# Apontamentos de Complementos

# de

# Álgebra Linear e Geometria Analítica

Licenciatura em Tecnologias de Informação Visual
1º Ano, 2º Semestre, Ano Lectivo de 2004/2005

**Departamento de Matemática da FCTUC**

Joana Teles Luís Nunes Vicente

25 de Maio de 2005

# Índice

# Capítulo 1

# Complementos sobre valores e vectores próprios

## 1.1 Valores próprios e vectores próprios de matrizes

**Revisões de ALGA:**

Dada uma matriz $A$ em $M_{n\times n}(\mathbb{C})$ e um vector não nulo $v \in \mathbb{C}^n$, diz-se que $v$ é um **vector próprio** de $A$ quando $Av$ é múltiplo escalar de $v$, isto é, quando $Av = \lambda v$ para algum $\lambda \in \mathbb{C}$. Nesse caso, $\lambda$ diz-se o **valor próprio** de $A$ associado ao vector próprio $v$.

$$\begin{aligned} \lambda \text{ é valor próprio de } A &\iff N(A-\lambda I) \neq \{0\} \\ &\iff A-\lambda I \text{ não é invertível} \\ &\iff \boxed{\det(A-\lambda I) = 0} \end{aligned}$$

Portanto, um número $\lambda$ é valor próprio de $A$ se $\lambda$ é solução da equação

$$\det(A-\lambda I) = 0$$

designada **equação característica** da matriz $A$. Ao polinómio $\det(A-\lambda I)$ chama-se **polinómio característico** da matriz $A$. Assim, um número $\lambda$ é valor próprio de $A$ se e só se $\lambda$ é raiz do polinómio característico de $A$.

**Teorema Fundamental da Álgebra** (Gauss, 1799)
Qualquer polinómio $p(x)$ de grau $n$ e com coeficientes complexos é da forma

$$p(x) = k(x-\lambda_1)(x-\lambda_2)\dots(x-\lambda_n)$$

para certas constantes $k, \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$. Isto é, a soma das multiplicidades algébricas das raízes de um polinómio de grau $n$ é $n$. Ou seja, qualquer polinómio não constante tem raízes em $\mathbb{C}$ ($\lambda_1, \lambda_2, \dots, \lambda_n$ são as raízes do polinómio $p(x)$).

Em particular, *qualquer matriz tem valores próprios em* $\mathbb{C}$.

Seja $\lambda_0$ um valor próprio de $A$.

A **multiplicidade algébrica** de $\lambda_0$ é a multiplicidade de $\lambda_0$ como raiz do polinómio característico $\det(A - \lambda I)$.

$v \neq 0$ é vector próprio de $A$ associado ao valor próprio $\lambda \iff (A - \lambda I)v = 0$
$\iff v \in N(A - \lambda I)$

O subespaço próprio de $A$ associado ao valor próprio $\lambda$ é $\boxed{N(A - \lambda I)}$

Seja $\lambda_0$ um valor próprio de $A$.

A **multiplicidade geométrica** de $\lambda_0$ é a dimensão do respectivo subespaço próprio $N(A - \lambda_0 I)$.

A multiplicidade geométrica de $\lambda_0$ é sempre menor ou igual à multiplicidade algébrica de $\lambda_0$.

**Matrizes semelhantes**: Suponha-se que $A$ e $B$ são matrizes semelhantes, isto é, que existe uma matriz invertível $T$ tal que $A = TBT^{-1}$. Então:

- os determinantes são iguais, $\boxed{\det A = \det B}$
- os polinómios característicos são iguais, $\boxed{\det(A - \lambda I) = \det\,(B - \lambda I)}$
- os valores próprios são os mesmos e com as mesmas multiplicidades algébricas e geométricas;
- os vectores próprios *correspondem-se mas não são necessariamente os mesmos*; se $v$ é vector próprio de $B$ associado ao valor próprio $\lambda$, então $Tv$ é vector próprio de $A$ associado ao valor próprio $\lambda$.

**Exemplos:**

1. Os valores próprios da matriz $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ são 1 e 3. O subespaço próprio associado a 1 é $N(A - I) = \mathcal{L}\left\{\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]\right\}$ e a 3 é $N(A - 3I) = \mathcal{L}\left\{\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]\right\}$.

2. O polinómio característico da matriz $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{bmatrix}$ é $\det\,(B - \lambda I) = -\lambda(\lambda - 1)^2$. Os valores próprios de $B$ são $\lambda_1 = 0$ e $\lambda_2 = 1$. O valor próprio 1 tem multiplicidade algébrica dois.

O subespaço próprio associado a $\lambda_1 = 0$ é $N(B) = \mathcal{L}\left\{\left[\begin{smallmatrix} -2 \\ -\frac{1}{2} \\ 1 \end{smallmatrix}\right]\right\}$. O subespaço próprio associado a $\lambda_2 = 1$ é $N(B-I) = \mathcal{L}\left\{\left[\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right]\right\}$.

## 1.2 Matrizes diagonalizáveis

**Revisões de ALGA:**

A matriz $A$ diz-se **diagonalizável** quando $A$ é semelhante a uma matriz diagonal.

**A matriz $A$ em $M_{n\times n}(\mathbb{C})$ é diagonalizável se e só se $A$ tiver $n$ vectores próprios linearmente independentes.**
Suponha-se que $v_1, \ldots, v_n$ são vectores próprios de $A$ linearmente independentes e que $\lambda_1, \ldots, \lambda_n$ são os respectivos valores próprios associados, isto é, $Av_j = \lambda_j v_j$ (os $\lambda_j$'s podem ser repetidos). Então $\boxed{A = VDV^{-1}}$ onde

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \quad \text{e} \quad V = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}.$$

**Exemplo:** A matriz $B$ do exemplo da subsecção anterior não é diagonalizável porque tem no máximo dois vectores próprios linearmente independentes.

- **Vectores próprios associados a valores próprios distintos são linearmente independentes**.
- Se $A$ tem $n$ valores próprios distintos dois a dois, então existem $n$ vectores próprios de $A$ linearmente independentes.
- Se as multiplicidades geométricas dos valores próprios de $A$ somam $n$, então existem $n$ vectores próprios de $A$ linearmente independentes.

**Diagonalizar uma matriz**: verificar se a matriz $A$ é diagonalizável e, se for, encontrar uma matriz invertível $V$ tal que $A = VDV^{-1}$ onde $D$ é diagonal. As etapas são as seguintes:

- Determinar os valores próprios de $A$, isto é, determinar as raízes do polinómio característico
$$\det(A - \lambda I);$$
- Para cada valor próprio $\lambda_j$, determinar uma base do seu subespaço próprio
$$N(A - \lambda_j I);$$
- A matriz $A$ é diagonalizável se e só se as dimensões dos subespaços próprios somam $n$. Nesse caso, obtêm-se $n$ vectores linearmente independentes $v_1, \ldots, v_n$ coleccionando bases dos subespaços próprios e a matriz $V$ é a matriz cujas colunas são os $v_j$'s.

**Exemplo:** A matriz $A$ do exemplo da subsecção anterior é diagonalizável. O conjunto de vectores próprios $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$ é linearmente independente, logo sendo $V = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$ e $D = \begin{bmatrix}1 & 0\\0 & 3\end{bmatrix}$ tem-se $A = VDV^{-1}$.

### Exemplo de aplicação da diagonalizabilidade

Uma equação diferencial é uma igualdade em que aparece uma função desconhecida e as suas derivadas, procurando-se determinar as funções que satisfazem essa igualdade. Talvez a equação diferencial mais simples seja

$$f'(t) = af(t),$$

onde $a$ é um número. Qualquer função da forma $Ke^{at}$ é uma solução desta equação diferencial
$$\left(Ke^{at}\right)' = K\left(e^{at}\right)' = Kae^{at} = aKe^{at}.$$
É simples ver que não há outras soluções: se $g(t)$ for uma outra solução tem-se

$$\left(g(t)e^{-at}\right)' = 0$$

(verificar como exercício), e portanto $g(t)e^{-at} = K$, $K$ constante, logo $g(t) = Ke^{at}$.

Tendo-se, não uma, mas duas funções desconhecidas, $f_1(t)$ e $f_2(t)$ satisfazendo as duas igualdades seguintes:

$$\begin{aligned} f_1'(t) &= af_1(t) + bf_2(t) \\ f_2'(t) &= cf_1(t) + df_2(t), \end{aligned}$$

onde $a, b, c, d$ são números, obtém-se um sistema linear de equações diferenciais. Este sistema pode ser escrito na forma matricial como

$$x'(t) = Ax(t)$$

onde $x(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix}$ e $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Dado um tal sistema, não é evidente como se pode resolvê-lo, isto é, encontrar as funções $f_1(t)$ e $f_2(t)$. Se a matriz $A$ for diagonalizável tudo fica mais simples. Seja $A = VDV^{-1}$ e considere-se

$$y(t) = V^{-1}x(t).$$

O sistema dado é equivalente a

$$y'(t) = \big(V^{-1}x(t)\big)' = V^{-1}x'(t) = V^{-1}Ax(t) = DV^{-1}x(t),$$

isto é,

$$y'(t) = Dy(t),$$

sendo $y(t) = \begin{bmatrix} g_1(t) \\ g_2(t) \end{bmatrix}$ e $D$ uma matriz diagonal, logo da forma $\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$.

Tem-se assim duas equações separadas

$$g_1'(t) = d_1 g_1(t) \quad \text{e} \quad g_2'(t) = d_2 g_2(t)$$

cujas soluções são

$$g_1(t) = K_1 e^{d_1 t} \quad \text{e} \quad g_2(t) = K_2 e^{d_2 t},$$

para certas constantes $K_1$ e $K_2$. Então tem-se

$$y(t) = \begin{bmatrix} K_1 e^{d_1 t} \\ K_2 e^{d_2 t} \end{bmatrix}$$

e como $x(t) = Vy(t)$, obtêm-se as funções $f_1(t)$ e $f_2(t)$ procuradas.

Com um raciocínio exactamente análogo consegue-se resolver qualquer sistema de equações diferenciais da forma

$$x'(t) = Ax(t)$$

com $A$ em $M_{n\times n}(\mathbb{C})$ desde que a matriz seja diagonalizável.

