\right],\left[\begin{array}{cc} 0 & -2 \\ 1 & -1 \end{array}\right],\left[\begin{array}{cc} 3 & 0 \\ 3 & 3 \end{array}\right]\right\}$$

de $M_{2\times 2}$. Para verificar se S é linearmente independente, temos de resolver o sistema

$$c_1\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]+c_2\left[\begin{array}{cc} 1 & -1 \\ 1 & 0 \end{array}\right]+c_3\left[\begin{array}{cc} 0 & -2 \\ 1 & -1 \end{array}\right]+c_4\left[\begin{array}{cc} 6 & 0 \\ 3 & 3 \end{array}\right]=\mathbf{0_{2\times 2}}$$

ou, equivalentemente

$$\left[\begin{array}{cc} c_1+c_2+6c_4 & c_1-c_2-2c_3 \\ c_2+c_3+3c_4 & c_1-c_3+3c_4 \end{array}\right]=\mathbf{0_{2\times 2}}\,.$$

Esta equação equivale a exigir que todas as entradas de ambas as matrizes coincidam, obtendo-se o seguinte sistema de equações.

$$\begin{cases} c_1+c_2+6c_4=0 \\ c_1-c_2-2c_3=0 \\ c_2+c_3+3c_4=0 \\ c_1-c_3+3c_4=0 \end{cases} \longrightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & 6 & 0 \\ 1 & -1 & -2 & 0 & 0 \\ 0 & 1 & 1 & 3 & 0 \\ 1 & 0 & -1 & 3 & 0 \end{array}\right] \begin{array}{c} \\ l_2-l_1 \\ \\ l_4-l_1 \end{array}$$

$$\longrightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 0 \\ 0 & -2 & -2 & -6 & 0 \\ 0 & 1 & 1 & 3 & 0 \\ 0 & -1 & -1 & -3 & 0 \end{array}\right] \begin{array}{c} \\ \\ 2l_3+l_2 \\ 2l_4-l_2 \end{array} \longrightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 0 \\ 0 & -2 & -2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Uma vez que este sistema é possível mas indeterminado (só restam duas equações para quatro variáveis), concluimos que existem várias combinações lineares que permitem obter a matriz nula a partir de elementos de S. Logo S não é linearmente independente.

4. Vamos agora ver se o subconjunto

$$S=\left\{\begin{bmatrix}1&1&2\\0&1&0\end{bmatrix},\begin{bmatrix}1&-1&0\\1&0&-1\end{bmatrix},\begin{bmatrix}1&1&1\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&-1\\-1&-1&-1\end{bmatrix}\right\}$$

de $M_{2\times 3}$ é linearmente independente. Para tal, temos de estudar as soluções de

$$c_1\begin{bmatrix}1&1&2\\0&1&0\end{bmatrix}+c_2\begin{bmatrix}1&-1&0\\1&0&-1\end{bmatrix}+c_3\begin{bmatrix}1&1&1\\0&0&0\end{bmatrix}+c_4\begin{bmatrix}1&0&-1\\-1&-1&-1\end{bmatrix}=\mathbf{0_{2\times 3}}$$

ou, equivalentemente, de

$$\begin{bmatrix}c_1+c_2+c_3+c_4 & c_1-c_2+c_3 & 2c_1+c_3-c_4\\ c_2-c_4 & c_1-c_4 & -c_2-c_4\end{bmatrix}=\mathbf{0_{2\times 3}}\,.$$

Tendo em conta que duas matrizes são iguais quando todas as suas entradas coincidem, obtemos o seguinte sistema de equações.

$$\begin{cases}c_1+c_2+c_3+c_4=0\\ c_1-c_2+c_3=0\\ 2c_1+c_3-c_4=0\\ c_2-c_4=0\\ c_1-c_4=0\\ -c_2-c_4=0\end{cases}\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&0\\1&-1&1&0&0\\2&0&1&-1&0\\0&1&0&-1&0\\1&0&0&-1&0\\0&-1&0&-1&0\end{array}\right]\begin{array}{l}\\ l_2-l_1\\ l_3-2l_1\\ \\ l_5-l_1\\ \\ \end{array}$$

$$\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&0\\0&-2&0&-1&0\\0&-2&-1&-3&0\\0&1&0&-1&0\\0&-1&-1&-2&0\\0&-1&0&-1&0\end{array}\right]\begin{array}{l}\\ \\ l_3-l_2\\ 2l_4+l_2\\ 2l_5-l_2\\ 2l_6-l_2\end{array}\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&0\\0&2&0&1&0\\0&0&-1&-2&0\\0&0&0&-3&0\\0&0&-2&-3&0\\0&0&0&-1&0\end{array}\right]\begin{array}{l}\\ \\ \\ \\ l_5-2l_3\\ \\ \end{array}$$

$$\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&0\\0&2&0&1&0\\0&0&-1&-2&0\\0&0&0&-3&0\\0&0&0&1&0\\0&0&0&-1&0\end{array}\right]\begin{array}{l}\\ \\ \\ \\ 3l_5+l_4\\ 3l_6-l_4\end{array}\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&0\\0&2&0&1&0\\0&0&-1&-2&0\\0&0&0&-3&0\\0&0&0&0&0\\0&0&0&0&0\end{array}\right]$$

que é um sistema possível determinado. Conclui-se portanto que o conjunto S é linearmente dependente.

Exercício 17. Determine se os seguintes subconjuntos de $\mathbb{R}^2$ são linearmente independentes.

(a) $\{(2,3)\}$ (b) $\{(1,0),(0,1)\}$ (c) $\{(2,2),(2,-2)\}$ (d) $\{(1,2),(2,1)\}$

Exercício 18. Determine se os seguintes subconjuntos de $\mathbb{R}^3$ são linearmente independentes.

(a) $\{(1,0,0),(-2,0,0)\}$

(b) $\{(2,1,1),(-1,-2,1),(1,0,0),(2,1,4)\}$

(c) $\{(1,-2,6),(2,-1,0),(-1,1,-2)\}$

(d) $\{(2,1,1),(1,-2,1),(1,0,0)\}$

Exercício 19. Determine se os seguintes subconjuntos de P_n são linearmente independentes.

(a) $\{x+3, x-3\}$ (b) $\{x+3, x, x^2+x+3\}$ (c) $\{2x+3, x-1, 3\}$

Por definição, qualquer subconjunto de um conjunto linearmente independente é linearmente independente. De facto, se $S \subseteq T$ e $c_1s_1 + c_2s_2 + \cdots + c_ns_n$ é uma combinação linear de elementos de S, então também é uma combinação linear de elementos de T; logo, se S for linearmente dependente, existe uma combinação linear de elementos de S com coeficientes não nulos que corresponde ao vector nulo. Esta combinação linear também é uma combinação linear de elementos de T, donde T também é linearmente dependente.

Podemos também caracterizar a independência linear apenas à custa dos elementos do próprio conjunto S.

Proposição 28. Sejam V um espaço linear e S um subconjunto não vazio de V. O conjunto S é linearmente dependente se e só se pelo menos um dos elementos de S é combinação linear dos restantes elementos de S.

Demonstração. Suponha-se que o conjunto S é linearmente dependente. Então existem elementos $s_1, s_2, \ldots, s_n \in S$ e reais $c_1, c_2, \ldots, c_n$ não simultaneamente nulos tais que

$$c_1s_1 + c_2s_2 + \cdots + c_ns_n = 0\,.$$

Podemos assumir que $c_1 \neq 0$ (eventualmente trocando os nomes dados aos coeficientes). Então, lendo a relação acima como uma equação linear e resolvendo-a em ordem a s_1, obtemos

$$s_1 = -\frac{c_2}{c_1}s_2 - \frac{c_3}{c_1}s_3 - \cdots - \frac{c_n}{c_1}s_n\,.$$

Logo, s_1 é combinação linear de $s_2, s_2, \ldots, s_n$.

Reciprocamente, suponha-se que $s \in S$ é combinação linear dos elementos $s_1, s_2, \ldots, s_n$ de S, ou seja, que $s = c_1s_1 + c_2s_2 + \cdots + c_ns_n$ para alguns coeficientes não nulos $c_1, \ldots, c_n$.

Então, temos que

$$c_1s_1 + c_2s_2 + \cdots + c_ns_n - s = 0\,,$$

o que mostra que a combinação linear $c_1s_1 + c_2s_2 + \cdots + c_ns_n + (-1)s$ é uma combinação linear de elementos de S com coeficientes não todos nulos que representa o vector nulo. Logo S é linearmente dependente. □

Podemos ler este teorema por contra-recíproco, obtendo uma caracterização de conjuntos linearmente independentes: um conjunto S é linearmente independente se e só se nenhum dos seus elementos for combinação linear dos restantes. Na prática, há situações em que esta caracterização é simples de usar.

Exemplo.

1. Considere-se o subconjunto $S = \{(1,0),(0,1),(2,2)\}$ de $\mathbb{R}^2$. Uma vez que se tem a relação $(2,2) = 2(1,0)+2(0,1)$, existe um elemento de S que é combinação linear dos restantes. Logo S é linearmente dependente.

2. Considere-se o subconjunto $S = \{(1,1,1),(2,2,2),(1,0,-3)\}$ de $\mathbb{R}^3$. Tem-se a relação $(2,2,2) = 2(1,1,1)$, pelo que há um elemento de S que é combinação linear dos restantes. Logo S é linearmente dependente.

3. Considere-se o subconjunto $S = \{x^2+1, x+1, x^2+x+2\}$ de P_2. Uma vez que se tem $(x^2+1)+(x+1) = x^2+x+2$, existe um elemento de S que é combinação linear dos restantes. Logo S é linearmente dependente.

4. Considere-se o subconjunto $S = \{\mathbf{I_2}, 2\mathbf{I_2}\}$ de $M_{2\times 2}$. Uma vez que $2\mathbf{I_2}$ é combinação linear de $\mathbf{I_2}$, o conjunto S é linearmente dependente.

5. Considere-se o subconjunto $S = \left\{\cos^2(x), \sin^2(x), 1\right\}$ de $\mathcal{F}(\mathbb{R})$. Uma vez que a relação $\cos^2(x)+\sin^2(x) = 1$ é válida para todos os reais, a terceira função de S é combinação linear das duas anteriores. Logo S é linearmente dependente.

Em geral, este resultado só é útil quando há um elemento de S que é claramente combinação linear de outros (tipicamente, um elemento é múltiplo de outro); quando tal não sucede, é à partida mais expedito formular a questão directamente como resolução dum sistema de equações.

Este resultado tem consequências importantes. Se S é um conjunto linearmente dependente, então podemos escolher um elemento $s \in S$ que seja combinação linear dos restantes; o conjunto $S \setminus \{s\}$ que se obtém é um conjunto com menos elementos que ainda gera o mesmo espaço.

Proposição 29. Sejam S um conjunto linearmente dependente e s um elemento de S que possa ser escrito como combinação linear dos restantes elementos de S. Então $L(S) = L(S')$, onde $S' = S \setminus \{s\}$.

Demonstração. Seja $s = c_1s_1 + c_2s_2 + \cdots + c_ns_n$ e seja u um elemento de $L(S)$. Então u é combinação linear de elementos de S, por exemplo

$$u = d_1t_1 + d_2t_2 + \cdots + d_mt_m$$

com $t_1, t_2, \ldots, t_m \in S$.

Se nenhum dos t_i for o elemento s, então u é combinação linear de elementos de S', donde $u \in L(S')$. Suponha-se sem perda de generalidade que $t_1 = s$. Então

$$\begin{aligned} u &= d_1t_1 + d_2t_2 + \cdots + d_mt_m \\ &= d_1\left(c_1s_1 + c_2s_2 + \cdots + c_ns_n\right) + d_2t_2 + \cdots + d_mt_m \\ &= (c_1d_1)\,s_1 + (c_2d_1)\,s_2 + \cdots + (c_nd_1)\,s_n + d_2t_2 + \cdots + d_mt_m \end{aligned}$$

donde u é ainda combinação linear de elementos de S'.

Conclui-se portanto que $L(S) = L(S')$. □

2.4 Bases, coordenadas e dimensão

Na secção anterior, vimos como podemos usar combinações lineares para gerar espaços vectoriais a partir de conjuntos com poucos elementos e discutimos a redundância que pode existir numa tal representação. Nesta secção, vamos discutir o conceito de *base*: um conjunto gerador dum espaço vectorial tão pequeno quanto possível. Veremos que esta minimalidade corresponde não só a exigir que o conjunto seja linearmente independente, como a garantir que cada elemento pode ser escrito duma única forma como combinação linear de elementos da base. É esta propriedade que nos vai permitir trabalhar com muitos espaços vectoriais como trabalhamos com $\mathbb{R}^n$, abstraindo da sua estrutura concreta.

2.4.1 Bases de um espaço linear

Seja V um espaço linear e S um seu subconjunto que gere V, ou seja, tal que $L(S) = V$. Então todo o elemento de V é combinação linear dos elementos de S, donde conhecendo apenas o conjunto S podemos determinar todos os elementos de V. Mas será que precisamos de conhecer *todos* os elementos de S?

Suponhamos que o conjunto S é linearmente dependente. Então pelo menos um dos seus elementos pode ser escrito como combinação linear dos restantes elementos de S. Vimos no final da secção anterior que retirando esse elemento de S obtemos um conjunto mais pequeno que ainda gera S. Por outro lado, se não houver nenhum elemento nessas condições, é claro que não podemos retirar elementos a S: o elemento que retirássemos seria um elemento de V que não era gerado pelo novo conjunto.

Obtemos então a seguinte definição fundamental.

Definição. Seja V um espaço linear. Um subconjunto S de V é uma *base* de V se S é linearmente independente e $L(S) = V$.

Diz-se que o espaço linear V tem *dimensão finita* se admite uma base S com um número finito de elementos, que tem *dimensão infinita* se admite uma base com um número infinito de elementos, e que tem *dimensão nula* se admite uma base vazia (ou seja, se $V = \{0\}$).

Estas definições não trazem nada de novo, a não ser a unificação entre diferentes conceitos que já abordámos nas secções anteriores. Vejamos alguns exemplos.

Exemplo.

1. O conjunto $S = \{(1,0),(0,1)\}$ é uma base de $\mathbb{R}^2$.

   - S é linearmente independente.

   $$c_1(1,0) + c_2(0,1) = (0,0) \longrightarrow (c_1,c_2) = (0,0) \longrightarrow \begin{cases} c_1 = 0 \\ c_2 = 0 \end{cases}$$

   - $L(S) = \mathbb{R}^2$: temos de mostrar que qualquer vector (x,y) pode ser escrito como combinação linear dos elementos de S.

   $$c_1(1,0) + c_2(0,1) = (x,y) \longrightarrow (c_1,c_2) = (x,y) \longrightarrow \begin{cases} c_1 = x \\ c_2 = y \end{cases}$$

   Logo S é uma base de $\mathbb{R}^2$, dita a *base canónica* de $\mathbb{R}^2$.

2. O conjunto $S = \{(1,1,1),(1,1,0),(1,0,0)\}$ é uma base de $\mathbb{R}^3$.

   - S é linearmente independente.

   $$c_1(1,1,1)+c_2(1,1,0)+c_3(1,0,0)=(0,0,0) \longrightarrow (c_1+c_2+c_3, c_1+c_2, c_1)=(0,0,0)$$

   $$\longrightarrow \begin{cases} c_1+c_2+c_3=0 \\ c_1+c_2=0 \\ c_1=0 \end{cases} \longrightarrow \begin{cases} c_1=0 \\ c_2=0 \\ c_3=0 \end{cases}$$

   - $L(S)=\mathbb{R}^3$: temos de mostrar que todo o vector (x,y,z) de $\mathbb{R}^3$ pode ser escrito como combinação linear dos elementos de S.

   $$c_1(1,1,1)+c_2(1,1,0)+c_3(1,0,0)=(x,y,z)$$

   $$\longrightarrow (c_1+c_2+c_3, c_1+c_2, c_1)=(x,y,z) \longrightarrow \begin{cases} c_1+c_2+c_3=x \\ c_1+c_2=y \\ c_1=z \end{cases}$$

   $$\longrightarrow \begin{cases} z+(y-z)+c_3=x \\ c_2=y-z \\ c_1=z \end{cases} \longrightarrow \begin{cases} c_3=x-y \\ c_2=y-z \\ c_1=z \end{cases}$$

   Logo S é base de $\mathbb{R}^3$.

3. O conjunto

   $$S=\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$$

   é uma base do espaço $M_{2\times 2}$.

   - S é linearmente independente.

   $$c_1\begin{bmatrix}1&0\\0&0\end{bmatrix}+c_2\begin{bmatrix}0&1\\0&0\end{bmatrix}+c_3\begin{bmatrix}0&0\\1&0\end{bmatrix}+c_4\begin{bmatrix}0&0\\0&1\end{bmatrix}=\mathbf{0_{2\times 2}}$$

   $$\longrightarrow \begin{bmatrix}c_1&c_2\\c_3&c_4\end{bmatrix}=\mathbf{0_{2\times 2}} \longrightarrow \begin{cases} c_1=0, c_2=0 \\ c_3=0, c_4=0 \end{cases}$$

   - $L(S)=M_{2\times 2}$: temos de mostrar que qualquer matriz $\begin{bmatrix}a&b\\c&d\end{bmatrix}$ pode ser escrita como combinação linear dos elementos de S.

   $$c_1\begin{bmatrix}1&0\\0&0\end{bmatrix}+c_2\begin{bmatrix}0&1\\0&0\end{bmatrix}+c_3\begin{bmatrix}0&0\\1&0\end{bmatrix}+c_4\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$$

   $$\longrightarrow \begin{bmatrix}c_1&c_2\\c_3&c_4\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix} \longrightarrow \begin{cases} c_1=a, c_2=b \\ c_3=c, c_4=d \end{cases}$$

   Logo S é uma base de $M_{2\times 2}$, dita a *base canónica* de $M_{2\times 2}$.

4. O conjunto $S=\{1,x,x^2\}$ é uma base do espaço P_2 dos polinómios de grau menor ou igual a 2.

- S é linearmente independente.

$$c_1 1 + c_2 x + c_3 x^2 = 0 \longrightarrow \begin{cases} c_1 = 0 \\ c_2 = 0 \\ c_3 = 0 \end{cases}$$

- $L(S) = P_2$: temos de mostrar que todo o polinómio p de grau menor ou igual a 2 pode ser escrito como combinação linear dos elementos de S — o que é imediato, pois $p(x)$ tem a forma $ax^2 + bx + c$.

Logo S é base de P_2, dita a *base canónica* de P_2.

5. Vimos anteriormente que o conjunto $S = \{1 + x, -1 + 4x - x^2, -2 + 3x - x^2\}$ não é linearmente independente, logo não pode ser uma base de P_2.

Os exemplos acima permitem concluir que $\mathbb{R}^2$, $\mathbb{R}^3$, $M_{2\times 2}$ e P_2 são espaços de dimensão finita. Já os espaços $\mathbb{R}^{\mathbb{N}}$, $\mathcal{F}(\mathbb{R})$, $\mathcal{C}^1(\mathbb{R})$ e $\mathcal{P}$ são espaços vectoriais de dimensão infinita.

Exercício 20.

(a) Indique quais dos conjuntos do Exercício 17 são bases de $\mathbb{R}^2$.

(b) Indique quais dos conjuntos do Exercício 18 são bases de $\mathbb{R}^3$.

(c) Indique quais dos conjuntos do Exercício 19 são bases de P_2 e quais são bases de P_3.

A partir deste ponto, vamos concentrar o nosso estudo em espaços vectoriais de dimensão finita.

2.4.2 Vectores e coordenadas

O conceito de base generaliza o conceito de referencial em $\mathbb{R}^n$, no sentido em que cada elemento do espaço é representado por vectores de coordenadas numa base do espaço. No caso de $\mathbb{R}^n$, um vector é usualmente escrito em coordenadas, sendo a base o conjunto dos vectores unitários sobre os eixos. Por exemplo, em $\mathbb{R}^2$ temos a base $\{(1,0),(0,1)\}$, constituída pelos vectores unitários sobre os eixos dos xx e dos yy. Em $\mathbb{R}^3$, temos a base $\{(1,0,0),(0,1,0),(0,0,1)\}$ constituída pelos vectores unitários sobre os eixos dos xx, yy e zz. Às bases de $\mathbb{R}^n$ que generalizam esta construção chamamos *bases canónicas* de $\mathbb{R}^n$.

Escrever um vector de $\mathbb{R}^n$ em coordenadas não é mais do que escrevê-lo de forma resumida como combinação linear dos vectores da base, indicando apenas os coeficientes da combinação. Por exemplo, o vector $(2,1)$ em $\mathbb{R}^2$ corresponde à combinação linear $2(1,0) + 1(0,1)$; em $\mathbb{R}^3$, o vector $(1,0,-1)$ corresponde à combinação linear $1(1,0,0) + 0(0,1,0) + (-1)(0,0,1)$. Esta forma de representar vectores pode ser generalizada a qualquer espaço vectorial: fixada uma base deste espaço, qualquer vector pode ser escrito como combinação linear dessa base, servindo esses coeficientes como coordenadas do vector. Esta representação goza de uma propriedade fundamental, habitual em $\mathbb{R}^n$: é única.

Proposição 30. Sejam V um espaço linear e $B = \{v_1, v_2, \ldots, v_n\}$ uma base de V. Então qualquer vector $u \in V$ pode escrever-se de forma única como combinação linear de elementos de B.

Demonstração. Seja u um elemento de V. Como B é uma base de V, então u pode escrever-se como combinação linear dos elementos de S na forma

$$u = c_1v_1 + c_2v_2 + \cdots + c_nv_n$$

para determinados números reais $c_1, c_2, \ldots, c_n$.

Suponhamos agora que havia outra forma de escrever u como combinação linear dos vectores $v_1, v_2, \ldots, v_n$, por exemplo

$$u = d_1v_1 + d_2v_2 + \cdots + d_nv_n$$

com $d_1, d_2, \ldots d_n$ números reais. Subtraindo estas duas igualdades, concluímos que

$$\underbrace{u - u}_{0} = (c_1 - d_1)\, v_1 + (c_2 - d_2)\, v_2 + \cdots + (c_n - d_n)\, v_n\,.$$

Ora B é linearmente independente. Então a combinação linear do lado direito desta equação tem necessariamente todos os coeficientes nulos, donde $c_i - d_i = 0$ para $i = 1, 2, \ldots, n$. Então $c_1 = d_1, c_2 = d_2, \ldots, c_n = d_n$. Conclui-se portanto que só existe uma forma de escrever u como combinação linear de elementos de B. $\square$

Definição. Sendo V um espaço linear com base $B = \{v_1, v_2, \ldots, v_n\}$, existe uma única forma de escrever u como combinação linear de elementos de B. Seja

$$u = c_1v_1 + c_2v_2 + \cdots + c_nv_n\,.$$

Os números $c_1, c_2, \ldots, c_n$ chamam-se as *coordenadas* de u na base B e escrevemos

$$(u)_B = (c_1, c_2, \ldots, c_n)\ .$$

A determinação das coordenadas dum vector é novamente um problema de combinações lineares que se reduz à resolução dum sistema de equações lineares. Por exemplo, se quisermos saber quais as coordenadas do vector $(1, -2, 3)$ na base $B = \{(1, -1, 2), (0, 1, 1), (-2, 1, 1)\}$ de $\mathbb{R}^3$, temos de encontrar os coeficientes c_1, c_2 e c_3 tais que

$$c_1(1, -1, 2) + c_2(0, 1, 1) + c_3(-2, 1, 1) = (1, -2, 3)\,.$$

Uma vez que

$$c_1(1, -1, 2) + c_2(0, 1, 1) + c_3(-2, 1, 1) = (c_1 - 2c_3, -c_1 + c_2 + c_3, 2c_1 + c_2 + c_3)\ ,$$

obtemos o seguinte sistema de equações lineares.

$$\begin{cases} c_1 - 2c_3 = 1 \\ -c_1 + c_2 + c_3 = -2 \\ 2c_1 + c_2 + c_3 = 3 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ -1 & 1 & 1 & -2 \\ 2 & 1 & 1 & 3 \end{array}\right] \begin{array}{l} \\ l_1 + l_2 \\ l_3 - 2l_1 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & 5 & 1 \end{array}\right] \begin{array}{l} \\ \\ l_3 - l_2 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 6 & 2 \end{array}\right] \longrightarrow \begin{cases} c_1 - 2c_3 = 1 \\ c_2 - c_3 = -1 \\ 6c_3 = 2 \end{cases} \longrightarrow \begin{cases} c_1 - 2c_3 = 1 \\ c_2 - c_3 = -1 \\ c_3 = \frac{1}{3} \end{cases} \longrightarrow \begin{cases} c_1 = \frac{5}{3} \\ c_2 = -\frac{2}{3} \\ c_3 = \frac{1}{3} \end{cases}$$

donde $((1, -2, 3))_B = \left(\frac{5}{3}, -\frac{2}{3}, \frac{1}{3}\right)$.

É preciso ter algum cuidado quando se trabalha com bases e coordenadas. Vimos atrás que

$$B_1 = \{(1,1,1),(1,1,0),(1,0,0)\}$$

era uma base de $\mathbb{R}^3$; contudo, este espaço também admite a base canónica

$$B_2 = \{(1,0,0),(0,1,0),(0,0,1)\} .$$

Considere-se o vector $(4,3,5)$. Para o escrever em coordenadas da base B_1, temos de resolver o sistema de equações

$$c_1(1,1,1)+c_2(1,1,0)+c_3(1,0,0) = (4,3,5) \longrightarrow (c_1+c_2+c_3, c_1+c_2, c_1) = (4,3,5)$$
$$\longrightarrow \begin{cases} c_1+c_2+c_3 = 4 \\ c_1+c_2 = 3 \\ c_1 = 5 \end{cases} \longrightarrow \begin{cases} 3+c_3 = 4 \\ 5+c_2 = 3 \\ c_1 = 5 \end{cases} \longrightarrow \begin{cases} c_3 = 1 \\ c_2 = -2 \\ c_1 = 5 \end{cases}$$

donde se conclui que

$$(4,3,5) = 5(1,1,1) - 2(1,1,0) + (1,0,0)$$

e portanto

$$((4,3,5))_{B_1} = (5,-2,1) .$$

Por outro lado, tem-se trivialmente

$$(4,3,5) = 4(1,0,0) + 3(0,1,0) + 5(0,0,1) ,$$

donde na base canónica se tem

$$((4,3,5))_{B_2} = (4,3,5) .$$

Este exemplo aponta um pormenor importante: as coordenadas dum vector são únicas *fixada a base*. É fundamental ter o cuidado de, num mesmo problema, trabalhar sempre com a mesma base em cada espaço.

Vejamos mais alguns exemplos, agora envolvendo polinómios.

Exemplo.

1. Consideremos os seguintes elementos de P_2.

$$p(x) = x^2+3x-2 \qquad q(x) = (2x+3)(3x-1) \qquad r(x) = x^2-9$$

Fixemos a base canónica $B = \{1, x, x^2\}$ de P_2. A forma como habitualmente escrevemos os polinómios torna extremamente simples representá-los em coordenadas desta base.

$$\begin{aligned} (p(x))_B &= (-2,3,1) \\ (q(x))_B &= \left(6x^2+7x-3\right)_B = (-3,7,6) \\ (r(x))_B &= (-9,0,1) \end{aligned}$$

Observe-se que os polinómios de B surgem por ordem crescente de grau, pelo que as coordenadas surgem na ordem inversa à que estamos habituados.

2. Pensando agora na base $C = \{x^2, x, 1\}$ de P_2 (que difere da base canónica apenas na ordem pela qual os polinómios surgem), as coordenadas dos mesmos três polinómios serão agora

$$(p(x))_C = (1, 3, -2) \qquad (q(x))_C = (6, 7, -3) \qquad (r(x))_C = (1, 0, -9)\,.$$

3. Finalmente, tomando agora a base $D = \{x^2 - 1, 2x^2 + 2x, x + 4\}$, vamos escrever os mesmos três polinómios em coordenadas na base D. Para tal, temos de resolver três sistemas de equações lineares.

$$\begin{aligned}
&c_1\left(x^2-1\right) + c_2\left(2x^2+2x\right) + c_3(x+4) = p(x)\\
&\quad\longrightarrow (-c_1+4c_3) + (2c_2+c_3)\,x + (c_1+2c_2)\,x^2 = -2+3x+x^2\\
&\quad\longrightarrow \begin{cases} -c_1+4c_3 = -2\\ 2c_2+c_3 = 3\\ c_1+2c_2 = 1\end{cases}
\end{aligned}$$

$$\begin{aligned}
&c_1\left(x^2-1\right) + c_2\left(2x^2+2x\right) + c_3(x+4) = q(x)\\
&\quad\longrightarrow (-c_1+4c_3) + (2c_2+c_3)\,x + (c_1+2c_2)\,x^2 = -3+7x+6x^2\\
&\quad\longrightarrow \begin{cases} -c_1+4c_3 = -3\\ 2c_2+c_3 = 7\\ c_1+2c_2 = 6\end{cases}
\end{aligned}$$

$$\begin{aligned}
&c_1\left(x^2-1\right) + c_2\left(2x^2+2x\right) + c_3(x+4) = r(x)\\
&\quad\longrightarrow (-c_1+4c_3) + (2c_2+c_3)\,x + (c_1+2c_2)\,x^2 = -9+x^2\\
&\quad\longrightarrow \begin{cases} -c_1+4c_3 = -9\\ 2c_2+c_3 = 0\\ c_1+2c_2 = 1\end{cases}
\end{aligned}$$

Uma vez que estes sistemas têm todos a mesma matriz de coeficientes A, podemos começar por calcular a inversa dessa matriz usando por exemplo a regra dos cofactores.

$$A = \begin{bmatrix} -1 & 0 & 4\\ 0 & 2 & 1\\ 1 & 2 & 0\end{bmatrix} \qquad A^{-1} = \frac{1}{\det(A)}\operatorname{cof}(A)^T = -\frac{1}{6}\begin{bmatrix} -2 & 8 & -8\\ 1 & -4 & 1\\ -2 & 2 & -2\end{bmatrix}$$

Então

$$\begin{aligned}
(p(x))_D &= A^{-1}\begin{bmatrix} -2\\ 3\\ 1\end{bmatrix} = \left(-\frac{10}{3}, \frac{13}{6}, -\frac{4}{3}\right)\\
(q(x))_D &= A^{-1}\begin{bmatrix} -3\\ 7\\ 6\end{bmatrix} = \left(-\frac{7}{3}, \frac{25}{6}, -\frac{4}{3}\right)\\
(r(x))_D &= A^{-1}\begin{bmatrix} -9\\ 0\\ 1\end{bmatrix} = \left(-\frac{5}{3}, \frac{4}{3}, -\frac{8}{3}\right)
\end{aligned}$$

Exercício 21. Determine as coordenadas do vector $(3, 2)$ em relação a cada uma das seguintes bases de $\mathbb{R}^2$.

(a) $\{(1, 0), (0, 1)\}$ (c) $\{(1, 0), (1, 1)\}$ (e) $\{(2, 3), (3, 2)\}$ (g) $\{(2, 2), (0, 1)\}$

(b) $\{(1, 0), (0, 2)\}$ (d) $\{(1, 1), (1, -1)\}$ (f) $\{(1, 2), (-2, 1)\}$ (h) $\{(-1, -1), (-2, 1)\}$

Exercício 22. Determine as coordenadas da matriz

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

em relação a cada uma das seguintes bases de $M_{2\times 2}$.

(a) $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

(b) $\left\{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right\}$

(d) $\left\{\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

(e) $\left\{\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

(f) $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\}$

Por outro lado, sabendo as coordenadas de um vector numa base é extremamente simples determiná-lo — basta calcular a combinação linear respectiva. Por exemplo, o vector $\overrightarrow{u}$ cujas coordenadas na mesma base S são $(4, 1, 0)$ é

$$\overrightarrow{u} = 4(1, -1, 2) + 1(0, 1, 1) + 0(-2, 1, 1) = (4, -3, 9)\,.$$

Exemplo.

1. Considerando novamente o espaço P_2 e a sua base canónica B, os polinómios cujas coordenadas são $(1, 3, 2)$, $(-2, 1, 4)$ e $(3, -1, -2)$ são respectivamente

$$(1, 3, 2) = 1 + 3x + 2x^2 \qquad (-2, 1, 4) = -2 + x + 4x^2 \qquad (3, -1, -2) = 3 - x - 2x^2\,.$$

2. Pensando agora na base C para o mesmo espaço definida no exemplo anterior, as mesmas coordenadas correspondem agora aos seguintes polinómios.

$$(1, 3, 2) = x^2 + 3x + 2 \qquad (-2, 1, 4) = -2x^2 + x + 4 \qquad (3, -1, -2) = 3x^2 - x - 2$$

3. Finalmente, com a base D do exemplo anterior, os polinómios representados pelas mesmas coordenadas são agora os seguintes.

$$\begin{aligned}(1,3,2)&=\left(x^2-1\right)+3\left(2x^2+2x\right)+2(x+4)=7x^2+8x+7\\(-2,1,4)&=-2\left(x^2-1\right)+\left(2x^2+2x\right)+4(x+4)=10x+14\\(3,-1,-2)&=3\left(x^2-1\right)-\left(2x^2+2x\right)-2(x+4)=-x^2-4x-11\end{aligned}$$

Exercício 23. Determine o vector que tem coordenadas $(3,2)$ em relação a cada uma das bases do Exercício 21.

Exercício 24. Determine a matriz cujas coordenadas são $(1,1,0,-2)$ em relação a cada uma das bases do Exercício 22.

Trabalhar com coordenadas torna todos os espaços lineares semelhantes a $\mathbb{R}^n$. Consideremos novamente o espaço P_2 com a sua base canónica. Vimos no exemplo da página 111 que os polinómios p, q e r podiam ser expressos em coordenadas dessa base por

$$(p(x))_B=(-2,3,1)\qquad(q(x))_B=(-3,7,6)\qquad(r(x))_B=(-9,0,1)\,.$$

A partir desta representação, podemos trabalhar com estes vectores através das suas coordenadas *como se estivéssemos em* $\mathbb{R}^3$. Por exemplo, se quisermos calcular $3p(x)+5q(x)-2r(x)$, podemos simplesmente calcular

$$3p(x)+5q(x)-2r(x)=3(-2,3,1)+5(-3,7,6)-2(-9,0,1)=(-9,46,34)$$

obtendo um resultado em coordenadas da base B. Conclui-se portanto que

$$3p(x)+5q(x)-2r(x)=-9+46x+34x^2\,,$$

facto que podemos verificar através de um cálculo directo.

É importante salientar que este raciocínio não depende da base escolhida. No mesmo exemplo, definimos uma base $D=\{x^2-1,2x^2+2x,x+4\}$ do mesmo espaço e concluímos que as coordenadas de p, q e r nessa base eram

$$(p(x))_D=\left(-\frac{10}{3},\frac{13}{6},-\frac{4}{3}\right)\qquad(q(x))_D=\left(-\frac{7}{3},\frac{25}{6},-\frac{4}{3}\right)\qquad(r(x))_D=\left(-\frac{5}{3},\frac{4}{3},-\frac{8}{3}\right)\,.$$

Podemos novamente usar estas coordenadas para calcular

$$\begin{aligned}3p(x)+5q(x)-2r(x)&=3\left(-\frac{10}{3},\frac{13}{6},-\frac{4}{3}\right)+5\left(-\frac{7}{3},\frac{25}{6},-\frac{4}{3}\right)-2\left(-\frac{5}{3},\frac{4}{3},-\frac{8}{3}\right)\\&=\left(\frac{5}{3},\frac{74}{3},\frac{16}{3}\right)\end{aligned}$$

e de facto o polinómio cujas coordenadas na base D são $\left(\frac{5}{3},\frac{74}{3},\frac{16}{3}\right)$ é

$$\frac{5}{3}\left(x^2-1\right)+\frac{74}{3}\left(2x^2+2x\right)+\frac{16}{3}(x+4)=-9+46x+34x^2\,.$$

Exercício 25. Considere a base $S = \{1+x, 1-x^2, 1+x+x^2\}$ de P_2. Escreva o polinómio que se obtém somando os polinómios com as seguintes coordenadas nesta base.

(a) $(2,2,2)$ e $(1,-1,3)$ (b) $(2,8,-4)$ e $(0,2,-2)$ (c) $(-2,-2,-2)$ e $(2,6,-2)$

2.4.3 Dimensão dum espaço linear

Os exemplos apresentados no parágrafo anterior apresentam todos uma particularidade importante: escolhidas duas bases distintas para um espaço vectorial, a representação dum mesmo vector em coordenadas altera-se, mas todas as representações têm o mesmo número de coordenadas. Dito doutra forma, em todos os exemplos que vimos, bases diferentes para um mesmo espaço vectorial V tinham o mesmo número de elementos. O resultado seguinte indica que este facto não é uma coincidência: o número de elementos numa base dum espaço vectorial V é uma propriedade *do espaço* e não das suas bases.

Teorema 5 (Teorema da Dimensão)**.** Sejam V um espaço linear e $B = \{v_1, v_2, \ldots, v_n\}$ uma base de V. Então verificam-se as duas propriedades seguintes.

1. Qualquer subconjunto de V com mais de n elementos é linearmente dependente.

2. Nenhum subconjunto de V com menos de n elementos gera V.

Demonstração.

1. Seja $S = \{s_1, s_2, \ldots, s_m\}$ um subconjunto de V mais do que n elementos; por outras palavras, $m > n$. Como B é base de V, todo o elemento de S é combinação linear dos elementos de V. Podemos então escrever as seguintes m relações.

$$\begin{cases} s_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ s_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \qquad\vdots \\ s_m = a_{m1}v_1 + a_{m2}v_2 + \ldots + a_{mn}v_n \end{cases}$$

Queremos mostrar que S é linearmente dependente, ou seja, que existem números reais $c_1, c_2, \ldots, c_m$ não simultaneamente nulos tais que

$$c_1s_1 + c_2s_2 + \cdots + c_ms_m = 0\,.$$

Ora, substituindo $s_1, s_2, \ldots, s_m$ pelas respectivas representações na base B, obtemos

$$(c_1a_{11} + c_2a_{21} + \cdots + c_ma_{m1})\,v_1 + \cdots + (c_1a_{1n} + c_2a_{2n} + c_ma_{mn})\,v_n = 0\,.$$

Como B é linearmente independente, todos os coeficientes de $v_1, v_2, \ldots, v_n$ têm de ser simultaneamente nulos. A equação anterior é então equivalente ao seguinte sistema.

$$\begin{cases} a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m = 0 \\ a_{12}c_1 + a_{22}c_2 + \cdots + a_{m2}c_m = 0 \\ \qquad\vdots \\ a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m = 0 \end{cases}$$

Ora este sistema é um sistema homogéneo de equações lineares, logo é possível (vimos que admite sempre a solução identicamente nula). Porém, o sistema tem apenas n equações para m variáveis, com $n < m$, pelo que é um sistema possível indeterminado. Então existem várias combinações lineares de elementos de S que representam o vector nulo, donde S é linearmente dependente.

2. Seja agora $S = \{s_1, s_2, \ldots, s_m\}$ um subconjunto de V com menos do que n elementos (ou seja, $m < n$) e suponhamos que S gera V.

   Há duas possibilidades a considerar. Se S é linearmente independente, então S é base de V, donde a alínea anterior implica que B é linearmente dependente, uma vez que tem mais elementos que S. Mas B é uma base de V, logo não pode ser linearmente dependente; conclui-se então que S não pode ser linearmente independente.

   Mas sendo S linearmente dependente, podemos construir um seu subconjunto S' que é linearmente independente retirando os elementos de S que são combinação linear dos restantes (note-se que S é finito, logo este processo termina). Este conjunto S' é ainda um conjunto de menos do que n elementos (porque é subconjunto de S) que gera S (pela Proposição 29) e é linearmente independente, pelo que é novamente uma base de V. Pela alínea anterior, novamente isto contradiz o facto de B ser uma base de V, o que é absurdo.

   Conclui-se então que em qualquer dos casos chegamos a uma contradição. Logo S não pode ser um conjunto gerador de V.

$\square$

A consequência fundamental deste último resultado é que se V tem uma base com n elementos, então todas as bases de V têm precisamente n elementos. Em particular, todas as bases de um espaço linear de dimensão finita têm o mesmo número de elementos. Este número, que é uma propriedade do espaço vectorial V, chama-se *dimensão* de V.

Definição. Seja V um espaço linear de dimensão finita. A *dimensão* de V é o número de elementos de qualquer base de V, denotada por $\dim(V)$.

Vejamos alguns exemplos relativamente aos espaços que já conhecemos.

Exemplo.

1. A base canónica de $\mathbb{R}^2$, $\{(1,0),(0,1)\}$, tem 2 elementos, logo $\dim(\mathbb{R}^2) = 2$.

2. A base canónica de $\mathbb{R}^3$, $\{(1,0,0),(0,1,0),(0,0,1)\}$, tem 3 elementos, logo $\dim(\mathbb{R}^3) = 3$.

3. A base canónica de $\mathbb{R}^n$, $\{(1,0,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,0,\ldots,0,1)\}$, tem n elementos, logo $\dim(\mathbb{R}^n) = n$.

4. A base canónica de $M_{2\times 2}$ é

$$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$$

   logo $\dim(M_{2\times 2}) = 4$.

5. A base canónica de $M_{2\times 3}$ é
$$\left\{\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}\right\}$$
logo $\dim(M_{2\times 3}) = 6$.

6. A base canónica de $M_{n\times m}$ contém todas as matrizes que têm exactamente uma entrada igual a 1 e as restantes iguais a 0. Há $n \times m$ destas matrizes, logo $\dim(M_{n\times m}) = nm$.

7. O espaço P_2 dos polinómios de grau menor ou igual a 2 tem base canónica $\{1, x, x^2\}$, logo $\dim(P_2) = 3$.

8. O espaço P_3 dos polinómios de grau menor ou igual a 3 tem base canónica $\{1, x, x^2, x^3\}$, logo $\dim(P_3) = 4$.

9. Em geral, o espaço P_n dos polinómios de grau menor ou igual a n é $\{1, x, x^2, \ldots, x^n\}$, que contém $n + 1$ elementos; logo $\dim(P_n) = n + 1$.

Exercício 26. Escreva as bases canónicas dos seguintes espaços e verifique que a sua dimensão é a que foi indicada no exemplo anterior.

(a) P_1 (b) $\mathbb{R}^4$ (c) $M_{1\times 3}$ (d) $M_{3\times 2}$

O conceito de dimensão aplica-se também aos subespaços destes espaços vectoriais.

Exemplo.

1. Considere-se a recta R dos pontos do plano satisfazendo a condição $y = 2x$, que como vimos atrás é um subespaço vectorial de $\mathbb{R}^2$. Esta recta é gerada pelo conjunto $\{(1, 2)\}$, que sendo um conjunto singular é automaticamente linearmente independente. Logo uma base de R é o conjunto $\{(1, 2)\}$. Como este conjunto é singular, conclui-se que $\dim(R) = 1$.

2. Considere-se o conjunto P dos polinómios p de grau menor ou igual a 2 que não têm termo independente. Estes polinómios são da forma $ax^2 + bx$, pelo que o conjunto $\{x^2, x\}$ claramente gera P. Mais, sendo este conjunto um subconjunto da base canónica de P_2, é linearmente independente, logo é uma base de P. Então $\dim(P) = 2$.

Mais adiante veremos formas expeditas de encontrar uma base para $L(S)$, sendo S um conjunto arbitrário, que nos permitirão em particular determinar rapidamente a dimensão dum espaço gerado por um conjunto de vectores.

2.4.4 Mudança de base

Imaginemos agora que conhecemos duas bases B e C dum mesmo espaço vectorial V. Suponhamos ainda que conhecemos as coordenadas de um determinado vector u na base B e queremos calcular as suas coordenadas na base C.

Este problema pode ser resolvido em dois passos. Primeiro, calculamos explicitamente u, calculando a combinação linear cujos coeficientes são as coordenadas de u na base B. De seguida, conhecendo u, calculamos as suas coordenadas na base C resolvendo um sistema de equações lineares.

Este método funciona bem se pretendermos converter apenas um vector. Porém, se precisarmos de fazer esta mudança de coordenadas para vários vectores, a quantidade de cálculos necessários é demasiado elevada. Interessa então dispor de um método mais eficiente para realizar esta operação.

Fixemos um espaço vectorial V de dimensão n com duas bases $B = \{v_1, v_2, \ldots, v_n\}$ e $C = \{w_1, w_2, \ldots, w_n\}$ e seja u um elemento deste espaço cujas coordenadas na base B são

$$(u)_B = (c_1, c_2, \ldots, c_n) .$$

Uma vez que C também é uma base de V, podemos escrever os vectores v_1, v_2, …, v_n como vectores de coordenadas na base C, obtendo as relações

$$\begin{cases} v_1 = a_{11}w_1 + a_{12}w_2 + \cdots + a_{1n}w_n \\ v_2 = a_{21}w_1 + a_{22}w_2 + \cdots + a_{2n}w_n \\ \qquad \vdots \\ v_n = a_{n1}w_1 + a_{n2}w_2 + \cdots + a_{nn}w_n \end{cases}$$

onde os coeficientes a_{ij} são únicos.

Ora o facto de dispormos das coordenadas de u na base B indica-nos que u pode ser escrito como combinação linear dos vectores de B, nomeadamente como

$$u = c_1v_1 + c_2v_2 + \cdots + c_nv_n .$$

Substituindo as expressões anteriores para v_1, v_2, …, v_n nesta igualdade obtemos

$$\begin{aligned} u &= c_1 (a_{11}w_1 + a_{21}w_2 + \cdots + a_{n1}w_n) + \\ &\qquad + c_2 (a_{12}w_1 + a_{22}w_2 + \cdots + a_{n2}w_n) + \cdots + c_n (a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{nn}w_n) \\ &= (a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n) w_1 + \\ &\qquad + (a_{12}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n) w_2 + (a_{1n}c_1 + a_{n2}c_2 + \cdots + a_{nn}c_n) w_n \end{aligned}$$

e portanto as coordenadas de u na base C são

$$(u)_C = (a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n, a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n, a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nn}c_n) .$$

Designando estas coordenadas por c'_1, c'_2, …, c'_n, podemos escrever estas relações matricialmente como

$$\begin{bmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

ou, simbolicamente,

$$(u)_C = P_{CB}(u)_B ,$$

onde as colunas da matriz P_{CB} contêm as coordenadas dos vectores da base B quando escritos na base C.

Uma vez que este raciocínio é válido para qualquer vector u, a forma eficiente de resolver o problema de converter vários vectores expressos em coordenadas duma base B para coordenadas duma outra base C é começar por escrever a matriz P_{CB}, já que os cálculos ficam reduzidos a produtos matriciais. Obviamente que só há interesse em seguir esta técnica quando o número de vectores a converter é elevado (tipicamente superior a n); porém, a matriz P_{CB} tem outras aplicações, como veremos mais adiante.

Definição. Sejam V um espaço vectorial de dimensão finita e B e C duas bases para V. À matriz P_{CB} cujas colunas são os vectores da base B escritos em coordenadas da base C chama-se *matriz de mudança de base* de B para C.

É comum representar esquematicamente a matriz P_{CB} da forma seguinte.

$$P_{CB} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ (v_1)_C & (v_2)_C & \ldots & (v_n)_C \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

Os cálculos que efectuámos antes desta definição constituem a prova do seguinte resultado.

Teorema 6 (Mudança de base)**.** Sejam V um espaço linear de dimensão n, B e C duas bases para V e u um vector de V. Então $(u)_C = P_{CB}(u)_B$, onde P_{CB} é a matriz de mudança de base de B para C.

Note-se que a prova que fizemos deste resultado foi completamente dedutiva, concluindo-se também que P_{CB} tal como foi definida é a *única* matriz com aquela propriedade.

Exemplo. Conforme já discutimos atrás, a vantagem de usar coordenadas é tornar todos os espaços vectoriais semelhantes a $\mathbb{R}^n$. Assim, nestes exemplos vamos concentrar-nos nestes espaços, sabendo que o raciocínio a usar noutros espaços lineares é em tudo semelhante.

1. Consideremos as bases $B = \{(1,0),(1,1)\}$ e $C = \{(1,1),(1,-1)\}$ de $\mathbb{R}^2$. Escrevendo os vectores de B como combinação linear dos elementos de C obtemos

   $$(1,0) = \frac{1}{2}(1,1) + \frac{1}{2}(1,-1) \qquad (1,1) = 1(1,1) + 0(1,-1)\,,$$

   donde as coordenadas destes vectores na base C são

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

   Assim, a matriz P_{CB} é

   $$P_{CB} = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix}.$$

2. Considerando ainda as mesmas duas bases de $\mathbb{R}^2$, podemos escrever agora os vectores de C em coordenadas da base B, obtendo

   $$((1,1))_B = (0,1) \qquad ((1,-1))_B = (2,-1)\,,$$

   donde a matriz P_{BC} é a matriz

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

   Observe-se que $P_{BC}P_{CB} = P_{CB}P_{BC} = \mathbf{I_2}$.

3. Sejam $\overrightarrow{u}$ e $\overrightarrow{v}$ os vectores de $\mathbb{R}^2$ cujas coordenadas na base B são

$$(\overrightarrow{u})_B = (-4, 2) \text{ e } (\overrightarrow{v})_B = (1, 5)\,.$$

As coordenadas de $\overrightarrow{u}$ e $\overrightarrow{v}$ na base C podem agora ser calculadas facilmente.

$$(\overrightarrow{u})_C = P_{CB}\,(\overrightarrow{u})_B = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = (0, -2)$$
$$(\overrightarrow{v})_C = P_{CB}\,(\overrightarrow{v})_B = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \left(\frac{11}{2}, \frac{1}{2}\right)$$

Podemos verificar este resultado calculando explicitamente $\overrightarrow{u}$ e $\overrightarrow{v}$. De acordo com as coordenadas na base B, temos

$$\overrightarrow{u} = -4(1,0) + 2(1,1) = (-2,2) \qquad \overrightarrow{v} = 1(1,0) + 5(1,1) = (6,5)\,.$$

Usando as coordenadas dos mesmos vectores na base C, obtemos

$$\overrightarrow{u} = 0(1,1) - 2(1,-1) = (-2,2) \qquad \overrightarrow{v} = \frac{11}{2}(1,1) + \frac{1}{2}(1,-1) = (6,5)\,.$$

4. Reciprocamente, sejam agora $\overrightarrow{w_1}$ e $\overrightarrow{w_2}$ os vectores de $\mathbb{R}^2$ cujas coordenadas na base C são

$$(\overrightarrow{w_1})_C = (-4, 2) \text{ e } (\overrightarrow{w_2})_C = (1, 5)\,.$$

As coordenadas de $\overrightarrow{w_1}$ e $\overrightarrow{w_2}$ na base B obtêm-se de forma análoga.

$$(\overrightarrow{w_1})_B = P_{BC}\,(\overrightarrow{w_1})_C = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = (4, -6)$$
$$(\overrightarrow{w_2})_B = P_{BC}\,(\overrightarrow{w_2})_C = \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = (10, -4)$$

Podemos verificar novamente estes resultados calculando explicitamente $\overrightarrow{w_1}$ e $\overrightarrow{w_2}$. De acordo com as coordenadas na base B, temos

$$\overrightarrow{w_1} = 4(1,0) - 6(1,1) = (-2,-6) \qquad \overrightarrow{w_2} = 10(1,0) - 4(1,1) = (6,-4)\,.$$

Usando as coordenadas dos mesmos vectores na base C, obtemos

$$\overrightarrow{w_1} = -4(1,1) + 2(1,-1) = (-2,-6) \qquad \overrightarrow{w_2} = 1(1,1) + 5(1,-1) = (6,-4)\,,$$

novamente de acordo com o esperado.

