Álgebra Linear

Jorge Orestes Cerdeira

Instituto Superior de Agronomia

- 2012 -

Conteúdo

Capítulo 1

Cálculo matricial

As matrizes são objetos matemáticos que ocorrem em inúmeras situações. Neste capítulo vamos introduzir a noção de matriz de elementos reais e definir operações básicas com matrizes. Começamos por estudar os sistemas de equações lineares com várias variáveis.

1.1 Sistemas de equações lineares

Vamos recordar as noções de equação linear, sistema de equações lineares e resolução de um sistema de equações e introduzir um método para classificar e resolver sistemas de equações lineares que nos vai ajudar na interpretação geométrica do sistema.

Exemplos 1

1. A equação $3x - y = 1$ é linear com duas variáveis. Os pontos (x, y) do plano que a satisfazem definem a reta perpendicular ao vetor $(3, -1)$ e que inclui o ponto $(0, -1)$.

2. A equação $x - y + 5z = 0$ é linear com três variáveis. Os pontos (x, y, z) de $\mathbb{R}^3$ que a satisfazem definem um plano perpendicular ao vetor $(1, -1, 5)$ e que passa na origem do sistema de eixos.

Uma *equação linear* com n variáveis $x_1, x_2, \ldots, x_n$ é uma equação que pode ser escrita na forma

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

com $a_1, a_2, \ldots, a_n$, não todos nulos, e b números reais. Diz-se que $x_1, x_2, \ldots, x_n$ são as variáveis ou incógnitas, a_i é o coeficiente da variável x_i e b é membro direito ou termo constante.

Exemplos 2

1. As equações $x + 8y - 4z + 3t = 12$ e $x_1 - 3x_2 + 5x_4 = \frac{5}{2}$ são exemplos de equações lineares com quatro variáveis.

2. As equações $x^2 - 2xy + 3y = 1$, $x_1 - \frac{1}{x_2} + x_3 = 0$, $e^{x_1} - x_2 = \frac{1}{2}$ são exemplos de equações não lineares.

Uma *solução* de uma equação é uma sequência de valores que atribuídos às variáveis transforma a equação numa proposição verdadeira. Assim, $(0, 0, 1)$ e $(-\frac{1}{2}, 1, 1)$ são soluções da equação $2x_1 + x_2 + x_3 = 1$.
Resolver uma equação é determinar o conjunto das soluções. Para resolver a equação

$$2x_1 + x_2 + x_3 = 1$$

podemos escolher duas variáveis para tomarem valores arbitrários (variáveis livres) e resolver a equação em ordem à variável restante (variável dependente). O conjunto das soluções é pois o plano definido por

$$\Big\{(x_1, x_2, x_3): \begin{array}{l} x_1 = \forall \\ x_2 = 1 - 2x_1 - x_3 \\ x_3 = \forall \end{array} \Big\}.$$

Observação 1 Se uma equação linear tem uma única variável, não há variáveis livres, o conjunto de soluções é singular e consiste num único ponto de $\mathbb{R}$. Se é uma equação linear com duas variável, tem uma variável livre e o conjunto de soluções define uma reta em $\mathbb{R}^2$. Com três variáveis, há duas variáveis livres e define um plano de $\mathbb{R}^3$. Quando a equação tem $n > 3$ variáveis, existem $n - 1$ variáveis livres e diz-se que o conjunto das soluções é um *hiperplano* de $\mathbb{R}^n$.

Um *sistema de equações lineares* é uma coleção finita de equações lineares com as mesmas variáveis. O sistema de equações lineares

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

tem m equações e n variáveis. Diz-se do tipo m por n e escreve-se $m \times n$. Diz-se também que a_{ij} é o coeficiente da variável j na equação i e que b_i é o membro direito da equação i.

Uma *solução* do sistema é uma solução comum às m equações.

Exemplo 3 $(-2, 0, 6)$ e $(2, 1, 1)$ são soluções do sistema de equações lineares 2×3

$$\begin{array}{rcl} x_1 + x_2 + x_3 & = & 4 \\ 3x_1 - 2x_2 + 2x_3 & = & 6 \end{array}$$

$(1, 1, 2)$ não é solução.

Se o sistema admite uma única solução diz-se *determinado*. É *indeterminado* se tem mais do que uma solução. Se não existem soluções o sistema diz-se *impossível*.

Exercícios 1

1. Para que valores de b o sistema

   $$\begin{array}{rcl} 2x_1 + 3x_2 & = & 4 \\ 4x_1 + 6x_2 & = & b \end{array}$$

   é impossível?

2. Indique uma equação a juntar a

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

de forma a obter um sistema impossível.

Dois sistemas são *equivalentes* se têm as mesmas soluções. Se num sistema de equações

1. adicionar a uma equação um múltiplo de outra,
2. multiplicar uma equação por uma constante não nula, ou
3. trocar duas equações,

obtem um sistema equivalente. As *operações* 1, 2 e 3 chamam-se *elementares*.

Exemplo 4 Os sistemas

$$A = \begin{cases} x - y &= 1 \\ \frac{1}{2}x + y &= 2 \end{cases} \quad \text{e} \quad B = \begin{cases} x - \ y &= 1 \\ -\frac{1}{2}x + 2y &= 1 \end{cases},$$

representados geometricamente na Figura 1.1 a) e b), respetivamente, são equivalentes.

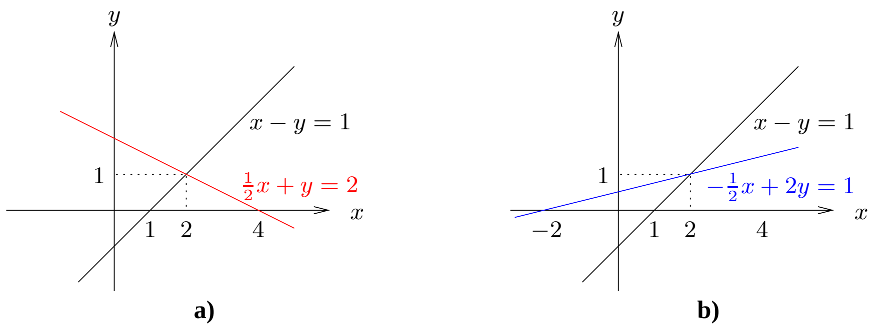


Figura 1.1: Representação das equações dos sistemas: a) A e b) B do Exemplo 4.

Note que a 2ª equação do sistema B obtem-se adicionando à 2ª equação de A a 1ª multiplicada por -1.

Resolver um sistema é determinar o conjunto das soluções. (O conjunto de soluções de um sistema impossível é vazio.) Para resolver um sistema identificam-se, caso existam, variáveis que podem tomar valores arbitrários (*variáveis livres*) e exprime-se, em função dessas variáveis, cada uma das restantes (*variáveis dependentes*).

O método de eliminação de Gauss resolve sistemas de equações lineares por aplicação sucessiva das operações elementares. O método decorre em duas fases. A fase descendente termina com um sistema em escada equivalente ao inicial. Num *sistema em escada* o 1º coeficiente não nulo de cada equação "está mais à direita" do que o 1º coeficiente não nulo da equação anterior.

Exemplo 5

$$S' = \left\{ \begin{array}{rcr} x_1 - 3x_2 + 3x_3 + \ x_4 + \ x_5 & = & 34 \\ -\,7x_3 - 4x_4 - 7x_5 & = & -76 \\ 6x_4 \qquad\quad & = & 30 \end{array} \right.$$

é um sistema em escada.

O sistema em escada S' do Exemplo 5 é equivalente ao sistema

$$S = \left\{ \begin{array}{rcr} x_1 - 3x_2 + 3x_3 + \ x_4 + \ \ x_5 & = & 34 \\ 2x_1 - 6x_2 - \ x_3 - 2x_4 - \ 5x_5 & = & -8 \\ 3x_1 - 9x_2 - 5x_3 + \ x_4 - 11x_5 & = & -20 \end{array} \right.$$

que vamos usar para ilustrar a descrição do método de eliminação de Gauss.

O método começa por utilizar a 1ª equação para "eliminar" a 1ª variável nas restantes equações, executando para este efeito operações elementares do tipo 1. Para eliminar a variável x_1 nas equações 2 e 3 de S substituem-se a equação 2 pela sua soma com -2 vezes a equação 1 e a equação 3 pela sua soma com -3 vezes a equação 1. Obtem-se assim o sistema

$$\left\{ \begin{array}{rcr} x_1 - 3x_2 + \ 3x_3 + \ x_4 + \ \ x_5 & = & 34 \\ -\ 7x_3 - 4x_4 - \ 7x_5 & = & -76 \\ -14x_3 - 2x_4 - 14x_5 & = & -122 \end{array} \right.$$

que é equivalente a S e em que a variável x_1 "não figura" nas equações 2 e 3.

Este procedimento é agora repetido com o sistema que resulta de ignorar a 1ª equação do sistema obtido na iteração anterior. Com o sistema anterior o procedimento conduz ao sistema em escada

$$S' = \begin{cases} x_1 - 3x_2 + 3x_3 + x_4 + x_5 & = 34 \\ \quad - 7x_3 - 4x_4 - 7x_5 & = -76 \\ \quad 6x_4 & = 30 \end{cases},$$

ficando assim concluída a fase descendente do método de Gauss.

Uma vez obtido um sistema em escada equivalente ao inicial, as variáveis livres e as dependentes são identificadas pelas posições dos *pivots*. Os *pivots* de um sistema em escada são os primeiros coeficiente não nulos das equações. As correspondentes variáveis chamam-se *variáveis pivot*. Os *pivots* do sistema S' são 1, -7, 6 e x_1, x_3, x_4 são as variáveis *pivot*. As variáveis *pivot* vão ser selecionadas para dependentes e as restantes para variáveis livres.

A fase ascendente do método de eliminação de Gauss começa com o sistema em escada determinado na fase descendente e termina com a obtenção de um sistema reduzido equivalente. Um *sistema reduzido* é um sistema em escada com todos os *pivots* iguais a 1 e em que cada equação não inclui mais do que uma variável *pivot*.

Exemplo 6

$$R = \begin{cases} x_1 - 3x_2 \quad - 2x_5 & = 5 \\ \quad x_3 + x_5 & = 8 \\ \quad x_4 & = 5 \end{cases}$$

é um sistema reduzido.

Na fase ascendente atribui-se o valor 1 ao *pivot* da última equação e utiliza-se esta equação para "eliminar" a correspondente variável *pivot* nas restantes equações. Para fazer o *pivot* igual a 1 executa-se uma operação elementar do tipo 2 e com operações do tipo 1 "elimina-se" a variável *pivot* das outras equações.

Relativamente ao sistema S', em primeiro lugar multiplica-se a última equação por $\frac{1}{6}$, fazendo assim o *pivot* igual a 1 na equação que a substitui. Em seguida, a equação 1 vai ser substituída pela sua soma com -1 vezes a última equação, ficando assim "eliminada" a variável x_4 nessa equação. Também a equação 2 vai ser substituída pela sua soma com 4 vezes a última, o que leva à "eliminação" de x_4 dessa equação. Tem-se pois

$$\left\{\begin{array}{rcr} x_1 - 3x_2 + 3x_3 + \;\; x_4 + \;\; x_5 & = & 34 \\ -\,7x_3 - 4x_4 - 7x_5 & = & -76 \\ x_4 \qquad\quad & = & 5 \end{array}\right. \longrightarrow \left\{\begin{array}{rcr} x_1 - 3x_2 + 3x_3 \qquad + \;\; x_5 & = & 29 \\ -\,7x_3 \qquad - 7x_5 & = & -56 \\ x_4 \qquad\quad & = & 5 \end{array}\right. .$$

A fase ascendente continua repetindo aquele procedimento com o sistema que resulta de ignorar a última equação do sistema produzido na iteração anterior. O método termina quando o sistema obtido é reduzido. Assim,

$$\left\{\begin{array}{rcr} x_1 - 3x_2 + 3x_3 \qquad + x_5 & = & 29 \\ x_3 \qquad + x_5 & = & 8 \\ x_4 \qquad\quad & = & 5 \end{array}\right. \longrightarrow R = \left\{\begin{array}{rcr} x_1 - 3x_2 \qquad\qquad - 2x_5 & = & 5 \\ x_3 \qquad + \;\; x_5 & = & 8 \\ x_4 \qquad\quad & = & 5 \end{array}\right. .$$

O método de eliminação de Gauss determinou o sistema reduzido R que é equivalente a S. Identificar o conjunto das soluções de um sistema reduzido é trivial. Tudo o que há a fazer é isolar num dos membros de cada equação a única variável dependente que figura nessa equação. O conjunto das soluções do sistema R, e portanto de S, é

$$\Big\{(x_1, x_2, x_3, x_4, x_5): \begin{array}{l} x_1 = 5 + 3x_2 + 2x_5 \\ x_2 = \forall \\ x_3 = 8 - x_5 \\ x_4 = 5 \\ x_5 = \forall \end{array}\Big\}.$$

É de notar que, durante a fase descendente, poderá haver necessidade de efetuar operações elementares do tipo 3 (troca de equações). Esta situação ocorre no exemplo que se apresenta em seguida.

Exemplo 7 Vai-se aplicar o método de Gauss para resolver o sistema linear 3×4

$$S = \begin{cases} x_1 + 2x_2 + \ x_3 + \ x_4 = 4 \\ 2x_1 + 4x_2 - \ x_3 + 2x_4 = 11 \\ x_1 \qquad + 2x_3 + 3x_4 = 1 \end{cases}.$$

$$S = \begin{cases} x_1 + 2x_2 + x_3 + x_4 = 4 \\ 2x_1 + 4x_2 - x_3 + 2x_4 = 11 \\ x_1 + 2x_3 + 3x_4 = 1 \end{cases} \longrightarrow \begin{cases} x_1 + 2x_2 + x_3 + x_4 = 4 \\ -3x_3 = 3 \\ -2x_2 + x_3 + 2x_4 = -3 \end{cases} \xrightarrow{(*)}$$

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

$$\begin{cases} x_1 + 2x_2 + x_4 = 5 \\ -2x_2 + 2x_4 = -2 \\ x_3 = -1 \end{cases} \longrightarrow \begin{cases} x_1 + 2x_2 + x_4 = 5 \\ x_2 - x_4 = 1 \\ x_3 = -1 \end{cases} \longrightarrow$$

$$R = \begin{cases} x_1 + 3x_4 = 3 \\ x_2 - x_4 = 1 \\ x_3 = -1 \end{cases}.$$ $(*)$ - realizou-se a troca das equações 2 e 3.

O conjunto de soluções do sistema é $\left\{(x_1, x_2, x_3, x_4) : \begin{array}{l} x_1 = 3 - 3x_4 \\ x_2 = 1 + x_4 \\ x_3 = -1 \\ x_4 = \forall \end{array} \right\}.$

Exemplo 8 Vamos agora aplicar o método de eliminação de Gauss ao sistema A, do tipo 2×2, do Exemplo 4. Na Figura 1.2 apresentam-se as representações geométricas das equações dos sistemas obtidos durante a execução do método.

$$A = \begin{cases} x - y = 1 \\ \frac{1}{2}x + y = 2 \end{cases} \longrightarrow A' = \begin{cases} x - y = 1 \\ \frac{3}{2}y = \frac{3}{2} \end{cases} \longrightarrow A'' = \begin{cases} x - y = 1 \\ y = 1 \end{cases} \longrightarrow$$

$$A''' = \left\{ \begin{array}{rcl} x & = & 2 \\ y & = & 1 \end{array} \right.$$

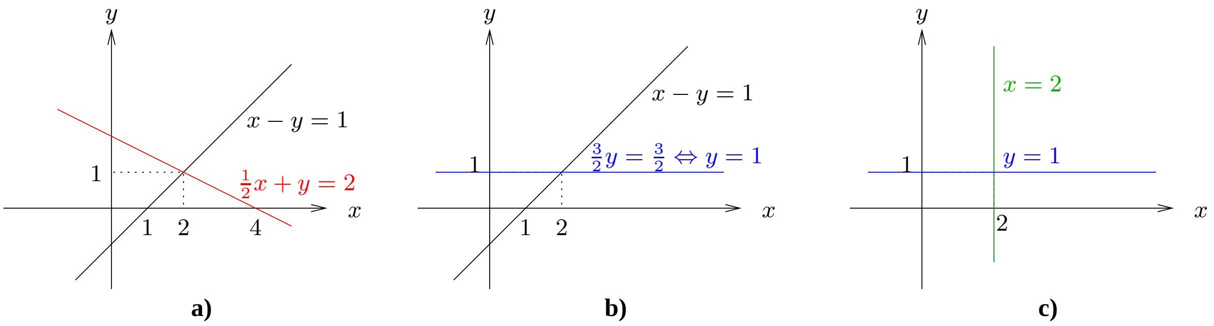


Figura 1.2: Representação das equações dos sistemas: a) A, b) A' e A'' e c) A''' do Exemplo 8.

Exercícios 2 Resolva cada um dos seguintes sistemas de equações lineares.

1. $\left\{ \begin{array}{rcl} x + 2y + 3z & = & 0 \\ x + y + z & = & 10 \\ x \quad\quad + 2z & = & 0 \end{array} \right.$

2. $\left\{ \begin{array}{rcl} x_1 + 2x_2 + 3x_3 & = & 6 \\ 2x_1 + 5x_2 + x_3 & = & 9 \\ x_1 + 4x_2 - 6x_3 & = & 1 \end{array} \right.$

3. $\left\{ \begin{array}{rcl} x_1 + 2x_2 + x_3 + x_4 & = & 4 \\ 2x_1 + 4x_2 - x_3 + 2x_4 & = & 11 \\ x_1 + x_2 + 2x_3 + 3x_4 & = & 1 \end{array} \right.$

4. $\left\{ \begin{array}{rcl} x_1 + 2x_2 + 3x_3 + x_4 - x_5 & = & 2 \\ x_3 + x_4 + x_5 & = & -1 \\ -x_1 - 2x_2 - x_3 + 2x_4 + x_5 & = & 0 \end{array} \right.$

5. $\left\{ \begin{array}{rcl} x_1 + x_2 + x_3 + x_4 & = & 1 \\ 2x_1 + x_2 - x_3 + 2x_4 & = & 9 \\ x_1 + 2x_2 + x_3 - x_4 & = & -6 \\ x_1 + x_2 - 2x_3 + x_4 & = & 7 \end{array} \right. .$

Vejamos como o método de eliminação de Gauss se comporta com sistemas impossíveis e perante a existência de equações redundantes.

O sistema

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

é impossível. De facto, a equação $3x_1 + 3x_2 - 5x_3 = 6$, que é incompatível com a 3ª equação do sistema, é 2 vezes a 2ª equação menos a 1ª. Aplicando o método de Gauss tem-se

$$\begin{cases} x_1 - x_2 + 3x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 4 \\ 3x_1 + 3x_2 - 5x_3 &= 1 \end{cases} \longrightarrow \begin{cases} x_1 - x_2 + 3x_3 &= 2 \\ 3x_2 - 7x_3 &= 0 \\ 6x_2 - 14x_3 &= -5 \end{cases} \longrightarrow \begin{cases} x_1 - x_2 + 3x_3 &= 2 \\ 3x_2 - 7x_3 &= 0 \\ 0 &= -5 \end{cases},$$

em que a equação 3 deu lugar à proposição falsa $0 = -5$.

Como anteriormente foi referido a equação $3x_1 + 3x_2 - 5x_3 = 6$ obtem-se subtraindo a 1ª equação ao dobro da 2ª. Assim, se em S substituirmos a 3ª equação por $3x_1+3x_2-5x_3 = 6$, esta equação torna-se redundante no sistema resultante. Ao aplicar o método de Gauss a este sistema obtem-se o seguinte resultado.

$$\begin{cases} x_1 - x_2 + 3x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 4 \\ 3x_1 + 3x_2 - 5x_3 &= 6 \end{cases} \longrightarrow \cdots \longrightarrow \begin{cases} x_1 - x_2 + 3x_3 &= 2 \\ 3x_2 - 7x_3 &= 0 \\ 0 &= 0 \end{cases}.$$

A equação redundante deu lugar à proposição verdadeira $0 = 0$ e o método permite concluir que o sistema inicial com três equações é equivalente ao sistema em escada com apenas 2 equações.

Em geral, quando o sistema é impossível é gerada uma proposição falsa do tipo $0 = a$, com $a \neq 0$. Equações redundantes originam proposições verdadeiras do tipo $0 = 0$.

Na Figura 1.3 apresenta-se um esquema para classificar sistemas de equações lineares, a partir do sistema em escada obtido no fim da fase descendente do método de Gauss.

· Sejam S um sistema de equações lineares e S' o sistema em escada obtido no fim da fase descendente do método de Gauss.

· Se S' inclui alguma proposição falsa ($0 = a$, com $a \neq 0$) $\Rightarrow S$ é **impossível**.
Caso contrário S é **possível**.

  · Se só há variáveis *pivot* $\Rightarrow S$ é **determinado**.
  Caso contrário S é **indeterminado**, com tantas variáveis livres quanto o número de variáveis sem *pivot*.

Figura 1.3: Esquema para classificar o sistema de equações lineares S.

Exemplo 9 Para classificar o sistema

$$S = \begin{cases} x_1 - x_2 = 1 \\ x_2 + 2x_3 = 1 \\ 2x_1 - 3x_2 = 0 \end{cases}$$

aplicámos a fase descendente do método de eliminação de Gauss, de que resultou

$$S \longrightarrow \begin{cases} x_1 - x_2 = 1 \\ x_2 + 2x_3 = 1 \\ -x_2 = -2 \end{cases} \longrightarrow S' = \begin{cases} x_1 - x_2 = 1 \\ x_2 + 2x_3 = 1 \\ 2x_3 = -1 \end{cases}.$$

Uma vez que o sistema em escada S' não inclui proposições falsas, podemos concluir que S é um sistema possível. Como todas as variáveis são *pivot*, o sistema é determinado e portanto a intersecção dos três planos definidos pelas equações de S ocorre num único ponto de $\mathbb{R}^3$. Note que, substituindo os membros direitos de S por quaisquer outros valores, o sistema resultante seria também possível e determinado.

Exercícios 3

1. Discuta, para todos os valores dos parâmetros, cada um dos seguintes sistemas.

a) $\begin{cases} x \quad - \quad z = 1 \\ y + \ az = 0 \\ -x + y + 2az = 1 \end{cases}$, $a \in \mathbb{R}$ b) $\begin{cases} x_1 + \ x_2 - \ x_3 = 1 \\ 2x_2 + 2x_3 = \gamma \\ x_1 + \gamma x_2 + \gamma x_3 = 1 \end{cases}$, $\gamma \in \mathbb{R}$.

c) $\begin{cases} ax \quad + 2z = 2 \\ x + 2y \quad = 1 \\ x - 2y + bz = 3 \end{cases}$, $a, b \in \mathbb{R}$ d) $\begin{cases} 2x \quad + 4y + bz = 2 \\ x + (d+2)y \quad = 1 \\ x \quad + 2y + bz = 1 \\ x \quad + 2y \quad = c \end{cases}$, $b, c, d \in \mathbb{R}$.

2. Seja S um sistema de equações lineares do tipo $m \times n$. Diga, justificando, se cada uma das seguintes afirmações é verdadeira ou falsa.

   a) Se $m < n$, então S é indeterminado.

   b) Se S é possível e $m < n$, então é indeterminado com exatamente $m-n$ variáveis livres.

   c) Se $m > n$, então S é impossível.

   d) Se S é possível e $m > n$, então S é determinado.

   e) Se S é possível e $m = n$, então S é determinado.

Um sistema com os membros direitos todos nulos chama-se *homogéneo*. Os sistemas homogéneos são possíveis pois admitem a *solução trivial*, i.e., com todas as variáveis iguais a zero. Note que um sistema homogéneo com menos equações do que variáveis é indeterminado.

1.2 Matrizes e vetores

Um sistema de equações lineares pode ser sucintamente representado (a menos dos nomes das variáveis) registando de modo organizado os números envolvidos no sistema. Os coeficientes e os membros direitos do sistema

$$\begin{aligned} x_1 + x_2 + 2x_3 - x_4 &= 8 \\ -x_1 - 2x_2 + 3x_3 + x_4 &= 1 \\ 3x_1 - 7x_2 + 4x_3 - 2x_4 &= 10 \end{aligned}$$

podem ser registados na *matriz dos coeficientes* e no *vetor do membros direitos*, respetivamente

$$A = \begin{bmatrix} 1 & 1 & 2 & -1 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & -2 \end{bmatrix} \text{ e } b = \begin{bmatrix} 8 \\ 1 \\ 10 \end{bmatrix}.$$

Cada *elemento* da matriz é referenciado pelo número de linha e número de coluna que ocupa. O elemento (i, j) é o que figura na linha i e coluna j. Uma matriz com m linhas e n colunas diz-se do *tipo* m por n e escreve-se $m \times n$. A matriz A é do tipo 3×4. O elemento $(1, 1)$ é 1, o elemento $(1, 2)$ é 1, ..., o elemento $(3, 4)$ é -2.

Duas matrizes são iguais se são do mesmo tipo e têm elementos homólogos iguais.

Uma matriz do tipo $n \times n$ diz-se *quadrada* de ordem de n. Numa matriz quadrada de ordem n os elementos $(1, 1)$, $(2, 2)$, ..., (n, n) são os da *diagonal principal*. Se são nulos os elementos por baixo (por cima) da diagonal principal, a matriz diz-se *triangular superior* (*inferior*). A matriz é *diagonal* se são nulos os elementos fora da diagonal principal. Uma matriz diagonal com os elementos da diagonal principal iguais a 1 chama-se *matriz identidade* e representa-se por I.