Se $A$ não for diagonalizável, o estudo de um sistema linear de equações diferenciais da forma

$$x'(t) = Ax(t)$$

pode ainda basear-se numa factorização do tipo $A = VJV^{-1}$, interessando encontrar $J$ tão simples quanto possível.

## 1.3 O caso das matrizes simétricas reais

**Revisões de ALGA:**

A **matriz transposta** de uma matriz $m \times n$ $A$ é a matriz $n \times m$, $A^\top$, cuja linha $j$ é a coluna $j$ de $A$.
- $(A^\top)^\top = A$ • $(A+B)^\top = A^\top + B^\top$ • $(AB)^\top = B^\top A^\top$
- Se $A$ é invertível, $A^\top$ também o é e $(A^\top)^{-1} = (A^{-1})^\top$.

A matriz $A$ diz-se **simétrica** se $A = A^\top$.
Uma matriz quadrada diz-se **ortogonal** se for invertível e a sua inversa coincidir com a sua transposta.

- Os valores próprios de uma matriz simétrica real são números reais.
- Qualquer matriz simétrica real é diagonalizável com uma matriz diagonalizante ortogonal.
- Os vectores próprios de uma matriz simétrica real correspondentes a valores próprios distintos são ortogonais.

**Exemplo:** Os valores próprios da matriz simétrica real $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$ são -2 e 1. O subespaço próprio associado a 1 é $\mathcal{L}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right\}$. Dois vectores deste subespaço ortogonais são $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ e $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$. O subespaço próprio associado a -2 é $\mathcal{L}\left\{\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}$. As colunas da matriz diagonalizante ortogonal da matriz $A$ são os vectores unitários múltiplos escalares dos vectores próprios ortogonais, isto é, $V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$.

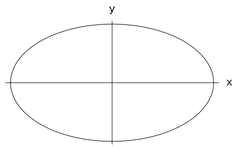
## 1.4 Curvas e superfícies do 2º grau

**Definição 1.4.1** *Uma cónica é o lugar geométrico dos pontos do plano cujas coordenadas cartesianas satisfazem uma equação do segundo grau em duas variáveis*

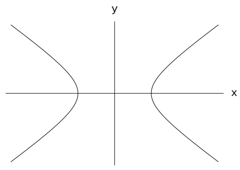
$$a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + b_1x + b_2y + c = 0.$$

**Exemplos**

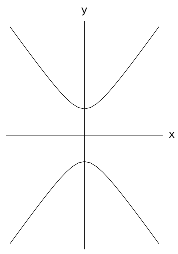
1. Elipse
   $$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$



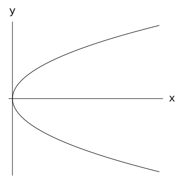

2. Hipérbole
   $$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$



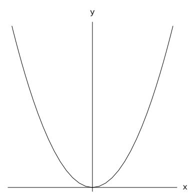

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$$




3. Parábola
   $$y^2 = 2px$$




$$x^2 = 2py$$



4. Cónicas degeneradas

   (a) Um ponto (elipse degenerada)
   $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$

   (b) Duas rectas concorrentes (hipérbole degenerada)
   $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$

   (c) Duas rectas paralelas (parábola degenerada)
   $x^2 = c^2$
   Duas rectas coincidentes (parábola degenerada)
   $x^2 = 0$

   (d) Conjunto vazio, por exemplo, $\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$ (elipse degenerada)

A caracterização das cónicas reduz-se à escolha de um referencial adequado de modo que a equação da cónica assuma uma das formas descritas nos exemplos anteriores – **equações reduzidas**.

A equação

$$a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + b_1x + b_2y + c = 0$$

pode escrever-se na forma matricial

$$X^\top AX + BX + c = 0$$

sendo $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$, $B = \begin{bmatrix} b_1 & b_2 \end{bmatrix}$ e $X = \begin{bmatrix} x \\ y \end{bmatrix}$.

Vamos procurar uma equação reduzida em duas etapas.

**Etapa 1**: Eliminar o termo $2a_{12}xy$

**Etapa 2**: Eliminar termos de grau 1

**Etapa 1**
Uma vez que a matriz $A$ é simétrica real, existe uma matriz ortogonal $Q$ tal que

$$A = QDQ^\top,$$

onde $D$ é uma matriz diagonal. As colunas de $Q$ são vectores próprios ortonormados de $A$ e os elementos diagonais $\lambda_1, \lambda_2 \in \mathbb{R}$ de $D$ são os valores próprios de $A$.

A equação da cónica é

$$X^\top QDQ^\top X + BX + c = 0.$$

Fazendo $X' = \begin{bmatrix} x' \\ y' \end{bmatrix} = Q^\top X$, tem-se $X'^\top = X^\top Q$ e assim a equação é

$$X'^\top D X' + BQX' + c = 0,$$

ou seja,

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d_1 x' + d_2 y' + c = 0,$$

com $BQ = \begin{bmatrix} d_1 & d_2 \end{bmatrix}$.

A mudança de coordenadas é dada por $X' = Q^\top X$. Quais são os novos eixos? O primeiro eixo é gerado pelo vector cujas novas coordenadas são 1 e 0, isto é, pelo vector $X$ que satisfaz $e_1 = Q^\top X$. Logo, a primeira coluna de $Q$. Analogamente, o segundo dos novos eixos é gerado pela segunda coluna de $Q$. Ou seja, passámos para um referencial ortonormado cujos eixos são gerados por vectores próprios ortonormados de $A$ (fez-se uma **rotação** dos eixos coordenados).

**Etapa 2**

Interessa agora eliminar os termos do primeiro grau, o que se faz por uma mudança de coordenadas do tipo

$$X'' = X' + K,$$

com $K \in \mathbb{R}^2$ (esta mudança corresponde a uma **translação**).

(i) Suponhamos $\lambda_1 \neq 0$ e $\lambda_2 \neq 0$.

"Completando quadrados", chegamos a uma equação do tipo

$$\lambda_1 (x' + a')^2 + \lambda_2 (y' + b')^2 + c' = 0.$$

Fazendo agora as substituições $x'' = x' + a'$ e $y'' = y' + b'$, obtemos a equação

$$\lambda_1 x''^2 + \lambda_2 y''^2 + c' = 0,$$

que é já uma equação reduzida.

(ii) Se $\lambda_1 = 0$ (o caso $\lambda_2 = 0$ é semelhante).

"Completando quadrados", chegamos a uma equação do tipo

$$\lambda_2 (y' + b')^2 + d_1 x' + c' = 0.$$

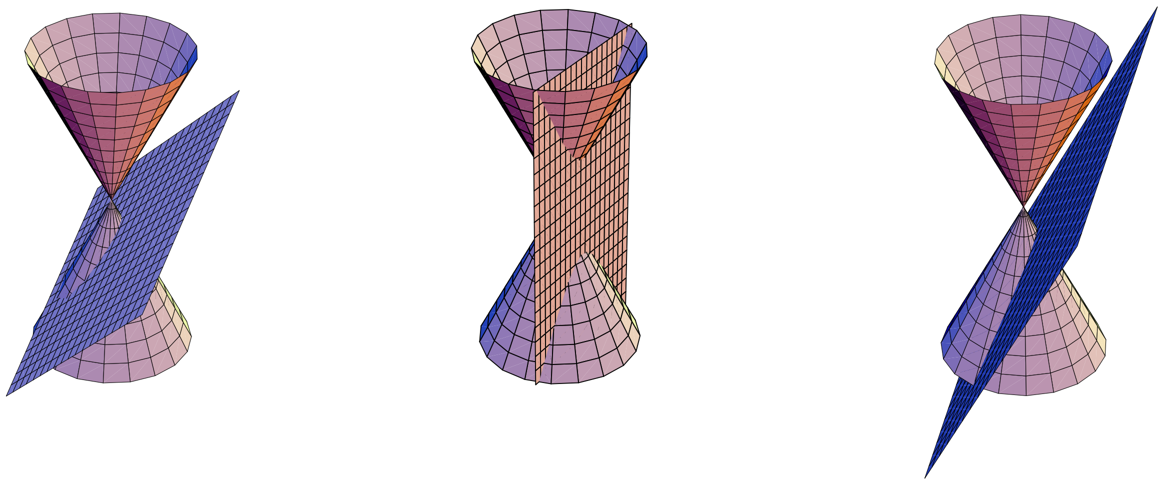
Fazendo a substituição $y'' = y' + b'$, obtemos

$$\lambda_2 y''^2 + d_1 x' + c' = 0.$$

Se $d_1 = 0$, já temos uma equação reduzida, caso contrário faz-se $x'' = x' + \frac{c'}{d_1}$ e obtém-se a equação reduzida

$$\lambda_2 y''^2 + d_1 x'' = 0.$$

**Nota:** Porquê o nome **cónica**? Por poder ser obtida da intersecção de uma superfície cónica com um plano. As cónicas mais importantes são elipses, hipérboles e parábolas.



**Definição 1.4.2** *Uma quádrica é o lugar geométrico dos pontos do espaço cujas coordenadas cartesianas satisfazem uma equação do segundo grau em três variáveis*

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + b_1x + b_2y + b_3z + c = 0.$$

Esta equação tem a seguinte expressão matricial

$$X^\top AX + BX + c = 0$$

sendo $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$, $B = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}$ e $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Tal como no caso das cónicas, podemos da equação geral chegar a uma equação reduzida da quádrica.

Para tal suprimimos primeiro os termos mistos da "parte quadrática", isto é, os monómios de grau dois envolvendo duas variáveis. Seguidamente anulamos todos os termos de grau um que nos for possível. Estas operações são feitas recorrendo a mudanças de variáveis dos tipos:

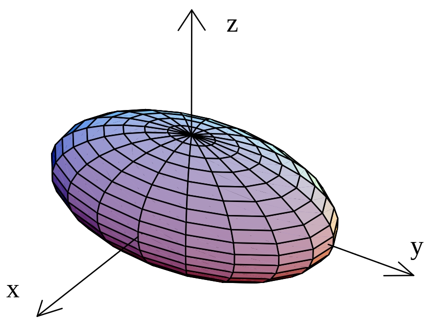
(i) $X' = Q^\top X$, onde $Q$ é uma matriz ortogonal que diagonaliza $A$.