5. Consideremos as bases $B = \{(1,0,0), (1,1,0), (1,1,1)\}$ e $C = \{(0,0,1), (0,1,0), (1,0,0)\}$ de $\mathbb{R}^3$. É simples obter as coordenadas dos vectores de B em termos da base C.

$$((1,0,0))_C = (0,0,1) \qquad ((1,1,0))_C = (0,1,1) \qquad ((1,1,1))_C = (1,1,1)$$

Então a matriz de mudança de base P_{CB} é a matriz

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

Se escrevermos os vectores de C em coordenadas da base B, obtemos

$$((0,0,1))_B = (0,-1,1) \qquad ((0,1,0))_B = (-1,1,0) \qquad ((1,0,0))_B = (1,0,0)\,,$$

donde a matriz de mudança de base P_{BC} é a matriz

$$P_{BC} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

6. Sejam $\overrightarrow{u} = (2,1,3)$, $\overrightarrow{v} = (-1,1,9)$ e $\overrightarrow{w} = (2,-3,1)$. Para obter as coordenadas destes vectores na base B, podemos começar por os escrever em coordenadas da base C e usar a matriz de mudança de base para obter o resultado.

$$(\overrightarrow{u})_B = P_{BC}\,(\overrightarrow{u})_C = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = (1,-2,3)$$

$$(\overrightarrow{v})_B = P_{BC}\,(\overrightarrow{v})_C = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 9 \end{bmatrix} = (8,2,-1)$$

$$(\overrightarrow{w})_B = P_{BC}\,(\overrightarrow{w})_C = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = (4,-5,2)$$

Exercício 27. Escreva a matriz de mudança de base de cada uma das seguintes bases para a base canónica de $\mathbb{R}^3$.

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

(b) $\{(2,0,0),(0,1,0),(0,0,-1)\}$

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

(d) $\{(1,-1,1),(2,1,0),(0,0,1)\}$

(e) $\{(1,2,3),(4,5,6),(1,2,1)\}$

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

Use estes resultados para converter o vector de coordenadas $(1,2,3)$ em cada uma daquelas bases para coordenadas da base canónica.

Exercício 28. Escreva a matriz de mudança de base de cada uma das seguintes bases para a base canónica de P_2.

(a) $\{1,x,x^2\}$

(b) $\{x^2,x,1\}$

(c) $\{1,1+x,1+x+x^2\}$

(d) $\{1,1+x,1+x^2\}$

(e) $\{x+1,x-1,x^2\}$

(f) $\{x^2-1,x^2+1,2x\}$

(g) $\{1+x,x+x^2,x^2+1\}$

(h) $\{x^2-x-1,2x^2+x,x+2\}$

(i) $\{x^2,2x,3\}$

Os exemplos anteriores sugerem algumas propriedades das matrizes de mudança de base. De facto, estas têm muitas propriedades interessantes e relativamente simples de demonstrar.

Proposição 31. Sejam V um espaço vectorial de dimensão finita e B, C e D bases para V.

1. A matriz de mudança de base P_{CB} é invertível e ${P_{CB}}^{-1} = P_{BC}$.
2. A matriz de mudança de base P_{DB} satisfaz a propriedade $P_{DB} = P_{DC}P_{CB}$.

Demonstração. Ambos os factos recorrem a propriedades muito simples das matrizes.

1. Sendo u um vector arbitrário de V, sabemos que $(u)_C = P_{CB}(u)_B$ e $(u)_B = P_{BC}(u)_C$. Daqui segue que
$$(u)_C = P_{CB}P_{BC}(u)_C \text{ e } (u)_B = P_{BC}P_{CB}(u)_B$$
para todos os vectores u. A única matriz que tem esta propriedade é a matriz identidade $\mathbf{I_n}$, donde P_{CB} e P_{BC} são inversas.

2. Sendo u um vector arbitrário de V, sabemos que $(u)_D = P_{DC}(u)_C$ e $(u)_C = P_{CB}(u)_B$. Juntando estas duas relações obtém-se
$$(u)_D = P_{DC}P_{CB}(u)_B\,,$$
donde $P_{DC}P_{CB}$ é uma matriz de mudança de base de B para D. Mas as matrizes de mudança de base são únicas, donde $P_{DC}P_{CB} = P_{DB}$.

□

Vejamos alguns exemplos da aplicação destes resultados.

Exemplo.

1. Consideremos os seguintes vectores de $\mathbb{R}^2$.
$$\overrightarrow{u} = (3,-2) \qquad \overrightarrow{v} = (5,1) \qquad \overrightarrow{w} = (1,-2)$$
Vamos escrever estes vectores na base $B = \{(1,1),(2,1)\}$. Considerando a base canónica $C = \{(1,0),(0,1)\}$ de $\mathbb{R}^2$, temos
$$((3,-2))_C = (3,-2) \qquad ((5,1))_C = (5,1) \qquad ((1,-2))_C = (1,-2)\,.$$
Precisamos então de construir a matriz mudança de base de C para B, P_{BC}. Ora as colunas desta matriz são as coordenadas de $(1,0)$ e $(0,1)$ na base B; porém, em vez de ir calcular estas coordenadas directamente, podemos usar a proposição anterior e escrever primeiro a matriz P_{CB}, já que é simples escrever as coordenadas dos vectores da base B na base canónica:
$$((1,1))_C = (1,1) \qquad ((2,1))_C = (2,1)\,.$$
Então
$$P_{CB} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$
e como $P_{BC} = (P_{CB})^{-1}$, obtemos (usando por exemplo a regra dos cofactores)
$$P_{BC} = \frac{1}{-1}\begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

Assim, temos que

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

2. Recordemos novamente o exemplo da página 111. Para calcular as coordenadas dos três polinómios p, q e r em relação à base $D = \{x^2-1, 2x^2+2x, x+4\}$, tivemos de resolver três sistemas de equações com a mesma matriz de coeficientes

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

Olhando para as colunas desta matriz, observamos que correspondem precisamente às coordenadas dos vectores da base D em termos da base canónica; ou seja, a matriz A é precisamente a matriz P_{BD}, sendo B a base canónica de P_2. O método de resolução usado nesse sistema correspondeu a calcular a matriz inversa de A, que sabemos já ser ${P_{BD}}^{-1} = P_{DB}$, e a multiplicá-la pelas coordenadas de p, q e r na base canónica — ou seja, a usar precisamente as relações $(p)_D = P_{DB}(p)_B$ e análogas para q e r.

Exercício 29. Use as matrizes calculadas no Exercício 28 para escrever o polinómio de coordenadas $(1,2,1)$ na base canónica de P_2 em cada uma das bases apresentadas nesse exercício.

2.5 Espaços associados a uma matriz

Quando discutimos o produto de matrizes (Secção 1.3), vimos que uma matriz de dimensão $n \times m$ pode ser vista como uma lista de n vectores com m coordenadas (vendo cada linha como um vector) ou uma lista de m vectores com n coordenadas (vendo cada coluna como um vector). Nalgumas aplicações ao longo deste capítulo, nomeadamente quando discutimos matrizes de mudança de base, encontrámos situações em que construímos matrizes desta forma — colocando coordenadas de vectores em linhas ou colunas.

Esta forma de olhar para uma matriz vai ser extremamente útil para resolver questões relacionadas com a determinação de dimensões e bases de subespaços de um espaço linear V que são gerados a partir dum conjunto de vectores.

2.5.1 Definições

Definição. Seja A uma matriz $n \times m$ genérica.

- O *espaço das linhas* de A é o subespaço $\mathrm{Lin}(A)$ de $\mathbb{R}^m$ gerado pelas linhas da matriz A:

$$\mathrm{Lin}(A) = L(\{A_{1*}, A_{2*}, \ldots, A_{n*}\})$$

onde $A_{i*} = (a_{i1}, a_{i2}, \ldots, a_{im})$ é a linha i da matriz A.

- O *espaço das colunas* de A é o subespaço $\mathrm{Col}(A)$ de $\mathbb{R}^n$ gerado pelas colunas da matriz A:

$$\mathrm{Col}(A) = L\left(\{A_{*1}, A_{*2}, \ldots, A_{*m}\}\right)$$

  onde $A_{*i} = (a_{1i}, a_{2i}, \ldots, a_{ni})$ é a coluna i da matriz A.

- O *núcleo* de A é o conjunto das soluções do sistema homogéneo associado à matriz A, denotado por $\mathrm{Ker}(A)$ ou $\mathcal{N}(A)$:

$$\mathrm{Ker}(A) = \mathcal{N}(A) = \left\{\underline{x} \in \mathbb{R}^m \mid A\underline{x} = \underline{0}\right\}.$$

É imediato verificar que $\mathrm{Lin}(A)$ e $\mathrm{Col}(A)$ são espaços lineares: ambos estão definidos como a expansão linear de um conjunto de vectores de $\mathbb{R}^m$ e $\mathbb{R}^n$, respectivamente.

Também é simples verificar que o núcleo de A é um subespaço linear de $\mathbb{R}^m$: basta verificar que qualquer combinação linear de vectores em $\mathcal{N}(A)$ é ainda um vector de $\mathcal{N}(A)$. Ora, sendo u e v vectores tais que $Au = 0$ e $Av = 0$, tem-se pelas propriedades do cálculo matricial que

$$A(\alpha u + \beta v) = \alpha(Au) + \beta(Bv) = 0\,,$$

donde $\alpha u + \beta v \in \mathcal{N}(A)$.

Para verificar se um vector pertence ao núcleo duma matriz A, basta aplicar directamente a definição.

Exemplo.

1. Seja A a matriz seguinte.

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

   Tomando $u = (2,2,0)$, $v = (1,3,1)$ e $w = (2,0,2)$, temos que

$$Au = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

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

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

   donde $u, v \notin \mathcal{N}(A)$ mas $w \in \mathcal{N}(A)$.

2. Seja B a matriz seguinte.

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

Tomando $u = (1, 1, 2)$, $v = (0, 1, 0)$ e $w = (1, 1, -1)$, temos que

$$Bu = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 1 \\ 7 \end{bmatrix}$$

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

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

donde $u, v, w \notin \mathcal{N}(B)$.

3. Para qualquer matriz $M_{n\times m}$, tem-se trivialmente $M\mathbf{0_{m\times 1}} = \mathbf{0_{n\times 1}}$, donde o vector nulo pertence ao núcleo de qualquer matriz.

Exercício 30. Considere a seguinte matriz A.

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

Verifique se os seguintes vectores pertencem ao núcleo de A.

(a) $(1, 1, 0)$ (b) $(-2, 1, -1)$ (c) $(1, -1, 3)$ (d) $(1, 2, 4)$

O raciocínio sobre $\text{Lin}(A)$ e $\text{Col}(A)$ é semelhante ao raciocínio sobre qualquer espaço gerado por um conjunto de vectores.

Exemplo.

1. Seja A a seguinte matriz.

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

Então

$$\text{Lin}(A) = L(\{(2, 1, -3), (1, 0, 1)\}) \qquad \text{Col}(A) = L(\{(2, 1), (1, 0), (-3, 1)\}).$$

Destas relações concluímos, por exemplo, que

$$\begin{array}{rr} (2, 1, -3) \in \text{Lin}(A) & 2(2, 1, 3) + 5(1, 0, 1) = (9, 2, 11) \in \text{Lin}(A) \\ (2, 1) + 2(1, 0) = (4, 1) \in \text{Col}(A) & 3(2, 1) - 2(1, 0) + 5(-3, 1) = (-11, 8) \in \text{Col}(A). \end{array}$$

Para determinar se um dado vector pertence a $\text{Lin}(A)$ ou a $\text{Col}(A)$, temos novamente de resolver um problema de combinações lineares.

(a) O vector $(2, 2, 4)$ pertence a $\mathrm{Lin}(A)$: a condição

$$c_1(2, 1, 3) + c_2(1, 0, 1) = (2, 2, 4)$$

gera o sistema de equações lineares

$$\begin{cases} 2c_1 + c_2 = 2 \\ c_1 = 2 \\ 3c_1 + c_2 = 4 \end{cases} \longrightarrow \begin{cases} c_2 = -2 \\ c_1 = 2 \\ 4 = 4 \end{cases}$$

que é possível e determinado.

(b) O vector $(1, 3, -2)$ não pertence a $\mathrm{Lin}(A)$: a condição

$$c_1(2, 1, 3) + c_2(1, 0, 1) = (1, 3, -2)$$

gera o sistema de equações lineares

$$\begin{cases} 2c_1 + c_2 = 1 \\ c_1 = 3 \\ 3c_1 + c_2 = -2 \end{cases} \longrightarrow \begin{cases} c_2 = -5 \\ c_1 = 3 \\ 4 = -2 \end{cases}$$

que é impossível.

(c) O vector $(2, 2)$ pertence a $\mathrm{Col}(A)$: a condição

$$c_1(2, 1) + c_2(1, 0) + c_3(-3, 1) = (2, 2)$$

gera o sistema de equações lineares

$$\begin{cases} 2c_1 + c_2 - 3c_3 = 2 \\ c_1 + c_3 = 2 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -3 & 2 \\ 1 & 0 & 1 & 2 \end{array}\right] \; 2l_2 - l_1 \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -3 & 2 \\ 0 & -1 & 5 & 2 \end{array}\right]$$

que é possível indeterminado.

(d) O vector $(5, -1)$ também pertence a $\mathrm{Col}(A)$: a condição

$$c_1(2, 1) + c_2(1, 0) + c_3(-3, 1) = (5, -1)$$

gera o sistema de equações lineares

$$\begin{cases} 2c_1 + c_2 - 3c_3 = 5 \\ c_1 + c_3 = -1 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -3 & 5 \\ 1 & 0 & 1 & -1 \end{array}\right] \; 2l_2 - l_1 \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -3 & 5 \\ 0 & -1 & 5 & -7 \end{array}\right]$$

que é possível indeterminado.

(e) Na realidade, podemos confirmar que $\mathrm{Col}(A) = \mathbb{R}^2$: sendo (x, y) um vector genérico de $\mathbb{R}^2$, a condição

$$c_1(2, 1) + c_2(1, 0) + c_3(-3, 1) = (x, y)$$

gera o sistema de equações lineares

$$\begin{cases} 2c_1 + c_2 - 3c_3 = x \\ c_1 + c_3 = y \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -3 & x \\ 1 & 0 & 1 & y \end{array}\right] \; 2l_2 - l_1 \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -3 & x \\ 0 & -1 & 5 & 2y - x \end{array}\right]$$

que é sempre possível indeterminado.

2. Considere-se agora a seguinte matriz B.

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

Os espaços das linhas e das colunas de B são os seguintes.

$$\begin{aligned} \mathrm{Lin}(B) &= L(\{(1,2,-1),(2,4,-2),(1,-1,0),(4,-1,-1)\}) \\ \mathrm{Col}(B) &= L(\{(1,2,1,4),(2,4,-1,-1),(-1,-2,0,-1)\}) \end{aligned}$$

Para determinar se um dado vector pertence a estes espaços, temos novamente de resolver um problema de combinações lineares.

(a) O vector $(2,1,-1)$ pertence a $\mathrm{Lin}(B)$: a condição

$$c_1(1,2,-1)+c_2(2,4,-2)+c_3(1,-1,0)+c_4(4,-1,-1)=(2,1,-1)$$

gera o seguinte sistema de equações lineares.

$$\begin{cases} c_1+2c_2+c_3+4c_4=2 \\ 2c_1+4c_2-c_3-c_4=1 \\ -c_1-2c_2-c_4=-1 \end{cases} \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 2 \\ 2 & 4 & -1 & -1 & 1 \\ -1 & -2 & 0 & -1 & -1 \end{array}\right] \begin{array}{l} \\ l_2-2l_1 \\ l_3+l_1 \end{array}$$

$$\longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 2 \\ 0 & 0 & -3 & -9 & -3 \\ 0 & 0 & 1 & 3 & 1 \end{array}\right] \begin{array}{l} \\ \\ 3l_3+l_2 \end{array} \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 2 \\ 0 & 0 & -3 & -9 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\longrightarrow \begin{cases} c_1+2c_2+c_3+4c_4=2 \\ -3c_3-9c_4=-3 \end{cases}$$

Este sistema é possível indeterminado, logo $(2,1,-1)\in \mathrm{Lin}(B)$.

(b) O vector $(5,1,7)$ não pertence a $\mathrm{Lin}(B)$: a condição

$$c_1(1,2,-1)+c_2(2,4,-2)+c_3(1,-1,0)+c_4(4,-1,-1)=(5,1,7)$$

gera o seguinte sistema de equações lineares.

$$\begin{cases} c_1+2c_2+c_3+4c_4=5 \\ 2c_1+4c_2-c_3-c_4=1 \\ -c_1-2c_2-c_4=7 \end{cases} \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 5 \\ 2 & 4 & -1 & -1 & 1 \\ -1 & -2 & 0 & -1 & 7 \end{array}\right] \begin{array}{l} \\ l_2-2l_1 \\ l_1+l_3 \end{array}$$

$$\longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 5 \\ 0 & 0 & -3 & -9 & -9 \\ 0 & 0 & 1 & 3 & 12 \end{array}\right] \begin{array}{l} \\ \\ 3l_3+l_2 \end{array} \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 5 \\ 0 & 0 & -3 & -9 & -9 \\ 0 & 0 & 0 & 0 & 27 \end{array}\right]$$

$$\longrightarrow \begin{cases} c_1+2c_2+c_3+4c_4=5 \\ -3c_3-9c_4=-9 \\ 0=27 \end{cases}$$

Este sistema é impossível, logo $(5,1,7)\notin \mathrm{Lin}(B)$.

(c) O vector $(2,4,0,2)$ pertence a $\mathrm{Col}(B)$: a condição

$$c_1(1,2,1,4)+c_2(2,4,-1,-1)+c_3(-1,-2,0,-1)=(2,4,0,2)$$

gera o seguinte sistema de equações lineares.

$$\begin{cases} c_1+2c_2-c_3=2\\ 2c_1+4c_2-2c_3=4\\ c_1-c_2=0\\ 4c_1-c_2-c_3=2 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2\\ 2 & 4 & -2 & 4\\ 1 & -1 & 0 & 0\\ 4 & -1 & -1 & 2 \end{array}\right] \begin{array}{l} \\ l_2-2l_1\\ l_3-l_1\\ l_4-4l_1 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2\\ 0 & 0 & 0 & 0\\ 0 & -3 & 1 & -2\\ 0 & -9 & 3 & -6 \end{array}\right] l_2 \leftrightarrow l_4 \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2\\ 0 & -9 & 3 & -6\\ 0 & -3 & 1 & -2\\ 0 & 0 & 0 & 0 \end{array}\right] 3l_3-l_2$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2\\ 0 & -9 & 3 & -6\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \begin{cases} c_1+2c_2-c_3=2\\ 9c_2-3c_3=6 \end{cases}$$

Este sistema é possível indeterminado, logo $(2,4,0,2)\in\mathrm{Lin}(B)$.

Os exemplos acima revelam um padrão que será essencial para as aplicações que faremos destes espaços: para determinar se $v\in\mathrm{Col}(A)$, resolvemos o sistema $A\underline{x}=v$; para determinar se $w\in\mathrm{Lin}(A)$, resolvemos o sistema $A^T\underline{x}=w$.

De facto, se $v\in\mathrm{Col}(A)$, então v é combinação linear das colunas de A, ou seja, existem coeficientes $c_1,c_2,\ldots,c_m$ tais que

$$v=c_1A_{*1}+c_2A_{*2}+\cdots+c_mA_{*m}\,,$$

e este produto é precisamente o produto

$$A\begin{bmatrix} c_1\\ c_2\\ \vdots\\ c_m \end{bmatrix}.$$

Da mesma forma, se $w\in\mathrm{Lin}(A)$, então v é combinação linear das linhas de A, ou seja, existem coeficientes $d_1,d_2,\ldots,d_n$ tais que

$$v=d_1A_{1*}+d_2A_{2*}+\cdots+d_nA_{n*}\,,$$

e este produto é precisamente o produto

$$A^T\begin{bmatrix} d_1\\ d_2\\ \vdots\\ d_n \end{bmatrix}.$$

Encontrar os coeficientes destas combinações lineares corresponde precisamente a resolver aqueles sistemas de equações lineares em ordem aos coeficientes $c_1,c_2,\ldots,c_m$ ou $d_1,d_2,\ldots,d_n$.

Uma das consequências desta observação é a seguinte caracterização dos sistemas de equações lineares.

Proposição 32. O sistema $A\underline{x} = \underline{b}$ tem solução se e só se $\underline{b} \in \mathrm{Col}(A)$.

Exercício 31. Considere a seguinte matriz A.

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

Verifique se os seguintes vectores pertencem $\mathrm{Lin}(A)$ e/ou a $\mathrm{Col}(A)$.

(a) $(1, 1, 0)$ (b) $(-2, 1, -1)$ (c) $(1, -1, 3)$ (d) $(1, 2, 4)$

2.5.2 Bases para o espaço das linhas e para o núcleo

Vamos agora debruçar-nos sobre o problema de encontrar bases para o espaço das linhas e para o núcleo duma matriz arbitrária A. O método que nos vai permitir realizar esta tarefa é, mais uma vez, o método de eliminação de Gauss. O problema correspondente para o espaço das colunas é um pouco diferente e será abordado posteriormente.

Observe-se que as três operações elementares sobre matrizes (troca de linhas, multiplicação duma linha por uma constante e soma duma linha com um múltiplo de outra) são operações que não alteram o conjunto das soluções do sistema homogéneo associado à matriz; foi aliás esta a justificação para utilizar o método de eliminação de Gauss na resolução de sistemas de equações lineares. Assim, claramente nenhuma destas operações altera o núcleo da matriz a que são aplicadas.

De forma semelhante, é possível verificar que as operações elementares não alteram o espaço das linhas da matriz sobre a qual são aplicadas. Para a troca de linhas isto é imediato; para a multiplicação por uma constante também (estamos a substituir um gerador por um seu múltiplo; qualquer combinação linear das linhas originais pode ser escrita com as novas linhas dividindo o coeficiente da linha alterada pelo multiplicador usado).

Para a substituição de l_i por $l'_i = l_i + \alpha l_j$, basta observar que qualquer combinação linear $c_1 l_1 + c_2 l_2 + \cdots + c_n l_n$ que pudesse ser escrita antes da operação elementar continua a poder ser representada, uma vez que $c_i l_i = c_i l'_i - \alpha l_j$.

Obtemos então o resultado seguinte.

Proposição 33. Sejam A uma matriz $n \times m$ e B a matriz que se obtém de A pelo método de eliminação de Gauss. Então $\mathcal{N}(A) = \mathcal{N}(B)$ e $\mathrm{Lin}(A) = \mathrm{Lin}(B)$.

Este resultado permite obter bases para o espaço das linhas duma matriz de forma muito simples. De facto, é fácil ver que numa matriz em escada de linhas todas as linhas não nulas são linearmente independentes: partindo da primeira linha e descendo na matriz, cada nova linha não pode ser escrita como combinação das anteriores, já que tem todas as entradas correspondentes aos pivots das outras linhas nulas (o que requer que os coeficientes das outras linhas sejam todos 0) mas o seu pivot não é nulo (e portanto a linha não é nula).

Proposição 34. Seja A uma matriz de dimensão $n \times m$ e B a matriz que se obtém de A pelo método de eliminação de Gauss. Então as linhas não nulas de B formam uma base para o espaço das linhas de A.

Este resultado permite-nos determinar simultaneamente a dimensão do espaço das linhas duma matriz: é simplesmente o número de pivots após aplicar o método de eliminação de Gauss.

Vejamos alguns exemplos com matrizes que já usámos atrás.

Exemplo.

1. Seja A a seguinte matriz.
$$A = \begin{bmatrix} 2 & 1 & -3 \\ 1 & 0 & 1 \end{bmatrix}$$
Aplicando o método de eliminação de Gauss a A, obtemos
$$\begin{bmatrix} 2 & 1 & -3 \\ 1 & 0 & 1 \end{bmatrix} \begin{matrix} \\ 2l_2 - l_1 \end{matrix} \longrightarrow \begin{bmatrix} 2 & 1 & -3 \\ 0 & -1 & 5 \end{bmatrix}$$
donde uma base para $\mathrm{Lin}(A)$ é $\{(2,1,-3),(0,-1,5)\}$ e a dimensão deste espaço é 2.

2. Considere-se a matriz
$$B = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 1 & -1 & 0 \\ 4 & -1 & -1 \end{bmatrix}.$$
Aplicando o método de eliminação de Gauss a B, obtemos
$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 1 & -1 & 0 \\ 4 & -1 & -1 \end{bmatrix} \begin{matrix} \\ l_2 - 2l_1 \\ l_3 - l_1 \\ l_4 - 4l_1 \end{matrix} \longrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & -9 & 3 \end{bmatrix} l_2 \leftrightarrow l_4 \longrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & -9 & 3 \\ 0 & -3 & 1 \\ 0 & 0 & 0 \end{bmatrix} 3l_3 - l_2$$
$$\longrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & -9 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
donde $\mathrm{Lin}(B)$ tem base $\{(1,2,-1),(0,-9,3)\}$ e dimensão 2.

3. Considere-se ainda a matriz
$$C = \begin{bmatrix} 3 & 0 & 1 & 3 \\ 1 & 0 & 2 & -1 \\ -3 & 0 & 0 & 1 \end{bmatrix}.$$
Para determinar uma base para $\mathrm{Lin}(C)$, começamos por aplicar o método de eliminação de Gauss a C.
$$\begin{bmatrix} 3 & 0 & 1 & 3 \\ 1 & 0 & 2 & -1 \\ -3 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ 3l_2 - l_1 \\ l_1 + l_3 \end{matrix} \longrightarrow \begin{bmatrix} 3 & 0 & 1 & 3 \\ 0 & 0 & 5 & -6 \\ 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \\ \\ 5l_3 - l_2 \end{matrix} \longrightarrow \begin{bmatrix} 3 & 0 & 1 & 3 \\ 0 & 0 & 5 & -6 \\ 0 & 0 & 0 & 26 \end{bmatrix}$$
Daqui concluímos que uma base para $\mathrm{Lin}(C)$ é
$$\{(3,0,1,3),(0,0,-5,6),(0,0,0,26)\}.$$

Exercício 32. Determine bases para o espaço das linhas de cada uma das seguintes matrizes.

(a) $\begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 1 \\ 4 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$

Observe-se ainda que as linhas que vão sendo anuladas ao longo do método de eliminação de Gauss são linhas que podem ser escritas como combinação linear das anteriores (já que são anuladas somando-se-lhes uma combinação linear das linhas anteriores). Então as linhas que não se anulam ao longo do método são linearmente independentes; podemos assim encontrar uma base para o espaço das linhas duma matriz formada por linhas dessa matriz.

Exemplo.

1. Ao reduzir a matriz A acima, concluímos que o espaço das suas linhas tinha dimensão 2. Como A tem duas linhas, estas serão forçosamente linearmente independentes. Então outra base para $\operatorname{Lin}(A)$ é $\{(2,1,-3),(1,0,1)\}$.

2. No caso da matriz B, a segunda linha anulou-se no primeiro passo da eliminação de Gauss; após a troca, a terceira linha também se anulou, sobrando apenas duas linhas: a primeira e a que originalmente era a quarta. Então a primeira e a quarta linhas da matriz B formam um base para o espaço das suas linhas, nomeadamente $\{(1,2,-1),(4,-1,-1)\}$.

3. Finalmente, nenhuma das linhas da matriz C se anulou durante a aplicação do método de eliminação de Gauss, pelo que outra base possível para $\operatorname{Lin}(C)$ seria

$$\{(3,0,1,3),(1,0,2,-1),(-3,0,0,1)\}.$$

Exercício 33. Para cada uma das matrizes do exercício anterior, indique uma base para o seu espaço das linhas constituída unicamente por vectores nas linhas dessa matriz.

Esta estratégia permite-nos ainda determinar bases para subespaços de espaços vectoriais. O caso de $\mathbb{R}^n$ é bastante simples: se S é um subconjunto finito de $\mathbb{R}^n$, para determinar uma base para $L(S)$ basta escrever uma matriz que tenha $L(S)$ como espaço das linhas — um exemplo simples é a matriz A_S cujas linhas são os elementos de S. O método acima permite determinar uma base para $\operatorname{Lin}(A_S)$ — que é precisamente uma base para $L(S)$.

Para subespaços doutros espaços vectoriais, começamos por escrever S em coordenadas duma base do espaço (tipicamente a sua base canónica); a partir daí podemos raciocinar como se estivéssemos em $\mathbb{R}^n$.

Exemplo.

1. Para determinar uma base para o subespaço $W = L(\{(2,1,1),(1,-1,-1),(0,3,3)\})$ de $\mathbb{R}^3$, escrevemos os geradores de W nas linhas duma matriz e aplicamos-lhe eliminação de Gauss.
$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 3 & 3 \end{bmatrix} \begin{matrix} \\ 2l_2 - l_1 \\ \\ \end{matrix} \longrightarrow \begin{bmatrix} 2 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & 3 & 3 \end{bmatrix} \begin{matrix} \\ \\ l_3 + l_2 \end{matrix} \longrightarrow \begin{bmatrix} 2 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$
Esta matriz tem duas linhas não nulas, fornecendo a base $\{(2,1,1),(0,-3,-3)\}$ para W. Conclui-se ainda que as duas primeiras linhas da matriz eram linearmente independentes, pelo que uma outra base para W, constituída por geradores, é $\{(2,1,1),(1,-1,-1)\}$.

2. Para determinar uma base do subespaço W de P_3 gerado por $(x-1)^2$, $x^3 - 2x + 1$, $2x^2(x+1)$ e $2x - 1$, começamos por escrever estes polinómios em coordenadas da base canónica B de P_3.
$$\begin{aligned} \left((x-1)^2\right)_B = \left(x^2 - 2x + 1\right)_B &= (1,-2,1,0) \\ \left(x^3 - 2x + 1\right)_B &= (1,-2,0,1) \\ \left(2x^2(x+1)\right)_B = \left(2x^3 + 2x^2\right)_B &= (0,0,2,2) \\ (2x-1)_B &= (-1,2,0,0) \end{aligned}$$
Construindo uma matriz com estas coordenadas nas linhas e reduzindo-a por eliminação de Gauss, obtemos
$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 1 & -2 & 0 & 1 \\ 0 & 0 & 2 & 2 \\ -1 & 2 & 0 & 0 \end{bmatrix} \begin{matrix} \\ l_2 - l_1 \\ \\ l_4 + l_1 \end{matrix} \longrightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} \\ \\ l_3 + 2l_2 \\ l_4 + l_2 \end{matrix}$$
$$\longrightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ 4l_4 - l_3 \end{matrix} \longrightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
que tem três linhas não nulas. Lendo-as em coordenadas da base canónica de P_3, encontramos a base
$$\left\{1 - 2x + x^2, -x^2 + x^3, 4x^3\right\}$$
para W. Uma vez que não houve trocas de linhas, concluímos também que os primeiros três geradores são linearmente independentes, pelo que outra base para W é
$$\left\{(x-1)^2, x^3 - 2x + 1, 2x^2(x+1)\right\} .$$

3. Consideremos o subespaço de $M_{2\times 2}$ definido como
$$W = L\left(\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 2 & 5 \end{bmatrix}\right\}\right) .$$
Para determinar uma base para W, começamos por escrever os geradores de W em coordenadas da base canónica B de $M_{2\times 2}$, obtendo os seguintes vectores.
$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)_B = (1,0,0,1) \quad \left(\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}\right)_B = (1,0,-1,0) \quad \left(\begin{bmatrix} 3 & 0 \\ 2 & 5 \end{bmatrix}\right)_B = (3,0,2,5)$$

Escrevendo estes vectores nas linhas duma matriz e aplicando eliminação de Gauss, obtemos

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 3 & 0 & 2 & 5 \end{bmatrix} \begin{array}{l} \\ l_1 - l_2 \\ l_3 - 3l_1 \end{array} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix} \begin{array}{l} \\ \\ l_3 - 2l_2 \end{array} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

que tem duas linhas não nulas. Lendo-as em coordenadas da base canónica de $M_{2\times 2}$, encontramos a base

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right\}$$

de W. Uma vez que não houve trocas de linhas, os dois primeiros geradores são linearmente independentes, e outra base para W é

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \right\}.$$

Exercício 34. Para cada um dos seguintes subespaços de $\mathbb{R}^4$, determine uma sua base constituída por geradores e escreva os restantes geradores como combinação linear dos vectores da base escolhida.

(a) $L(\{(1,1,-4,-3),(2,0,2,-2),(2,-1,3,2)\})$

(b) $L(\{(-1,1,-2,0),(3,3,6,0),(9,0,0,3)\})$

(c) $L(\{(1,0,1,1),(-3,3,7,1),(-1,3,9,3),(-5,3,5,-1)\})$

(d) $L(\{(1,1,0,0),(0,0,1,1),(-2,0,2,2),(0,-3,0,3)\})$

(e) $L(\{(1,-2,0,3),(2,-4,0,6),(-1,1,2,0),(0,-1,2,3)\})$

(f) $L(\{(1,0,0,0),(1,1,0,0),(1,1,1,0),(1,1,1,1)\})$

Para encontrar bases para o núcleo duma matriz, a técnica é um pouco diferente. Tal como no caso anterior, começamos por aplicar eliminação de Gauss à matriz, que vimos já preservar o seu núcleo. Na matriz em escada de linhas, as colunas com pivots correspondem às coordenadas cujo valor é completamente determinado pelas restantes. Então, uma forma de encontrar um conjunto linearmente independente de geradores é escolher as diferentes formas de atribuir 1 a uma das coordenadas independentes e 0 às restantes (calculando as coordenadas dependentes em função destas).

Vejamos um exemplo simples. Consideremos novamente a matriz

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

que sabemos já reduzir-se por eliminação de Gauss a

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

Esta matriz tem pivots nas duas primeiras colunas. Isto significa que a única coordenada independente é a terceira — fixada esta, os valores das duas primeiras estão completamente determinados. De facto, um vector (x, y, z) no núcleo de A com $z = 1$ tem necessariamente de satisfazer $y = 5$ e $x = -1$. Por outras palavras, uma base para $\mathcal{N}(A)$ é $\{(-1, 5, 1)\}$, e este espaço tem dimensão 1.

Exemplo.

1. Consideremos novamente a matriz

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

que se reduz por eliminação de Gauss a

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

Esta matriz tem pivots nas duas primeiras colunas. Dando à variável independente o valor 1, encontramos o vector $\left(\frac{1}{3}, \frac{1}{3}, 1\right)$. Uma base possível para $\mathcal{N}(B)$ é então $\left\{\left(\frac{1}{3}, \frac{1}{3}, 1\right)\right\}$, ou, se multiplicarmos este vector por 3 para o desembaraçar de denominadores, $\{(1, 1, 3)\}$.

2. No caso da matriz

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

esta reduz-se por eliminação de Gauss a

$$\begin{bmatrix} 3 & 0 & 1 & 3 \\ 0 & 0 & 5 & -6 \\ 0 & 0 & 0 & 26 \end{bmatrix}.$$

Esta matriz tem pivots na primeira, terceira e quarta colunas, pelo que a única coordenada independente é a segunda. Dando a esta o valor 1, verificamos que todas as outras têm de valer 0; então o núcleo de C tem por base $\{(0, 1, 0, 0)\}$.

3. Vejamos agora um exemplo em que a dimensão do núcleo é superior a 1. Seja D a matriz

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

Aplicando eliminação de Gauss a D, obtemos

$$\begin{bmatrix} 2 & 1 & 3 & -1 \\ -4 & -2 & -1 & 3 \end{bmatrix} \begin{matrix} \\ l_2 + 2l_1 \end{matrix} \longrightarrow \begin{bmatrix} 2 & 1 & 3 & -1 \\ 0 & 0 & 5 & 1 \end{bmatrix}.$$

Esta matriz em escada tem pivots na primeira e na terceira colunas. Então podemos escolher os valores da segunda e quarta coordenadas. Para encontrar uma base, basta escolher combinações linearmente independentes destas duas coordenadas; seguindo a técnica que descrevemos, vamos escolher vectores com uma delas igual a 1 e a outra igual a 0. Fazendo a segunda igual a 1, obtemos $\left(-\frac{1}{2}, 1, 0, 0\right)$; fazendo a quarta igual a 1, obtemos $\left(\frac{4}{5}, 0, -\frac{1}{5}, 1\right)$. Escolhendo múltiplos inteiros destes vectores, encontramos a base $\{(-1, 2, 0, 0), (4, 0, -1, 5)\}$ para $\mathcal{N}(D)$.

Exercício 35. Determine bases para o núcleo de cada uma das matrizes do Exercício 32.

2.5.3 Bases para o espaço das colunas

Antes de prosseguir, é importante salientar que o método de eliminação de Gauss *não* preserva o espaço das colunas duma matriz. Por exemplo, seja

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

O espaço das colunas de A é $\mathrm{Col}(A) = L(\{(1, 2), (3, 6)\})$; uma vez que o segundo destes vectores é múltiplo do primeiro, podemos escrever simplesmente $\mathrm{Col}(A) = L(\{(1, 2)\})$.

Aplicando o método de eliminação de Gauss a A obtemos

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

mas $\mathrm{Col}(B) = L(\{(1, 0), (3, 0)\}) \neq \mathrm{Col}(A)$.

Contudo, os espaços das colunas da matriz original e da matriz em escada de linhas estão relacionados, conforme o seguinte resultado indica.

Proposição 35. Seja A uma matriz de dimensão $n \times m$ e B a matriz que se obtém de A pelo método de eliminação de Gauss.

Então um subconjunto das colunas de A é linearmente independente de $\mathrm{Col}(A)$ se e só se o correspondente subconjunto das colunas de B for linearmente independente para $\mathrm{Col}(B)$.

Demonstração. Sejam $\{c_1, c_2, \ldots, c_k\}$ uma base para $\mathrm{Col}(A)$ formada por colunas de A e $\{c'_1, c'_2, \ldots, c'_k\}$ o correspondente subconjunto de colunas de B e considere-se a matriz A^0 que contém apenas as colunas $c_1, c_2, \ldots, c_k$. Aplicando a mesma sequência de operações elementares que usámos para reduzir A a B à matriz A^0, vamos obter uma matriz B^0 contendo precisamente as colunas $c'_1, c'_2, \ldots, c'_k$.

Ora as operações elementares preservam as soluções do sistema homogéneo associado a uma matriz. Então a solução geral de $A^0\underline{x} = \underline{0}$ é a solução geral de $B^0\underline{x} = \underline{0}$. Uma vez que as soluções destes sistemas são coeficientes de combinações lineares que produzem o vector nulo a partir, respectivamente, das colunas de A^0 e das colunas de B^0, conclui-se que as colunas de A^0 são linearmente independentes se e só se as colunas de B^0 o forem. $\square$

O argumento que usámos para justificar que as linhas que contêm os pivots numa matriz em escada de linhas são linearmente independentes pode ser adaptado, mostrando que as colunas que contêm os pivots também são linearmente independentes. Encontrar uma base para o espaço das colunas duma matriz em escada de linhas é então simples — basta seleccionar as colunas que contêm os pivots. Uma vez que o método de eliminação de Gauss nunca troca colunas, as colunas correspondentes da matriz original são linearmente independentes e formam uma base para o espaço das colunas dessa matriz.

Proposição 36. Seja A uma matriz $n \times m$ e B a matriz que se obtém por aplicar o método de eliminação de Gauss a A.

Então o conjunto formado pelas colunas de A correspondentes às colunas de B que contêm os pivots é uma base para $\mathrm{Col}(A)$.

Vamos aplicar este resultado para determinar bases para o espaço das colunas das matrizes que usámos como exemplo na secção anterior.

Exemplo.

1. Vimos atrás que a matriz
$$A = \begin{bmatrix} 2 & 1 & -3 \\ 1 & 0 & 1 \end{bmatrix}$$
se reduzia por eliminação de Gauss a
$$\begin{bmatrix} 2 & 1 & -3 \\ 0 & 1 & -5 \end{bmatrix}.$$
Esta matriz tem pivots nas duas primeiras colunas. Então as duas primeiras colunas de A formam uma base para $\mathrm{Col}(A)$, ou seja, este espaço admite a base $\{(2,1),(1,0)\}$.

   Observe-se que como $\mathrm{Col}(A)$ é um subespaço de $\mathbb{R}^2$ com dimensão 2 necessariamente coincide com o próprio $\mathbb{R}^2$, pelo que a base canónica deste espaço também é uma base para $\mathrm{Col}(A)$.

2. No caso da matriz
$$B = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 1 & -1 & 0 \\ 4 & -1 & -1 \end{bmatrix},$$
esta reduz-se por eliminação de Gauss a
$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 9 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$
Esta matriz tem novamente pivots nas duas primeiras colunas. Então $\mathrm{Col}(B)$ admite uma base formada pelas duas primeiras colunas de B, que é $\{(1,2,1),(2,4,-1)\}$.

3. A matriz
$$C = \begin{bmatrix} 3 & 0 & 1 & 3 \\ 1 & 0 & 2 & -1 \\ -3 & 0 & 0 & 1 \end{bmatrix}$$
reduz-se por eliminação de Gauss a
$$\begin{bmatrix} 3 & 0 & 1 & 3 \\ 0 & 0 & 5 & -6 \\ 0 & 0 & 0 & 26 \end{bmatrix},$$

que tem pivots na primeira, terceira e quarta colunas. Seleccionando as colunas correspondentes de C, encontramos a base $\{(3, 1, -3), (1, 2, 0), (3, -1, 1)\}$ para $\mathrm{Col}(C)$. Também aqui este espaço coincide com o próprio $\mathbb{R}^3$.

4. Finalmente, a matriz

$$D = \begin{bmatrix} 2 & 1 & 3 & -1 \\ -4 & -2 & -1 & 3 \end{bmatrix}$$

reduz-se por eliminação de Gauss a

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

que tem pivots na primeira e na terceira colunas. Então uma base para $\mathrm{Col}(D)$ é $\{(2, -4), (3, -1)\}$, constituída pelas colunas correspondentes de D.

Exercício 36. Determine bases para o espaço das colunas de cada uma das matrizes do Exercício 32.

2.5.4 Teorema da dimensão

Os resultados das últimas secções permitem obter algumas relações entre as dimensões dos vários espaços associados a uma matriz. O primeiro é imediato: os espaços das linhas e das colunas duma matriz têm a mesma dimensão.

Proposição 37. Seja A uma matriz. Então $\dim(\mathrm{Lin}(A)) = \dim(\mathrm{Col}(A)) = \mathrm{car}(A)$.

Demonstração. Podemos obter quer uma base de $\mathrm{Lin}(A)$, quer uma de $\mathrm{Col}(A)$, escolhendo um vector por cada pivot da matriz. $\square$

À dimensão do núcleo duma matriz A é habitual chamar *nulidade* de A, sendo frequente designá-la por $\mathrm{nul}(A)$. A característica e a nulidade duma matriz estão directamente relacionadas.

Teorema 7 (Teorema da dimensão)**.** Seja A uma matriz de dimensão $n \times m$. Então

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

Demonstração. Por definição, a característica de A corresponde ao número dos seus pivots. Por outro lado, cada coluna sem pivots está associada a uma coordenada livre, que gera um vector na base de $\mathcal{N}(A)$, conforme discutimos na Secção 2.5.2. Assim, cada coluna de A contribui ou com um pivot ou com um vector na base de $\mathcal{N}(A)$, donde a soma da característica de A com a sua nulidade é precisamente o número de colunas da matriz. $\square$

Exercício 37. Verifique o Teorema da Dimensão para as matrizes do Exercício 32.

Podemos nesta altura apresentar uma versão mais detalhada da Proposição 32.

Proposição 38. Sejam A uma matriz de dimensão $n \times m$ e $\underline{b} \in \mathbb{R}^n$. Então o sistema de equações lineares $A\underline{x} = \underline{b}$ é:

- impossível se $\underline{b} \notin \mathrm{Col}(A)$;
- possível e determinado se $\underline{b} \in \mathrm{Col}(A)$ e $\mathrm{car}(A) = m$;
- possível e indeterminado se $\underline{b} \in \mathrm{Col}(A)$ e $\mathrm{car}(A) < m$.

Demonstração. Vimos já (Proposição 32) que $A\underline{x} = \underline{b}$ é possível se e só se $\underline{b} \in \mathrm{Col}(A)$. A primeira alínea é um caso particular desta proposição.

Por outro lado, o conjunto de todas as soluções do sistema $A\underline{x} = \underline{b}$ pode ser obtido como a soma duma solução particular com a solução geral do sistema homogéneo correspondente $A\underline{x} = \underline{0}$. Ora o conjunto das soluções deste sistema é precisamente $\mathcal{N}(A)$, sendo em particular um conjunto singular (o que significa que $A\underline{x} = \underline{b}$ é determinado) quando $\mathrm{nul}(A) = 0$; pelo Teorema da Dimensão, nesta situação tem-se $\mathrm{car}(A) = m$ (segunda alínea). No caso em que o sistema é indeterminado, tem-se $\mathrm{nul}(A) > 0$ e portanto necessariamente $\mathrm{car}(A) < m$. □

2.6 Espaços euclideanos

Nesta secção vamos ver uma primeira aplicação prática da Álgebra Linear: o cálculo de aproximações. Para tal, vamos precisar de generalizar a estrutura geométrica de $\mathbb{R}^n$ a outros espaços vectoriais.

2.6.1 Produtos internos

Recorde-se que em $\mathbb{R}^n$ podemos definir um produto interno entre dois vectores — foi aliás generalizando este produto interno que definimos o produto de matrizes.

Definição. O *produto interno euclideano* em $\mathbb{R}^n$ é uma operação $\cdot \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definida por

$$\overrightarrow{u} \cdot \overrightarrow{v} = (u_1, u_2, \ldots, u_n) \cdot (v_1, v_2, \ldots, v_n) = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

para $\overrightarrow{u}$ e $\overrightarrow{v}$ em $\mathbb{R}^n$.

Exemplo. Em $\mathbb{R}^4$, o produto interno entre $(1, 3, 2, -1)$ e $(2, -1, 0, 2)$ é

$$(1, 3, 2, -1) \cdot (2, -1, 0, 2) = 1 \times 2 + 3 \times (-1) + 2 \times 0 + (-1) \times 2 = -3\,.$$

Exercício 38. Calcule o produto interno dos seguintes pares de vectores, para o produto interno usual em $\mathbb{R}^2$ e $\mathbb{R}^3$.

(a) $(1, 0)$ e $(0, 1)$

(b) $(2, 4)$ e $(-1, -2)$

(c) $(2, 1)$ e $(-2, 1)$

(d) $(3, 1, -5)$ e $(-2, 0, 1)$

(e) $(1, 2, 1)$ e $(-1, 2, 1)$

(f) $(0, 2, 1)$ e $(-2, -2, -2)$

Este produto interno satisfaz o seguinte conjunto de propriedades, que são consequência directa das propriedades da soma e multiplicação de números reais.

Proposição 39. As seguintes propriedades são verificadas por quaisquer vectores $\overrightarrow{u}$ e $\overrightarrow{v}$ de $\mathbb{R}^n$ e número real α.

1. $\overrightarrow{u} \cdot \overrightarrow{v} = \overrightarrow{v} \cdot \overrightarrow{u}$
2. $\overrightarrow{u} \cdot \overrightarrow{u} \geq 0$
3. $(\alpha \overrightarrow{u}) \cdot \overrightarrow{v} = \alpha (\overrightarrow{u} \cdot \overrightarrow{v})$
4. $\overrightarrow{u} \cdot (\overrightarrow{v} + \overrightarrow{w}) = \overrightarrow{u} \cdot \overrightarrow{v} + \overrightarrow{u} \cdot \overrightarrow{w}$.

Tal como anteriormente abstraímos as propriedades de $\mathbb{R}^n$ a outros conjuntos para obter um conceito geral de espaço vectorial, vamos partir das propriedades do produto interno euclideano em $\mathbb{R}^n$ para definir um conceito geral de produto interno em qualquer espaço vectorial.

Definição. Sejam V um espaço linear e $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ uma operação. A operação $\langle \cdot, \cdot \rangle$ diz-se um *produto interno* em V se satisfizer as seguintes propriedades para quaisquer $u, v, w \in V$ e $\alpha \in \mathbb{R}$.

1. Simetria: $\langle u, v \rangle = \langle v, u \rangle$
2. Positividade: $\langle u, u \rangle \geq 0$, verificando-se a igualdade apenas para $u = 0_V$
3. Linearidade: $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$
4. Distributividade: $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$

Um espaço linear V com um produto interno diz-se um *espaço euclideano*.

Mesmo em $\mathbb{R}^n$, o produto interno euclideano não é a única operação com estas propriedades. Consideremos por exemplo a operação em $\mathbb{R}^2$ definida por

$$\langle (a, b), (c, d) \rangle = ac + 2bd\,.$$

Esta operação satisfaz as quatro propriedades acima indicadas, pelo que define um outro produto interno em $\mathbb{R}^2$.

Exemplo. No espaço $M_{n \times m}$ das matrizes de dimensão $n \times m$, pode definir-se o produto interno

$$\langle A, B \rangle = A_{1*} \cdot B_{1*} + A_{2*} \cdot B_{2*} + \cdots + A_{n*} \cdot B_{n*}$$

onde A_{i*} é a linha i da matriz A e B_{i*} é a linha i da matriz B. Repare-se que esta operação consiste na realidade em ver as matrizes A e B como vectores de dimensão $n \times m$ (com todas as entradas escritas numa única linha) e calcular o seu produto interno euclideano. Assim, esta operação satisfaz trivialmente as propriedades dum produto interno.

Tomem-se as matrizes

$$A = \begin{bmatrix} 1 & 9 & 7 \\ 2 & 8 & 0 \end{bmatrix} \text{ e } \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 9 \end{bmatrix}$$

em $M_{2 \times 3}$. O seu produto interno é dado por

$$\begin{aligned} \langle A, B \rangle &= \left\langle \begin{bmatrix} 1 & 9 & 7 \\ 2 & 8 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 9 \end{bmatrix} \right\rangle \\ &= (1, 9, 7) \cdot (1, 0, 2) + (2, 8, 0) \cdot (0, 3, 9) \\ &= 1 \times 1 + 9 \times 0 + 7 \times 2 + 2 \times 0 + 8 \times 3 + 0 \times 9 \\ &= 1 + 14 + 24 = 39\,. \end{aligned}$$

Observe-se que se escrevermos os elementos de cada matriz como um vector de $\mathbb{R}^6$ e calcularmos o produto interno neste espaço obtemos o mesmo resultado.

$$\begin{aligned}(1,9,7,2,8,0)\cdot(1,0,2,0,3,9) &= 1\times 1+9\times 0+7\times 2+2\times 0+8\times 3+0\times 9\\ &= 1+14+24=39\end{aligned}$$

Exercício 39. Calcule o produto interno dos seguintes pares de matrizes, para o produto interno usual em $M_{m\times n}$.

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

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

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

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

Em muitos contextos de relevância prática, há interesse em aproximar funções por outras mais simples. Por forma a poder utilizar as técnicas que introduziremos mais adiante, é necessário definir produtos internos em espaços de funções.

Um dos produtos internos mais comuns, definido no espaço $\mathcal{C}^0([a,b])$ das funções contínuas no intervalo $[a,b]$, é definido como

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt\,.$$

Proposição 40. A operação $\langle\cdot,\cdot\rangle : \mathcal{C}^0([a,b])\times\mathcal{C}^0([a,b])\to\mathbb{R}$ definida por

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt$$

é um produto interno em $\mathcal{C}^0([a,b])$.

Demonstração. Temos de verificar as quatro propriedades do produto interno.

1. Simetria: pela definição, temos que

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt = \int_a^b g(t)f(t)\,dt = \langle g,f\rangle$$

atendendo à comutatividade do produto de números reais.

2. Positividade: o produto interno duma função f consigo própria é

$$\langle f,f\rangle = \int_a^b f^2(t)\,dt$$

que é positivo porque $f^2(t)\geq 0$.

3. Linearidade: sendo α um número real,

$$\langle \alpha f, g\rangle = \int_a^b (\alpha f)(t)g(t)\ dt = \int_a^b \alpha f(t)g(t)\ dt = \alpha \int_a^b f(t)g(t)\ dt = \alpha \langle f, g\rangle$$

pela linearidade do integral.

4. Distributividade:

$$\begin{aligned}\langle f, g+h\rangle &= \int_a^b f(t)(g+h)(t)\ dt &&= \int_a^b f(t)(g(t)+h(t))\ dt \\ &= \int_a^b f(t)g(t)+f(t)h(t)\ dt &&= \int_a^b f(t)h(t)\ dt + \int_a^b g(t)h(t)\ dt \\ &= \langle f, h\rangle + \langle g, h\rangle\end{aligned}$$

Uma vez que a operação satisfaz estas quatro propriedades, temos que $\mathcal{C}^0([a,b])$ com aquele produto interno é um espaço euclideano. □

Exemplo. O produto interno de $f(x) = \cos x$ e $g(x) = 1$ em $\mathcal{C}^0\left(\left[0, \frac{\pi}{2}\right]\right)$ é

$$\int_0^{\frac{\pi}{2}} \cos x\ dx = [\sin(x)]_0^{\frac{\pi}{2}} = 1\,.$$

Em particular, como os polinómios são funções contínuas, este produto interno é um produto interno em P_n, para qualquer n. Assim, os espaços P_n também são espaços euclideanos.