Um *vetor* é uma matriz com uma só coluna. A *componente* i de um vetor é o elemento da linha i. O vetor b tem 3 componentes.

O método de eliminação de Gauss para classificar ou resolver um sistema, em que A é a matriz de coeficientes e b o vetor membro direito, pode ser aplicado diretamente à *matriz* *ampliada* $[A|b]$.

Exemplo 10 A matriz ampliada do sistema 3×4

$$S = \begin{cases} x_1 + \ \ x_2 + 2x_3 - \ \ x_4 = 8 \\ -x_1 - 2x_2 + 3x_3 + \ \ x_4 = 1 \\ 3x_1 - 7x_2 + 4x_3 - 2x_4 = 10 \end{cases} \quad \text{é} \quad [A|b] = \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 8 \\ -1 & -2 & 3 & 1 & 1 \\ 3 & -7 & 4 & -2 & 10 \end{array}\right],$$

à qual aplicamos o método de Gauss, de que resultou o seguinte.

$$[A|b] \longrightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 8 \\ 0 & -1 & 5 & 0 & 9 \\ 0 & -10 & -2 & 1 & -14 \end{array}\right] \longrightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 8 \\ 0 & -1 & 5 & 0 & 9 \\ 0 & 0 & -52 & 1 & -104 \end{array}\right] = [A'|b'].$$

A obtenção desta matriz em escada indica o fim da fase descendente do método. Numa *matriz em escada* o primeiro elemento não nulo de uma linha - o *pivot* - está mais à direita do que o primeiro não nulo da linha anterior.

A não existência de linhas $[0\ 0\ 0\ 0|\ a]$, com $a \neq 0$, permite concluir que S é um sistema possível. A existência em A' de colunas sem *pivots* (coluna 4) indica que o sistema é indeterminado.

A fase ascendente do método prossegue com a matriz

$$[A'|b'] \longrightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 2 & -1 & 8 \\ 0 & -1 & 5 & 0 & 9 \\ 0 & 0 & 1 & -\frac{1}{52} & 2 \end{array}\right] \longrightarrow \left[\begin{array}{rrrr|r} 1 & 1 & 0 & -\frac{50}{52} & 4 \\ 0 & -1 & 0 & \frac{5}{52} & -1 \\ 0 & 0 & 1 & -\frac{1}{52} & 2 \end{array}\right] \longrightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -\frac{45}{52} & 3 \\ 0 & 1 & 0 & -\frac{5}{52} & 1 \\ 0 & 0 & 1 & -\frac{1}{52} & 2 \end{array}\right]$$

e termina com esta matriz reduzida. Uma *matriz reduzida* é uma matriz em escada em que nas coluna com *pivot* os elementos são todos nulos exceto o *pivot* que é igual a 1.

A matriz reduzida que resultou do método de Gauss representa o sistema reduzido

$$\begin{aligned} x_1 \qquad\qquad - \tfrac{45}{52}x_4 &= 3 \\ x_2 \qquad - \tfrac{5}{52}x_4 &= 1 \\ x_3 - \tfrac{1}{52}x_4 &= 2, \end{aligned}$$

que é equivalente ao sistema dado e cujas soluções são dadas por $x_1 = 3 + \frac{45}{52}x_4$, $x_2 = 1 + \frac{5}{52}x_4$, $x_3 = 2 + \frac{1}{52}x_4$, $x_4 = \forall$.

Exercícios 4

1. Discuta e interprete geometricamente o sistema de equações lineares correspondente a cada uma das seguintes matrizes ampliadas.

$$\text{a)} \left[\begin{array}{rr|r} 1 & 2 & 2 \\ 3 & -1 & 7 \\ 4 & 1 & 9 \\ 2 & -3 & 3 \end{array}\right] \qquad \text{b)} \left[\begin{array}{rrr|r} 1 & 1 & 3 & 5 \\ 2 & -1 & 4 & 11 \\ 0 & -1 & 1 & 3 \end{array}\right] \qquad \text{c)} \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 3 & 2 & 1 & 0 \\ 5 & 3 & 3 & 0 \end{array}\right].$$

2. Considere os sistemas de equações lineares cujas correspondentes matrizes ampliadas são

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & b_1 \\ 0 & 1 & a & b_2 \\ -1 & 1 & 2a & b_3 \end{array}\right], \text{ com } \; a, b_1, b_2, b_3 \in \mathbb{R}.$$

   a) Para que valores de a os sistemas são possíveis, independentemente dos valores dos parâmetros b_1, b_2, b_3?

   b) Para que valores de b_1, b_2, b_3 os sistemas são possíveis, independentemente do valor do parâmetro a?

   c) Atribua a a, b_1, b_2, b_3 valores que façam o sistema

      c1) impossível,

      c2) indeterminado.

3. É correto afirmar que um sistema de equações lineares do tipo $n \times n$ é possível e determinado se e só se a matriz reduzida que se obtem quando se aplica o método de Gauss à matriz dos coeficientes é a matriz identidade? Justifique.

4. Seja E uma matriz em escada do tipo $m \times n$.

a) Quantos *pivots* podem existir em E?

b) Qual é a relação entre o número de *pivots* e o número de linhas nulas de E?

As matrizes não aparecem apenas no contexto dos sistemas de equações lineares. São objetos matemáticos com uma aritmética própria. Em seguida apresentam-se algumas operações com matrizes.

1.3 Operações com matrizes

Considere

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & \vdots & & \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \text{ e } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

matrizes genéricas do tipo $m \times n$ e um vetor com n componentes e seja λ um escalar.

Definição 1 A *transposta* de A é a matriz

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

i.e., a matriz do tipo $n \times m$, cuja coluna j é a linha j de A.

Exemplo 11

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

O seguinte resultado é óbvio.

Proposição 1.1 $(A^\top)^\top = A$.

Uma matriz diz-se *simétrica* se é igual à transposta.

EXEMPLO 12

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

é uma matriz simétrica.

Definição 2 O *produto escalar* de λ por A é

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

EXEMPLO 13

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

O produto escalar verifica as seguintes propriedades.

Proposição 1.2 *Se A é uma matriz e λ e μ são escalares, tem-se*

1. $(\lambda\mu)A = \lambda(\mu A)$.

2. $(\lambda A)^\top = \lambda A^\top$.

Apresenta-se agora a definição de soma de matrizes do mesmo tipo.

Definição 3 A *soma* de A e B é

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

Exemplo 14

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

É fácil provar os seguintes resultados.

Proposição 1.3 *Sejam A, B e C matrizes do tipo $m \times n$, e λ e μ escalares.*

1. $A+B=B+A$.
2. $A+(B+C)=(A+B)+C$.
3. *Se* 0 *é a matriz nula do tipo* $m \times n$ *e* $-A=-1A$, *tem-se* $A+0=A$ *e* $A+(-A)=A-A=0$.
4. $(A+B)^\top = A^\top + B^\top$.
5. *Se* Q *é uma matriz quadrada, a matriz* $Q+Q^\top$ *é simétrica.*
6. $\lambda(A+B)=\lambda A+\lambda B$.
7. $(\lambda+\mu)A=\lambda A+\mu A$.

Exercícios 5

1. Para

$$A = \begin{bmatrix} 3 & -1 & 0 \\ 2 & 7 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 5 & 4 & 1 \\ 2 & -3 & -4 \end{bmatrix}, \ C = \begin{bmatrix} 0 & -2 \\ 1 & -1 \\ -3 & -4 \end{bmatrix} \text{ e } D = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

   calcule, sempre que possível, o valor de cada uma das seguintes expressões.

   a) $(5A - A) - (B - 2B)$ b) $(2A - B)^\top - C$ c) $(2(A^\top - C)^\top + B)^\top$

   d) $(B^\top - C)^\top + 2B^\top$ e) $D + D^\top$ f) $D - D^\top$.

2. Identifique, se existirem, escalares α e β tais que

$$\alpha \begin{bmatrix} 1 & 0 \\ -2 & 4 \end{bmatrix} + \beta \begin{bmatrix} 6 & 2 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 8 & -12 \end{bmatrix}.$$

A multiplicação de matrizes é uma operação um pouco mais complexa do que as apresentadas anteriormente. Começamos por definir a multiplicação de matrizes por vetores.

Definição 4 O *produto* de A por x é

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

Exemplo 15

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

Nas observações seguintes estabelece-se a notação matricial de sistemas de equações lineares e apresenta-se uma nova interpretação geométrica da resolução de sistemas.

Observações 2

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

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

De um modo geral tem-se

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

i.e., o vetor Ax é soma de múltiplos das colunas da matriz A, em que coluna j de A é multiplicada pela componente j de x.

2. O sistema

$$\begin{aligned} x_1 + \ x_2 + 2x_3 - \ x_4 &= \ 8 \\ -x_1 - 2x_2 + 3x_3 + \ x_4 &= \ 1 \\ 3x_1 - 7x_2 + 4x_3 - 2x_4 &= 10 \end{aligned}$$

escreve-se matricialmente na forma

$$\begin{bmatrix} 1 & 1 & 2 & -1 \\ -1 & -2 & 3 & 1 \\ 3 & -7 & 4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \\ 10 \end{bmatrix}.$$

De um modo geral o sistema de equações lineares

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

é representado pela equação matricial $Ax = b$, em que A é a matriz dos coeficientes e b o vetor membro direito.

3. Como consequência das duas observações anteriores tem-se a seguinte interpretação geométrica para a classificação e resolução de sistemas de equações lineares. O sistema $Ax = b$ é possível sse "percorrendo" as direções das colunas de A é possível "atingir" o vetor b. Cada solução é a quantificação do "percurso" em cada uma das direções.

   O sistema

$$S = \begin{cases} x + \ \ y = 6 \\ x - 2y = 0 \end{cases} \Leftrightarrow \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \Leftrightarrow x \begin{bmatrix} 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

   é possível pois "percorrendo" a direção do vetor $(1, 1)$ e a direção do vetor $(1, -2)$ "atinge-se" o membro direito $(6, 0)$ (ver Figura 1.4). De facto, tem-se

$$4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

É fácil deduzir as seguintes propriedades.

Proposição 1.4 *Sejam A e B matrizes do tipo $m \times n$, I a matriz identidade de ordem n, x e y vetores com n componentes e λ um escalar. Tem-se*

1. $Ix = x$,

2. $A(x + y) = Ax + Ay$,

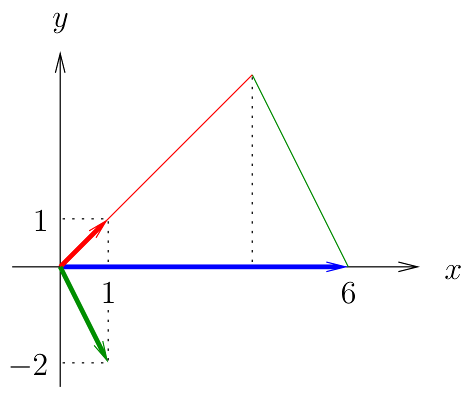


Figura 1.4: Interpretação geometrica da resolução do sistema S da Observação 2.3.

3. $(A+B)x = Ax + Bx$,

4. $A(\lambda x) = \lambda(Ax) = (\lambda A)x$.

EXERCÍCIOS 6

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

   a) Calcule $Ab + Ib$, $(A+I)b$, $(A+A^\top)2b$ e $b^\top b$.

   b) Resolva a equação matricial $Ax = 3x + b$, com $x \in \mathbb{R}^3$.

2. Considere a matriz $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ e o vetor genérico de $\mathbb{R}^2$ $v = \begin{bmatrix} x \\ y \end{bmatrix}$.

   a) Calcule, em função de x e y, o vetor Av e represente geometricamente v e Av.

   b) Qual é a relação entre entre os vetores v e Av?

3. Considere $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 2 & 5 \\ 1 & 0 & 0 \end{bmatrix}$ e o sistema homogéneo $(A - \lambda I)x = \vec{0}$, com $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ e $\lambda \in \mathbb{R}$.

a) Para que valores de λ o sistema é indeterminado?

b) Mostre que se $v \in \mathbb{R}^3$ é solução do sistema, então $Av = \lambda v$.

c) Resolva o sistema considerando $\lambda = -1$. Interprete geometricamente o conjunto das soluções e a relação estabelecida na alínea b).

A multiplicação de matrizes realiza-se efetuando sucessivas multiplicações de matrizes por vetores. Sejam

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \text{ e } B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ & \vdots & & \\ b_{n1} & b_{n2} & \cdots & b_{nr} \end{bmatrix}$$

matrizes do tipo $m \times n$ e $n \times r$, respetivamente. As matrizes A e B dizem-se *encadeadas* pois o número de colunas de A é igual ao número de linhas de B.

Definição 5 O *produto* de A por B é uma matriz do tipo $m \times r$, cuja coluna j, $j = 1, \ldots, r$, é o produto da matriz A pela coluna j de B.

Exemplo 16

$$\begin{bmatrix} 1 & 1 & 2 & -1 \\ -1 & -2 & 3 & 1 \\ 3 & 0 & 4 & -2 \end{bmatrix} \begin{bmatrix} 2 & 5 & 1 \\ -1 & -2 & 0 \\ 0 & 1 & -4 \\ 3 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 5 & -6 \\ 3 & 2 & -14 \\ 0 & 19 & -11 \end{bmatrix}.$$

Observação 3 O elemento (i, j) da matriz $P = AB$ é o "produto da linha i de A pela coluna j de B", i.e., $p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

É fácil verificar que a multiplicação de matrizes satisfaz as seguintes propriedades.

Proposição 1.5 *Sejam A, B, C matrizes, I matriz identidade e λ um escalar. Sempre que as operações se possam realizar, tem-se*

1. $AI = A$.

2. $(AB)C = A(BC)$.

3. $A(B+C) = AB + AC$.

4. $(A+B)C = AC + BC$.

5. $A(\lambda B) = (\lambda A)B = \lambda(AB)$.

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

Observações 4

1. Em geral $AB \neq BA$, i.e., a multiplicação de matrizes não é comutativa. Verifique que $AB \neq BA$, para $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ e $B = \begin{bmatrix} -1 & 0 \\ 2 & 4 \end{bmatrix}$.

   Quando $AB = BA$ as matrizes dizem-se *permutáveis*. Qualquer matriz quadrada é permutável com a matriz identidade da mesma ordem.

2. Em $\mathbb{R}$ tem-se $ab = 0 \Rightarrow a = 0$ ou $b = 0$ (lei do anulamento do produto). Para matrizes $AB = 0 \not\Rightarrow A = 0$ ou $B = 0$. Verifique que $AB = 0$, com $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

3. Em $\mathbb{R}$ tem-se $ab = ac$ e $a \neq 0 \Rightarrow b = c$ (lei do corte). Para matrizes $AB = AC$ e $A \neq 0 \not\Rightarrow B = C$. Verifique que $AB = AC$, para $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \neq C = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix}$.

Exercício 7 Considere

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

Calcule, se possível, AB, BA, $BA^\top$, CC, $AA^\top$, $a^\top a$ e $a\,a^\top$.

Uma vez definida a multiplicação de matrizes é natural atribuir significado à potência de expoente inteiro não negativo de uma matriz quadrada.

Definição 6 Sejam A uma matriz quadrada de ordem n e I a matriz identidade de ordem n. A *potência* de expoente inteiro $k \geq 0$ da matriz A é

$$A^k = \begin{cases} \underbrace{A \times A \times \cdots \times A}_{k \text{ vezes}} & \text{se } k \geq 1 \\ I & \text{se } k = 0 \end{cases}.$$

Exercício 8 Calcule B^3 com $B = \begin{bmatrix} 2 & 0 & 1 \\ -4 & 0 & 3 \\ 1 & -2 & 0 \end{bmatrix}$.

A noção de inversa de uma matriz é análoga à do inverso de um número real. O inverso do real a é $b \in \mathbb{R}$ tal que $ab = 1$. O inverso de a representa-se por a^{-1} e, para $a \neq 0$, tem-se $a^{-1} = \frac{1}{a}$.

Para matrizes tem-se a seguinte definição.

Definição 7 A *inversa* de uma matriz quadrada A de ordem n, é uma matriz quadrada B de ordem n, tal que $AB = BA = I$. A matriz inversa de A representa-se por A^{-1}.

Exercício 9 Verifique que $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & -1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 & 3 \\ -1 & 1 & 2 \\ 3 & -2 & -6 \end{bmatrix}$.

Se uma matriz tem inversa diz-se *invertível* ou *não singular*. Caso contrário diz-se *singular*.

Observação 5 Da definição de inversa decorre diretamente que se A é invertível A^{-1} também é invertível e $(A^{-1})^{-1} = A$ (i.e., A é a inversa da inversa de A).

De acordo com a Definição 7, para verificar se uma dada matriz B é a inversa de A, há que efetuar os produtos AB e BA e ver se ambos são iguais à matriz identidade. O seguinte resultado permite concluir que essa verificação não requer mais do que um daqueles produtos.

Teorema 1.6 *Se $AB = I$, então $BA = I$, i.e., uma matriz e a sua inversa são permutáveis.*

Facilmente se provam os seguintes resultados.

Proposição 1.7

1. *Uma matriz não singular tem uma única inversa.*
2. *Se A e B são matrizes não singulares da mesma ordem, então AB é não singular e $(AB)^{-1} = B^{-1}A^{-1}$ (a inversa do produto é o produto das inversas por ordem inversa).*
3. $(A^k)^{-1} = (A^{-1})^k$, *para* $k \in \mathbb{Z}_0^+$.
4. $(A^{-1})^\top = (A^\top)^{-1}$.

Exercício 10 Prove os resultados da Proposição 1.7.

Vejamos agora como determinar a inversa de uma matriz ou decidir que a matriz não é invertível. Para isso consideremos a matriz

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

Pretende-se determinar uma matriz $B = \begin{bmatrix} x & y & z \\ | & | & | \end{bmatrix}$ tal que

$$A \begin{bmatrix} x & y & z \\ | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

ou seja resolver os três sistemas de equações:

$$Ax = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$Ay = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$Az = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Para resolver estes três sistemas, que têm a mesma matriz de coeficientes, aplica-se o método de eliminação de Gauss à matriz ampliada

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 3 & 0 & 1 & 0 \\ 3 & 4 & 5 & 0 & 0 & 1 \end{array}\right] = [A|I] \to \cdots \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & 0 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right].$$

Tem-se pois

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

e assim

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

de onde se conclui que

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

Vejamos o resultado do procedimento anterior com a matriz

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

$$[A|I] = \left[\begin{array}{ccc|ccc} -1 & 3 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} -1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

A matriz em escada obtida permite concluir que os sistemas

$$Ax = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ e } Ax = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

são impossíveis e portanto não existe uma matriz B tal que $AB = I$, i.e., A é singular.

Na Figura 1.5 sistematiza-se este método para identificar a inversa de uma matriz quadrada ou decidir que a matriz é singular.

· Sejam A uma matriz quadrada, $[A|I]$ a matriz A ampliada com a identidade e $[A'|I']$ a matriz em escada resultante de aplicar a fase descendente do método de Gauss a $[A|I]$. (Note que I' não tem linhas nulas.)

· Se A' tem alguma linha nula $\Rightarrow A$ é **singular**.
Caso contrário A é **invertível**.
Para determinar A^{-1} aplique a fase ascendente do método de Gauss à matriz $[A'|I']$. A matriz reduzida resultante é do tipo $[I|A'']$ e $A^{-1} = A''$.

Figura 1.5: Esquema para determinar a inversa da matriz A ou decidir que A é singular.

Exercícios 11

1. Determine, caso exista, a inversa de cada uma das seguintes matrizes.

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

2. Mostre que a matriz $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 1 & 3 & 5 \end{bmatrix}$ é não singular e utilize A^{-1} para resolver o sistema $Ax = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$.

3. Sejam A, B e C matrizes invertíveis da mesma ordem.

a) É correto afirmar que $A+B$ é invertível?

b) Será que a matriz A^3BC^{-1} é invertível?

c) Mostre que $A^{-1}(A+B)B^{-1} = A^{-1} + B^{-1}$.

d) Prove que se $AB =$, então $B = C$.

4. Sejam A uma matriz quadrada de ordem 3 invertível e b e c vetores de $\mathbb{R}^3$.

a) Classifique os sistemas $Ax = b$ e $A^{-1}x = c$.

b) Prove que os sistemas $Ax = b$ e $A^{-1}x = c$ são equivalentes sse $b = A^2c$.

c) Sejam u, v e w as soluções dos sistemas

$$Ax = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, Ax = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ e } Ax = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix},$$

respetivamente. Determine, em termos dos vetores u, v e w, a matriz inversa de A.

5. Seja A uma matriz quadrada de ordem n. Mostre que as proposições seguintes são equivalentes.

a) A é invertível.

b) $Ax = 0 \Leftrightarrow x = 0$.

c) O sistema $Ax = b$ é possível para todo o vetor b de $\mathbb{R}^n$.

6. Calcule $A^2 + 3bb^\top$, com $A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 5 \\ -1 & 0 & 0 \end{bmatrix}$ e $b = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$.

7. Considere $A = \begin{bmatrix} \alpha & -2 & 4 \\ 1 & 0 & 1 \\ -1 & -1 & -\alpha \end{bmatrix}$ e $b = \begin{bmatrix} 4 \\ 1 \\ 3+\beta \end{bmatrix}$, com $\alpha, \beta \in \mathbb{R}$.

a) Discuta o sistema $Ax = b$ para todos os valores de α e a β.

b) Resolva o sistema $Ax = b$, considerando $\alpha = 0$ e $\beta = -3$.

c) Indique, justificando, um valor de α para o qual a matriz A é invertível.

8. Seja $Ax = b$ um sistema que admite as soluções não nulas u e v. Em que condições o vetor $u + v$ ainda é solução de $Ax = b$? Justifique.

Capítulo 2

Espaços vetoriais

Vamos agora estudar certas estruturas definidas sobre vetores de $\mathbb{R}^m$.

2.1 Subespaços vetoriais

Definição 8 Um conjunto V de vetores de $\mathbb{R}^m$ diz-se:

. *fechado para a adição* se para todo o par de vetores $x, y \in V$, o vetor $x + y \in V$;

. *fechado para a multiplicação escalar* se para todo o vetor $x \in V$ e para todo o escalar $\lambda \in \mathbb{R}$, o vetor $\lambda x \in V$.

Exercícios 12

1. Quais dos seguintes conjuntos são fechados para a adição e multiplicação escalar?

   a) $\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$.

   b) $\{(x_1, x_2) \in \mathbb{R}^2 : 2x_1 - x_2 = 0\}$.

   c) $\{(x_1, x_2) \in \mathbb{R}^2 : 2x_1 - x_2 = 1\}$.

   d) $\{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 \in \mathbb{Z}\}$.

e) $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 x_2 \geq 0\}$.

2. Quais são os subconjuntos de $\mathbb{R}^2$ e de $\mathbb{R}^3$ fechados para a adição e multiplicação escalar?

Definição 9 Um conjunto de vetores $V \subseteq \mathbb{R}^m$ diz-se *subespaço vetorial* se $V \neq \emptyset$ e é fechado para a adição e multiplicação escalar.

Observações 6

1. $\{\vec{0}\}$ é um subespaço vetorial minimal (subespaço trivial).

2. $\mathbb{R}^m$ é um subespaço vetorial maximal.

3. Se V é um subespaço vetorial, tem-se

   a) $\vec{0} \in V$, i.e., todo o subspaço vetorial inclui o vetor nulo.

   b) Se $x \in V \Rightarrow -x \in V$

Considere uma matriz A do tipo $m \times n$ e represente por $\mathcal{N}(A)$ o conjunto das soluções do sistema homogéneo $Ax = \vec{0}$, i.e., $\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = \vec{0}\}$. É claro que o vetor nulo de $\mathbb{R}^n$ pertence a $\mathcal{N}(A)$.
Pode provar-se o seguinte.

Teorema 2.1 *Se A é uma matriz do tipo $m \times n$, $\mathcal{N}(A)$ é um subespaço vetorial de $\mathbb{R}^n$. Chama-se* espaço nulo *da matriz A.*

Demonstração:

Já vimos que o vetor nulo (de $\mathbb{R}^n$) pertence a $\mathcal{N}(A)$, e portanto $\mathcal{N}(A) \neq \emptyset$.

Se x e y são vetores de $\mathcal{N}(A)$, tem-se $Ax = \vec{0}$ e $Ay = \vec{0}$. Somando membro a membro as duas equações, vem $Ax + Ay = \vec{0} + \vec{0} \Leftrightarrow A(x + y) = \vec{0}$, i.e., $x + y$ pertence a $\mathcal{N}(A)$, o que mostra que $\mathcal{N}(A)$ é fechado para a adição.

Se $x \in \mathcal{N}(A)$ e $\lambda \in \mathbb{R}$, tem-se $\lambda(Ax) = \lambda\vec{0} \Leftrightarrow A(\lambda x) = \vec{0}$, i.e., λx pertence a $\mathcal{N}(A)$, o que mostra que $\mathcal{N}(A)$ é fechado para a multiplicação escalar. $\square$

Exemplo 17 O espaço nulo da matriz $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix}$ é

$$\mathcal{N}(A) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \}.$$

Aplicando o método de Gauss para resolver o sistema, obtem-se

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \longrightarrow \begin{cases} x_1 = x_3 \\ x_2 = -x_3 \\ x_3 = \forall \end{cases} \text{ e portanto}$$

$$\mathcal{N}(A) = \{ \begin{bmatrix} a \\ -a \\ a \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \forall a \in \mathbb{R}\},$$

que é a reta com a direção do vetor $(1, -1, 1)$ que passa na origem.