(ii) $X'' = X' + K$, com $K \in \mathbb{R}^3$, "completando quadrados".

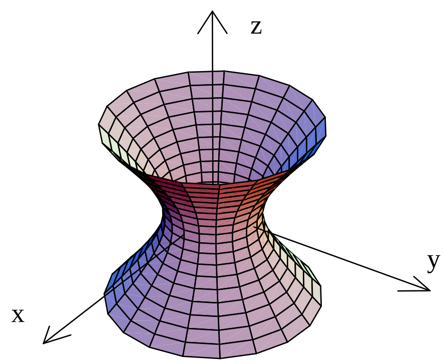
No primeiro caso fazemos uma rotação dos eixos, sendo os novos eixos gerados por vectores próprios de $A$ ortonormados (colunas da matriz diagonalizante $Q$). No segundo caso efectuamos uma translação do sistema de eixos.

Apresentamos de seguida uma lista das equações reduzidas das principais quádricas.

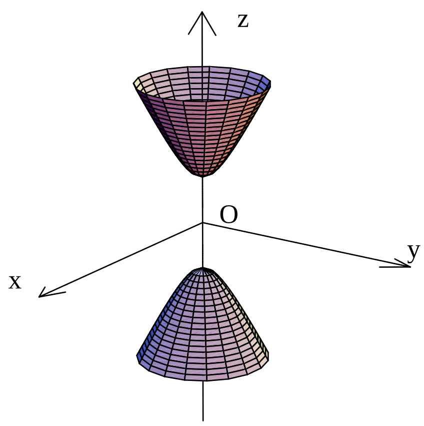
1. Elipsóide
   $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$



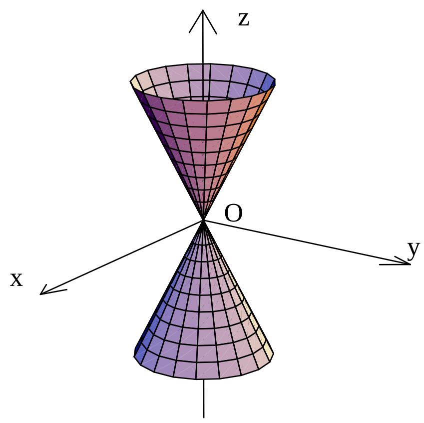

2. Hiperbolóide de uma folha
   $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$



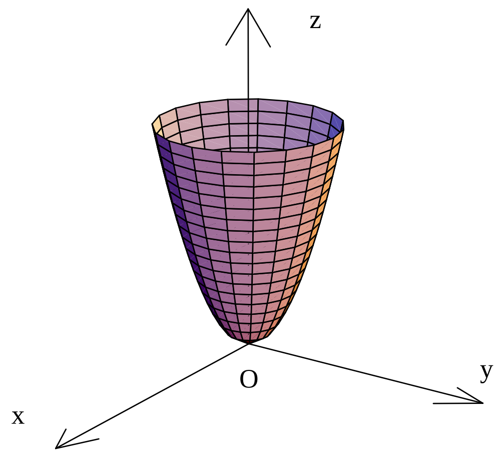

3. Hiperbolóide de duas folhas
   $-\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$



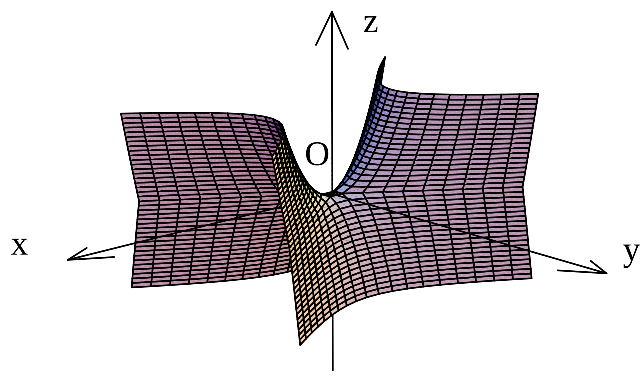
4. Superfície cónica
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$




5. Parabolóide elíptico
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2pz \ (p \neq 0)$$



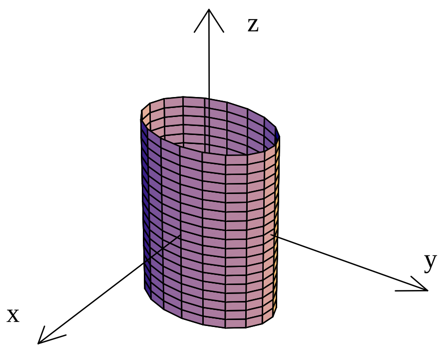

6. Parabolóide hiperbólico
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2pz \ (p \neq 0)$$



7. No espaço, uma equação reduzida em que uma das variáveis não aparece representa uma superfície de tipo cilíndrico. Por exemplo,

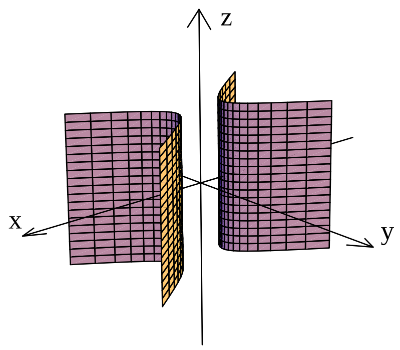
Superfície cilíndrica elíptica

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$



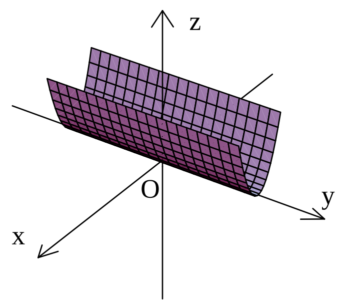

Superfície cilíndrica hiperbólica

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$




Superfície cilíndrica parabólica

$z = ax^2$




## 1.5 O caso das matrizes normais

Recorde-se que a matriz adjunta ou transconjugada de uma matriz $A$ é dada por $A^* = \bar{A}^\top = \overline{A^\top}$. É indiferente transpor a matriz primeiro e depois conjugar os seus elementos ou começar por conjugar antes de aplicar a transposição. Se $A$ for $m \times n$ então $A^*$ é $n \times m$.

**Definição 1.5.1** *Uma matriz $A \in M_{n\times n}(\mathbb{C})$ diz-se hermítica se*

$$A = A^*.$$

Por exemplo, a seguinte matriz é hermítica,

$$A = \begin{bmatrix} 1 & 1-i \\ 1+i & 2 \end{bmatrix},$$

uma vez que

$$\bar{A} = \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix} = A^\top.$$

Este exemplo motiva o enunciado da seguinte proposição, cuja demonstração é deixada como exercício.

**Proposição 1.5.1** *Uma matriz hermítica tem elementos reais ao longo da sua diagonal principal. O seu elemento na posição $i, j$ é o conjugado do seu elemento na posição $j, i$, para índices $i, j \in \{1, \ldots, n\}$.*

As matrizes hermíticas em $M_{n\times n}(\mathbb{C})$ correspondem às matrizes simétricas em $M_{n\times n}(\mathbb{R})$, no sentido em que uma matriz hermítica com elementos reais é simétrica.

Uma outra generalização de $M_{n\times n}(\mathbb{R})$ para $M_{n\times n}(\mathbb{C})$ é o das matrizes unitárias. As matrizes unitárias com elementos reais são ortogonais.

**Definição 1.5.2** *Uma matriz $A \in M_{n\times n}(\mathbb{C})$ diz-se unitária se*

$$A^* = A^{-1}.$$

Assim sendo, uma matriz é unitária se $UU^* = I$ (e necessariamente $U^*U = I$) ou se $U^*U = I$ (e de certeza que $UU^* = I$).

É fácil confirmar que é unitária a seguinte matriz:

$$U = \begin{bmatrix} \frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}i \end{bmatrix}.$$

Para o efeito, basta mostrar uma das igualdades $UU^* = I$ ou $U^*U = I$.

A noção de matriz unitária permite-nos expandir o conceito de matrizes ortogonalmente semelhantes para matrizes com elementos complexos. Ambas as noções são particularizações da noção de equivalência entre matrizes (em que se assume somente que a matriz $U$ em baixo é invertível) .

**Definição 1.5.3** *Duas matrizes $A, B \in M_{n\times n}(\mathbb{C})$ dizem-se unitariamente semelhantes se existir uma matriz $U \in M_{n\times n}(\mathbb{C})$ unitária tal que*

$$B = U^*AU.$$

Desta definição resulta que $A = UBU^*$. Este raciocínio é importante e aparecerá variadas vezes. Em primeiro lugar multiplicam-se ambos os membros de $B = U^*AU$, à esquerda, por $U$, obtendo-se, de $UU^* = I$, a igualdade $UB = AU$. A seguir multiplicam-se ambos os membros de $UB = AU$, à direita, por $U^*$, concluindo-se que $A = UBU^*$.

As matrizes unitariamente semelhantes com elementos reais são ortogonalmente semelhantes. Neste caso, a matriz $U = Q$ desta definição tem entradas reais e é ortogonal ($QQ^\top = Q^\top Q = I$).

Sabe-se que duas matrizes semelhantes $A$ e $B$ têm os mesmos valores próprios. Logo, duas matrizes (ortogonalmente ou) unitariamente semelhantes partilham os mesmos valores próprios.

### 1.5.1 O Teorema de Schur

Qualquer matriz é unitariamente semelhante a uma matriz triangular. Este resultado, conhecido como Teorema de Schur, é enunciado e demonstrado de seguida, escolhendo como matrizes triangulares as triangulares superiores. Como veremos mais adiante, este teorema tem várias implicações relevantes em diagonalização de matrizes.