Exemplo. Apresentam-se de seguida alguns exemplos de produtos internos de polinómios. Assume-se sempre que estamos a trabalhar em $P_n([-1,1])$ com n suficientemente elevado para conter os polinómios em questão.

1. O produto interno de $p(x) = 2x+1$ e $q(x) = x$ é dado por

$$\int_{-1}^1 (2x+1)x\ dx = \int_{-1}^1 \left(2x^2+x\right)\ dx = \left[\frac{2x^3}{3} + \frac{x^2}{2}\right]_{-1}^1 = \frac{4}{3}\,.$$

2. O produto interno de $p(x) = 2x+1$ com $r(x) = 2x-1$ é

$$\int_{-1}^1 (2x+1)(2x-1)\ dx = \int_{-1}^1 \left(4x^2-1\right)\ dx = \left[\frac{4x^3}{3} - x\right]_{-1}^1 = \frac{2}{3}\,.$$

3. O produto interno de $p(x) = 2x+1$ com $s(x) = 1$ é

$$\int_{-1}^1 (2x+1)\ dx = \left[x^2+x\right]_{-1}^1 = 2\,.$$

Podíamos ter chegado a este resultado sem efectuar os cálculos observando que

$$s(x) = 1 = 2x - (2x-1) = 2q(x) - r(x)$$

e aplicando a linearidade do produto interno para concluir que

$$\langle p, s\rangle = \langle p, 2q-r\rangle = 2\langle p, q\rangle - \langle p, r\rangle = 2\times\frac{4}{3} - \frac{2}{3} = 2\,.$$

Exercício 40. Calcule o produto interno dos seguintes pares de funções em $P_2([-1,1])$.

(a) x e 1 (b) $x-1$ e $x+1$ (c) x^2-1 e x (d) $2x^2-1$ e x

2.6.2 Comprimentos e distâncias

Voltando mais uma vez a $\mathbb{R}^n$, sabemos que podemos determinar o comprimento dum vector (distância do seu extremo à origem) aplicando o teorema de Pitágoras. Em $\mathbb{R}^2$, o comprimento de (x,y) é $\sqrt{x^2+y^2}$ (Figura 2.7, à esquerda); em $\mathbb{R}^3$, o comprimento de (x,y,z) é $\sqrt{x^2+y^2+z^2}$ (Figura 2.7, à direita); e, de forma geral, em $\mathbb{R}^n$ podemos calcular o comprimento de $\overrightarrow{v}=(v_1,v_2,\ldots,v_n)$ por

$$\|\overrightarrow{v}\| = \sqrt{(v_1)^2+(v_2)^2+\cdots+(v_n)^2} = \sqrt{\overrightarrow{v}\cdot\overrightarrow{v}}\,.$$

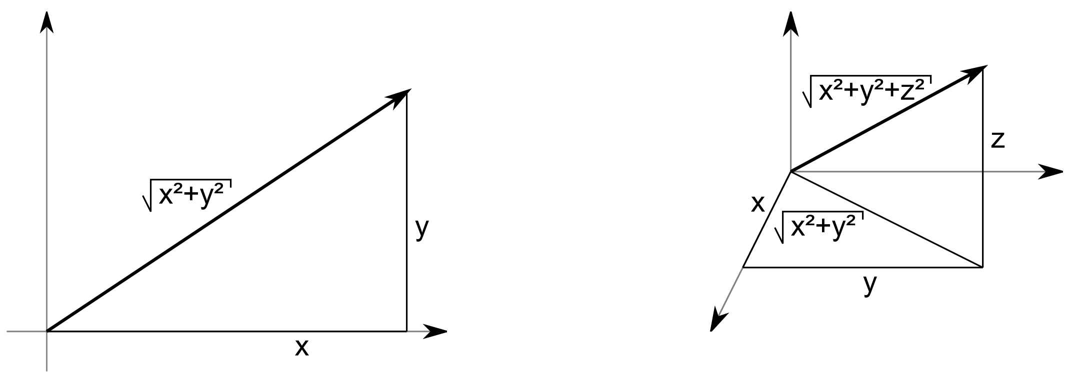


Figura 2.7: Determinação do comprimento de vectores em $\mathbb{R}^2$ e $\mathbb{R}^3$.

Tal como atrás, podemos generalizar esta relação a qualquer espaço euclideano, definindo a norma da seguinte forma.

Definição. Seja V um espaço euclideano com produto interno $\langle\cdot,\cdot\rangle$. A *norma* dum vector u de V é definida como
$$\|u\| = \sqrt{\langle u,u\rangle}\,.$$

Exercício 41. Calcule a norma (usual) de cada um dos seguintes vectores de $\mathbb{R}^2$ e $\mathbb{R}^3$.

(a) $(0,1)$ (c) $(2,1)$ (e) $(-2,-1)$ (g) $(0,2,1)$ (i) $(1,1,1)$

(b) $(-1,-2)$ (d) $(-2,1)$ (f) $(-1,2,1)$ (h) $(-2,2,2)$ (j) $(-1,3,2)$

O facto de ser definida à custa do produto interno faz com que haja um conjunto de propriedades da norma que se verificam em qualquer espaço euclideano. Por exemplo, uma vez que $\|u\|$ é definida como uma raiz quadrada, automaticamente $\|u\|\geq 0$ para qualquer u, sabendo-se ainda da positividade do produto interno que a igualdade se verifica apenas quando $u = 0_V$. De forma semelhante,

$$\|\alpha u\| = \sqrt{\langle \alpha u, \alpha u\rangle} = \sqrt{\alpha^2 \langle u, u\rangle} = |\alpha| \|u\|$$

Proposição 41. Sejam V um espaço euclideano com produto interno $\langle \cdot, \cdot\rangle$ e norma $\|.\|$ definida pela relação $\|u\| = \sqrt{\langle u, u\rangle}$. Então as seguintes igualdades são verificadas por quaisquer $u, v \in V$ e $\alpha \in \mathbb{R}$.

1. Positividade: $\|u\| \geq 0$
2. Zero: $\|u\| = 0$ sse $u = 0_V$
3. Linearidade: $\|\alpha u\| = |\alpha| \|u\|$
4. Desigualdade triangular: $\|u + v\| \leq \|u\| + \|v\|$

A única propriedade que não foi verificada foi a desigualdade triangular. Em $\mathbb{R}^n$, esta propriedade reduz-se simplesmente ao facto de o comprimento de qualquer lado dum triângulo (dado por $\|u + v\|$) ser inferior à soma dos outros dois (dada por $\|u\| + \|v\|$). Nos outros espaços vectoriais, é consequência da *desigualdade de Cauchy–Schwarz*, válida em qualquer espaço vectorial, e que afirma que, para quaisquer vectores u e v, se tem

$$| \langle u, v\rangle | \leq \|u\| \|v\| .$$

Não vamos demonstrar a desigualdade de Cauchy–Schwarz. A partir dela, contudo, é fácil mostrar que a desigualdade triangular se verifica para quaisquer vectores u e v:

$$\begin{aligned}
\|u + v\| \leq \|u\| + \|v\| &\iff \|u + v\|^2 \leq (\|u\| + \|v\|)^2 \iff \|u + v\|^2 \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| \\
&\iff \langle u + v, u + v\rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| \\
&\iff \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| \\
&\iff \|u\|^2 + 2 \langle u, v\rangle + \|v\|^2 \leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| \\
&\iff 2 \langle u, v\rangle \leq 2\|u\| \|v\| \iff \langle u, v\rangle \leq \|u\| \|v\|
\end{aligned}$$

e a última relação é consequência da desigualdade de Cauchy–Schwarz, pois $\langle u, v\rangle \leq | \langle u, v\rangle |$.

Exemplo. Vejamos alguns exemplos destas relações no espaço $\mathbb{R}^3$. Sejam $\overrightarrow{u} = (1, 2, -1)$ e $\overrightarrow{v} = (3, -1, 2)$. Da definição de produto interno, temos as seguintes igualdades.

$$\begin{aligned}
\langle \overrightarrow{u}, \overrightarrow{u}\rangle &= (1, 2, -1) \cdot (1, 2, -1) = 1 \times 1 + 2 \times 2 + (-1) \times (-1) = 6 \\
\langle \overrightarrow{u}, \overrightarrow{v}\rangle &= (1, 2, -1) \cdot (3, -1, 2) = 1 \times 3 + 2 \times (-1) + (-1) \times 2 = -1 \\
\langle \overrightarrow{v}, \overrightarrow{v}\rangle &= (3, -1, 2) \cdot (3, -1, 2) = 3 \times 3 + (-1) \times (-1) + 2 \times 2 = 14
\end{aligned}$$

Daqui se conclui que $\|u\| = \sqrt{\langle u, u\rangle} = \sqrt{6}$ e $\|v\| = \sqrt{14}$. A desigualdade de Cauchy–Schwarz toma neste caso a forma

$$| \langle u, v\rangle | \leq \|u\| \|v\| \iff 1 \leq \sqrt{6}\sqrt{14} ,$$

que é claramente verdadeira.

Observe-se que, considerando dois vectores (a, b, c) e (x, y, z) genéricos, a desigualdade de Cauchy–Schwarz afirma que

$$|ax + by + cz| \leq \sqrt{a^2 + b^2 + c^2}\sqrt{d^2 + e^2 + f^2} ,$$

relação que é portanto válida para quaisquer reais a, b, c, x, y e z.

Noutros espaços vectoriais, o raciocínio é semelhante.

Exemplo.

1. Consideremos o espaço linear $M_{3\times 2}$ e sejam A, B e C as seguintes matrizes.

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

Temos então as seguintes relações.

$$\begin{aligned} \langle A, B\rangle &= \left\langle \begin{bmatrix} -1 & 5 \\ 2 & -3 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix} \right\rangle \\ &= (-1)\times 1 + 5\times 0 + 2\times 2 + (-3)\times(-1) + 1\times 3 + 0\times(-2) = 9 \end{aligned}$$

$$\begin{aligned} \langle A, C\rangle &= \left\langle \begin{bmatrix} -1 & 5 \\ 2 & -3 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 2 & 2 \end{bmatrix} \right\rangle \\ &= (-1)\times 2 + 5\times(-1) + 2\times 1 + (-3)\times 3 + 1\times 2 + 0\times 2 = -12 \end{aligned}$$

$$\begin{aligned} \langle B, C\rangle &= \left\langle \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 2 & 2 \end{bmatrix} \right\rangle \\ &= 1\times 2 + 0\times(-1) + 2\times 1 + (-1)\times 3 + 3\times 2 + (-2)\times 2 = 3 \end{aligned}$$

Por simetria, tem-se ainda que

$$\langle B, A\rangle = \langle A, B\rangle = 9 \qquad \langle C, A\rangle = \langle A, C\rangle = -12 \qquad \langle C, B\rangle = \langle B, C\rangle = 3\,.$$

A norma de cada uma destas matrizes é dada por

$$\begin{aligned} \|A\| = \sqrt{\langle A, A\rangle} &= \sqrt{\left\langle \begin{bmatrix} -1 & 5 \\ 2 & -3 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 5 \\ 2 & -3 \\ 1 & 0 \end{bmatrix} \right\rangle} \\ &= \sqrt{(-1)^2 + 5^2 + 2^2 + (-3)^2 + 1^2 + 0^2} = \sqrt{40} \end{aligned}$$

$$\begin{aligned} \|B\| = \sqrt{\langle B, B\rangle} &= \sqrt{\left\langle \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix} \right\rangle} \\ &= \sqrt{1^2 + 0^2 + 2^2 + (-1)^2 + 3^2 + (-2)^2} = \sqrt{19} \end{aligned}$$

$$\begin{aligned} \|C\| = \sqrt{\langle C, C\rangle} &= \sqrt{\left\langle \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 2 & 2 \end{bmatrix} \right\rangle} \\ &= \sqrt{2^2 + (-1)^2 + 1^2 + 3^2 + 2^2 + 2^2} = \sqrt{23} \end{aligned}$$

A desigualdade de Cauchy–Schwarz implica as três desigualdades

$$
\begin{aligned}
|\langle A,B\rangle| \leq \|A\|\|B\| &\Longleftrightarrow 9 \leq \sqrt{40}\sqrt{19} = \sqrt{760}\\
|\langle A,C\rangle| \leq \|A\|\|C\| &\Longleftrightarrow 12 \leq \sqrt{40}\sqrt{23} = \sqrt{920}\\
|\langle B,C\rangle| \leq \|B\|\|C\| &\Longleftrightarrow 3 \leq \sqrt{19}\sqrt{23} = \sqrt{437}
\end{aligned}
$$

que são verdadeiras.

2. Em espaços de funções com o produto interno dado pelo integral, os cálculos são um pouco mais complexos mas obtemos o mesmo tipo de resultados. Vejamos alguns exemplos em $P_2([0,1])$; sejam p, q e r os polinómios seguintes.

$$p(x) = x+1 \qquad q(x) = 2x-1 \qquad r(x) = 3x^2-2$$

Temos então as seguintes relações.

$$
\begin{aligned}
\langle p,q\rangle = \langle x+1, 2x-1\rangle &= \int_0^1 (x+1)(2x-1)\,dx = \int_0^1 \left(2x^2+x-1\right)\,dx\\
&= \left[\frac{2x^3}{3}+\frac{x^2}{2}-x\right]_0^1 = \frac{2}{3}+\frac{1}{2}-1 = \frac{1}{6}
\end{aligned}
$$

$$
\begin{aligned}
\langle p,r\rangle = \left\langle x+1, 3x^2-2\right\rangle &= \int_0^1 (x+1)\left(3x^2-2\right)\,dx = \int_0^1 \left(3x^3+3x^2-2x-2\right)\,dx\\
&= \left[\frac{3x^4}{4}+x^3-x^2-2x\right]_0^1 = \frac{3}{4}+1-1-2 = -\frac{5}{4}
\end{aligned}
$$

$$
\begin{aligned}
\langle q,r\rangle = \left\langle 2x-1, 3x^2-2\right\rangle &= \int_0^1 (2x-1)\left(3x^2-2\right)\,dx = \int_0^1 \left(6x^3-3x^2-4x+2\right)\,dx\\
&= \left[\frac{3x^4}{2}-x^3-2x^2+2x\right]_0^1 = \frac{3}{2}-1-2+2 = \frac{1}{2}
\end{aligned}
$$

Por simetria, tem-se ainda que

$$\langle p,q\rangle = \langle q,p\rangle = \frac{1}{6} \qquad \langle p,r\rangle = \langle r,p\rangle = -\frac{5}{4} \qquad \langle q,r\rangle = \langle r,q\rangle = \frac{1}{2}\,.$$

A norma de cada uma destas funções é dada por

$$
\begin{aligned}
\|p\| = \sqrt{\langle p,p\rangle} &= \sqrt{\int_0^1 (x+1)^2\,dx} = \sqrt{\int_0^1 \left(x^2+2x+1\right)\,dx}\\
&= \sqrt{\left[\frac{x^3}{3}+x^2+x\right]_0^1} = \sqrt{\frac{1}{3}+1+1} = \sqrt{\frac{7}{3}}
\end{aligned}
$$

$$
\begin{aligned}
\|q\| = \sqrt{\langle q,q\rangle} &= \sqrt{\int_0^1 (2x-1)^2\,dx} = \sqrt{\int_0^1 \left(4x^2-4x+1\right)\,dx}\\
&= \sqrt{\left[\frac{4x^3}{3}-2x^2+x\right]_0^1} = \sqrt{\frac{4}{3}-2+1} = \sqrt{\frac{1}{3}}
\end{aligned}
$$

$$\begin{aligned}\|r\| = \sqrt{\langle r, r\rangle} &= \sqrt{\int_0^1 \left(3x^2-2\right)^2\, dx} = \sqrt{\int_0^1 \left(9x^4-12x^2+4\right)\, dx} \\ &= \sqrt{\left[\frac{9x^5}{5} - 4x^3 + 4x\right]_0^1} = \sqrt{\frac{9}{5} - 4 + 4} = \frac{3}{\sqrt{5}}\end{aligned}$$

A desigualdade de Cauchy–Schwarz implica as três desigualdades

$$\begin{aligned}|\langle p, q\rangle| \leq \|p\|\|q\| &\iff \frac{1}{6} \leq \sqrt{\frac{7}{3}}\sqrt{\frac{1}{3}} = \frac{\sqrt{7}}{3} \\ |\langle p, r\rangle| \leq \|p\|\|r\| &\iff \frac{5}{4} \leq \sqrt{\frac{7}{3}}\sqrt{\frac{9}{5}} = \sqrt{\frac{21}{5}} \\ |\langle q, r\rangle| \leq \|q\|\|r\| &\iff \frac{1}{2} \leq \sqrt{\frac{1}{3}}\sqrt{\frac{9}{5}} = \sqrt{\frac{3}{5}}\end{aligned}$$

que são verdadeiras.

A desigualdade de Cauchy–Schwarz para este produto interno em espaços de funções implica a seguinte relação entre integrais definidos.

$$\left|\int_a^b f(t)g(t)\, dt\right| \leq \int_a^b f^2(t)\, dt \int_a^b g^2(t)\, dt$$

Exercício 42. Verifique a desigualdade de Cauchy–Schwarz para os pares de vectores:

(a) $(1, 2, 1)$ e $(0, 2, -1)$ em $\mathbb{R}^3$;

(b) $(0, 1, -1)$ e $(1, 2, 3)$ em $\mathbb{R}^3$;

(c) $\begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$ e $\begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$ em $M_{2\times 2}$;

(d) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ e $\begin{bmatrix} -3 & 0 \\ 1 & 2 \end{bmatrix}$ em $M_{2\times 2}$;

(e) $x^2 + 1$ e $-x + 1$ em $P_2([-1, 1])$;

(f) $x - 1$ e $x^2 + x$ em $P_2([-1, 1])$.

O conceito de norma pode ser usado para definir distância entre dois vectores em qualquer espaço vectorial. Voltemos ao caso de $\mathbb{R}^n$; a distância entre dois vectores $\overrightarrow{u}$ e $\overrightarrow{v}$ é o comprimento do vector que une as suas extremidades, dado por $\overrightarrow{v} - \overrightarrow{u}$ (ver Figura 2.8).

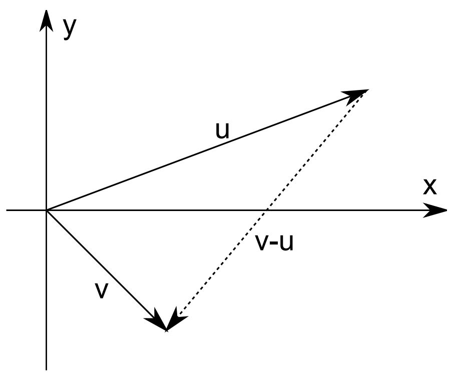


Figura 2.8: Distância entre dois vectores em $\mathbb{R}^n$.

Podemos generalizar esta relação para qualquer espaço vectorial.

Definição. Sejam V um espaço vectorial e u, v vectores de V. A *distância* entre u e v é

$$d(u,v) = \|v-u\| .$$

A distância goza de propriedades muito semelhantes às da norma.

Proposição 42. Sejam V um espaço euclideano com produto interno $\langle \cdot, \cdot \rangle$ e distância dada por $d(u,v) = \|v-u\|$. Então as seguintes relações verificam-se para quaisquer $u, v, w \in V$ e $\alpha \in \mathbb{R}$.

1. Positividade: $d(u,v) \geq 0$
2. Zero: $d(u,v) = 0$ sse $u = v$
3. Linearidade: $d(\alpha u, v) = d(u, \alpha v) = \alpha\, d(u,v)$
4. Desigualdade triangular: $d(u,v) \leq d(u,w) + d(w,v)$

É possível definir distâncias directamente em espaços vectoriais não euclideanos, exigindo--se apenas que satisfaçam estas propriedades (à semelhança da forma como definimos um produto interno num espaço linear arbitrário). Nesta exposição, porém, apenas vamos considerar distâncias definidas a partir da norma.

Observe-se que, expandindo a definição de norma, podemos encontrar uma fórmula explícita para a distância entre dois vectores à custa do produto interno.

$$d(u,v) = \|v-u\| = \sqrt{\langle v-u, v-u \rangle} = \sqrt{\langle v,v \rangle + \langle u,u \rangle - 2\langle u,v \rangle}$$

Porém, na prática é mais simples calcular directamente o vector $v-u$ do que recorrer a estas relações.

Vejamos alguns exemplos.

Exemplo.

1. A distância entre os vectores $\overrightarrow{u} = (1,2,-1)$ e $\overrightarrow{v} = (3,-1,2)$ de $\mathbb{R}^3$ considerados no exemplo anterior é dada por

$$d\left(\overrightarrow{u}, \overrightarrow{v}\right) = \|v-u\| = \|(3,-1,2)-(1,2,-1)\| = \|(2,-3,3)\| = \sqrt{2^2+(-3)^2+3^2} = \sqrt{22}\,.$$

2. Para as matrizes A, B e C de $M_{3\times 2}$ do exemplo anterior, temos

$$\begin{aligned} d(A,B) = \|B-A\| &= \left\| \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix} - \begin{bmatrix} -1 & 5 \\ 2 & -3 \\ 1 & 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 2 & -5 \\ 0 & 2 \\ 2 & -2 \end{bmatrix} \right\| \\ &= \sqrt{2^2+(-5)^2+0^2+2^2+2^2+(-2)^2} = \sqrt{35} \end{aligned}$$

$$\begin{aligned} d(A,C) = \|C-A\| &= \left\| \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} -1 & 5 \\ 2 & -3 \\ 1 & 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 3 & -6 \\ -1 & 6 \\ 1 & 2 \end{bmatrix} \right\| \\ &= \sqrt{3^2+(-6)^2+(-1)^2+6^2+1^2+2^2} = \sqrt{87} \end{aligned}$$

$$\begin{aligned} d(B,C) = \|C-B\| &= \left\| \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1 & -1 \\ -1 & 4 \\ -1 & 4 \end{bmatrix} \right\| \\ &= \sqrt{1^2 + (-1)^2 + (-1)^2 + 4^2 + (-1)^2 + 4^2} = \sqrt{36} = 6\,, \end{aligned}$$

o que significa que, geometricamente, a matriz A está mais próxima de B do que de C. Podemos interpretar este facto dizendo que A é uma melhor aproximação de B do que C.

3. Para os polinómios p, q e r de $P_2([0,1])$ considerados atrás, temos as relações seguintes.

$$\begin{aligned} d(p,q) = \|q-p\| &= \|(2x-1)-(x+1)\| = \|x-2\| \\ &= \sqrt{\int_0^1 (x-2)^2\,dx} = \sqrt{\int_0^1 \left(x^2-4x+4\right)\,dx} \\ &= \sqrt{\left[\frac{x^3}{3} - 2x^2 + 4x\right]_0^1} = \sqrt{\frac{1}{3} - 2 + 4} = \sqrt{\frac{7}{3}} \end{aligned}$$

$$\begin{aligned} d(p,r) = \|r-p\| &= \left\|\left(3x^2-2\right)-(x+1)\right\| = \left\|3x^2-x-3\right\| \\ &= \sqrt{\int_0^1 \left(3x^2-x-3\right)^2\,dx} = \sqrt{\int_0^1 \left(9x^4-6x^3-17x^2+6x+9\right)\,dx} \\ &= \sqrt{\left[\frac{9x^5}{5} - \frac{3x^4}{2} - \frac{17x^3}{3} + 3x^2 + 9x\right]_0^1} = \sqrt{\frac{9}{5} - \frac{3}{2} - \frac{17}{3} + 3 + 9} \\ &= \sqrt{\frac{199}{30}} \end{aligned}$$

$$\begin{aligned} d(q,r) = \|r-q\| &= \left\|\left(3x^2-2\right)-(2x-1)\right\| = \left\|3x^2+2x-1\right\| \\ &= \sqrt{\int_0^1 \left(3x^2+2x-1\right)^2\,dx} = \sqrt{\int_0^1 \left(9x^4+12x^3-2x^2-4x+1\right)\,dx} \\ &= \sqrt{\left[\frac{9x^5}{5} + 3x^4 - \frac{2x^3}{3} - 2x^2 + x\right]_0^1} = \sqrt{\frac{9}{5} + 3 - \frac{2}{3} - 2 + 1} \\ &= \sqrt{\frac{47}{15}} \end{aligned}$$

Podemos interpretar estes resultados geometricamente como indicando que p está mais próximo de q do que de r. Isto significa que a área entre os gráficos de p e q é menor do que a área entre os gráficos de p e r (ver Figura 2.9).

De facto, com esta definição de produto interno entre duas funções, a norma duma função está relacionada com a área sob o gráfico da função (embora não corresponda ao valor dessa área); assim, duas funções estão tanto mais próximas quanto menor for a área entre os seus gráficos — o que intuitivamente corresponde de facto a uma noção correcta de proximidade.

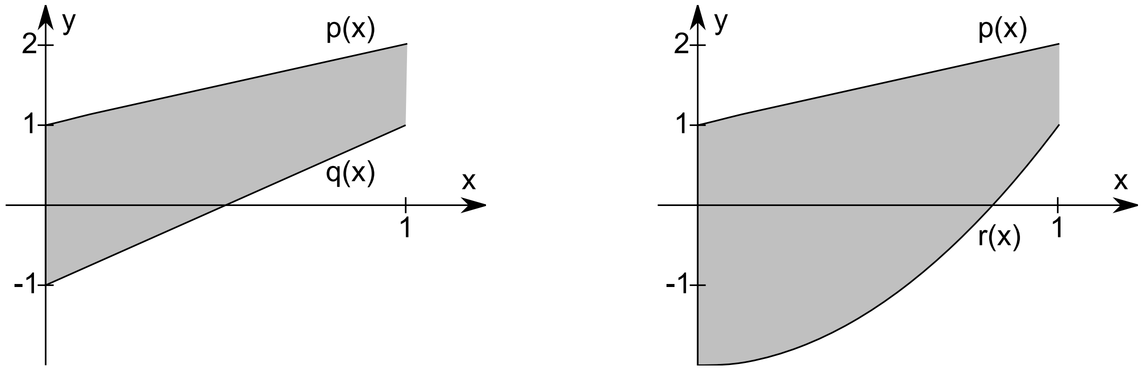


Figura 2.9: Comparação dos gráficos de $p(x)$, $q(x)$ e $r(x)$. A área entre p e q é menor do que a área entre p e r.

Exercício 43. Qual dos seguintes vectores é melhor aproximação de $(1,1,1)$?

(a) o vector $(1,2,3)$

(b) o vector $(1,0,1)$

Exercício 44. Qual das seguintes matrizes é melhor aproximação de $\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$?

(a) a matriz $\begin{bmatrix} -2 & 2 \\ 3 & 0 \end{bmatrix}$

(b) a matriz $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

Exercício 45. Qual dos seguintes polinómios é melhor aproximação de x^2+1?

(a) o polinómio $2x$

(b) o polinómio $x+2$

2.6.3 Ângulos e projecções ortogonais

Consideremos dois vectores arbitrários $\overrightarrow{u}$ e $\overrightarrow{v}$ em $\mathbb{R}^2$. Sendo α o ângulo formado por $\overrightarrow{u}$ com o eixo dos xx e β o ângulo formado por $\overrightarrow{v}$ com o eixo dos yy, as coordenadas de $\overrightarrow{u}$ e $\overrightarrow{v}$ podem ser escritas na forma

$$\overrightarrow{u} = (\|\overrightarrow{u}\|\cos(\alpha), \|u\|\sin(\alpha)) \qquad \overrightarrow{v} = (\|\overrightarrow{v}\|\cos(\beta), \|v\|\sin(\beta))$$

conforme ilustrado na Figura 2.10.

Se usarmos esta fórmula para calcular o produto interno entre os vectores $\overrightarrow{u}$ e $\overrightarrow{v}$, obtemos

$$\begin{aligned}
\overrightarrow{u}\cdot\overrightarrow{v} &= (\|\overrightarrow{u}\|\cos\alpha, \|\overrightarrow{u}\|\sin\alpha)\cdot(\|\overrightarrow{v}\|\cos\beta, \|\overrightarrow{v}\|\sin\beta) \\
&= \|\overrightarrow{u}\|\cos\alpha\,\|\overrightarrow{v}\|\cos\beta + \|\overrightarrow{u}\|\sin\alpha\,\|\overrightarrow{v}\|\sin\beta \\
&= \|\overrightarrow{u}\|\,\|\overrightarrow{v}\|\,(\cos\alpha\cos\beta+\sin\alpha\sin\beta) = \|\overrightarrow{u}\|\,\|\overrightarrow{v}\|\cos(\alpha-\beta) = \|\overrightarrow{u}\|\,\|\overrightarrow{v}\|\cos\theta
\end{aligned}$$

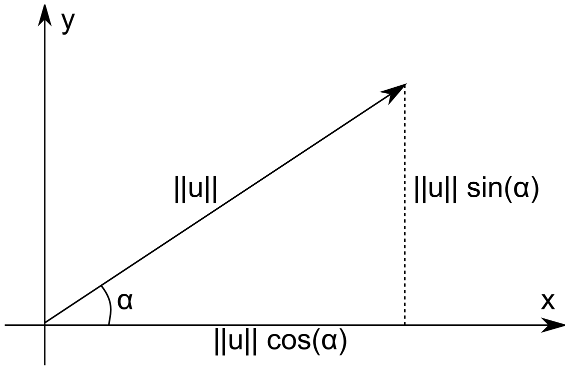


Figura 2.10: Outra forma de determinar as coordenadas dum vector em $\mathbb{R}^2$.

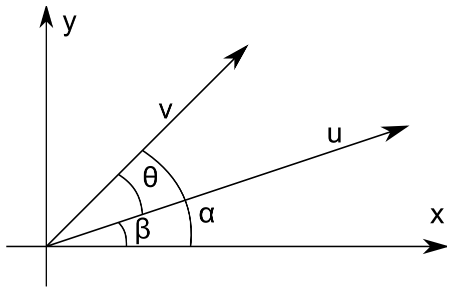


Figura 2.11: Ângulo entre dois vectores.

onde $\theta = \alpha - \beta$ é o ângulo formado pelos vectores $\overrightarrow{u}$ e $\overrightarrow{v}$ (ver Figura 2.11).

Decorre daqui que

$$\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\|\,\|\overrightarrow{v}\|\cos\theta \Longleftrightarrow \cos\theta = \frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\|\overrightarrow{u}\|\,\|\overrightarrow{v}\|} \Longleftrightarrow \theta = \arccos\frac{\overrightarrow{u} \cdot \overrightarrow{v}}{\|\overrightarrow{u}\|\,\|\overrightarrow{v}\|}$$

Mais uma vez, podemos usar esta relação (válida aliás em qualquer espaço $\mathbb{R}^n$ de dimensão superior a 2) para definir ângulo entre dois vectores em qualquer espaço euclideano V.

Definição. Sejam V um espaço euclideano com produto interno $\langle\cdot,\cdot\rangle$ e u, v vectores de V. O *ângulo* entre os vectores u e v é definido pela expressão

$$\angle(u,v) = \arccos\frac{\langle u,v\rangle}{\|u\|\|v\|}\,.$$

Devido à desigualdade de Cauchy–Schwarz, esta expressão está sempre bem definida: uma vez que $|\langle u,v\rangle| \leq \|u\|\|v\|$, tem-se necessariamente $-1 \leq \frac{\langle u,v\rangle}{\|u\|\|v\|} \leq 1$, pelo que aquele valor corresponde sempre ao coseno de algum ângulo.

Vejamos alguns exemplos, continuando com os espaços e vectores que considerámos atrás.

Exemplo.

1. O ângulo entre os vectores $\overrightarrow{u} = (1,2,-1)$ e $\overrightarrow{v} = (3,-1,2)$ de $\mathbb{R}^3$ é dado por

$$\angle(\overrightarrow{u},\overrightarrow{v}) = \arccos\frac{\langle\overrightarrow{u},\overrightarrow{v}\rangle}{\|\overrightarrow{u}\|\,\|\overrightarrow{v}\|} = \arccos\frac{-1}{\sqrt{6}\sqrt{14}}\,.$$

Este valor corresponde aproximadamente a -83^{o}.

2. As matrizes A, B e C de $M_{3\times 2}$ definidas atrás formam os seguintes ângulos entre si.

$$\begin{aligned}
\angle(A,B) &= \arccos\frac{\langle A,B\rangle}{\|A\|\|B\|} = \arccos\frac{9}{\sqrt{760}} \approx 70^{\text{o}}\\
\angle(A,C) &= \arccos\frac{\langle A,C\rangle}{\|A\|\|C\|} = \arccos\frac{-12}{\sqrt{920}} \approx 66.7^{\text{o}}\\
\angle(B,C) &= \arccos\frac{\langle B,C\rangle}{\|B\|\|C\|} = \arccos\frac{3}{\sqrt{437}} \approx 81.7^{\text{o}}
\end{aligned}$$

3. Para os polinómios p, q e r de $P_2([0,1])$ considerados atrás, temos os seguintes ângulos.

$$\begin{aligned}
\angle(p,q) &= \arccos\frac{\langle p,q\rangle}{\|p\|\|q\|} = \arccos\frac{\frac{1}{6}}{\frac{\sqrt{7}}{3}} = \arccos\frac{1}{2\sqrt{7}} \approx 79^{\text{o}}\\
\angle(p,r) &= \arccos\frac{\langle p,r\rangle}{\|p\|\|r\|} = \arccos\frac{-\frac{5}{4}}{\sqrt{\frac{21}{5}}} = \arccos\left(-\sqrt{\frac{125}{336}}\right) \approx -52.4^{\text{o}}\\
\angle(q,r) &= \arccos\frac{\langle q,r\rangle}{\|q\|\|r\|} = \arccos\frac{\frac{1}{2}}{\sqrt{\frac{3}{5}}} = \arccos\sqrt{\frac{5}{12}} \approx 49.8^{\text{o}}
\end{aligned}$$

Exercício 46. Calcule o ângulo formado pelos dois vectores de cada alínea dos Exercícios 38 e 40. Quais desses pares são perpendiculares?

Há alguns casos particulares de interesse. Em primeiro lugar, é útil observar que dois vectores formam um ângulo diferente de 0^{o} precisamente quando são linearmente independentes. Por semelhança com o que se passa em $\mathbb{R}^n$, dizemos que dois vectores são ortogonais quando fazem um ângulo de 90^{o} entre si.

Definição. Seja V um espaço euclideano. Dois vectores u e v dizem-se *ortogonais*, e escreve-se $u \perp v$, quando $\angle(u,v) = \frac{\pi}{2}$, ou, equivalentemente, quando $\langle u,v\rangle = 0$.

Por exemplo, os vectores $(1,1)$ e $(1,-1)$ de $\mathbb{R}^2$ são ortogonais, uma vez que

$$(1,1)\cdot(1,-1) = 1\times 1 + 1\times(-1) = 0\,.$$

Os vectores $(2,1,3)$ e $(3,0,-2)$ de $\mathbb{R}^3$ também são ortogonais, uma vez que

$$(2,1,3)\cdot(3,0,-2) = 2\times 3 + 1\times 0 + 3\times(-2) = 0\,.$$

Em ambos os casos, podemos verificar (desenhando os vectores) que de facto o ângulo entre eles é um ângulo recto.

O mesmo conceito funciona noutros espaços vectoriais.

Exemplo.

1. As matrizes $\mathbf{I_2}$ e $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ de $M_{2\times 2}$ são ortogonais, pois

$$\left\langle \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\rangle = 1\times 0 + 0\times 1 + 0\times(-1) + 1\times 0 = 0\,.$$

2. As matrizes
$$\begin{bmatrix} 2 & 1 \\ 3 & 1 \\ -1 & 5 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 0 & -1 \end{bmatrix}$$
de $M_{3\times 2}$ também são ortogonais, uma vez que
$$\left\langle \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ -1 & 5 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 0 & -1 \end{bmatrix} \right\rangle = 2\times 1+1\times(-1)+3\times 2+1\times(-2)+(-1)\times 0+5\times(-1) = 0\,.$$

3. Duas funções podem também ser ortogonais. Por exemplo, em $\mathcal{C}^0\left([0,2\pi]\right)$ as funções $f(x) = \sin(x)$ e $g(x) = 1$ são ortogonais, uma vez que
$$\langle \sin(x), 1\rangle = \int_0^{2\pi} \sin(x)\,dx = [-\cos(x)]_0^{2\pi} = 0\,.$$

4. Os polinómios $p(x) = x+1$ e $q(x) = x-a$ são ortogonais em $P_1([-1,1])$ se
$$\begin{aligned}\langle x+1, x-a\rangle &= \int_{-1}^{1} (x+1)(x-a)\,dx = \int_{-1}^{1} \left(x^2+(1-a)x-a\right)\,dx \\ &= \left[\frac{x^3}{3} + \frac{(1-a)x^2}{2} - ax\right]_{-1}^{1} = \frac{2}{3} - 2a\end{aligned}$$
for igual a 0, ou seja, se $a = \frac{1}{3}$.

5. Já em $P_1([0,1])$, para os polinómios $p(x) = x+1$ e $q(x) = x-a$ serem ortogonais tem de se ter
$$\begin{aligned}\langle x+1, x-a\rangle &= \int_{0}^{1} (x+1)(x-a)\,dx = \int_{0}^{1} \left(x^2+(1-a)x-a\right)\,dx \\ &= \left[\frac{x^3}{3} + \frac{(1-a)x^2}{2} - ax\right]_{0}^{1} = \frac{1}{3} + \frac{1-a}{2} - a\end{aligned}$$
igual a 0, ou seja
$$\frac{5}{6} - \frac{3}{2}a = 0\,.$$
donde $a = \frac{5}{9}$.

Exercício 47. Para cada um dos seguintes pares de vectores de $\mathbb{R}^2$, determine os valores de k que os tornam perpendiculares com o produto interno usual.

(a) $(1,0)$ e $(-1,k)$ (b) $(2,k)$ e $(-1,k)$ (c) $(2,k)$ e $(-2,-k)$

O conceito de ortogonalidade é extraordinariamente útil em aplicações porque nos permite definir a noção de projecção ortogonal. Em $\mathbb{R}^n$, a projecção ortogonal dum vector $\overrightarrow{u}$ sobre outro vector $\overrightarrow{v}$ obtém-se traçando uma recta perpendicular a $\overrightarrow{v}$ que passa pelo extremo de $\overrightarrow{u}$; a projecção de $\overrightarrow{u}$ sobre $\overrightarrow{v}$ é o vector com extremidade nessa intersecção (ver Figura 2.12).

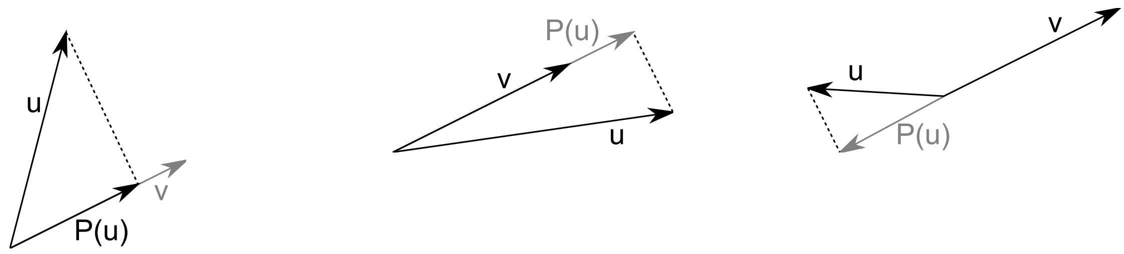


Figura 2.12: Três exemplos de projecção de um vector u sobre outro vector v.

Com um pouco de geometria analítica, conseguimos escrever explicitamente a expressão para a projecção ortogonal de $\overrightarrow{u}$ sobre $\overrightarrow{v}$. O comprimento dessa projecção é dado por $\|\overrightarrow{u}\|\cos\theta$, onde θ é o ângulo entre $\overrightarrow{u}$ e $\overrightarrow{v}$; a direcção é simplesmente a direcção de $\overrightarrow{v}$. Para obtermos um vector com a direcção de $\overrightarrow{v}$ e o comprimento que desejamos, vamos dividir $\overrightarrow{v}$ pelo seu comprimento (obtendo um vector unitário) e multiplicar pelo comprimento desejado. Usando a expressão que já conhecemos para $\cos\theta$, dada também pelo produto interno, concluímos que a projecção de $\overrightarrow{u}$ sobre $\overrightarrow{v}$ é o vector $P_{\overrightarrow{v}}\left(\overrightarrow{u}\right)$ dado por

$$P_{\overrightarrow{v}}\left(\overrightarrow{u}\right) = \frac{\overrightarrow{v}}{\|\overrightarrow{v}\|}\|\overrightarrow{u}\|\cos\theta = \frac{\overrightarrow{v}}{\|\overrightarrow{v}\|}\|\overrightarrow{u}\|\frac{\overrightarrow{u}\cdot\overrightarrow{v}}{\|\overrightarrow{u}\|\|\overrightarrow{v}\|} = \frac{\overrightarrow{u}\cdot\overrightarrow{v}}{\|\overrightarrow{v}\|^2}\overrightarrow{v}\,.$$

No caso em que $\|v\| = 1$ esta expressão reduz-se simplesmente a

$$P_{\overrightarrow{v}}\left(\overrightarrow{u}\right) = \left(\overrightarrow{u}\cdot\overrightarrow{v}\right)\overrightarrow{v}\,.$$

Uma vez que qualquer espaço euclideano tem esta estrutura geométrica, podemos reproduzir este raciocínio para todos eles, obtendo um conceito geral de projecção ortogonal.

Definição. Sejam V um espaço euclideano com produto interno $\langle\cdot,\cdot\rangle$ e u, v dois vectores de V. A *projecção ortogonal* de u sobre v é o vector $P_v(u)$ definido por

$$P_v(u) = \frac{\langle u, v\rangle}{\|v\|^2}v\,.$$

A noção de projecção ortogonal tem muito mais aplicações do que as que são evidentes à partida. Olhando novamente para a Figura 2.12, vemos que a projecção de $\overrightarrow{u}$ sobre $\overrightarrow{v}$ é o vector pertencente ao espaço linear gerado por $\overrightarrow{v}$ que melhor aproxima $\overrightarrow{u}$: a distância de $\overrightarrow{u}$ aos pontos de $L\left(\{\overrightarrow{v}\}\right)$ é mínima para o ponto $P_{\overrightarrow{v}}\left(\overrightarrow{u}\right)$.

Isto sugere as seguintes duas definições.

Definição. Sejam V um espaço euclideano com produto interno $\langle\cdot,\cdot\rangle$, $W = L(\{v\})$ um seu subespaço linear de dimensão 1 e $u \in V$.

- A *melhor aproximação* de u em W é o vector $P_v(u)$.

- A *distância* de u a W é dada por $\|u - P_v(u)\|$.

Note-se que estas definições são independentes da escolha de v (repare-se que $v \neq 0_V$), uma vez que a projecção ortogonal foi definida à custa do comprimento de u e da *direcção* de v (que é comum a todos os elementos de W).

Exemplo. Para determinar a distância de $(1,1)$ à recta $y = 2x$, em $\mathbb{R}^2$, começamos por escrever esta recta como espaço gerado por um vector de $\mathbb{R}^2$, por exemplo $L(\{(1,2)\})$. A projecção de $(1,1)$ sobre a recta é

$$P_{(1,2)}\left((1,1)\right) = \frac{(1,1)\cdot(1,2)}{\|(1,2)\|^2}(1,2) = \frac{3}{5}(1,2) = \left(\frac{3}{5},\frac{6}{5}\right).$$

A distância de u àquela recta é então

$$\left\|(1,1) - P_{(1,2)}\left((1,1)\right)\right\| = \left\|\left(\frac{2}{5},-\frac{1}{5}\right)\right\| = \sqrt{\frac{1}{5}}.$$

Podemos aplicar um raciocínio semelhante para calcular a distância de um ponto $\overrightarrow{u}$ a uma recta que não passe pela origem. De facto, uma recta $y = mx + b$ pode ser vista como uma "soma" do subespaço linear $y = mx$ com o vector $(0,b)$; subtrair esse vector a $\overrightarrow{u}$ e à recta corresponde a aplicar uma translação à figura, transformação essa que preserva as distâncias (ver Figura 2.13). A distância de $\overrightarrow{u}$ a $y = mx + b$ é então igual à distância de $\overrightarrow{u} - (0,b)$ a $y = mx$, que vimos acima como calcular.

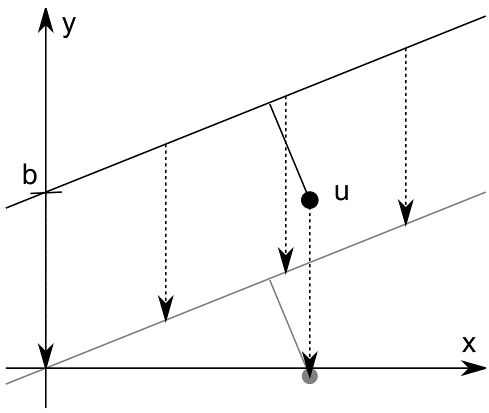


Figura 2.13: Uma translação permite transformar qualquer recta num subespaço linear de $\mathbb{R}^2$.

Exemplo.

1. Para calcular a distância, em $\mathbb{R}^2$, de $(2,3)$ à recta $y = -x$, começamos por encontrar um vector que gere esta recta — por exemplo $(1,-1)$. De seguida, calculamos a projecção de $(2,3)$ sobre esse vector.

$$P_{(1,-1)}\left((2,3)\right) = \frac{(2,3)\cdot(1,-1)}{\|(1,-1)\|^2}(1,-1) = \frac{-1}{2}(1,-1) = \left(-\frac{1}{2},\frac{1}{2}\right)$$

   A distância de $(2,3)$ a este vector é

$$\left\|(2,3) - \left(-\frac{1}{2},\frac{1}{2}\right)\right\| = \left\|\left(\frac{5}{2},\frac{5}{2}\right)\right\| = \frac{5}{\sqrt{2}}.$$

   Observe-se que o resultado seria o mesmo independentemente do vector usado para gerar a recta. Por exemplo, tomando o vector $(-2,2)$, obteríamos

$$P_{(-2,2)}\left((2,3)\right) = \frac{(2,3)\cdot(-2,2)}{\|(-2,2)\|^2}(-2,2) = \frac{2}{8}(-2,2) = \left(-\frac{1}{2},\frac{1}{2}\right),$$

   que é a mesma projecção determinada acima.

2. Se quiséssemos agora calcular a distância de $(2,3)$ à recta $y=-x+2$, bastaria proceder de forma semelhante para determinar a distância de $(2,3)-(0,2)=(2,1)$ a $y=-x$. Obteríamos os seguintes resultados.

$$P_{(1,-1)}\left((2,1)\right)=\frac{(2,1)\cdot(1,-1)}{\|(1,-1)\|^2}(1,-1)=\frac{1}{2}(1,-1)=\left(\frac{1}{2},\frac{1}{2}\right)$$
$$\left\|(2,1)-\left(\frac{1}{2},\frac{1}{2}\right)\right\|=\left\|\left(\frac{3}{2},\frac{1}{2}\right)\right\|=\sqrt{\frac{5}{2}}$$

3. Em dimensões superiores, o raciocínio é perfeitamente análogo. Em $\mathbb{R}^3$, para calcular a distância de $(2,1,-2)$ à recta definida por $y=2x$ e $z=x$, começamos por encontrar um vector que gere esta recta — por exemplo $(1,2,1)$. De seguida, calculamos a projecção de $(2,1,-2)$ sobre esse vector e a sua distância ao ponto $(2,1,-2)$.

$$P_{(1,2,1)}\left((2,1,-2)\right)=\frac{(2,1,-2)\cdot(1,2,1)}{\|(1,2,1)\|^2}(1,2,1)=\frac{2}{6}(1,2,1)=\left(\frac{1}{3},\frac{2}{3},\frac{1}{3}\right)$$
$$\left\|(2,1,-2)-\left(\frac{1}{3},\frac{2}{3},\frac{1}{3}\right)\right\|=\left\|\left(\frac{5}{3},\frac{1}{3},-\frac{7}{3}\right)\right\|=\frac{5\sqrt{3}}{3}$$

4. De forma análoga, a forma para determinar a distância, em $\mathbb{R}^4$, de $(-1,0,-1,1)$ à recta definida pelas condições $y=2x+1$, $z=-2x$ e $w=-x+2$ seria escrever esta recta como

$$(0,1,0,2)+L(\{(1,2,-2,-1)\})$$

e calcular a distância de $(-1,0,-1,1)-(0,1,0,2)=(-1,-1,-1,-1)$ à sua projecção sobre $(1,2,-2,-1)$.

$$P_{(1,2,-2,-1)}\left((-1,-1,-1,-1)\right)=\frac{(-1,-1,-1,-1)\cdot(1,2,-2,-1)}{\|(1,2,-2,-1)\|^2}(1,2,-2,-1)=\overrightarrow{0}$$
$$\|(-1,-1,-1,-1)-\overrightarrow{0}\|=\|(-1,-1,-1,-1)\|=2$$

Exercício 48. Calcule a distância do vector $(1,1)$ em cada um dos seguintes subespaços de $\mathbb{R}^2$, munido do produto interno usual.

(a) $L(\{(1,0)\})$ (b) $L(\{(1,1)\})$ (c) $L(\{(0,-1)\})$ (d) $L(\{(1,2)\})$ (e) $L(\{(2,1)\})$

Exercício 49. Calcule a distância do vector $(1,-2)$ a cada uma das seguintes rectas, com o produto interno usual em $\mathbb{R}^2$.

(a) $y=2x$ (b) $y=2x+1$ (c) $y=2x-1$ (d) $2x+3y=0$ (e) $2x+3y=-1$

As seguintes propriedades são consequência das propriedades do produto interno e mostram que a projecção ortogonal funciona bem com as operações do espaço vectorial.

Proposição 43. Sejam V um espaço euclideano e u, v e w vectores de V. Então as seguintes igualdades são satisfeitas.

1. $P_v(\alpha u) = \alpha P_v(u)$
2. $P_{\alpha v}(u) = P_v(u)$
3. $P_v(u+w) = P_v(u) + P_v(w)$
4. $P_v(u) = u$ se $u = \alpha v$ para algum escalar α.

Voltaremos a falar de projecções ortogonais mais adiante, num contexto mais geral, que nos permitirá discutir o problema de determinar a distância de um vector a um subespaço e de determinar a melhor aproximação dum vector num subespaço linear. As técnicas aqui descritas serão aplicadas na prática em disciplinas mais avançadas.

2.6.4 Bases ortogonais e ortonormadas

Vimos na secção anterior que podemos definir o ângulo entre dois vectores em qualquer espaço vectorial que disponha de um produto interno. Nesta secção, vamos estudar as propriedades de bases constituídas apenas por vectores ortogonais — as chamadas bases ortogonais. Estas bases permitem simplificar muitos dos cálculos envolvendo projecções, tornando os espaços euclideanos ainda mais semelhantes a $\mathbb{R}^n$.

Definição. Seja V um espaço euclideano. Um subconjunto S de V diz-se *ortogonal* se todos os seus elementos são ortogonais dois a dois. No caso de todos os elementos de S terem ainda norma igual a 1, o conjunto diz-se *ortonormal* ou *ortonormado*.

Um conjunto S ortogonal (respectivamente ortonormado) que seja uma base de V diz-se uma *base ortogonal* (respectivamente uma *base ortonormada*) de V.

Exemplo. Consideremos o espaço $\mathbb{R}^3$ com o produto interno usual. A base canónica deste espaço é o conjunto $S_1 = \{(1,0,0),(0,1,0),(0,0,1)\}$, que é um conjunto ortonormado. De facto, tem-se

$$\begin{array}{rrr} (1,0,0)\cdot(0,1,0)=0 & \quad (1,0,0)\cdot(0,0,1)=0 & \quad (0,1,0)\cdot(0,0,1)=0 \\ \|(1,0,0)\|=1 & \|(0,1,0)\|=1 & \|(0,0,1)\|=1\,. \end{array}$$

Em contrapartida, o conjunto $S_2 = \{(1,0,0),(0,2,0),(0,0,3)\}$ é um conjunto ortogonal que não é ortonormado.

$$\begin{array}{rrr} (1,0,0)\cdot(0,2,0)=0 & \quad (1,0,0)\cdot(0,0,3)=0 & \quad (0,2,0)\cdot(0,0,3)=0 \\ \|(1,0,0)\|=1 & \|(0,2,0)\|=2 & \|(0,0,3)\|=3\,. \end{array}$$

A base canónica de $\mathbb{R}^n$ (tal como a base canónica de $M_{n\times m}$) é uma base ortonormada. Contudo, o mesmo não se passa noutros espaços: a base canónica de $P_n([a,b])$ não é em geral ortonormada, nem sequer para P_1:

$$\langle 1,x\rangle = \int_a^b x\,dx = \left[\frac{x^2}{2}\right]_a^b = \frac{b^2-a^2}{2}\,,$$

que só vale 0 quando $a = \pm b$. Em particular, a base canónica de $P_1([0,1])$ não é ortonormada. A base canónica de $P_1([-1,1])$ é ortonormada, mas a de $P_1([-2,2])$ é ortogonal não ortonormada. Já a base canónica de $P_2([a,b])$ nunca é ortogonal, pois

$$\left\langle 1, x^2 \right\rangle = \int_a^b x^2\, dx = \left[\frac{x^3}{3}\right]_a^b = \frac{b^3 - a^3}{3},$$

que só vale 0 no caso impossível $a = b$.