Exercícios 13 Identifique geometricamente $\mathcal{N}(A)$, com $A = [a \ \ b]$, $A = [a \ \ b \ \ c]$, $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$ e $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

Observação 7 $\{x \in \mathbb{R}^n : Ax = b\}$ é subespaço vetorial sse $b = \vec{0}$.

Em seguida introduz-se a noção de combinação linear de vetores.

Definição 10 Um vetor $w \in \mathbb{R}^m$ é *combinação linear* dos vetores $v_1, v_2, \ldots, v_n$ de $\mathbb{R}^m$ se existem escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tais que $w = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$, i.e., o sistema $\left[\begin{array}{cccc|c} v_1 & v_2 & \ldots & v_n & w \end{array}\right]$ é possível.

Observações 8

1. As combinações lineares do vetor v são os vetores λv, com $\lambda \in \mathbb{R}$, i.e., os vetores múltiplos de v, ou seja a reta com a direção de v que passa na origem, no caso de $v \neq \vec{0}$.

2. Todo o vetor de $\mathbb{R}^3$ é combinação linear dos vetores $(1,0,0)$, $(0,1,0)$, $(0,0,1)$.

Exemplos 18 Sejam $u = (1,2,-1)$ e $v = (6,4,2)$. Mostre que:

a) $w = (9,2,7)$ é combinação linear de u e v.

O problema converte-se na resolução do sistema de equações lineares

$$\left[\begin{array}{cc|c} u & v & w \end{array}\right] = \left[\begin{array}{cc|c} 1 & 6 & 9 \\ 2 & 4 & 2 \\ -1 & 2 & 7 \end{array}\right] \longrightarrow \ldots \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right] \longrightarrow \begin{cases} \lambda_1 = -3 \\ \lambda_2 = 2 \end{cases}$$

De facto, tem-se $-3\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 2 \\ 7 \end{bmatrix}$, ou seja $-3u + 2v = w$.

b) $w' = (4,-1,8)$ não é combinação linear de u e v.

Aplicando o método de Gauss ao sistema

$$\left[\begin{array}{cc|c} u & v & w' \end{array}\right] = \left[\begin{array}{cc|c} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 8 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 6 & 4 \\ 0 & -8 & -9 \\ 0 & 8 & 12 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 6 & 4 \\ 0 & -8 & -9 \\ 0 & 0 & 3 \end{array}\right]$$

conclui-se que é o sistema impossível, i.e., não existem escalares λ_1, λ_2 tais que $\lambda_1 u + \lambda_2 v = w'$ e assim w' não é combinação linear de u e v.

Teorema 2.2 *Seja A uma matriz do tipo $m \times n$. O conjunto de todas as combinações lineares das n colunas de A é um subespaço vetorial de $\mathbb{R}^m$, que se chama* espaço das colunas *de A e se representa por $\mathcal{C}(A)$.*

Note que um vetor $w \in \mathbb{R}^m$ pertence a $\mathcal{C}(A)$ sse o sistema $Ax = w$ é possível.

Demonstração (do Teorema 2.2):

Como sistema homogéneo $Ax = \vec{0}\,(\in \mathbb{R}^m)$ é possível, tem-se $\vec{0} \in \mathcal{C}(A)$, garantindo que $\mathcal{C}(A) \neq \emptyset$.

Se w e w' são vetores de $\mathcal{C}(A)$ é porque os sistemas $Ax = w$ e $Ax = w'$ são possíveis. Sejam $u \in \mathbb{R}^n$ solução de $Ax = w$ e $u' \in \mathbb{R}^n$ solução $Ax = w'$, i.e., $Au = w$ e $Au' = w'$. Somando membro a membro as duas últimas igualdades, obtem-se $Au + Au' = w + w' \Leftrightarrow A(u + u') = w + w'$, que permite concluir que o sistema $Ax = w + w'$ é possível ($u + u'$ é uma solução) e assim que $w + w'$ pertence a $\mathcal{C}(A)$.

Se $w \in \mathcal{C}(A)$, tem-se $Au = w$ para algum vetor $u \in \mathbb{R}^n$. Multiplicando ambos os membros da equação por $\lambda \in \mathbb{R}$, vem $\lambda(Au) = \lambda w \Rightarrow A(\lambda u) = \lambda w$, i.e., o sistema $Ax = \lambda w$ é possível (λu é uma solução) e portanto $\lambda w \in \mathcal{C}(A)$. $\square$

Exemplo 19 O espaço das colunas da matriz $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -4 & -8 & 4 \end{bmatrix}$ é

$\mathcal{C}(A) = \{w = (w_1, w_2, w_3) \in \mathbb{R}^3 : \text{o sistema} Ax = w \text{ é possível}\}$.

Ao aplicar o método de Gauss para classificar o sistema $Ax = w$ vem

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & w_1 \\ 2 & 4 & -2 & w_2 \\ -4 & -8 & 4 & w_3 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & w_1 \\ 0 & 0 & 0 & w_2 - 2w_1 \\ 0 & 0 & 0 & w_3 + 4w_1 \end{array}\right]$$, que permite concluir que $Ax = w$

é possível sse $\begin{cases} w_2 - 2w_1 = 0 \\ w_3 + 4w_1 = 0 \end{cases} \Leftrightarrow \begin{cases} w_1 = \forall \\ w_2 = 2w_1 \\ w_3 = -4w_1 \end{cases}$

Para interpretar geometricamente o subspaço $\mathcal{C}(A)$, podemos escrever $(w_1, 2w_1, -4w_1) = w_1(1, 2, -4)$ e assim reconhecer que $\mathcal{C}(A)$ é a reta de $\mathbb{R}^3$ que passa na origem e tem a direção do vetor $(1, 2, -4)$.

Para determinar o subspaço das colunas de uma matriz arbitrária A do tipo $m \times n$ procede-se como no exemplo anterior. Mais precisamente,

. Define-se a matriz ampliada $[A|w]$, com $w = (w_1, w_2, \ldots, w_m)$ vetor genérico de $\mathbb{R}^m$.

. Aplica-se o método de Gauss (fase descendente) a $[A|w]$. Seja $[A'|w']$ a matriz em escada resultante.

. Se A' não tem linhas nulas (i.e., o sistema $Ax = w$ é possível, $\forall w \in \mathbb{R}^m$) $\Rightarrow \mathcal{C}(A) = \mathbb{R}^m$.
Caso contrário (cada linha nula de A' introduz uma restrição aos membros direitos para os quais o sistema $Ax = w$ é possível.)
Se i é linha nula de A', tem-se a restrição $w'_i = 0$.

Observação 9 Quando A' tem linhas nulas, o algoritmo identifica $\mathcal{C}(A)$ com espaço do nulo de uma matriz com m colunas e tantas linhas quantas as linhas nulas de A'.

Definição 11 Chama-se *espaço gerado* por um conj. de vetores $V = \{v_1, v_2, \ldots, v_n\}$, e representa-se por $< V >$, o conj. de todas as combinações lineares desses vetores, i.e., o espaço das colunas da matriz $[v_1\ v_2\ \ldots\ v_n]$.

Exemplo 20 Determinar o espaço gerado por $V = \{(1, 3, 4), (2, 2, 4), (3, 5, 8), (1, 1, 2)\}$, é determinar o espaço das colunas da matriz $A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 3 & 2 & 5 & 1 \\ 4 & 4 & 8 & 2 \end{bmatrix}$. Aplica-se pois o método de Gauss à matriz $[A|w]$.

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 1 & w_1 \\ 3 & 2 & 5 & 1 & w_2 \\ 4 & 4 & 8 & 2 & w_3 \end{array}\right] \longrightarrow \ldots \longrightarrow \left[\begin{array}{cccc|c} 1 & 2 & 3 & 1 & w_1 \\ 0 & -4 & -4 & -2 & -3w_1 + w_2 \\ 0 & 0 & 0 & 0 & -w_1 - w_2 + w_3 \end{array}\right]$$

O vetor w pertence ao espaço das colunas da matriz A, ou seja pertence ao espaço gerado por V sse $-w_1 - w_2 + w_3 = 0$.
Assim, $< V >= \mathcal{N}([-1 \ -1 \ \ 1])$ que é o plano ortogonal ao vetor $(-1,-1,1)$ que passa na origem.
Note que $< V >=< \{(1,3,4),(2,2,4),(3,5,8),(1,1,2)\} >=< \{(1,3,4),(2,2,4)\} >$.

De uma forma geral tem-se o seguinte.

Observação 10 Se A' é uma matriz em escada resultante de aplicar o método de Gauss à matriz A, $\mathcal{C}(A)$ é o espaço gerado pelas colunas de A que correspondem às colunas *pivot* de A'.

Exercícios 14

1. Determine os espaços nulo e das colunas das seguintes matrizes.

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

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

   g) $\begin{bmatrix} 1 & -2 & 2 \\ 4 & -6 & 2 \end{bmatrix}$ h) $\begin{bmatrix} 1 & -2 & 2 \\ 3 & -6 & 6 \end{bmatrix}$ i) $\begin{bmatrix} 3 & -1 \\ 9 & -3 \\ 6 & -2 \end{bmatrix}$.

2. Verifique se o vetor $(-3,12,12)$ é combinação linear dos vetores $v_1 = (-1,3,1)$, $v_2 = (0,2,4)$, $v_3 = (1,0,2)$.

3. Verifique se o vetor $(3,1)$ está no espaço das colunas da matriz $\begin{bmatrix} 2 & -1 & 5 \\ 4 & 3 & 9 \end{bmatrix}$.

4. Verifique se o vetor $(0, 1, 4)$ está no espaço das colunas da matriz $\begin{bmatrix} 1 & 2 \\ 3 & 9 \\ -5 & 4 \end{bmatrix}$.

5. Em cada uma das alíneas seguintes, verifique se o vetor u é combinação linear dos vetores de V.

   a) $u = (3, -5), \qquad V = \{(1, 2), (-2, 6)\}$;

   b) $u = (1, 1, 1), \qquad V = \{(1, 0, 1), (0, 3, 5)\}$;

   c) $u = (2, -2, \frac{1}{6}, \frac{1}{6}), \quad V = \{(1, -1, 0, 0), (2, 0, 1, 1), (0, 3, 1, 1)\}$;

   d) $u = (0, 1, 0, 1, 0), \qquad V = \{(1, 2, 2, 1, 1), (\frac{2}{3}, 1, \frac{4}{3}, 1, \frac{2}{3})\}$.

2.2 Independência linear

Vamos agora introduzir o conceito fundamental de independência linear de vetores.

Definição 12 Um conjunto de vetores $V = \{v_1, v_2, \ldots, v_n\}$ de $\mathbb{R}^m$ é *linearmente independente* se todas as colunas da matriz em escada resultante de aplicar o método de Gauss à matriz $[v_1\ v_2\ \ldots\ v_n]$ são *pivot*. Se V não é linearmente independente diz-se *linearmente dependente*.

Observações 11

1. $\{v\}$ é linearmente independente sse $v \neq \vec{0}$.

2. Um conjunto que inclua o vetor nulo é linearmente dependente.

3. Se o conjunto de vetores $V = \{v_1, v_2, \ldots, v_n\}$ de $\mathbb{R}^m$ é linearmente independente, então $n \leq m$, i.e., um conj. linearmente independente de vetores de $\mathbb{R}^m$ não inclui mais do que m vetores.

EXEMPLO 21 Decida sobre a independência linear $V = \{\underbrace{(1,2,0,1)}_{v_1}, \underbrace{(0,-1,3,1)}_{v_2}, \underbrace{(4,2,1,0)}_{v_3}\}$.
Aplicando o método de Gauss à matriz cujas colunas são os vetores do conjunto, tem-se

$$\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 4 \\ 2 & -1 & 2 \\ 0 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix} \longrightarrow \ldots \longrightarrow \begin{bmatrix} 1 & 0 & 4 \\ 0 & -1 & -6 \\ 0 & 0 & -17 \\ 0 & 0 & 0 \end{bmatrix}.$$

Uma vez que toda a coluna da matriz em escada é *pivot*, conclui-se que V é linearmente independente.

EXERCÍCIO 15 Decida sobre a independência linear de $U = \{(1,2,-1),(0,2,1),(2,-1,3)\}$ e $U' = \{(1,2,-1),(0,2,1),(2,-1,3),(4,5,-2)\}$.

O resultado seguinte caracteriza independência linear de vetores em termos de sistemas de equações lineares.

Teorema 2.3 *O conjunto de vetores $V = \{v_1, v_2, \ldots, v_n\}$ de $\mathbb{R}^m$ é linearmente independente sse $\mathcal{N}[v_1\ v_2\ \ldots\ v_n] = \{\vec{0}\}$, i.e., $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = \vec{0} \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. (Só se obtem uma combinação linear nula dos vetores $v_1, v_2, \ldots, v_n$ anulando os coeficientes.)*

EXEMPLO 22 Mostre que $V = \{\underbrace{(1,0,1,1)}_{v_1}, \underbrace{(0,1,2,1)}_{v_2}, \underbrace{(2,-1,0,1)}_{v_3} \underbrace{(0,0,3,3)}_{v_4}\}$ é linearmente dependente.

$$A = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 2 & 0 & 3 \\ 1 & 1 & 1 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A'.$$

A coluna 3 de A' não é *pivot*, logo V é linearmente dependente. De facto, o sistema homogéneo $Ax = \vec{0}$ é indeterminado. $\mathcal{N}(A) = \{(-2a, a, a, 0), \forall a \in \mathbb{R}\} \neq \{\vec{0}\}$ e portanto

$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4 = \vec{0} \not\Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. Por exemplo $-2v_1 + v_2 + v_3 + 0v_4 = \vec{0}$.

Note que, como a coluna 3 de A' não é *pivot*, a coluna 3 de A é combinação linear das colunas 1 e 2 de A, i.e., o sistema $[v_1\ v_2|v_3]$ é possível ($(2, -1)$ é solução).

De uma forma geral, se a coluna j da matriz em escada que resulta de aplicar o método de Gauss à matriz A não é *pivot*, então a coluna j de A é combinação linear das restantes colunas de A. Tem-se pois o seguinte resultado

Teorema 2.4 *Um conjunto com dois ou mais vetores é linearmente dependente sse um dos vetores do conjunto é combinação linear dos restantes.*

Exercícios 16

1. Quais dos seguintes conjuntos de vetores são linearmente independentes?

   a) $\{(3, 1), (4, -2)\}$

   b) $\{(3, 1), (4, -2), (7, 2)\}$

   c) $\{(-1, 2, 0, 2), (5, 0, 1, 1), (8, -6, 1, -5)\}$

   d) $\{(0, -3, 1), (2, 4, 1), (-2, 8, 5)\}$.

2. Mostre que o conjunto de vetores $\{(1, 0, 3, 1), (-1, 1, 0, 1), (2, 3, 0, 0), (1, 1, 6, 3)\}$ é linearmente dependente.

   Pode cada um dos vetores ser expresso como uma combinação linear dos restantes?

3. Discuta em função dos parâmetros $\alpha, \beta, \gamma \in \mathbb{R}$ a independência linear dos seguintes conjuntos de vetores.

   a) $\{(1, -2), (\alpha, -1)\}$

   b) $\{(\alpha, 1, 1), (1, \alpha, 1), (1, 1, \alpha)\}$

   c) $\{0, \gamma, -\beta), (-\gamma, 0, \alpha), (\beta, -\alpha, 0)\}$.

4. Sabendo que $\{v_1, v_2, v_3\}$ é um conjunto de vetores de $\mathbb{R}^3$ linearmente independente, decida sobre a independência linear do conjunto $\{v_1 + v_2, v_1 + v_3, v_2 + v_3\}$.

2.3 Base

Definição 13 Sejam $S \neq \{\vec{0}\}$ um subespaço vetorial e $V = \{v_1, v_2, \ldots, v_n\}$ um conjunto de vetores de S. Diz-se que V é uma *base* de S se:

1. V é linearmente independente, e

2. V gera S, i.e., $< V >= S$.

Convenciona-se que $\emptyset$ é base do subespaço $\{\vec{0}\}$.

Observações 12

1. Todo o vetor de um subespaço vetorial exprime-se de forma única como combinação linear dos vetores da base.

2. $\{(1, 0, 0), (0, 1, 0)\}$ é uma base do plano X0Y. Outra base é $\{(1, 0, 0), (1, 1, 0)\}$. O conjunto $\{(1, 0, 0), (0, 1, 0), (1, 1, 0)\}$ não é base desse plano.

3. Uma base da reta de $\mathbb{R}^3$ que passa na origem e no ponto $(1, 1, 1)$ é $\{(1, 1, 1)\}$. Os conjuntos $\{(-1, -1, -1)\}$ e $\{(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})\}$ também são bases.

Exercícios 17 Indique uma base para cada um dos seguintes conjuntos.

1. $\mathbb{R}^3$.

2. O plano de $\mathbb{R}^3$ definido por $2x_1 + 4x_2 - 2x_3 = 0$.

3. O hiperplano de $\mathbb{R}^5$ definido por $3x_1 - 6x_2 + 3x_3 - 2x_4 + 9x_5 = 0$.

O exemplo seguinte mostra como obter uma base do espaço nulo de uma matriz.

EXEMPLO 23 Indique uma base do espaço nulo da matriz $A = \begin{bmatrix} 1 & 2 & 1 & -1 & 3 \\ 2 & 4 & 3 & 0 & 2 \\ 3 & 6 & 4 & -1 & 5 \end{bmatrix}$.

Começa-se por resolver o sistema homogéneo $Ax = \vec{0}$ aplicando o método de Gauss à matriz A.

$$A = \begin{bmatrix} 1 & 2 & 1 & -1 & 3 \\ 2 & 4 & 3 & 0 & 2 \\ 3 & 6 & 4 & -1 & 5 \end{bmatrix} \longrightarrow \ldots \longrightarrow \begin{bmatrix} 1 & 2 & 0 & -3 & 7 \\ 0 & 0 & 1 & 2 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Tem-se então $\mathcal{N}(A) = \{ \begin{bmatrix} -2x_2 + 3x_4 - 7x_5 \\ x_2 \\ -2x_4 + 4x_5 \\ x_4 \\ x_5 \end{bmatrix}, x_2 = \forall, x_4 = \forall, x_5 = \forall \}$.

Fazendo cada uma das variáveis livres igual a 1 e as restantes iguais a 0, obtem-se o seguinte conj. de 3 vetores de $\mathcal{N}(A)$:

$$V = \{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 4 \\ 0 \\ 1 \end{bmatrix} \},$$

que é linearmente independente e que gera $\mathcal{N}(A)$ uma vez que

$$\begin{bmatrix} -2x_2+3x_4-7x_5 \\ x_2 \\ -2x_4+4x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -7 \\ 0 \\ 4 \\ 0 \\ 1 \end{bmatrix}.$$

Conclui-se pois que V é uma base de $\mathcal{N}(A)$.

De uma forma geral tem-se o seguinte procedimento para determinar uma base do espaço nulo de uma matriz arbitrária A.

. Aplica-se o método de Gauss a A. Seja R a matriz reduzida resultante.

. Se toda a coluna de R é *pivot* (i.e., o sistema $Ax = \vec{0} \Leftrightarrow Rx = \vec{0}$ só tem a solução trivial) $\Rightarrow \mathcal{N}(A) = \{\vec{0}\}$ e a base é $\emptyset$.
Caso contrário o conjunto das soluções dos sistema $Ax = \vec{0} \Leftrightarrow Rx = \vec{0}$ que se obtêm fazendo cada uma das variáveis livres igual a 1 e as restantes iguais a 0 é uma base de $\mathcal{N}(A)$. (A cardinalidade da base é pois o número de variáveis livres, i.e., o número de colunas não *pivot* de R.)

Vejamos agora como obter uma base para o espaço das colunas de uma matriz.

Exemplo 24 Indique uma base de $\mathcal{C}(A)$, com $A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 \\ 2 & -6 & 9 & -1 & 8 \\ 2 & -6 & 9 & -1 & 9 \\ -1 & 3 & -4 & 2 & -5 \end{bmatrix}$.

Aplicando o método de eliminação de Gauss à matriz A vem

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 \\ 2 & -6 & 9 & -1 & 8 \\ 2 & -6 & 9 & -1 & 9 \\ -1 & 3 & -4 & 2 & -5 \end{bmatrix} \longrightarrow \ldots \longrightarrow \begin{bmatrix} 1 & -3 & 4 & -2 & 5 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = A'.$$

Tem-se $\mathcal{C}(A) =< V >$, com $V = \{ \begin{bmatrix} 1 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \\ 9 \\ -4 \end{bmatrix}, \begin{bmatrix} 5 \\ 8 \\ 9 \\ -5 \end{bmatrix} \}$,

que é o conjunto das colunas de A que correspondem às colunas *pivot* de A'. Como V é linearmente independente, V é uma base de $\mathcal{C}(A)$.

Em geral para obter uma base do espaço das colunas de uma matriz A procede-se da seguinte forma.

. Aplica-se o método de Gauss a A (fase descendente). Seja A' a matriz em escada resultante.

. O conjunto das colunas de A que correspondem às colunas *pivot* de A' é uma base de $\mathcal{C}(A)$. (A cardinalidade da base é pois o número de colunas *pivot* de A'.)

O teorema seguinte enuncia uma característica muito particular da independência linear de vetores.

Teorema 2.5 *Se $V = \{v_1, v_2, \ldots, v_n\}$ é uma base de um subespaço vetorial S, todas as bases de S têm n vetores. Equivalentemente,*

Qualquer conjunto de vetores de S com mais do que n vetores é linearmente dependente.

Qualquer conjunto de vetores de S com menos do que n vetores não gera S.

O seguinte resultado é um consequência imediata do teorema anterior.

Teorema 2.6 *Seja V um conjunto não vazio de vetores de um subespaço vetorial S.*

Se V é linearmente independente e $\exists u \in S \backslash < V >$, então $V \cup \{u\}$ é linearmente independente. (Todo o independente pode ser ampliado até constituir uma base.)

Se $< V >= S$ e $\exists v \in V$ que é combinação linear dos outros vetores de V, então $< V \setminus \{v\} >= S$. (Todo o gerador pode ser reduzido até constituir uma base.)

Exercícios 18

1. Determine uma base para o espaço nulo e para o espaço das colunas de cada uma das seguintes matrizes.

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

d) $\begin{bmatrix} 1 & 2 & 1 & -1 & 3 \\ 2 & 4 & 3 & 0 & 2 \\ 3 & 6 & 4 & -1 & 5 \end{bmatrix}$ e) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

2. Construa uma base de $\mathbb{R}^3$ que inclua o vetor $(1, 1, 1)$.

3. Considere a matriz $A = \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & 0 & 3 \\ 2 & 1 & 1 & 6 \end{bmatrix}$.

   Verifique que $v = (0, 3, 3, -1) \in \mathcal{N}(A)$ e indique uma base de $\mathcal{N}(A)$ que inclua v.

2.4 Dimensão e característica

O facto de todas as bases de um subspaço vetorial terem o mesmo número de vetores, valida a definição seguinte.

Definição 14 Se S é um subespaço vetorial, a *dimensão* de S, representada por dim$S$, é a cardinalidade de uma base de S.

Observações 13

1. $\dim \mathbb{R}^3 = 3$.
2. O plano $X0Y$ tem dimensão 2. Todo o plano de $\mathbb{R}^3$ que passa na origem tem dimensão 2.
3. A reta de $\mathbb{R}^3$ que passa na origem e no ponto $(1, 1, 1)$ tem dimensão 1. Toda a recta de $\mathbb{R}^3$ que passa na origem tem dimensão 1.
4. O hiperplano $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$, com $a_1, a_2, \ldots, a_n$ em $\mathbb{R}$ não todos nulos, tem dimensão $n - 1$.
5. Se A' é uma matriz em escada que resulta de aplicar o método de Gauss à matriz A do tipo $m \times n$,

   a) a dimensão de $\mathcal{N}(A)$ é o número de colunas não *pivot* de A';

   b) a dimensão de $\mathcal{C}(A)$ é o número de colunas *pivot* de A'.

   Tem-se portanto $n = \dim \mathcal{N}(A) + \dim \mathcal{C}(A)$.

Exercício 19

1. Calcule $\dim S$, com $S =< \{(1, 0, 1), (1, -1, 0), (3, -1, 2)\} >$ e $S = \{(x, y, z, t) \in \mathbb{R}^4 : x - 2y + z - t = 0\}$.
2. Para que valores de α a dimensão do subspaço $S =< \{(1, \alpha, -1), (-1, 1, 1), (\alpha, 0, -1)\} >$ é 3?

Definição 15 Chama-se *característica* de uma matriz A, e representa-se por $\operatorname{car} A$, a dimensão de $\mathcal{C}(A)$.

Pode provar-se o seguinte.

Teorema 2.7 *Para toda a matriz A tem-se* $\operatorname{car} A = \operatorname{car} A^\top$.

Observações 14 Seja A uma matriz do tipo $m \times n$.

1. $\operatorname{car} A = \dim \mathcal{C}(A) = \dim \mathcal{L}(A)$, em que $L(A)$ é o subespaço de $\mathbb{R}^n$ gerado pelas m linhas de A.

2. $\dim \mathcal{N}(A) = n - \operatorname{car} A$.

Os teoremas que se seguem estabelecem relações entre várias noções que foram apresentadas.