**Teorema 1.5.1** *Toda a matriz $A \in M_{n\times n}(\mathbb{C})$ é unitariamente semelhante a uma matriz triangular superior, ou seja, qualquer que seja $A$ existe uma matriz unitária $U \in M_{n\times n}(\mathbb{C})$ tal que*

$$U^*AU \;=\; T,$$

*em que $T \in M_{n\times n}(\mathbb{C})$ satisfaz $t_{ij} = 0, i > j$. Os elementos diagonais de $T$ são, desta forma, os valores próprios de $A$.*

**Demonstração.** Seja $v_1$ um vector próprio de $A$ normalizado, ou seja,

$$Av_1 \;=\; \lambda_1 v_1, \quad \|v_1\| = 1.$$

Para efeitos desta demonstração, seja

$$\{v_1, w_2, \ldots, w_n\}$$

uma base ortonormada de $\mathbb{C}^n$. Esta base pode ser obtida acrescentando vectores a $\{v_1\}$, de forma linearmente independente, até ser alcançada uma base para $\mathbb{C}^n$ e, depois, aplicando o processo de ortogonalização de Gram-Schmidt a esta última base.

Ponha-se $U_1 = [\, v_1 \; w_2 \; \cdots \; w_n \,]$. Assim sendo, tem-se que

$$\begin{aligned} U_1^* A U_1 \;&=\; U_1^*[\, Av_1 \; Aw_2 \; \cdots \; Aw_n \,] \\ &=\; \begin{bmatrix} v_1^* \\ w_2^* \\ \vdots \\ w_n^* \end{bmatrix} [\, \lambda_1 v_1 \; Aw_2 \; \cdots \; Aw_n \,] \\ &=\; \begin{bmatrix} \lambda_1 & r_1^* \\ 0 & A_1 \end{bmatrix}, \end{aligned}$$

com $r_1 \in \mathbb{C}^{n-1}$ e $A_1 \in M_{n-1\times n-1}(\mathbb{C})$. Daqui resulta que as matrizes $A$ e

$$\begin{bmatrix} \lambda_1 & r_1^* \\ 0 & A_1 \end{bmatrix}$$

têm os mesmos valores próprios. Se $n$ fosse igual a 2 a demonstração estaria concluída. Suponhamos, então, que $n \geq 2$.

Agora, apliquemos a $A_1$ um raciocínio idêntico, garantindo a existência de uma matriz unitária $U_2 \in M_{n-1\times n-1}(\mathbb{C})$ tal que

$$U_2^* A_1 U_2 \;=\; \begin{bmatrix} \lambda_2 & r_2^* \\ 0 & A_2 \end{bmatrix},$$

com $r_2 \in \mathbb{C}^{n-2}$ e $A_2 \in M_{n-2\times n-2}(\mathbb{C})$. A matriz unitária que permite continuar o raciocínio é dada por

$$V_2 \;=\; \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}.$$

É relegado para os exercícios provar que $V_2$ é unitária. Continuando, veja-se que

$$\begin{aligned}
V_2^*(U_1^* A U_1)V_2 \;&=\; \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & r_1^* \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \\
&=\; \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & r_1^* U_2 \\ 0 & A_1 U_2 \end{bmatrix} \\
&=\; \begin{bmatrix} \lambda_1 & r_1^* U_2 \\ 0 & U_2^* A_1 U_2 \end{bmatrix} \\
&=\; \begin{bmatrix} \lambda_1 & (r_1^* U_2)_{1,1} & (r_1^* U_2)_{1,2:n} \\ 0 & \lambda_2 & r_2^* \\ 0 & 0 & A_2 \end{bmatrix}.
\end{aligned}$$

A demonstração é concluída aplicando o mesmo argumento sucessivamente até se obter uma matriz $A_{n-1} \in M_{1\times 1}(\mathbb{C})$. No final, chega-se a

$$U^* A U \;=\; T \quad \text{com} \quad U \;=\; U_1 V_2 \cdots V_{n-1}$$

e $t_{ii} = \lambda_i$, $i = 1, \ldots, n$, em que $T$ é uma matriz triangular superior. Fica como exercício provar que $U$ é uma matriz unitária. ■

As matrizes $T_1$ e $T_2$, dadas por

$$T_1 \;=\; \begin{bmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{e} \quad T_2 \;=\; \begin{bmatrix} 2 & -1 & 3\sqrt{2} \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 3 \end{bmatrix},$$

são unitariamente (aliás, ortogonalmente) semelhantes. A matriz $U$ que mostra este facto é dada por

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

As contas que validam a igualdade $U^* T_1 U = T_2$ são deixadas como exercício. Deste exemplo, nomeadamente do facto de serem válidas ambas as igualdades

$$U^* T_1 U = T_2 \quad \text{e} \quad I^* T_1 I = T_1,$$

é possível concluir que nem as matrizes unitárias nem as matrizes triangulares do Teorema de Schur são únicas.

A demonstração do Teorema de Schur pode ser feita em aritmética real se a matriz $A$ tiver elementos reais. Neste caso obtém-se

$$Q^\top A Q = T,$$

em que $T$ é uma matriz triangular superior e $Q$ é uma matriz ortogonal ($A, Q, T \in M_{n\times n}(\mathbb{R})$).

### 1.5.2 Matrizes diagonalizáveis unitariamente

Como se demonstrará mais adiante as matrizes unitariamente diagonalizáveis são conhecidas por matrizes normais.

**Definição 1.5.4** *Uma matriz $A \in M_{n\times n}(\mathbb{C})$ diz-se normal se*

$$A^* A = A A^*,$$

*ou seja, se $A$ comutar com a sua adjunta ou transconjugada.*

São várias as matrizes normais já conhecidas.

- Todas as matrizes hermíticas são normais.
- Todas as matrizes unitárias são normais.
- Todas as matrizes hemi-hermíticas ($A^* = -A$) são normais.

A demonstração destes factos é simples. Por exemplo, se $U$ é unitária, então $U^*U = I = UU^*$, como vem directamente da definição de matriz unitária. Os outros casos são deixados como exercício.

Estas três classes de matrizes (hermíticas, unitárias, hemi-hermíticas) não esgotam todas as matrizes normais. A matriz

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

é normal mas não pertence a nenhuma destas classes. As contas ficam, uma vez mais, ao cuidado do leitor.

Por outro lado, apresentamos um exemplo de uma matriz que não é normal ($A^*A \neq AA^*$):

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

Neste exemplo, tem-se que

$$A^*A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad AA^* = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

A importância das matrizes normais resulta, como foi dito anteriormente, do facto destas coincidirem com as matrizes que admitem uma diagonalização através de uma matriz diagonalizante unitária.

**Teorema 1.5.2** *Uma matriz $A \in M_{n\times n}(\mathbb{C})$ é normal se e só se existir uma matriz unitária $U \in M_{n\times n}(\mathbb{C})$ tal que*

$$U^*AU = D \quad \textit{com} \quad D \in M_{n\times n}(\mathbb{C}) \textit{ diagonal.}$$

**Demonstração.** Comecemos por provar a implicação mais imediata. Assuma-se, assim, que $A$ admite uma diagonalização unitária. Utilizando um argumento já conhecido, prova-se, a partir de $U^*AU = D$, que $A = UDU^*$. Daqui vem que

$$A^*A = (UDU^*)^* UDU^* = (U^*)^* D^*(U^*U)DU^* = U\,D^*D\,U^*.$$

Da mesma forma, provar-se-ia que

$$AA^* = \cdots = U\,DD^*\,U^*.$$

Esta parte da demonstração fica rematada pelo facto de $D^*D = DD^*$ (uma vez que o produto de matrizes diagonais é comutativo).

A implicação recíproca ($A$ normal então $A$ unitariamente diagonalizável) baseia-se no Teorema de Schur. Este último resultado é aplicável a qualquer matriz quadrada e, em particular, a uma matriz normal. Deste modo, temos que

$$T = U^*AU$$

com $T$ matriz triangular superior e $U$ matriz unitária, ambas em $M_{n\times n}(\mathbb{C})$. Se soubéssemos que $T$ era diagonal não haveria mais nada a provar. O resto desta demonstração resume-se, então, a mostrar que a matriz $T$ é diagonal.

Em primeiro lugar constatamos que $T$ é uma matriz normal. Efectuando um tipo de cálculos que começam a ser rotineiros, mostra-se que

$$TT^* = U^*\,AA^*\,U = U^*\,A^*A\,U = T^*T,$$

com a igualdade do meio a ser verdadeira pelo facto de se estar a assumir, nesta implicação recíproca, que $A$ é uma matriz normal.

A igualdade $T^*T = TT^*$, quando escrita por extenso, apresenta a forma

$$\begin{bmatrix} \overline{t_{11}} & & & \\ \overline{t_{12}} & \overline{t_{22}} & & \\ \vdots & \vdots & \ddots & \\ \overline{t_{1n}} & \overline{t_{2n}} & \cdots & \overline{t_{nn}} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ & t_{22} & \cdots & t_{2n} \\ & & \ddots & \vdots \\ & & & t_{nn} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ & t_{22} & \cdots & t_{2n} \\ & & \ddots & \vdots \\ & & & t_{nn} \end{bmatrix} \begin{bmatrix} \overline{t_{11}} & & & \\ \overline{t_{12}} & \overline{t_{22}} & & \\ \vdots & \vdots & \ddots & \\ \overline{t_{1n}} & \overline{t_{2n}} & \cdots & \overline{t_{nn}} \end{bmatrix}.$$

Efectuando, em ambos os membros, o produto da primeira linha pela primeira coluna, vem que

$$\overline{t_{11}}t_{11} \;=\; t_{11}\overline{t_{11}} + t_{12}\overline{t_{12}} + \cdots + t_{1n}\overline{t_{1n}}$$

ou seja, que

$$0 \;=\; |t_{12}|^2 + \cdots + |t_{1n}|^2.$$

Deste primeiro cálculo obtemos que $t_{12} = \cdots = t_{1n} = 0$. Um segundo cálculo (segunda linha pela segunda coluna em ambos os membros), permite escrever

$$|t_{12}|^2 + |t_{22}|^2 \;=\; |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2.$$

Mas, como já se sabe que $t_{12} = 0$, esta igualdade reduz-se a

$$0 \;=\; |t_{23}|^2 + \cdots + |t_{2n}|^2,$$

o que implica a nulidade dos elementos $t_{23}, \ldots, t_{2n}$. Repetindo este argumento (até ao produto entre as linhas $n-1$ e as colunas $n-1$ de $T^*T$ e $TT^*$) provar-se-ia, sucessivamente, que todos os elementos não diagonais de $T$ são nulos. ■

Pelo facto de serem normais, as matrizes hermíticas admitem, à luz deste teorema, uma diagonalização unitária. Além disso, as matrizes hermíticas apresentam uma diagonalização com elementos diagonais reais, ou seja, têm valores próprios reais. Estes dois factos são resumidos no seguinte corolário.