Trabalhar com bases ortogonais simplifica muito a determinação de coordenadas. Conforme vimos na Secção 2.4.2, para encontrar as coordenadas de um vector v em relação a uma base B é necessário resolver um sistema de equações lineares. Se B for uma base ortonormada, porém, há uma alternativa muito mais simples.

Concretamente, sejam $B = \{v_1, v_2, \dots, v_n\}$ e $v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$. Se calcularmos o produto interno $\langle v, v_1 \rangle$, obtemos

$$\begin{aligned}
\langle v, v_1 \rangle &= \langle c_1 v_1 + c_2 v_2 + \cdots + c_n v_n, v_1 \rangle \\
&= \langle c_1 v_1, v_1 \rangle + \langle c_2 v_2, v_1 \rangle + \cdots + \langle c_n v_n, v_1 \rangle \\
&= c_1 \langle v_1, v_1 \rangle + c_2 \langle v_2, v_1 \rangle + \cdots + c_n \langle v_n, v_1 \rangle \\
&= c_1 \langle v_1, v_1 \rangle = c_1 \|v_1\|^2
\end{aligned}$$

aplicando sucessivamente a linearidade do produto interno e o facto de B ser ortogonal — ou seja, $\langle v_1, v_i \rangle = 0$ para $i \neq 1$.

Mas $\langle v, v_1 \rangle = c_1 \|v_1\|^2$ equivale a dizer que

$$c_1 = \frac{\langle v, v_1 \rangle}{\|v_1\|^2}.$$

Então a coordenada c_1 pode ser calculada através daquele produto interno. Um raciocínio semelhante permite-nos obter fórmulas análogas para as restantes coordenadas.

No caso particular de B ser uma base ortonormada, aquela expressão simplifica-se ainda mais, uma vez que $\|v_1\| = 1$, e obtém-se simplesmente $c_1 = \langle v, v_1 \rangle$.

Este resultado está resumido na proposição seguinte.

Proposição 44. Sejam V um espaço euclideano com produto interno e $B = \{v_1, v_2, \dots, v_n\}$ uma base ortogonal de V. Então as coordenadas de qualquer vector v de V são dadas por $(v)_B = (c_1, c_2, \dots, c_n)$ onde

$$c_i = \frac{\langle v, v_i \rangle}{\|v_i\|^2}.$$

Se B for ortonormada, as coordenadas de v são simplesmente $c_i = \langle v, v_i \rangle$.

Contrariamente ao que a proposição acima possa sugerir, na prática não há muitas vezes grande vantagem em trabalhar com bases ortonormadas relativamente a bases ortogonais, uma vez que normalizar os vectores introduz muitas vezes coeficientes irracionais. É frequentemente mais simples usar uma base ortogonal, tendo o cuidado de dividir pela sua norma sempre que necessário.

Exemplo. A base $B = \{(1,1), (1,-1)\}$ de $\mathbb{R}^2$ é uma base ortogonal que não é ortonormada, uma vez que

$$(1,1) \cdot (1,-1) = 0 \text{ mas } \|(1,1)\| = \|(1,-1)\| = \sqrt{2}.$$

A base ortonormal correspondente seria

$$B' = \left\{\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)\right\}$$

(obtida dividindo cada vector pela sua norma), base esta que é substancialmente mais complicada para usar em cálculos.

As coordenadas de $(1,0)$ em relação a B são dadas por

$$((1,0))_B = \left(\frac{(1,0)\cdot(1,1)}{\|(1,1)\|^2}, \frac{(1,0)\cdot(1,-1)}{\|(1,-1)\|^2}\right) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

Já as coordenadas de $(1,0)$ em relação a B' são dadas por

$$((1,0))_{B'} = \left((1,0)\cdot\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), (1,0)\cdot\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)\right) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

O recurso a uma base ortogonal simplifica também o cálculo de projecções ortogonais. Sendo V um espaço euclideano e W um seu subespaço linear com base ortogonal B, a projecção de um vector $v \in V$ sobre W é dada pela soma das projecções de v sobre cada um dos elementos de B. Essencialmente, estamos a escrever a projecção de v sobre W em coordenadas de W — algo que se reduz a calcular um produto interno, devido à ortogonalidade de B.

Exemplo. Seja W o plano de $\mathbb{R}^3$ gerado pelos vectores $(1,2,-1)$ e $(1,-1,-1)$. Então W admite a base $\{(1,2,-1),(1,-1,1)\}$, que é ortogonal:

$$(1,2,-1)\cdot(1,-1,-1) = 0\,.$$

Para calcular a distância do ponto $(1,2,3)$ a W, calculamos a sua projecção sobre W somando as projecções sobre cada um dos elementos da base.

$$P_{(1,2,-1)}\left((1,2,3)\right) = \frac{(1,2,3)\cdot(1,2,-1)}{\|(1,2,-1)\|^2}(1,2,-1) = \frac{2}{6}(1,2,-1) = \left(\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right)$$

$$P_{(1,-1,-1)}\left((1,2,3)\right) = \frac{(1,2,3)\cdot(1,-1,-1)}{\|(1,-1,-1)\|^2}(1,-1,-1) = -\frac{4}{3}(1,-1,-1) = \left(-\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right)$$

$$P_W\left((1,2,3)\right) = \left(\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right) + \left(-\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right) = (-1,2,1)$$

$$\|(1,2,3)-(-1,2,1)\| = \|(2,0,2)\| = 2\sqrt{2}$$

A distância de $(1,2,3)$ a W é $\frac{\sqrt{232}}{3} \approx 5.077$.

Exercício 50. Determine a projecção do vector $(1,0,1)$ sobre os seguintes subespaços de $\mathbb{R}^3$.

(a) $L(\{(1,1,1),(1,1,-2)\})$ (b) $L(\{(-1,0,2),(2,2,1)\})$ (c) $L(\{(2,1,0),(0,0,2)\})$

Antes de apresentarmos mais exemplos, há uma questão que se coloca. Neste caso, o espaço W era gerado por dois vectores perpendiculares, pelo que foi fácil encontrar uma base ortogonal para ele; mas como é que se pode em geral encontrar uma base ortogonal para um espaço linear?

Consideremos um conjunto linearmente independente $S = \{u, v\}$ e suponhamos que u e v não são ortogonais. Podemos calcular a projecção de v sobre u, $P_u(v)$, que é um vector paralelo a u; subtraindo este vector a v obtemos um vector v' perpendicular a u (ver Figura 2.14) e é fácil ver que u e v' ainda são linearmente independentes.

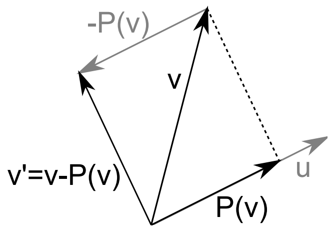


Figura 2.14: Passagem duma base arbitrária a uma base ortogonal (caso bidimensional).

O conjunto $\{u, v'\}$ é ortogonal:

$$\begin{aligned}\langle u, v'\rangle &= \langle u, v - P_u(v)\rangle = \langle u, v\rangle - \langle u, P_u(v)\rangle = \langle u, v\rangle - \left\langle u, \frac{\langle v, u\rangle}{\|u\|^2} u\right\rangle \\ &= \langle u, v\rangle - \frac{\langle v, u\rangle}{\|u\|^2} \underbrace{\langle u, u\rangle}_{\|u\|^2} = \langle u, v\rangle - \langle u, v\rangle = 0\end{aligned}$$

Para espaços de dimensão superior, podemos iterar este método, obtendo uma base ortogonal. Mais: nem precisamos de começar com uma base do espaço — os vectores linearmente dependentes coincidem com a sua projecção ortogonal sobre o espaço gerado pelos restantes, pelo que quando lhes retiramos esta projecção se transformam no vector nulo.

Definição. Seja V um espaço euclideano e $S = \{v_1, v_2, v_3, \ldots, v_n\}$ um subconjunto de V. O *método de ortogonalização de Gram–Schmidt* consiste em calcular a seguinte sequência de vectores $v'_1, v'_2, v'_3, \ldots, v'_n$.

$$\begin{aligned}v'_1 &= v_1 \\ v'_2 &= v_2 - P_{v'_1}(v_2) \\ v'_3 &= v_3 - P_{v'_1}(v_3) - P_{v'_2}(v_3) \\ &\vdots \\ v'_n &= v_n - P_{v'_1}(v_n) - P_{v'_2}(v_n) - P_{v'_2}(v_n) - \cdots - P_{v'_{n-1}}(v_n)\end{aligned}$$

O conjunto S' constituído pelos vectores não nulos de entre $v'_1, v'_2, v'_3, \ldots, v'_n$ é uma base ortogonal para W.

Uma consequência imediata deste método é o seguinte resultado.

Proposição 45. Todo o espaço euclideano de dimensão finita tem uma base ortonormada.

Exemplo. Consideremos o espaço vectorial $\mathbb{R}^3$ com o produto interno usual e o seu subespaço linear $W = L(\{(1,0,1),(-1,1,0)\})$. Para determinar uma base ortogonal para W, vamos aplicar o método de ortogonalização de Gram–Schmidt.

$$\begin{aligned}
v_1 &= (1,0,1)\\
v_2 &= (-1,1,0) - P_{(1,0,1)}\left((-1,1,0)\right)\\
&= (-1,1,0) - \frac{(-1,1,0)\cdot(1,0,1)}{\|(1,0,1)\|^2}(1,0,1) = (-1,1,0) - \frac{-1}{2}(1,0,1)\\
&= (-1,1,0) - \left(-\frac{1}{2},0,-\frac{1}{2}\right) = \left(-\frac{1}{2},1,\frac{1}{2}\right)
\end{aligned}$$

O conjunto $B = \left\{(1,0,1),\left(-\frac{1}{2},1,\frac{1}{2}\right)\right\}$ é uma base ortogonal para W. Para determinar uma base ortonormada para este conjunto, temos que transformar cada um dos vectores num vector unitário. Para tal, vamos dividir cada vector pela sua norma.

$$\begin{aligned}
w_1 &= \frac{v_1}{\|v_1\|} = \frac{(1,0,1)}{\|(1,0,1)\|} = \left(\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right)\\
w_2 &= \frac{v_2}{\|v_2\|} = \frac{\left(-\frac{1}{2},1,\frac{1}{2}\right)}{\left\|\left(-\frac{1}{2},1,\frac{1}{2}\right)\right\|} = \left(-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}}\right)
\end{aligned}$$

A base $B' = \{w_1, w_2\}$ é uma base ortonormada para S.

Para determinar a distância do ponto $(1,2,1)$ a W, podemos usar as bases B ou B'. Em geral, é mais simples usar a base ortogonal B.

$$P_{(1,0,1)}\left((1,2,1)\right) = \frac{(1,2,1)\cdot(1,0,1)}{\|(1,0,1)\|^2}(1,0,1) = \frac{2}{2}(1,0,1) = (1,0,1)$$

$$P_{\left(-\frac{1}{2},1,\frac{1}{2}\right)}\left((1,2,1)\right) = \frac{(1,2,1)\cdot\left(-\frac{1}{2},1,\frac{1}{2}\right)}{\left\|\left(-\frac{1}{2},1,\frac{1}{2}\right)\right\|^2}\left(-\frac{1}{2},1,\frac{1}{2}\right) = \frac{2}{\frac{6}{4}}\left(-\frac{1}{2},1,\frac{1}{2}\right) = \left(-\frac{2}{3},\frac{4}{3},\frac{2}{3}\right)$$

$$P_W\left((1,2,1)\right) = (1,0,1) + \left(-\frac{2}{3},\frac{4}{3},\frac{2}{3}\right) = \left(\frac{1}{3},\frac{4}{3},\frac{5}{3}\right)$$

$$\|(1,2,1) - P_W\left((1,2,1)\right)\| = \left\|(1,2,1) - \left(\frac{1}{3},\frac{4}{3},\frac{5}{3}\right)\right\| = \left\|\left(\frac{2}{3},\frac{2}{3},-\frac{4}{3}\right)\right\| = \frac{2\sqrt{6}}{3}$$

A distância de $(1,2,1)$ a W é $\frac{2\sqrt{3}}{6} \approx 1.155$.

Exercício 51. Encontre bases ortonormadas para cada um dos seguintes subespaços de $\mathbb{R}^3$, considerando o produto interno usual.

(a) $L(\{(1,-1,0),(0,0,-1)\})$ (b) $L(\{(-1,-1,-1),(0,0,0)\})$ (c) $L(\{(1,3,1),(2,2,2)\})$

2.6.5 Produto externo em $\mathbb{R}^3$

Vamos agora discutir brevemente uma operação definida no espaço $\mathbb{R}^3$, que é muito usada em Física pelas suas propriedades geométricas.

Em $\mathbb{R}^3$, precisamos muitas vezes de nos referir a rectas e planos. Pela sua natureza e pelo contexto em que surgem, estes são muitas vezes fáceis de caracterizar em termos de vectores perpendiculares. No que diz respeito a planos, em particular, há várias formas de os definir: ou pela sua equação analítica, que é simplesmente uma equação linear em $\mathbb{R}^3$; ou por um ponto e um vector perpendicular ao plano; ou por três pontos não colineares; ou por duas rectas paralelas contidas nesse plano; ou por duas rectas concorrentes contidas nesse plano.

A forma mais simples de trabalhar com planos é, claro está, a partir da equação que os define. Por sua vez, a forma mais expedita de a obter é a partir dum ponto do plano e um vector que lhe é perpendicular. Se $P_0 = (x_0, y_0, z_0)$ for um ponto do plano e $\overrightarrow{n} = (A, B, C)$ for um vector perpendicular ao plano, então para um ponto genérico $P = (x, y, z)$ do plano o vector $\overrightarrow{PP_0}$ pertence ao plano e, portanto, é perpendicular a $\overrightarrow{n}$. Então

$$\overrightarrow{PP_0} \cdot \overrightarrow{n} = 0\,.$$

Expandindo a equação acima, obtemos $((x, y, z) - (x_0, y_0, z_0)) \cdot (A, B, C) = 0$, que se reduz à equação geral do plano

$$Ax + By + Cz = Ax_0 + By_0 + Cz_0\,.$$

Exemplo. Para determinar a equação do plano que passa na origem e é perpendicular ao vector $(1, 1, 1)$, basta expandir a equação

$$((x, y, z) - (0, 0, 0)) \cdot (1, 1, 1) = 0$$

e obtém-se

$$x + y + z = 0\,.$$

O problema torna-se substancialmente mais complicado quando não conhecemos um vector perpendicular ao plano. Por exemplo, como é que se pode determinar a equação do plano que contém o ponto $(1, 0, 1)$ e os vectores $\overrightarrow{u} = (1, 1, 2)$ e $\overrightarrow{v} = (0, 1, 1)$? A partir dos vectores $\overrightarrow{u}$ e $\overrightarrow{v}$, podemos procurar um vector (x, y, z) que seja perpendicular aos dois, o que se resume a resolver o sistema

$$\begin{cases} (x,y,z)\cdot(1,1,2) = 0 \\ (x,y,z)\cdot(0,1,1) = 0 \end{cases} \longrightarrow \begin{cases} x+y+2z = 0 \\ y+z = 0 \end{cases} \longrightarrow \begin{cases} x+y+2z = 0 \\ z = -y \end{cases}$$
$$\longrightarrow \begin{cases} x-y = 0 \\ z = -y \end{cases} \longrightarrow \begin{cases} x = y \\ z = -y \end{cases}$$

Assim, um vector perpendicular ao plano é, por exemplo, o vector $(1, 1, -1)$. A partir deste, pode-se determinar a equação do plano em causa como no exemplo acima, obtendo-se agora $x + y - z = 0$.

Um processo expedito para calcular vectores perpendiculares a outros dois vectores de $\mathbb{R}^3$ consiste em usar o produto externo de vectores, que se baseia em determinantes de matrizes de dimensão 3.

Definição. Sejam $\overrightarrow{u} = (u_1, u_2, u_3)$ e $\overrightarrow{v} = (v_1, v_2, v_3)$ dois vectores de $\mathbb{R}^3$. O *produto externo* de $\overrightarrow{u}$ por $\overrightarrow{v}$, denotado por $\overrightarrow{u} \times \overrightarrow{v}$, é o vector

$$\overrightarrow{u} \times \overrightarrow{v} = \det \begin{bmatrix} (1,0,0) & (0,1,0) & (0,0,1) \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix}$$

A matriz na definição do produto externo é um exemplo duma matriz que não contém apenas números reais: a primeira linha da matriz contém vectores. Apesar disso, o determinante calcula-se da maneira habitual.

No exemplo anterior, queríamos determinar um vector perpendicular a $(1, 1, 2)$ e $(0, 1, 1)$. Calculando o produto externo destes dois vectores, obtemos

$$\overrightarrow{n} = (1,1,2) \times (0,1,1) = \det \begin{bmatrix} (1,0,0) & (0,1,0) & (0,0,1) \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}.$$

Aplicando a regra de Laplace à primeira linha, ficamos com

$$\begin{aligned} \overrightarrow{n} &= (1,0,0) \underbrace{\det \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}}_{-1} - (0,1,0) \underbrace{\det \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}}_{1} + (0,0,1) \underbrace{\det \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{1} \\ &= -(1,0,0) - (0,1,0) + (0,0,1) = (-1,-1,1) \end{aligned}$$

que é o simétrico do vector determinado anteriormente.

Exemplo. Vamos determinar um vector perpendicular a $(1, 2, -2)$ e a $(3, 0, 1)$. Usando o produto externo, temos

$$\begin{aligned} (1,2,-2) \times (3,0,1) &= \det \begin{bmatrix} (1,0,0) & (0,1,0) & (0,0,1) \\ 1 & 2 & -2 \\ 3 & 0 & 1 \end{bmatrix} \\ &= (1,0,0) \det \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix} - (0,1,0) \det \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} + (0,0,1) \det \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \\ &= 2(1,0,0) - 7(0,1,0) - 6(0,0,1) = (2,-7,-6) \end{aligned}$$

Podemos verificar que de facto

$$(1,2,-2) \cdot (2,-7,-6) = 2 - 14 - 12 = 0 \text{ e } (3,0,1) \cdot (2,-7,-6) = 6 + 0 - 6 = 0\,;$$

obtivemos portanto um vector perpendicular a $(1, 2, -2)$ e a $(3, 0, 1)$.

Exercício 52. Calcule o produto externo dos seguintes pares de vectores de $\mathbb{R}^3$.

(a) $(1, 2, 3)$ e $(3, 2, 1)$ (b) $(1, -1, 1)$ e $(2, 1, 1)$ (c) $(-2, 4, 1)$ e $(0, 0, 1)$

Verifique que em todos os casos os vectores obtidos são sempre perpendiculares aos dois vectores originais.

Também podemos determinar uma expressão explícita para o produto externo. Dados vectores $\overrightarrow{u} = (u_1, u_2, u_3)$ e $\overrightarrow{v} = (v_1, v_2, v_3)$, o seu produto externo $\overrightarrow{u} \times \overrightarrow{v}$ é

$$\begin{aligned} \overrightarrow{n} = \overrightarrow{u} \times \overrightarrow{v} &= \det \begin{bmatrix} (1,0,0) & (0,1,0) & (0,0,1) \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix} \\ &= (u_2v_3 - v_2u_3, v_1u_3 - u_1v_3, u_1v_2 - v_1u_2)\,. \end{aligned}$$

Vamos ver que, de facto, $\overrightarrow{n}$ é perpendicular a $\overrightarrow{u}$:

$$\begin{aligned} \overrightarrow{n} \cdot \overrightarrow{u} &= (u_2v_3 - v_2u_3, v_1u_3 - u_1v_3, u_1v_2 - v_1u_2) \cdot (u_1, u_2, u_3) \\ &= u_1u_2v_3 - u_1v_2u_3 + u_2v_1u_3 - u_2u_1v_3 + u_3u_1v_2 - u_3v_1u_2 = 0 \end{aligned}$$

Prova-se de maneira análoga que $\overrightarrow{n}$ é perpendicular a $\overrightarrow{v}$. Temos assim o seguinte resultado.

Proposição 46. Sejam $\overrightarrow{u}$ e $\overrightarrow{v}$ vectores de $\mathbb{R}^3$. Então o vector $\overrightarrow{u} \times \overrightarrow{v}$ é perpendicular a $\overrightarrow{u}$ e a $\overrightarrow{v}$.

Não vamos estudar detalhadamente produtos externos; o interesse de os referir deveu-se a serem uma boa ferramenta para determinar vectores perpendiculares a dois outros. Resumimos de seguida as propriedades mais importantes desta operação, que podem ser verificadas por cálculo directo.

Proposição 47. Sejam $\overrightarrow{u}$, $\overrightarrow{v}$ e $\overrightarrow{w}$ vectores em $\mathbb{R}^3$ e $\alpha \in \mathbb{R}$. Então verificam-se as seguintes igualdades.

1. $\overrightarrow{u} \times \overrightarrow{v} = -\overrightarrow{v} \times \overrightarrow{u}$
2. $\overrightarrow{u} \times \overrightarrow{u} = \overrightarrow{0}$
3. $\overrightarrow{u} \times (\overrightarrow{v} + \overrightarrow{w}) = \overrightarrow{u} \times \overrightarrow{v} + \overrightarrow{u} \times \overrightarrow{w}$
4. $\alpha(\overrightarrow{u} \times \overrightarrow{v}) = (\alpha\overrightarrow{u}) \times \overrightarrow{v} = \overrightarrow{u} \times (\alpha\overrightarrow{v})$
5. $\overrightarrow{u} \times \overrightarrow{0} = \overrightarrow{0}$

Por cálculo directo, prova-se também que o produto externo entre dois vectores $\overrightarrow{u}$ e $\overrightarrow{v}$ de $\mathbb{R}^3$ se relaciona com a norma de ambos, de maneira semelhante ao que acontece com o produto interno.

Proposição 48. Sejam $\overrightarrow{u}$ e $\overrightarrow{v}$ vectores não nulos de $\mathbb{R}^3$ e θ o ângulo formado por ambos. Então

$$\|\overrightarrow{u} \times \overrightarrow{v}\| = \|\overrightarrow{u}\|\|\overrightarrow{v}\| \sin(\theta)\,.$$

Como consequência deste resultado, a área do paralelogramo de lados $\overrightarrow{u}$ e $\overrightarrow{v}$ é dada por $\overrightarrow{u} \times \overrightarrow{v}$. De facto, considerando o paralelogramo de lados $\overrightarrow{u}$ e $\overrightarrow{v}$ e escolhendo para base o vector $\overrightarrow{u}$, a sua altura é dada por $\overrightarrow{v}\sin(\theta)$, donde a sua área é dada pela expressão acima.

Em particular, se $\overrightarrow{u}$ e $\overrightarrow{v}$ forem vectores paralelos, a área do paralelogramo definido por eles é zero, pelo que o produto externo de vectores paralelos é zero.

2.6.6 Aplicações

Para terminar, vamos apresentar alguns exemplos de aplicações de espaços vectoriais com produto interno.

Uma das aplicações mais óbvias é à Geometria Analítica. A noção de produto interno em $\mathbb{R}^n$, juntamente com os conceitos derivados que discutimos nesta secção, permite-nos calcular ângulos, projecções e distâncias de forma relativamente simples, conforme exemplificámos nas secções correspondentes.

Outra grande utilidade dos produtos internos é em problemas de aproximação. Aqui não estamos tão interessados em calcular distâncias e ângulos, mas sim em verificar ortogonalidade e calcular projecções, uma vez que (como acima discutimos) a projecção dum vector v num subespaço linear W é a melhor aproximação de v por um vector em W.

Exemplo. Uma matriz quadrada A diz-se *simétrica* se $A = A^T$; dito doutra forma, A é simétrica se $a_{ij} = a_{ji}$ para quaisquer i, j: a matriz é invariante por reflexão sobre a diagonal principal.

Exemplos de matrizes simétricas são

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

enquanto as matrizes

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

não são matrizes simétricas.

O espaço $S_{2\times 2}$ das matrizes simétricas de dimensão 2×2 é um subespaço vectorial de $M_{2\times 2}$, conforme pode ser facilmente verificado. Uma base para este subespaço é o conjunto

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

que é uma base ortogonal para $S_{2\times 2}$ (embora não seja ortonormada, já que $\left\| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\| = \sqrt{2}$).

Seja X uma matriz quadrada de dimensão 2 qualquer. A matriz simétrica X_s que melhor aproxima X é a projecção de X sobre $S_{2\times 2}$. Podemos calcular esta projecção para alguns casos particulares.

$$\begin{aligned} X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} : X_s &= \left\langle \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &+ \left\langle \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{5}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{5}{2} \\ \frac{5}{2} & 4 \end{bmatrix} \end{aligned}$$

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} : X_s = \left\langle \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$+ \left\langle \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & -1 \end{bmatrix}$$

$$X = \begin{bmatrix} 2 & 1 \\ -3 & 5 \end{bmatrix} : X_s = \left\langle \begin{bmatrix} 2 & 1 \\ -3 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 2 & 1 \\ -3 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$+ \left\langle \begin{bmatrix} 2 & 1 \\ -3 & 5 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= 2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 5 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}$$

Se tomarmos uma matriz genérica $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, verificamos que

$$X_s = \left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\rangle \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$+ \left\langle \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\rangle \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{b+c}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{bmatrix}$$

Caso não tivéssemos obtido imediatamente uma base ortogonal de $S_{2\times 2}$, poderíamos ter começado com uma base qualquer e aplicado o método de Gram–Schmidt.

Exercício 53. Seja $W = L\left(\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\}\right)$ o espaço das matrizes triangulares superiores de dimensão 2×2. Calcule a melhor aproximação em W de cada uma das seguintes matrizes.

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

Exercício 54. Calcule a melhor aproximação do vector $(1, 2, 1, 1, 0)$ em cada um dos seguintes subespaços lineares de $\mathbb{R}^5$.

(a) $L(\{(1, 0, 1, 0, 0), (0, 0, 1, 0, 0)\})$

(b) $L(\{(1, 2, 3, 0, 0), (0, 0, 0, 0, 1)\})$

(c) $L(\{(1, 0, 0, 0, 0), (1, 1, 0, 0, 0), (1, 1, 1, 0, 0)\})$

(d) $L(\{(1, 2, 0, 0, 0), (0, 0, 0, 1, 1), (0, 0, 0, 0, 1)\})$

Outros exemplos importantes são: a aproximação de funções por polinómios; a teoria das séries de Fourier; e o método dos mínimos quadrados.

2.7 Exercícios e problemas

Espaços e subespaços vectoriais

55. Indique, justificando, se os seguintes subconjuntos de $M_{2\times 2}$ são espaços vectoriais.
    (a) O conjunto das matrizes A que satisfazem $A = A^{-1}$.
    (b) O conjunto das matrizes A que satisfazem $\operatorname{tr}(A) \geq 0$.

56. Indique, justificando, se os seguintes subconjuntos de $M_{3\times 3}$ são espaços vectoriais.
    (a) O conjunto das matrizes invertíveis.
    (b) O conjunto das matrizes A que satisfazem $\operatorname{tr}(A) = 2$.

57. Indique, justificando, se os seguintes subconjuntos de $M_{4\times 4}$ são espaços vectoriais.
    (a) O conjunto das matrizes A que satisfazem $A = A^T$.
    (b) O conjunto das matrizes A que satisfazem $\det(A) = 0$.

58. Determine quais das seguintes estruturas são espaços vectoriais reais.
    (a) $\mathbb{R}^2$ com a soma usual e o produto definido por $\alpha(x, y) = (2\alpha x, 2\alpha y)$.
    (b) $\mathbb{R}^+$ com a soma definida por $x + y = xy$ e $\alpha x = x^\alpha$.
    (c) $\mathbb{R}^-$ com a soma definida por $x + y = xy$ e $\alpha x = x^\alpha$.
    (d) $M_{2\times 2}$ com a soma definida por $A + B = AB$ e o produto usual.

59. Indique quais dos seguintes conjuntos são subespaços vectoriais de $\mathbb{R}^n$.
    (a) $\{(x, y) \in \mathbb{R}^2 \mid x = 0\}$
    (b) $\{(x, y) \in \mathbb{R}^2 \mid x \geq 0\}$
    (c) $\{(x, y) \in \mathbb{R}^2 \mid x + y = 2\}$
    (d) $\{(x, y) \in \mathbb{R}^2 \mid x + y \leq 2\}$
    (e) $\{(x, y) \in \mathbb{R}^2 \mid x + y \neq 2\}$
    (f) $\{(x, y) \in \mathbb{R}^2 \mid x \times y = 0\}$
    (g) $\{(x, y) \in \mathbb{R}^2 \mid x \times y = 1\}$
    (h) $\{(x, y) \in \mathbb{R}^2 \mid x \times y \geq 0\}$
    (i) $\{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 3z = 0\}$
    (j) $\{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 3z = 2\}$
    (k) $\{(x, y, z) \in \mathbb{R}^3 \mid y = x + z\}$
    (l) $\{(x, y, z) \in \mathbb{R}^3 \mid y = x + z + 1\}$
    (m) $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + 2y + 3z = 0\}$
    (n) $\{(x, y, z) \in \mathbb{R}^3 \mid x = 0 \text{ ou } y - z = 0\}$
    (o) $\{(x, y, z, w) \in \mathbb{R}^4 \mid x \neq 1\}$
    (p) $\{(x, y, z, w) \in \mathbb{R}^4 \mid x = 0\}$
    (q) $\{(x, y, z, w) \in \mathbb{R}^4 \mid xyzw = 0\}$
    (r) $\{(x, y, z, w) \in \mathbb{R}^4 \mid x + y = 0 \text{ ou } z + w \neq 2\}$

60. Mostre que os seguintes conjuntos são espaços vectoriais de $\mathbb{R}^n$.
    (a) As soluções dum sistema homogéneo $Ax = 0$, com A uma matriz não singular.
    (b) As soluções dum sistema homogéneo $Ax = 0$, com A uma matriz 2×3.

61. Mostre que os seguintes conjuntos não são espaços vectoriais de $\mathbb{R}^n$.

(a) Uma parábola de $\mathbb{R}^2$ que passe pela origem.

(b) As soluções dum sistema não homogéneo $Ax = b$, com A uma matriz não singular.

(c) As soluções dum sistema não homogéneo $Ax = b$, com A uma matriz 2×3.

(d) O conjunto dos valores de a para os quais a matriz $\begin{bmatrix} a & 1 \\ 1 & a \end{bmatrix}$ é não-singular.

62. Determine quais dos seguintes espaços de matrizes são subespaços de $M_{n\times m}(\mathbb{R})$.

(a) $\{A_{3\times 3} \mid A^T = -A\}$

(b) $\{A_{3\times 3} \mid \det(A) = 1\}$

(c) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbb{Z}\right\}$

(d) $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a + b + c + d = 0\right\}$

(e) $\left\{\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \mid a, b, d \in \mathbb{R}\right\}$

(f) $\{A_{3\times 3} \mid A$ é uma matriz em escada de linhas$\}$

63. Determine quais dos seguintes conjuntos de polinómios são subespaços vectoriais de P_n.

(a) $\{ax + b \mid a \times b < 0\}$

(b) $\{ax^2 + bx + c \mid b^2 - ac = 0\}$

(c) $\{ax^3 + bx^2 + cx + d \mid a + b + c + d = 0\}$

(d) $\{ax^3 + d \mid a, d \in \mathbb{R}\}$

(e) $\{ax^n + b \mid a, b \in \mathbb{R}\}$

(f) $\{p(x) \in P_n \mid p(2) = 0\}$

(g) $\{p(x) \in P_n \mid p(-x) = p(x)\}$

64. Determine quais dos seguintes conjuntos de funções são subespaços vectoriais de $\mathcal{F}(\mathbb{R})$.

(a) $\{f \in \mathcal{F}(\mathbb{R}) \mid f(0) = 2\}$

(b) $\{f \in \mathcal{F}(\mathbb{R}) \mid f(2) = 0\}$

(c) $\{a\sin(x) + b\cos(x) \mid a, b \in \mathbb{R}\}$

(d) $\{f \in \mathcal{F}(\mathbb{R}) \mid f$ é contínua$\}$

(e) $\{f \in \mathcal{F}(\mathbb{R}) \mid f''$ é constante$\}$

(f) $\{f \in \mathcal{F}(\mathbb{R}) \mid f''(-1) = 0\}$

(g) $\{f \in \mathcal{F}(\mathbb{R}) \mid f(x) + 2f'(x) = 0\}$

(h) $\{f \in \mathcal{F}(\mathbb{R}) \mid xf'(x) + f(x) = 2f''(x)\}$

Representação de vectores

65. Escreva os seguintes vectores como combinação linear de elementos de

$$S = \{(2, 1, 4), (1, -1, 3), (3, 2, 5)\}.$$

(a) $(-9, -7, -15)$

(b) $(6, 11, 6)$

(c) $(0, 0, 0)$

(d) $(7, 8, 9)$

(e) $(6, 2, 12)$

(f) $(2, 1, 4)$

(g) $(4, -1, 10)$

(h) $(5, 3, 9)$

(i) $(1, 2, 1)$

(j) $(6, 12, 6)$

(k) $(1, 0, 0)$

(l) $(0, 0, 1)$

66. Considere o espaço P_2 dos polinómios de grau menor ou igual a 2.

(a) Calcule os valores de k para os quais o polinómio $2x^2 + x + k$ é combinação linear de $p(x) = x^2 + 2x - 1$ e $q(x) = x^2 - x + 3$.

(b) Calcule $3p(x) - 2q(x)$.

67. Seja S o conjunto

$$S = \left\{ \begin{bmatrix} 4 & 0 \\ -2 & -2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 4 \end{bmatrix} \right\}.$$

Quais dos seguintes vectores são combinação linear de elementos de S?

(a) $\begin{bmatrix} 6 & -8 \\ -1 & -8 \end{bmatrix}$ (c) $\begin{bmatrix} 6 & 0 \\ 3 & 8 \end{bmatrix}$ (e) $\begin{bmatrix} 7 & 7 \\ -3 & 5 \end{bmatrix}$ (g) $\begin{bmatrix} -3 & 3 \\ 6 & 13 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & 5 \\ 7 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix}$ (h) $\begin{bmatrix} 10 & 6 \\ -8 & -4 \end{bmatrix}$

68. Determine se os seguintes subconjuntos de $\mathbb{R}^2$ são linearmente independentes.

(a) $\{(2,3),(2,5)\}$ (c) $\{(1,-1),(0,1)\}$ (e) $\{(1,2),(0,1),(2,1)\}$

(b) $\{(1,2),(-2,-4)\}$ (d) $\{(a,2),(-1,a)\}$ (f) $\{(2,-1),(2,3),(1,5)\}$

69. Determine se os seguintes subconjuntos de $\mathbb{R}^3$ são linearmente independentes.

(a) $\{(2,3,1),(1,2,0)\}$ (e) $\{(-2,1,2),(2,2,2),(0,3,0)\}$

(b) $\{(1,0,1),(0,1,0)\}$ (f) $\{(2,2,1),(-1,1,0),(0,4,1)\}$

(c) $\{(-1,-1,-1),(-1,1,1),(-1,-1,1)\}$ (g) $\{(2,1,1),(1,4,2),(-1,2,-1)\}$

(d) $\{(1,0,2),(0,0,0)\}$ (h) $\{(2,1,1),(-2,1,1),(1,0,0)\}$

70. Determine se os seguintes subconjuntos de P_n são linearmente independentes.

(a) $\{x^3,x^2,1\}$ (g) $\{2+4x^2,2+5x,2+10x-4x^2\}$

(b) $\{x^2+x,x^3+x\}$ (h) $\{3+x+x^2,x+5x^2,4-3x^2\}$

(c) $\{3-2x+x^2,6-4x+2x^2\}$ (i) $\{6-x^2,1+x+4x^2\}$

(d) $\{x^2-1,x+1,x-1\}$ (j) $\{2x+3,x-2,3x+1\}$

(e) $\{x^3-x,x^2-x,x^3-x^2\}$ (k) $\{1,x,x+x^2\}$

(f) $\{1,1+x,1+x+x^2\}$ (l) $\{2x+x^2+x^3,3x^2-2x,1+3x^3\}$

71. Determine se os seguintes subconjuntos de $\mathcal{F}(\mathbb{R})$ são linearmente independentes.

(a) $\{x,x^2,e^x\}$ (c) $\{e^x,e^{2x},e^{3x}\}$ (e) $\{\sin^2(x),\cos^2(x),2\}$

(b) $\{\cos(x),x^2+1,e^{2x}\}$ (d) $\{\sin(x),\cos(x),\sin(2x)\}$ (f) $\{\log(x),2x\}$

72. Indique quais dos conjuntos dos Exercícios 68, 69 e 70 são bases de $\mathbb{R}^2$, $\mathbb{R}^3$, P_2 e P_3.

73. Determine as coordenadas do vector $(3,2,2)$ em relação às seguintes bases de $\mathbb{R}^3$.

(a) $\{(1,0,0),(0,1,0),(0,0,1)\}$ (d) $\{(1,-1,1),(2,1,0),(0,0,1)\}$

(b) $\{(2,0,0),(0,1,0),(0,0,-1)\}$ (e) $\{(1,2,3),(4,5,6),(1,2,1)\}$

(c) $\{(1,0,1),(1,1,0),(0,1,1)\}$ (f) $\{(1,0,0),(1,1,0),(1,1,1)\}$

74. Determine as coordenadas do polinómio $x^2 - 1$ em relação às seguintes bases de P_2.

(a) $\{1, x, x^2\}$
(b) $\{x^2, 2x, 3\}$
(c) $\{1, 1 + x, 1 + x + x^2\}$
(d) $\{1, 1 + x, 1 + x^2\}$
(e) $\{x + 1, x - 1, x^2\}$
(f) $\{x^2 - 1, x^2 + 1, 2x\}$
(g) $\{1 + x, x + x^2, x^2 + 1\}$
(h) $\{x^2 - x - 1, 2x^2 + x, x + 2\}$

75. Indique a dimensão e a base de cada um dos seguintes subespaços de $\mathbb{R}^3$.

(a) O plano $3x - 2y + 5z = 0$.
(b) O plano $x - y = 0$.
(c) A recta $(2t, -t, 4t)$ com $t \in \mathbb{R}$.
(d) Os vectores (x, y, z) tais que $y = x+z$.
(e) Os vectores (x, y, z) com $2x - y = 3z$.
(f) Os vectores $(a, 2b, b - a)$ com $a, b \in \mathbb{R}$.

76. Indique a dimensão e a base de cada um dos seguintes subespaços de $\mathbb{R}^4$.

(a) Os vectores da forma $(x, y, z, 0)$ com $x, y, z \in \mathbb{R}$.
(b) Os vectores da forma (x, y, z, w) com $z = x + y$ e $w = x - y$.
(c) Os vectores da forma (x, y, z, w) com $x = y = z = w$.
(d) Os vectores da forma (x, y, z, w) com $x = y$ e $z = w$.
(e) Os vectores da forma (x, y, z, w) com $3x - 2y = z + w$.
(f) Os vectores da forma $(x, 0, z, w)$ com $x + 2z + w = 0$.

77. Indique a dimensão e a base de cada um dos seguintes espaços de polinómios.

(a) $\{p(x) \in P_2 \mid p(0) = 0\}$
(b) $\{p(x) \in P_3 \mid p(0) = 0\}$
(c) $\{p(x) \in P_2 \mid p(0) = p(1) = 0\}$
(d) $\{p(x) \in P_3 \mid p(0) = p(1) = 0\}$
(e) Polinómios de grau 2 com derivada 0 em $x = 0$.
(f) Polinómios de grau 3 sem termo em x^2.
(g) Polinómios de grau 3 que são funções ímpares.

78. Indique a dimensão e a base de cada um dos seguintes espaços de funções.

(a) Todas as funções da forma $a \sin(x) + b \cos(x)$ com $a, b \in \mathbb{R}$.
(b) Todos os polinómios de grau menor ou igual a 4.
(c) Todos os polinómios de grau menor ou igual a n.
(d) Todos os polinómios de grau menor ou igual a 4 tais que $p(x) = p(-x)$.
(e) Todos os polinómios de grau menor ou igual a $2n$ tais que $p(x) = p(-x)$.
(f) Todas as funções da forma $a_1e^x + a_2e^{2x} + a_3e^{3x}$, $a_i \in \mathbb{R}$.
(g) Todas as funções da forma $a_1e^x + a_2e^{2x} + \cdots + a_ne^{nx}$ com n fixo, $a_i \in \mathbb{R}$.
(h) Todas as funções da forma $\log{(ax^2)} + \log(bx) + \log(c)$ com $a, b, c \in \mathbb{R}$.

Espaços associados a uma matriz

79. Para cada um dos seguintes sistemas de equações lineares, determine a dimensão e uma base do espaço das suas soluções.

(a) $\begin{cases} x+y-z=0 \\ -2x-y+2z=0 \\ -x+z=0 \end{cases}$

(b) $\begin{cases} x-3y+z=0 \\ 2x-6y+2z=0 \\ 3x-9y+3z=0 \end{cases}$

(c) $\begin{cases} 2x+y+3z=0 \\ x+5z=0 \\ y+z=0 \end{cases}$

(d) $\begin{cases} 3x+y+z+w=0 \\ 5x-y+z-w=0 \end{cases}$

(e) $\begin{cases} x-4y+3z-w=0 \\ 2x-8y+6z-2w=0 \end{cases}$

(f) $\begin{cases} x+y+z=0 \\ 3x+2y-2z=0 \\ 4x+3y-z=0 \\ 6x+5y+z=0 \end{cases}$

80. Para cada uma das matrizes do exercício anterior, indique uma base para o espaço das linhas constituída unicamente por vectores das linhas das respectivas matrizes.

81. Encontre matrizes 2×2 e 3×3 cujo núcleo seja:

(a) um ponto; (b) uma recta; (c) todo o espaço.

82. Encontre uma matriz 3×3 cujo núcleo seja um plano.

83. Considere matrizes com as seguintes características.

Matriz	A	B	C	D	E	F	G	H
Dimensão	3×3	3×3	3×3	5×9	9×5	4×4	6×2	4×4
Característica	3	2	1	2	2	0	2	2

Para cada uma destas matrizes, determine as dimensões do seu espaço das linhas, do seu espaço das colunas e do seu núcleo.

84. Indique a característica máxima e a nulidade mínima duma matriz de dimensão:

(a) 4×4; (b) 3×5; (c) 5×3; (d) $m \times n$.

85. Indique os valores de α para os quais o sistema de equações lineares

$$\begin{cases} x+y+\alpha z=0 \\ x+\alpha y+z=0 \\ \alpha x+y+z=0 \end{cases}$$

tem como soluções:

(a) a origem; (b) uma recta; (c) um plano; (d) todo o espaço.

86. Determine a característica das seguintes matrizes em termos dos parâmetros α e β.

(a) $\begin{bmatrix} 1 & 1 & \alpha \\ 1 & \alpha & 1 \\ \alpha & 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} \alpha & 3 & -1 \\ 3 & 6 & -2 \\ -1 & -3 & \alpha \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha-2 & 2 \\ 0 & \beta-1 & \beta+2 \\ 0 & 0 & 3 \end{bmatrix}$

Espaços euclideanos

87. Calcule o produto interno dos seguintes pares de vectores, para o produto interno usual em $\mathbb{R}^2$ e $\mathbb{R}^3$.

    (a) $(1,2)$ e $(-1,2)$
    (b) $(1,1)$ e $(-1,0)$
    (c) $(2,1)$ e $(-2,4)$
    (d) $(-2,2,0)$ e $(-1,1,1)$
    (e) $(1,1,1)$ e $(-1,3,-2)$
    (f) $(2,1,-1)$ e $(-1,2,5)$

88. Calcule o ângulo formado pelos dois vectores de cada alínea do Exercício 87. Quais desses pares são perpendiculares?

89. Para cada um dos seguintes pares de vectores, determine os valores de k que os tornam perpendiculares com o produto interno usual em $\mathbb{R}^2$ e $\mathbb{R}^3$.

    (a) $(2,1)$ e $(-1,k)$
    (b) $(1,0)$ e $(0,k)$
    (c) (k,k) e $(-2,1)$
    (d) $(-2,2,0)$ e $(-1,k,k)$
    (e) $(2,k,-1)$ e $(1,1,0)$
    (f) $(1,2,1)$ e $(-1,k,k)$

90. Calcule a norma (usual) de cada um dos seguintes vectores.

    (a) $(1,2)$
    (b) $(-1,2)$
    (c) $(1,1)$
    (d) $(2,4)$
    (e) $(-1,0)$
    (f) $(1,0)$
    (g) $(-2,4)$
    (h) $(-2,2,0)$
    (i) $(-1,1,1)$
    (j) $(3,1,-5)$
    (k) $(-2,0,1)$
    (l) $(5,0,-5)$
    (m) $(1,2,1)$
    (n) $(2,1,-1)$
    (o) $(-1,2,5)$

91. Determine a projecção ortogonal do vector $(1,1)$ em cada um dos seguintes subespaços de $\mathbb{R}^2$, munido do produto interno usual.

    (a) $L(\{(1,-1)\})$
    (b) $L(\{(-1,3)\})$
    (c) $L(\{(1,3)\})$
    (d) $L(\{(0,3)\})$

92. Verifique que as bases canónicas de $\mathbb{R}^2$, $\mathbb{R}^3$ e $\mathbb{R}^4$ são ortonormadas. Generalize o resultado para $\mathbb{R}^n$.

93. Encontre bases ortonormadas para cada um dos seguintes subespaços de $\mathbb{R}^3$, considerando o produto interno usual.

    (a) $L(\{(1,0,1),(1,1,0)\})$
    (b) $L(\{(1,0,0),(0,0,2)\})$
    (c) $L(\{(2,1,1),(1,2,2),(1,1,1))\})$
    (d) $L(\{(1,1,1),(1,1,-1)\})$
    (e) $L(\{(2,1,1),(1,2,2),(1,1,1))\})$
    (f) $L(\{(2,1,0),(0,1,2)\})$

94. Considere novamente o espaço $\mathbb{R}^4$ munido do produto interno usual. Use o método de ortogonalização de Gram–Schmidt para encontrar uma base ortonormal para cada um dos subespaços do Exercício 34 e determine as projecções ortogonais dos vectores $(1,0,0,0)$ e $(1,1,1,1)$ em cada um desses subespaços.

95. Considere o espaço $\mathcal{C}[-1,1]$ com o produto interno habitual.

    (a) Aplicando o processo de ortogonalização de Gram–Schmidt, encontre uma base ortonormal para o subespaço $L(\{1,x,x^2,x^3\})$.
    (b) Escreva a expressão para encontrar o polinómio de grau menor ou igual a 3 que melhor aproxima uma função f naquele intervalo.

Problemas

96. Considere o espaço $\mathbb{R}^3$ com o produto interno usual. Considere ainda o subespaço S de $\mathbb{R}^3$ gerado pelos vectores $(1,0,1)$ e $(1,1,0)$.

    (a) Calcule o ângulo entre os vectores $(1,0,1)$ e $1,1,0$.

    (b) Determine uma base ortogonal para S.

    (c) Calcule a projecção do vector $(3,1,-2)$ no espaço S.

97. Considere o subespaço S de $\mathbb{R}^3$ gerado pelos vectores $(2,1,0)$ e $(1,0,1)$ com o produto interno usual em $\mathbb{R}^3$.

    (a) Indique quais dos vectores $(-1,-2,3)$, $(3,2,0)$ e $(2,2,-1)$ pertencem a S.

    (b) Calcule o ângulo entre os vectores $(2,1,0)$ e $(1,0,1)$.

    (c) Encontre uma base ortonormal para S.

    (d) Escreva os vectores $(2,1,0)$ e $(1,0,1)$ em coordenadas da base determinada na alínea anterior.

98. Considere o espaço $M_{2\times 2}$ das matrizes quadradas de dimensão 2.

    (a) Verifique se o conjunto

    $$S = \left\{ \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 5 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$

    é linearmente independente.

    (b) Escreva

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

    como combinação linear dos elementos de S.

99. Considere o espaço P_2 dos polinómios de grau menor ou igual a 2.

    (a) Verifique se o conjunto $S = \{x^2+2x-1, 2x^2+x, x+3\}$ é um subconjunto linearmente independente de P_2.

    (b) Escreva o polinómio $2x^2+5x+2$ como combinação linear de elementos de S.

100. Considere o subespaço S de $M_{2\times 2}$ constituído pelas matrizes da forma

    $$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$$

    com o produto interno definido por

    $$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \cdot \begin{bmatrix} d & e \\ 0 & f \end{bmatrix} = ad + be + cf\,.$$

    (a) Calcule o ângulo entre as matrizes $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ e $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

(b) Verifique que a base $\left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$ de S é ortonormada.

(c) Encontre o elemento de S que melhor aproxima a matriz $\begin{bmatrix}2 & 1\\-1 & 1\end{bmatrix}$.

101. Seja S o subconjunto de $\mathbb{R}^3$ gerado pelos vectores $\overrightarrow{v_1} = (1,1,1)$, $\overrightarrow{v_2} = (-1,-1,2)$ e $\overrightarrow{v_3} = (1,1,4)$.

(a) Calcule o ângulo formado pelos vectores $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$.

(b) Verifique que v_3 é combinação linear de $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$.

(c) Determine uma base ortonormada para S.

(d) Indique a dimensão de S.

102. Considere o conjunto P_3 dos polinómios de grau 3 ou inferior.

(a) Mostre que os polinómios p de grau 3 tais que $p(1) = 0$ formam um subespaço vectorial de P_3.

Considere agora a base $B = \{1, (x-1), (x-1)^2, (x-1)^3\}$ de P_3.

(b) Escreva a matriz de mudança de base que converte coordenadas em B para coordenadas na base canónica de P_3.
Nota: $(x-1)^2 = x^2 - 2x + 1$ e $(x-1)^3 = x^3 - 3x^2 + 3x - 1$.

(c) Use a matriz determinada na alínea anterior para escrever $x^3 - 1$ e $x^2 + 2x + 1$ em coordenadas de B.

(d) Calcule os polinómios cujas coordenadas na base B são $(0,2,3,0)$ e $(1,2,1,0)$.

103. Sejam u e v vectores num espaço vectorial V com produto interno tais que

$$u \cdot u = 4 \qquad u \cdot v = 3 \qquad v \cdot v = 9\,.$$

Calcule:

(a) a norma de u e a norma de v;

(b) o ângulo entre u e v;

(c) as projecções ortogonais de v sobre u e de u sobre v;

(d) a distância entre u e v.
Sugestão: escreva $d(u,v) = \sqrt{(u-v)\cdot(u-v)}$ e aplique a linearidade do produto interno.

104. Indique, justificando, matrizes A de dimensão 3×3 satisfazendo as condições seguintes.

(a) O determinante de A é 2 e A tem no máximo dois zeros.

(b) A inversa de A é a matriz $\begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\\1 & 0 & 1\end{bmatrix}$.

(c) O vector $(1,2,1)$ pertence ao espaço das linhas de A, o vector $(2,1,-1)$ pertence ao espaço das colunas de A e A é invertível.

105. Determine a matriz diagonal que melhor aproxima

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

e calcule a distância de A a essa aproximação, considerando o produto interno e a norma usuais em $M_{2\times 2}$.

106. Sejam B, C e D bases de $M_{2\times 2}$ tais que as matrizes de mudança de base de B para C e de C para D são, respectivamente,

$$P_{CB} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad P_{DC} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

e seja M a matriz tal que $(M)_D = (2, 0, 0, 1)$. Determine $(M)_B$.

107. Seja V o subconjunto de P_3 constituído pelos polinómios p de grau menor ou igual a 3 tais que $p(x) = ax^3 + bx + c$.