Teorema 2.8 *Sejam A uma matriz do tipo $m \times n$ e b um vetor de $\mathbb{R}^m$. As seguintes proposições são equivalentes.*

1. *O sistema $Ax = b$ é possível.*

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

Teorema 2.9 *Seja A uma matriz do tipo $m \times n$. As seguintes proposições são equivalentes.*

1. *O sistema $Ax = b$ é possível para todo o vetor $b \in \mathbb{R}^m$.*

2. $\operatorname{car} A = m$.

Teorema 2.10 *Seja A uma matriz quadrada de ordem n. As seguintes proposições são equivalentes.*

1. *A é invertível.*

2. $\operatorname{car} A = n$.

3. $\dim \mathcal{N}(A) = 0$.

4. *O sistema $Ax = b$ é possível e determinado para todo o vetor $b \in \mathbb{R}^n$.*

Exercícios 20

1. Determine uma base e a dimensão dos subespaços de $\mathbb{R}^4$ gerados pelos seguintes conjuntos de vetores.

   a) $\{(1,0,2,3),(7,4,-2,1),(5,2,4,1),(3,2,0,1)\}$

   b) $\{(1,3,2,-1),(2,0,-1,0),(1,1,1,1),(1,2,0,0),(5,6,2,0)\}$

2. Considere o subconjunto de $\mathbb{R}^4$,

$$V = \{(x_1,x_2,x_3,x_4) : x_1 = x_2 - 3x_3,\ x_3 = 2x_4\}.$$

   a) Mostre que V é subespaço vetorial.

   b) Indique uma base de V.

3. Indique quais dos seguintes conjuntos são bases de $\mathbb{R}^2$:

   a) $V = \{(1,-1),(3,0)\}$

   b) $U = \{(1,1),(0,2),(2,3)\}$

   c) $W = \{(1,1),(0,8)\}$.

4. Indique quais dos seguintes conjuntos são bases de $\mathbb{R}^3$.

   a) $V = \{(1,1,1),(0,2,3),(1,0,2)\}$

   b) $U = \{(1,0,1),(2,4,8)\}$

   c) $W = \{(3,0,1),(1,1,1),(4,1,2)\}$.

5. Considere em $\mathbb{R}^3$ os vetores $v_1 = (\alpha,6,-1)$, $v_2 = (1,\alpha,-1)$ e $v_3 = (2,\alpha,-3)$, com $\alpha \in \mathbb{R}$.

   a) Determine α de modo que $\{v_1,v_2,v_3\}$ seja uma base de $\mathbb{R}^3$.

   b) Para um dos valores de α determinados em a), determine as componentes do vetor $(-1,1,2)$ em relação à base correspondente.

6. Seja A uma matriz do tipo $m \times n$. Para cada um dos casos considerados na tabela seguinte, determine as dimensões de $\mathcal{C}(\mathcal{A})$, $\mathcal{N}(\mathcal{A})$ e $\mathcal{N}(\mathcal{A}^T)$.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
$m \times n$	3×3	3×3	3×3	5×9	9×5	4×4	6×2
car(A)	3	2	1	2	2	0	2

7. Responda às alíneas seguintes utilizando a informação, respeitante a uma matriz A do tipo $m \times n$, fornecida na tabela seguinte.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
$m \times n$	3×3	3×3	3×3	5×9	9×5	4×4	6×2
car(A)	3	2	1	2	2	0	2
car($A\|b$)	3	3	1	2	3	0	2

   a) Classifique os sistemas lineares $Ax = b$.

   b) Indique o número de variáveis livres dos sistemas $Ax = 0$.

   c) Qual é a dimensão de $\mathcal{N}(\mathcal{A})$?

8. Seja A uma matriz quadrada de ordem 3, cujo espaço das colunas define um plano de $\mathbb{R}^3$ que passa na origem. Pode o espaço nulo de A determinar um plano que passa na origem? Justifique.

9. Seja V o espaço vetorial gerado pelo conjunto de vetores de $\mathbb{R}^3$

$$\{(1, 0, 5), (1, 1, 1), (0, 3, 1), (-3, 0, -2)\}.$$

   a) Mostre que $V = \mathbb{R}^3$.

   b) Determine uma base de $\mathbb{R}^3$ contida no conjunto de vetores dado.

   c) Escreva o vetor $(-2, 3, 4)$ como combinação linear dos vetores da base obtida em b).

10. Considere a matriz $\begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 5 & 0 \\ 1 & -1 & 3 & -1 \\ 0 & 1 & -1 & 2 \end{bmatrix}$.

    a) Resolva o sistema homogéneo $Ax = \vec{0}$ e indique a dimensão do espaço nulo da matriz A.

    b) Mostre que o espaço nulo de A é gerado pelos vetores $(1, 2, 0, -1)$ e $(-1, 3, 1, -1)$.

    c) Verifique que $v = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ é solução do sistema $Ax = \begin{bmatrix} 4 \\ 6 \\ 2 \\ 2 \end{bmatrix}$, e mostre que se u é um vetor do espaço nulo de A, entao $v + u$ é também solução do sistema.

11. Considere $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ e $S = \{x \in \mathbb{R}^3 : Ax = Bx\}$.

    a) Mostre que S é um espaço vetorial de $\mathbb{R}^3$.

    b) Indique uma base de S.

    c) Determine um vetor não nulo do espaço nulo de A que pertença a S.

    d) Mostre que se y é um vetor que pertence simultaneamente a S e ao espaço nulo de A, então y também pertence ao espaço nulo de B.

12. Considere o sistema $Ax = b$ em que $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & -1 \end{bmatrix}$ e $b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

    a) Determine o conjunto das soluções do sistema $Ax = b$.

    b) Utilizando a resposta da alínea anterior, indique o espaço nulo de A. Interprete geometricamente o resultado obtido.

13. Sejam $A = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix}$ e $b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

    a) Determine uma base $\mathcal{N}(\mathcal{A})$.

    b) Determine uma solução do sistema $Ax = b$.

    c) Seja x_0 a solução obtida em b). Verifique que para todo o vetor $u \in \mathcal{N}(\mathcal{A})$, $x_0 + u$ é solução de $Ax = b$.

    d) Interprete geometricamente os resultados obtidos nas alíneas anteriores e conclua que não existem mais soluções para o sistema $Ax = b$.

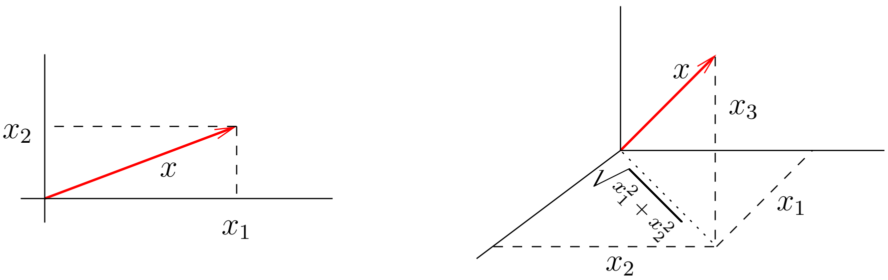
Capítulo 3

Norma, produto interno, ângulo, ortogonalidade e projeção

Neste capítulo vamos estender a $\mathbb{R}^m$ as noções conhecidas em $\mathbb{R}^2$ e $\mathbb{R}^3$ de norma de um vetor, produto interno e ângulo de vetores, estudar certas propriedades da ortogonalidade e formalizar a noção de projeção de um vetor sobre um subespaço vetorial.

3.1 Norma, produto interno, ângulo

Se $x = (x_1, x_2)$ é um vetor de $\mathbb{R}^2$, a norma de x é $\|x\| = \sqrt{x_1^2 + x_2^2}$ (ver Figura 3.1 a)).



a) **b)**

Figura 3.1: Um vetor a) $x = (x_1, x_2)$ de $\mathbb{R}^2$ e b) $x = (x_1, x_2, x_3)$ de $\mathbb{R}^3$.

Se $x = (x_1, x_2, x_3)$ é um vetor de $\mathbb{R}^3$, a norma de x é $\|x\| = \sqrt{(\sqrt{x_1^2 + x_2^2}\,)^2 + x_3^2} = \sqrt{x_1^2 + x_2^2 + x_3^2}$ (ver Figura 3.1 b)).
Em geral tem-se o seguinte.

Definição 16 Se $x = (x_1, x_2, \ldots, x_m)$ é um vetor de $\mathbb{R}^m$, a *norma* de x é

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_m^2}.$$

Exemplos 25

1. $\|(0, 1, 2)\| = \sqrt{0^2 + 1^2 + 2^2} = \sqrt{5}$.
2. $\|(1, -1, 1, 2)\| = \sqrt{1^2 + (-1)^2 + 1^2 + 2^2} = \sqrt{7}$.

A norma verifica as seguintes propriedades.

Proposições 3.1

1. $\|x\| \geq 0$ e $\|x\| = 0$ sse $x = \vec{0}$.
2. $\|\lambda x\| = |\lambda| \|x\|, \forall \lambda \in \mathbb{R}$.

Chama-se *vetor unitário* a um vetor de norma igual a 1. Note que se x é um vetor não nulo, $\frac{x}{\|x\|}$ é um vetor unitário, colinear e com o mesmo sentido do que x. Diz-se que $\frac{x}{\|x\|}$ é o *versor* do vetor x. Os versores dos vetores $(\frac{1}{2}, \frac{1}{2})$ e $(0, 1, 2)$ são $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ e $(0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$, respetivamente.
Se x e y são vetores de $\mathbb{R}^2$, a distância entre x e y, é $d(x, y) = \|x - y\|$ (ver Figura 3.2).

Definição 17 Se $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$ são vetores de $\mathbb{R}^m$, a *distância* entre x e y é

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_m - y_m)^2}.$$

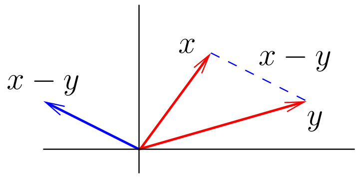


Figura 3.2: Vetores x e y de $\mathbb{R}^2$ e o vetor $x-y$.

Exemplo 26 A distância entre os vetores $(1,2,0,-1)$ e $(2,0-2,-1)$ é $\|(-1,2,2,0)\| = \sqrt{1+4+4+0} = 3$.

Observação 15 A norma de um vetor x, $\|x\| = \|x-\vec{0}\|$, é a distância de x à origem.

Consideremos os vetores $x = (x_1,x_2,x_3)$ e $y = (y_1,y_2,y_3)$ de $\mathbb{R}^3$. Tem-se o seguinte

$$\|x-y\|^2 = (x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2 = \|x\|^2+\|y\|^2-2(x_1y_1+x_2y_2+x_3y_3). \quad (3.1)$$

Isto é, o quadrado da distância entre x e y é a soma dos quadrados das normas menos o dobro de $x_1y_1+x_2y_2+x_3y_3$. Suponhamos que x e y são vetores não nulos e ortogonais (ver Figura 3.3). Nesse caso, por aplicação do Teorema de Pitágoras, tem-se

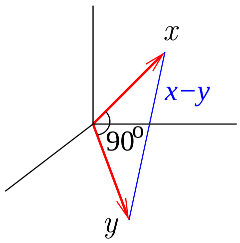


Figura 3.3: Dois vetores x e y de $\mathbb{R}^3$ ortogonais.

$$\|x-y\|^2 = \|x\|^2+\|y\|^2. \quad (3.2)$$

Combinando as expressões (3.1) e (3.2) temos a condição de ortogonalidade para vetores de $\mathbb{R}^3$: $x \perp y$ sse $x_1y_1+x_2y_2+x_3y_3 = 0$. Ao escalar $x_1y_1+x_2y_2+x_3y_3$ dá-se o nome de produto interno ou escalar dos vetores x e y. Vamos generalizar a $\mathbb{R}^m$ estes dois conceitos.

Definição 18 Dois vetores $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$ de $\mathbb{R}^m$, dizem-se *ortogonais* e representa-se por $x \perp y$, se $x_1y_1 + x_2y_2 + \cdots + x_my_m = 0$.

Exemplos 27 Os seguintes pares de vetores são ortogonais.

1. $(1, 1, 0, 1)$ e $(1, -1, 3, 0)$.

2. $(-2, 1, 0, 3, -1)$ e $(2, 1, 2, \frac{1}{3}, -2)$.

Definição 19 Sejam $x = (x_1, x_2, \ldots, x_m)$ e $y = (y_1, y_2, \ldots, y_m)$ vetores de $\mathbb{R}^m$. Chama-se *produto interno* (ou escalar) de x e y ao escalar

$$x|y = x_1y_1 + x_2y_2 + \cdots + x_my_m.$$

Observação 16 $x|y = x_1y_1 + x_2y_2 + \cdots + x_my_m = [x_1\, x_2\, \ldots\, x_m] \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = x^\top y.$

É fácil verificar que o produto interno satisfaz as seguintes propriedades.

Proposições 3.2

1. $x|y = y|x$.

2. $x|(y + z) = x|y + x|z$.

3. $\lambda(x|y) = \lambda x|y = x|\lambda y$.

4. $x|x = \|x\|^2$.

5. $x|x = 0$ sse $x = \vec{0}$ ($\vec{0}$ é o único vetor ortogonal a si mesmo).

Observação 17 Como $\|x + y\|^2 = (x + y)|(x + y) = \|x\|^2 + 2(x|y) + \|y\|^2$, tem-se $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ sse $x|y = 0$, que é o enunciado do Teorema de Pitágoras em $\mathbb{R}^m$.

A inequação que se apresenta de seguida relaciona o produto interno com o produto das normas, é considerada uma das desigualdades mais importantes em Álgebra Linear.

Teorema 3.3 *(Desigualdade de Cauchy-Schwarz.) Para quaisquer vetores x e y de $\mathbb{R}^m$ tem-se $|x|y| \leq \|x\|\|y\|$, com igualdade sse x e y são colineares.*

Demonstração: Se $x = \vec{0}$ ou $y = \vec{0}$, tem-se $x|y = \|x\|\|y\| = 0$ e nada mais há a provar.
Consideremos pois que os vetores x e y são ambos não nulos.
A chave da prova é a desigualdade seguinte, que é válida para todo o escalar λ.

$$\|x\|^2 - 2\lambda x|y + \lambda^2\|y\|^2 \geq 0. \tag{3.3}$$

De facto, $\|x\|^2 - 2\lambda x|y + \lambda^2\|y\|^2 = (x - \lambda y)|(x - \lambda y) = \|x - \lambda y\|^2 \geq 0, \forall\lambda \in \mathbb{R}$.
Substituindo na inequação (3.3) λ por $\frac{x|y}{\|y\|^2}$, tem-se $\|x\|^2 - 2\frac{(x|y)^2}{\|y\|^2} + \frac{(x|y)^2}{\|y\|^4}\|y\|^2 \geq 0 \Rightarrow \|x\|^2 - \frac{(x|y)^2}{\|y\|^2} \geq 0 \Rightarrow \|x\|^2\|y\|^2 \geq (x|y)^2 \Rightarrow \|x\|\|y\| \geq |x|y|$, ficando assim provada a primeira parte do teorema.
Se os vetores x e y não são colineares, i.e., $x \neq \lambda y, \forall\lambda \in \mathbb{R} \Rightarrow \|x - \lambda y\|^2 > 0$ e portanto $\|x\|\|y\| > |x|y|$.
Se $x = \alpha y$, para algum $\alpha \in \mathbb{R}$, então $|x|y| = |\alpha y|y| = |\alpha|\|y\|^2 = |\alpha|\|y\|\|y\| = |\alpha|\|\frac{1}{\alpha}x\|\|y\| = |\alpha|\frac{1}{|\alpha|}\|x\|\|y\| = \|x\|\|y\|$. $\square$

Da desigualdade de Cauchy-Schwarz facilmente se deduz o seguinte resultado conhecido por desigualdade triangular.

Teorema 3.4 *(Desigualdade $\triangle_{\text{lar}}$.) Para todo o par de vetores $x, y \in \mathbb{R}^m$, $\|x + y\| \leq \|x\| + \|y\|$.*

Demonstração: $\|x+y\|^2 = (x+y)|(x+y) = \|x\|^2 + 2x|y + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$. Logo, $\|x + y\| \leq \|x\| + \|y\|$. $\square$

Recorde que o ângulo θ entre dois vetores não nulos x e y de $\mathbb{R}^2$ é definido por $\cos\theta = \frac{x|y}{\|x\|\|y\|}$ e $\theta \in [0, \pi]$.

Para vetores não nulos x e y de $\mathbb{R}^m$, a desigualdade Cauchy-Schwarz $|x|y| \leq \|x\|\|y\|$ é equivalente a $-1 \leq \frac{x|y}{\|x\|\|y\|} \leq 1$. É pois legítimo associar à razão $\frac{x|y}{\|x\|\|y\|}$ o coseno de um ângulo e assim generalizar a noção de ângulo entre vetores.

Definição 20 O *ângulo* entre dois vetores $x, y \in \mathbb{R}^m$ é o ângulo $\theta \in [0, \pi]$, tal que $x|y = \|x\|\|y\| \cos\theta$.

Observação 18 $x, y \neq \vec{0}, x|y = \|x\|\|y\| \cos\theta = 0$ $(x \perp y)$ sse $\cos\theta = 0$, i.e., $\theta = \frac{\pi}{2}$.

Exemplo 28 O ângulo entre $x = (1,1,0,1)$ e $y = (1,1,1,1)$ é $\theta = \arccos \frac{x|y}{\|x\|\|y\|} = \arccos \frac{3}{\sqrt{3}\sqrt{4}} = \arccos \frac{\sqrt{3}}{2} = \frac{1}{6}\pi$.

Exercícios 21

1. Calcule as normas dos seguintes vetores.

   a) $(1, -1, 2)$ b) $(-1, 0, \pi, 0)$ c) $(5, 0, 1, 0, 1, 3)$.

2. Calcule as distância entre os seguintes pares de vetores.

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

   c) $(5, 0, 1, 0, 1, 3)$ e $(-1, 2, 0, 1, 1, 0)$.

3. Determine todos os vetores unitários que fazem ângulos de $\frac{\pi}{3}$ com cada um dos seguintes pares de vetores.

   a) $(1, 0, 0)$ e $(0, 1, 0)$ b) $(1, 0, 1)$ e $(0, 1, 0)$.

4. Identifique um vetor não nulo que seja ortogonal a ambos os vetores de cada um dos seguintes pares.

   a) $(1, 0, 1)$ e $(1, 1, -1)$ b) $(1, -1, 2)$ e $(2, 1, -1)$.

5. Indique dois vetores não nulos ortogonais entre si e ortogonais ao vetor de cada uma das alíneas seguintes.

   a) $(1, 1, 1)$ b) $(1, 2, 1, -3)$.

6. Sejam x e y vetores de $\mathbb{R}^m$. Prove os seguintes resultados.

   a) $\|x+y\| = \|x-y\|$ sse x e y são ortogonais.

   b) Os vetores $x-y$ e $x+y$ são ortogonais sse $\|x\| = \|y\|$.

   c) Se x e y são ortogonais, então $\|x-y\|^2 = \|x\|^2 + \|y\|^2$.

   d) Se x e y são unitários e ortogonais, então $\|x-y\| = \sqrt{2}$.

3.2 Ortogonalidade

Atrás definimos ortogonalidade de dois vetores de $\mathbb{R}^m$. Vamos agora estender a noção de ortogonalidade a subespaços vetoriais. Começamos com a noção de ortogonalidade de um conjunto de vetores.

Definição 21 Um conjunto $\{v_1, v_2, \ldots, v_k\}$ de vetores de $\mathbb{R}^m$ é *ortogonal* se $v_i|v_j = 0$, com $i \neq j = 1, \ldots, k$. Se adicionalmente $\|v_i\| = 1$, para $i = 1, \ldots, k$, o conj. diz-se *ortonormal*.

Observações 19

1. A base canónica de $\mathbb{R}^m$ é um conjunto ortonormal de m vetores.

2. Se $\{v_1, \ldots, v_k\}$ é um conjunto ortogonal de vetores não nulos, $\{\frac{v_1}{\|v_1\|}, \ldots, \frac{v_k}{\|v_k\|}\}$ é ortonormal.

O resultado seguinte relaciona ortogonalidade e independência linear de vetores.

Teorema 3.5 *Um conjunto ortogonal de vetores não nulos é linearmente independente.*

Demonstração: Seja $\{v_1, v_2, \ldots, v_k\}$ um conjunto ortogonal vetores não nulos, e consideremos uma combinação linear nula

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = \vec{0} \tag{3.4}$$

desses vetores. Pretende provar-se que a combinação linear anterior apenas é realizada com os coeficientes $\lambda_1, \lambda_2, \ldots, \lambda_k$ todos nulos.

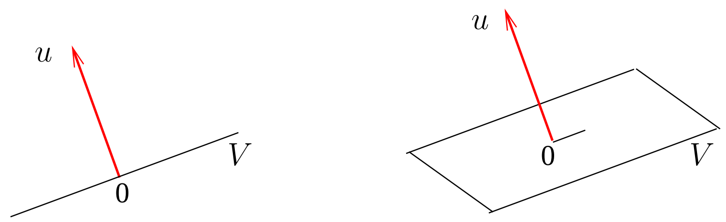
Se fizermos o produto interno por v_1 de cada um dos membros da equação (3.4) obtem-se $v_1|(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k) = v_1|\vec{0} \Leftrightarrow \lambda_1(v_1|v_1) + \lambda_2(v_1|v_2) + \cdots + \lambda_k(v_1|v_k) = 0$. Uma vez que $v_1|v_1 \neq 0$ e $v_2|v_2 = \cdots = v_k|v_k = 0$, necessariamente $\lambda_1 = 0$.

Para provar que $\lambda_2, \ldots, \lambda_k$ também são iguais a zero, procede-se de forma análoga tomando, respetivamente, o produto interno por $v_2, \ldots, v_k$ de ambos os membros da equação (3.4). $\square$

Do resultado anterior e do facto de que todo o conjunto linearmente independente de $\mathbb{R}^m$ não inclui mais do que m vetores decorre o seguinte.

Corolário 3.6 *Um conjunto ortogonal de vetores não nulos de $\mathbb{R}^m$ não inclui mais do que m vetores.*

A figura seguinte representa um vetor u ortogonal a dois subespaços vetoriais V: uma reta e a um plano que incluem a origem.



De um forma geral tem-se a seguinte definição.

Definição 22 Um vetor $u \in \mathbb{R}^m$ é *ortogonal a um subespaço V* de $\mathbb{R}^m$, e representa-se por $u \perp V$, se u é ortogonal a todo o vetor de V, i.e., $u|v = 0, \forall\, v \in V$.

Exemplo 29 O vetor $(-1, 1, 2)$ é ortogonal a $\{(x, y, z) \in \mathbb{R}^3 : x - y - 2z = 0\}$.

Observação 20 O único vetor ortogonal a $\mathbb{R}^m$ é o vetor nulo.

O resultado seguinte estabelece que para decidir se um vetor é ortogonal a um subespaço vetorial basta testar se é ortogonal a um conjunto finito de vetores do subespaço.

Proposição 3.7 *Sejam $u \in \mathbb{R}^m$ e V o subespaço de $\mathbb{R}^m$ gerado por $\{v_1, v_2, \ldots, v_n\}$. O vetor u é ortogonal a V sse u é ortogonal a cada um dos vetores v_i, com $i = 1, \ldots, n$.*

Demonstração: É óbvio que se $u \perp V \Rightarrow u \perp v_i, i = 1, \ldots, n$.

Para provar a implicação no outro sentido consideremos um vetor abritrário $v \in V$. Como $\{v_1, v_2, \ldots, v_n\}$ gera o subspaço, existem escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tais que $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Fazendo o produto interno por u de cada um dos membros da igualdade anterior, e tendo em conta que $u \perp v_1, u \perp v_2, \ldots, u \perp v_n$, tem-se

$$u|v = \lambda_1 \underbrace{u|v_1}_{0} + \lambda_2 \underbrace{u|v_2}_{0} + \cdots + \lambda_n \underbrace{u|v_n}_{0}$$

e portanto $u|v = 0$, como queria provar-se. □

Observação 21 Se V é um subespaço vetorial, $u \perp V$ sse u é ortogonal a uma base de V.

Exercício 22 Mostre que o vetor $(2, 1, 1, -1)$ é ortogonal ao espaço das colunas da matriz

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

Vejamos como obter o conjunto dos vetores ortogonais a um subespaço vetorial.

Sejam V um subespaço de $\mathbb{R}^m$ e $\{v_1, v_2, \ldots, v_n\}$ um conjunto gerador de V. O vetor x de $\mathbb{R}^m$ é ortogonal a V sse $\begin{cases} v_1|x = 0 \\ v_2|x = 0 \\ \quad\vdots \\ v_n|x = 0 \end{cases}$. Equivalentemente, se definirmos a matriz $A := \begin{bmatrix} v_1 & v_2 & \ldots & v_n \\ | & | & & | \end{bmatrix}$, tem-se que x é ortogonal a V sse $A^\top x = \vec{0}$. Por outras palavras, o vetor x é ortogonal a $V = \mathcal{C}(A)$ sse $x \in \mathcal{N}(A^\top)$. Tem-se assim provado os pontos 1, 2 e 3 do seguinte teorema.

Teorema 3.8 *Seja V um subespaço vetorial de $\mathbb{R}^m$.*

1. *O conj. dos vetores de $\mathbb{R}^m$ ortogonais a V, que se chama* complemento ortogonal *de V e se representa por $V^\perp$, é um subespaço vetorial de $\mathbb{R}^m$.*

2. $\dim V^\perp = m - \dim V$.

3. $V \cap V^\perp = \{\vec{0}\}$.