**Corolário 1.5.1** *Se $A \in M_{n\times n}(\mathbb{C})$ for uma matriz hermítica então é unitariamente diagonalizável com valores próprios reais.*

**Demonstração.** A matriz $A$ é normal porque é hermítica. Logo, pelo teorema anterior, existe uma matriz diagonalizante unitária $U \in M_{n\times n}(\mathbb{C})$ tal que

$$U^*AU \;=\; D \quad \text{com} \quad D \in M_{n\times n}(\mathbb{C}) \text{ diagonal.}$$

Se transconjugarmos $D$ obtemos, ao aplicarmos $A^* = A$, que

$$D^* \;=\; (U^*AU)^* \;=\; U^*A^*(U^*)^* \;=\; U^*AU \;=\; D,$$

mostrando que os elementos diagonais de $D$ (os valores próprios de $A$) são números reais. ■

O facto das matrizes simétricas reais terem valores próprios reais e serem diagonalizáveis através de matrizes diagonalizantes ortogonais é um caso particular deste corolário.

A diagonalização de matrizes pode ser resumida através do seguinte esquema.

$$\begin{array}{c} \text{matrizes diagonalizáveis em } M_{n\times n}(\mathbb{C}) \\ \Uparrow \\ \text{matrizes unitariamente diagonalizáveis (ou normais) em } M_{n\times n}(\mathbb{C}) \\ \Uparrow \\ \text{matrizes hermíticas em } M_{n\times n}(\mathbb{C}) \\ \Uparrow \\ \text{matrizes simétricas em } M_{n\times n}(\mathbb{R}) \end{array}$$

## Exercícios

1. Sejam $U_1, \ldots, U_m$ matrizes unitárias em $M_{n\times n}(\mathbb{C})$ (com $m$ um número inteiro positivo). Prove que o seu produto $U = U_1 \cdots U_m$ é, também, uma matriz unitária.

2. Seja $U$ uma matriz unitária em $M_{n_2\times n_2}(\mathbb{C})$. Mostre que também é unitária a matriz

$$\begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix},$$

em que $I$ é a matriz identidade de $M_{n_1\times n_1}(\mathbb{C})$ (e $n_1$ e $n_2$ números inteiros positivos).

3. Destes dois primeiros exercícios conclua que a matriz $U$ da demonstração do teorema de Schur é unitária.

4. Mostre que se $U \in M_{n\times n}(\mathbb{C})$ for uma matriz unitária então também são unitárias as matrizes $\bar{U}$, $U^\top$ e $U^*$.

5. Demonstre que se $U \in M_{n\times n}(\mathbb{C})$ for uma matriz unitária então [$x$ e $y$ são ortogonais se e só se $Ux$ e $Uy$ são ortogonais, qualquer que sejam $x, y \in \mathbb{C}^n$].

6. Utilizando as propriedades dos determinantes, mostre que o módulo do determinante de uma matriz unitária vale sempre um.

7. Utilizando as propriedades dos determinantes, prove que duas matrizes unitariamente semelhantes têm os mesmos valores próprios.

8. Classifique as seguintes matrizes, indicando quais são: simétricas; hermíticas; ortogonais; unitárias; normais.

$$\text{(a)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(c)} \begin{bmatrix} 1 & i \\ i & 0 \end{bmatrix} \quad \text{(d)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}i \\ 0 & -\frac{\sqrt{2}}{2}i & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

9. Considere a seguinte matriz em $M_{3\times 3}(\mathbb{C})$:

$$A = \begin{bmatrix} 1 & 0 & i \\ 0 & 1 & 0 \\ a_{31} & 0 & 1 \end{bmatrix}$$

em que $a_{31}$ é um parâmetro complexo.

   (a) Indique, se possível, valores para $a_{31}$ de forma a que $A$ seja: simétrica; hermítica; ortogonal; unitária; normal.

   (b) Faça $a_{31} = 0$. Indique os valores próprios de $A$.

   (c) A matriz $A$ é diagonalizável (com $a_{31} = 0$)?

10. Prove que se $A$ for hermítica (respectivamente hemi-hermítica) então $iA$ é hemi-hermítica (respectivamente hermítica).

11. Mostre que a matriz $A \in M_{n\times n}(\mathbb{C})$ é normal se e só se $\|Ax\| = \|A^*x\|$ para todo o $x \in \mathbb{C}^n$. (As matrizes normais são aquelas em que $A$ e $A^*$ preservam distâncias entre si.)

12. Mostre que a matriz $A \in M_{n\times n}(\mathbb{C})$ é normal se e só se $(Ax)^*(Ax) = (A^*x)^*(A^*x)$ para todo o $x \in \mathbb{C}^n$. (As matrizes normais são aquelas em que $A$ e $A^*$ preservam ângulos entre si.)

13. Prove que se $A \in M_{n\times n}(\mathbb{C})$ for uma matriz normal então

$$Ax = 0 \quad \text{se e só se} \quad A^*x = 0.$$

(Quando $A$ é normal o seu espaço nulo coincide com o da sua adjunta.)

Mostre, ainda, que a implicação recíproca não é verdadeira, recorrendo à matriz

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

## 1.6 O caso das matrizes circulantes

Uma das classes de matrizes normais mais importantes em aplicações é formada pelas matrizes circulantes. Mostrar-se-á que uma matriz circulante é normal ao exibir-se a sua diagonalização unitária, utilizando, assim, o resultado da secção anterior. As matrizes circulantes aparecem como matrizes de sistemas de equações lineares de diversos problemas relacionados com o processamento de sinal e imagem.

Uma matriz circulante $n$-por-$n$ é definida por um vector de $n$ componentes. A sua primeira coluna é definida por este vector. A segunda coluna define-se deslocando, inferiormente, as componentes do vector, trazendo para a primeira posição a última componente do vector. A terceira coluna da matriz forma-se aplicando à segunda coluna um procedimento análogo ao que transformou a primeira na segunda. E assim se formariam, sucessivamente, as restantes colunas.

Por exemplo, no caso $n = 4$, uma matriz circulante apresenta a seguinte forma

$$H = \begin{bmatrix} h_0 & h_3 & h_2 & h_1 \\ h_1 & h_0 & h_3 & h_2 \\ h_2 & h_1 & h_0 & h_3 \\ h_3 & h_2 & h_1 & h_0 \end{bmatrix}.$$

O vector que define esta matriz é $h = [\,h_0 \;\; h_1 \;\; h_2 \;\; h_3\,]^\top \in \mathbb{C}^4$. Por uma questão de conveniência, relacionada com a notação desta secção, a primeira componente $h_0$ de $h$ é identificada com o índice 0. Vemos, então, como os elementos das colunas desta matriz vão sendo deslocados inferiormente, num movimento do tipo circular.

As matrizes circulantes surgem, também, da *discretização* de equações diferenciais. A título ilustrativo, considere-se o seguinte problema de Poisson unidimensional:

$$-u''(x) = f(x), \quad x \in\, ]0,1[, \tag{1.1}$$
$$u(0) = u(1), \tag{1.2}$$
$$u'(0) = u'(1), \tag{1.3}$$

em que $f$ é uma função definida em $[0,1]$ (que se assume conhecida no contexto do problema em causa). Pretende-se descobrir a função $u$, definida em $[0,1]$, que verifica a equação diferencial (1.1) e as *condições de fronteira periódicas* (1.2) e (1.3).

Sejam $n$ um número inteiro não inferior a 2 e $p$ um número racional dado por $p = 1/(n-1)$. Considere-se um vector $\bar{u} \in \mathbb{R}^n$ cujas componentes $\bar{u}_k$, $k = 0, 1, \ldots, n-1$, são *aproximações* do valor $u(kp)$ da função $u$ no ponto $kp$. Um esquema de *discretização* permite calcular os elementos de $\bar{u}$ através da solução de um sistema de equações lineares.

Vamos considerar uma discretização de (1.1)-(1.3) definida pelas equações lineares

$$-\bar{u}_{k-1} + d\bar{u}_k - \bar{u}_{k+1} = p^2 f(kp), \quad k = 0, \ldots, n-1, \tag{1.4}$$
$$\bar{u}_{-1} = \bar{u}_{n-1}, \tag{1.5}$$
$$\bar{u}_n = \bar{u}_0, \tag{1.6}$$

com $d$ um parâmetro real.

Este sistema de equações lineares (1.4)-(1.6) escreve-se, eliminando os valores de $\bar{u}_{-1}$ e de $\bar{u}_n$, na seguinte forma matricial

$$\begin{bmatrix} d & -1 & 0 & \cdots & 0 & 0 & -1 \\ -1 & d & -1 & & 0 & 0 & 0 \\ 0 & -1 & d & & 0 & 0 & 0 \\ \vdots & & & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & d & -1 & 0 \\ 0 & 0 & 0 & & -1 & d & -1 \\ -1 & 0 & 0 & & 0 & -1 & d \end{bmatrix} \begin{bmatrix} \bar{u}_0 \\ \bar{u}_1 \\ \bar{u}_2 \\ \vdots \\ \bar{u}_{n-3} \\ \bar{u}_{n-2} \\ \bar{u}_{n-1} \end{bmatrix} = p^2 \begin{bmatrix} f(0) \\ f(p) \\ f(2p) \\ \vdots \\ f((n-3)p) \\ f((n-2)p) \\ f(1) \end{bmatrix}.$$

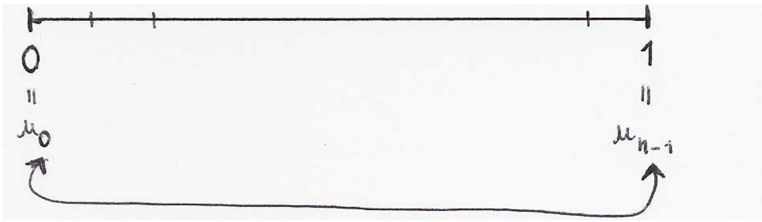



Figura 1.6.1: Discretização do intervalo $[0,1]$ em $n-1$ subintervalos de igual amplitude.