   (a) Mostre que V é um subespaço linear de P_3.

   (b) Mostre que V é gerado por $S = \{x^3 - x, x^3 + x + 2, x^3 + 3x, 3x - 1, x^3\}$.

   (c) A partir de S, determine uma base B para V.

   (d) * Determine as coordenadas de $3x^3 + 2x - 5$ na base determinada na alínea anterior.

108. Determine a distância do ponto $(1, 0, 1, 0)$ ao plano de $\mathbb{R}^4$ definido pelo seguinte sistema de equações.

$$\begin{cases} x + y = 0 \\ x - z + w = 1 \end{cases}$$

109. Sejam B e C as duas bases seguintes de P_2.

$$B = \left\{x^2 + 1, x^2 - 1, 2x\right\} \qquad C = \left\{x^2, x^2 + x, x^2 + x + 1\right\}$$

   (a) Determine a matriz de mudança de base P_{CB} que transforma vectores de coordenadas relativamente à base B em vectores de coordenadas relativamente à base C.

   (b) A partir da matriz determinada na alínea anterior, calcule $(p)_C$, sabendo que p é um polinómio tal que $(p)_B = (1, 1, 1)$.

110. Seja A a matriz seguinte.

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

   (a) Verifique se $(2, 0, 1, 2)$ pertence ao espaço das colunas de A.

   (b) Verifique se $(1, -1, 0, 1)$ pertence ao núcleo de A.

   (c) Determine uma base para o espaço das linhas de A.

(d) Determine uma base ortogonal para o espaço das linhas de A.

(e) Calcule a projecção de $(1, 0, -2, 2)$ sobre o espaço das linhas de A.

111. Determine a distância entre a matriz

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

e a matriz com traço 0 que melhor aproxima A, considerando o produto interno e a norma usuais em $M_{2\times 2}$.

Capítulo 3

Transformações lineares

No capítulo anterior estudámos os espaços que são os objectos de estudo por excelência da Álgebra Linear: os espaços lineares. Vimos que a estrutura destes, com as operações de soma e produto por escalares, era adequada para capturar da forma mais geral o conceito de linearidade.

Neste capítulo vamos estudar as funções entre espaços lineares que se comportam bem com a linearidade: as transformações lineares. Embora a apresentação que aqui se faz apenas aborde uma pequena parte do estudo destas funções, teremos oportunidade de discutir os aspectos fundamentais para um aprofundamento posterior destas matérias.

3.1 Definição e primeiros exemplos

Uma transformação linear entre dois espaços vectoriais não é mais do que uma função entre esses espaços que respeita as operações de soma e produto por escalares.

Definição. Sejam V e W espaços lineares. Uma aplicação $T : V \to W$ diz-se uma *transformação linear* se verificar as seguintes propriedades para quaisquer u, v em V e α real.

1. $T(u + v) = T(u) + T(v)$
2. $T(\alpha u) = \alpha\, T(u)$

Esta definição significa simplesmente que a imagem da soma de dois vectores é a soma das imagens dos vectores; e que a imagem do produto dum vector por um escalar é o produto da imagem do vector pelo escalar. A Figura 3.1 ilustra estas propriedades para um caso particular.

É importante salientar desde já que este comportamento *não* se verifica para qualquer função. Seria perfeitamente aceitável definir uma função f como na Figura 3.2 — mas f não é uma transformação linear.

Antes de ver exemplos, vejamos algumas propriedades imediatas que serão úteis para caracterizar transformações lineares. Primeiro, atendendo a que $0_V = 0v$ para qualquer vector v no espaço vectorial V, temos que

$$T0_V = T(0v) = 0T(v) = 0_W\,,$$

e portanto T preserva zeros. De forma semelhante, sendo $v \in V$ arbitrário, sabemos que $-v = (-1)v$; então temos por um raciocínio análogo as igualdades

$$T(-v) = T((-1)v) = (-1)Tv = -Tv\,,$$

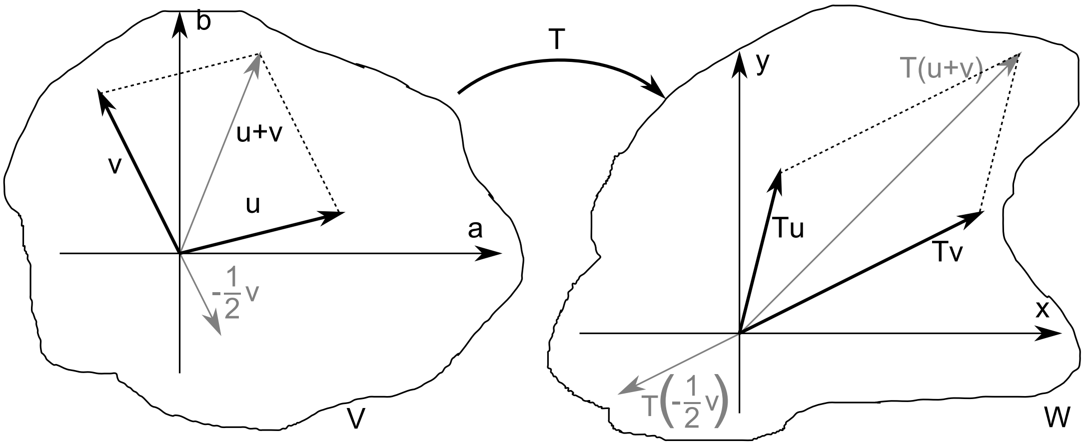


Figura 3.1: Comportamento duma transformação linear face à soma e produto por escalares.

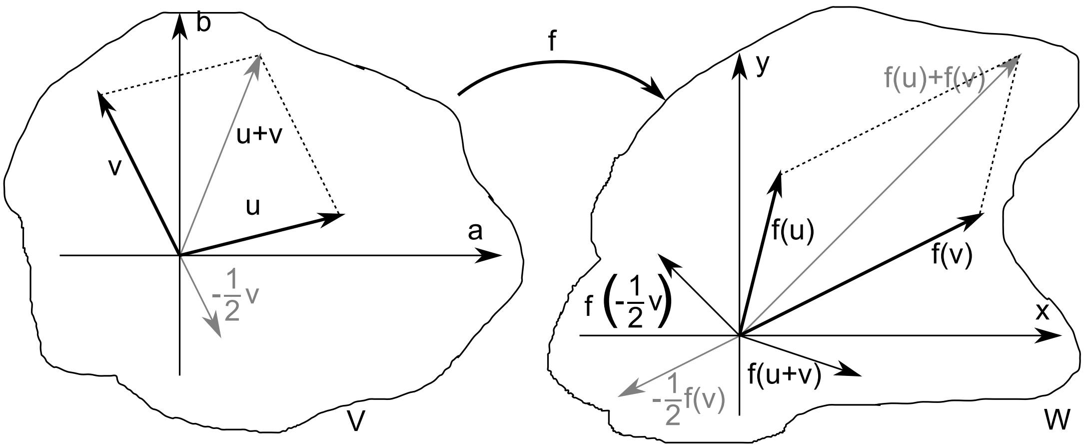


Figura 3.2: Uma função entre dois espaços vectoriais que não é transformação linear.

donde T preserva simétricos. Estas propriedades são por vezes úteis para verificar que uma função *não* é transformação linear: uma função que não transforme zeros em zeros ou simétricos em simétricos não pode nunca ser linear.

As transformações lineares podem ainda ser caracterizadas pela propriedade

$$T(\alpha u + \beta v) = \alpha\, T(u) + \beta\, T(v)$$

para quaisquer $u, v \in V$ e $\alpha, \beta \in \mathbb{R}$. A verificação deste facto é muito semelhante à prova da Proposição 27.

Vejamos alguns exemplos com funções reais (ou seja, do espaço linear $\mathbb{R}$ para si próprio).

Exemplo.

1. A função $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = 2x$ é uma transformação linear:

   - $f(u+v) = 2(u+v)$ e $f(u)+f(v) = 2u+2v = 2(u+v)$, logo $f(u+v) = f(u)+f(v)$;
   - $f(\alpha u) = 2\alpha u$ e $\alpha\, f(u) = \alpha 2u = 2\alpha u$, logo $f(\alpha u) = \alpha\, f(u)$.

2. A função $g : \mathbb{R} \to \mathbb{R}$ definida por $g(x) = 2x + 1$ não é uma transformação linear:

$$\begin{aligned} g(u+v) &= 2(u+v)+1 = 2u+2v+1 \\ g(u)+g(v) &= 2u+1+2v+1 = 2u+2v+2 \end{aligned}$$

donde $g(u+v) \neq g(u)+g(v)$. Portanto g não é uma transformação linear.

Em alternativa, poderíamos ter observado que $g(0) = 1$, donde g não preserva zeros e portanto não pode ser transformação linear.

3. A função $h : \mathbb{R} \to \mathbb{R}$ dada por $h(x) = e^x$ não é transformação linear:

$$\begin{aligned} h(u+v) &= e^{u+v} = e^u e^v \\ h(u)+h(v) &= e^u + e^v \end{aligned}$$

donde em geral $h(u+v) \neq h(u)+h(v)$, logo h não é transformação linear. Novamente, temos que $h(0) = 1$.

4. A função $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = \sin(x)$ também não é transformação linear:

$$\begin{aligned} f(u+v) &= \sin(u+v) \\ f(u)+f(v) &= \sin(u)+\sin(v) \end{aligned}$$

pelo que em geral $f(u+v) \neq f(u)+f(v)$, e portanto f não é transformação linear. Neste caso, $f(0) = 0$, donde não poderíamos concluir nada sobre f apenas a partir do seu valor neste ponto.

Exercício 1. Quais das seguintes funções de $\mathbb{R}$ para $\mathbb{R}$ são transformações lineares?

(a) $f(x) = 3x+2$

(b) $g(x) = -2x$

(c) $g(x) = x + x^2 - 1$

(d) $h(x) = 3$

(e) $f(x) = 0$

(f) $f(x) = 2x - 1$

(g) $h(x) = \log(x)$

(h) $g(x) = (x^2+1) - (x+1)^2$

(i) $g(x) = -5x$

Antes de prosseguirmos, vamos introduzir algumas abreviaturas que serão usadas adiante. Em geral, quando tratamos de transformações lineares omitimos os parêntesis exteriores sempre que possível. Assim, escreveremos Tu e Tv em vez de $T(u)$ e $T(v)$; escreveremos $T\overrightarrow{u}$ e $T\overrightarrow{v}$ em vez de $T\left(\overrightarrow{u}\right)$ e $T\left(\overrightarrow{v}\right)$; e, nomeadamente em $\mathbb{R}^n$, escreveremos $T(x,y)$ e $T(x,y,z)$ em vez de $T((x,y))$ e $T((x,y,z))$.

Vejamos agora alguns exemplos de transformações entre outros espaços vectoriais.

Exemplo.

1. A transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x,y) = (2x+y, -x)$ é uma transformação linear.

Em primeiro lugar, verificamos que T transforma somas em somas. Sejam (x_1, y_1) e (x_2, y_2) dois vectores arbitrários de $\mathbb{R}^2$.

$$\begin{aligned} T\left((x_1, y_1) + (x_2, y_2)\right) &= T\left(x_1 + x_2, y_1 + y_2\right) \\ &= \left(2\left(x_1 + x_2\right) + \left(y_1 + y_2\right), -\left(x_1 + x_2\right)\right) \\ &= \left(2x_1 + 2x_2 + y_1 + y_2, -x_1 - x_2\right) \end{aligned}$$

$$\begin{aligned} T\left(x_1, y_1\right) + T\left(x_2, y_2\right) &= \left(2x_1 + y_1, -x_1\right) + \left(2x_2 + y_2, -x_2\right) \\ &= \left(2x_1 + y_1 + 2x_2 + y_2, -x_1 - x_2\right) \end{aligned}$$

Concluímos que $T\left((x_1, y_1) + (x_2, y_2)\right) = T\left(x_1, y_1\right) + T\left(x_2, y_2\right)$, ou seja, T preserva a soma.

De seguida, vamos ver se T se comporta bem com produtos por escalares.

$$T\left(\alpha(x, y)\right) = T(\alpha x, \alpha y) = (2\alpha x + \alpha y, -\alpha x)$$

$$\begin{aligned} \alpha\, T(x, y) &= \alpha(2x + y, -x) = (\alpha(2x + y), -\alpha x) \\ &= (2\alpha x + \alpha y, -\alpha x) \end{aligned}$$

Logo $T\left(\alpha(x, y)\right) = \alpha\, T(x, y)$, ou seja, T preserva produtos por escalares. Assim, T é uma transformação linear.

2. A transformação $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (2x, 3y - x, x + 2y)$ também é uma transformação linear. Vamos verificar directamente que T preserva combinações lineares. Sejam (x_1, y_1) e (x_2, y_2) dois vectores arbitrários de $\mathbb{R}^2$.

$$\begin{aligned} &T\left(\alpha\left(x_1, y_1\right) + \beta\left(x_2, y_2\right)\right) = T\left(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2\right) \\ &\quad = \left(2\left(\alpha x_1 + \beta x_2\right), 3\left(\alpha y_1 + \beta y_2\right) - \left(\alpha x_1 + \beta x_2\right), \left(\alpha x_1 + \beta x_2\right) + 2\left(\alpha y_1 + \beta y_2\right)\right) \\ &\quad = \left(2\alpha x_1 + 2\beta x_2, 3\alpha y_1 + 3\beta y_2 - \alpha x_1 - \beta x_2, \alpha x_1 + \beta x_2 + 2\alpha y_1 + \beta y_2\right) \end{aligned}$$

$$\begin{aligned} &\alpha\, T\left(x_1, y_1\right) + \beta\, T\left(x_2, y_2\right) \\ &\quad = \alpha\left(2x_1, 3y_1 - x_1, x_1 + 2y_1\right) + \beta\left(2x_2, 3y_2 - x_2, x_2 + 2y_2\right) \\ &\quad = \left(2\alpha x_1, 3\alpha y_1 - \alpha x_1, \alpha x_1 + 2\alpha y_1\right) + \left(2\beta x_2, 3\beta y_2 - \beta x_2, \beta x_2 + 2\beta y_2\right) \\ &\quad = \left(2\alpha x_1 + 2\beta x_2, 3\alpha y_1 - \alpha x_1 + 3\beta y_2 - \beta x_2, \alpha x_1 + 2\alpha y_1 + \beta x_2 + 2\beta y_2\right) \end{aligned}$$

donde se conclui que $T\left(\alpha\left(x_1, y_1\right) + \beta\left(x_2, y_2\right)\right) = \alpha\, T\left(x_1, y_1\right) + \beta\, T\left(x_2, y_2\right)$ e portanto T é transformação linear.

3. Em espaços de polinómios, a transformação $T : P_1 \to P_2$ dada por $T(p(x)) = xp(x)$ é linear.

Sejam p e q dois polinómios arbitrários. Vejamos o comportamento de T relativamente à soma $p + q$.

$$T\left(p(x) + q(x)\right) = x(p(x) + q(x)) = xp(x) + xq(x) = T(p(x)) + T(q(x))$$

Conclui-se que $T(p(x) + q(x)) = T(p(x)) + T(q(x))$, ou seja, T preserva somas.

Sejam agora p um polinómio e α um número real. Vejamos o comportamento de T relativamente ao produto $\alpha\, p$.

$$T(\alpha p(x)) = x(\alpha p(x)) = \alpha x p(x) = \alpha\, T(p(x))$$

Logo $T(\alpha p(x)) = \alpha\, T(p(x))$, ou seja, T preserva produtos por escalares.

Assim, T é uma transformação linear.

4. Na alínea anterior, não usámos em nenhum passo a informação de que os argumentos de T eram polinómios de grau 1. Assim, o mesmo raciocínio poderia ter sido usado para mostrar que $T : P_4 \to P_5$ definida por $T(p(x)) = xp(x)$ era uma transformação linear.

   Aliás, podemos mesmo definir uma transformação linear T no espaço $\mathcal{P}$ de *todos* os polinómios por $T(p(x)) = xp(x)$. A prova de que T é linear é, mais uma vez, idêntica à do exemplo anterior.

5. A transformação $T : P_2 \to P_2$ definida por $Tp = p'(0) + 1$ não é uma transformação linear. De facto, se p e q forem polinómios de grau 2, temos que

$$\begin{aligned} T(p+q) &= (p+q)'(0) + 1 = p'(0) + q'(0) + 1 \\ Tp + Tq &= (p'(0)+1) + (q'(0)+1) = p'(0) + q'(0) + 2 \end{aligned}$$

   e estes valores não coincidem.

6. A transformação $T : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$ dada por $Tf = f'$ é linear. De facto, escolhendo duas funções f e g diferenciáveis e reais α e β, sabemos das regras de derivação que

$$T(\alpha f + \beta g) = (\alpha f + \beta g)' = \alpha f' + \beta g' = \alpha\, Tf + \beta\, Tg$$

   donde T é uma transformação linear.

7. A transformação $T : \mathcal{C}^0([a,b]) \to \mathbb{R}$ definida por $Tf = \int_a^b f(x)\, dx$ é uma transformação linear. De facto, escolhendo duas funções f e g contínuas em $[a,b]$ e dois reais α e β, as propriedades do integral definido garantem que

$$\begin{aligned} T(\alpha f + \beta g) &= \int_a^b (\alpha f + \beta g)(x)\, dx = \int_a^b (\alpha f(x) + \beta g(x))\, dx \\ &= \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx = \alpha\, Tf + \beta\, Tg \end{aligned}$$

   donde T é uma transformação linear.

8. A transformação $T : M_{2\times 2} \to \mathbb{R}$ dada por $TA = \det(A)$ não é uma transformação linear. De facto, se escolhermos as matrizes

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

   temos $\det(A) = \det(B) = 0$, enquanto $A + B = \mathbf{I_2}$ e portanto $\det(A+B) = 1 \neq 0 + 0$.

9. A transformação $T : \mathbb{R}^3 \to P_2$ definida por $T(a,b,c) = ax^2 + 2ax + bc + 1$ não é transformação linear: para o vector $0_{\mathbb{R}^3}$ tem-se

$$T\vec{0} = T(0,0,0) = 1 \neq 0_{P_2}$$

   donde T não preserva zeros.

Exercício 2. Quais das seguintes funções são transformações lineares?

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$
$T(x, y) = (x + 2y, 3x - y)$

(b) $T : \mathbb{R}^3 \to \mathbb{R}^2$
$T(x, y, z) = (2x - y + z, y - 4z + 1)$

(c) $T : M_{2\times 2} \to M_{2\times 2}$
$T(A) = A \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$

(d) $T : M_{2\times 2} \to \mathbb{R}^2$
$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (3a + 4b, c - d)$

(e) $T : M_{2\times 2} \to \mathbb{R}$
$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a^2 + b^2$

(f) $T : P_2 \to P_2$
$T\left(ax^2 + bx + c\right) = a(x+1)^2 + b(x+1) + c$

(g) $T : P_2 \to P_2$
$T\left(ax^2 + bx + c\right) = c(x+1)^2 + b(x+1) + a$

(h) $T : P_2 \to P_2$
$T\left(ax^2 + bx + c\right) =$
$= (a + 1)x^2 + (b + 1)x + (c + 1)$

No decurso desta exposição, encontrámos já várias operações definidas em espaços vectoriais particulares que preservavam a estrutura de espaço linear; na altura chamámos até a atenção para esse facto. Essas operações fornecem outros exemplos de transformações lineares em espaços que não $\mathbb{R}$.

Exemplo.

1. Sejam V e W espaços lineares arbitrários. A transformação $T : V \to W$ definida por $Tv = 0_W$ para todo o $v \in V$ é linear.

   - $T(u + v) = 0_W$ e $Tu + Tv = 0_W + 0_W = 0_W$, logo $T(u + v) = Tu + Tv$.
   - $T(\alpha u) = 0_W$ e $\alpha\, Tu = \alpha 0_W = 0_W$, logo $T(\alpha u) = \alpha\, Tu$.

   Assim T é uma transformação linear, dita *transformação nula* de V para W.

2. Num espaço linear qualquer V, a transformação $T : V \to V$ definida por $Tu = u$ para todo $u \in V$ é linear. De facto, $T(\alpha u + \beta v) = \alpha u + \beta v = \alpha\, Tu + \beta\, Tv$, donde T preserva combinações lineares. A transformação T diz-se *transformação identidade* em V e denota-se por vezes por $\mathbf{I_V}$.

3. Num espaço linear qualquer V, a transformação $T : V \to V$ dada por $Tu = k\, u$ para qualquer $u \in V$ (com $k \in \mathbb{R}$ fixo) é linear.

   - $T(u + v) = k(u + v)$ e $Tu + Tv = ku + kv = k(u + v)$, logo $T(u + v) = Tu + Tv$.
   - $T(\alpha u) = k\alpha u$ e $\alpha\, Tu = \alpha k u = k\alpha u$, logo $T(\alpha u) = \alpha\, Tu$.

   Assim, T é uma transformação linear. Diz-se que T é uma *dilatação* se $|k| > 1$ e uma *contracção* se $0 < |k| < 1$. Se $k = 0$, T é a transformação nula, enquanto que T é a transformação identidade para $k = 1$. Se $k = -1$, T diz-se uma *simetria*.

   Estes nomes vêm da interpretação geométrica destas transformações em $\mathbb{R}^n$. A Figura 3.3 ilustra estes casos para $\mathbb{R}^2$.

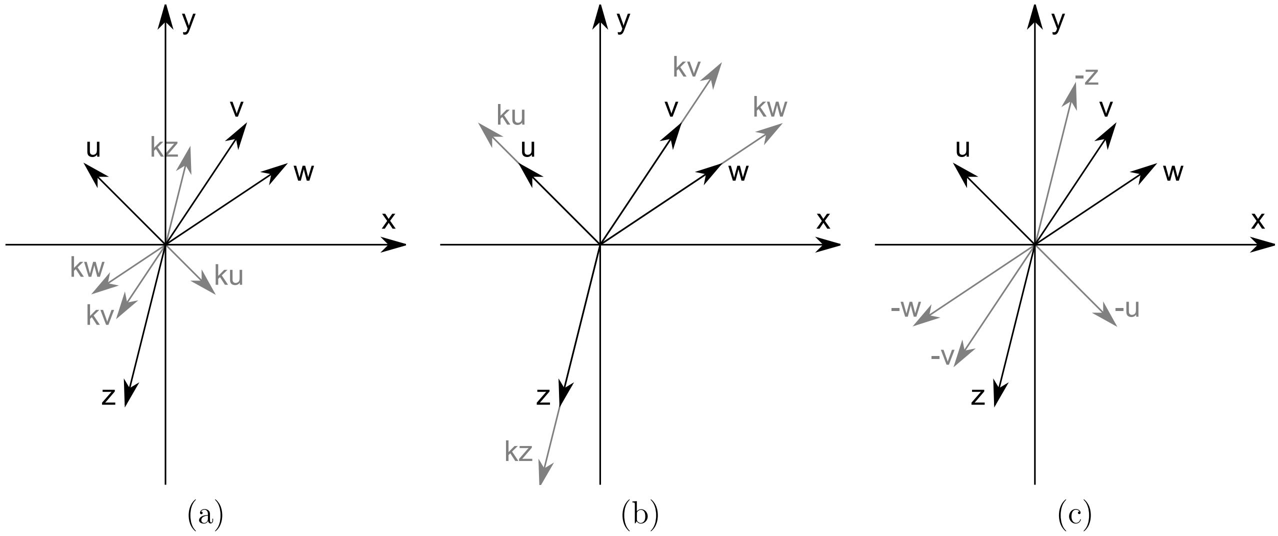


Figura 3.3: Diferentes aspectos da transformação $T(u) = k\ u$ em $\mathbb{R}^2$, para $|k| < 1$ (a), $|k| > 1$ (b) e $k = -1$ (c).

4. Sejam A uma matriz $n \times m$ e $T : \mathbb{R}^m \to \mathbb{R}^k$ a transformação definida por $T\underline{x} = A\underline{x}$. Das propriedades do produto de matrizes, sabemos que

$$T\left(\alpha \underline{u} + \beta \underline{v}\right) = A\left(\alpha \underline{u} + \beta \underline{v}\right) = \alpha A\underline{u} + \beta A\underline{v} = \alpha\ T\underline{u} + \beta\ T\underline{v}$$

para quaisquer $u, v \in \mathbb{R}^m$ e $\alpha, \beta \in \mathbb{R}$, donde T é transformação linear. Esta transformação linear T chama-se *transformação matricial*.

5. Num espaço euclideano V, seja v um vector e $Tu = \langle u, v\rangle$ a transformação que aplica cada vector u no seu produto interno com v. Das propriedades que vimos do produto interno, temos que

$$T(\alpha u + \beta w) = \langle \alpha u + \beta w, v\rangle = \alpha \langle u, v\rangle + \beta \langle w, v\rangle = \alpha\ Tu + \beta\ Tw\,,$$

donde o produto interno é uma transformação linear.

6. Num espaço euclideano V, seja v um vector e $Tu = P_v\left(u\right)$ a projecção ortogonal de u sobre v. Das propriedades que vimos da projecção ortogonal, temos que

$$T(\alpha u + \beta w) = P_v\left(\alpha u + \beta w\right) = \alpha P_v\left(u\right) + \beta P_v\left(w\right) = \alpha\ Tu + \beta\ Tw\,,$$

donde a projecção ortogonal é uma transformação linear.

Exercício 3. Quais das seguintes funções são transformações lineares?

(a) $T : \mathbb{R}^3 \to \mathbb{R}$
$T(u) = \|u\|$

(b) $T : M_{3\times3} \to \mathbb{R}$
$T(A) = \mathrm{tr}A$

(c) $T : M_{3\times3} \to \mathbb{R}$
$T(A) = A + \mathbf{I_3}$

(d) $T : M_{3\times3} \to \mathbb{R}$
$T(A) = \det(A)$

(e) $T : M_{m\times n} \to M_{n\times m}$
$T(A) = A^T$

(f) $T : \mathcal{F}\left(\mathbb{R}\right) \to \mathcal{F}\left(\mathbb{R}\right)$
$T(f(x)) = f(x+1)$

(g) $T : \mathbb{R}^3 \to \mathbb{R}$
$T(u) = u \cdot a$, com $a \in \mathbb{R}^3$ fixo

(h) $T : M_{2\times 2} \to M_{2\times 3}$
$T(A) = AB$, com $B_{2\times 3}$ fixa

(i) $T : \mathcal{F}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$
$T(f(x)) = f(x) + 1$

(j) $T : \mathcal{C}^2\mathbb{R} \to \mathcal{F}(\mathbb{R})$
$T(f(x)) = 3f(x) + f''(x)$

É importante observar que uma transformação linear $T : V \to W$ também é uma transformação linear quando a restringimos a um subespaço vectorial V' de V. De facto, para T ser transformação linear em V' tem de satisfazer duas propriedades para todos os vectores de V'; ora como $V' \subseteq V$ e T satisfaz essas propriedades para todos os vectores de V, automaticamente T é transformação linear entre V' e W.

Por exemplo: vimos atrás que a diferenciação e o integral definido (ou seja, as transformações lineares que levam uma função f para f' e para $\int_a^b f(x)\, dx$, respectivamente) são exemplos de transformações lineares definidas em $\mathcal{C}^1(\mathbb{R})$ e $\mathcal{C}^0([a, b])$, respectivamente. Então sabemos, por exemplo, que:

- a diferenciação é uma transformação linear entre P_3 e P_2;
- a diferenciação é uma transformação linear entre P_5 e P_4;
- a diferenciação é uma transformação linear entre $\mathcal{P}$ e $\mathcal{P}$;
- o integral definido é uma transformação linear entre $P_3[a, b]$ e $\mathbb{R}$;
- o integral definido é uma transformação linear entre $P_5[a, b]$ e $\mathbb{R}$;
- o integral definido é uma transformação linear entre $\mathcal{P}[a, b]$ e $\mathbb{R}$.

Exercício 4. Verifique que as transformações lineares do Exercício 2 ainda são transformações lineares se substituirmos:

- $\mathbb{R}^2$ por $L(\{(1, 0)\})$ (alíneas (a), (b) e (d));
- $\mathbb{R}^3$ por $L(\{(1, 0, 0), (0, 0, 1)\})$ (alínea (b));
- $M_{2\times 2}$ pelo conjunto das matrizes diagonais de dimensão 2 (alíneas (c), (d) e (e));
- P_2 por P_1 (alíneas (f), (g) e (h)).

3.2 Bases e representação matricial

Vimos atrás que as transformações lineares preservam somas e produtos por escalares. Como consequência disso, preservam toda a estrutura de espaço vectorial — zeros, simétricos e combinações lineares. Vamos ver que esta propriedade nos permite, por um lado, trabalhar com transformações lineares usando relativamente pouca informação; e por outro reduzindo os problemas de transformações lineares a problemas de cálculo matricial.

3.2.1 Transformações lineares e combinações lineares

Vimos atrás que qualquer transformação linear $T : V \to W$ preserva zeros e simétricos. Em geral, sendo $c_1v_1 + c_2v_2 + \cdots + c_nv_n$ uma combinação linear em V, temos mesmo que

$$T(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = T(c_1v_1) + T(c_2v_2) + \cdots + T(c_nv_n) = c_1Tv_1 + c_2Tv_2 + \cdots + c_nTv_n\,,$$

donde T preserva combinações lineares arbitrárias.

Uma consequência importante desta propriedade é o facto de uma transformação linear $T : V \to W$ ficar completamente determinada pelos seus valores numa base (qualquer) de V. De facto, seja $B = \{v_1, v_2, \ldots, v_n\}$ uma base de V. Então todo o elemento v de V pode ser escrito como combinação linear dos elementos de B na forma

$$v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$$

com $c_1, c_2, \ldots, c_n \in \mathbb{R}$. Mas uma vez que T é transformação linear, temos necessariamente pela proposição anterior que

$$Tv = T(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = c_1Tv_1 + c_2Tv_2 + \cdots + c_nTv_n\,.$$

Portanto, conhecendo apenas as imagens dos elementos de uma base do espaço de partida, podemos conhecer a imagem de qualquer objecto de V.

Por exemplo, seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ uma transformação linear tal que $T(1,0,0) = (1,1)$, $T(0,1,0) = (2,-1)$ e $T(0,0,1) = (3,4)$ e suponhamos que pretendíamos determinar o valor de $T(2,1,3)$. Escrevendo $(2,1,3)$ como combinação linear de $(1,0,0)$, $(0,1,0)$ e $(0,0,1)$, o facto de T ser uma transformação linear fornece-nos informação suficiente, uma vez que

$$\begin{aligned} T(2,1,3) &= T(2(1,0,0) + (0,1,0) + 3(0,0,1)) \\ &= 2T(1,0,0) + T(0,1,0) + 3T(0,0,1) \\ &= 2(1,1) + (2,-1) + 3(3,4) \\ &= (13,13) \end{aligned}$$

O mesmo raciocínio funciona mesmo que os pontos conhecidos não sejam os da base canónica e, até, mesmo que não formem uma base do espaço — desde que o vector pretendido seja combinação linear daqueles.

Exemplo.

1. Considere-se uma transformação linear $T : \mathbb{R}^3 \to P_1$ tal que $T(1,1,0) = x^2 + 1$ e $T(0,1,1) = 2x$. Embora esta informação não seja suficiente para caracterizar completamente T, podemos usá-la para calcular valores de T nos pontos que são combinação linear de $(1,1,0)$ e $(0,1,1)$. Por exemplo, uma vez que $(1,3,2) = (1,1,0) + 2(0,1,1)$, temos que

$$\begin{aligned} T(1,3,2) &= T((1,1,0) + 2(0,1,1)) = T(1,1,0) + 2T(0,1,1) \\ &= \left(x^2 + 1\right) + 2(2x) = x^2 + 4x + 1\,. \end{aligned}$$

2. Uma situação frequente em que recorremos sistematicamente a esta propriedade é o cálculo de derivadas. Vimos atrás que a operação de derivação é uma transformação linear de $\mathcal{C}^1(\mathbb{R})$ para $\mathcal{C}^0(\mathbb{R})$; a linearidade permite-nos derivar combinações lineares de funções cujas derivadas são conhecidas.

Por exemplo,

$$\left(3x^2 + \sin(x) - 3e^x + \sqrt{2}\log x\right)' = 3\left(x^2\right)' + (\sin(x))' - 3\left(e^x\right)' + \sqrt{2}\left(\log x\right)'$$
$$= 6x + \cos(x) - 3e^x + \frac{\sqrt{2}}{x}\,.$$

3. Analogamente, uma vez que o integral definido é uma transformação linear de $\mathcal{C}^0\left([a,b]\right)$ para $\mathbb{R}$, podemos recorrer a integrais de valor conhecido para calcular valores de novas expressões. Por exemplo, sabendo que

$$\int_0^1 x^2\,dx = \frac{1}{3} \qquad \int_0^1 \log(x+1)\,dx = 2\log 2 - 1 \qquad \int_0^1 \sin(\pi x)\,dx = \frac{2}{\pi}\,,$$

temos que

$$\int_0^1 \left(3x^2 - 2\log(x+1) + 5\sin(\pi x)\right)\,dx$$
$$= 3\int_0^1 x^2\,dx - 2\int_0^1 \log(x+1)\,dx + 5\int_0^1 \sin(\pi x)\,dx$$
$$= 3 \times \frac{1}{3} - 2 \times (2\log 2 - 1) + 5 \times \frac{2}{\pi} = 3 - 4\log 2 + \frac{10}{\pi}\,.$$

Exercício 5. Para cada uma das seguintes transformações lineares definidas num espaço vectorial V, calcule o valor pedido sabendo os valores nos pontos indicados.

(a) $T : V \to \mathbb{R}$
$\begin{cases} T(u) = 2 \\ T(v) = -2 \end{cases}$
Qual o valor de $T(3u - v)$?

(b) $T : V \to P_2$
$\begin{cases} T(u) = 3x - 1 \\ T(v) = x^2 - 4 \end{cases}$
Qual o valor de $T(3u - 2v)$?

(c) $T : V \to M_{2\times 2}$
$\begin{cases} T(u) = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \\ T(v) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\ T(w) = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix} \end{cases}$
Qual o valor de $T(-2u + v + 2w)$?

Usando o mesmo raciocínio, podemos determinar a expressão geral duma transformação linear, conhecidos os seus valores em suficientes pontos.

Exemplo.

1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear tal que $T(1,2) = (0,1)$ e $T(1,1) = (1,1)$. Vamos ver que com esta informação conseguimos determinar a expressão de $T(x,y)$ para qualquer vector $\overrightarrow{v} = (x,y) \in \mathbb{R}^2$.

   Para aplicar o raciocínio anterior, temos de conhecer os valores de T numa base de $\mathbb{R}^2$; vamos então começar por verificar que $\{(1,2),(1,1)\}$ é base de $\mathbb{R}^2$. Para tal, colocamos os vectores nas linhas de uma matriz e fazemos eliminação de Gauss.

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{matrix} \\ l_2 - l_1 \end{matrix} \longrightarrow \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$$

Esta matriz tem característica 2, logo $B = \{(1,2),(1,1)\}$ é uma base de $\mathbb{R}^2$. Então qualquer vector $(x,y) \in \mathbb{R}^2$ pode ser escrito como combinação linear desta base. Sendo $(x,y) = c_1(1,2) + c_2(1,1)$, temos

$$\begin{aligned} T(x,y) &= T\left(c_1(1,2) + c_2(1,1)\right) = c_1 T(1,2) + c_2 T(1,1) \\ &= c_1(0,1) + c_2(1,1) = (0,c_1) + (c_2,c_2) \\ &= (c_2, c_1 + c_2)\ . \end{aligned}$$

Para obter uma expressão em termos de x e y, precisamos ainda de conhecer as coordenadas (c_1,c_2) de (x,y).

$$(x,y) = c_1(1,2) + c_2(1,1) \longrightarrow (x,y) = (c_1 + c_2, 2c_1 + c_2) \longrightarrow \begin{cases} c_1 + c_2 = x \\ 2c_1 + c_2 = y \end{cases}$$

$$\longrightarrow \left[\begin{array}{cc|c} 1 & 1 & x \\ 2 & 1 & y \end{array}\right] \begin{array}{c} \\ l_2 - 2l_1 \end{array} \longrightarrow \left[\begin{array}{cc|c} 1 & 1 & x \\ 0 & -1 & y - 2x \end{array}\right] \longrightarrow \begin{cases} c_1 + c_2 = x \\ -c_2 = y - 2x \end{cases}$$

$$\longrightarrow \begin{cases} c_1 + 2x - y = x \\ c_2 = 2x - y \end{cases} \longrightarrow \begin{cases} c_1 = y - x \\ c_2 = 2x - y \end{cases}$$

Então a expressão geral de T é

$$T(x,y) = (c_2, c_1 + c_2) = (2x - y, x)\,.$$

2. Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ uma transformação linear tal que $T(1,1,1) = (0,1)$, $T(1,0,1) = (2,2)$ e $T(1,1,0) = (1,-1)$.

   Tal como atrás, para encontrarmos uma expressão geral para $T(x,y,z)$ temos de conhecer os valores de T numa base de $\mathbb{R}^3$. Vamos então usar novamente eliminação de Gauss para verificar que $\{(1,1,1),(1,0,1),(1,1,0)\}$ é uma base de $\mathbb{R}^3$ construindo uma matriz com estes vectores em linhas.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{array}{c} \\ l_2 - l_1 \\ l_3 - l_1 \end{array} \longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

   Esta matriz tem característica 3, logo $B = \{(1,1,1),(1,0,1),(1,1,0)\}$ é uma base de $\mathbb{R}^3$. Então qualquer vector $(x,y) \in \mathbb{R}^2$ pode ser escrito como combinação linear desta base. Sendo $(x,y,z) = c_1(1,1,1) + c_2(1,0,1) + c_3(1,1,0)$, temos

$$\begin{aligned} T(x,y,z) &= T\left(c_1(1,1,1) + c_2(1,0,1) + c_3(1,1,0)\right) \\ &= c_1 T(1,1,1) + c_2 T(1,0,1) + c_3 T(1,1,0) \\ &= c_1(0,1) + c_2(2,2) + c_3(1,-1) \\ &= (2c_2 + c_3, c_1 + 2c_2 - c_3)\ . \end{aligned}$$

   Para obter uma expressão em termos de x, y e z, precisamos finalmente de conhecer as coordenadas (c_1,c_2,c_3) de (x,y,z).

$$(x,y,z) = c_1(1,1,1) + c_2(1,0,1) + c_3(1,1,0) \longrightarrow (x,y,z) = (c_1+c_2+c_3, c_1+c_3, c_1+c_2)$$

$$\longrightarrow \begin{cases} c_1+c_2+c_3 = x \\ c_1+c_3 = y \\ c_1+c_2 = z \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & x \\ 1 & 0 & 1 & y \\ 1 & 1 & 0 & z \end{array}\right] \begin{array}{c} \\ l_2 - l_1 \\ l_3 - l_1 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & x \\ 0 & -1 & 0 & y-x \\ 0 & 0 & -1 & z-x \end{array}\right]$$

$$\longrightarrow \begin{cases} c_1+c_2+c_3 = x \\ -c_2 = y-x \\ -c_3 = z-x \end{cases} \longrightarrow \begin{cases} c_1 + x - y + x - z = x \\ c_2 = x-y \\ c_3 = x-z \end{cases} \longrightarrow \begin{cases} c_1 = z+y-x \\ c_2 = x-y \\ c_3 = x-z \end{cases}$$

Então a expressão geral de T é

$$T(x,y,z) = (2c_2+c_3, c_1+2c_2-c_3) = (3x-2y-z, 2z-y)\,.$$

3. Seja agora $T : P_2 \to \mathbb{R}^3$ uma transformação linear satisfazendo $T(2x+1) = (0,1,1)$, $T\left(x^2-1\right) = (2,-1,0)$ e $T(x) = (0,0,-1)$.

   Novamente, temos de começar por verificar que dispomos duma base de P_2. Recordando que este espaço tem dimensão 3, basta verificar que $\{2x+1, x^2-1, x\}$ é uma base de P_2. Para proceder como atrás, vamos escrever estes vectores em coordenadas da base canónica de P_2 e usar estas coordenadas como linhas duma matriz a que aplicaremos eliminação de Gauss.

$$\begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \; l_2 + l_1 \longrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \; 2l_3 - l_2 \longrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

   Esta matriz tem característica 3, logo $B = \{2x+1, x^2-1, x\}$ é uma base de P_2. Então qualquer polinómio $p(x) = a + bx + cx^2 \in P_2$ pode ser escrito como combinação linear desta base. Sendo $p(x) = c_1(2x+1) + c_2\left(x^2-1\right) + c_3 x$, temos que

$$\begin{aligned} T(p(x)) &= T\left(c_1(2x+1) + c_2\left(x^2-1\right) + c_3 x\right) \\ &= c_1 T(2x+1) + c_2 T\left(x^2-1\right) + c_3 T(x) \\ &= c_1(0,1,1) + c_2(2,-1,0) + c_3(0,0,-1) \\ &= (2c_2, c_1-c_2, c_1-c_3)\;. \end{aligned}$$

   Para obter uma expressão geral em termos das variáveis a, b e c, precisamos ainda de conhecer as coordenadas (c_1, c_2, c_3) do polinómio original $p(x) = a + bx + cx^2$ em termos da base B. Este problema reduz-se mais uma vez a resolver um sistema de equações lineares.

$$\begin{aligned} a + bx + cx^2 &= c_1(2x+1) + c_2\left(x^2-1\right) + c_3 x \\ &\longrightarrow a + bx + cx^2 = 2c_1 x + c_1 + c_2 x^2 - c_2 + c_3 x \\ &\longrightarrow a + bx + cx^2 = (c_1 - c_2) + (2c_1 + c_3)\,x + c_2 x^2 \end{aligned}$$

$$\longrightarrow \begin{cases} c_1 - c_2 = a \\ 2c_1 + c_3 = b \\ c_2 = c \end{cases} \longrightarrow \begin{cases} c_1 = a + c \\ 2(a+c) + c_3 = b \\ c_2 = c \end{cases} \longrightarrow \begin{cases} c_1 = a + c \\ c_3 = b - 2a - 2c \\ c_2 = c \end{cases}$$

   Então a expressão geral de T é

$$T\left(a + bx + cx^2\right) = (2c_2, c_1-c_2, c_1-c_3) = (2c, a, 3a-b+3c)\,.$$

4. Para terminar, seja $T : M_{2\times 2} \to \mathbb{R}$ uma transformação linear tal que

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

O espaço $M_{2\times 2}$ tem dimensão 4. Vamos ver que as quatro matrizes em que conhecemos valores de T formam de facto uma base deste espaço. Para tal, vamos escrever estas matrizes em coordenadas da base canónica e formar uma matriz com essas coordenadas nas suas linhas.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \begin{array}{l} \\ l_2 - l_1 \\ l_3 - l_1 \\ l_4 - l_1 \end{array} \longrightarrow \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & -2 \\ 0 & -2 & -2 & -2 \end{bmatrix}$$

Por troca de linhas ($l_2 \leftrightarrow l_4$) obtemos uma matriz de característica 4, pelo que

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

é uma base de $M_{2\times 2}$. Seja

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = c_1\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + c_2\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} + c_4\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$$

uma matriz genérica. Então

$$\begin{aligned} TA &= T\left(c_1\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + c_2\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} + c_4\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}\right) \\ &= c_1T\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + c_2T\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} + c_3T\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} + c_4T\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \\ &= 2c_1\,. \end{aligned}$$

Para obter uma expressão em termos das entradas de A, precisamos ainda de conhecer as coordenadas (c_1, c_2, c_3, c_4) de A — mais concretamente, a coordenada c_1, já que as restantes não ocorrem na expressão de TA.

$$\begin{aligned} &\begin{bmatrix} a & b \\ c & d \end{bmatrix} = c_1\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + c_2\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} + c_4\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \\ &\longrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c_1+c_2+c_3+c_4 & c_1+c_2+c_3-c_4 \\ c_1+c_2-c_3-c_4 & c_1-c_2-c_3-c_4 \end{bmatrix} \\ &\longrightarrow \begin{cases} c_1+c_2+c_3+c_4 = a \\ c_1+c_2+c_3-c_4 = b \\ c_1+c_2-c_3-c_4 = c \\ c_1-c_2-c_3-c_4 = d \end{cases} \longrightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & a \\ 1 & 1 & 1 & -1 & b \\ 1 & 1 & -1 & -1 & c \\ 1 & -1 & -1 & -1 & d \end{array}\right] \begin{array}{l} \\ l_2 - l_1 \\ l_3 - l_1 \\ l_4 - l_1 \end{array} \\ &\longrightarrow \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & a \\ 0 & 0 & 0 & -2 & b-a \\ 0 & 0 & -2 & -2 & c-a \\ 0 & -2 & -2 & -2 & d-a \end{array}\right] \longrightarrow \begin{cases} c_1+c_2+c_3+c_4 = a \\ -2c_4 = b-a \\ -2c_3-2c_4 = c-a \\ -2c_2-2c_3-2c_4 = d-a \end{cases} \end{aligned}$$

Da última equação tiramos que $c_2 + c_3 + c_4 = \frac{a-d}{2}$; substituindo este valor na primeira, obtemos $c_1 + \frac{a-d}{2} = a$, donde $c_1 = \frac{a+d}{2}$.

Esta é uma situação em que poderia ter sido útil usar a regra de Cramer, uma vez que só estamos interessados em obter o valor duma coordenada.

Então a expressão geral de T é

$$T \begin{bmatrix} a & b \\ c & d \end{bmatrix} = 2c_1 = a + d\,.$$

Exercício 6. Para cada uma das seguintes transformações lineares, escreva a sua forma geral sabendo os valores nos pontos indicados.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$
$\begin{cases} T(1,0) = (1,-2) \\ T(0,1) = (-4,1) \end{cases}$

(b) $T : \mathbb{R}^2 \to \mathbb{R}^2$
$\begin{cases} T(1,1) = (1,-2) \\ T(1,0) = (-4,1) \end{cases}$

(c) $T : \mathbb{R}^2 \to \mathbb{R}^3$
$\begin{cases} T(-2,1) = (1,-2,0) \\ T(1,3) = (0,-3,5) \end{cases}$

(d) $T : \mathbb{R}^3 \to \mathbb{R}^3$
$\begin{cases} T(1,1,1) = (2,-1,4) \\ T(1,1,0) = (3,0,1) \\ T(1,0,0) = (-1,5,1) \end{cases}$

(e) $T : \mathbb{R}^3 \to \mathbb{R}^2$
$\begin{cases} T(1,2,1) = (1,0) \\ T(2,9,0) = (-1,1) \\ T(3,3,4) = (0,1) \end{cases}$

3.2.2 Representação matricial de transformações lineares

Podemos desenvolver o raciocínio anterior um pouco mais. Sendo $T : V \to W$ uma transformação linear e fixada uma base $B = \{v_1, v_2, \ldots, v_n\}$ para o espaço linear V, vimos acima que podemos escrever qualquer vector $v \in V$ como combinação linear dos elementos de B, na forma

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n\,,$$

ou, equivalentemente,

$$v_B = (c_1, c_2, \ldots, c_n)\,,$$

tendo-se

$$Tv = T(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n) = c_1 T v_1 + c_2 T v_2 + \cdots + c_n T v_n\,.$$

Ora os vectores $Tv_1, Tv_2, \ldots, Tv_n$ são elementos de W. Então, se fixarmos uma base $C = \{w_1, w_2, \ldots, w_m\}$, cada Tv_i também pode ser escrito como combinação linear dos vectores de C, ou seja,

$$Tv_i = a_{1i} w_1 + a_{2i} w_2 + \cdots + a_{mi} w_m\,,$$

com $a_{1i}, a_{2i}, \ldots, a_{mi} \in \mathbb{R}$.

Substituindo estes valores na equação anterior, temos que

$$\begin{aligned}
T(v) &= c_1\left(a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m\right) + c_2\left(a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_2\right) \\
&\quad + \cdots + c_n\left(a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m\right) \\
&= \underbrace{\left(c_1a_{11} + c_2a_{12} + \ldots + c_na_{1n}\right)}_{d_1} w_1 + \underbrace{\left(c_1a_{21} + c_2a_{22} + c_na_{2n}\right)}_{d_2} w_2 \\
&\quad + \cdots \underbrace{\left(c_1a_{m1} + c_2a_{m2} + \cdots + c_na_{mn}\right)}_{d_m} w_m \\
&= d_1w_1 + d_2w_2 + \cdots + d_mw_m\,.
\end{aligned}$$

Logo as coordenadas de $T(v)$ na base C são dadas por

$$(Tv)_C = (d_1, d_2, \ldots, d_m)\ .$$

Da definição de $d_1, d_2, \ldots, d_m$, temos que

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}}_{T_{CB}} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

ou, de forma abreviada,

$$(Tv)_C = T_{CB}(v)_B\,.$$

Definição. Sejam V e W espaços lineares com bases respectivamente $B = \{v_1, v_2, \ldots, v_n\}$ e $C = \{w_1, w_2, \ldots, w_m\}$ e $T : V \to W$ uma transformação linear. A matriz T_{CB} cujas colunas são as coordenadas de Tv_i na base C diz-se a *representação matricial da transformação T em relação às bases B e C*.

Podemos escrever a matriz T_{CB} na seguinte forma sugestiva.

$$T_{CB} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ (Tv_1)_C & (Tv_2)_C & \ldots & (Tv_n)_C \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

Proposição 49. Sejam V e W espaços vectoriais com bases B e C e $T : V \to W$ uma transformação linear. Então, para qualquer vector $v \in V$, tem-se a relação

$$(Tv)_C = T_{CB}(v)_B\,.$$

Estes resultados permitem representar qualquer transformação linear entre espaços vectoriais de dimensão finita como uma matriz. Na prática, tal como as bases nos permitem trabalhar com qualquer espaço vectorial como se de $\mathbb{R}^n$ se tratasse, a representação matricial de transformações lineares permite-nos reduzir todo o trabalho com estas transformações a operações sobre matrizes. Em particular, o cálculo de valores de T resume-se a um produto duma matriz por um vector.

Por exemplo, se quisermos não determinar valores de T, mas sim resolver equações envolvendo T, o nosso problema fica reduzido à resolução de um sistema de equações lineares. De facto, se pretendermos resolver uma equação $T(v) = u$, basta substituir a transformação pela sua representação matricial e os elementos pelos correspondentes vectores de coordenadas:

$$T(v) = u \Longleftrightarrow T_{CB}(v)_B = (u)_C\,,$$

e estamos perante um sistema de equações lineares cujas técnicas de resolução são já bem conhecidas.

O diagrama que se segue ilustra a situação: qualquer dos dois caminhos entre os vértices opostos dos quadrados em baixo produz ao mesmo resultado.

$$\begin{tikzcd}
u \arrow[rrr, bend left] \arrow[ddd] & & & T(u) \arrow[ddd] \\
& V \arrow[r, "T"] \arrow[d, "(\cdot)_B"'] & W \arrow[d, "(\cdot)_C"] & \\
& \mathbb{R}^n \arrow[r, "T_{CB}"'] & \mathbb{R}^m & \\
(u)_B \arrow[rrr, bend right] & & & (T(u))_C
\end{tikzcd}$$

Figura 3.4: Transformações lineares e coordenadas.

Vejamos um exemplo extremamente simples. Consideremos a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (2x - y, 3x + 2y)$ e procuremos determinar a sua representação matricial em relação à base canónica B de $\mathbb{R}^2$. Para tal, vamos escrever as imagens dos vectores desta base em coordenadas dela própria.

$$(T(1,0))_B = ((2,3))_B = (2,3) \qquad (T(0,1))_B = ((-1,2))_B = (-1,2)$$

A matriz T_{BB} é a matriz que tem estes vectores em colunas, ou seja,

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

Suponhamos agora que fixávamos outras bases para $\mathbb{R}^2$. Concretamente, vamos trabalhar com a base $C = \{(1,1),(1,-1)\}$ para o espaço de partida e com a base $D = \{(1,1),(0,1)\}$ para o espaço de chegada. (Embora este exemplo possa parecer algo artificial, teremos ocasião de ver mais adiante contextos em que de facto ocorrem situações semelhantes.) Procedemos como atrás.

$$(T(1,1))_D = ((1,5))_D = (1,4) \qquad (T(1,-1))_D = ((3,1))_D = (3,-2)$$

Omitimos o cálculo detalhado das coordenadas, que fica como exercício.

A matriz que representa T em relação a estas bases é então

$$T_{CD} = \begin{bmatrix} 1 & 3 \\ 4 & -2 \end{bmatrix},$$

obtida escrevendo aqueles vectores em colunas.