4. *Se $\{v_1, v_2, \ldots, v_r\}$ é uma base de V e $\{w_1, w_2, \ldots, w_{m-r}\}$ é uma base de $V^\perp$, então $\{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{m-r}\}$ é uma base de $\mathbb{R}^m$.*

Demonstração: Para provar 4, considere uma combinação linear nula dos m vetores $v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{m-r}$

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r + \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_{m-r} w_{m-r} = \vec{0} \Leftrightarrow$$

$$\Leftrightarrow \underbrace{\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r}_{\in V} = \underbrace{-(\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_{m-r} w_{m-r})}_{\in V^\perp}.$$

Tendo em conta que $\vec{0}$ é o único vetor que pertence a V e $V^\perp$, e que $\{v_1, v_2, \ldots, v_r\}$ e $\{w_1, w_2, \ldots, w_{m-r}\}$ são conjuntos linearmente independentes, conclui-se que

$$\begin{cases} \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = \vec{0} & \Rightarrow \quad \lambda_1 = \cdots = \lambda_r = 0 \\ \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_{m-r} w_{m-r} = \vec{0} & \Rightarrow \quad \alpha_1 = \cdots = \alpha_{m-r} = 0. \end{cases}$$

i.e., que os coeficientes da combinação linear nula dos m vetores $v_1, v_2, \ldots, v_r, w_1, w_2, \ldots,$ w_{m-r} são necessariamente todos nulos, o que prova que os vetores constituem uma base de $\mathbb{R}^m$. $\square$

Observações 22

1. $\mathcal{C}^\perp(A) = \mathcal{N}(A^\top)$.

2. Em $\mathbb{R}^3$,

   a) o complemento ortogonal de $\{\vec{0}\}$ é $\mathbb{R}^3$;

   b) o complemento ortogonal de uma reta que passa na origem é o plano que passa na origem e é perpendicular à reta;

   c) o complemento ortogonal de um plano que passa na origem é a reta que passa na origem e é perpendicular ao plano;

   d) o complemento ortogonal de $\mathbb{R}^3$ é $\{\vec{0}\}$.

Exercícios 23

1. Determine os complementos ortogonais do espaço das colunas de cada uma das seguintes matrizes.

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

   d) $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ e) $\begin{bmatrix} 1 & 1 \\ -1 & -2 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}$.

2. Verifique que o vetor $(4, 2, -1)$ é ortogonal ao espaço das colunas da matriz

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

   Qual é o complemento ortogonal de $\mathcal{C}(A)$?

3. Determine os complementos ortogonais dos subespaços gerados por $\{(1, 2, 2, 1), (1, 0, 2, 0)\}$ e por $\{(1, 1, 2, -1)\}$.

4. Calcule a dimensão e indique uma base do complemento ortogonal para cada um dos seguintes subespaços.

   a) $< \{(1,1)\} >$ b) $< \{(1,1,3),(1,1,2)\} >$ c) $< \{(1,1,0,0),(0,2,4,5)\} >$

   d) $< \{(2,2,1,0),(2,4,0,1),(4,-2,1,-1)\} >$.

5. Construa uma base de $\mathbb{R}^3$ que inclua vetores do subespaço gerado por $\{(1,1,3),(1,1,2)\}$ e do seu complemento ortogonal.

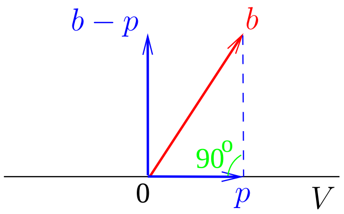
6. Uma matriz quadrada A de ordem n diz-se *ortogonal* se as colunas são unitárias e quaisquer duas colunas distintas são ortogonais. Prove os seguintes resultados.

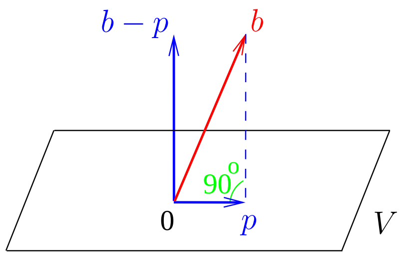
   a) A matriz A é ortogonal sse $A^{-1} = A^\top$.

   b) Se a matriz A é ortogonal, então é simétrica sse $A^2 = I$.

3.3 Projeção ortogonal

A figura seguinte representa a projeção ortogonal p do vetor b sobre dois subespaços vetoriais V: uma reta e um plano que passam na origem.





Note que $p \in V$ e $b - p \in V^\perp$.

Pode provar-se o seguinte.

Teorema 3.9 *Seja V é um subespaço vetorial de $\mathbb{R}^m$.*

1. *Para todo o vetor $b \in \mathbb{R}^m$, existe um e um só vetor $p \in V$, tal que $b - p \in V^\perp$.*

2. *Como $b = p + (b - p)$, os vetores p e $b - p$ dão uma decomposição única de b como soma de um vetor em V com outro em $V^\perp$.*

Demonstração: Sejam $\{v_1, v_2, \ldots, v_r\}$ uma base de V e $\{w_1, w_2, \ldots, w_{m-r}\}$ uma base de $V^\perp$. Como $\{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_{m-r}\}$ é uma base de $\mathbb{R}^m$, existem escalares $\lambda_1, \lambda_2, \ldots, \lambda_r, \alpha_1, \alpha_2, \ldots, \alpha_{m-r}$ tais que

$$b = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r + \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_{m-r} w_{m-r}.$$

Se chamarmos p ao vetor $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r$, tem-se $b = p + (b - p)$, com $p \in V$ e $b - p \in V^\perp$.

Suponha que, para além de p, existe $q \in V$ tal que $b - q \in V^\perp$. Tem-se então, tendo em conta que V e $V^\perp$ são subespaços vetoriais e portanto fechados para a adição e multiplicação escalar,

$$p + (b - p) = q + (b - q) \Leftrightarrow \underbrace{p - q}_{\in V} = \underbrace{(b - q) - (b - p)}_{\in V^\perp},$$

o que implica $p - q = \vec{0}$, i.e., $p = q$. $\square$

O teorema anterior valida a seguinte definição.

Definição 23 Sejam V um subespaço vetorial de $\mathbb{R}^m$ e $b \in \mathbb{R}^m$. A *projeção (ortogonal)* de b sobre V, representada por $\text{proj}_V b$, é o único vetor $p \in V$ tal que $b - p \in V^\perp$.

Exemplo 30 Verifique que $p = (4, 2, -1)$ é a projeção do vetor $b = (7, -1, 5)$ sobre o espaço das colunas da matriz $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$.

. Em primeiro lugar mostra-se que $p = (4, 2, -1) \in \mathcal{C}(A)$, i.e., que o sistema $Ax = p$ é possível.

$$[A|p] = \left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 1 & 2 \\ -1 & 0 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 1 & 2 \\ 0 & \frac{1}{2} & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$ de onde se conclui que o sistema é possível.

($x_1 = 1, x_2 = 2$ é a solução. De facto, $\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}$.)

. Em seguida verifica-se que $b - p = (3, -3, 6) \in V^\perp$. Para isso calcula-se o produto interno do vetor $(3, -3, 6)$ com cada uma das colunas de A. Tem-se $(3, -3, 6)|(2, 0, -1) = 0$ e $(3, -3, 6)|(1, 1, 0) = 0$, i.e., $b - p = (3, -3, 6)$ é ortogonal a cada uma das colunas de A e portanto $b - p \perp \mathcal{C}(A) \Leftrightarrow b - p \in \mathcal{C}^\perp(A)$.

Observação 23 Se p é a projeção de b sobre V, $b - p$ é a projeção de b sobre $V^\perp$. (De facto, $b - p \in V^\perp$ e $b - (b - p) = p \in V = (V^\perp)^\perp$.)

O teorema seguinte estabelece uma definição alternativa de projeção de um vetor sobre um subspaço vetorial.

Teorema 3.10 *Sejam V um subespaço vetorial de $\mathbb{R}^m$ e $b \in \mathbb{R}^m$. A projeção de b sobre V é o vetor de V a menor distância de b.*

Demonstração: Sejam p a projeção de b sobre V e q um qualquer vetor de V. Como V é um subespaço vetorial, $p - q \in V$ e portanto $p - q \perp b - p$. Pelo teorema de Pitágoras (se $x \perp y \Rightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2$), vem $\|(b - p) + (p - q)\|^2 = \|b - p\|^2 + \|p - q\|^2 \Leftrightarrow \|b - q\|^2 = \|b - p\|^2 + \|p - q\|^2 \Rightarrow \|b - q\|^2 \geq \|b - p\|^2$, que permite concluir que a distância de p a b é menor ou igual do que a distância de q a b. $\square$

Em seguida descreve-se um procedimento para determinar a projeção de um vetor $b \in \mathbb{R}^m$ sobre um subespaço vetorial V de $\mathbb{R}^m$.

. Identifica-se uma base $\{v_1, v_2, \ldots, v_n\}$ de V, e define-se a matriz $A := \begin{bmatrix} v_1 & v_2 & \ldots & v_n \\ | & | & & | \end{bmatrix}$

. Identifica-se uma base $\{w_1, w_2, \ldots, w_{m-n}\}$ de $\mathcal{N}(A^\top) = V^\perp$.

. Resolve-se o sistema (determinado) $\left[\begin{array}{cccccc|c} v_1 & \ldots & v_n & w_1 & \ldots & w_{m-n} & b \\ | & & | & | & & | & | \end{array}\right]$.

Seja $(\lambda_1, \ldots, \lambda_n,\ \alpha_1, \ldots, \alpha_{m-n})$ a solução do sistema.

. Tem-se $b = \underbrace{\lambda_1 v_1 + \cdots + \lambda_n v_n}_{p \in V} + \underbrace{\alpha_1 w_1 + \cdots + \alpha_{m-n} w_{m-n}}_{b-p \in V^\perp}$

$$\begin{array}{ccc} || & \qquad\qquad & || \\ \mathrm{proj}_V b & & \mathrm{proj}_{V^\perp} b \end{array}$$

Exemplo 31 Determine a projeção do vetor $(7, -1, 5)$ sobre o subespaço gerado por $\{(2, 0, -1), (1, 1, 0)\}$.

Seja $V =< \{(2, 0, -1), (1, 1, 0)\} >$. Como $\{(2, 0, -1), (1, 1, 0)\}$ é linearmente independente, é uma base de V. Vamos definir $A := \begin{bmatrix} 2 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$ e identificar uma base de $\mathcal{N}(A^\top)$.

$$A^\top = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} \longrightarrow \ldots \longrightarrow \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{bmatrix},\ \mathcal{N}(A^\top) = \{ \begin{bmatrix} \frac{1}{2}x_3 \\ -\frac{1}{2}x_3 \\ x_3 = \forall \end{bmatrix} \}.$$

$\{(1, -1, 2)\}$ é um a base de $\mathcal{N}(A^\top) = V^\perp$.

Vamos agora resolver o sistema $\left[\begin{array}{ccc|c} 2 & 1 & 1 & 7 \\ 0 & 1 & -1 & -1 \\ -1 & 0 & 2 & 5 \end{array}\right] \longrightarrow \ldots \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$.

Assim, $\begin{bmatrix} 7 \\ -1 \\ 5 \end{bmatrix} = \underbrace{1\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\substack{\begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix} \in V \\ \| \\ \mathrm{proj}_V b}} + \underbrace{3\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}}_{\substack{\begin{bmatrix} 3 \\ -3 \\ 6 \end{bmatrix} \in V^\perp \\ \| \\ \mathrm{proj}_{V^\perp} b}}$

Observação 24 O vetor p é a projeção de b sobre $\mathcal{C}(A)$ sse

(i) $p \in \mathcal{C}(A)$, i.e., $\exists \bar{x} : p = A\bar{x}$, e

(ii) $b - p \in \mathcal{N}(A^\top)$, i.e., $A^\top(b-p) = \vec{0} \overset{(i)}{\Rightarrow} A^\top(b - A\bar{x}) = \vec{0} \Leftrightarrow A^\top A\bar{x} = A^\top b$.

Tem-se pois a seguinte forma alternativa de determinar a projeção de b sobre $\mathcal{C}(A)$.

. Resolve-se o sistema $A^\top Ax = A^\top b$ (*equações normais*). Seja $\bar{x}$ uma solução.

. Define-se $p = \mathrm{proj}_{\mathcal{C}(A)} b := A\bar{x}$.

Exercício 24 Determine a projeção do vetor $(4, -1, 1)$ sobre $V =< \{(1, 0, 1), (1, 1, 0)\} >$.

Observação 25 A *projeção* de b *sobre um vetor* $v \neq \vec{0}$ é a projeção de b sobre o subespaço (a reta) gerado por v. Tem-se pois, $A := \begin{bmatrix} v \\ | \end{bmatrix}$, $\underbrace{A^\top A}_{v|v} x = \underbrace{A^\top b}_{v|b} \Rightarrow \bar{x} = \frac{v|b}{v|v}$, e portanto $p = \mathrm{proj}_v b = A\bar{x} = \frac{v|b}{v|v} v = \frac{v|b}{\|v\|\,\|v\|} v = \|b\| \frac{v|b}{\|b\|\,\|v\|} \frac{v}{\|v\|} = \|b\| \cos\theta \frac{v}{\|v\|}$, em que θ é o ângulo entre b e v.

Exemplo 32 A projeção de $b = (2, 1, 6, 3)$ sobre $v = (1, 1, 1, 1)$ é $\frac{v|b}{v|v} v = \frac{12}{4}(1, 1, 1, 1) = (3, 3, 3, 3)$.

O resultado seguinte estabelece uma condição necessária e suficiente para a invertibilidade da matriz $A^\top A$.

Proposição 3.11 *Uma matriz A, do tipo $m \times n$, tem característica n sse a matriz $A^\top A$ é invertível.*

Demonstração: Suponha que $\operatorname{car} A < n$. Então, para algum vetor não nulo $x \in \mathbb{R}^n$ tem-se $Ax = \vec{0} \Rightarrow (A^\top A)x = \vec{0}$, garantindo que $A^\top A$ não é invertível. Fica assim provado que, se a matriz $A^\top A$ é invertível, então $\operatorname{car} A = n$.
Considere agora $x \in \mathcal{N}(A^\top A)$, i.e., x tal que $(A^\top A)x = \vec{0} \Leftrightarrow A^\top(Ax) = \vec{0} \Rightarrow Ax \in \mathcal{N}(A^\top)$. Por outro lado, obviamente $Ax \in \mathcal{C}(A)$.
Assim, se $x \in \mathcal{N}(A^\top A)$, $Ax \in \mathcal{N}(A^\top) \cap \mathcal{C}(A)$ e como $\mathcal{N}(A^\top) = \mathcal{C}^\perp(A)$, tem-se $Ax = \vec{0} \overset{\operatorname{car}(A)=n}{\Rightarrow} x = \vec{0}$. A equivalência $\mathcal{N}(A^\top A) = \vec{0}$ sse a matriz $A^\top A$ é invertível, completa a demonstração. □

Pode então concluir-se que, se o conjunto das colunas de A é linearmente independente, a única solução das equações normais $A^\top Ax = A^\top b$ é $\bar{x} = (A^\top A)^{-1}A^\top b$ e, consequentemente,

$$\operatorname{proj}_{\mathcal{C}(A)} b = A\,(A^\top A)^{-1}A^\top b.$$

Definição 24 Se o conjunto das colunas da matriz A é linearmente independente ($A^\top A$ é invertível), chama-se a $A\,(A^\top A)^{-1}A^\top$ *matriz de projeção (ortogonal)* sobre o espaço das colunas de A.

Observação 26 No caso em que o conjunto das colunas de $A = \begin{bmatrix} w_1 & w_2 & \dots & w_n \\ | & | & & | \end{bmatrix}$ é ortonormal, tem-se $A^\top A = \begin{bmatrix} w_1 & — \\ w_2 & — \\ \vdots & \\ w_n & — \end{bmatrix} \begin{bmatrix} w_1 & w_2 & \dots & w_n \\ | & | & & | \end{bmatrix} = I$ e portanto a matriz de projeção sobre $\mathcal{C}(A)$ é $AA^\top$. Dado um vetor arbitrário $b \in \mathbb{R}^m$,

$$\text{proj}_{\mathcal{C}(A)}b = (AA^\top)b = A(A^\top b) = A \begin{bmatrix} w_1|b \\ w_2|b \\ \vdots \\ w_n|b \end{bmatrix} = \underbrace{w_1|b\ w_1}_{\text{proj. de } b \text{ sobre } w_1} + \underbrace{w_2|b\ w_2}_{\text{proj. de } b \text{ sobre } w_2} + \cdots + \underbrace{w_n|b\ w_n}_{\text{proj. de } b \text{ sobre } w_n}$$

Tem-se assim o seguinte teorema.

Teorema 3.12 *Sejam $\{w_1, w_2, \ldots, w_n\}$ uma base ortonormal de um subespaço V de $\mathbb{R}^m$ e $b \in \mathbb{R}^m$. A projeção de b sobre V é $p = w_1|b\ w_1 + w_2|b\ w_2 + \cdots + w_n|b\ w_n$, a soma das projeções de b sobre cada um dos vetores da base.*

No caso em que a base é ortogonal, tem-se um resultado semelhante.

Teorema 3.13 *Sejam $\{w_1, w_2, \ldots, w_n\}$ uma base ortogonal de um subespaço V de $\mathbb{R}^m$ e $b \in \mathbb{R}^m$. A projeção de b sobre V é $p = \frac{w_1|b}{w_1|w_1}\ w_1 + \frac{w_2|b}{w_2|w_2}\ w_2 + \cdots + \frac{w_n|b}{w_n|w_n}\ w_n$, a soma das projeções de b sobre cada um dos vetores da base.*

Exemplo 33 Determine a projeção do vetor $b = (1, 3, 1)$ sobre o subespaço V gerado por $w_1 = (0, 1, 1)$, $w_2 = (-1, 1, -1)$.
Como $\{(0, 1, 1), (-1, 1, -1)\}$ é uma base ortogonal de V, tem-se $p = \text{proj}_V b = \frac{w_1|b}{w_1|w_1}\ w_1 + \frac{w_2|b}{w_2|w_2}\ w_2 = \frac{4}{2}\ (0, 1, 1) + \frac{1}{3}\ (-1, 1, -1) = (-\frac{1}{3}, \frac{7}{3}, \frac{5}{3})$.

É pois importante identificar bases ortogonais dos subespaços vetoriais. O método que a seguir se descreve, chamado *processo de ortogonalização de Gram-Schmidt*, tem esta finalidade.
Seja $V \neq \{\vec{0}\}$ um subespaço vetorial de $\mathbb{R}^m$.

. Identifique-se $v_1 \neq \vec{0} \in V$. Se $\dim V > 1$, então

. identifique-se $u_2 \in V\backslash < \{v_1\} >$ e defina-se $v_2 := u_2 - \frac{v_1|u_2}{v_1|v_1}v_1$. (Claramente $v_2 \in V$ e $v_2|v_1 = u_2|v_1 - \frac{v_1|u_2}{v_1|v_1}v_1|v_1 = 0$.) Se $\dim V > 2$, então

. identifique-se $u_3 \in V\backslash < \{v_1, v_2\} >$ e defina-se $v_3 := u_3 - \frac{v_1|u_3}{v_1|v_1}v_1 - \frac{v_2|u_3}{v_2|v_2}v_2$. (Claramente $v_3 \in V$ e $v_3|v_1 = u_3|v_1 - \frac{v_1|u_3}{v_1|v_1}v_1|v_1 - \frac{v_2|u_3}{v_2|v_2}v_2|v_1 = 0$, $v_3|v_2 = u_3|v_2 - \frac{v_1|u_3}{v_1|v_1}v_1|v_2 - \frac{v_2|u_3}{v_2|v_2}v_2|v_2 = 0$.) Se $\dim V > 3$, então

. ...

O conjunto $\{v_1, v_2, \ldots, v_{\dim V}\}$ assim obtido é uma base ortogonal de V.

Exemplo 34 Determine uma base ortogonal de $\mathbb{R}^3$ que inclua o vetor $(1, 1, 1)$.

. $v_1 = (1, 1, 1)$.

. $u_2 = (1, 0, 0) \in \mathbb{R}^3\backslash < \{(1, 1, 1)\} >$.

$$v_2 := u_2 - \frac{v_1|u_2}{v_1|v_1}v_1 = (1, 0, 0) - \tfrac{1}{3}(1, 1, 1) = \tfrac{1}{3}(2, -1, -1) \rightsquigarrow (2, -1, -1).$$

. $u_3 = (1, -1, 1) \in \mathbb{R}^3\backslash < \{(1, 1, 1), (2, -1, -1)\} >$.

$$v_3 := u_3 - \frac{v_1|u_3}{v_1|v_1}v_1 - \frac{v_2|u_3}{v_2|v_2}v_2 = (1, -1, 1) - \tfrac{1}{3}(1, 1, 1) - \tfrac{2}{6}(2, -1, -1) = (0, -1, 1).$$

$\{(1, 1, 1), (2, -1, -1), (0, -1, 1)\}$ é uma base ortogonal de $\mathbb{R}^3$ que inclui o vetor $(1, 1, 1)$.

Exemplo 35 Determine uma base ortogonal de $\mathbb{R}^3$ que inclua o vetor $(1, 1, 1)$.

. $v_1 = (1, 1, 1)$.

. $u_2 = (1, 0, 0) \in \mathbb{R}^3\backslash < \{(1, 1, 1)\} >$.

$$v_2 := u_2 - \frac{v_1|u_2}{v_1|v_1}v_1 = (1, 0, 0) - \tfrac{1}{3}(1, 1, 1) = \tfrac{1}{3}(2, -1, -1) \rightsquigarrow (2, -1, -1).$$

. $u_3 = (1, -1, 1) \in \mathbb{R}^3\backslash < \{(1, 1, 1), (2, -1, -1)\} >$.

$$v_3 := u_3 - \frac{v_1|u_3}{v_1|v_1}v_1 - \frac{v_2|u_3}{v_2|v_2}v_2 = (1, -1, 1) - \tfrac{1}{3}(1, 1, 1) - \tfrac{2}{6}(2, -1, -1) = (0, -1, 1).$$

$\{(1, 1, 1), (2, -1, -1), (0, -1, 1)\}$ é uma base ortogonal de $\mathbb{R}^3$ que inclui o vetor $(1, 1, 1)$.

Exercícios 25

1. Determine a projeção do vetor $(2, 3)$ sobre o vetor $(3, 1)$.

2. Determine a projeção do vetor $(6, 5, 4)$ sobre a reta $< (1, -1, 3) >$.

3. Identifique o vetor do subespaço vetorial $< \{(1, 0, 1), (1, 1, 0)\} >$ a menor distância do vetor $(1, 2, 3)$.

4. Considere o vetor $b = (1, 1, 1)$ e os seguintes subespaços vetoriais de $\mathbb{R}^3$.

$$V =< \{(1, 0, 1), (0, 1, 1)\} > \quad \text{e} \quad U = \{(x, y, z) : x + y + z = 0\}.$$

   a) Determine a projeção ortogonal de b sobre o vetor (1,0,1).

   b) Determine as projeções ortogonais de b sobre V, U, $V^\perp$ e $U^\perp$.

   c) Calcule as distâncias de b a V e a U.

5. Determine a projeção do vetor $(0, 2, 5, -1)$ sobre o subespaço vetorial de $\mathbb{R}^4$ gerado pelos vetores $(1, 1, 0, 2)$ e $(-1, 0, 0, 1)$.

6. Considere o subespaço vetorial de $\mathbb{R}^4$

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

   e o vetor $v = (2, 1, 0, 1)$. Determine as projeções ortogonais de v sobre U e sobre complemento ortogonal de U.

7. Defina a matriz de projeção sobre o plano de equação $x + 2y + 3z = 0$.

8. Considere o vetor $w = (1, -2, 2, 2)$ e o subespaço $V =< \{(1, 2, 0, 0), (1, 0, 1, 1)\} >$.

   a) Defina a matriz de projeção sobre o subespaço V.

   b) Determine a projeção de w sobre V.

9. Verifique que $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ é a matriz de projeção sobre o subespaço vetorial $W = \{(x, y, z, t) : x = y,\ z = t\}$.

10. Sejam A uma matriz do tipo $m \times n$, com característica n e $P = A\,(A^\top A)^{-1}A^\top$ a matriz de projeção sobre $\mathcal{C}(A)$. Prove os seguintes resultados.

    a) $P^\top = P$.

    b) $P^2 = P$.

11. Considere os vetores $u = (1, -1, 0, 1)$, $v = (0, 1, 0, 1)$ e $b = (2, -1, 0, 1)$.

    a) Calcule o ângulo definido pelos vetores u e v.

    b) Determine a projeção ortogonal do vetor b sobre o subespaço de $\mathbb{R}^4$ gerado pelos vetores u e v.

12. Considere os vetores $a = (1, -1, 1)$, $b = (-1, 1, 2)$ e $c = (1, 1, 0)$.

    a) Mostre que o conjunto $\{a, b, c\}$ é uma base ortogonal de $\mathbb{R}^3$.

    b) Escreva o vetor $(0, 2, 4)$ como combinação linear dos vetores a, b e c. Interprete geometricamente os coeficientes da combinação linear.

13. a) Utilizando o processo de ortogonalização de Gram-Schmidt, determine uma base ortogonal de $\mathbb{R}^3$ que inclua o vetor $(1, 0, 1)$.

    b) Transforme a base obtida na alínea anterior numa base ortonormal de $\mathbb{R}^3$.