Constata-se, imediatamente, que a matriz deste sistema é circulante. Se $n$ for igual a 4 então esta matriz é

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

É deixada ao leitor, de seguida, a definição formal de matriz circulante.

**Definição 1.6.1** *Uma matriz circulante $H \in M_{n\times n}(\mathbb{C})$ é uma matriz da forma*

$$H = \begin{bmatrix} h_0 & h_{n-1} & \cdots & h_1 \\ h_1 & h_0 & \cdots & h_2 \\ h_2 & h_1 & \cdots & h_3 \\ \vdots & \vdots & & \vdots \\ h_{n-2} & h_{n-3} & \cdots & h_{n-1} \\ h_{n-1} & h_{n-2} & \cdots & h_0 \end{bmatrix}.$$

O vector que define uma matriz circulante aparece como a sua primeira coluna:

$$h = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{n-2} \\ h_{n-1} \end{bmatrix}.$$

### 1.6.1 A matriz de permutação de deslocamento inferior

Entre as mais conhecidas matrizes circulantes encontram-se as matrizes de permutação de deslocamento inferior (PDI). Quando $n=4$, a matriz de PDI é dada por

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

O vector $h$ que define esta matriz circulante é $[0\ 1\ 0\ 0]^\top$. No caso geral, o vector $h$ coincide com a segunda coluna da matriz identidade de ordem $n$.

**Definição 1.6.2** *Uma matriz de permutação de deslocamento inferior (PDI) é uma matriz circulante $C \in M_{n\times n}(\mathbb{R})$ da forma*

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

Estas matrizes de PDI apresentam algumas propriedades atraentes. Por exemplo, se calcularmos $C^2$, obtemos, no caso $n = 4$,

$$C^2 = C\,C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

que continua a ser uma matriz de permutação. Se continuarmos, temos

$$C^3 = C\,C^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

e

$$C^4 = C\,C^3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Observa-se, assim, que $C^4 = I$. Por definição, tem-se que $C^0 = I$. Provar-se-ia, facilmente, no caso geral, que $C^n = I$.

**Proposição 1.6.1** *Dada uma matriz $C$ de PDI em $M_{n\times n}(\mathbb{R})$, tem-se que*

$$C^n = I = C^0.$$

O efeito de multiplicar um vector, à esquerda, por $C$ consiste em *deslocar*, *inferiormente*, os elementos do vector, passando o último a tomar a posição do primeiro. Veja-se o seguinte exemplo:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ h_3 \end{bmatrix} = \begin{bmatrix} h_3 \\ h_0 \\ h_1 \\ h_2 \end{bmatrix}.$$

(Este efeito foi já visível nos produtos $C\,C$, $C\,C^2$ e $CC^3$ acima descritos...) Surge, assim, com naturalidade, o seguinte resultado, cuja demonstração é deixada como exercício.

**Proposição 1.6.2** *Dada uma matriz circulante $H$ em $M_{n\times n}(\mathbb{C})$, definida pelo vector $h$, tem-se que*

$$H \;=\; [\,C^0h\; Ch\; \cdots\; C^{n-1}h\,],$$

*em que $C$ é a matriz de PDI em $M_{n\times n}(\mathrm{I\!R})$.*

É possível escrever, de uma forma diferente da anterior, uma matriz circulante em função das potências de uma matriz de PDI. Vejamos o caso $n = 3$ para depois enunciarmos o caso geral.

$$\begin{bmatrix} h_0 & h_2 & h_1 \\ h_1 & h_0 & h_2 \\ h_2 & h_1 & h_0 \end{bmatrix} \;=\; h_0 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + h_1 \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + h_2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

**Proposição 1.6.3** *Dada uma matriz circulante $H$ em $M_{n\times n}(\mathbb{C})$, definida pelo vector $h$, tem-se que*

$$H \;=\; h_0C^0 + h_1C + \cdots + h_{n-1}C^{n-1},$$

*em que $C$ é a matriz de PDI em $M_{n\times n}(\mathrm{I\!R})$.*

Esta última forma de exprimir uma matriz circulante torna-se útil se soubermos, de antemão, uma diagonalização para $C$. Em jeito de antecipação, veja-se que se $VCV^{-1} = D$ e $n = 3$, então

$$VHV^{-1} \;=\; V(h_0C^0 + h_1C + h_2C^2)V^{-1} \;=\; h_0VC^0V^{-1} + h_1VCV^{-1} + h_2VC^2V^{-1}.$$

Como $VC^2V^{-1} = D^2$, vem que

$$VHV^{-1} \;=\; h_0D^0 + h_1D + h_2D^2.$$

Concluímos que $V$ também diagonaliza $H$ e que os valores próprios de $H$ são os elementos diagonais da matriz diagonal $h_0I + h_1D + h_2D^2$.

### 1.6.2 A matriz da transformada discreta de Fourier

Conforme veremos mais à frente, a matriz diagonalizante de uma matriz circulante é um múltiplo de uma matriz designada por matriz da transformada discreta de Fourier.

**Definição 1.6.3** *A matriz da transformada discreta de Fourier (TDF) é a matriz*

$$F \;=\; \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & \cdots & w^{n-1} \\ 1 & w^2 & w^4 & \cdots & w^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)^2} \end{bmatrix} \in M_{n\times n}(\mathbb{C}),$$

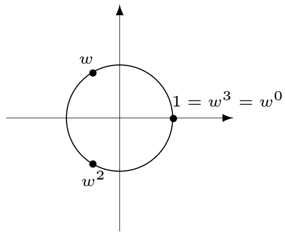



Figura 1.6.2: A distribuição das raízes da equação $w^3 = 1$ no círculo unitário.

*em que*

$$w \;=\; e^{\frac{i2\pi}{n}} \;=\; \text{cis}\left(\frac{2\pi}{n}\right) \;=\; \cos\left(\frac{2\pi}{n}\right) + \text{sen}\left(\frac{2\pi}{n}\right) i.$$

Repare-se que $w$ está no círculo unitário $\{w \in \mathbb{C} : \; |w| = 1|\}$ do plano complexo, fazendo um ângulo de $2\pi/n$ com a horizontal ou, por outras palavras, $\arg(w) = 2\pi/n$.

Para $n = 2$ resultam as igualdades:

$$w \;=\; e^{i\pi} \;= -1 \qquad \text{e} \qquad F \;=\; \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Para $n = 3$, vem que

$$w \;=\; \cos\left(\frac{2\pi}{3}\right) + \text{sen}\left(\frac{2\pi}{3}\right) i.$$

Logo,

$$w^2 \;=\; \left(e^{\frac{i2\pi}{3}}\right)^2 \;=\; e^{\frac{i4\pi}{3}} \;=\; \cos\left(\frac{4\pi}{3}\right) + \text{sen}\left(\frac{4\pi}{3}\right) i$$

e

$$w^3 \;=\; \left(e^{\frac{i2\pi}{3}}\right)^3 \;=\; e^{i2\pi} \;=\; 1.$$

Os números complexos $1 = w^0$, $w$ e $w^2$ distribuem-se no círculo unitário de acordo com a Figura 1.6.2.

A matriz da TDF, quando $n = 3$, é dada por

$$F \;=\; \begin{bmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w^4 \end{bmatrix} \;=\; \begin{bmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w \end{bmatrix}.$$

Neste último exemplo, utilizou-se a igualdade $w^3 = 1$ para derivar $w^4 = ww^3 = 1$. De uma forma geral, temos que

$$w^n \;=\; \left(e^{\frac{i2\pi}{n}}\right)^n \;=\; e^{i2\pi} \;=\; 1,$$

o que permite escrever uma propriedade para $w$ em tudo semelhante à enunciada para a matriz de PDI na Proposição 1.6.1.

**Proposição 1.6.4** *Se* $w = e^{\frac{i2\pi}{n}}$ *então*

$$w^n \;=\; 1 \;=\; w^0.$$

Apresenta-se, ainda, o caso $n = 4$, em que

$$w \;=\; e^{\frac{i\pi}{2}} \;=\; i, \quad w^2 \;=\; -1, \quad w^3 \;=\; -i, \quad w^4 = 1,$$

e

$$F \;=\; \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 \\ 1 & w^2 & w^4 & w^6 \\ 1 & w^3 & w^6 & w^9 \end{bmatrix} \;=\; \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}.$$

A matriz $F$ da TDF é não singular. A sua inversa $F^{-1}$ é um múltiplo da sua conjugada $\bar{F}$ (que coincide com a sua transconjugada $F^*$ por $F$ ser simétrica), conforme se prova de seguida.

**Teorema 1.6.1** *A matriz da TDF de ordem* $n$ *satisfaz*

$$F\bar{F} \;=\; nI.$$

**Demonstração.** O elemento $s_{pq}$ de $F\bar{F}$ na posição $p, q$ pode ser escrito na forma

$$\begin{aligned} s_{pq} \;&=\; \sum_{k=0}^{n-1}(w^{k(p-1)})(\overline{w^{k(q-1)}}) \;=\; \sum_{k=0}^{n-1}(w^{k(p-1)})(\bar{w}^{k(q-1)}) \\ &=\; \sum_{k=0}^{n-1}(w^{k(p-1)})(w^{-k(q-1)}) \;=\; \sum_{k=0}^{n-1} w^{k(p-q)}. \end{aligned}$$

Quando $p = q$ vem que

$$s_{pq} \;=\; \sum_{k=0}^{n-1} w^0 \;=\; n,$$

mostrando o pretendido para os elementos diagonais de $F\bar{F}$.

Para provar que os elementos não diagonais de $F\bar{F}$ são nulos, multipliquemos $s_{pq}$ por $1 - w^{p-q}$. Como estamos interessados, agora, nas posições $p \neq q$, podemos utilizar o facto de $1 - w^{p-q}$ ser diferente de zero. Assim, $s_{pq}$ é nulo se e só se $(1 - w^{p-q})s_{pq}$ o for. Recorrendo à expressão para $s_{pq}$ provada acima, resultam os cálculos

$$(1 - w^{p-q})s_{pq} \;=\; \sum_{k=0}^{n-1} w^{k(p-q)} - \sum_{k=1}^{n} w^{k(p-q)} \;=\; w^0 - w^{n(p-q)} \;=\; 1 - 1^{p-q} \;=\; 0,$$

com os quais a demonstração fica concluída. ■

Ficámos a saber, deste modo, que $F^{-1} = (1/n)\bar{F} = (1/n)F^*$. O produto entre $F$ e $F^*(= \bar{F})$ não é igual à matriz identidade, mas não fica muito longe desse objectivo. Escalonando $F$ de forma apropriada obtém-se uma matriz unitária.