Os seguintes exemplos envolvem todos transformações já discutidas anteriormente, definidas em $\mathbb{R}^n$ ou subespaços destes, em que todos estes cálculos são por regra bastante directos.

Exemplo.

1. Considere-se a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $(2x + y, -x)$. As imagens dos vectores da base canónica B, expressas em coordenadas dessa base, são

$$(T(1,0))_B = ((2,-1))_B = (2,-1) \qquad (T(0,1))_B = ((1,0))_B = (1,0)\,,$$

donde a representação matricial de T em relação à base B é

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

Podemos verificar que, por exemplo,

$$(T(2,3))_B = T_{BB}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = (7,-2)\,,$$

que corresponde ao que se obteria usando directamente a expressão de T.

2. Considere-se a transformação $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x,y) = (2x, 3y-x, x+2y)$. Em relação às bases canónicas B de $\mathbb{R}^2$ e C de $\mathbb{R}^3$, a representação matricial de T calcula-se calculando as imagens dos vectores da base B em coordenadas da base C,

$$(T(1,0))_C = ((2,-1,1))_C = (2,-1,1) \qquad (T(0,1))_C = ((0,3,2))_C = (0,3,2)$$

e escrevendo estas em colunas da matriz

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

Novamente, podemos verificar que

$$(T(2,-1))_C = T_{CB}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = (4,-5,0)\,,$$

que corresponde novamente ao que se obteria segundo a expressão de T.

3. Consideremos outra vez a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(1,2) = (0,1)$ e $T(1,1) = (1,1)$. Determinámos atrás a expressão geral de T em relação à base canónica de $\mathbb{R}^2$; porém, uma vez que $B = \{(1,2),(1,1)\}$ e $C = \{(0,1),(1,1)\}$ são duas bases deste espaço, podemos escrever directamente a representação matricial de T em relação a estas bases, observando que

$$(T(1,2))_C = ((0,1))_C = (1,0) \qquad (T(1,1))_C = ((1,1))_C = (0,1)\,,$$

pelo que $T_{CB} = \mathbf{I_2}$.

Tomemos por exemplo $\overrightarrow{v}$ tal que $(\overrightarrow{v})_B = (1,-1)$. Então $\overrightarrow{v} = (1,2) - (1,1) = (0,1)$. Temos

$$(T(0,1))_C = T_{CB}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{I_2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (1,-1)\,,$$

e este vector (em coordenadas de C) é $(0,1) - (1,1) = (-1,0)$. Então $T(0,1) = (-1,0)$.

4. Da mesma forma semelhante, podemos determinar facilmente a representação matricial da transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T(1,1,1) = (0,1)$, $T(1,0,1) = (2,2)$ e $T(1,1,0) = (1,-1)$ em relação a bases adequadas de $\mathbb{R}^2$ e $\mathbb{R}^3$ sem passar pela sua expressão explícita. Contudo, aqui há que ter cuidado: os vectores $(1,1,1)$, $(1,0,1)$ e $(1,1,0)$ formam uma base B de $\mathbb{R}^3$, mas os vectores $(0,1)$, $(2,2)$ e $(1,-1)$ não são linearmente independentes, pois $\mathbb{R}^2$ tem dimensão 2. Para encontrar uma base C de $\mathbb{R}^2$ temos portanto de escolher dois vectores linearmente independentes dentre estes; tomemos por exemplo $C = \{(0,1),(2,2)\}$. Temos então

$$\begin{aligned}
(T(1,1,1))_C &= ((0,1))_C = (1,0)\\
(T(1,0,1))_C &= ((2,2))_C = (0,1)\\
(T(1,1,0))_C &= (1,-1)_C = \left(-2,\frac{1}{2}\right),
\end{aligned}$$

donde

$$T_{CB} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & \frac{1}{2} \end{bmatrix}.$$

Mais uma vez, sendo $\overrightarrow{v}$ o vector cujas coordenadas na base B são $(3,-2,2)$, temos

$$\overrightarrow{v} = 3(1,1,1) - 2(1,0,1) + 2(1,1,0) = (3,5,1).$$

Para calcular $T\overrightarrow{v}$, usamos a representação matricial:

$$(T\overrightarrow{v})_C = T_{CB}\begin{bmatrix} 3 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 2 \end{bmatrix} = (-1,-1)$$

donde $T\overrightarrow{v} = -(0,1) - (2,2) = (-2,-3)$.

Exercício 7. Represente matricialmente as seguintes transformações lineares relativamente às bases canónicas dos espaços de partida e de chegada.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ com $T(x,y) = (3x-2y, 2x)$

(b) $T : \mathbb{R}^2 \to \mathbb{R}^2$ com $T(x,y) = (x+y, x-y)$

(c) $T : \mathbb{R}^3 \to \mathbb{R}^2$ com $T(x,y,z) = (z-2x, 3y-2z)$

Vejamos agora exemplos envolvendo transformações lineares entre outros espaços vectoriais.

Exemplo.

1. Considere-se a transformação linear $T : P_1 \to P_2$ definida por

$$T(p(x)) = p(x) + \left(1+x^2\right)p'(x).$$

Sejam $B = \{1,x\}$ e $C = \{1,x,x^2\}$ as bases canónicas de P_1 e P_2 respectivamente. Vamos determinar a representação de T em relação a estas bases. Esta matriz T_{CB} tem por colunas as imagens dos elementos de B escritos em coordenadas da base C.

Vamos começar por calcular $T(1)$ e $T(x)$ em coordenadas da base C.

$$\begin{aligned}(T(1))_C &= \left(1+\left(1+x^2\right)1'\right) = (1)_C = (1,0,0)\\ (T(x))_C &= \left(x+\left(1+x^2\right)x'\right) = \left(x+1+x^2\right)_C = (1,1,1)\end{aligned}$$

Assim, temos que

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

Apesar de esta transformação ter um aspecto complicado, a sua representação nas bases canónicas é extremamente simples. Por outro lado, esta representação através duma matriz acentua a semelhança com as transformações entre espaços $\mathbb{R}^n$.

Vamos usar esta representação para calcular $T(5+2x)$. Para tal, começamos por escrever este polinómio em coordenadas da base canónica de P_1 para de seguida multiplicar pela matriz T_{CB}.

$$(T(5+2x))_C = T_{CB}(5+2x)_B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = (7,2,2)$$

Assim, $T(5+2x) = 7+2x+2x^2$, facto que pode ser comprovado recorrendo à expressão explícita de T.

2. Consideremos agora a transformação $T : M_{2\times 2} \to P_2$ tal que

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a+d)x^2 + (c-b)\,.$$

Tomando as bases canónicas B e C de $M_{2\times 2}$ e P_2, respectivamente, podemos determinar a representação matricial T_{CB} desta transformação em relação a essas bases.

Começamos como habitualmente por calcular a imagem dos vectores de B, escrevendo-os em coordenadas de C.

$$\left(T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)_C = \left(x^2\right)_C = (0,0,1) \qquad \left(T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right)_C = (-1)_C = (-1,0,0)$$
$$\left(T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right)_C = (1)_C = (1,0,0) \qquad \left(T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)_C = \left(x^2\right)_C = (0,0,1)$$

Então a matriz pretendida é

$$T_{CB} = \begin{bmatrix} 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Podemos usar esta representação para calcular a imagem da matriz

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

Em coordenadas da base B, esta matriz corresponde ao vector $(2,-1,-3,5)$, donde

$$\left(T\begin{bmatrix} 2 & -1 \\ -3 & 5 \end{bmatrix}\right)_C = T_{CB}\begin{bmatrix} 2 \\ -1 \\ -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ -3 \\ 5 \end{bmatrix} = (-2,0,7)\,,$$

e portanto $TA = -2+7x^2$.

3. Finalmente, seja $T : P_3 \to M_{2\times 2}$ a transformação linear definida por

$$T(p(x)) = \begin{bmatrix} p(0) & p'(2) \\ -p''(-2) & 3p(1)+p''(-1) \end{bmatrix}.$$

Vamos considerar novamente as bases canónicas $B = \{1, x, x^2, x^3\}$ de P_3 e

$$C = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$$

de $M_{2\times 2}$. As imagens dos vectores de B em coordenadas da base C são agora

$$\begin{aligned} (T(1))_C &= \left(\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\right)_C = (1,0,0,3) \\ (Tx)_C &= \left(\begin{bmatrix} 0 & 1 \\ 0 & 3 \end{bmatrix}\right)_C = (0,1,0,3) \\ \left(Tx^2\right)_C &= \left(\begin{bmatrix} 0 & 4 \\ -2 & 5 \end{bmatrix}\right)_C = (0,4,-2,5) \\ \left(Tx^3\right)_C &= \left(\begin{bmatrix} 0 & 12 \\ 12 & -3 \end{bmatrix}\right)_C = (0,12,12,-3) \end{aligned}$$

donde a matriz que representa T em relação a estas bases é

$$T_{CB} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 4 & 12 \\ 0 & 3 & -2 & 12 \\ 3 & 3 & 5 & -3 \end{bmatrix}.$$

Vamos novamente usar esta representação para calcular Tp para um polinómio particular. Tomemos $p(x) = 3+2x-2x^2-x^3$; observe-se que, para prosseguir pela definição de T, teríamos de calcular a segunda derivada de p; contudo, através da representação matricial o caminho é muito mais simples:

$$(Tv)_C = T_{CB}(v)_B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 4 & 12 \\ 0 & 3 & -2 & 12 \\ 3 & 3 & 5 & -3 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ -2 \\ -1 \end{bmatrix} = (3,-18,-2,8)$$

donde

$$Tv = \begin{bmatrix} 3 & -18 \\ -2 & 8 \end{bmatrix}.$$

Estes exemplos mostram que o caso em que a transformação linear é dada através de valores num conjunto de pontos que forma uma base do espaço de partida é mais simples do que o caso em que a transformação é dada através da sua expressão, uma vez que, para efeitos práticos, o primeiro passo já está feito.

Exercício 8. Represente matricialmente as seguintes transformações lineares relativamente às bases canónicas dos espaços de partida e de chegada.

(a) $T : P_2 \to P_1$ com $T(ax^2 + bx + c) = -(3a + 2b)x + (b + c)$

(b) $T : P_2 \to P_2$ com $T(ax^2 + bx + c) = a(x-1)^2 + b(x-1) + c$

(c) $T : P_2 \to P_3$ com $T(p(x)) = xp(x)$

(d) $T : M_{2\times 2} \to \mathbb{R}$ com $T(A) = 3\,\mathrm{tr}(A)$

(e) $T : M_{2\times 2} \to M_{2\times 2}$ com $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 2c & a+c \\ b-2c & d \end{bmatrix}$

Exercício 9. Represente matricialmente a transformação linear $T : M_{2\times 2} \to M_{2\times 2}$ com expressão

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

relativamente à base canónica de $M_{2\times 2}$.

Existe alguma relação entre a matriz que representa T e a matriz $\begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}$?

3.2.3 Transformações lineares dadas por matrizes

Acabámos de ver a forma de, dada uma transformação linear, determinar uma sua representação matricial. Vamos agora discutir o problema contrário: dada uma matriz que representa uma transformação linear, como determinar a expressão dessa transformação.

Vimos atrás que a transformação linear $T : \mathbb{R}^m \to \mathbb{R}^n$ definida por $T\underline{x} = A\underline{x}$, onde A é uma matriz $n \times m$, é uma transformação linear. Veremos nesta secção que, num sentido preciso, estas são as transformações lineares representativas de todas as transformações lineares entre espaços vectoriais de dimensão finita — da mesma forma que os espaços $\mathbb{R}^n$ são representativos de todos os espaços lineares de dimensão finita.

Se $T : V \to W$ é representada pela matriz T_{CB}, fixadas bases $B = \{v_1, v_2, \ldots, v_n\}$ de V e $C = \{w_1, w_2, \ldots, w_m\}$ de W, então por definição de representação podemos calcular a imagem de qualquer vector $v \in V$ através da relação $(Tv)_C = T_{CB}(v)_B$. Se conseguirmos escrever esta expressão para um vector genérico v, encontramos a expressão geral da transformação T. De certa forma, estamos a fazer o percurso descrito na Figura 3.4 no sentido inverso.

Vejamos um exemplo concreto. Consideremos a transformação $T : \mathbb{R}^3 \to \mathbb{R}^4$ cuja representação em relação às bases canónicas B e C de $\mathbb{R}^3$ e $\mathbb{R}^4$ é

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

Para determinar $T(x,y)$, vamos usar a propriedade expressa acima. Pela forma como a representação matricial matriz T_{CB} é construída, sabemos que

$$T_{CB} = \begin{bmatrix} \vdots & \vdots & \vdots \\ (T(1,0,0))_C & (T(0,1,0))_C & (T(0,0,1))_C \\ \vdots & \vdots & \vdots \end{bmatrix};$$

ora isto significa que podemos ler directamente nas colunas de T_{CB} as imagens dos vectores da base canónica de $\mathbb{R}^3$ em coordenadas da base canónica de $\mathbb{R}^4$: $T(1,0,0) = (1,2,3,0)$, $T(0,1,0) = (-2,1,1,2)$ e $T(0,0,1) = (0,-1,5,-2)$. A partir daqui, podemos calcular a expressão geral de T usando o facto de qualquer vector se exprimir de forma trivial em coordenadas da base canónica.

$$T(x,y,z) = T_{CB}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 2 & 1 & -1 \\ 3 & 1 & 5 \\ 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x-2y, 2x+y-z, 3x+y+5z, 2y-2z)$$

É esta a expressão geral de T.

No caso em que a representação matricial não é relativa às bases canónicas, temos de aplicar as técnicas descritas nas secções anteriores. Tome-se por exemplo como $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear cuja representação na base $B = \{(1,3),(-1,4)\}$ é

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

Para determinar a expressão geral de T, vamos começar por ler os valores de $T(1,3)$ e $T(-1,4)$, em coordenadas de B, nas colunas de T_{BB}.

$$T(1,3) = 1(1,3) - 2(-1,4) = (3,-5) \qquad T(-1,4) = 3(1,3) + 5(-1,4) = (-2,29)$$

Agora estamos perante o problema abordado na Secção 3.2.1: determinar a expressão geral de T conhecidos os seus valores numa base do espaço de partida. Sabemos já como proceder: qualquer vector (x,y) de $\mathbb{R}^2$ pode ser escrito como combinação linear dos elementos de B.

$$(x,y) = c_1(1,3) + c_2(-1,4) \longrightarrow (x,y) = (c_1 - c_2, 3c_1 + 4c_2) \longrightarrow \begin{cases} c_1 - c_2 = x \\ 3c_1 + 4c_2 = y \end{cases}$$

$$\longrightarrow \left[\begin{array}{cc|c} 1 & -1 & x \\ 3 & 4 & y \end{array}\right] \; l_2 - 3l_1 \longrightarrow \left[\begin{array}{cc|c} 1 & -1 & x \\ 0 & 7 & y-3x \end{array}\right] \longrightarrow \begin{cases} c_1 - c_2 = x \\ 7c_2 = -3x + y \end{cases}$$

$$\longrightarrow \begin{cases} c_1 = \frac{y+4x}{7} \\ c_2 = \frac{y-3x}{7} \end{cases}$$

Então, a expressão geral de $T(x,y)$ é

$$\begin{aligned} T(x,y) &= T\left(c_1(1,3) + c_2(-1,4)\right) = c_1 T(1,3) + c_2 T(-1,4) = c_1(3,-5) + c_2(-2,29) \\ &= (3c_1 - 2c_2, -5c_1 + 29c_2) = \left(3\frac{4x+y}{7} - 2\frac{-3x+y}{7}, -5\frac{4x+y}{7} + 29\frac{-3x+y}{7}\right) \\ &= \left(\frac{18x+y}{7}, \frac{-107x+24y}{7}\right) \end{aligned}$$

Vejamos mais alguns exemplos deste processo.

Exemplo.

1. Seja $T : \mathbb{R}^2 \to P_2$ a transformação linear representada pela matriz

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

   em relação às bases $B = \{(1,1),(1,-1)\}$ de $\mathbb{R}^2$ e $C = \{1, x, x^2\}$ (base canónica) de P_2. Então

$$T(1,1) = 1 + x^2 \qquad T(1,-1) = 1 + 2x + x^2 .$$

   Sendo (a,b) um vector genérico de $\mathbb{R}^2$, temos que as suas coordenadas em relação a B são dadas por

$$(a,b) = c_1(1,1) + c_2(1,-1) \longrightarrow \begin{cases} c_1 = \frac{a+b}{2} \\ c_2 = \frac{a-b}{2} \end{cases}$$

   donde a expressão de $T(a,b)$ é

$$\begin{aligned} T(a,b) &= T\left(c_1(1,1) + c_2(1,-1)\right) = c_1T(1,1) + c_2T(1,-1) \\ &= \frac{a+b}{2}\left(1 + x^2\right) + \frac{a-b}{2}\left(1 + 2x + x^2\right) = a + (a-b)x + ax^2 . \end{aligned}$$

2. Seja $T : P_2 \to \mathbb{R}^3$ a transformação linear representada pela matriz

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

   em relação às bases canónicas B e C de P_2 e $\mathbb{R}^3$, respectivamente. Então

$$T(1) = (1,2,1) \qquad Tx = (-1,1,0) \qquad Tx^2 = (2,-1,1)$$

   donde a expressão de Tp para um polinómio p genérico é

$$\begin{aligned} T\left(a + bx + cx^2\right) &= aT(1) + bTx + cTx^2 = a(1,2,1) + b(-1,1,0) + c(2,-1,1) \\ &= (a - b + 2c, 2a + b - c, a + c) . \end{aligned}$$

3. Seja $T : P_3 \to P_2$ a transformação linear cuja representação, em relação às bases canónicas B e C dos dois espaços, é

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

   Podemos ler as imagens dos vectores da base canónica de P_3 nas colunas desta matriz, expressas em coordenadas da base canónica de P_2. Temos então que

$$\begin{aligned} (T(1))_C &= (0,0,0,0) \text{ donde } T(1) = 0 \\ (Tx)_C &= (1,0,0,0) \text{ donde } Tx = 1 \\ \left(Tx^2\right)_C &= (0,2,0,0) \text{ donde } Tx^2 = 2x \\ \left(Tx^3\right)_C &= (0,0,3,0) \text{ donde } Tx^3 = 3x^2 . \end{aligned}$$

Então $T\left(a+bx+cx^2+dx^3\right)=aT(1)+bTx+cTx^2+dTx^3=b+2cx+3dx^2$.

A transformação linear T não é mais do que o operador derivação entre os espaços P_3 e P_2. A sua representação matricial, porém, permite-nos calcular derivadas de polinómios duma forma completamente diferente da habitual.

Exercício 10. Represente matricialmente as transformações lineares do Exercício 7 relativamente às bases B e C apresentadas para os espaços de partida e chegada.

(a) $B=C=\{(1,1),(-1,0)\}$

(b) $B=\{(1,0),(0,1)\}$ e $C=\{(1,1),(0,-1)\}$

(c) $B=\{(1,1,1),(1,1,0),(1,0,0)\}$ e $C=\{(1,1),(1,0)\}$

Exercício 11. Represente matricialmente as transformações lineares do Exercício 8 relativamente às bases B e C apresentadas para os espaços de partida e chegada.

(a) $B=\{x^2,x,1\}$ e $C=\{1,2x\}$

(b) $B=\{1,x,x^2\}$ e $C=\{1,x-1,x^2-x\}$

(c) $B=\{1,1+x,1+x+x^2\}$ e $C=\{1,1+x,1+x+x^2,1+x+x^3\}$

(d) $B=\left\{\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}1&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\1&1\end{bmatrix}\right\}$ e $C=\{3\}$

(e) $B=C=\left\{\begin{bmatrix}1&0\\0&1\end{bmatrix},\begin{bmatrix}1&0\\0&-1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix},\begin{bmatrix}0&1\\-1&0\end{bmatrix}\right\}$

3.2.4 Transformações definidas em subespaços

Da exposição da secção anterior a questão que provavelmente mais fica no ar é: qual o interesse de representar matricialmente uma transformação linear em relação a uma base que não a base canónica? Na realidade, há várias respostas; uma das vantagens, conforme exemplificámos, é encontrar bases que tornam essa representação mais simples (o que simplifica substancialmente o nosso trabalho se estivermos interessados em estudar propriedades da transformação, mais do que em calcular os seus valores).

O que sucede muitas vezes, contudo, é que estamos a definir transformações não nos espaços canónicos $\mathbb{R}^n$, $M_{m\times n}$ ou P_n, mas sim em seus subespaços para os quais não há bases canónicas. Nestas situações, não temos outro recurso senão usar bases menos simples para escrever vectores de coordenadas e representar transformações.

Vejamos um exemplo. Consideremos o subespaço $S\subseteq P_3$ dos polinómios p tais que $p(2)=0$. É fácil verificar que S é de facto um subespaço linear; uma vez que qualquer polinómio de S se pode escrever na forma $p(x)=(x-2)p^*(x)$, com p^* de grau menor ou igual a 2, temos que a dimensão de S é 3 e uma base de S é formada pelos produtos $(x-2)1$, $(x-2)x$ e $(x-2)x^2$, ou seja, S admite a base

$$B = \left\{x-2, x^2-2x, x^3-2x^2\right\} .$$

Consideremos a transformação linear $T : S \to P_2$ tal que $Tp = p^*$. Por outras palavras, T divide um polinómio de grau menor ou igual a 3 por $(x-2)$, estando definida apenas quando aquela divisão é exacta.

Fixando a base canónica C de P_2, podemos representar matricialmente T de forma muito simples como

$$T_{CB} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I_3} ;$$

para determinar a imagem de um polinómio arbitrário (de S) basta escrevê-lo em coordenadas de B e obtemos imediatamente a sua imagem em coordenadas de C, uma vez que

$$(Tp)_C = T_{CB}(p)_B = \mathbf{I_3}(p)_B = (p)_B .$$

Seja por exemplo $p(x) = x^3 - 4x^2 + 7x - 6$. Para escrever p em coordenadas de B, temos de resolver um sistema de equações lineares.

$$\begin{aligned} p(x) &= c_1(x-2) + c_2\left(x^2-2x\right) + c_3\left(x^3-2x^2\right) \\ &\longrightarrow x^3 - 4x^2 + 7x - 6 = c_1x - 2c_1 + c_2x^2 - 2c_2x + c_3x^3 - 2c_3x^2 \\ &\longrightarrow -6 + 7x - 4x^2 + x^3 = -2c_1 + (c_1-2c_2) + (c_2-2c_3) + c_3 \\ &\longrightarrow \begin{cases} -2c_1 = -6 \\ c_1 - 2c_2 = 7 \\ c_2 - 2c_3 = -4 \\ c_3 = 1 \end{cases} \quad \longrightarrow \begin{cases} c_1 = 3 \\ c_2 = -2 \\ c_3 = 1 \end{cases} \end{aligned}$$

Concluímos portanto que $Tp = p^* = 3 - 2x + x^2$. Observe-se que, na prática, fizemos a divisão de dois polinómios resolvendo um sistema de equações lineares, em alternativa aplicar a outros algoritmos como a regra de Ruffini.

Exemplo. Consideremos o plano $S = L(\{(1,2,1),(1,0,-1)\})$ e a recta $R = L(\{(1,1,1)\})$, ambos subespaços de $\mathbb{R}^3$, com a transformação linear $T : S \to R$ definida como a projecção ortogonal sobre R.

$$T\overrightarrow{v} = P_{(1,1,1)}\left(\overrightarrow{v}\right)$$

Uma vez que S é gerado por dois vectores linearmente independentes, estes formam uma base B deste espaço. Calculando a imagem destes dois vectores e escrevendo-a em coordenadas de $C = \{(1,1,1)\}$, obtemos as seguintes igualdades.

$$\begin{aligned} T(1,2,1) &= P_{(1,1,1)}\left((1,2,1)\right) = \frac{(1,2,1)\cdot(1,1,1)}{\|(1,1,1)\|^2}(1,1,1) = \frac{4}{3}(1,1,1) \\ T(1,0,-1) &= P_{(1,1,1)}\left((1,0,-1)\right) = \frac{(1,0,-1)\cdot(1,1,1)}{\|(1,1,1)\|^2}(1,1,1) = \overrightarrow{0} \end{aligned}$$

A representação matricial de T é então

$$T_{CB} = \begin{bmatrix} \frac{4}{3} & 0 \end{bmatrix} .$$

Com esta matriz, podemos calcular projecções ortogonais de qualquer ponto de S sobre R resolvendo um sistema de equações lineares. Por exemplo:

$$\begin{aligned} P_R\left((5,4,-1)\right) &= T(5,4,-1) = \left(T_{CB}((5,4,-1))_B\right)(1,1,1) \\ &= \left(\begin{bmatrix} \frac{4}{3} & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right)(1,1,1) = \frac{8}{3}(1,1,1) = \left(\frac{8}{3},\frac{8}{3},\frac{8}{3}\right). \end{aligned}$$

3.2.5 Mudança de base

Vimos atrás que, dado um espaço linear V de dimensão n com bases $B = \{v_1, v_2, \ldots, v_n\}$ e $C = \{w_1, w_2, \ldots, w_n\}$, a matriz mudança de base P_{CB} que transforma coordenadas de um objecto na base B para as suas coordenadas na base C é uma matriz que satisfaz

$$(v)_C = P_{CB}(v)_B$$

para qualquer vector v de V. Esta matriz é obtida colocando nas colunas as coordenadas dos elementos da base B escritos na base C.

$$P_{CB} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ (v_1)_C & (v_2)_C & \cdots & (v_n)_C \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

Vimos ainda que toda a matriz mudança de base é invertível, tendo-se

$$(P_{CB})^{-1} = P_{BC}\,.$$

Sejam V e W espaços vectoriais, B e B' bases de V, C e C' bases de W e $T : V \to W$ uma transformação linear. Consideremos as representações matriciais T_{CB} e $T_{C'B'}$ de T em relação, respectivamente, às bases B e C e às bases B' e C'.

Tomando $v \in V$ arbitrário, temos que

$$(Tv)_{C'} = P_{C'C}(T(v))_C \qquad (Tv)_C = T_{CB}(v)_B \qquad (v)_B = P_{BB'}(v)_{B'}$$

e, portanto,

$$(Tv)_{C'} = P_{C'C}(T(v))_C = P_{C'C}T_{CB}(v)_B = P_{C'C}T_{CB}P_{BB'}(v)_{B'}$$

donde necessariamente

$$P_{C'C}T_{CB}P_{BB'} = T_{C'B'}\,.$$

Graficamente, esta relação é simples de compreender. Por um lado, a matriz $T_{C'B'}$ calcula $(Tv)_{C'}$ directamente a partir de $(v)_{B'}$. Por outro, partindo de $(v)_{B'}$, a matriz $P_{BB'}$ converte estas coordenadas em $(v)_B$; a matriz T_{CB} calcula o resultado $(Tv)_C$; e a matriz $P_{C'C}$ converte novamente este resultado para coordenadas de C'. Este processo está representado na Figura 3.5

Esta técnica é bastante útil quando é simples calcular a representação matricial duma transformação linear em relação a uma base que não a que nos interessa. Voltemos ao exemplo da transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ satisfazendo $T(1,2) = (0,1)$ e $T(1,1) = (1,1)$. Vimos atrás que a representação de T em relação às bases $B = \{(1,2),(1,1)\}$ e $C = \{(0,1),(1,1)\}$ de $\mathbb{R}^2$ se obtinha directamente: $T_{CB} = \mathbf{I_2}$.

$$\begin{tikzcd}
(v)_{B'} \arrow[r, bend left, "P_{BB'}"] \arrow[rrr, bend right, "T_{C'B'}"'] & (v)_B \arrow[r, bend left, "T_{CB}"] & (Tv)_C \arrow[r, bend left, "P_{C'C}"] & (Tv)_{C'}
\end{tikzcd}$$

Figura 3.5: Transformações lineares e mudança de base.

Para determinar a representação T_{DD} em relação à base canónica D de $\mathbb{R}^2$, vamos usar a fórmula anterior. Temos que

$$T_{DD} = P_{DC}T_{CB}P_{BD}$$

pelo que basta calcularmos as matrizes de mudança de base P_{DC} e P_{BD}. O cálculo de P_{DC} é directo: basta escrever os vectores de C em coordenadas da base canónica de $\mathbb{R}^2$ e colocá-los nas colunas desta matriz.

$$P_{DC} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Para calcular P_{BD}, o mais simples é inverter a matriz P_{DB}, que se escreve directamente de forma semelhante.

$$P_{BD} = (P_{DB})^{-1} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}^{-1} = -\begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$$

Então temos que

$$T_{DD} = P_{DC}T_{CB}P_{BD} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{I_2} \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$$

é a representação de T em relação à base canónica de $\mathbb{R}^2$.

Exemplo.

1. Consideremos a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ anteriormente discutida definida por $T(x, y) = (2x + y, -x)$, que vimos ter representação matricial relativamente à base canónica B de $\mathbb{R}^2$ dada por
$$T_{BB} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}.$$
Vamos determinar a representação matricial de T em relação à base $C = \{(1, 1), (1, -1)\}$ de $\mathbb{R}^2$. Para tal, vamos simplesmente determinar as matrizes de mudança de base P_{BC} e P_{CB}, recorrendo depois à relação
$$T_{CC} = P_{CB}T_{BB}P_{BC}\,.$$
Ora a matriz P_{BC} tem por colunas os vectores de C em coordenadas de B, que são trivialmente conhecidos, enquanto P_{CB} é a sua inversa. Temos então que
$$\begin{aligned} T_{CC} &= P_{CB}T_{BB}P_{BC} = (P_{BC})^{-1}\,T_{BB}P_{BC} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= -\frac{1}{2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} -2 & 0 \\ -4 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}. \end{aligned}$$

2. Outro exemplo que abordámos envolveu a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $T(1,1,1) = (0,1)$, $T(1,0,1) = (2,2)$ e $T(1,1,0) = (1,-1)$. Vimos que, escolhendo as bases $B = \{(1,1,1),(1,0,1),(1,1,0)\}$ para $\mathbb{R}^3$ e $C = \{(0,1),(2,2)\}$ para $\mathbb{R}^2$, tínhamos a representação matricial
$$T_{CB} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & \frac{1}{2} \end{bmatrix}$$
para T.

   Vamos agora determinar a matriz $T_{C'B'}$ que representa T em relação às bases canónicas B' de $\mathbb{R}^3$ e C' de $\mathbb{R}^2$. Para tal, vamos usar a relação
$$T_{C'B'} = P_{C'C}T_{CB}P_{BB'}$$
onde $P_{C'C}$ e $P_{BB'}$ são as matrizes de mudança de base adequadas.

   A matriz $P_{C'C}$ calcula-se directamente, escrevendo nas suas colunas os vectores de C em coordenadas de C'. Uma vez que C' é a base canónica de $\mathbb{R}^2$, temos que
$$P_{C'C} = \begin{bmatrix} 0 & 2 \\ 1 & 2 \end{bmatrix}.$$

   Para calcular $P_{BB'}$, vamos escrever a matriz $P_{B'B}$ (cujas colunas são os vectores de B em coordenadas de B') e invertê-la usando uma das técnicas para inverter matrizes de dimensão 3×3. Obtemos
$$P_{BB'} = (P_{B'B})^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

   Temos então que
$$\begin{aligned} T_{C'B'} = P_{C'C}T_{CB}P_{BB'} &= \begin{bmatrix} 0 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 & 1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \end{aligned}$$

3. Num exemplo anterior, vimos que a transformação linear $T : P_1 \to P_2$ definida por $T(p(x)) = p(x) + (1+x^2)p'(x)$ era representada em relação às bases canónicas B de P_1 e C de P_2 por
$$T_{CB} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

   Dadas agora as bases $B' = \{2, 3+2x\}$ e $C' = \{1+x, 1+x+x^2, 1-x^2\}$ de P_1 e P_2, respectivamente, vamos determinar a representação de T nessas bases. Em vez de construir esta representação pela definição, podemos mais uma vez recorrer às matrizes de mudança de base visto já conhecermos a representação de T em relação às bases canónicas.
$$T_{C'B'} = P_{C'C}T_{CB}P_{BB'}$$

Novamente, a matriz $P_{BB'}$ é fácil de obter, pois precisamos apenas de escrever as coordenadas dos polinómios de B' na base canónica.

$$P_{BB'} = \begin{bmatrix} \vdots & \vdots \\ (2)_B & (3+2x)_B \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$$

Para determinar $P_{C'C}$, vamos usar outra vez o facto de $P_{CC'}$ transformar coordenadas da base C' em coordenadas da base canónica, sendo portanto mais simples de obter.

$$P_{CC'} = \begin{bmatrix} \vdots & \vdots & \vdots \\ (1+x)_C & (1+x+x^2)_C & (1-x^2)_C \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

Tal como antes, $P_{C'C} = (P_{CC'})^{-1}$. Calculando esta inversa pela regra dos cofactores ou pelo método de Gauss–Jordan, obtemos

$$(P_{CC'})^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

Assim

$$\begin{aligned} T_{C'B'} = P_{C'C}T_{CB}P_{BB'} &= \begin{bmatrix} -1 & 2 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 2 & 5 \\ 2 & 3 \end{bmatrix}. \end{aligned}$$

Exercício 12. Para cada uma das alíneas seguintes, escreva as matrizes de mudança de base das bases B e C indicadas relativamente à base canónica. Usando essas matrizes, escreva a matriz T_{CB} que representa a transformação linear T indicada em relação a essas bases.

(a) $B = C = \{(2,1), (-3,4)\}$
$T(x,y) = (x-2y, -y)$

(b) $B = C = \{(1,0,0), (1,1,0), (1,1,1)\}$
$T(x,y,z) = (x+2y-z, -y, x+7z)$

(c) $B = \{2, 3+2x\}$
$C = \{1+x, 1+x+x^2, 1-x^2\}$
$T(p(x)) = p(x) + (1+x^2)p'(x)$

3.3 Operações algébricas

Tal como podemos realizar operações algébricas com funções, também faz sentido definir algumas operações entre transformações lineares — nomeadamente, mais uma vez, aquelas operações que estamos a considerar desde o início: soma e produto por escalares. Por outro lado, e uma vez que é de funções que estamos a falar, também podemos efectuar a composição de uma transformação linear com imagens em W com outra que receba argumentos em W.

Nesta secção vamos discutir estas operações. Teremos em simultâneo ocasião de comprovar que a representação matricial definida na secção anterior é de facto uma ferramenta poderosa: as operações sobre transformações lineares correspondem de forma muito directa a operações sobre matrizes.

3.3.1 Operações lineares

A soma e produto por escalar de transformações lineares definem-se de forma natural usando as operações no espaço de chegada, tal como é hábito no caso geral de soma de funções.

Definição. Sejam $T, S : V \to W$ duas transformações lineares entre os mesmos espaços vectoriais V e W, e seja α um real.

A *soma* de T com S é a transformação $(T + S) : V \to W$ tal que

$$(T + S)(v) = Tv + Sv$$

para qualquer $v \in V$ (*soma pontual* de V e W).

O *produto* de α por T é a transformação $\alpha T : V \to W$ tal que

$$(\alpha T)(v) = \alpha\, Tv$$

para qualquer $v \in V$.

Note-se que, à partida, nada garante que $T + V$ e αT sejam transformações lineares, já que, por definição, apenas podemos garantir que são funções entre V e W. Contudo, podemos verificar facilmente que não só o são, como têm uma representação matricial definida à custa das representações matriciais de T e S da forma natural.

Proposição 50. Sejam $T : V \to W$ e $S : V \to W$ transformações lineares, α um real e B e C bases de V e W, respectivamente. Então:

1. $T + S$ é transformação linear;
2. $(T + S)_{CB} = T_{CB} + S_{CB}$;
3. αT é transformação linear;
4. $(\alpha T)_{CB} = \alpha T_{CB}$.

Demonstração.

1. Tomando dois vectores arbitrários $u, v \in V$ e dois reais β, γ arbitrários, temos que

$$\begin{aligned}
(T + S)(\beta u + \gamma v) &= T(\beta u + \gamma v) + S(\beta u + \gamma v) && \text{definição de } (T + S) \\
&= \beta Tu + \gamma Tv + \beta Su + \gamma Sv && T \text{ e } S \text{ são lineares} \\
&= \beta(Tu + Su) + \gamma(Tv + Sv) && \text{propriedades da soma em } W \\
&= \beta(T + S)(u) + \gamma(T + S)(v) && \text{definição de } (T + S)
\end{aligned}$$

donde $T + S$ é transformação linear.

2. Sejam T_{CB} e S_{CB} as matrizes que representam T e S em relação às bases B e C de V e W, respectivamente, e escolha-se $v \in V$ arbitrário. Então tem-se

$$\begin{aligned}
((T+S)(v))_C &= (Tv+Sv)_C && \text{definição de } (T+S)\\
&= (Tv)_C + (Sv)_C && \text{propriedades das coordenadas em } W\\
&= T_{CB}(v)_B + S_{CB}(v)_B && \text{definição de representação matricial}\\
&= (T_{CB}+S_{CB})\,(v)_B && \text{distributividade do produto de matrizes}
\end{aligned}$$

donde $T_{CB}+S_{CB}$ é a representação matricial de $T+S$ em relação às bases B e C dos espaços V e W.

3. Tomando novamente dois vectores arbitrários $u, v \in V$ e dois reais β, γ arbitrários, temos

$$\begin{aligned}
(\alpha T)(\beta u+\gamma v) &= \alpha T(\beta u+\gamma v) && \text{definição de } (\alpha T)\\
&= \alpha(\beta Tu)+\alpha(\gamma Tv) && T \text{ é transformação linear}\\
&= \beta(\alpha Tu)+\gamma(\alpha Tv) && \text{propriedades das operações em } W\\
&= \beta(\alpha T)(u)+\gamma(\alpha T)(v) && \text{definição de } (\alpha T)
\end{aligned}$$

donde αT também é transformação linear.

4. Seja T_{CB} a matriz que representa T em relação às bases B e C de V e W, respectivamente, e escolha-se $v \in V$ arbitrário. Então tem-se

$$\begin{aligned}
((\alpha T)(v))_C &= (\alpha Tv)_C && \text{definição de } (\alpha T)\\
&= \alpha(Tv)_C && \text{propriedades das coordenadas em } W\\
&= \alpha\,(T_{CB}(v)_B) && \text{definição de representação matricial}\\
&= (\alpha T_{CB})\,(v)_B && \text{associatividade do produto de matrizes}
\end{aligned}$$

donde αT_{CB} representa αT em relação às bases B e C de V e W. □

Sejam $T, S : \mathbb{R}^2 \to \mathbb{R}^3$ as transformações lineares definidas por

$$T(x,y) = (2x+y, x-y, 0) \qquad S(x,y) = (x-2y, x+y, 2x)\,.$$

A sua soma $T+S$ é a transformação linear tal que

$$\begin{aligned}
(T+S)(x,y) &= T(x,y)+S(x,y)\\
&= (2x+y,x-y,0)+(x-2y,x+y,2x) = (3x-y,2x,2x)\,.
\end{aligned}$$

Tomando as bases canónicas B de $\mathbb{R}^2$ e C de $\mathbb{R}^3$, encontramos as seguintes representações matriciais de T e de S.

$$T_{CB} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix} \qquad S_{CB} = \begin{bmatrix} 1 & -2 \\ 1 & 1 \\ 2 & 0 \end{bmatrix}$$

Somando estas duas matrizes, obtemos

$$T_{CB}+S_{CB} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 2 & 0 \end{bmatrix}$$

enquanto que, calculando directamente $(T+S)_{CB}$, obtemos

$$(T+S)_{CB} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \\ 2 & 0 \end{bmatrix},$$

que é a matriz calculada anteriormente.

Considerando ainda a mesma transformação T e tomando $\alpha = 3$, temos que $\alpha T = 3T$ é a transformação linear tal que

$$(3T)(x,y) = 3T(x,y) = (6x+3y, 3x+3y, 0)$$

cuja representação matricial em relação às bases canónicas B e C de $\mathbb{R}^2$ e $\mathbb{R}^3$ é

$$(3T)_{CB} = \begin{bmatrix} 6 & 3 \\ 3 & 3 \\ 0 & 0 \end{bmatrix} = 3T_{CB}\,.$$

O resultado anterior implica a seguinte relação entre transformações lineares e combinações lineares.

Proposição 51. Sejam $T_1, T_2, \ldots, T_n : V \to W$ transformações lineares e $c_1, c_2, \ldots, c_n$ números reais. Então $c_1T_1+c_2T_2+\cdots+c_nT_n$ é uma transformação linear de V para W cuja representação matricial em relação a bases B e C destes espaços é

$$(c_1T_1 + c_2T_2 + \cdots + c_nT_n)_{CB} = c_1\,(T_1)_{CB} + c_2\,(T_2)_{CB} + \cdots + c_n\,(T_n)_{CB}\,.$$

A título de exemplo, este resultado implica que a transformação $5T-2S$, com T e S como atrás, é uma transformação linear entre $\mathbb{R}^2$ e $\mathbb{R}^3$, com representação matricial em relação às bases canónicas destes espaços dada por

$$(5T-2S)_{CB} = 5T_{CB} - 2S_{CB} = 5\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix} - 2\begin{bmatrix} 1 & -2 \\ 1 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 9 \\ 3 & -7 \\ -4 & 0 \end{bmatrix}.$$

Exemplo.

1. Sejam $T, S : P_1 \to P_2$ as tranformações lineares definidas por

   $$T(p(x)) = p(x) + \left(1+x^2\right)p'(x) \qquad S(p(x)) = xp(x)\,.$$

   Podemos definir a soma destas transformações como sendo $T+S : P_1 \to P_2$ com expressão

   $$\begin{aligned}(T+S)(p(x)) &= T(p(x)) + S(p(x)) = p(x) + \left(1+x^2\right)p'(x) + xp(x) \\ &= (1+x)p(x) + \left(1+x^2\right)p'(x)\,.\end{aligned}$$

   As transformações lineares T e S são representadas em relação às bases canónicas B de P_1 e C de P_2 pelas matrizes

   $$T_{CB} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ e } S_{CB} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Assim, a representação matricial de $T+S$ em relação às bases canónicas de P_1 e P_2 é dada por

$$(T+S)_{CB} = T_{CB} + S_{CB} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 2 \end{bmatrix},$$

conforme se pode comprovar calculando directamente $(T+S)_{CB}$ através da definição.

2. Com a mesma transformação T do exemplo acima, a transformação $-T = (-1)T$ é a transformação entre P_1 e P_2 definida por

$$(-T)(p(x)) = -T(p(x)) = -p(x) - \left(1+x^2\right) p'(x)\,.$$

A representação matricial desta transformação em relação às bases canónicas B de P_1 e C de P_2 é

$$(-T)_{CB} = \begin{bmatrix} -1 & -1 \\ 0 & -1 \\ 0 & -1 \end{bmatrix} = -T_{CB}\,,$$

confirmando mais uma vez o resultado estabelecido atrás.

3. Consideremos agora transformações lineares $T, S : M_{2\times2} \to \mathbb{R}$ definidas por $TA = \operatorname{tr}(A)$ e $SA = a_{12} - a_{21}$. Consideremos a combinação linear $3T - 2S : M_{2\times2} \to \mathbb{R}$, que também é uma transformação linear; esta transformação tem uma representação matricial em relação às bases canónicas B e C de $M_{2\times2}$ e $\mathbb{R}$ que é obtida pela correspondente combinação linear das representações matriciais de T e S, isto é,

$$\begin{aligned}(3T-2S)_{CB} &= 3T_{CB} - 2S_{CB} \\ &= 3\begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} - 2\begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 2 & 3 \end{bmatrix}.\end{aligned}$$

Daqui podemos concluir que a expressão de $(3T-2S)$ é

$$(3T-2S)\begin{bmatrix} a & b \\ c & d \end{bmatrix} = 3a - 2b + 2c + 3d\,,$$

expressão equivalente (mas mais útil na prática) à que obteríamos calculando directamente $(3T-2S)(A) = 3TA - 2SA = 3\operatorname{tr}(A) - 2a_{12} + 2a_{21}$.

Com base nestes resultados, pode-se ainda observar que as transformações lineares entre V e W formam um espaço linear equivalente ao espaço linear $M_{m\times n}$, identificando cada transformação linear com a sua representação matricial em bases fixas daqueles espaços. Das propriedades da soma de matrizes e do produto de matrizes por escalares obtemos directamente propriedades análogas para as correspondentes operações sobre transformações lineares; em particular, a soma de transformações lineares é associativa e comutativa e tem a transformação nula como elemento neutro. Dito de outra forma, o conjunto das transformações lineares entre dois espaços vectoriais V e W é ele próprio um espaço vectorial.

Exercício 13. Considere as seguintes transformações lineares.

$T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ $T_1(x,y) = (x+y, x-y)$	$T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ $T_2(x,y) = (3x-2y, 2x)$	$T_3 : \mathbb{R}^2 \to \mathbb{R}^3$ $T_3(x,y) = (x+2y, -x, 0)$
$T_4 : M_{2\times2} \to M_{2\times2}$ $T_4A = A^T$	$T_5 : M_{2\times2} \to M_{2\times2}$ $T_5 \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2c & a+c \\ b-2c & d \end{bmatrix}$	$T_6 : M_{2\times2} \to M_{2\times2}$ $T_6X = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} X + X \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$
$T_7 : P_2 \to P_1$ $T_7(p(x)) = p'(x)$	$T_8 : P_2 \to P_2$ $T_8(p(x)) = p(x-1)$	$T_9 : P_2 \to P_3$ $T_9(p(x)) = xp(x)$

(a) Indique para que valores de n e m é que $T_n + T_m$ está definida e calcule a sua expressão.

(b) Verifique o resultado da alínea anterior recorrendo à representação matricial das transformações em causa em relação às bases canónicas dos espaços envolvidos.

3.3.2 Composição de transformações lineares

A composição de transformações lineares é simplesmente uma composição de funções, de acordo com o esquema da Figura 3.6.

$$\begin{array}{ccccc} V & \xrightarrow{T} & U & \xrightarrow{S} & W \\ \\ v & \longrightarrow & Tv & \longrightarrow & STv \end{array}$$

Figura 3.6: Composição de transformações lineares.

À regra "calcular Tv e aplicar-lhe S" corresponde uma função, designada por *composição* de T com S. Mais uma vez, esta função vai ser uma transformação linear.

Definição. Sejam $T : V \to U$ e $S : U \to W$ transformações lineares. A *composição* de T com S é a transformação $S \circ T$ (também chamada de S *após* T) entre V e W definida por $(S \circ T)(v) = S(Tv)$ para todo $v \in V$.

Antes de vermos exemplos, vamos apenas verificar as propriedades desta operação.

Proposição 52. Sejam $T : V \to U$ e $S : U \to W$ transformações lineares e B, C e D bases de V, U e W, respectivamente. Então $S \circ T$ é uma transformação linear e a sua representação em relação às bases B e D é

$$(S \circ T)_{DB} = S_{DC} T_{CB} ,$$

sendo S_{DC} e T_{CB} as representações matriciais das transformações S e T, respectivamente, em relação às bases dos espaços correspondentes.

Demonstração. A verificação destes factos é muito simples. Por um lado, tem-se que, para $u, v \in V$ arbitrários e α, β reais,

$$\begin{aligned}
(S \circ T)(\alpha u + \beta v) &= S(T(\alpha u + \beta v)) && \text{definição de } S \circ T \\
&= S(\alpha T u + \beta T v) && T \text{ é transformação linear} \\
&= \alpha(S(Tu)) + \beta(S(Tv)) && S \text{ é transformação linear} \\
&= \alpha((S \circ T)(u)) + \beta((S \circ T)(v)) && \text{definição de } S \circ T
\end{aligned}$$

donde $S \circ T$ é uma transformação linear. Por outro lado, sendo T_{CB} e S_{DC} como no enunciado, tem-se que

$$\begin{aligned}
((S \circ T)(v))_D &= (S(Tv))_D && \text{definição de } S \circ T \\
&= S_{DC}(Tv)_C && \text{definição de representação} \\
&= S_{DC}\left(T_{CB}(v)_B\right) && \text{definição de representação} \\
&= \left(S_{DC}T_{CB}\right)(v)_B && \text{associatividade do produto de matrizes}
\end{aligned}$$

para $v \in V$ arbitrário, donde $S_{DC}T_{CB}$ é a representação matricial de $S \circ T$ em relação às bases B e D. □

Este resultado ilustra mais uma vez a íntima relação entre transformações lineares e a sua representação matricial; por outro lado, fornece mais uma prova de que a operação de produto de matrizes, embora inicialmente pudesse não ser muito intuitiva, está de facto definida da forma correcta.

Observe-se ainda que a restrição dimensional para o produto de matrizes também faz todo o sentido do ponto de vista da composição de transformações lineares: para que $S \circ T$ esteja definida, é necessário que o espaço de chegada de T seja o espaço de partida de S. Ora a dimensão do espaço de chegada de T é o número de linhas de T_{CB} (já que cada coluna contém as coordenadas dum vector desse espaço), enquanto que a dimensão do espaço de partida de S é o número de colunas de S_{DC} (já que as colunas desta matriz correspondem a uma base de S). Obtemos mais uma confirmação de que a restrição dimensional do produto de matrizes não é uma arbitrariedade, estando antes intimamente relacionada com as diferentes transformações que podem ser capturadas através desta operação.

Vejamos um exemplo. Sejam $T : \mathbb{R}^3 \to \mathbb{R}^2$ e $S : \mathbb{R}^2 \to \mathbb{R}^4$ as transformações lineares definidas por

$$T(x, y, z) = (2x - z, x + y + z) \qquad S(x, y) = (x + y, x - y, 2x - y, -3x)\,.$$

Para calcular directamente a composição de T com S, recorremos à definição. Uma vez que $S \circ T : \mathbb{R}^3 \to \mathbb{R}^4$, vamos escolher um vector arbitrário (x, y, z) de $\mathbb{R}^3$ e calcular a sua imagem. Este cálculo faz-se em dois passos (ver Figura 3.6): primeiro calculamos $T(x, y, z)$, e depois aplicamos S ao resultado; ou seja,

$$\begin{aligned}
(S \circ T)(x, y, z) &= S\,\underbrace{(2x - z, x + y + z)}_{T(x,y,z)} \\
&= ((2x - z) + (x + y + z), (2x - z) - (x + y + z), \\
&\qquad 2(2x - z) - (x + y + z), -3(2x - z)) \\
&= (3x + y, x - y - 2z, 3x - y - 3z, -6x + 3z)
\end{aligned}$$

A representação matricial de T em relação às bases canónicas B de $\mathbb{R}^3$ e C de $\mathbb{R}^2$ é

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

enquanto que a representação matricial de S em relação às bases canónicas C e D de $\mathbb{R}^2$ e $\mathbb{R}^4$ é dada por

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

O produto destas duas matrizes é

$$S_{DC}T_{CB} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 2 & -1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & -1 & -2 \\ 3 & -1 & -3 \\ -6 & 0 & 3 \end{bmatrix}$$

que corresponde, como facilmente se verifica, à matriz $(S \circ T)_{DB}$.

Este exemplo ilustra outra particularidade especialmente relevante: enquanto que, a nível de somas e produtos por escalares, utilizar a representação matricial não era uma mais-valia especialmente notória (especialmente se apenas dispúnhamos da expressão das transformações lineares a somar ou multiplicar), já no caso da composição a situação é diferente. Por norma, compensa escrever as representações matriciais das transformações envolvidas e calcular o seu produto, eventualmente reescrevendo o resultado em forma de expressão, face a calcular directamente a composição. A razão para isto é um pouco semelhante à razão por que introduzimos a representação matricial de sistemas de equações: a forma como o produto de matrizes é feito garante que a informação fica mais organizada e que os cálculos são feitos de forma sistemática, minimizando a probabilidade de erros.

Dito isto, há excepções. Considerem-se as transformações lineares $T : P_2 \to P_1$ definida por $Tp = p'$ e $S : P_1 \to P_2$ tal que $S(p(x)) = p(x) + (1 + x^2)\, p'(x)$. Para calcular a expressão composição $S \circ T$ basta substituir p por p' (já que é este o valor de Tp) na expressão de Sp, obtendo-se

$$(S \circ T)(p(x)) = p'(x) + \left(1 + x^2\right) p''(x)\,,$$

sendo $S \circ T : P_2 \to P_2$. As matrizes que representam estas transformações em relação às bases canónicas B de P_2 e C de P_1 são

$$T_{CB} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \text{ e } S_{BC} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

cujo produto é

$$S_{BC}T_{CB} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{bmatrix},$$

e embora esta matriz corresponda à representação matricial $(S \circ T)_{BB}$ de $S \circ T$ em relação à base B de P_2, é muito menos informativa do que a representação explícita determinada acima.