14. Seja $V = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4\ :\ x_1 + x_2 - x_3 + x_4 = 0\}$.

    a) Utilizando o processo de ortogona1ização de Gram-Schmidt determine uma base ortogonal de V.

b) Seja $b = (2, 1, 0, 1)$. Calcule a projeção de b sobre o subespaço V.

Capítulo 4

Introdução à programação linear

A programação linear trata de uma grande variedade de problemas *lineares* em que se pretende obter o maior proveito possível na utilização de recursos limitados. Apresenta-se uma breve introdução à programação linear, mostrando que é um assunto fortemente relacionado com os sistemas de equações lineares.

4.1 Exemplos

Considere o seguinte problema. Uma exploração agrícola dispõe de 80 hectares de terra para produzir tomate e trigo. Os recursos suscetíveis de limitar a produção das duas culturas são a terra (80 ha), a água para rega e o trabalho, de acordo com os valores indicados no quadro seguinte.

recursos	Necessidades (por ha)		Disponibilidade
	tomate	trigo	
Água (m^3)	8000	0	320000
Trabalho (DH)	40	20	2000

As receitas resultantes de cada hectare de tomate e de trigo são 300 e 200 Euros, respetivamente. Quais as áreas a destinar a cada uma das culturas de modo que a receita total seja máxima?

Para descrever matematicamente o problema, vamos usar as variáveis x, que indica o número de hectares que vão ser destinados à cultura do tomate e y, que indica o número de hectares destinados à cultura do trigo. Com estas variáveis o problema consiste em

$$\begin{array}{llcll}
\text{maximizar} & z = 300x + 200y & & & (4.1) \\
\text{sujeito a} & x + y & \leq & 80 & (4.2) \\
 & 8000x & \leq & 320000 & (4.3) \\
 & 40x + 20y & \leq & 2000 & (4.4) \\
 & x, y & \geq & 0 & (4.5)
\end{array}$$

As *desigualdades lineares* (4.2), (4.3) e (4.4) traduzem as limitações decorrentes dos recursos terra, água e trabalho, respetivamente. As *restrições de sinal* (4.5) invalidam opções matematicamente válidas, mas sem nexo (em que áreas negativas são atribuídas a uma das culturas). A *função objetivo* (4.1) estabelece o valor, em euros, da receita correspondente à opção x ha destinados à cultura do tomate e y ha à cultura do trigo.

O problema (4.1)-(4.5) é um caso particular do problema de programação linear.

A programação linear trata da maximização ou minimização de funções lineares, em que as variáveis satisfazem restrições lineares. O problema geral de programação linear consiste em determinar valores para as variáveis $x_1, x_2, \ldots, x_n$ que otimizam (maximizam ou minimizam) uma função linear

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

chamada *função objetivo*, sujeito a restrições lineares do tipo

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \geq, \leq \text{ ou } = b$$

e restrições de sinal das variáveis

$$x_j \geq, \leq \text{ ou } \gtrless 0 \quad (\gtrless \text{ significa maior, menor ou igual})$$

Muitos problemas podem ser formulados desta forma. Vejamos alguns exemplos.

Exemplos 36

1. Um criador de porcos pretende estabelecer com custo mínimo um concentrado composto que satisfaça certos requisitos nutricionais para a alimentação dos seus animais, de acordo com os dados fornecidos na seguinte tabela.

	Unidades contidas em cada Kg				
ingredientes	milho	farinhas de carne e osso	luzerna	valores min. diários	valores max. diários
hidratos carbono	90	20	40	200	
proteína	30	80	60	180	
fibra	15	0	30		70
vitaminas	10	20	60	150	270
aminoácidos	10	25	15	15	
custo (Euros/Kg)	0.30	0.25	0.18		

Para modelar esta situação utilizamos as variáveis x_1, x_2, x_3 que indicam as quantidades em Kgs de milho, farinhas de carne e osso e luzerna, respetivamente, que integram a mistura. Com essas variáveis o problema pode ser formulado linearmente do seguinte modo.

$$\text{minimizar} \qquad z = 0.30x_1 + 0.25x_2 + 0.18x_3$$

$$
\begin{array}{lrcl}
\text{sujeito a} & 90x_1 + 20x_2 + 40x_3 & \geq & 200 \\
& 30x_1 + 80x_2 + 60x_3 & \geq & 180 \\
& 15x_1 \qquad\quad + 30x_3 & \leq & 70 \\
& 10x_1 + 20x_2 + 60x_3 & \geq & 150 \\
& 10x_1 + 20x_2 + 60x_3 & \leq & 270 \\
& 10x_1 + 25x_2 + 15x_3 & \geq & 15 \\
& x_1, x_2, x_3 & \geq & 0
\end{array}
$$

2. A água que abastece três regiões no Alentejo, R1, R2 e R3, provém das barragens B1 e B2. Estima-se que as regiões R1, R2 e R3 necessitam anualmente, pelo menos, 10, 5 e 15 milhões de metros cúbicos de água, respetivamente. Sabe-se que no próximo ano as barragens B1 e B2 poderão fornecer até 14 e 16 milhões de metros cúbicos, respetivamente. O custo, em 10^4 Euros, do abastecimento de cada 10^6m^3 de água de cada barragem para cada uma das regiões, é indicado no quadro seguinte.

	R1	R2	R3
B1	14	10	9
B2	13	11	12

Qual é o plano abastecimento de água às três regiões a que corresponde o menor custo?

Se representarmos por x_{ij}, com $1 \leq i \leq 2$ e $1 \leq j \leq 3$, a quantidade de água, em milhões de metros cúbicos, que a barragem Bi fornece à região Rj, o problema pode ser formulado da seguinte forma.

$$
\text{minimizar} \qquad z = 14x_{11} + 10x_{12} + 9x_{13} + 13x_{21} + 11x_{22} + 12x_{23}
$$

$$\begin{array}{lrcl}
\text{sujeito a} & x_{11}+x_{21} & \geq & 10 \\
& x_{12}+x_{22} & \geq & 5 \\
& x_{13}+x_{23} & \geq & 15 \\
& x_{11}+x_{12}+x_{13} & \leq & 14 \\
& x_{21}+x_{22}+x_{23} & \leq & 16 \\
& x_{11},x_{12},x_{13},x_{21},x_{22},x_{23} & \geq & 0
\end{array}$$

4.2 Geometria da programação linear

Há uma terminologia própria da programação linear. Assim, chama-se *região admissível* ao conjunto dos pontos que satisfazem todas as restrições. Um ponto da região admissível chama-se *solução admissível*. É *solução ótima* a que corresponde ao maior ou menor valor da função objetivo, consoante o problema é de maximização ou de minimização.

Quando só há duas variáveis, o problema de programação linear pode ser facilmente solucionado recorrendo a argumentos geométricos. Consideremos o problema (4.1)-(4.5), que pode equivalentemente ser escrito da seguinte forma.

$$\text{maximizar} \quad z = 300x + 200y \tag{4.6}$$

$$\begin{array}{lrcl}
\text{sujeito a} & x+y & \leq & 80 \qquad (4.7) \\
& x & \leq & 40 \qquad (4.8) \\
& 2x+y & \leq & 100 \qquad (4.9) \\
& x,y & \geq & 0 \qquad (4.10)
\end{array}$$

O conjunto das soluções admissíveis é o *poliedro* $\mathcal{R}$ representado na Figura 4.1. (Um poliedro é a intersecção de um número finito de inequações lineares.)

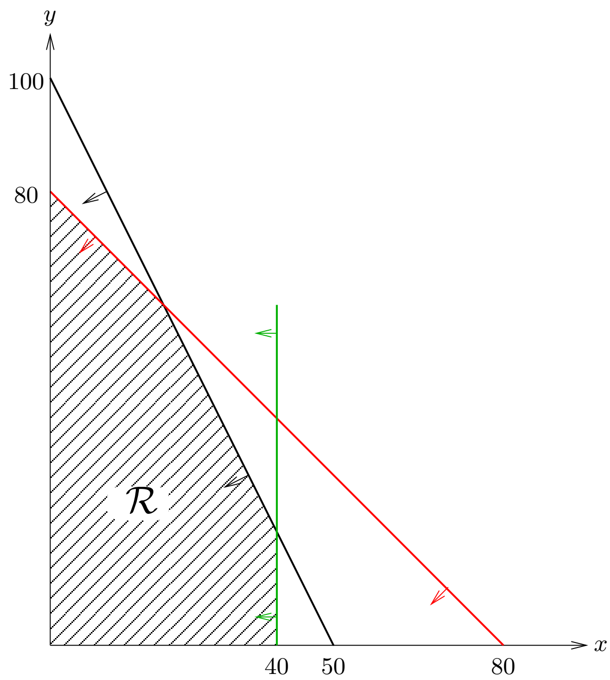


Figura 4.1: Região admissível definida pelas restrições (4.7), (4.8), (**4.9**) e (4.10).

Para identificar uma solução a que corresponde o maior valor de z, comecemos por notar que se fixarmos o valor de z, por exemplo fazendo $z = 6000$, obtemos

$$300x + 200y = 6000$$

que é uma equação da reta r representada na Figura 4.2. Todas as soluções admissíveis que estão nessa reta têm portanto valor da função objetivo igual a 6000. Estão nesta situação as duas opções admissíveis extremas que consistem em destinar 20 ha do terreno para a cultura do tomate e zero para o trigo ($x = 20, y = 0$), e a que reserva 30 ha para trigo e zero para tomate ($x = 0, y = 30$). A ambas (e todas as opções do segmento que as une) corresponde uma receita total de 6000 Euros.

Se queremos um lucro maior, igualamos z ao valor (maior do que 6000) desejado e procuramos soluções admissíveis que satisfaçam essa relação. Por exemplo, se pretendemos

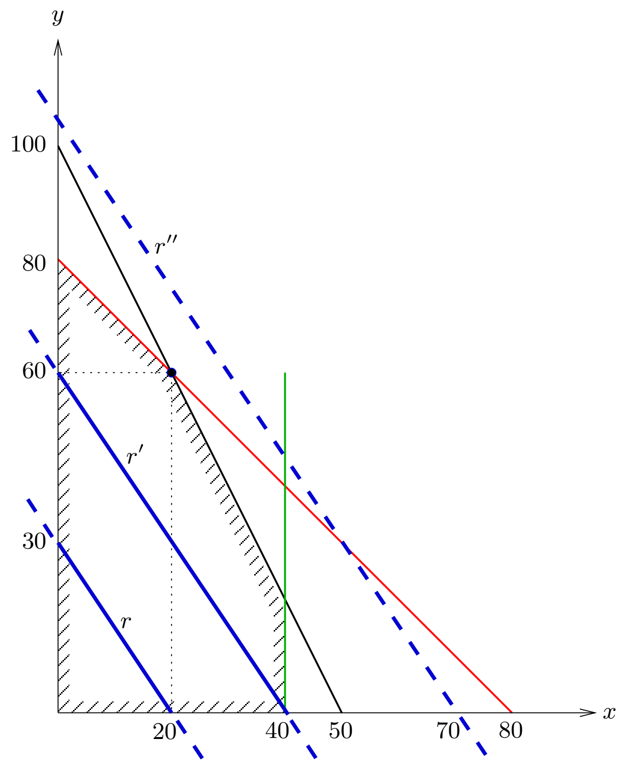


Figura 4.2: Representação das soluções admissíveis do problema (4.6)-(4.10) com valores da função objetivo iguais a 6000 (r) e a 12000 (r').

obter um lucro de 12000 Euros, procuramos soluções admissíveis que satisfaçam a equação

$$300x + 200y = 12000.$$

A equação define a reta r' da Figura 4.2 e as soluções admissíveis que a satisfazem estão no segmento cujos extremos são os pontos $(40, 0)$, que consiste a atribuir 40 ha à exploração do tomate e zero à cultura do trigo e $(0, 60)$, que destina ao trigo 60 ha e zero ao tomate. Se formos demasiado ambiciosos em relação às receitas totais, corremos o risco de não exitirem soluções admissíveis correpondentes aos valores pretendidos. Por exemplo, nenhum dos pontos da reta de equação

$$300x + 200y = 21000,$$

representada por r'' na Figura 4.2, é admissível. Por outras palavras, não é viável atingir uma receita total de 21000 Euros com as presentes limitações dos recursos que condicionam a produção agrícola.

Do que ficou dito pode concluir-se que o valor z^* da receita total máxima é tal que a reta de equação

$$300x + 200y = z^*$$

intersecta a região admissível, mas para todo o $\delta > 0$, a reta de equação

$$300x + 200y = z^* + \delta$$

já não inclui qualquer solução admissível. É evidente, e este é o facto que se pretende destacar, que o valor desejado z^* ocorre num *vértice* da região admissível.

No exemplo que estamos a analisar o maior valor de (4.6) ocorre no vértice $(20, 60)$, que corresponde a reservar 20 ha de terreno para o tomate e 60 ha para o trigo, a que está associado o lucro máximo de 18000 Euros. Note que esta solução utiliza na totalidade os recursos disponíveis respeitantes à dimensão da exploração (restrição (4.7)) e trabalho (restrição (4.9)). Diz-se então que as restrições (4.7) e (4.9) estão *saturadas*. São estas as restrições que limitam o valor da receita total. Já a quantidade de água disponível não é toda utilizada na opção ótima.

Neste caso $(20, 60)$ é a única solução ótima. Note que, se em lugar de (4.6), a função objetivo fosse

$$z = 400x + 200y$$

então haveria mais do que uma alternativa ótima. De facto, os dois vértices $(20, 60)$ e $(40, 20)$, bem como todos os pontos do segmento que os ligam, teriam valores da função objetivo iguais a 20000, que é o valor máximo.

Julgamos que o exemplo que acabámos de tratar é suficientemente esclarecedor para justificar que o facto da função objetivo ser linear garante que a otimalidade ocorre em pelo menos um vértice da região admissível. Assim, uma solução ótima pode ser encontrada enumerando os vértices do poliedro, calculando para cada um deles o correspondente valor da função objetivo e selecionar para solução ótima um vértice a que corresponda o maior ou menor (consoante se trate de maximizar ou minimizar) valor.

Exercício 26 Considere o seguinte problema de problema de programação linear.

$$
\begin{array}{lrcl}
\text{maximizar} & z = x_1 + 2x_2 & & \\
\text{sujeito a} & x_1 + 3x_2 & \leq & 24 \\
 & x_1 + \ \ x_2 & \leq & 10 \\
 & x_1 \qquad & \leq & 8 \\
 & x_1, x_2 & \geq & 0
\end{array}
$$

a) Represente geometricamente a região admissível.

b) Indique uma solução ótima, o valor da função objetivo nesse ponto e identifique as restrições *saturadas* (satisfeitas com igualdades).

c) Indique o maior intervalo de variação do membro direito da terceira restrição que mantém ótima a solução que refiriu na alínea b).

d) Dê exemplo de uma outra função objetivo relativamente à qual se mantém ótima a solução que indicou na alínea b).

Argumentos geométricos análogos aos utilizados para o caso de duas variáveis também validam que o procedimento anterior possa ser utilizado para resolver problemas de programação linear que envolvem três variáveis. Consideremos o problema seguinte.

$$
\begin{array}{lrcl}
\text{maximizar} & z = x_1 + 2x_3 & & \qquad (4.11) \\
\text{sujeito a} & x_1 + x_2 + x_3 & \leq & 4 \qquad (4.12) \\
 & x_1 \qquad\qquad & \leq & 2 \qquad (4.13) \\
 & x_3 & \leq & 3 \qquad (4.14) \\
 & 3x_2 + x_3 & \leq & 6 \qquad (4.15) \\
 & x_1, x_2, x_3 & \geq & 0 \qquad (4.16)
\end{array}
$$

A região admissível é o poliedro representado na Figura 4.3.

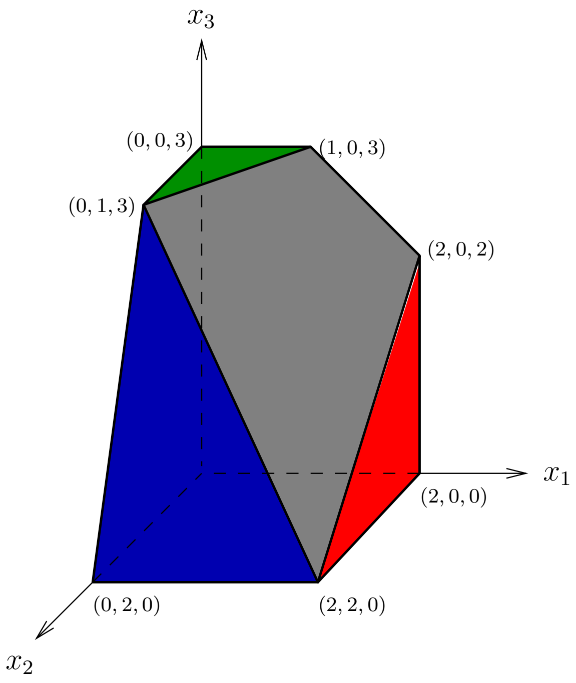


Figura 4.3: Região admissível correspondente às restrições (**4.12**), (4.13), (4.14), (4.15), e (4.16).

Os vértices do poliedro e os correspondentes valores da função objetivo estão indicados na Tabela 4.1.

De acordo com o procedimento anterior $x_1 = 1, x_2 = 0, x_3 = 3$ é solução ótima, pois é o vértice a que corresponde o maior valor da função objetivo.

De facto, quando igualamos a função objetivo a um certo valor v, obtemos a equação

$$x_1 + 0x_2 + 2x_3 = v$$

de um plano ortogonal ao vetor $(1, 0, 1)$. Na Figura 4.4 estão representadas as intersecções do poliedro da admissibilidade com três desses planos, resultantes de diferentes concretizações de v. (A intensidade da intersecção é tanto mais acentuada quanto maior for o valor de v.) Os pontos do mesmo plano têm igual valor da função objetivo. Os que estão

vértice	correspondente valor da função objetivo $x_1 + 2x_3$
$(0, 0, 0)$	0
$(0, 2, 0)$	0
$(2, 2, 0)$	2
$(2, 0, 0)$	2
$(2, 0, 2)$	6
$(1, 0, 3)$	7
$(0, 0, 3)$	6
$(0, 1, 3)$	6

Tabela 4.1: Vértices do poliedro e respetivos valores da função objetivo do problema (4.11)-(4.16).

no plano mais escuro têm maiores valores. Se deslocarmos este plano no sentido definido pelo vetor representado na figura, estamos a obter pontos com valores da função objetivo cada vez maiores. O plano 'mais deslocado' no sentido desse vetor, e que inclui soluções admissíveis, intersecta o poliedro apenas no ponto $(1, 0, 3)$. Pode então concluir-se que $(1, 0, 3)$ é uma solução admissível com o valor máximo, e portanto que a otimalidade ocorre num vértice.

Mesmo quando há mais do que três variáveis envolvidas, a otimalidade pode ser procurada nos vértices do poliedro da admissibilidade. De facto, tem-se o seguinte resultado.

Teorema 4.1 *Se $\mathcal{P}$ é um poliedro não vazio e limitado de $\mathbb{R}^n$ e $c = (c_1, c_2, \ldots, c_n)$ é um vetor de $\mathbb{R}^n$,*

1. *Existe um vértice de $\mathcal{P}$ que é solução ótima do problema de programação linear*

$$\begin{array}{ll} \textit{maximizar (minimizar)} & z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \textit{sujeito a} & x = (x_1, x_2, \ldots, x_n) \in \mathcal{P} \end{array}$$

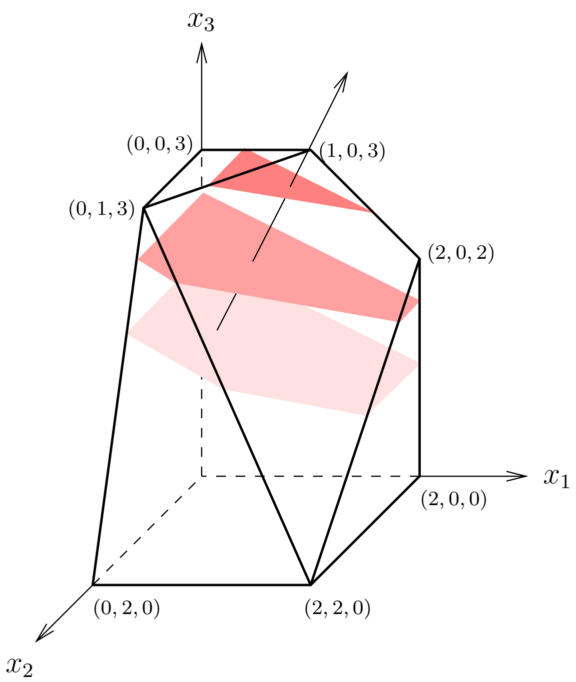


Figura 4.4: Intersecções do poliedro da Figura 4.3 com três planos ortogonais ao vetor $(1,0,1)$.

2. *Se os vértices* $v_1, v_2, \ldots, v_k$ *são soluções ótimas, então* $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k$*, com* $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$ *e* $\lambda_1, \lambda_2, \ldots, \lambda_k \geq 0$*, também é solução ótima.*

4.3 Forma *standard*

O teorema anterior remete-nos para a identificação dos vértices de um poliedro. Em seguida vamos ver, o que não é assim tão surpreendente, que os vértices podem ser determinados resolvendo sistemas de equações lineares. Note que, por exemplo, o vértice $(20,60)$ do poliedro representado na Figura 4.1 é a intersecção das retas definidas pelas equações que se obtêm substituindo o símbolo "$\leq$"por "$=$"nas restrições (4.7) e (4.9). Por outras palavras, $(20, 60)$ é a solução do sistema linear

$$\begin{cases} x + y = 80 \\ 2x + y = 100 \end{cases}.$$

Também o vértice $(40, 0)$ é definido pelas equações

$$\begin{cases} x \quad = 40 \\ \quad y = 0 \end{cases}$$

correspondentes à inequação (4.8) e a uma das restrições (4.10) de não negatividade das variáveis.

O vértice $(1, 0, 3)$ do poliedro da Figura 4.3 é a intersecção do plano cinzento, com o verde e com o plano definido pelos vetores x_1 e x_3, i.e., é a solução do sistema

$$\begin{cases} x_1 + x_2 + x_3 = 4 \\ x_3 = 3 \\ x_2 = 0 \end{cases}.$$

Antes de explicitar a correspondência entre vértices e soluções de sistemas lineares, vamos mostrar que a região admissível de qualquer problema de programação linear pode ser descrita por um sistema de equações lineares com variáveis não negativas.

Comecemos por ver que é possível condicionar todas as variáveis a tomar valores não negativos.

Se uma dada variável x_j está limitada a assumir valores não positivos, i.e., se na formulação está presente a restrição $x_j \leq 0$, então tudo o que há a fazer é substituir na formulação a variável x_j por $-\bar{x}_j$ e exigir que $\bar{x}_j \geq 0$. Desta forma, qualquer que seja o valor $p > 0$ atribuído a $\bar{x}_j$, equivale a fazer $x_j = -p < 0$.

Se $x_j \gtrless 0$, i.e., a variável x_j não tem restrição de sinal, então podemos substituir na formulação a variável x_j pela diferença $x_j^1 - x_j^2$ e impor que $x_j^1, x_j^2 \geq 0$. Assim, x_j assume um valor positivo se o valor atribuído a x_j^1 for maior do que o atribuído a x_j^2. Caso contrário, o valor de x_j será negativo ou nulo.

Assim, qualquer problema pode ser transformado de forma a que todas as variáveis sejam não negativas.

Note também que uma inequação linear do tipo

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

pode ser substituída pelas condições

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n + x_{n+1} = b \ \text{ e } \ x_{n+1} \geq 0,$$

e que a desigualdade

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

é equivalente a

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n - x_{n+1} = b \ \text{ e } \ x_{n+1} \geq 0.$$

Em ambos os casos as desigualdades foram tranformadas em equações pela introdução da variável não negativa x_{n+1}. A essa variável adicional dá-se o nome de *variável de folga*.

Em conclusão, todo o problema de programação linear pode ser apresentado na forma de minimização ou maximização de uma função linear com variáveis não negativas, que satisfazem um sistema de equações lineares. Um problema escrito neste formato diz-se na forma *standard*.

Uma formulação *standard* do problema (4.11)-(4.16) é

$$\begin{array}{ll} \text{maximizar} & x_1 + 2x_3 \\ \text{sujeito a} & \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 3 \\ 6 \end{bmatrix} \end{array} \tag{4.17}$$

$$x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0, \tag{4.18}$$

em que x_4, x_5, x_6, x_7 são as variáveis de folga.

É óbvia a correspondência entre soluções admissíveis de um dado problema de programação linear e do correspondente problema na forma *standard*. A solução admissível $x_1 = 1, x_2 = 0, x_3 = 2, x_4 = 1, x_5 = 1, x_6 = 1, x_7 = 1$ de (4.17),(4.18) corresponde à solução admissível $x_1 = 1, x_2 = 0, x_3 = 2$ de (4.11)-(4.16). A solução $x_1 = 0, x_2 = 1, x_3 = 0, x_4 = 3, x_5 = 2, x_6 = 3, x_7 = 3$ corresponde a $x_1 = 0, x_2 = 1, x_3 = 0$.

De uma forma geral tem-se o seguinte. Sejam P um problema de programação linear com n variáveis não negativas $x_1, x_2, \ldots, x_n$ e

$$Ax = b, \text{ com } x = (x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+k}) \tag{4.19}$$

o sistema de equações do problema *standard* associado. À solução não negativa $(\bar{x}_1, \ldots, \bar{x}_n, \bar{x}_{n+1}, \ldots, \bar{x}_{n+k})$ de (4.19) corresponde a solução $(\bar{x}_1, \ldots, \bar{x}_n)$ do problema P.