**Corolário 1.6.1** *A matriz*
$$U = \frac{1}{\sqrt{n}}F$$
*é unitária (em que $F$ é a matriz da TDF de ordem $n$).*

### 1.6.3 Diagonalização de matrizes circulantes

A diagonalização de matrizes circulantes assenta na diagonalização de matrizes de PDI. O lema seguinte dá-nos uma primeira pista sobre esta diagonalização.

**Lema 1.6.1** *A matriz $C$ de PDI e a matriz $F$ da TDF satisfazem a seguinte igualdade*
$$FC = DF \qquad (\textit{ equivalente a } FCF^{-1} = D),$$
*onde $D \in M_{n\times n}(\mathbb{C})$ é a matriz diagonal definida por*
$$D = \begin{bmatrix} 1 & & & & \\ & w & & & \\ & & w^2 & & \\ & & & \ddots & \\ & & & & w^{n-1} \end{bmatrix}.$$

Vamos ver o que acontece quando $n = 3$. Verifica-se, facilmente, que
$$FC = \begin{bmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w^4 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ w & w^2 & 1 \\ w^2 & w & 1 \end{bmatrix}$$
e
$$DF = \begin{bmatrix} 1 & 0 & 0 \\ 0 & w & 0 \\ 0 & 0 & w^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w^4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ w & w^2 & 1 \\ w^2 & w & 1 \end{bmatrix}.$$
A demonstração do caso geral, que se deixa como exercício, segue os cálculos do caso $n = 3$ e baseia-se no facto de $w^n = 1$.

Neste momento podemos afirmar que $C$ admite uma diagonalização unitária da forma
$$UCU^* = D \quad \text{com} \quad U = \frac{1}{\sqrt{n}}F.$$
A matriz $C$ é, consequentemente, uma matriz normal (ver Teorema 1.5.2).

A diagonalização unitária de uma matriz circulante $H$ resulta do desenvolvimento de $H$ em potências da matriz $C$ e da sua diagonalização unitária.

**Teorema 1.6.2** *Seja $H \in M_{n\times n}(\mathbb{C})$ uma matriz circulante definida pelo vector $h = [\, h_0\; h_1\; \cdots h_{n-1}\,]^\top \in \mathbb{C}^n$. Então*

$$FHF^{-1} \;=\; \Lambda,$$

*onde $F$ é a matriz da TDF e $\Lambda \in M_{n\times n}(\mathbb{C})$ é a matriz diagonal definida por*

$$\Lambda \;=\; \sum_{k=0}^{n-1} h_k \begin{bmatrix} 1 & & & & \\ & w & & & \\ & & w^2 & & \\ & & & \ddots & \\ & & & & w^{n-1} \end{bmatrix}^k .$$

**Demonstração.** Recorrendo à Proposição 1.6.3, ao Lema 1.6.1 e às propriedades do produto de matrizes, escreve-se

$$FHF^{-1} \;=\; F\left(\sum_{k=0}^{n-1} h_k C^k\right)F^{-1} \;=\; \sum_{k=0}^{n-1} h_k F C^k F^{-1} \;=\; \sum_{k=0}^{n-1} h_k D^k,$$

onde $D$ é a matriz diagonal formada pelos elementos diagonais $1, w, w^2, \ldots, w^{n-1}$. ■

Conclui-se, também, que qualquer matriz circulante $H$ admite uma diagonalização unitária da forma

$$UHU^* \;=\; \Lambda \quad \text{com} \quad U \;=\; \frac{1}{\sqrt{n}}F.$$

Logo, qualquer matriz circulante é normal (ver Teorema 1.5.2).

### 1.6.4 Resolução de sistemas com matrizes circulantes

Conhecida a forma de diagonalizar (unitariamente) uma matriz circulante, vamos, agora, ver como é que se resolve um sistema de equações lineares com este tipo de matrizes. Considere-se um sistema de equações lineares escrito na forma

$$Hg \;=\; f,$$

em que

$$H \;=\; \begin{bmatrix} h_0 & h_{n-1} & \cdots & h_1 \\ h_1 & h_0 & \cdots & h_2 \\ h_2 & h_1 & \cdots & h_3 \\ \vdots & \vdots & & \vdots \\ h_{n-2} & h_{n-3} & \cdots & h_{n-1} \\ h_{n-1} & h_{n-2} & \cdots & h_0 \end{bmatrix}, \quad g \;=\; \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_{n-2} \\ g_{n-1} \end{bmatrix} \quad \text{e} \quad f \;=\; \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{bmatrix}.$$

Dada a matriz $H$ e o termo independente $f$, pretende-se calcular $g$.

**Comentário 1.6.1** *A notação $H$, $g$ e $f$ não foi escolhida ao acaso. O termo independente $f$ é, frequentemente, a discretização de uma função $f$ (por exemplo, um sinal). A solução $g$ pode, também, corresponder à discretização de uma função, como um sinal processado.*

*O produto $d = F^{-1}c$ designa-se por transformada discreta de Fourier (TDF) de $c$, o que, diga-se de passagem, está em sintonia com a designação dada a $F$. É possível estabelecer um paralelo entre a TDF e a transformada contínua de Fourier. Para entendermos melhor o que se está a passar, façamos*

$$c \;=\; \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} \;\in\; \mathbb{R}^n \quad e \quad d \;=\; \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{n-1} \end{bmatrix} \;\in\; \mathbb{R}^n.$$

*Tem-se, então, que a $j$-ésima componente do vector $d$ é expressa como*

$$d_j \;=\; (F^{-1}c)_j \;=\; \frac{1}{n}(\bar{F}c)_j \;=\; \frac{1}{n}\sum_{k=0}^{n-1} c_k\overline{w^{jk}},$$

*ou seja, como*

$$d_j \;=\; \frac{1}{n}\sum_{k=0}^{n-1} c_k e^{-\frac{i2\pi}{n}jk}.$$

*Começa a tomar forma, deste modo, a analogia entre $d_j$ e a transformada contínua de Fourier da função cuja discretização no intervalo $[0, 2\pi]$ é armazenada pelo vector $c$. Considere-se que $c_k = c(x_k)$ em que $x_k = k(2\pi)/(n-1)$, $k = 0, \ldots, n-1$, e $c$ é uma função definida em $[0, 2\pi]$. Assumindo as devidas hipóteses de integrabilidade sobre a função $c$ e tomando $n$ a tender para $+\infty$, obter-se-ia*

$$d_j \;=\; \frac{1}{2\pi}\int_0^{2\pi} c(x)e^{-ijx}dx.$$

Regressamos ao sistema de equações lineares $Hg = f$. São verdadeiras as seguintes equivalências:

$$\begin{aligned} Hg = f \;&\Longleftrightarrow\; FHg = Ff \\ &\Longleftrightarrow\; (FHF^{-1})(Fg) = Ff \\ &\Longleftrightarrow\; \Lambda F_g = F_f, \end{aligned}$$

com

$$F_g \;=\; Fg \quad \text{e} \quad F_f \;=\; Ff.$$

Se $H$ for não singular então $\Lambda$ é não singular e

$$Hg \;=\; f \quad\Longleftrightarrow\quad F_g \;=\; \Lambda^{-1}F_f$$

ou, em termos da solução $g$,

$$Hg = f \iff g = F^{-1}\Lambda^{-1}F_f.$$

A inversa de $F$ é dada por $F^{-1} = (1/n)\bar{F}$. Organizamos estes cálculos, então, na forma do seguinte algoritmo.

**Algoritmo 1.6.1 [Resolver $Hg = f$ com $H$ circulante e não singular.]**

1. *Calcular* $F_f = Ff$.

2. *Calcular* $F_g = \Lambda^{-1}F_f$.

3. *Calcular a transformada discreta de Fourier de* $F_g$*:*

$$g = \frac{1}{n}\bar{F}F_g.$$

Os Passos 1 e 3 deste algoritmo consistem em produtos entre matrizes e vectores. Constata-se, facilmente, que o número de operações elementares (adições, subtracções, multiplicações e divisões) para efectuar o produto entre uma matriz em $M_{n\times n}(\mathbb{C})$ e um vector em $\mathbb{C}^n$ é um polinómio de grau 2 em $n$. É possível mostrar, aliás, que o primeiro termo deste polinómio é $8n^2$.

O Passo 2 do Algoritmo 1.6.1 resume-se a $n$ divisões entre números complexos (uma vez que a matriz $\Lambda$ é diagonal). Assim sendo, o número total de operações elementares que este algoritmo faz para resolver o sistema circulante $Hg = f$ é da ordem de $n^2$ (ou seja, é um polinómio de grau 2 em $n$). Como mencionaremos no parágrafo seguinte, este desempenho pode, ainda, ser melhorado. Mas mesmo um número de operações da ordem de $n^2$ constitui uma significativa melhoria em relação a um número da ordem de $n^3$, que seria necessário se fosse utilizado o método de eliminação de Gauss.

Os produtos complexos matriz-vector envolvendo as matrizes $F$ ou $\bar{F}$ podem ser organizados de forma a que as operações elementares passem a ser, em número, dominadas por $5n\log_2 n$. Este tipo de cálculos é conhecido por transformada rápida de Fourier, em inglês *fast Fourier transform (FFT)*. Os ganhos de $5n\log_2 n$ em relação a $8n^2$ são maiores do que aparentam. A Tabela 1.1 ilustra o quociente entre $8n^2$ e $5n\log_2 n$ para diversos valores de $n$.

O desempenho computacional dos algoritmos baseados em FFTs está na base da sua ampla utilização em equipamentos que processam sinal ou imagem.