Em contrapartida, uma vez que as transformações S e T também são componíveis na ordem inversa, podemos tentar calcular a transformação $T \circ S : P_1 \to P_1$. Pela definição, temos que

$$(T \circ S)(p(x)) = T\underbrace{\left(p(x) + \left(1 + x^2\right) p'(x)\right)}_{S(p(x))} = \left(p(x) + \left(1 + x^2\right) p'(x)\right)' .$$

Aqui o cálculo desta expressão não é trivial. Recorrendo à representação matricial, porém, temos que

$$\begin{aligned}(T \circ S)_{CC} &= T_{CB} S_{BC} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix},\end{aligned}$$

donde $(T \circ S)(a + bx) = b + 2bx$, e neste caso beneficiámos muito de recorrer à representação matricial.

Exercício 14. Considere novamente as transformações lineares do Exercício 13.

(a) Indique para que valores de n e m é que $T_n \circ T_m$ está definida e calcule a sua expressão.

(b) Verifique o resultado da alínea anterior recorrendo à representação matricial das transformações em causa em relação às bases canónicas dos espaços envolvidos.

Os resultados desta secção expressam um paralelismo entre transformações lineares e matrizes. A tabela seguinte resume a correspondência entre operações sobre transformações lineares e operações entre as suas representações.

Transformação	Representação
$T + S$	$T_{CB} + S_{CB}$
αT	αT_{CB}
$S \circ T$	$S_{DC} T_{CB}$

Tabela 3.1: Correspondência entre operações sobre transformações lineares e operações sobre matrizes.

Exercício 15. Sejam $T_1, T_2 : V \to U$ e $S_1, S_2 : U \to W$ transformações lineares cujas representações em relação às bases B, C e D de V, U e W, respectivamente, são T_{1CB}, T_{2CB}, S_{1DC} e S_{2DC}. Escreva as expressões que caracterizam as representações matriciais das seguintes transformações lineares.

(a) $T_1 + 2T_2$

(b) $T_2 - 3T_1$

(c) $2S_1 - 2S_2$

(d) $S_1 \circ T_1$

(e) $S_2 \circ T_1$

(f) $S_2 \circ (T_1 + T_2)$

(g) $(S_2 + 2S_1) \circ (3T_2)$

(h) $(3S_1 - 2S_2) \circ (T_1 - T_2)$

(i) $(2S_2 + 5S_1) \circ (2T_1 + T_2)$

Exercício 16. Nas condições do Exercício 15, suponha que as matrizes T_{1CB}, T_{2CB}, S_{1DC} e S_{2DC} são as seguintes.

$$T_{1CB} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 2 & 0 \end{bmatrix} \qquad T_{2CB} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$S_{1DC} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \qquad S_{2DC} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

(a) Calcule a representação matricial de cada uma das transformações definidas nesse exercício.

(b) Considere agora que $V = \mathbb{R}^2$, $U = \mathbb{R}^3$ e $W = \mathbb{R}^2$, com B, C e D as bases canónicas destes espaços. Escreva as expressões dessas transformações recorrendo às matrizes acima determinadas.

(c) Considere agora que $V = \mathbb{R}^2$, $U = \mathbb{R}^3$ e $W = L\{(1,0,1,0),(0,1,0,-1) \subseteq \mathbb{R}^4\}$, com B e C as bases canónicas de V e U e D a base $\{(1,0,1,0),(0,1,0,-1)\}$ de W. Escreva as expressões dessas transformações recorrendo às matrizes acima determinadas.

(d) Considere agora que $V = P_1$, $U = P_2$ e $W = P_1$. Escreva as expressões dessas transformações recorrendo às matrizes acima determinadas.

(e) Considere agora que V, U e W são os conjuntos de polinómios gerados respectivamente pelas bases $B = \{1, x^2\}$, $c = \{1, 1+x, 1+x+x^2\}$ e $D = \{x, x^2\}$. Escreva as expressões dessas transformações recorrendo às matrizes acima determinadas.

(f) Considere agora que V, U e W os conjuntos de matrizes com bases

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \right\}$$
$$C = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$
$$D = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$$

Escreva as expressões dessas transformações recorrendo às matrizes acima determinadas.

Nas próximas secções teremos ocasião de ver outras operações sobre transformações lineares que também têm uma correspondência directa a nível das suas representações.

3.4 Núcleo e imagem

Nas secções anteriores vimos as propriedades das transformações lineares enquanto funções que respeitam a estrutura de espaço vectorial existente nos conjuntos em que são aplicadas. Nesta secção vamos estudar outro tipo de propriedades relacionadas com a linearidade: subespaços dos conjuntos de partida e de chegada que estão associados a uma transformação linear.

3.4.1 Núcleo

Qualquer espaço vectorial W tem um elemento 0_W, elemento neutro da soma, que é um elemento especial do ponto de vista da estrutura linear de W. Dada uma transformação linear $T : V \to W$, o conjunto dos elementos de V que são mapeados em 0_W é um subconjunto importante de V — relacionado, como veremos, com a solução do sistema homogéneo associado à representação matricial de T, no caso em que V e W têm dimensão finita.

Definição. Seja $T : V \to W$ uma transformação linear. O *núcleo* de T é o conjunto $\mathcal{N}(T)$ ou $\mathrm{Ker}(T)$ dos elementos v de V tais que $Tv = 0_W$.

Simbolicamente, podemos escrever

$$\mathcal{N}(T) = \{v \in V \mid Tv = 0_W\} .$$

A primeira observação a fazer é que combinações lineares de vectores no núcleo duma transformação linear $T : V \to W$ geram vectores que ainda estão no núcleo de T. De facto, se u e v são vectores em $\mathcal{N}(T)$, então $Tu = Tv = 0_W$, donde

$$T(\alpha u + \beta v) = \alpha Tu + \beta Tv = \alpha 0_W + \beta 0_W = 0_W .$$

Concluímos portanto que $\mathcal{N}(T)$ é um subespaço vectorial de V. À dimensão deste espaço chamamos *nulidade* de T: $\mathrm{nul}(T) = \dim \mathcal{N}(T)$.

Vejamos dois casos muito simples. Se $T : V \to W$ for a transformação nula, definida por $Tv = 0_W$ para qualquer $v \in V$, então automaticamente todos os vectores de V estão no núcleo de T, donde $\mathcal{N}(T) = V$ e $\mathrm{nul}(T) = \dim(V)$. Por outro lado, se $T : V \to V$ for a transformação identidade $\mathbf{I_V}$ que aplica cada vector nele próprio, então o único vector com imagem nula é o próprio vector 0_V, donde $\mathcal{N}(T) = \{0_V\}$ e $\mathrm{nul}(T) = 0$. Estes dois casos geram portanto como núcleo os subespaços triviais do espaço vectorial de partida V.

Em geral, a determinação do núcleo duma transformação linear implica perceber um pouco do funcionamento da transformação.

Exemplo.

1. Seja $T : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$ a transformação linear que transforma cada função na sua derivada. Então o núcleo de T são as funções f tais que $Tf = 0$, ou seja, tais que $f' = 0$. Dos resultados da Análise, sabemos que estas funções são as funções constantes, que são precisamente os polinómios de grau zero. Ou seja: $\mathcal{N}(T) = P_0$ e $\mathrm{nul}(T) = 1$.

2. Seja $T : \mathcal{P} \to \mathcal{P}$ a transformação linear que transforma cada polinómio na sua segunda derivada. Então o núcleo de T são os polinómios p cuja segunda derivada p'' é nula. Ora se $p'' = 0$ sabemos que p' é constante, donde p é um polinómio de primeiro grau (ou seja, $p(x) = a + bx$ para determinadas constantes a e b). Então $\mathcal{N}(T) = P_1$ e $\mathrm{nul}(T) = 2$.

3. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear que transforma cada vector na sua projecção sobre a recta $y = x$. Então o núcleo de T são os vectores cuja projecção ortogonal sobre essa recta é a origem. Uma vez que a projecção ortogonal dum ponto P sobre uma recta R é calculada (geometricamente) como a intersecção de R com a recta perpendicular a R que passa por P, concluímos que os pontos no núcleo de T são os pontos sobre a recta perpendicular a $y = x$ que passa pela origem (ver Figura 3.7), ou seja, os pontos da recta $y = -x$. Então $\mathcal{N}(T) = L(\{(1, -1)\})$ e $\mathrm{nul}(T) = 1$.

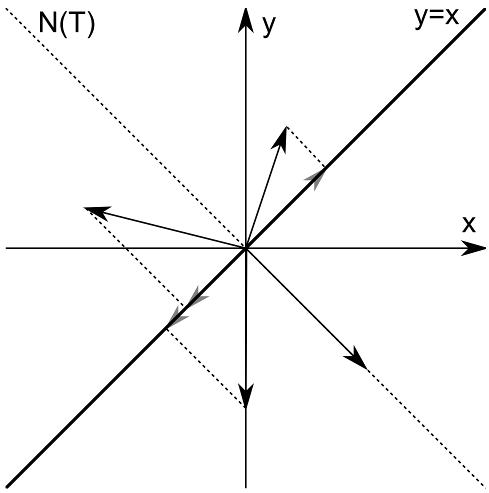


Figura 3.7: Núcleo da projecção ortogonal sobre a recta $y = x$.

4. Analogamente, se $T : \mathbb{R}^3 \to \mathbb{R}^3$ for a transformação linear que transforma cada vector na sua projecção ortogonal sobre o plano $x + y + z = 0$, o mesmo raciocínio permite concluir que o núcleo de T é o conjunto dos vectores sobre a recta perpendicular a esse plano que passa pela origem. Esta recta é a recta com vector director $(1, 1, 1)$, ou seja, $\mathcal{N}(T) = L(\{(1, 1, 1)\})$ e $\text{nul}(T) = 1$.

Em geral, porém, a determinação do núcleo reduz-se à resolução dum sistema homogéneo de equações lineares. O caso em que isto é mais óbvio é o caso duma transformação matricial $T : \mathbb{R}^m \to \mathbb{R}^n$ definida por $T\overrightarrow{u} = A_{m\times n}\underline{u}$. Por definição, o núcleo de T é o conjunto dos vectores $\overrightarrow{u}$ tais que $T\overrightarrow{u} = \overrightarrow{0}$, ou seja, tais que $A\underline{u} = \underline{0}$; este conjunto é precisamente o núcleo da matriz A, tendo-se portanto a relação

$$\mathcal{N}(T) = \mathcal{N}(A) .$$

Veremos de seguida que esta situação específica é muito semelhante ao caso geral.

Exemplo. Seja $T : P_1 \to P_2$ a transformação linear definida por

$$T(p(x)) = p(x) + \left(1 + x^2\right) p'(x) .$$

O núcleo de T é o conjunto dos polinómios p tais que $Tp = 0$. Uma vez que p tem grau 1, podemos escrevê-lo na forma $p(x) = a+bx$; a condição anterior transforma-se então na seguinte.

$$T(p(x)) = 0 \longrightarrow p(x) + \left(1 + x^2\right) p'(x) = 0 \longrightarrow a + bx + \left(1 + x^2\right) b = 0$$

$$\longrightarrow a + b + bx + bx^2 = 0 \longrightarrow \begin{cases} a + b = 0 \\ b = 0 \\ b = 0 \end{cases} \longrightarrow \begin{cases} a = 0 \\ b = 0 \end{cases}$$

Logo se $Tp = 0$ é porque $p = 0$; então $\mathcal{N}(T) = \{0\}$.

Vejamos como podemos usar uma qualquer representação matricial de T para obter este mesmo resultado. Conforme vimos num exemplo anterior, a representação de T em relação às bases canónicas B e C de P_1 e P_2 é dada pela matriz

$$T_{CB} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} ,$$

tendo-se ainda a relação

$$(Tp)_C = T_{CB}(p)_B\,.$$

Ora o polinómio nulo tem todas as coordenadas nulas (em qualquer base), pelo que

$$\begin{aligned}
&T(p(x)) = 0 \longrightarrow T_{CB}(p(x))_B = (0)_C\\
&\qquad\longrightarrow \begin{bmatrix} 1 & 1\\ 0 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix} a\\ b\end{bmatrix} = \begin{bmatrix} 0\\ 0\\ 0\end{bmatrix} \longrightarrow \begin{bmatrix} a+b\\ b\\ b\end{bmatrix} = \begin{bmatrix} 0\\ 0\\ 0\end{bmatrix} \longrightarrow \begin{cases} a=0\\ b=0\end{cases}
\end{aligned}$$

O raciocínio no final do exemplo anterior aplica-se em qualquer situação. Concretamente, seja $T : V \to W$ uma transformação linear cuja representação em relação às bases B e C de V e W, respectivamente, é T_{CB}. Então $v \in V$ pertence ao núcleo de T precisamente quando se verificam as seguintes condições.

$$\begin{aligned}
v \in \mathcal{N}(T) &\Longleftrightarrow Tv = 0_W && \text{definição de núcleo}\\
&\Longleftrightarrow (Tv)_C = (0_W)_B &&\\
&\Longleftrightarrow T_{CB}(v)_B = \underline{0} && \text{definição de representação matricial}\\
&\Longleftrightarrow (v)_B \in \mathcal{N}(T_{CB}) && \text{definição de núcleo duma matriz}
\end{aligned}$$

Concluímos então que $\mathcal{N}(T) = \mathcal{N}(T_{CB})$, donde, como consequência, $\mathrm{nul}(T) = \mathrm{nul}(T_{CB})$, identificando na primeira igualdade os vectores de V com as suas coordenadas em relação à base B.

Exercício 17. Considere as transformações lineares $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ definidas por

$$T_1(x,y) = (2x - y, -8x + 4y) \qquad T_2(x,y) = (x + y, x - y)\,.$$

(a) Verifique se os vectores $(0,1)$, $(1,-4)$, $(1,2)$, $(5,0)$ e $(-2,-3)$ pertencem ao núcleo de cada uma das transformações.

(b) Determine bases para $N(T_1)$ e $N(T_2)$.

Exercício 18. Considere as transformações lineares $T_1, T_2 : P_3 \to P_3$ definidas por

$$T_1(p(x)) = p'(x) \qquad T_2(p(x)) = (x+1)p''(x)\,.$$

(a) Verifique se os vectores $x^2 + 3x$, $2x^3 - 1$, 2, $3x - 1$ e $3x^3 + 2x - 1$ pertencem ao núcleo de cada uma das transformações.

(b) Determine bases para $N(T_1)$ e $N(T_2)$.

Exercício 19. Encontre o núcleo de cada uma das transformações lineares consideradas nos Exercícios 7 e 8 recorrendo às suas representações matriciais.

3.4.2 Imagem

Ao definirmos uma transformação linear $T : V \to W$, os elementos do espaço vectorial W ficam divididos em duas categorias: aqueles que são da forma Tv, para algum $v \in V$, e os restantes.

Definição. Seja $T : V \to W$ uma transformação linear. Chama-se *imagem* ou *contradomínio* de T ao conjunto $\mathrm{im}(T)$ ou $T(V)$ dos vectores de W que são da forma Tv, para algum $v \in V$.

Simbolicamente,

$$T(V) = \{w \in W \mid T(v) = w \text{ para algum } v \in V\} .$$

O termo *contradomínio* é o termo usado de forma genérica para o conjunto de valores de qualquer função. No contexto da Álgebra Linear, é mais usual usar-se a designação de *imagem*, pelo que será esta que adoptaremos daqui em diante.

Tal como sucedia com o núcleo, também a imagem de T é um espaço vectorial, agora subespaço de W. Para tal, vamos aplicar a linearidade de T no sentido inverso ao habitual. Se w_1 e w_2 forem vectores de $T(V)$, então por definição de imagem existem vectores v_1 e v_2 em V tais que $Tv_1 = w_1$ e $Tv_2 = w_2$. Tomando uma combinação linear $\alpha w_1 + \beta w_2$ destes vectores, temos que

$$\alpha w_1 + \beta w_2 = \alpha (Tv_1) + \beta (Tv_2) = T (\alpha v_1) + T (\beta v_2) = T (\alpha v_1 + \beta v_2) ,$$

donde $\alpha w_1 + \beta w_2$ é ainda um elemento da imagem de T.

Concluímos então que $T(V)$ é subespaço linear de W e à dimensão deste espaço chamamos *característica* de T: $\mathrm{car}(T) = \dim(T(V))$.

Consideremos para começar os dois casos muito simples de há pouco. Para a transformação nula $T : V \to W$, temos que $Tv = 0_W$ para qualquer $v \in V$; então o único vector que é da forma Tv para algum v é o vector nulo, donde $T(V) = \{0_W\}$ e $\mathrm{car}(T) = 0$. Por outro lado, para a transformação identidade $\mathbf{I_W} : W \to W$ definida por $\mathbf{I_W} w = w$, temos que todos os vectores são imagem de algum vector (nomeadamente, são imagem deles próprios), pelo que $T(W) = W$ e $\mathrm{car}(T) = \dim(W)$. Estas duas transformações geram portanto como imagem os subespaços triviais do espaço vectorial de chegada W.

Tal como antes, o cálculo da imagem duma transformação linear requer que se comece por perceber o seu funcionamento. Voltemos a analisar as transformações lineares discutidas atrás.

Exemplo.

1. Seja $T : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$ a transformação linear que transforma cada função na sua derivada. Sabemos da Análise Matemática que toda a função contínua f tem uma primitiva F tal que $F' = f$, ou seja, $TF = f$. Então toda a função em $\mathcal{C}^0(\mathbb{R})$ é a derivada de alguma função, pelo que $T(\mathcal{C}^1(\mathbb{R})) = \mathcal{C}^0(\mathbb{R})$ e $\mathrm{car}(T) = \infty$.

2. Seja $T : \mathcal{P} \to \mathcal{P}$ a transformação linear que transforma cada polinómio na sua segunda derivada. Novamente, para cada polinómio p conseguimos encontrar um outro polinómio P cuja segunda derivada é p: basta primitivar p duas vezes. Assim, $T(\mathcal{P}) = \mathcal{P}$ e $\mathrm{car}(T) = \infty$.

3. Não é contudo verdade que, em geral, a imagem duma transformação linear seja todo o espaço de chegada. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear que transforma cada vector

na sua projecção sobre a recta $y = x$. Então qualquer vector de $\mathbb{R}^2$ é transformado num vector daquela recta, pelo que só esses vectores é que podem estar na imagem de T. Reciprocamente, qualquer vector da forma (x, x) é a sua própria imagem por T; então $T\left(\mathbb{R}^2\right) = L(\{(1,1)\})$ e $\operatorname{car}(T) = 1$ (ver Figura 3.7).

4. Analogamente, se $T : \mathbb{R}^3 \to \mathbb{R}^3$ for a transformação linear que transforma cada vector na sua projecção ortogonal sobre o plano $x + y + z = 0$, o mesmo raciocínio permite concluir que a imagem de T é precisamente esse plano. Para obter uma base desse plano, precisamos de encontrar dois vectores linearmente independentes que satisfaçam a equação $x+y+z = 0$. Tomando $x = 1$ e $y = 0$, obtemos o vector $(1, 0, -1)$; tomando $x = 0$ e $y = -1$, obtemos $(0, 1, -1)$. Então $T\left(\mathbb{R}^3\right) = L(\{(1, 0, -1), (0, 1, -1)\})$ e $\operatorname{car}(T) = 2$.

Mais uma vez, é possível em muitas situações reduzir o cálculo da imagem duma transformação linear T a um problema de cálculo matricial sobre a representação de T. Consideremos novamente o exemplo simples da transformação matricial $T : \mathbb{R}^n \to \mathbb{R}^m$ com $T\overrightarrow{u} = A_{m\times n}\underline{u}$. A imagem de T é o conjunto dos vectores que são da forma $A_{m\times n}\underline{u} = \underline{b}$ para algum $\overrightarrow{u}$; então estamos precisamente a falar dos vectores $\overrightarrow{v}$ para os quais o sistema $A\underline{u} = \underline{v}$ é possível.

Dos resultados do capítulo anterior, conclui-se que $T\left(\mathbb{R}^m\right) = \operatorname{Col}(A)$.

Podemos novamente fazer um raciocínio semelhante no caso geral, desde que $T : V \to W$ seja uma transformação linear entre espaços de dimensão finita. Sejam B e C bases de V e W e T_{CB} a matriz que representa T em relação a essas bases; então um vector $w \in W$ pertence a $T(V)$ se existir $v \in V$ tal que $Tv = w$, ou seja,

$$
\begin{aligned}
w \in T(V) &\iff Tv = w \text{ para algum } v && \text{definição de imagem}\\
&\iff (Tv)_C = (w)_C \text{ para algum } v\\
&\iff T_{CB}(v)_B = (w)_C \text{ para algum } v && \text{definição de representação matricial}\\
&\iff T_{CB}(v)_B = (w)_C \text{ é possível}\\
&\iff (w)_C \in \operatorname{Col}\left(T_{CB}\right) && \text{Proposição 32}
\end{aligned}
$$

donde $T(V) = \operatorname{Col}\left(T_{CB}\right)$ e consequentemente $\operatorname{car}(T) = \operatorname{car}\left(T_{CB}\right)$. Novamente, a primeira igualdade deve ser entendida identificando agora os vectores de W com as suas coordenadas na base C.

Assim, em muitos casos a imagem duma transformação pode ser calculada recorrendo aos resultados sobre o espaço das colunas duma matriz — embora esta forma não seja necessariamente a mais simples.

Vejamos mais um exemplo. Seja $T : P_1 \to P_2$ a transformação linear definda por

$$T(p(x)) = p(x) + \left(1 + x^2\right) p'(x)\,.$$

Para determinar directamente a imagem de T, precisamos de determinar para que polinómios $p(x) \in P_2$ é que $T(q(x)) = p(x)$ é uma equação possível. Sendo $p(x) = a_0 + a_1x + a_2x^2$ e $q(x) = b_0 + b_1x$, temos que

$$
\begin{aligned}
T(q(x)) = p(x) &\longrightarrow q(x) + \left(1 + x^2\right) q'(x) = p(x)\\
&\longrightarrow b_0 + b_1x + \left(1 + x^2\right) b_1 = a_0 + a_1x + a_2x^2\\
&\longrightarrow b_0 + b_1 + b_1x + b_1x^2 = a_0 + a_1x + a_2x^2 \longrightarrow \begin{cases} b_0 + b_1 = a_0\\ b_1 = a_1\\ b_1 = a_2 \end{cases}
\end{aligned}
$$

Para o sistema ser possível temos que ter $a_1 = a_2$, pelas duas últimas equações, ficando os valores de b_0 e b_1 determinados pelo valor de a_0 e $a_1 = a_2$. Concluímos assim que

$$T(P_1) = \left\{p(x) = a_0 + a_1 x + a_2 x^2 \in P_2 \mid a_1 = a_2\right\} = \left\{a + bx + bx^2 \mid a, b \in \mathbb{R}\right\} .$$

Em particular, $\mathrm{car}(T) = 2$.

Vejamos agora como poderíamos determinar a imagem de T a partir duma sua qualquer representação matricial. Vamos fazer o cálculo usando a já conhecida representação de T nas bases canónicas de P_1 e P_2. Fazendo $(p(x))_B = (a_0, a_1, a_2)$ e $(q(x))_B = (b_0, b_1)$, temos

$$\begin{aligned} T(q(x)) = p(x) &\longrightarrow T_{CB}(q(x))_B = (p(x))_C \\ &\longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \longrightarrow \begin{bmatrix} b_0 + b_1 \\ b_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \\ &\longrightarrow \begin{cases} b_0 + b_1 = a_0 \\ b_1 = a_1 \\ b_1 = a_2 \end{cases} \end{aligned}$$

donde retiraríamos as mesmas conclusões que atrás.

Exercício 20. Considere as transformações lineares $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ definidas por

$$T_1(x, y) = (2x - y, -8x + 4y) \qquad T_2(x, y) = (x + y, x - y) .$$

(a) Verifique se os vectores $(5, 10)$, $(3, 2)$, $(1, 1)$, $(4, 0)$, $(1, -4)$ e $(-1, -2)$ pertencem à imagem de cada uma das transformações.

(b) Determine bases para $\mathrm{im}(T_1)$ e $\mathrm{im}(T_2)$.

Exercício 21. Considere as transformações lineares $T_1, T_2 : P_3 \to P_3$ definidas por

$$T_1(p(x)) = p'(x) \qquad T_2(p(x)) = (x + 1)p''(x) .$$

(a) Verifique se os vectores $3x^3 + x$, $x^2 - 2x + 1$, 0, $4x + 4$ e $2x + 3$ pertencem à imagem de cada uma das transformações.

(b) Determine bases para $\mathrm{im}(T_1)$ e $\mathrm{im}(T_2)$.

Exercício 22. Encontre a imagem de cada uma das transformações lineares consideradas nos Exercícios 7 e 8 recorrendo às suas representações matriciais.

3.4.3 Teorema da dimensão

Do teorema da dimensão para matrizes, obtemos a relação $\mathrm{nul}(T) + \mathrm{car}(T) = \dim(V)$ para qualquer transformação linear $T : V \to W$ com V e W espaços vectoriais de dimensão finita.

Na verdade, pode-se mostrar em geral que essa relação é válida mesmo para espaços de dimensão infinita (convencionando, como habitualmente, $\infty + n = n + \infty = \infty + \infty = \infty$). Tem-se assim o seguinte resultado.

Teorema 8 (Teorema da Dimensão)**.** Seja $T : V \to W$ uma transformação linear. Então

$$\mathrm{nul}(T) + \mathrm{car}(T) = \dim(V)\,.$$

Como consequência, temos várias maneiras de calcular a nulidade e característica de uma transformação linear:

- calcular o núcleo e a imagem de T pela definição e determinar a sua dimensão;

- determinar a nulidade e a característica de uma representação matricial de T, T_{CB}, com B base de V e C base de W;

- sabendo a característica, aplicar o Teorema da Dimensão para obter a nulidade;

- sabendo a nulidade, aplicar o Teorema da Dimensão para obter a característica.

O Teorema da Dimensão para transformações lineares tem muitas consequências importantes. Uma delas é a relação

$$\mathrm{car}(T) \leq \dim(V)\,,$$

que nos indica que a imagem de T é um espaço vectorial com dimensão no máximo igual à de V. Repare-se que já podíamos ter deduzido esta propriedade — uma vez que T preserva combinações lineares, a imagem de qualquer base de V é um conjunto gerador de $T(V)$ — mas temos agora uma forma sintética e expedita de a mostrar.

Por outro lado, conhecer este resultado permite-nos resolver vários problemas de formas mais simples do que as técnicas apresentadas acima. Suponhamos, por exemplo, que pretendemos determinar uma base para a imagem duma transformação $T : \mathbb{R}^2 \to \mathbb{R}^3$ cujo núcleo tem dimensão 1. Pelo Teorema da Dimensão, a imagem de T também tem dimensão 1, pelo que é gerada por um único vector. Por regra, é muito mais simples procurar um vector em $T\left(\mathbb{R}^2\right)$ (qualquer vector serve) do que determinar o espaço das colunas duma matriz que represente T. Em particular, basta calcular $T(0,1)$ e, se este for o vector nulo, calcular $T(1,0)$ (que já não poderá ser $\overrightarrow{0}$).

Exercício 23. Para cada uma das seguintes transformações lineares, indica-se a sua característica. Determine a nulidade de cada uma delas.

(a) $T : \mathbb{R}^3 \to \mathbb{R}^3$ com $\mathrm{car}(T) = 3$

(b) $T : P_2 \to \mathbb{R}^3$ com $\mathrm{car}(T) = 2$

(c) $T : P_2 \to P_2$ com $\mathrm{car}(T) = 1$

(d) $T : P_4 \to M_{3\times 3}$ com $\mathrm{car}(T) = 2$

Exercício 24. Para cada uma das seguintes transformações lineares, indica-se a sua nulidade. Determine a característica de cada uma delas.

(a) $T : M_{3\times 3} \to \mathbb{R}^5$ com $\mathrm{nul}(T) = 2$

(b) $T : M_{2\times 2} \to M_{2\times 2}$ com $\mathrm{nul}(T) = 0$

(c) $T : P_5 \to \mathbb{R}^2$ com $\mathrm{nul}(T) = 4$

(d) $T : M_{2\times 2} \to \mathbb{R}^4$ com $\mathrm{nul}(T) = 2$

3.5 Invertibilidade

Tal como no contexto geral das funções, tem particular interesse saber quando é que uma transformação linear é invertível. Por invertibilidade significa que podemos "desfazer" a aplicação de T, ou seja, a partir de Tv encontrar o vector argumento v.

Embora as transformações não invertíveis também tenham grande utilidade, há um conjunto de resultados que só se aplicam a transformações invertíveis. Uma das grandes vantagens é podermos resolver de forma expedita equações da forma $Tv = w$ por uma aplicação directa da transformação inversa. Mais uma vez, há aqui uma grande semelhança entre esta situação e a resolução de sistemas de equações lineares.

Se for possível obter v a partir de Tv, para qualquer v, podemos pensar na regra operatória que corresponde a essa associação. Assim, faz sentido a seguinte definição.

Definição. Uma transformação linear $T : V \to W$ diz-se *invertível* se existir uma aplicação $S : T(V) \to V$ satisfazendo as duas condições seguintes.

- $(S \circ T)(v) = v$ para qualquer $v \in V$.
- $(T \circ S)(w) = w$ para qualquer $w \in T(V)$.

Uma função S nestas condições diz-se uma *inversa* de T.

Note-se que não exigimos nada relativamente a S a não ser que seja uma aplicação. Veremos dentro de pouco que na realidade S é sempre uma transformação linear; antes, contudo, há algumas observações importantes a fazer.

A inversa de T, quando existe, não é uma aplicação em W, mas apenas na imagem de T. Esta condição não é essencial, já que podemos sempre reduzir o conjunto de chegada ao subespaço apropriado; mas na prática faz mais sentido aceitar esta definição mais permissiva. Considere-se por exemplo uma aplicação $T : \mathbb{R}^2 \to \mathbb{R}^3$. A imagem de T é no máximo um plano (uma vez que $\mathbb{R}^2$ tem dimensão 2 e a dimensão da imagem de T é necessariamente menor ou igual a esse valor); contudo, no caso em que $T\left(\mathbb{R}^2\right)$ é um plano, faz todo o sentido dizer que T é invertível e falar da sua transformação inversa, apesar de esta não estar definida para todos os vectores de $\mathbb{R}^3$.

Exemplo. Sejam $T : \mathbb{R}^2 \to \mathbb{R}^3$ a transformação linear definida por $T(x, y) = (x + y, x - y, 2x)$ e $S : W \to \mathbb{R}^2$ a aplicação tal que $S(x, y, z) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right)$, onde W é o subespaço de $\mathbb{R}^3$ constituído pelos vectores (x, y, z) tais que $z = x + y$.

Vejamos que S é inversa de T. Em primeiro lugar, observe-se que $W = T\left(\mathbb{R}^2\right)$: a imagem de T é gerada por

$$\{T(1, 0), T(0, 1)\} = \{(1, 1, 2), (1, -1, 0)\}$$

e estes vectores satisfazem a condição $z = x+y$; por outro lado, o subespaço de $\mathbb{R}^3$ dos vectores que satisfazem esta condição tem dimensão 2, portanto $W = T(\mathbb{R}^2)$.

Tomando um vector arbitrário (x, y) de $\mathbb{R}^2$, temos que

$$\begin{aligned}(S \circ T)(x, y) &= S(T(x, y)) = S(x+y, x-y, 2x)\\ &= \left(\frac{(x+y)+(x-y)}{2}, \frac{(x+y)-(x-y)}{2}\right) = (x, y)\,,\end{aligned}$$

enquanto que, para um vector arbitrário $(x, y, x+y)$ de W, temos que

$$\begin{aligned}(T \circ S)(x, y, x+y) &= T(S(x, y, x+y)) = T\left(\frac{x+y}{2}, \frac{x-y}{2}\right)\\ &= \left(\frac{x+y}{2} + \frac{x-y}{2}, \frac{x+y}{2} - \frac{x-y}{2}, 2\frac{x+y}{2}\right) = (x, y, x+y)\,.\end{aligned}$$

Conclui-se portanto que S é inversa de T.

Exercício 25. Verifique que os pares de transformações abaixo são inversas.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(x, y) = (2x+y, y-x)$ e $S : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $S(x, y) = \left(\frac{x-y}{3}, \frac{x+2y}{3}\right)$.

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $T(x, y, z) = (x+y+z, x+y, x+z)$ e $S : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $S(x, y, z) = (y+z-x, x-z, x-y)$.

(c) $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(x, y) = (x+y, 2x, 2y)$ e $S : W \to \mathbb{R}^2$ com $W = \left\{\left(\frac{y+z}{2}, y, z\right) \mid y, z \in \mathbb{R}\right\}$ e $S(x, y, z) = \left(\frac{y}{2}, \frac{z}{2}\right)$.

(d) $T : W \to P_2$ com $W = \{p \in P_3 \mid p(0) = 0\}$ tal que $Tp = p'$ e $S : P_2 \to P_3$ com $S\left(a + bx + cx^2\right) = ax + \frac{b}{2}x^2 + \frac{c}{3}x^3$.

(e) $T : M_{2\times 2} \to M_{2\times 2}$ tal que $TA = A^T$ e $S = T$.

(f) $T : \mathbb{R}_3 \to M_{3\times 3}$ tal que $T\overrightarrow{v}$ é a matriz diagonal que tem os elementos de $\overrightarrow{v}$ na diagonal principal e $S : D_{3\times 3} \to \mathbb{R}^3$ com $D_{3\times 3}$ o espaço das matrizes diagonais de dimensão 3×3 e $SA = (a_{11}, a_{22}, a_{33})$.

Em todos os exemplos e exercícios acima, a transformação inversa de T é ainda uma transformação linear. De facto, se S é inversa de $T : V \to W$, então S preserva combinações lineares: tomando $w_1, w_2 \in T(V)$ e dois reais α e β, existem elementos v_1 e v_2 em V tais que $Tv_1 = w_1$ e $Tv_2 = w_2$ e tem-se

$$\begin{aligned}S(\alpha w_1 + \beta w_2) &= S(\alpha T v_1 + \beta T v_2) = S(T(\alpha v_1 + \beta v_2))\\ &= \alpha v_1 + \beta v_2 = \alpha S w_1 + \beta S w_2\,.\end{aligned}$$

Proposição 53. Seja $T : V \to W$ uma transformação linear e $S : T(V) \to V$ uma inversa de T. Então S é transformação linear.

Tal como já estamos habituados noutros contextos, a transformação inversa, quando existe, é única. Mais, as condições da definição de inversa são na realidade demasiado fortes: basta exigirmos que $(S \circ T)(v) = v$ para todo o v para garantir que S é inversa de T.

Proposição 54. Seja $T : V \to W$ uma transformação linear e $S : T(V) \to V$ uma aplicação tal que $(S \circ T)(v) = v$ para qualquer $v \in V$. Então S é a única aplicação nessas condições e T é invertível.

Demonstração. Sejam T e S funções nas condições do enunciado e suponha-se que a função $S' : T(V) \to V$ satisfaz $(S' \circ T)(v) = v$ para qualquer $v \in V$. Então $S'(Tv) = S(Tv)$ para qualquer $v \in V$, donde S' é precisamente S.

Por outro lado, seja $w \in T(V)$. Vamos mostrar que necessariamente $(T \circ S)(w) = w$. Por definição, w é Tv para algum $v \in V$; tem-se então

$$(T \circ S)(w) = T(Sw) = T(S(Tv)) = Tv = w\,,$$

donde S é inversa de T. □

Este resultado justifica que se use uma notação especial para a (única) inversa de T, quando esta existe: esta função denota-se por T^{-1}.

Exercício 26. Para cada uma das transformações lineares T do Exercício 25, quem é T^{-1}?

Há várias formas de caracterizar transformações invertíveis, uma vez que várias das consequências da invertibilidade são de facto equivalentes a essa propriedade. Suponhamos por agora que $T : V \to W$ é uma transformação linear.

Vejamos primeiro que se T for invertível então é injectiva, ou seja, transforma elementos diferentes em imagens diferentes: se $u \neq v$ então $Tu \neq Tv$. Outra forma de escrever isto é exigir que, se $Tu = Tv$, então $u = v$. Ora se existe uma função T^{-1} que transforma Tv em v, então necessariamente este v é único (caso contrário T^{-1} não seria uma função). Logo T é injectiva.

Reciprocamente, se T for injectiva então para cada $w \in T(V)$ existe um único elemento v tal que $Tv = w$; a correspondência de Tv para v é funcional, pelo que define uma função T^{-1} que é trivialmente inversa de T.

A injectividade de T pode ainda ser expressa em termos do seu núcleo: T é injectiva precisamente quando $\mathcal{N}(T) = \{0_V\}$. Observe-se que $0_V \in \mathcal{N}(T)$ em qualquer situação, pois T é uma transformação linear e portanto $T(0_V) = 0_W$.

Se T for injectiva, então existe apenas um vector v tal que $Tv = 0_W$; uma vez que 0_V satisfaz esta equação, concluímos que é forçosamente o único vector no núcleo de T.

Para a implicação recíproca, vamos começar por assumir que $\mathcal{N}(T) = \{0_V\}$ e vejamos que T é injectiva. Tomando u e v tais que $Tu = Tv$, temos que $Tu - Tv = 0_W$; ora $Tu - Tv = T(u-v)$, pois T é transformação linear, donde $(u - v) \in \mathcal{N}(T)$. Mas $\mathcal{N}(T) = \{0_V\}$, logo $u - v = 0_V$, donde $u = v$ e portanto T é injectiva.

Temos então os seguintes resultados.

Proposição 55. Seja $T : V \to W$ uma transformação linear. Então as seguintes afirmações são equivalentes.

- T é invertível.

- T é injectiva.

- $\mathcal{N}(T) = \{0_V\}$.

No caso de V e W terem dimensão finita e bases B e C, respectivamente, a última propriedade é ainda equivalente a $\text{nul}T_{CB} = 0$.

Pelo Teorema da Dimensão para transformações lineares, temos que

$$\text{nul}(T) + \text{car}(T) = \dim(V) .$$

Se T é invertível, então $\text{nul}(T) = 0$, donde $\text{car}(T) = \dim(V)$. Uma vez que a característica é a dimensão da imagem de T, temos o seguinte resultado.

Teorema 9. Seja $T : V \to W$ uma transformação linear invertível. Então

$$\dim(V) = \dim(T(V)) .$$

Este resultado, aliado a propriedades já discutidas atrás, é bastante útil para caracterizar a imagem duma transformação linear. Se B é uma base de V, então $T(B)$ gera $T(V)$; no caso em que T é invertível e V tem dimensão finita, $T(B)$ é um conjunto com $\dim(V)$ elementos que gera $T(V)$, logo, como $\dim(T(V)) = \dim(V)$, $T(B)$ é uma base de $T(V)$.

Outra consequência imediata é que se $T : V \to W$ é tal que $\dim(W) < \dim(V)$, então necessariamente $\text{nul}(T) > 0$ e portanto T não pode ser invertível.

Exemplo.

1. Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ uma transformação linear. Uma vez que $\dim(\mathbb{R}^3) = 3 < \dim(\mathbb{R}^2) = 2$, a transformação T não pode ser invertível.

2. Seja $T : M_{2\times 2} \to \mathbb{R}^3$. Uma vez que $\dim(M_{2\times 2}) = 4 < \dim(\mathbb{R}^3) = 3$, a transformação T não pode ser invertivel.

3. Seja $T : P_3 \to P_2$. Uma vez que $\dim(P_3) = 4 < \dim(P_2) = 3$, a transformação T não pode ser invertivel.

4. Seja $T : \mathcal{P} \to P_5$. Uma vez que $\dim(\mathcal{P}) = \infty < \dim(P_5) = 6$, a transformação T não pode ser invertivel.

Observe-se que nestes casos nem precisamos de conhecer a expressão de T para determinar a sua invertibilidade. Claro está que se $\dim(V) \leq \dim(W)$ não podemos concluir nada sem analisar a transformação.

Exercício 27. Sendo $T : V \to W$ uma transformação linear, indique para quais das seguintes combinações de V e W é que T garantidamente *não* é invertível.

(a) $V = \mathbb{R}^2$, $W = \mathbb{R}$

(b) $V = \mathbb{R}^3$, $W = \mathbb{R}^3$

(c) $V = \mathbb{R}^3$, $W = \mathbb{R}$

(d) $V = \mathbb{R}^3$, $W = M_{2\times 2}$

(e) $V = \mathbb{R}^5$, $W = M_{2\times 2}$

(f) $V = M_{2\times 3}$, $W = P_4$

(g) $V = P_3$, $W = M_{2\times 2}$

(h) $V = P_2$, $W = \mathbb{R}^2$

(i) $V = \mathcal{F}(\mathbb{R})$, $W = P_3$

(j) $V = P_3$, $W = \mathcal{F}(\mathbb{R})$

O resultado seguinte também é bastante útil para o mesmo tipo de aplicação.

Proposição 56. Sejam $T : V \to W$ uma transformação linear invertível, B uma base de V e C uma base de $T(V)$. Então T_{CB} é invertível e

$$\left(T^{-1}\right)_{BC} = \left(T_{CB}\right)^{-1} .$$

Demonstração. Se T é invertível, então $\dim(V) = \mathrm{car}(T) = \dim\left(\mathrm{Col}\left(T_{CB}\right)\right)$, donde T_{CB} é uma matriz quadrada cujas colunas são linearmente independentes; então $\det\left(T_{CB}\right) \neq 0$, donde T_{CB} é invertível.

Então a transformação linear cuja representação matricial em relação às bases C de $T(V)$ e B de V é $\left(T_{CB}\right)^{-1}$ é uma inversa de T, já que $\left(T_{CB}\right)^{-1} T_{CB} = \mathbf{I}_{\dim(\mathbf{V})}$. Logo esta transformação é precisamente T^{-1}, donde $\left(T^{-1}\right)_{BC} = \left(T_{CB}\right)^{-1}$. □

Obtemos então outra forma de determinar se uma transformação é invertível: calcular o determinante da sua representação matricial. Vamos ver alguns exemplos destes diferentes processos de resolução aplicados a transformações lineares concretas.

Exemplo. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida por

$$T(x, y) = (x + y, x - y) .$$

Para determinar se T é invertível, podemos proceder de diferentes maneiras.

1. T é invertível se e só se o seu núcleo contém apenas o vector $(0, 0)$. Ora

$$T(x, y) = (0, 0) \longrightarrow (x + y, x - y) = (0, 0) \longrightarrow \begin{cases} x + y = 0 \\ x - y = 0 \end{cases} \longrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}$$

   donde T é invertível.

2. Uma vez que $\dim\left(\mathbb{R}^2\right) = 2$, sabemos que T é invertível se e só se $\mathrm{car}(T) = 2$. Ora $T(1, 0) = (1, 1)$ e $T(0, 1) = (1, -1)$, que são vectores linearmente independentes. Logo $\dim(T(V)) = 2$, donde T é invertível.

3. A representação matricial de T em relação à base canónica B de $\mathbb{R}^2$ é

$$T_{BB} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} ,$$

   que tem determinante $2 \neq 0$. Então T é invertível.

Para determinar a inversa de T, também podemos utilizar diferentes métodos.

1. Por definição, a inversa de T pega num vector $(a, b) = T(x, y)$ e devolve o vector original (x, y). Temos portanto de resolver esta equação para a e b genéricos.

$$T(x, y) = (a, b) \longrightarrow (x + y, x - y) = (a, b) \longrightarrow \begin{cases} x + y = a \\ x - y = b \end{cases}$$

$$\longrightarrow \left[\begin{array}{cc|c} 1 & 1 & a \\ 1 & -1 & b \end{array}\right] \underset{l_2 - l_1}{} \longrightarrow \left[\begin{array}{cc|c} 1 & 1 & a \\ 0 & -2 & b - a \end{array}\right] \longrightarrow \begin{cases} x + y = a \\ -2y = b - a \end{cases}$$

$$\longrightarrow \begin{cases} x + y = a \\ y = \frac{a-b}{2} \end{cases} \longrightarrow \begin{cases} x = \frac{a+b}{2} \\ y = \frac{a-b}{2} \end{cases}$$

Assim, $T^{-1}(a,b) = \left(\frac{a+b}{2}, \frac{a-b}{2}\right)$ ou, usando as variáveis usuais de $\mathbb{R}^2$,

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

2. Em alternativa, podemos recorrer à representação matricial de T em relação à base canónica. Pela proposição acima, a representação de T^{-1} na mesma base é $(T_{BB})^{-1}$, ou seja,

$$\left(T^{-1}\right)_{BB} = (T_{BB})^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = -\frac{1}{2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

A partir da representação matricial de T^{-1} podemos determinar a sua expressão. Das colunas da matriz acima, obtemos $T^{-1}(1,0) = \left(\frac{1}{2}, \frac{1}{2}\right)$ e $T^{-1}(0,1) = \left(\frac{1}{2}, -\frac{1}{2}\right)$, donde

$$\begin{aligned} T^{-1}(x,y) &= T^{-1}(x(1,0) + y(0,1)) = xT^{-1}(1,0) + yT^{-1}(0,1) \\ &= x\left(\frac{1}{2}, \frac{1}{2}\right) + y\left(\frac{1}{2}, -\frac{1}{2}\right) = \left(\frac{x+y}{2}, \frac{x-y}{2}\right) \end{aligned}$$

Exemplo.

1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ a transformação linear definida por

$$T(x,y) = (2x+y, x, x+y).$$

Vamos começar por verificar que T é invertível mostrando que $\mathcal{N}(T) = \{(0,0)\}$.

$$T(x,y) = (0,0,0) \longrightarrow (2x+y, x, x+y) = (0,0,0) \longrightarrow \begin{cases} 2x+y=0 \\ x=0 \\ x+y=0 \end{cases} \longrightarrow \begin{cases} x=0 \\ y=0 \end{cases}$$

Logo T é invertível, existindo portanto a sua inversa $T^{-1} : T(\mathbb{R}^2) \to \mathbb{R}^2$. Como $\dim(T(\mathbb{R}^2)) = \dim(\mathbb{R}^2) = 2$, a imagem de T é um subespaço próprio de $\mathbb{R}^3$; podemos determiná-lo explicitamente procurando directamente os vectores (a,b,c) de $\mathbb{R}^3$ que são imagem de $T(x,y)$.

$$\begin{aligned} T(a,b) = (x,y,z) &\longrightarrow (2a+b, a, a+b) = (x,y,z) \\ &\longrightarrow \begin{cases} 2a+b=x \\ a=y \\ a+b=z \end{cases} \longrightarrow \left[\begin{array}{cc|c} 2 & 1 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{array}\right] \begin{array}{c} \\ 2l_2 - l_1 \\ 2l_3 - l_1 \end{array} \\ &\longrightarrow \left[\begin{array}{cc|c} 2 & 1 & x \\ 0 & -1 & 2y-x \\ 0 & 1 & 2z-x \end{array}\right] \begin{array}{c} \\ \\ l_3 + l_2 \end{array} \longrightarrow \left[\begin{array}{cc|c} 2 & 1 & x \\ 0 & -1 & 2y-x \\ 0 & 0 & 2x-2y-2z \end{array}\right] \end{aligned}$$

Assim, o sistema é possível apenas se $2x - 2y - 2z = 0$ que é equivalente a $z = x - y$. Temos portanto que

$$T\left(\mathbb{R}^2\right) = \left\{(x,y,z) \in \mathbb{R}^3 \mid z = x - y\right\} = \{(x,y,x-y) \mid x,y \in \mathbb{R}\}.$$

Para determinar a expressão de T^{-1}, temos de resolver a equação $T(x,y) = (a,b,a-b)$.

$$\begin{aligned} T(x,y) = (a,b,a-b) &\longrightarrow (2x+y, x, x+y) = (a,b,a-b) \\ &\longrightarrow \begin{cases} 2x+y=a \\ x=b \\ x+y=a-b \end{cases} \longrightarrow \begin{cases} y=a-2b \\ x=b \\ b+(a-2b)=a-b \end{cases} \end{aligned}$$

Logo $(x,y) = (b, a-2b) = T^{-1}(a,b,a-b)$ e, portanto, a expressão de T^{-1} é

$$T^{-1}(x,y,x-y) = (y, x-2y)\,.$$

2. Consideremos novamente a mesma transformação linear T do exemplo anterior e vamos resolver o mesmo problema recorrendo à sua representação matricial.

   Uma vez que $T(1,0) = (2,1,1)$ e $T(0,1) = (1,0,1)$ e estes vectores são linearmente independentes, concluímos que a imagem de T tem dimensão 2, pelo que T é injectiva. A representação de T em relação à base canónica B de $\mathbb{R}^2$ e à base $C = \{(2,1,1),(1,0,1)\}$ de $T\left(\mathbb{R}^2\right)$ é, trivialmente, a matriz identidade; temos então que T^{-1} também é representada pela matriz identidade em relação às mesmas bases.

   Para determinar a expressão geral de T^{-1}, temos de resolver um problema de combinações lineares. Um vector genérico (x,y,z) de $T\left(\mathbb{R}^2\right)$ é combinação linear de $(2,1,1)$ e $(1,0,1)$, donde

   $$(x,y,z) = c_1(2,1,1) + c_2(1,0,1) = (2c_1+c_2, c_1, c_1+c_2)$$

   que é uma equação possível se

   $$(x,y,z) = (2c_1+c_2, c_1, c_1+c_2) \longrightarrow \begin{cases} 2c_1+c_2=x \\ c_1=y \\ c_1+c_2=z \end{cases} \longrightarrow \begin{cases} 2y+(z-y)=x \\ c_1=y \\ c_2=z-y \end{cases}$$

   for um sistema possível. A primeira equação corresponde a $y+z=x$, ou $z=x-y$, e portanto $T\left(\mathbb{R}^2\right)$ é o conjunto dos vectores $\{(x,y,z) \in \mathbb{R}^3 \mid z = x-y\}$. Tomando um vector nestas condições, temos que

   $$\begin{aligned} T^{-1}(x,y,z) &= T^{-1}\left(c_1(2,1,1) + c_2(1,0,1)\right) = c_1T^{-1}(2,1,1) + c_2T^{-1}(1,0,1) \\ &= c_1(1,0) + c_2(0,1) = (c_1,c_2) = (y, z-y) = (y, x-2y)\,. \end{aligned}$$

3. Seja $T : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$ a transformação linear dada por $Tf = f'$. Vimos atrás que o núcleo de T é

   $$\mathcal{N}(T) = \{f \mid f(x) \text{ é constante}\}\,.$$

   Assim, o núcleo de T contém várias funções para além da função nula, donde T não é invertível.

Exercício 28. Considere novamente as transformações lineares do Exercício 13.

(a) Indique para que valores de n é que T_n é invertível e calcule a expressão de T_n^{-1}.

(b) Verifique o resultado da alínea anterior recorrendo à representação matricial das transformações em causa em relação às bases canónicas dos espaços envolvidos.

3.6 Valores e vectores próprios

Nas secções anteriores, discutimos um conjunto de propriedades e espaços associados a transformações lineares: estudámos operações algébricas com transformações lineares, discutimos como determinar núcleo e imagem duma transformação, e estudámos transformações lineares quanto à invertibilidade.

Uma das técnicas a que temos recorrido repetidamente, como forma sistemática de resolver vários destes problemas usando técnicas muitas vezes mais simples, foi a representação matricial de transformações lineares. Reduzindo os problemas sobre transformações lineares a problemas sobre matrizes, conseguimos aplicar métodos já conhecidos (como o método de eliminação de Gauss) para responder de forma eficiente a todas essas questões.

Dos exemplos que vimos, é claro que a complexidade de trabalhar com a representação matricial depende da complexidade da própria matriz. Em casos em que esta é mais simples, a resposta às perguntas anteriores será também mais simples. Por exemplo, se T tiver uma representação por uma matriz em escada de linhas, o cálculo da sua característica e a determinação da sua invertibilidade serão muito simples; se T tiver uma representação por uma matriz diagonal, até a determinação da sua inversa é imediata. Nesta secção vamos analisar as transformações lineares que podem ser representadas por matrizes diagonais.

Começamos por notar que para a representação de $T : V \to W$ ser uma matriz diagonal é necessário que $\dim(V) = \dim(W)$. Nesta secção vamos mesmo ser mais restritivos, e considerar apenas transformações cujo espaço de chegada coincide com o espaço de partida.