Vamos admitir que a característica da matriz A é igual ao número de linhas, que designamos por m. Seja $\beta = \{A_{j_1}, A_{j_2}, \ldots, A_{j_m}\}$ um conjunto de m colunas de A linearmente independente, i.e., uma base do espaço das colunas de A. Chama-se *solução básica* à solução que se obtem resolvendo o sistema apenas com as variáveis associadas às colunas de β, e igualando a zero as restantes variáveis.

As colunas 4, 5, 6 e 7 da matriz dos coeficientes do sistema (4.17) são linearmente independentes. A solução básica corresponde a essa escolha de colunas é $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 4, x_5 = 2, x_6 = 3, x_7 = 6$. Também as colunas 2, 5, 6, 7 formam um conjunto de vetores linearmente independente. A correspondente solução básica é $x_1 = 0, x_2 = 4, x_3 = 0, x_4 = 0, x_5 = 2, x_6 = 3, x_7 = -6$.

Uma *solução básica* é *admissível* (sba) se não tem componentes negativas. A solução $(0, 0, 0, 4, 2, 3, 6)$ de (4.17) é básica admissível, a solução básica $(0, 4, 0, 0, 2, 3, -6)$ é não admissível.

O resultado seguinte estabelece a correspondência entre as sba_s e os vértices do poliedro do problema original.

Vértice do poliedro do problema (4.11)-(4.16)	correspondente sba de (4.17)
(0,0,0)	(0,0,0,4,2,3,6)
(0,2,0)	(0,2,0,2,2,3,0)
(2,2,0)	(2,2,0,0,0,3,0)
(2,0,0)	(2,0,0,2,0,3,6)
(2,0,2)	(2,0,2,0,0,1,4)
(1,0,3)	(1,0,3,0,1,0,3)
(0,0,3)	(0,0,3,1,2,0,3)
(0,1,3)	(0,1,3,0,2,0,0)

Tabela 4.2: Vértices do poliedro de admissibilidade do problema (4.11)-(4.16) e correspondentes sba_s de (4.17).

Teorema 4.2 *Sejam $\mathcal{P}$ o poliedro da admissibilidade de um problema de programação linear com variáveis não negativas e $Ax = b$ o sistema de equações lineares do problema na forma* standard *que lhe está associado. Os vértices de $\mathcal{P}$ e as sba_s de $Ax = b$ estão em correspondência biunívoca.*

Na Tabela 4.2 listam-se os vértices do poliedro da Figura 4.3 e as correspondentes sba_s de (4.17).

Assim, pode obter-se uma solução ótima de um problema de programação linear resolvendo um número finito de sistemas de equações lineares. O método do *simplex* é geralmente bastante eficaz na resolução de problemas de programação linear, ao selecionar criteriosamente os diferentes sistemas de equações a resolver. Estão disponíveis para utilização implementações do método do *simplex* integradas em programas informáticos para cálculo matemático (ver por exemplo o pacote *simplex* do programa Maxima ou o pacote *linprog* do programa R).

Exercícios 27

1. Uma câmara municipal pretende rentabilizar um parque de 100 ha destinando-o a área florestal, parque de campismo e reserva de caça. Para a manutenção do parque dispõe anualmente de uma verba de 30000 Euros e de 20000 horas de trabalho. O quadro seguinte indica o capital e horas de trabalho necessários à manutenção anual de cada ha, consoante o tipo de ocupação de solo.

	capital (Euros)	horas de trabalho
floresta	100	100
caça	300	150
campismo	400	500

Prevê-se um lucro anual de 40, 60 e 80 Euros por cada ha de terreno destinado a área florestal, parque de campismo e reserva de caça, respetivamente. Pretende determinar-se o número de ha a destinar a cada tipo de ocupação de solo de forma a maximizar o lucro.

a) Formule linearmente o problema atribuindo significado às variáveis utilizadas.

b) Converta à forma *standard* a formulação anterior e indique uma solução básica admissível.

c) Determine uma solução que proporcione o maior lucro quando 40 ha do terreno são destinados a reserva de caça.

2. Formule e resolva o seguinte problema.

Um distribuidor de cafés vai misturar numa certa proporção os grãos provenientes do Brasil, Quénia e Jamaica, que dispõe em armazém, para fazer dois lotes de café A e B. A composição e o preço de venda de cada um dos lotes, assim como a quantidade existente em armazém de cada um dos tipos de café estão indicados no quadro seguinte.

	lote A	lote B	quant. disponível (kg)
Brasil	0.25	0.25	100
Quénia	0.75	0.25	150
Jamaica	0.0	0.5	175
preço de venda (Euros/Kg)	3.5	5.0	

Sabendo que todo o café será vendido, pretende determinar-se a quantidade de cada um dos lotes a que corresponde a maior receita bruta.

3. Uma fábrica tem que reduzir a emissão dos seus 3 principais poluentes atmosféricos: as partículas, os óxidos sulfúricos e os hidrocarbonetos, em pelo menos 72, 50 e 24 milhares de quilos por ano, respetivamente. Para este efeito a fábrica vai modificar a chaminé, aumentando a altura e/ou a área dos filtros. Estas modificações permitem reduzir a emissão anual dos poluentes nos valores (em milhares de quilos) indicados na tabela seguinte.

	Aumentar 1 m altura da chaminé	Aumentar 1 m^2 área dos filtros
Partículas	9	18
Óxidos sulfúricos	10	10
Hidrocarbonetos	12	4

Os custos de aumentar 1 m a altura e 1 m^2 a área dos filtros da chaminé são, respetivamente, 10 e 7 mil euros. A fábrica pretende determinar os valores dos aumentos da altura e da área dos filtros de modo a atingir o objetivo proposto com o menor custo possível.

a) Formule linearmente o problema, atribuindo significado às variáveis.

b) Represente graficamente a região admissível.

c) Determine a solução ótima e a correspondente solução básica admissível. Qual é o custo que corresponde a esta solução?

4. Um avião de combate a incêndios florestais pode transportar dois tipos de produtos, P1 e P2. Uma tonelada de P1 ocupa 0.5 m^3, permite combater uma área de incêndio de 1.5 ha e custa 2000 Euros. Uma tonelada de P2 ocupa 2 m^3, permite combater uma área de 4 ha e custa 3000 Euros. O peso e espaço reservados para o transporte desses produtos não pode ultrapassar os 1.5 toneladas e 1.0 m^3. Pretende determinar-se a quantidade a transportar de cada produto de modo a combater incêndios numa área de pelo menos 2.5 ha e minimizando os custos.

   a) Formule linearmente o problema, indicando os signicado das variáveis intervenientes.

   b) Mostre que 1 tonelada de P1 e 0.25 toneladas de P2 é uma solução admissível e determine a área de incêndio que esta opção permite combater.

5. Um estabelecimento comercial pretende obter o máximo lucro disponibilizando 150 m^2 para armazenar, durante 3 meses, materiais dos tipos A, B, C e D. O processo de armazenagem terá que decorrer em não mais do que 10 horas e o compromisso de armazenar pelo menos 2 toneladas do material A terá que ser respeitado. Cada tonelada de A, B, C e D requer para ser armazenado 1, 4, 1 e 2 horas e ocupa 15, 16, 20 e 30 m^2, respetivamente. São cobrados 200, 300, 400 e 700 Euros, respetivamente, por cada tonelada de A, B, C e D.

   a) Formule o problema em termos de Programação linear, atribuindo significado às variáveis utilizadas.

   b) Converta à forma *standard* a formulação anterior e atribua significado às variáveis de folga.

c) Mostre que é admissível a opção que consiste em armazenar 2 toneladas de A, 0 de B, 3 de C e 2 de D, mas não corresponde a um vértice da região admissível.

6. Uma empresa de distribuição foi encarregue de abastecer 3 clientes com uma mercadoria existente nos armazéns A e B. O armazém A pode disponibilizar até 60 toneladas (t) dessa mercadoria e o armazém B até 30 t. O cliente 1 requereu exatamente 20 t. Os clientes 2 e 3 estão dispostos a receber qualquer quantidade da mercadoria, mas a empresa comprometeu-se apenas com o cliente 2 a fornecer-lhe pelo menos 50 t.

   A tabela seguinte indica o lucro (em dezenas de euros) resultante da distribuição de uma tonelada de mercadoria de cada armazém para cada um dos clientes.

	Cliente		
Armazém	1	2	3
A	8	5	7
B	6	4	10

   A empresa pretende determinar a quantidade de mercadoria a transportar de cada armazém para cada cliente de modo a obter o maior lucro.

   a) Formule o problema em termos de Programação linear, atribuindo significado às variáveis.

   b) Verifique que é admissível a opção descrita na tabela seguinte

	Cliente		
Armazém	1	2	3
A	20	40	0
B	0	10	20

   Qual é o lucro resultante desta opção?

c) Converta à forma *standard* a formulação anterior.

d) Mostre que a opção da alínea b) corresponde a um vértice da região admissível.

7. Uma empresa decidiu iniciar a produção dos produtos P_1 e P_2, dispondo para isso de mão-de-obra equivalente a 80 horas semanais. Semanalmente, cada tonelada de P_1 e P_2 dá um lucro de 12 e 8 euros e requer 5 e 2 horas de mão-de-obra, respetivamente. Sabe-se que a procura semanal do produto P_1 é não limitada, mas a de P_2 não ultrapassa as 30 toneladas. A empresa pretende determinar a quantidade a produzir semanalmente de cada produto, de forma a obter o lucro máximo.

    a) Descreva o problema de forma linear, atribuindo significado às variáveis utilizadas.

    b) Represente graficamente a região admissível.

    c) Identifique uma solução ótima e a correspondente solução básica admissível.

    d) Determine os valores que poderá assumir o lucro resultante da venda de cada tonelada de produto P_1 de forma a manter ótima a solução determinada na alínea anterior.

8. Considere o problema de programação linear seguinte.

$$\text{maximizar} \;\; 2x_1 + x_2 - x_3 + 3x_4$$

$$\begin{array}{rrcl}
\text{com } x \in \mathcal{P} = \{(x_1,x_2,x_3,x_4) : & x_2 - 2x_3 + x_4 & \geq & 3 \\
& x_1 \quad\; - 2x_3 + x_4 & \geq & 2 \\
& x_1 \quad\; + x_3 \quad\quad & \leq & 3 \\
& x_1 + x_2 - 2x_3 + x_4 & = & 5 \\
& x_1, x_2, x_3, x_4 & \geq & 0\}.
\end{array}$$

    a) Estabeleça as restrições lineares que definem a região admissível $\mathcal{F} \subset \mathbb{R}^7$ do correspondente problema linear na forma *standard*.

b) Verifique que $v = (2, 3, 0, 0)$ é vértice de $\mathcal{P}$ e indique o valor da função objetivo em v.

9. Considere o problema

$$\begin{array}{rrcr} \text{maximizar} & 20x_1 + 30x_2 & & \\ \text{sujeito a} & x_1 + 2x_2 & \leq & 120 \\ & x_1 \qquad & \leq & 60 \\ & x_2 & \leq & 50 \\ & x_1, x_2 & \geq & 0 \end{array}$$

a) Represente graficamente a região admissível e as soluções admissíveis a que correspondem valores da função objetivo iguais a 600.

b) Indique uma solução ótima e a correspondente solução básica admissível.

c) Se os coeficientes da função objetivo coincidissem e fossem positivos, quais seriam as soluções ótimas?

Capítulo 5

Determinantes

Vamos associar a cada matriz quadrada um valor que se define da seguinte forma.

Definição 25 Sejam A uma matriz quadrada de ordem n e A' a matriz em escada que se obtem de A por aplicação da fase descendente do método de eliminação de Gauss, utilizando exclusivamente as operações elementares de troca de linhas e substituição de uma linha por soma desta com um múltiplo de outra linha. Chama-se *determinante* de A e representa-se por $\det A$ ou $|A|$, o valor

$$\det A = |A| = \delta a'_{11} a'_{22} \cdots a'_{nn},$$

em que $a'_{11}, a'_{22}, \ldots, a'_{nn}$ são os elementos da diagonal principal da matriz A' e

$$\delta = \begin{cases} 1 & \text{se é par o n}^{\text{o}} \text{ de trocas de linhas efetuadas no processo } A \to \cdots A', \\ -1 & \text{caso contrário.} \end{cases}$$

Exemplos 37

1. $A = \begin{bmatrix} 1 & 2 \\ -1 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} = A'$ e portanto $\det A = 1 \times (-3) = -3$.

2. $A = \begin{bmatrix} 0 & 5 \\ -3 & 4 \end{bmatrix} \to \begin{bmatrix} -3 & 4 \\ 0 & 5 \end{bmatrix} = A'$ e portanto $\det A = -1 \times (-3) \times 5 = 15$.

3. De uma forma geral, tem-se $\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$.

4. $A = \begin{bmatrix} 0 & 5 & -1 \\ -3 & 4 & 2 \\ -1 & 1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 1 & 3 \\ -3 & 4 & 2 \\ 0 & 5 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 1 & 3 \\ 0 & 1 & -7 \\ 0 & 5 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 1 & 3 \\ 0 & 1 & -7 \\ 0 & 0 & 34 \end{bmatrix} = A'$

   e portanto $\det A = -1 \times (-1) \times 1 \times 34 = 34$.

5. $A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 4 & 6 \\ -1 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 4 \\ 0 & -2 & -2 \\ 0 & 4 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 4 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{bmatrix} = A'$ e portanto $\det A =$ $1 \times (-2) \times 0 = 0$.

Exercícios 28 Prove os seguintes resultados.

1. O determinante de uma matriz triangular é o produto dos elementos da diagonal principal.

2. Uma matriz com uma linha ou uma coluna de zeros tem determinante igual a zero.

3. É nulo o determinante de uma matriz com linhas proporcionais.

O determinante satisfaz a seguinte propriedade.

Proposição 5.1 *Se A e B são matrizes quadradas da mesma ordem, tem-se* $\det(AB) = \det A \det B$, *i.e., o determinante do produto de matrizes é igual ao produto dos determinantes.*

É claro que não poderá haver grandes expectativas relativamente à quantidade de informação que o determinante contem da matriz. De facto, não é razoável admitir que um único valor possa reter muito conhecimento sobre os n^2 elementos de uma matriz de ordem n. No entanto, o determinante permite caracterizar a invertibilidade de matrizes.

Proposição 5.2 *Uma matriz quadrada A é invertível sse* $\det A \neq 0$. *Se a matriz A é invertível, então* $\det A^{-1} = \frac{1}{\det A}$.

Demonstração: É sabido que uma matriz quadrada A é invertível sse todas as colunas de A', a matriz em escada obtida aplicando a A a fase descendente do método de Gauss, têm *pivots*. Como os *pivots* são os elementos não nulos da diagonal principal de A' e $\det A$ é, a menos do sinal, o produto dos elementos da diagonal principal de A', tem-se $\det A \neq 0$ sse A é invertível.
Se A é invertível, $\det(AA^{-1}) = 1 = \det A \det A^{-1}$, donde se conclui que $\det A^{-1} = \frac{1}{\det A}$.
□

Vamos agora apresentar uma forma alternativa de calcular o determinante. Para isso precisamos da seguinte definição.

Definição 26 Chama-se *complemento algébrico* ou *co-fator* do elemento (i, j) da matriz A e representa-se por Δ_{ij} o valor $\Delta_{ij} = (-1)^{i+j} A_{ij}$, em que A_{ij} é o determinante da matriz que se obtem de A eliminando a linha i e a coluna j.

Exemplo 38 Se $A = \begin{bmatrix} 2 & 4 & -2 \\ 1 & 3 & 0 \\ -1 & 2 & 5 \end{bmatrix}$, $\Delta_{11} = (-1)^2 \det \begin{bmatrix} 3 & 0 \\ 2 & 5 \end{bmatrix} = 15$, $\Delta_{21} = (-1)^{2+1} \det \begin{bmatrix} 4 & -2 \\ 2 & 5 \end{bmatrix} = (-1) \times 24 = -24$.

Teorema 5.3 *(Teorema de Laplace) Sejam i e j, respetivamente, uma linha e uma coluna arbitrárias da matriz A de ordem n. Tem-se*

$$\det A = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \cdots + a_{in}\Delta_{in} = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \cdots + a_{nj}\Delta_{nj}.$$

Exemplos 39

1. $\det \begin{bmatrix} 2 & 3 & -1 \\ 4 & 5 & 7 \\ 0 & -2 & 1 \end{bmatrix} = 2 \det \begin{bmatrix} 5 & 7 \\ -2 & 1 \end{bmatrix} - 3 \det \begin{bmatrix} 4 & 7 \\ 0 & 1 \end{bmatrix} - 1 \det \begin{bmatrix} 4 & 5 \\ 0 & -2 \end{bmatrix} = 2 \times 19 -$

$3 \times 4 - (-8) = 34.$

2. $\det \begin{bmatrix} 0 & 5 & 3 & 2 \\ 1 & 4 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} = -2 \det \begin{bmatrix} 0 & 5 & 2 \\ 1 & 4 & 0 \\ 2 & 3 & 0 \end{bmatrix} = -2 \times 2 \det \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} = -2 \times 2 \times (-5) = 20.$

Terminamos a matéria sobre determinantes com uma curiosa aplicação, que nos vai permitir obter de forma expedita um vetor que é ortogonal a cada um de dois vetores dados de $\mathbb{R}^3$.

Definição 27 Sejam $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ vetores de $\mathbb{R}^3$. Chama-se *produto externo* de x e y e representa-se por $x \times y$, o vetor de $\mathbb{R}^3$

$$x \times y = (x_2 y_3 - x_3 y_2, -x_1 y_3 + x_3 y_1, x_1 y_2 - x_2 y_1).$$

O vetor $x \times y$ pode ser memorizado "aplicando" da seguinte forma o Teorema de Laplace à "matriz" $\begin{bmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$, em que $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$. Tem-se pois,

$$x \times y = \text{“}\det \begin{bmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}\text{”} = \det \begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \det \begin{bmatrix} x_1 & x_3 \\ y_1 & y_3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Exemplo 40 $(1,-2,0) \times (1,0,1) = \text{“}\det \begin{bmatrix} e_1 & e_2 & e_3 \\ 1 & -2 & 0 \\ 1 & 0 & 1 \end{bmatrix}\text{”} = (-2,-1,2).$

Para mostrar que o vetor produto externo $x \times y$ é ortogonal a x e a y, consideremos a seguinte definição.

Definição 28 Sejam $z = (z_1, z_2, z_3)$, $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ vetores de $\mathbb{R}^3$. Chama-se *produto misto* de z, x e y ao produto interno de z por $x \times y$, i.e.,

$$z|x\times y = z_1 \det \begin{bmatrix} x_2 & x_3 \\ y_2 & y_3 \end{bmatrix} - z_2 \det \begin{bmatrix} x_1 & x_3 \\ y_1 & y_3 \end{bmatrix} + z_3 \det \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} = \det \begin{bmatrix} z_1 & z_2 & z_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}.$$

Tem-se então o seguinte resultado.

Proposição 5.4 *Sejam $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ vetores de $\mathbb{R}^3$. O* produto externo *de x e y é um vetor ortogonal a x e a y.*

Demonstração: $x|x \times y = \det \begin{bmatrix} x_1 & x_2 & x_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} = 0$, pois é o determinante de uma matriz com duas linhas iguais e portanto $x \perp x \times y$.
O mesmo raciocínio permite concluir que $y|x \times y = 0$, i.e., $y \perp x \times y$. $\square$

Assim, se $\{x, y\}$ é linearmente independente, $x \times y$ é ortogonal ao plano gerado por x e y (ver a Figura 5).

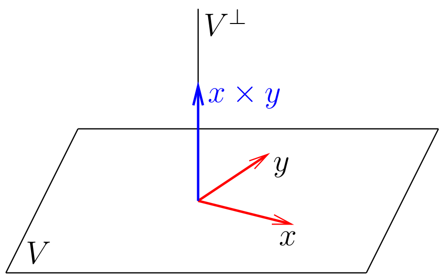


Figura 5.1: O vetor produto externo de dois vetores x e y que geram um plano V de $\mathbb{R}^3$

A norma do vetor produto externo é dada pelo seguinte resultado.

Proposição 5.5 *Se x e y vetores de $\mathbb{R}^3$, $\|x \times y\| = \|x\|\|y\|| \sin \theta|$.*

Demonstração: Por cálculo algébrico não há dificuldade em estabelecer que $\|x \times y\|^2 = \|x\|^2\|y\|^2 - (x|y)^2$. A demonstração prossegue tendo em conta que $\|x\|^2\|y\|^2 - (x|y)^2 =$

$\|x\|^2\|y\|^2 - \|x\|^2\|y\|^2\cos^2\theta = \|x\|^2\|y\|^2(1-\cos^2\theta) = \|x\|^2\|y\|^2 sin^2\theta \Rightarrow \|x\times y\| = \|x\|\|y\||\sin\theta|$. $\square$

Assim, a norma do produto externo $\|x\times y\|$ é a área do paralelogramo de lados x e y (ver a Figura 5).

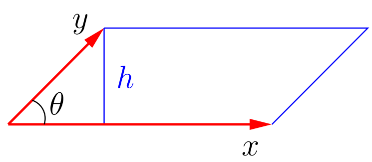


Figura 5.2: $\|x\times y\| = \|x\|\|y\||\sin\theta| = \|x\|h$, é a área do paralelogramo de lados x e y.

Também se pode concluir que o valor absoluto do produto misto $|x\times y|z|$ é o volume do paralelipípedo definido por x, y e z (ver a Figura 5).

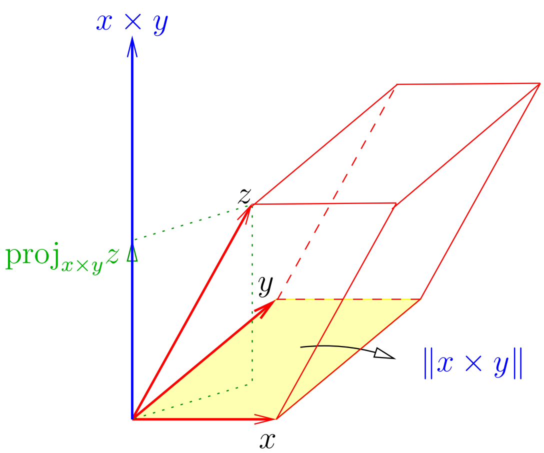


Figura 5.3: $|x\times y|z| = \|x\times y\|\|z\||\cos\theta| = \|x\times y\|\|\mathrm{proj}_{x\times y} z\|$, é o volume do paralelipípedo definido por x, y e z.

Exercícios 29

1. Calcule o determinante de cada uma das seguintes matrizes indicando se é invertível.

a) $\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}, \alpha \in \mathbb{R}$ b) $\begin{bmatrix} 2 & 4 & 18 \\ 1 & 3 & 15 \\ 1 & 0 & 6 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 1 & 2 & 0 \\ -1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ e) $\begin{bmatrix} 2 & 0 & 2 & 1 \\ 2 & -2 & 1 & 0 \\ 0 & 1 & 0 & -2 \\ 5 & 1 & -1 & 3 \end{bmatrix}$ f) $\begin{bmatrix} 1 & 3 & 4 & 3 \\ 2 & 0 & 0 & 2 \\ 5 & 6 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

2. Utilizando a noção de produto externo, indique

   a) um vetor ortogonal aos vetores $u = (1, 1, 2)$ e $v = (1, 0, 1)$,

   b) uma equação cartesiana do plano definido por

$$(x, y, z) = (1, 2, 3) + \lambda (1, 1, 2) + \mu (1, 0, 1), \ \forall \lambda, \mu \in \mathbb{R}.$$

3. Sejam $P_0 = (x_0, y_0, z_0)$ um ponto de $\mathbb{R}^3$ e $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ vetores linearmente independentes de $\mathbb{R}^3$. Mostre que a equação

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0$$

define o plano que passa no ponto P_0 e que contém as direções dos vetores u e v.

Capítulo 6

Valores e vetores próprios

6.1 Valores e vetores próprios

Definição 29 Seja A uma matriz quadrada de ordem n. Um vetor v não nulo de $\mathbb{R}^n$ é *vetor próprio* de A se existir um número λ tal que $Av = \lambda v$. O número λ chama-se *valor próprio* associado ao vetor próprio v.

Exemplo 41 $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Diz-se pois que $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ é vetor próprio da matriz $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ e 3 é o valor próprio associado.

Note que os vetores próprios associados a λ são os vetores v, não nulos, tais que $Av = \lambda v \Leftrightarrow Av = \lambda I v \Leftrightarrow (A - \lambda I)v = \vec{0}$, i.e., são os vetores de $\mathcal{N}(A - \lambda I) \setminus \{\vec{0}\}$.
Tem-se pois provado os seguintes resultados.

Teorema 6.1 *Seja A uma matriz quadrada.*

1. *λ é valor próprio de A sse o espaço nulo da matriz $A - \lambda I$ inclui vetores não nulos, i.e., $\mathcal{N}(A - \lambda I) \neq \{\vec{0}\}$.*

2. *Se λ é valor próprio de A, os vetores próprios associados a λ são os vetores não nulos de $\mathcal{N}(A - \lambda I)$.*

Definição 30 Se λ é valor próprio da matriz A, o espaço nulo de $A - \lambda I$ chama-se *subespaço próprio* de λ e representa-se por $E(\lambda)$.

EXEMPLO 42 Consideremos a matriz $A = \begin{bmatrix} 2 & -5 & 5 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}$.

Para decidir se 2 é valor próprio de A, vamos ver se o espaço nulo da matriz $A - \lambda I$, com $\lambda = 2$, inclui vetores não nulos, i.e., se existem soluções não nulas do sistema homogéneo $(A - 2I)x = \vec{0}$.

Aplicando a fase descendente do método de Gauss à matriz $A - 2I$, tem-se

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

Como a matriz em escada obtida tem colunas sem *pivots*, podemos concluir que o sistema $(A - 2I)x = \vec{0}$ tem soluções não nulas, o que permite concluir que 2 é valor próprio da matriz A.

Para identificar os vetores próprios associados ao valor próprio 2, vamos determinar o espaço próprio $E(2)$, que é o conjunto das soluções do sistema $(A - 2I)x = \vec{0}$. Para isso aplica-se a fase ascendente do método de Gauss à matriz em escada obtida anteriormente. Assim,

$$\begin{bmatrix} 0 & -5 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ e portanto } E(2) = \left\{ v = (v_1, v_2, v_3) : \begin{array}{l} v_1 = \forall \\ v_2 = v_3 \\ v_3 = \forall \end{array} \right\}.$$

Os vetores próprios associados ao valor próprio 2 são os vetores não nulos de $E(2)$, i.e., os vetores não nulos de $\mathbb{R}^3$ que têm a segunda componente igual à terceira.

Podemos facilmente verificar que, se v é um qualquer vetor de $E(2)$, i.e., $v = (a, b, b)$, tem-se

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

O ponto 2 do Teorema 6.1 indica (e o Exemplo 42 ilustra) como se podem identificar os vetores próprios associados a cada valor próprio. Vamos agora ver como é que se determinam os valores próprios de uma matriz.

O ponto 1 do Teorema 6.1 estabelece que λ é valor próprio da matriz A sse o sistema homogéneo $(A - \lambda I)x = 0$ é indeterminado, que como sabemos é equivalente à não existência de inversa da matriz $A - \lambda I$, ou ainda ao facto do determinante de $A - \lambda I$ ser igual a zero. Tem-se pois o seguinte resultado.

Proposição 6.2 *λ é valor próprio de A sse* $\det(A - \lambda I) = 0$.

Assim, os valores próprios de A são os valores de λ que anulam a função $p(\lambda) = \det(A-\lambda I)$. Vamos ver que a função $p(\lambda)$ é um polinómio na variável λ.

Se $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ é uma matriz genérica de ordem 2, $A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}$ e $p(\lambda) = \det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} - a_{22})\lambda - a_{12}a_{21}$ é um polinómio de grau 2.

Se $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ é uma matriz de ordem 3, $A-\lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{bmatrix}$

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) = \quad & (a_{11} - \lambda) \det \begin{bmatrix} a_{22} - \lambda & a_{23} \\ a_{32} & a_{33} - \lambda \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} - \lambda \end{bmatrix} + \\ & + \quad a_{13} \det \begin{bmatrix} a_{21} & a_{22} - \lambda \\ a_{31} & a_{32} \end{bmatrix}. \end{aligned}$$

Uma vez que o 1º termo é um polinómio de grau 3 e os 2º e 3º termos são polinómios de grau 1, $p(\lambda)$ é um polinómio de grau 3.

Repetindo este raciocínio para matrizes genéricas de ordens $4, 5, \ldots$, conclui-se o seguinte.

Proposição 6.3 *Se A é uma matriz quadrada de ordem n, a função $p(\lambda) = \det(A - \lambda I)$ é um polinómio de grau n, que se chama* polinómio característico *de A.*

Os valores próprios são portanto os zeros do polinómio característico.

EXEMPLOS 43

1. Para determinar os valores próprios da matriz $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ considera-se o polinómio característico
$$p(\lambda) = \det \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 2 & 1-\lambda \end{bmatrix} = (1-\lambda)((1-\lambda)^2 - 1) = (1-\lambda)(-\lambda)(2-\lambda).$$
Os valores próprios de A são 0, 1 e 2, pois são os valores de λ que anulam o polinómio característico.

2. O polinómio característico da matriz $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ é
$$p(\lambda) = \det \begin{bmatrix} -\lambda & 1 & 0 \\ -1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix} = (1-\lambda)(\lambda^2 + 1).$$
Os valores próprios de A são $\lambda = 1$ e os zeros de $\lambda^2 + 1$, que são os números imaginários i e $-i$.

Uma matriz A de ordem n tem n valores próprios, reais e/ou complexos, distintos ou não. O número de vezes que λ aparece como zero do polinómio é a multiplicidade algébrica de λ. Assim, por exemplo, os zeros de $(2-\lambda)^2\lambda(1+\lambda)^3$ são 2, 0 e -1 com multiplicidades algébricas iguais a 2, 1 e 3, respetivamente.
Note que em cada um dos Exemplos 43 a soma dos valores próprios é igual à soma dos elementos da diagonal principal da matriz A. Também o determinante de cada matriz e o produto dos correspondentes valores próprios são iguais. Tal facto não é uma coincidência, como estipulam os dois resultados seguintes, que permitem de alguma forma averiguar eventuais erros cometidos no cálculo dos valores próprios.

Proposição 6.4 *Sejam $\lambda_1, \lambda_2, \ldots, \lambda_n$ os valores próprios de uma matriz A de ordem n.*

1. *A soma dos valores própios é igual ao* traço *da matriz, i.e., $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$.*

2. *O produto do valores próprios é igual ao determinante da matriz, i.e., $\lambda_1\lambda_2 \ldots \lambda_n = \det A$.*

Resulta diretamente do ponto 2 da Proposição 6.4 a seguinte caracterização da invertibilidade de matrizes em termos de valores próprios.

Proposição 6.5 *Uma matriz é singular (i.e., não é invertível) sse zero é valor próprio.*

Exercícios 30

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

    a) Verifique que $(1, 5, 10)$ é vetor próprio.

    b) Verifique que 1 é valor próprio.

2. Verifique que -1 é valor próprio da matriz $A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$ e determine os vetores próprios associados a -1.

3. Determine os valores próprios e correspondentes vetores próprios de cada uma das seguintes matrizes, indicando em cada caso, uma base e a dimensão do subespaço próprio associado a cada valor próprio.

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\ C = \begin{bmatrix} 1 & 0 & 0 \\ -7 & 1 & 0 \\ 4 & -3 & 1 \end{bmatrix},\ D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix},$$

$$E = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix},\ F = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix},\ G = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

4. Considere a matriz $A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 2 & 0 \\ 1 & a & a \end{bmatrix}$, com $a \in \mathbb{R}$.

   a) Determine os valores do parâmetro a para os quais a matriz A admite o valor próprio zero.

   b) Para cada um dos valores de a obtidos na alínea anterior calcule os valores próprios de A e identifique os correspondentes vetores próprios.

   c) Discuta, em função do parâmetro a, a invertibilidade da matriz A.

5. Seja v um vetor próprio associado ao valor próprio λ de uma matriz A.

   a) Mostre que, para todo o real α, v é um vetor próprio da matriz $A-\alpha I$ e indique o valor próprio associado.

b) Mostre que, para todo o inteiro n, v é vetor próprio da matriz A^n e indique o valor próprio associado.

6.2 Diagonalização

Uma questão importante no estudo dos valores e vetores próprios é a diagonalização de matrizes. Começamos esta secção com a definição de matrizes semelhantes.

Definição 31 Duas matrizes quadradas da mesma ordem A e B são *semelhantes* se existir uma matriz invertível P, tal que $B = P^{-1}AP$.

EXEMPLO 44 As matrizes $A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -3 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 0 \\ 4 & -5 & 3 \end{bmatrix}$ são semelhantes.

De facto, tomando a matriz invertível $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ -1 & 0 & 0 \end{bmatrix}$ tem-se

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

Pode provar-se o seguinte.

Proposição 6.6 *Matrizes semelhantes têm os mesmos valores próprios.*

Demonstração: Se A e B são matrizes semelhantes, existe uma matriz invertível P tal que $B = P^{-1}AP$. Assim, tem-se $\det(B - \lambda I) = \det(P^{-1}AP - \lambda I) = \det(P^{-1}AP - \lambda P^{-1}P) = \det(P^{-1}(AP - \lambda P)) = \det(P^{-1}(A - \lambda I)P) = \det P^{-1} \det(A - \lambda I) \det P = \det(A - \lambda I)$, i.e., as matrizes A e B têm o mesmo polinónimo característico e, consequentemente, os mesmos valores próprios. $\square$

Definição 32 Uma matriz quadrada A é *diagonalizável* se é semelhante a uma matriz diagonal, i.e., existe uma matriz P invertível, tal que $D = P^{-1}AP$ é uma matriz diagonal. Diz-se que P é *matriz de diagonalização.*

OBSERVAÇÃO 27 Se a matriz A é semelhante à matriz diagonal D,

1. os valores próprios A são os elementos da diagonal principal de D;

2. como $D = P^{-1}AP \Leftrightarrow A = PDP^{-1}$, tem-se,
   para todo o $k \in \mathbb{Z}^+$, $A^k = \underbrace{(PDP^{-1})(PDP^{-1})...(PDP^{-1})}_{k \text{ vezes}} = PD^kP^{-1}$.

O resultado seguinte estabelece uma condição necessária e suficiente para uma matriz ser diagonalizável.

Teorema 6.7 *Uma matriz quadrada A de ordem n é diagonalizável sse existem n vetores próprios de A que formam um conjunto linearmente independente.*

Demonstração:

(i) Se A é diagonalizável, existe uma matriz P invertível, tal que $D = P^{-1}AP$ é matriz diagonal.

Sejam $P = \begin{bmatrix} w_1 & w_2 & \dots & w_n \\ | & | & & | \end{bmatrix}$ e $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \lambda_n \end{bmatrix}$.

Note que $D = P^{-1}AP \Leftrightarrow PD = AP$.

Como $AP = \begin{bmatrix} Aw_1 & Aw_2 & \dots & Aw_n \\ | & | & & | \end{bmatrix}$ e $PD = \begin{bmatrix} \lambda_1 w_1 & \lambda_2 w_2 & \dots & \lambda_n w_n \\ | & | & & | \end{bmatrix}$,

$AP = PD$ significa que $Aw_1 = \lambda_1 w_1, Aw_2 = \lambda_2 w_2, \dots, Aw_n = \lambda_n w_n$, i.e., $w_1, w_2, \dots, w_n$ são n vetores próprios de A. Esses vetores próprios formam um conjunto linearmente independente uma vez que são as colunas da matriz invertível P.

Note também que os valores próprios associados às colunas de P são os elementos da diagonal principal matriz D.

(ii) Se $\{v_1, v_2, \ldots, v_n\}$ é um conjunto de vetores próprios de A linearmente independente e $\lambda_1, \lambda_2, \ldots, \lambda_n$ os valores próprios associados, vamos definir a matriz invertível

$$P := \begin{bmatrix} v_1 & v_2 & \ldots & v_n \\ | & | & & | \end{bmatrix} \text{ e a matriz diagonal } D := \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

As igualdades $Av_1 = \lambda_1 v_1, Av_2 = \lambda_2 v_2, \ldots, Av_n = \lambda_n v_n$ podem ser escritas matricialmente na forma

$$\begin{bmatrix} Av_1 & Av_2 & \ldots & Av_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \ldots & \lambda_n v_n \\ | & | & & | \end{bmatrix} \Leftrightarrow AP = PD \Leftrightarrow P^{-1}AP = D,$$

que permite conluir que A é diagonalizável. $\square$

Observação 28 Na demonstração do Teorema 6.7 constatou-se o seguinte. Se P é uma matriz de diagonalização da matriz A de ordem n,

1. as n colunas de P são vetores próprios de A que formam um conjunto linearmente independente;

2. o valor próprio associado à coluna i da matriz P é o elemento (i, i) da matriz diagonal $D = P^{-1}AP$.

Teorema 6.8 *Um conjunto de vetores próprios associados a valores próprios distintos é linearmente independente.*

Demonstração: Seja A uma matriz quadrada de ordem n.

(i) Sejam $\lambda_1 \neq \lambda_2$ valores próprios de A e v_1, v_2 vetores próprios correspondentes. Quer provar-se que a combinação linear nula $\alpha_1 v_1 + \alpha_2 v_2 = \vec{0}$ só é realizável com os coeficientes $\alpha_1 = \alpha_2 = 0$.

Ora, $\alpha_1 v_1 + \alpha_2 v_2 = \vec{0} \Rightarrow A(\alpha_1 v_1) + A(\alpha_2 v_2) = A\vec{0} \Leftrightarrow \alpha_1(Av_1) + \alpha_2(Av_2) = \vec{0} \Leftrightarrow$ $\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 = \vec{0} \Leftrightarrow$ (como $\alpha_2 v_2 = -\alpha_1 v_1$) $\alpha_1 \lambda_1 v_1 + \lambda_2(-\alpha_1 v_1) = \vec{0} \Leftrightarrow \alpha_1(\lambda_1 -$

$\lambda_2)v_1 = \vec{0} \Rightarrow$ (como $v_1 \neq \vec{0}$) $\alpha_1(\lambda_1 - \lambda_2) = 0 \Rightarrow$ (como $\lambda_1 \neq \lambda_2$) $\alpha_1 = 0 \Rightarrow \alpha_2 = 0$, e portanto $\{v_1, v_2\}$ é linearmente independente.

(ii) Sejam $\lambda_1, \lambda_2, \lambda_3$ valores próprios distintos de A e v_1, v_2, v_3 vetores próprios correspondentes. Quer provar-se que a combinação linear nula $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 \lambda_3 v_3 = \vec{0}$ só é realizável se $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = \vec{0} \overset{(\times A)}{\Rightarrow} \alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \alpha_3\lambda_3 v_3 = \vec{0} \Leftrightarrow \alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \lambda_3(-\alpha_1 v_1 - \alpha_2 v_2) = \vec{0} \Leftrightarrow \alpha_1(\lambda_1 - \lambda_3)v_1 + \alpha_2(\lambda_2 - \lambda_3)v_2 = \vec{0}$. Uma vez que v_1 e v_2 são vetores próprios associados a valores próprios distintos, de (i) resulta que a equação anterior só é satisfeita com $\alpha_1(\lambda_1 - \lambda_3) = \alpha_2(\lambda_2 - \lambda_3) = 0$. Tendo em conta que $\lambda_1 \neq \lambda_3$ e $\lambda_2 \neq \lambda_3$, tem-se $\alpha_1 = \alpha_2 = 0$. Como v_3 é um vetor não nulo, $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = \vec{0}$ só se verifica se também $\alpha_3 = 0$.

Temos assim provado que $\{v_1, v_2, v_3\}$ é linearmente independente.

O resultado para $k > 3$ valores próprios distintos prova-se de forma análoga. □

O teorema anterior permite concluir que, se uma matriz de ordem n tem n valores próprios distintos, então é diagonalizável. E se a matriz tem algum valor próprio com multiplicidade algébrica maior do que 1? A resposta a esta questão é dada utilizando o seguinte conceito.

Definição 33 Chama-se *multiplicidade geométrica* do valor próprio λ da matriz A à dimensão do subespaço próprio $E(\lambda) = \mathcal{N}(A - \lambda I)$.

A relação entre multiplicidades álgebrica e geométrica é estabelecida no resultado seguinte.

Proposição 6.9 *A multiplicidade geométrica de um valor próprio é menor ou igual do que a multiplicidade algébrica.*

O próximo teorema estabelece uma forma expedita de decidir sobre a diagonalização de matrizes.

Teorema 6.10 *Uma matriz é diagonalizável sse as multiplicidades geométrica e algébrica de cada valor próprio são iguais.*

Exemplo 45 Veja se a matriz $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 5 & 0 \\ 4 & -2 & 5 \end{bmatrix}$ é diagonalizável.

$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} -1-\lambda & 1 & 0 \\ 0 & 5-\lambda & 0 \\ 4 & -2 & 5-\lambda \end{bmatrix} = (5-\lambda)(-1-\lambda)(5-\lambda).$

$p(\lambda) = 0 \Leftrightarrow \lambda = -1$ (mult. alg. 1) ou $\lambda = 5$ (mult. alg. 2).

$E(5) = \mathcal{N}(A - 5I) = \mathcal{N} \underbrace{\begin{bmatrix} -6 & 1 & 0 \\ 0 & 0 & 0 \\ 4 & -2 & 0 \end{bmatrix}}_{\text{car}=2}.$

Assim, mult. geométrica de $5 = \dim E(5) = 3 - \text{car}(A - 5I) = 1 <$ mult. algébrica de 5 $= 2$, e portanto A não é diagonalizável.

Os valores e vetores próprios de matrizes simétricas têm propriedades interessantes.

Teorema 6.11 *Se A é uma matriz simétrica ($A = A^\top$),*

1. *os valores próprios são reais;*
2. *a matriz é diagonalizável;*
3. *vetores próprios associados a valores próprios distintos são ortogonais.*

Demonstração do ponto 3: Sejam $\lambda_1 \neq \lambda_2$ valores próprios da matriz simétrica A e v_1, v_2 vetores próprios correspondentes. Tem-se $\lambda_1 v_1 | v_2 = (Av_1)|v_2 = (Av_1)^\top v_2 = v_1^\top A^\top v_2 = v_1^\top A v_2 = v_1^\top \lambda_2 v_2 = \lambda_2 v_1 | v_2$. Ora, $\lambda_1 v_1|v_2 = \lambda_2 v_1 | v_2 \Leftrightarrow (\lambda_1 - \lambda_2) v_1 | v_2 = 0 \overset{\lambda_1 \neq \lambda_2}{\Rightarrow} v_1 | v_2 = 0$, i.e., $v_1 \perp v_2$. $\square$

Exemplo 46 Vamos verificar que os vetores próprios associados a valores próprios distintos da matriz simétrica $A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$ são ortogonais.

$p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 0 & 3-\lambda \end{bmatrix} = (2-\lambda)((3-\lambda)^2 - 1) = (2-\lambda)(2-\lambda)(4-\lambda)$.

Os valores próprios são 2 (m. alg = 2) e 4 (m. alg = 1).

$E(2) = \mathcal{N}(A - 2I) = \mathcal{N} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \mathcal{N} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \{ \begin{bmatrix} x_1 = -x_3 \\ x_2 = \forall \\ x_3 = \forall \end{bmatrix} \} = \{ \begin{bmatrix} -b \\ a \\ b \end{bmatrix} \}$.

$E(4) = \mathcal{N}(A - 4I) = \mathcal{N} \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \cdots = \mathcal{N} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \{ \begin{bmatrix} x_1 = x_3 \\ x_2 = 0 \\ x_3 = \forall \end{bmatrix} \} = \{ \begin{bmatrix} c \\ 0 \\ c \end{bmatrix} \}$.

Ora, $(-b, a, b)|(c, 0, c) = -bc + 0 + bc = 0$, i.e, quaisquer dois vetores próprios u e v, com $u \in E(2)$ e $v \in E(4)$, são ortogonais.

Teorema 6.12 *Uma matriz simétrica A do tipo $n \times n$ tem n vetores próprios ortogonais.*

Demonstração: Defina-se uma base ortogonal do subespaço próprio de cada valor próprio de A. Como A é diagonalizável (e portanto o número de vetores da base é igual à multiplicidade algébrica do correspondente valor próprio), a reunião destas bases é constituída por n vetores. Se dois destes vetores estão associados ao mesmo valor próprio, são ortogonais por construção. Se estão associados a valores próprios distintos, o ponto 3 do Teorema 6.11 estabelece que são ortogonais. □

Assim, uma matriz simétrica A, do tipo $n \times n$, tem n vetores próprios ortonormais (ortogonais de norma 1). Sejam $v_1, v_2, \ldots, v_n$ vetores próprios ortonormais, e $\lambda_1, \lambda_2, \ldots, \lambda_n$ os valores próprios correspondentes.

Se definirmos a matriz $P = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \\ | & | & & | \end{bmatrix}$, tem-se $P^{-1}AP = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \lambda_n \end{bmatrix} = D$.

Como $P^\top = P^{-1}$, i.e., P é *matriz ortogonal*, tem-se $D = P^{-1}AP = P^\top AP$. Diz-se que A é *ortogonalmente diagonalizável*, i.e., admite matrizes de diagonalização ortogonais. Tem-se assim provado o seguinte resultado

Teorema 6.13 *Matrizes simétricas são ortogonalmente diagonalizáveis*

Sejam A uma matriz simétrica do tipo $n \times n$, $v_1, v_2, \ldots, v_n$ vetores próprios ortonormais e $\lambda_1, \lambda_2, \ldots, \lambda_n$ os correspondente valores próprios.

Se definirmos $P = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \\ | & | & & | \end{bmatrix}$, $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \lambda_n \end{bmatrix}$, tem-se $P^{-1} = \begin{bmatrix} v_1 & — \\ v_2 & — \\ \vdots & \\ v_n & — \end{bmatrix}$

e $A = PDP^{-1}$.

Se tomarmos um vetor arbitrário x de $\mathbb{R}^n$, vem

$Ax = (PD)(P^{-1}x) = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \ldots & \lambda_n v_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} v_1|x \\ \vdots \\ v_n|x \end{bmatrix} = \lambda_1 v_1 v_1|x + \cdots + \lambda_n v_n v_n|x =$
$\lambda_1 v_1 v_1^\top x + \cdots + \lambda_n v_n v_n^\top x = (\lambda_1 v_1 v_1^\top + \cdots + \lambda_n v_n v_n^\top)x$.

Como o vetor x é arbitrário, pode concluir-se das igualdades anteriores que $A = \lambda_1 v_1 v_1^\top + \cdots + \lambda_n v_n v_n^\top$, i.e., a matriz A pode ser escrita à custa dos valores próprios e de vetores próprios ortonormais. Este resultado, conhecido como *Teorema da decomposição espectral*, é agora enunciado.

Teorema 6.14 *Sejam A uma matriz simétrica do tipo $n \times n$, $v_1, v_2, \ldots, v_n$ vetores próprios* *ortonormais e* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *os correspondentes valores próprios. A matriz A pode ser decomposta na forma seguinte* $A = \lambda_1 v_1 v_1^\top + \lambda_2 v_2 v_2^\top \cdots + \lambda_n v_n v_n^\top$.

Observação 29 O Teorema da decomposição espectral tem a seguinte interpretação. Toda a matriz simétrica do tipo $n \times n$ é uma combinação linear das matrizes de projeção sobre cada um de n vetores próprios ortonormais. Os coeficientes são os correspondentes valores próprios.

Exemplo 47 Como se viu no Exemplo 46, a matriz simétrica $A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$ admite os vetores próprios $(-b, a, b)$, correspondentes ao valor próprio 2 e $(c, 0, c)$, associados ao valor próprio 4. Fazendo cada uma das variáveis livres igual a 1 e as restantes iguais a 0, obtem-se o conjunto $\{(0,1,0), (-1,0,1), (1,0,1)\}$ de três vetores próprios linearmente independente. Como o conjunto é ortogonal, para obter três vetores próprios ortonormais basta tomar o versor de cada um deles, i.e., $v_1 = (0,1,0), v_2 = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), v_3 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$. As matrizes de projeção sobre cada um desses vetores são

$$v_1 v_1^\top = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; v_2 v_2^\top = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \; v_3 v_3^\top = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

A combinação linear destas matrizes, com coeficientes iguais aos correspondentes valores próprios, é a matriz

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

Exercícios 31

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

   a) Calcule os valores próprios de A e as respetivas multiplicidades algébricas.

   b) Indique um vetor próprio de A.

   d) Será que existe uma matriz quadrada P, de ordem 3, invertível tal que $P^{-1}AP$ é uma matriz diagonal? Justifique.

2. Indique, justificando, quais das seguintes matrizes são diagonalizáveis.

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ -7 & 1 & 0 \\ 4 & -3 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix},$$

$$E = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, F = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}, G = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 5 \end{bmatrix}, H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

3. Determine uma matriz de diagonalização de cada uma das seguintes matrizes

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

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

   a) verifique que o polinómio característico de A é $p(\lambda) = \lambda(1-\lambda)(\lambda - \frac{1}{4})$.

b) Determine uma matriz invertível P tal que $P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$.

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

   a) Indique uma matriz de diagonalização.

   b) Prove que $\lim_{n \to +\infty} A^n = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$.

6. Considere $A = \begin{bmatrix} 2 & 1 \\ a & b \end{bmatrix}$, com $a, b \in \mathbb{R}$.

   a) Para $a = 2$ e $b = 1$, indique uma matriz de diagonalização.

   b) Se $b = 2$, para que valores de a é A ortogonalmente diagonalizável?

   c) Se $b = 2$, existirá algum $a > 0$ tal que $\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$ e A sejam semelhantes? Justifique.

7. Seja A uma matriz quadrada de ordem 3 que admite o valor próprio 1, com de multiplicidade algébrica 2 e $(1, 0, -1)$, $(0, 1, 1)$ vetores próprios associados a 1.

   a) Justifique que A é diagonalizável.

   b) Determine $E(1)$.

   c) Sabendo que $(-1, 1, 0)$ é um vetor próprio de A associado a 2, determine a matriz A.

8. Indique uma matriz ortogonal de diagonalização da matriz $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$.

9. Prove os seguintes resultados.

   a) Matrizes ortogonalmente diagonalizáveis são simétricas.

   b) Se λ é um valor próprio real não nulo de uma matriz A e v um vetor próprio associado a λ, então λ tem o sinal de $v^T Av$.