3.6.1 Subespaços invariantes

Vejamos um exemplo que ilustra como a escolha de base pode condicionar a resolução do problema. Consideremos a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida como a projecção ortogonal sobre o vector $(2, 1, -4)$; ou seja, $T\overrightarrow{v} = P_{(2,1,-4)}(\overrightarrow{v})$.

Escolhendo a base canónica B de $\mathbb{R}^3$, as imagens dos seus vectores são

$$\begin{aligned}
T(1,0,0) &= P_{(2,1,-4)}((1,0,0)) = \frac{(1,0,0)\cdot(2,1,-4)}{\|(2,1,-4)\|}(2,1,-4) = \left(\frac{4}{21}, \frac{2}{21}, -\frac{8}{21}\right)\\
T(0,1,0) &= P_{(2,1,-4)}((0,1,0)) = \frac{(0,1,0)\cdot(2,1,-4)}{\|(2,1,-4)\|}(2,1,-4) = \left(\frac{2}{21}, \frac{1}{21}, -\frac{4}{21}\right)\\
T(0,0,1) &= P_{(2,1,-4)}((0,0,1)) = \frac{(0,0,1)\cdot(2,1,-4)}{\|(2,1,-4)\|}(2,1,-4) = \left(-\frac{8}{21}, -\frac{4}{21}, \frac{16}{21}\right)
\end{aligned}$$

donde T tem representação matricial em relação a esta base

$$T_{BB} = \begin{bmatrix} \frac{4}{21} & \frac{2}{21} & -\frac{8}{21} \\ \frac{2}{21} & \frac{1}{21} & -\frac{4}{21} \\ -\frac{8}{21} & -\frac{4}{21} & \frac{16}{21} \end{bmatrix}.$$

Pensemos, porém em termos geométricos. Uma vez que estamos a fazer projecções ortogonais sobre $(2, 1, -4)$, sabemos que todos múltiplos deste vector ficam inalterados; ou seja,

$$T(2,1,-4) = (2,1,-4).$$

Por outro lado, qualquer vector ortogonal a $(2, 1, -4)$ pertence ao núcleo de T; ora $(1, -2, 0)$ e $(2, 0, 1)$ são dois vectores ortogonais a $(2, 1, -4)$ que são linearmente independentes. Então $C = \{(2, 1, -4), (1, -2, 0), (2, 0, 1)\}$ é uma base de $\mathbb{R}^3$ e temos

$$\begin{aligned}
(T(2, 1, -4))_C &= ((2, 1, -4))_C = (1, 0, 0) \\
(T(1, -2, 0))_C &= \left(\overrightarrow{0}\right)_C = (0, 0, 0) \\
(T(2, 0, 1))_C &= \left(\overrightarrow{0}\right)_C = (0, 0, 0)
\end{aligned}$$

donde

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

Daqui é imediato que $\mathrm{car}(T) = 1$, que $\mathrm{nul}(T) = 2$ e que $T\left(\mathbb{R}^3\right) = L(\{(2, 1, -4)\})$, por exemplo. O facto de estarmos a utilizar uma base adequada permite-nos simplificar em muito o estudo da transformação linear T.

Geometricamente, o espaço gerado pelo vector $(2, 1, -4)$ tem uma propriedade interessante: para qualquer vector $v \in L(\{(2, 1, -4)\})$, tem-se que $Tv \in L(\{(2, 1, -4)\})$, ou seja, o espaço $L(\{(2, 1, -4)\})$ é um espaço que não varia sob a transformação T.

Definição. Seja $T : V \to V$ uma transformação linear. Um subespaço S de V tal que $Tv \in S$, para qualquer $v \in S$, diz-se um *subespaço invariante* de T.

No exemplo anterior, o espaço $S_1 = L(\{(2, 1, -4)\})$ é um subespaço invariante de T. Este não é o único subespaço de $\mathbb{R}^3$ com esta propriedade: $S_2 = \left\{\overrightarrow{0}\right\}$ e $S_3 = \mathbb{R}^3$ também são subespaços invariantes (embora pouco interessantes) de T. Outro subespaço invariante de T é $S_4 = L(\{(2, 1, -4), (2, 0, 1)\})$. De facto, se $\overrightarrow{v} = c_1(2, 1, -4) + c_2(2, 0, 1)$, temos que

$$T\overrightarrow{v} = T\left(c_1(2, 1, -4) + c_2(2, 0, 1)\right) = c_1T(2, 1, -4) + c_2T(2, 0, 1) = c_1(2, 1, -4) = c_1\overrightarrow{v}$$

e portanto $T\overrightarrow{v}$ ainda é combinação linear daqueles dois vectores.

Outro subespaço invariante por T é $S_5 = L(\{(1, -2, 0), (2, 0, 1)\})$. Aqui temos $T\overrightarrow{v} = \overrightarrow{0}$ para todo o vector que seja combinação linear de $(1, -2, 0)$ e $(2, 0, 1)$, e como $\overrightarrow{0}$ é um elemento de qualquer subespaço de $\mathbb{R}^3$ concluímos que S_5 é um subespaço invariante.

Os subespaços S_1 e S_5 têm uma outra propriedade que não é partilhada por S_2, S_3 e S_4: a imagem de qualquer vector desses subespaços é um múltiplo (único) desse vector. De facto, temos que, para $\overrightarrow{v} \in S_1$,

$$T\overrightarrow{v} = T\left(c_1(2, 1, -4)\right) = c_1T(2, 1, -4) = c_1(2, 1, -4) = \overrightarrow{v} = 1\overrightarrow{v},$$

enquanto que para $\overrightarrow{v} \in S_5$ já vimos que

$$T\overrightarrow{v} = \overrightarrow{0} = 0\overrightarrow{v}.$$

Dizemos que estes espaços são *subespaços próprios* de T. É esta propriedade que torna a representação matricial de T tão simples: conseguimos encontrar dois subespaços próprios que conjuntamente geram $\mathbb{R}^3$, pelo que conseguimos encontrar vectores suficientes para formar uma base em que a imagem de qualquer vector é um múltiplo desse vector — o que, conforme um pouco de reflexão mostra, corresponde a ter uma representação matricial diagonal para T.

Vejamos outro exemplo. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida por

$$T(x, y) = (2y, 2x)\,.$$

Esta transformação admite à mesma os subespaços invariantes triviais $\left\{\overrightarrow{0}\right\}$ e $\mathbb{R}^2$. Contudo, há outro subespaço invariante: o espaço S_1 dos vectores cujas coordenadas são iguais; de facto, estes vectores são transformados em vectores de coordenadas iguais. Por outras palavras, se $\overrightarrow{v} = (x, x)$, então

$$T\overrightarrow{v} = T(x, x) = (2x, 2x) = 2\overrightarrow{v}\,.$$

Há ainda outro subespaço invariante não trivial de T: o espaço S_2 constituído pelos vectores cujas coordenadas são simétricas. Neste caso, sendo $\overrightarrow{v}(x, -x)$ um tal vector, temos que

$$T\overrightarrow{v} = T(x, -x) = (-2x, 2x) = -2\overrightarrow{v}\,.$$

Uma vez que S_1 e S_2 são espaços próprios, vamos escolher uma base de $\mathbb{R}^2$ com vectores destes espaços. Como a dimensão de $\mathbb{R}^2$ é 2, basta escolher um vector em cada espaço, por exemplo $(1, 1)$ e $(1, -1)$. Seja B a base formada por estes vectores; então

$$(T(1, 1))_B = (2(1, 1))_B = (2, 0) \qquad (T(1, -1))_B = (-2(1, -1)) = (0, -2)\,,$$

donde

$$T_{BB} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Neste caso, a representação matricial de T em relação à base canónica C de $\mathbb{R}^2$ era a matriz

$$T_{CC} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix},$$

pelo que não ganhámos muito em termos do cálculo de núcleo, característica ou imagem. Contudo, inverter T_{BB} é directo, enquanto inverter T_{CC} é menos imediato.

Exercício 29. Verifique que $L(\{(1, 0, 1), (1, 0, 0)\})$ e $L(\{(0, 2, 0)\})$ são subespaços invariantes da transformação $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T(x, y, z) = (2x, y, 2z)\,.$$

Exercício 30. Determine se $(1, 2, 3)$ pertence a um subespaço invariante da transformação $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T(x, y, z) = (2x + y, z, z)\,.$$

Exercício 31. Determine os espaços próprios da transformação $T : M_{2\times 2} \to M_{2\times 2}$ definida por $TA = A^T$.

3.6.2 Valores e vectores próprios

Vimos na secção anterior que alguns subespaços próprios satisfazem a propriedade de Tv ser um múltiplo de v, para qualquer vector v nesses espaços. Também observámos nos exemplos anteriores (e é simples mostrar que isto se verifica sempre) que quando tal sucede se tem mesmo $Tv = \lambda v$ para um valor fixo λ, único para todos os vectores desse espaço.

Definição. Sejam $T : V \to V$ uma transformação linear, $v \in V$ um vector e $\lambda \in \mathbb{R}$ um real tais que $Tv = \lambda v$. O vector v diz-se um *vector próprio* de T e o valor λ diz-se um *valor próprio* de V. Diz-se ainda que v é um vector próprio *associado* ao valor próprio λ.

A importância dos valores e vectores próprios deve ser clara: se conseguirmos encontrar uma base de vectores próprios de T para o espaço linear V, então a representação matricial de T em relação a essa base é uma matriz diagonal e os valores na diagonal principal dessa matriz são precisamente os valores próprios associados aos vectores próprios da base.

Exemplo.

1. No caso da transformação $T : \mathbb{R}^3 \to \mathbb{R}^3$ que associa cada vector à sua projecção ortogonal sobre o vector $(2, 1, -4)$, temos que $(2, 1, -4)$ é um vector próprio de T associado ao valor próprio 1, enquanto que $(1, 2, 0)$ e $(2, 0, 1)$ são vectores próprios de T associados ao valor próprio 0. A representação matricial de T em relação à base formada por estes três vectores é precisamente a matriz diagonal contendo os valores próprios na diagonal principal.

2. No caso da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por
$$T(x, y) = (2y, 2x) ,$$
o vector $(1, 1)$ é um vector próprio de T associado ao valor próprio 2, enquanto que $(1, -1)$ é um vector próprio de T associado ao valor próprio -2. A representação matricial de T em relação à base formada por estes dois vectores é precisamente a matriz diagonal contendo os valores próprios na diagonal principal.

Os resultados das secções anteriores já nos permitem calcular um conjunto particular de vectores próprios de qualquer transformação. Por exemplo, se $v \in \mathcal{N}(T)$, então $Tv = 0_V = 0v$, donde v é um vector próprio associado ao valor próprio 0. Por outras palavras, o núcleo de T é sempre constituído pelos vectores próprios associados ao valor próprio 0.

Infelizmente, em geral esta informação não é muito útil: a única transformação que admite uma base de vectores próprios associados ao valor próprio 0 é a transformação nula. Porém, este resultado salienta um pormenor importante: o núcleo de T é um subespaço próprio de T. Em geral, temos o seguinte resultado.

Proposição 57. Sejam $T : V \to V$ uma transformação linear e λ um valor próprio de T. O conjunto E_λ dos vectores próprios associados ao valor próprio λ é um espaço linear que é subespaço próprio de T.

Demonstração. Claramente, se $v \in E_\lambda$ então $Tv = \lambda v$ pertence a qualquer espaço linear que contenha v. Vejamos então que E_λ é um subespaço linear de V. Quaisquer dois elementos u e

v em E_λ são vectores próprios de T associados ao valor próprio λ, donde $Tu = \lambda u$ e $Tv = \lambda v$. Então para quaisquer reais α e β

$$T(\alpha u + \beta v) = \alpha Tu + \beta Tv = \alpha(\lambda u) + \beta(\lambda v) = \lambda(\alpha u + \beta v)\,.$$

$\square$

À dimensão do subespaço próprio E_λ chama-se *multiplicidade geométrica de* λ, denotada por $m_g(\lambda)$.

Exercício 32. Qual a multiplicidade geométrica dos valores próprios dos exemplos anteriores?

Exercício 33. Verifique se o vector $(1,1,1)$ é um vector próprio das seguintes transformações lineares de $\mathbb{R}^3$ para $\mathbb{R}^3$. Em caso afirmativo, qual é o valor próprio que lhe está associado?

(a) $T(x,y,z) = (x+y, y+z, z+x)$

(b) $T(x,y,z) = (2x, 0, y)$

(c) $T(x,y,z) = (2x, y+z, 4x+y+z)$

(d) $T(x,y,z) = (y-2x, -z, x-y-z)$

3.6.3 Polinómio característico

As questões que se colocam neste ponto são duas.

- Como calcular todos os valores próprios de uma transformação linear?
- Como calcular todos os vectores próprios associados a um valor próprio?

Nesta secção vamos ver como responder a estas questões de uma forma sistemática, a partir da representação matricial da transformação linear.

Consideremos uma transformação linear $T : V \to V$ onde V é um espaço linear de dimensão finita n. Sendo B e C bases de V e T_{CB} a representação matricial de T em relação às bases B e C, sabemos que

$$Tu = v \Longleftrightarrow T_{CB}(u)_B = (v)_C\,.$$

Um vector próprio de T satisfaz a equação

$$Tv = \lambda v$$

e em termos de representação matricial obtemos as seguintes equivalências.

$$\begin{aligned} Tv = \lambda v &\Longleftrightarrow T_{CB}(v)_B = \lambda(v)_B \Longleftrightarrow T_{CB}(v)_B - \lambda(v)_B = 0 \\ &\Longleftrightarrow (T_{CB} - \lambda\mathbf{I_n})\,(v)_B = 0 \Longleftrightarrow (v)_B \in \mathcal{N}\,(T_{CB} - \lambda\mathbf{I_n}) \end{aligned}$$

Descontando o caso trivial $v = 0_V$, para que esta equação tenha soluções é necessário que $\mathcal{N}\,(T_{CB} - \lambda\mathbf{I_n})$ seja não trivial; então a matriz $T_{CB} - \lambda\mathbf{I_n}$ é uma matriz singular, pelo que o seu determinante é nulo. Obtemos então a seguinte caracterização dos valores e vectores próprios de T.

Proposição 58. Sejam V um espaço vectorial de dimensão finita n e $T : V \to V$ uma transformação linear cuja representação matricial em relação a bases B e C de V é T_{CB}. Então:

- λ é um valor próprio de T se e só se $\det(T_{CB} - \lambda \mathbf{I_n}) = 0$;
- nessa situação, $E_\lambda = \mathcal{N}(T_{CB} - \lambda \mathbf{I_n}) = \mathcal{N}(T - \lambda \mathbf{I_V})$, onde $\mathbf{I_V}$ é a transformação identidade em V.

Definição. A equação $\det(T_{CB} - \lambda \mathbf{I_n}) = 0$ diz-se *equação característica* de T e o polinómio que resulta de expandir esta equação característica chama-se *polinómio característico* de T, denotado por $p_T(\lambda)$ ou simplesmente $p(\lambda)$ quando a transformação T for clara do contexto.

Chama-se *multiplicidade aritmética* do valor próprio λ_0, denotada $m_a(\lambda_0)$, ao expoente do factor $(\lambda - \lambda_0)$ na factorização do polinómio característico $p(\lambda)$.

Uma vez que estes conceitos estão definidos para a representação matricial de T, podemos proceder de forma inversa ao habitual e definir os conceitos análogos para matrizes quadradas.

Definição. Seja A uma matriz quadrada de dimensão n. Um vector $\vec{v} \in \mathbb{R}^n$ diz-se *vector próprio* de A associado ao *valor próprio* λ se $A\underline{v} = \lambda\underline{v}$. O conjunto dos vectores próprios associados ao mesmo valor próprio λ diz-se *subespaço próprio* de A associado a λ; o polinómio $\det(A - \lambda \mathbf{I_n})$ diz-se polinómio característico de A.

Todas as propriedades que discutimos anteriormente para transformações lineares aplicam--se naturalmente aos conceitos análogos para matrizes. Em particular, os valores próprios de uma matriz A determinam-se resolvendo a equação

$$\det(A - \lambda \mathbf{I_n}) = 0\,.$$

Há no entanto um pormenor importante: o conceito de valor e vector próprio é um conceito geométrico que tem uma correspondência algébrica sobre matrizes. Porém, tratando-se de propriedades da *transformação*, mantêm-se inalterados independentemente da base escolhida para a sua representação matricial. Ou seja, se A e B são representações matriciais duma mesma transformação T em relação a bases diferentes, então os valores e vectores próprios de A e B são iguais (e iguais aos de T).

Exemplo.

1. Seja $T : P_1 \to P_1$ a transformação linear definida por $Tp = p + p'$. A representação matricial de T em relação à base canónica B de P_1 é a matriz

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

Assim, os valores próprios de T são as soluções da equação característica

$$\begin{aligned} \det(T_{BB} - \lambda \mathbf{I_2}) = 0 &\iff \det\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \\ &\iff \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = 0 \iff (1-\lambda)^2 = 0 \\ &\iff 1 - \lambda = 0 \iff \lambda = 1 \end{aligned}$$

sendo $(1-\lambda)^2$ é o polinómio característico de T (e da matriz T_{BB}). Logo T tem apenas um valor próprio $\lambda = 1$ com multiplicidade aritmética 2, $m_a(1) = 2$, uma vez que o factor $(\lambda - 1)$ ocorre com expoente 2 em $p(\lambda)$.

Vamos agora determinar os vectores próprios associados a $\lambda = 1$. Para tal, temos de determinar o núcleo da transformação $T - \lambda \mathbf{I_{P_1}} = T - \mathbf{I_{P_1}}$, ou seja, os polinómios p para os quais $(T - \mathbf{I_{P_1}})(p) = 0$.

$$\begin{aligned}(T - \mathbf{I_{P_1}})(p) = 0 &\iff Tp - \mathbf{I_{P_1}}p = 0 \iff p + p' - p = 0 \\ &\iff p' = 0 \iff p(x) = a\end{aligned}$$

para algum $a \in \mathbb{R}$.

Logo o subespaço próprio associado ao valor próprio 1 é o espaço que contém os polinómios constantes, pelo que $m_g(1) = 1$.

2. Vamos agora determinar os valores e vectores próprios da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ com $T(x,y) = (2y, 2x)$ acima discutida por via algébrica. A representação de T relativamente à base canónica B de $\mathbb{R}^2$ é
$$T_{BB} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$
e o polinómio característico desta matriz é
$$\begin{aligned}\det(T_{BB} - \lambda \mathbf{I_2}) &= \det\left(\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det \begin{bmatrix} -\lambda & 2 \\ 2 & -\lambda \end{bmatrix} \\ &= \lambda^2 - 4 = (\lambda - 2)(\lambda + 2)\end{aligned}$$
donde obtemos os valores próprios 2 e -2, ambos com multiplicidade algébrica 1.

Para $\lambda = 2$, o espaço próprio é dado pelo núcleo de $T_{BB} - 2\mathbf{I_2}$, ou seja, é o conjunto de soluções do seguinte sistema homogéneo.
$$\left[\begin{array}{cc|c} -2 & 2 & 0 \\ 2 & -2 & 0 \end{array}\right] \longrightarrow \left\{ x = y \right.$$
O subespaço próprio E_2 é portanto o espaço gerado pelo vector $(1,1)$, tendo-se $m_g(2) = 1$.

Para $\lambda = -2$, o espaço próprio é dado pelo núcleo de $T_{BB} + 2\mathbf{I_2}$, ou seja, é o conjunto de soluções do seguinte sistema homogéneo.
$$\left[\begin{array}{cc|c} 2 & 2 & 0 \\ 2 & 2 & 0 \end{array}\right] \longrightarrow \left\{ x = -y \right.$$
O subespaço próprio E_{-2} é portanto o espaço gerado pelo vector $(1,-1)$, tendo-se novamente $m_g(-2) = 1$.

Neste caso, o raciocínio algébrico é também bastante simples, não havendo grande complexidade na determinação dos valores e vectores próprios.

3. Voltemos finalmente a considerar a transformação $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida como a projecção ortogonal sobre o vector $(2,1,-4)$ e vamos determinar os seus valores e vectores próprios por via algébrica. Vimos já que a representação de T em relação à base canónica B de $\mathbb{R}^3$ é a matriz
$$T_{BB} = \begin{bmatrix} \frac{4}{21} & \frac{2}{21} & -\frac{8}{21} \\ \frac{2}{21} & \frac{1}{21} & -\frac{4}{21} \\ -\frac{8}{21} & -\frac{4}{21} & \frac{16}{21} \end{bmatrix}.$$

Vamos então calcular os zeros de $\det\left(T_{BB} - \lambda \mathbf{I_3}\right)$.

$$\begin{aligned}
\det\left(T_{BB} - \lambda \mathbf{I_3}\right) &= \det\left(\begin{bmatrix} \frac{4}{21} & \frac{2}{21} & -\frac{8}{21} \\ \frac{2}{21} & \frac{1}{21} & -\frac{4}{21} \\ -\frac{8}{21} & -\frac{4}{21} & \frac{16}{21} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\
&= \det\begin{bmatrix} \frac{4}{21} - \lambda & \frac{2}{21} & -\frac{8}{21} \\ \frac{2}{21} & \frac{1}{21} - \lambda & -\frac{4}{21} \\ -\frac{8}{21} & -\frac{4}{21} & \frac{16}{21} - \lambda \end{bmatrix} \\
&= \left(\frac{4}{21} - \lambda\right) \det\begin{bmatrix} \frac{1}{21} - \lambda & -\frac{4}{21} \\ -\frac{4}{21} & \frac{16}{21} - \lambda \end{bmatrix} - \frac{2}{21} \det\begin{bmatrix} \frac{2}{21} & -\frac{4}{21} \\ -\frac{8}{21} & \frac{16}{21} - \lambda \end{bmatrix} \\
&\qquad - \frac{8}{21} \det\begin{bmatrix} \frac{2}{21} & \frac{1}{21} - \lambda \\ -\frac{8}{21} & -\frac{4}{21} \end{bmatrix} \\
&= \left(\frac{4}{21} - \lambda\right)\left(\left(\frac{1}{21} - \lambda\right)\left(\frac{16}{21} - \lambda\right) - \frac{16}{441}\right) - \frac{2}{21}\left(\frac{2}{21}\left(\frac{16}{21} - \lambda\right) - \frac{32}{441}\right) \\
&\qquad - \frac{8}{21}\left(-\frac{8}{441} + \left(\frac{1}{21} - \lambda\right)\frac{8}{21}\right) \\
&= \left(\frac{4}{21} - \lambda\right)\left(\lambda^2 - \frac{17}{21}\lambda\right) + \frac{4}{441}\lambda + \frac{64}{441}\lambda \\
&= -\lambda^3 + \lambda^2 = \lambda^2(1 - \lambda)
\end{aligned}$$

cujas raízes são 0 (com multiplicidade aritmética 2) e 1 (com multiplicidade aritmética 1).

O espaço próprio associado ao valor próprio 0 é o núcleo de T, que corresponde ao núcleo de T_{BB}. Ora

$$\left[\begin{array}{ccc|c} \frac{4}{21} & \frac{2}{21} & -\frac{8}{21} & 0 \\ \frac{2}{21} & \frac{1}{21} & -\frac{4}{21} & 0 \\ -\frac{8}{21} & -\frac{4}{21} & \frac{16}{21} & 0 \end{array}\right] \begin{array}{l} \frac{21}{2}l_1 \\ l_1 - 2l_2 \\ 2l_1 + l_3 \end{array} \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow 2x = 4z - y$$

que é o espaço gerado por $\{(1, -2, 0), (2, 0, 1)\}$. Assim, temos que $m_g(0) = 2$. Quanto ao espaço próprio associado ao valor próprio 1, é o núcleo de $T_{BB} - \mathbf{I_3}$, e temos

$$\left[\begin{array}{ccc|c} \frac{4}{21} - 1 & \frac{2}{21} & -\frac{8}{21} & 0 \\ \frac{2}{21} & \frac{1}{21} - 1 & -\frac{4}{21} & 0 \\ -\frac{8}{21} & -\frac{4}{21} & \frac{16}{21} - 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} -\frac{17}{21} & \frac{2}{21} & -\frac{8}{21} & 0 \\ \frac{2}{21} & -\frac{20}{21} & -\frac{4}{21} & 0 \\ -\frac{8}{21} & -\frac{4}{21} & -\frac{5}{21} & 0 \end{array}\right] \begin{array}{l} 21l_1 \\ 21l_2 \\ 21l_3 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} -17 & 2 & -8 & 0 \\ 2 & -20 & -4 & 0 \\ -8 & -4 & -5 & 0 \end{array}\right] \begin{array}{l} \\ 2l_1 + 17l_2 \\ 8l_1 - 17l_3 \end{array} \longrightarrow \left[\begin{array}{ccc|c} -17 & 2 & -8 & 0 \\ 0 & -336 & -84 & 0 \\ 0 & 84 & 21 & 0 \end{array}\right] \begin{array}{l} \\ -\frac{1}{84}l_2 \\ 4l_3 + l_2 \end{array}$$

$$\longrightarrow \left[\begin{array}{ccc|c} -17 & 2 & -8 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \begin{cases} -17x + 2y - 8z = 0 \\ 4y + z = 0 \end{cases} \longrightarrow \begin{cases} x = 2y \\ z = -4y \end{cases}$$

donde este espaço é gerado pelo vector $(2, 1, -4)$ e portanto $m_g(1) = 1$.

Observe-se que neste exemplo o raciocínio geométrico é francamente mais simples do que a resolução algébrica, devido à complexidade da matriz T_{BB}.

Exercício 34. Para cada uma das seguintes matrizes, determine os seus valores próprios e os espaços próprios associados a cada valor próprio.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 4 & 0 & 1 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 4 & 2 & -1 \\ -6 & -4 & 3 \\ -6 & -6 & 5 \end{bmatrix}$

3.6.4 Diagonalização

Vamos agora usar os vectores e valores próprios para responder às questões colocadas no início deste estudo. Em particular, vamos ver como resolver o seguinte problema: se $T : V \to V$ é uma transformação linear, será que existe uma base de V tal que a representação matricial de T nessa base é o mais simples possível?

A representação matricial mais simples possível é claramente por uma matriz diagonal.

Definição. Seja $T : V \to V$ uma transformação linear. Se existir uma base B de V para a qual T_{BB} é uma matriz diagonal, a transformação T diz-se *diagonalizável*.

Teorema 10. Uma transformação $T : V \to V$ é diagonalizável se existir uma base de V formada apenas por vectores próprios de T.

Demonstração. Suponhamos que a matriz T é diagonalizável. Então existe uma base B de V tal que T_{BB} é uma matriz diagonal. Por definição de representação matricial, $(Tv)_B = T_{BB}(v)_B$ para qualquer vector v de V.

Em particular, esta relação é válida para os vectores da base B. Sendo $B = \{v_1, v_2, \ldots, v_n\}$, o vector v_i tem coordenadas $(0, \ldots, 0, 1, 0, \ldots, 0)$, com a única entrada não nula na posição i. Então $T_{BB}\left(v_i\right)_B$, é a coluna i de T_{BB}, que é da forma $(0, \ldots, 0, a_{ii}, 0, \ldots, 0) = a_{ii}v_i$. Portanto, v_i é um vector próprio de T associado ao valor próprio a_{ii}. □

Por outras palavras: no caso em que uma transformação $T : V \to V$, com $\dim(V) = n$, é diagonalizável, existe uma base $B = \{v_1, v_2, \ldots, v_n\}$ constituída por vectores próprios na qual a representação matricial de T é a matriz diagonal

$$T_{BB} = \begin{bmatrix} \lambda(v_1) & 0 & \ldots & 0 \\ 0 & \lambda(v_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda(v_n) \end{bmatrix}$$

onde $\lambda(v_i)$ é o valor próprio associado ao vector próprio v_i.

Já tínhamos aliás referido este resultado várias vezes em exemplos anteriores. A seguinte propriedade, que não demonstraremos, é muito útil para mostrar que determinadas transformações não são diagonalizáveis.

Proposição 59. Seja λ um valor próprio de uma transformação linear $T : V \to V$. Então $1 \leq m_g(\lambda) \leq m_a(\lambda)$.

Exercício 35. Seja A uma matriz 6×6 com polinómio característico $\lambda^2(\lambda - 1)(\lambda - 2)^3$. Determine os valores próprios de A e a multiplicidade associada a cada um. Quais as possíveis dimensões dos espaços próprios de A?

Observe-se que o polinómio característico de T é sempre um polinómio de grau n, em que n é a dimensão de V, pelo que a soma das multiplicidades aritméticas é sempre no máximo n (pelo Teorema Fundamental da Álgebra). Daqui temos duas consequências importantes.

- Se o polinómio característico de T tem menos do que n raízes reais (contando com as suas multiplicidades), então T não é diagonalizável.

- Se para algum valor próprio λ de T se tiver $m_g(\lambda) < m_a(\lambda)$, então T não é diagonalizável.

Em particular, se T tem exactamente n valores próprios, então T é diagonalizável.

Por exemplo: na secção anterior considerámos o exemplo da transformação linear $T : P_1 \to P_1$ definida por $Tp = p + p'$ e vimos que esta tinha apenas o valor próprio 1, com $m_a(1) = 2$ e $m_g(1) = 1$. Podemos então concluir T não é diagonalizável.

Exemplo.

1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear definida pela expressão $T(x, y) = (3x + y, -2x)$.

   A representação de T em relação à base canónica B de $\mathbb{R}^2$ é a matriz

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

   Para encontrar os valores próprios de T temos de resolver a equação $\det(T_{BB} - \lambda \mathbf{I_2}) = 0$.

$$\begin{aligned} \det(T_{BB} - \lambda \mathbf{I_2}) = 0 &\iff \det\left(\begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0 \\ &\iff \det\begin{bmatrix} 3-\lambda & 1 \\ -2 & -\lambda \end{bmatrix} = 0 \\ &\iff -\lambda(3-\lambda) + 2 = 0 \iff \lambda^2 - 3\lambda + 2 = 0 \iff \lambda = 1 \text{ ou } \lambda = 2 \end{aligned}$$

   Uma vez que a transformação T tem dois valores próprios distintos, sabemos imediatamente que é diagonalizável, uma vez que $\dim(\mathbb{R}^2) = 2$. Vamos agora determinar os espaços próprios associados a cada um dos valores próprios.

   O subespaço próprio E_1 corresponde ao núcleo de $T_{BB} - \lambda \mathbf{I_2}$, ou seja, ao conjunto das soluções do sistema homogéneo associado àquela matriz.

$$\left[\begin{array}{cc|c} 2 & 1 & 0 \\ -2 & -1 & 0 \end{array}\right] \begin{array}{l} \\ l_1 + l_2 \end{array} \longrightarrow \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \longrightarrow \left\{ y = -2x \right.$$

   Logo $E_1 = L(\{(1, -2)\})$. Este espaço próprio tem dimensão 1, logo vamos escolher um vector próprio associado a $\lambda = 1$, por exemplo $v_1 = (1, -2)$.

Para determinar E_2, procedemos de forma análoga.

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ -2 & -2 & 0 \end{array}\right] \underset{2l_1+l_2}{\longrightarrow} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \longrightarrow \left\{ y = -x \right.$$

Logo $E_2 = L(\{(1,-1)\})$. Este espaço próprio tem dimensão 1, logo vamos escolher um vector próprio associado a $\lambda = 2$, por exemplo $v_2 = (1,-1)$.

Obtemos assim a seguinte base de $\mathbb{R}^2$ formada por vectores próprios de T.

$$C = \{(1,-2),(1,-1)\}$$

A representação matricial de T na base C é a matriz

$$T_{CC} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

ou seja, é a matriz diagonal que tem na diagonal principal os valores próprios de T, pela ordem dos vectores a que estão associados na base C.

Exercício 36. Para cada uma das seguintes transformações lineares, determine os seus valores próprios, os espaços próprios associados a cada valor próprio e determine se a transformação é diagonalizável e, em caso afirmativo, qual a base em que a transformação é representada por uma matriz diagonal.

(a) $T : P^2 \to P^2$ com $T(ax^2 + bx + c) = (a-2c)x^2 + (b+2c)x + (a+4c)$

(b) $T : M_{2\times 2} \to M_{2\times 2}$ com $T(A) = A + A^T$

(c) $T : M_{2\times 2} \to M_{2\times 2}$ com $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} -2c & a+c \\ -b-2c & d \end{bmatrix}$

Seja agora $T : V \to V$ uma transformação linear diagonalizável cuja representação matricial em relação a uma base B (não necessariamente de vectores próprios) é a matriz T_{BB}, e seja C uma outra base de V tal que T_{CC} é diagonal. Se P_{CB} e $P_{BC} = (P_{CB})^{-1}$ forem as matrizes de mudança de base que permitem permutar coordenadas de B com coordenadas de C, tem-se a relação

$$T_{CC} = P_{CB}T_{BB}P_{BC} = P_{CB}T_{BB}(P_{CB})^{-1}.$$

Vamos usar esta relação para definir conceitos análogos para matrizes.

Definição. Duas matrizes quadradas A e B de dimensão n dizem-se *semelhantes* se existir uma matriz P, também quadrada e de dimensão n, tal que

$$A = PBP^{-1}.$$

Em particular, a matriz A diz-se *diagonalizável* se for semelhante a uma matriz diagonal.

As matrizes semelhantes têm esse nome por partilharem uma série de propriedades, como sejam o determinante, a nulidade ou a característica. Transpondo para matrizes os resultados acima expostos para transformações lineares, concluímos que uma matriz A de dimensão n é diagonalizável quando existe uma base $B = \{v_1, v_2, \ldots, v_n\}$ de $\mathbb{R}^n$ formada apenas por vectores próprios de A. Nesse caso, tem-se

$$A = \begin{bmatrix} \vdots & \vdots & & \vdots \\ v_1 & v_2 & \ldots & v_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \lambda(v_1) & 0 & \ldots & 0 \\ 0 & \lambda(v_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda(v_n) \end{bmatrix} \begin{bmatrix} \vdots & \vdots & & \vdots \\ v_1 & v_2 & \ldots & v_n \\ \vdots & \vdots & & \vdots \end{bmatrix}^{-1}$$

onde $\lambda(v_i)$ é o valor próprio associado ao vector próprio v_i.

Exemplo. Reanalisemos os exemplos anteriores em termos de matrizes.

1. A matriz
$$A = \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$$
é uma matriz diagonalizável, já que representa uma transformação diagonalizável de $\mathbb{R}^2$ para $\mathbb{R}^2$. De facto, temos que $A = PDP^{-1}$ sendo D a representação da mesma transformação em relação à base de valores próprios e P a matriz de mudança de base dessa base para a base canónica de $\mathbb{R}^2$.
$$D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad P = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}$$
Podemos verificar este facto:
$$P^{-1} = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} \qquad PDP^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 2 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix} = A\,.$$

Exercício 37. Determine se cada uma das matrizes dos Exercícios 34 e 36 é diagonalizável e, em caso afirmativo, indique qual a matriz diagonalizante.

3.7 Exercícios e problemas

Definição de transformações lineares

38. Para cada uma das seguintes transformações lineares definidas num espaço vectorial V, calcule o valor pedido sabendo os valores nos pontos indicados.

 (a) $T : V \to \mathbb{R}^3$
 $$\begin{cases} T(u) = (1, -1, 2) \\ T(v) = (0, 3, 2) \\ T(w) = (-3, 1, 2) \end{cases}$$
 Qual o valor de $2u - 3v + 4w$?

 (b) $T : V \to V$
 $$\begin{cases} T(u) = 3v + w \\ T(v) = 2u \\ T(w) = w \end{cases}$$
 Qual o valor de $T(2u - v - w)$?

(c) $T : V \to M_{3\times 2}$

$T(u) = \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$ e $T(v) = \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 1 & 1 \end{bmatrix}$

Qual o valor de $T(2u + v)$?

39. Seja $T : \mathbb{R}^2 \to P_2$ uma transformação tal que

$$T(1,0) = x^2 + 2x + 1 \qquad T(1,1) = x^2 - 3x + 3 \qquad T(2,3) = 2x^2 - 10x + 8.$$

Justifique que T não é uma transformação linear.

40. Represente matricialmente as seguintes transformações lineares relativamente às bases canónicas dos espaços de partida e de chegada.

(a) $T : \mathbb{R}^2 \to \mathbb{R}^3$ com $T(x, y) = (x + 2y, -x, 0)$

(b) $T : \mathbb{R}^2 \to \mathbb{R}^3$ com $T(x, y) = (2x - y, 2x + y, 3y)$

(c) $T : \mathbb{R}^3 \to \mathbb{R}^3$ com $T(x, y, z) = (x - y, y - z, x - z)$

(d) $T : M_{2\times 2} \to M_{2\times 2}$ com $T(A) = A^T$

(e) $T : M_{2\times 2} \to M_{2\times 2}$ com $T(X) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} X + X \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$

(f) $T : M_{2\times 3} \to M_{2\times 3}$ com $T(A) = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} A$

(g) $T : M_{2\times 2} \to P_2$ com $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ax^2 - d$

(h) $T : P_2 \to P_2$ com $T(p(x)) = p(2x + 1)$

(i) $T : P_2 \to P_3$ com $T(p(x)) = xp(x - 3)$

41. Use as matrizes determinadas no Exercício 40 para calcular os seguintes valores.

(a) $T(2 - 3x + 4x^2)$, com a transformação T definida no Exercício (40h)

(b) $T(1 + x + x^2)$, com a transformação T definida no Exercício (40h)

(c) $T(2 - 3x + 4x^2)$, com a transformação T definida no Exercício (40i)

(d) $T(1 + x + x^2)$, com a transformação T definida no Exercício (40i)

42. Represente matricialmente as transformações lineares do Exercício 40 relativamente às bases B e C apresentadas para os espaços de partida e chegada.

(a) $B = \{(1, 3), (-2, 4)\}$ e $C = \{(1, 1, 1), (2, 2, 0), (3, 0, 0)\}$

(b) $B = \{(2, 1), (-1, 1)\}$ e $C = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

(c) $B = C = \{(1, 0, 1), (0, 1, 1), (1, 1, 0)\}$

(d) $B = C = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$

(e) $B = C = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$

(g) $B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ e $C = \{x^2 + x + 1, x + 1, 1\}$

(h) $B = \{2x + 3, x^2 - 1, x + x^2 - 2\}$ e $C = \{1, x, x^2\}$

(i) $B = \{1, x, x^2\}$ e $C = \{x^3, x^2, x, 1\}$

43. Para cada uma das alíneas seguintes, seja T a transformação linear representada pela matriz apresentada em relação às bases indicadas. Escreva a expressão matricial de T em relação às bases canónicas dos espaços envolvidos.

(a) $B = \{(1, 3), (-1, 4)\}$
$T_{BB} = \begin{bmatrix} 1 & 3 \\ -2 & 5 \end{bmatrix}$

(b) $B = \{3x + 3x^2, -1 + 3x + 2x^2, 3 + 7x + 2x^2\}$
$T_{BB} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 5 \\ 6 & -2 & 4 \end{bmatrix}$

(c) $B = \{(0, 1, 1, 1), (2, 1, -1, -1), (1, 4, -1, 2), (6, 9, 4, 2)\}$
$C = \{(0, 8, 8), (-7, 8, 1), (-6, 9, 1)\}$
$T_{CB} = \begin{bmatrix} 3 & -2 & 1 & 0 \\ 1 & 6 & 2 & 1 \\ -3 & 0 & 7 & 1 \end{bmatrix}$

44. Seja V um espaço vectorial com base $B = \{v_1, v_2, v_3, v_4\}$ e T uma transformação linear tal que

$$T(v_1) = v_2 \qquad T(v_2) = v_3 \qquad T(v_3) = v_4 \qquad T(v_4) = v_1 .$$

Determine a matriz T_{BB}.

45. Para cada uma das alíneas seguintes, escreva as matrizes de mudança de base das bases B e C indicadas relativamente à base canónica. Usando essas matrizes, escreva a matriz T_{CB} que representa a transformação linear T indicada em relação a essas bases.

(a) $B = \{(1, 1), (2, 1)\}$
$C = \{(0, 1), (1, 1)\}$
$T : \mathbb{R}^2 \to \mathbb{R}^2$
$T(x, y) = (3x, 2y)$

(b) $B = \{(2, 2), (4, -1)\}$
$C = \{(1, 3), (-1, -1)\}$
$T : \mathbb{R}^2 \to \mathbb{R}^2$
$T(x, y) = (x + 7y, 3x - 4y)$

(c) $B = \{6 + 3x, 10 + 2x\}$
$C = \{2, 3 + 2x\}$
$T : P_2 \to P_2$
$T(p(x)) = p(x + 1)$

(d) $B = \{(1, -1), (1, 1)\}$
$C = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$
$T : \mathbb{R}^2 \to \mathbb{R}^3$
$T(x, y) = \left(x - y, x - y, \frac{3}{2}x - \frac{1}{2}y\right)$

Operações algébricas

46. Considere as seguintes transformações lineares.

$$T_1 : \mathbb{R}^2 \to \mathbb{R}^3 \qquad T_1(x,y) = (2x-y, 2x+y, 3y)$$

$$T_2 : \mathbb{R}^3 \to \mathbb{R}^2 \qquad T_2(x,y,z) = (z-2x, 3y-2z)$$

$$T_3 : \mathbb{R}^3 \to \mathbb{R}^3 \qquad T_3(x,y,z) = (x-y, y-z, x-z)$$

$$T_4 : P_2 \to P_2 \qquad T_4(p(x)) = p(2x+1)$$

$$T_5 : P_2 \to P_3 \qquad T_5(p(x)) = xp(x-3)$$

$$T_6 : P_2 \to P_3 \qquad T_6(p(x)) = x^2p'(x)$$

(a) Indique para que valores de n e m é que $T_n + T_m$ está definida e calcule a sua expressão.

(b) Indique para que valores de n e m é que $T_n \circ T_m$ está definida e calcule a sua expressão.

(c) Indique para que valores de n é que T_n é invertível e calcule a expressão de T_n^{-1}.

(d) Verifique os resultados obtidos nas alíneas anteriores recorrendo à representação matricial das transformações em causa.

Núcleo, imagem e invertibilidade

47. Considere a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$T(x,y) = (3x-2y, 0)\,.$$

(a) Verifique se os vectores $(0,1)$, $(1,-4)$, $(1,2)$, $(5,0)$ e $(-2,-3)$ pertencem ao núcleo de T.

(b) Determine uma base para $N(T)$.

(c) Verifique se os vectores $(5,10)$, $(3,2)$, $(1,1)$, $(4,0)$ e $(-1,-2)$ pertencem à imagem de T.

(d) Determine uma base para $\mathrm{im}(T)$.

48. Considere a transformação linear $T : \mathbb{R}^4 \to \mathbb{R}^3$ definida por

$$T(x,y,z,w) = (4x+y-2z-3w, 2x+y+z-4w, 6x-9z-9w)\,.$$

(a) Verifique se os vectores $(3,-8,2,0)$, $(0,0,0,1)$ e $(0,-4,1,0)$ pertencem ao núcleo da transformação.

(b) Determine uma base para $N(T)$.

(c) Verifique se os vectores $(0,0,6)$, $(1,3,0)$ e $(2,4,1)$ pertencem à imagem da transformação.

(d) Determine uma base para $\mathrm{im}(T)$.

49. Considere a transformação linear $T : P_3 \to P_3$ definida por

$$T(p(x)) = p(2x+1) + 2p'(x)\,.$$

   (a) Verifique se os vectores x^2+3x, $2x^3-1$, 2, $3x-1$ e $3x^3+2x-1$ pertencem ao núcleo de T.
   (b) Determine uma base para $N(T)$.
   (c) Verifique se os vectores $3x^3+x$, x^2-2x+1, 0, $4x$ e $2x+3$ pertencem à imagem de cada uma das transformações.
   (d) Determine uma base para $\mathrm{im}(T)$.

50. Considere as transformações lineares T_1, T_2, T_3 definidas num espaço vectorial arbitrário V por

$$T_1(v) = 3v \qquad T_2(v) = \frac{v}{3} \qquad T_3(v) = 0\,.$$

Determine o núcleo e a imagem de cada uma das transformações.

51. Recorrendo às representações matriciais determinadas no Exercício 40, encontre o núcleo e a imagem de cada uma das transformações lineares consideradas.

Valores e vectores próprios

52. Verifique se o vector $(1,1,1)$ é um vector próprio das seguintes transformações lineares de $\mathbb{R}^3$ para $\mathbb{R}^3$. Em caso afirmativo, qual é o valor próprio que lhe está associado?

   (a) $T(x,y,z) = (0, x-y, 2z-x-y)$
   (b) $T(x,y,z) = (x, z, x+y)$
   (c) $T(x,y,z) = (3x+y, x+3z, x+y+z)$
   (d) $T(x,y,z) = (3x+y, x+3z, x+y+2z)$

53. Para cada uma das seguintes matrizes, determine os seus valores próprios, os espaços próprios associados a cada valor próprio e determine se a matriz é diagonalizável e, em caso afirmativo, qual a matriz diagonalizante.

   (a) $\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$

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

   (c) $\begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}$

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

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

Problemas

54. Considere a matriz seguinte.

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

Encontre uma combinação de valores de α e β tal que:

   (a) a característica de A é 1;
   (b) a dimensão do espaço das colunas de A é 2;

(c) A é a representação em relação às bases canónicas de $\mathbb{R}^3$ e $\mathbb{R}^2$ da transformação linear T tal que

$$T(1,0,0) = (1,3) \qquad T(0,1,0) = (2,2) \qquad T(0,0,1) = (-1,1)\,.$$

55. Dê exemplos, justificando, de transformações lineares $T : \mathbb{R}^2 \to \mathbb{R}^2$ tais que:

(a) $T(1,0) = (0,1)$ e $T(0,1) = (1,0)$;

(b) $(2,2)$ é vector próprio de T;

(c) -1 é valor próprio de T.

56. Considere uma transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

$$T(1,1,1) = (2,2,2) \qquad T(1,1,-1) = (-1,-1,0) \qquad T(1,0,0) = (0,0,-1)\,.$$

(a) Justifique que $(1,1,1)$ é vector próprio de T e indique o valor próprio associado.

(b) Calcule $T(2,2,0)$, $T(0,0,1)$, $T(0,1,0)$ e $T(5,3,-1)$.

(c) Determine se $(2,2,0)$ e $(5,3,-1)$ pertencem ao núcleo de T.

(d) Determine se $(1,1,2)$ e $(1,0,0)$ pertencem à imagem de T.

(e) A transformação T é invertível?

(f) Determine a expressão de $T(x,y,z)$.

(g) Escreva a matriz de mudança de base S que converte as coordenadas de vectores na base $\{(1,1,1),(1,1,-1),(1,0,0)\}$ para coordenadas canónicas.

(h) Escreva uma expressão para a matriz que representa a transformação T na base $\{(1,1,1),(1,1,-1),(1,0,0)\}$.

57. Considere uma transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

$$T(1,0,0) = (0,0,1) \qquad T(1,1,0) = (0,2,2) \qquad T(1,1,1) = (3,3,3)\,.$$

(a) Calcule $T(0,1,0)$ e $T(0,0,1)$.

(b) Determine a expressão geral de $T(x,y,z)$.

(c) Encontre a matriz que representa T em relação à base canónica de $\mathbb{R}^3$.

(d) Determine os valores próprios de T e os espaços próprios que lhes estão associados.

(e) Existe alguma base de $\mathbb{R}^3$ em que T é representada por uma matriz diagonal?

(f) A transformação T é invertível?

58. Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ cuja representação, relativamente às bases canónicas $\{(1,0,0),(0,1,0),(0,0,1)\}$ de $\mathbb{R}^3$ e $\{(1,0),(0,1)\}$ de $\mathbb{R}^2$, é dada pela seguinte matriz.

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

(a) Calcule $T(1,0,2)$, $T(-3,2,1)$ e $T(3,1,0)$.

(b) Indique um vector no núcleo de T e outro na imagem de T.

(c) Encontre uma base para a imagem de T e para o núcleo de T.

(d) Determine se $x^2 - 3x + 1$ pertence à imagem de T.

(e) Determine a expressão geral de $T(x, y, z)$.

59. Seja S o subconjunto de $\mathbb{R}^3$ gerado pelos vectores $\overrightarrow{v_1} = (1, 1, 1)$, $\overrightarrow{v_2} = (-1, -1, 2)$ e $\overrightarrow{v_3} = (1, 1, 4)$. Considere a transformação $T : \mathbb{R}^3 \to \mathbb{R}^3$ que transforma cada vector $\overrightarrow{v}$ na sua projecção ortogonal sobre S. Observe que $T(v) = P_{\overrightarrow{v_1}}(\overrightarrow{v}) + P_{\overrightarrow{v_2}}(\overrightarrow{v})$

   (a) Mostre que T é transformação linear sabendo que a projecção ortogonal sobre um vector é uma transformação linear.

   (b) Mostre que $T(x, y, z)$ pode ser calculado pela seguinte expressão.

$$T(x, y, z) = \frac{x + y + z}{3}(1, 1, 1) + \frac{-x - y + 2z}{6}(-1, -1, 2)$$

   (c) Indique o núcleo e a imagem de T. Note que o vector $(2, -2, 1)$ é perpendicular a S.

   (d) Os valores próprios de T são 0 e 1. Encontre os vectores próprios associados a cada um destes vectores.

   (e) Indique uma possível representação matricial de T numa base de vectores próprios.

60. Considere uma transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ cuja representação matricial em relação às bases canónicas de $\mathbb{R}^3$ e $\mathbb{R}^2$ é dada por

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

   (a) Calcule $T(1, 0, 2)$ e $T\left(2, -1, \frac{1}{2}\right)$.

   (b) Determine a expressão de T para qualquer vector (x, y, z) de $\mathbb{R}^3$.

   (c) Determine o núcleo da transformação T.

   (d) A transformação T é injectiva? É invertível?

61. Considere a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T(x, y, z) = (x - y, 2x - 2y, -5x + 2y - 6z)\,.$$

   (a) Encontre a matriz que representa T em relação à base canónica de $\mathbb{R}^3$.

   (b) Indique uma base para a imagem de T.

   (c) Determine o núcleo da transformação T.

   (d) Determine os valores próprios de T e os espaços próprios que lhes estão associados.

   (e) Existe alguma base de $\mathbb{R}^3$ em que T é representada por uma matriz diagonal?

62. Sejam S o plano de $\mathbb{R}^3$ gerado pelos vectores $(1, 2, 0)$ e $(3, 1, 1)$ e $T : \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear que transforma cada vector de $\mathbb{R}^3$ na sua projecção ortogonal sobre S.

   Determine a representação matricial de T numa base de vectores próprios.

   Sugestão. Recorde que, dados dois vectores $\overrightarrow{u}$ e $\overrightarrow{v}$ de $\mathbb{R}^3$, o seu produto externo $\overrightarrow{u} \times \overrightarrow{v}$ é perpendicular quer a $\overrightarrow{u}$ quer a $\overrightarrow{v}$.

63. Sejam V o subconjunto de P_3 constituído pelos polinómios p de grau menor ou igual a 3 tais que $p(x) = ax^3 + bx + c$ e $T : V \to V$ a transformação linear definida por

$$T(p(x)) = xp'(x)\,.$$

   (a) Verifique se $q(x) = 3x - 2$ pertence ao núcleo de T.
   (b) Verifique se $r(x) = 6x^3 - x$ pertence à imagem de T.
   (c) Verifique que qualquer polinómio da forma $p(x) = ax$ é vector próprio de T e calcule o valor próprio que lhe está associado.
   (d) Determine a representação matricial de T em relação à base $\{1, x, x^3\}$ de V.
   (e) Determine a representação matricial de T em relação à base B de V.
   (f) Calcule as coordenadas de $T\left(3x^3 + 2x - 5\right)$ na base B.

64. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear tal que

$$T(1,1,0) = (3,1,2) \qquad T(0,1,1) = (2,3,1) \qquad T(1,2,-1) = (5,0,1)$$

Determine o espaço próprio associado ao valor próprio 3.

65. Considere a transformação linear $T : P_3 \to P_3$ cuja representação em relação à base

$$\left\{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\right\}$$

de P_3 é a matriz

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

   (a) Verifique se $2x^2 + 6x + 1$ pertence ao núcleo de T.
   (b) Determine a expressão geral de $T\left(a + bx + cx^2 + dx^3\right)$.

Bibliografia

[1] H. Anton e C. Rorres, *Elementary Linear Algebra*. Wiley, 2005.

[2] L.T. Magalhães, *Álgebra Linear como Introdução à Matemática Aplicada*. Texto Editora, 1992.