Tabela 1.1: Ganhos de um algoritmo baseado em FFTs em relação ao Algoritmo 1.6.1.

| $n$ | $\frac{8n^2}{5n\log_2 n}$ |
|---|---|
| 32 | $\simeq 10$ |
| 1024 | $\simeq 160$ |
| 32768 | $\simeq 3500$ |
| 1048574 | $\simeq 84000$ |

## Exercícios

1. Escreva a matriz de permutação de deslocamento inferior $C$, quando $n = 6$. Classifique esta matriz, indicando se é: simétrica; hermítica; ortogonal; unitária; normal; circulante. Escreva $C^3$ sem efectuar quaisquer cálculos.

2. Identifique a matriz de permutação de deslocamento superior e mostre o seu efeito quando multiplica, à esquerda, um vector de dimensão apropriada.

3. Escreva a matriz da transformada discreta de Fourier $F$, quando $n = 6$, em termos de $w = e^{i\pi/3}$ e das suas potências ($w^0$, $w^1$, $w^2$, $w^3$, $w^4$ e $w^5$).

4. Considere a seguinte matriz em $M_{3\times 3}(\mathbb{C})$:

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

   (a) Classifique esta matriz (indicando todas as propriedades matriciais conhecidas que ela satisfaça).

   (b) Escreva esta matriz como uma combinação linear de $C^0$, $C$ e $C^2$, em que $C$ designa a matriz de permutação de deslocamento inferior de ordem 3.

   (c) Indique os valores próprios de $C$ em função de $w = e^{2\pi i/3}$.

   (d) Indique os valores próprios de $A$ em função de $w = e^{2\pi i/3}$.

5. Considere a seguinte matriz em $M_{3\times 3}(\mathbb{C})$:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 1 & i & -1 \end{bmatrix}.$$

   (a) Complete esta matriz de forma a ser circulante.

   (b) Complete esta matriz de forma a ter valores próprios reais.

   (c) Complete esta matriz de forma a ser unitariamente diagonalizável.

(d) No caso da alínea (a), escreva a matriz na forma $[\,h \;\; C^7h \;\; C^5h\,]$, em que $C$ é a matriz de permutação de deslocamento inferior de ordem 3 e $h$ um vector em $\mathbb{C}^3$.

(e) Indique os valores próprios de $C^5$.

6. Escreva a matriz $U = (1/\sqrt{n})F$ quando $n = 4$, em que $F$ é a matriz da TDF. Mostre que as suas segunda e terceira colunas são ortogonais.

7. Calcule os valores e vectores próprios da matriz $C$ de PDI quando $n = 4$.

8. Considere a matriz circulante

$$H = \begin{bmatrix} d & -1 & 0 & -1 \\ -1 & d & -1 & 0 \\ 0 & -1 & d & -1 \\ -1 & 0 & -1 & d \end{bmatrix}.$$

(a) Mostre que os seus valores próprios são iguais a

$$\begin{aligned} d - 1 - 1 &= d - 2, \\ d - i - i^{-1} &= d, \\ d - i^2 - i^{-2} &= d + 2, \\ d - i^3 - i^{-3} &= d. \end{aligned}$$

(b) Verifica-se, assim, que os valores próprios de $H$ são reais se $d$ for um número real. Teria sido possível afirmar isto antes de ter calculado os valores próprios de $H$?

(c) Faça, agora, $d = 2$. Diga por que é que a matriz $H$ é singular. Mostre que $[\,1\;1\;1\;1\,]^\top$ é um vector próprio de $H$ associado ao valor próprio $\lambda = 0 = d - 2$.

# Capítulo 2

# Geometria analítica

**Revisões de ALGA:**

Seja $V$ um espaço vectorial real. Um **produto interno** em $V$ é uma operação que a cada par de vectores $x, y$ de $V$ faz corresponder um número real, denotado por $\langle x, y\rangle$ e chamado produto interno de $x$ por $y$, de forma que se verifiquem as seguintes propriedades

1. $\langle x, y\rangle = \langle y, x\rangle$
2. $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle, \quad x' \in V$
3. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle, \quad \alpha \in \mathbb{R}$
4. $x \neq 0 \Longrightarrow \langle x, x\rangle > 0$

Um espaço em que está definido um produto interno diz-se um espaço com produto interno ou **espaço euclidiano**.

Podem considerar-se produtos internos em espaços **complexos**. Nesse caso, a primeira propriedade deve ser substituída por $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

**Exemplos**

1. $\langle x, y\rangle = y^\top x = x_1y_1 + x_2y_2 + \cdots + x_ny_n$ é um produto interno no espaço vectorial real $\mathbb{R}^n$.
2. $\langle x, y\rangle = y^* x = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$ é um produto interno no espaço vectorial complexo $\mathbb{C}^n$.

$V$ espaço vectorial com produto interno, $x, y \in V$.
A **norma** (ou comprimento) de $x$ é $\|x\| = \sqrt{\langle x, x\rangle}$.
A **distância** entre $x$ e $y$ é $\|x - y\|$.
$x$ e $y$ são **ortogonais** (ou perpendiculares) se $\langle x, y\rangle = 0$ (notação: $x \perp y$).
**Desigualdade de Cauchy-Schwartz**: $|\langle x, y\rangle| \leq \|x\|\|y\|$.
**Desigualdade triangular**: $\|x + y\| \leq \|x\| + \|y\|$.
**Teorema de Pitágoras**: Se $x \perp y$ tem-se $\|x - y\|^2 = \|x\|^2 + \|y\|^2$.

A **projecção ortogonal** de um vector $y$ sobre um vector não nulo $x$ num espaço com produto interno define-se como o vector

$$\mathrm{proj}_x y = \frac{\langle y, x\rangle}{\|x\|^2} x.$$

O **ângulo** $\theta$ entre dois vectores não nulos $x$ e $y$ é $\theta = \arccos \dfrac{\langle x, y\rangle}{\|x\|\|y\|}$.

Seja $V$ um espaço euclidiano e $F$ um subespaço de $V$. Ao conjunto de vectores de $V$ que são ortogonais a todos os vectores de $F$ chama-se **complemento ortogonal** de $F$. A notação habitual é $F^\perp$. Simbolicamente:

$$F^\perp = \{v \in V : \langle v, u\rangle = 0, \text{ para todo o } u \in F\}.$$

É imediato verificar que $F^\perp$ é também um subespaço vectorial de $V$ (exercício). Tem-se $(F^\perp)^\perp = F$ e $F \cap F^\perp = \{0\}$.

**Exemplo:** Os subespaços de $\mathbb{R}^3$,

$$X = \mathcal{L}\{e_1\} \text{ e } Y = \mathcal{L}\{e_2\}$$

são ortogonais, mas não são complementos ortogonais um do outro, uma vez que

$$X^\perp = \mathcal{L}\{e_2, e_3\} \text{ e } Y^\perp = \mathcal{L}\{e_1, e_3\}.$$

Se $x \in V$, um vector $x_F$ diz-se a **projecção ortogonal de $x$ sobre $F$** se

$$x - x_F \in F^{\perp}.$$

Relativamente a uma base ortogonal $\{v_1, \ldots, v_n\}$ (vectores ortogonais dois a dois) de $F$ tem-se

$$x_F = \text{proj}_F x = \sum_{i=1}^{n} \frac{\langle x, v_i \rangle}{\|v_i\|^2} v_i = \sum_{i=1}^{n} \text{proj}_{v_i} x.$$

**Definição 2.0.4** *Sejam $F$ e $G$ subespaços do espaço vectorial $V$. Chama-se* ***soma dos subespaços $F$ e $G$*** *ao conjunto*

$$F + G = \{v + w : v \in F \text{ e } w \in G\}.$$

**Exemplo**: Sejam $F$ e $G$ duas rectas em $\mathbb{R}^2$ diferentes que passam pela origem. Verifique geometricamente que $\mathbb{R}^2 = F + G$.

**Teorema 2.0.3** *$F + G$ é ainda um subespaço de $V$.*

**Demonstração.** Exercício.

**Teorema 2.0.4** *dim $(F + G)$ = dim $F$ + dim $G$ - dim $(F \cap G)$.*

**Demonstração.** Sejam dim $F = k$, dim $G = m$, dim $(F \cap G) = s$ e $\{u_1, \ldots, u_s\}$ base de $F \cap G$. Estendamo-la a uma base de $F$, $\{u_1, \ldots, u_s, v_1, \ldots, v_{k-s}\}$ e a uma base de $G$ $\{u_1, \ldots, u_s, w_1, \ldots, w_{m-s}\}$. Resta mostrar que o conjunto

$$\{u_1, \ldots, u_s, v_1, \ldots, v_{k-s}, w_1, \ldots, w_{m-s}\}$$

é uma base de $F + G$ (exercício). ■

Se se tiver $F \cap G = \{0\}$ diz-se que a **soma** de $F$ com $G$ é **directa** e escreve-se $F \oplus G$. Tem-se ainda

$$\text{dim } (F \oplus G) = \text{dim } F + \text{dim } G,$$

e uma base de $F \oplus G$ obtém-se reunindo uma base de $F$ com uma base de $G$.

Se $V = F \oplus G$, dizemos que $F$ e $G$ são **subespaços complementares** ou que $F$ é um complemento de $G$ (e vice-versa).

Tem-se

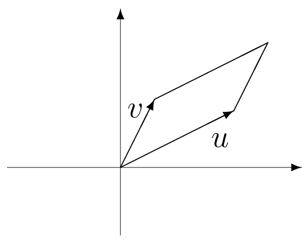
$$\boxed{V = F \oplus F^{\perp}}$$

para qualquer subespaço $F$.

## 2.1 Determinantes e medidas de paralelipípedos

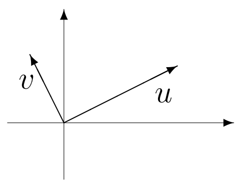
Dados dois vectores $u, v \in \mathbb{R}^2$, o paralelogramo definido por $u$ e $v$, denotado por $P[u, v]$, é o conjunto

$$P[u, v] = \{\alpha u + \beta v : 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1\}.$$




A **área** de $P[u, v]$ é igual a $|\det A|$ onde $A$ é a matriz cujas colunas são $u$ e $v$.

**Caso 1:** $u \perp v$
Neste caso, área de $P[u, v] = \|u\|\|v\|$.
Sendo a matriz $A = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}$, tem-se

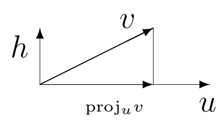



$$A^\top A = \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} = \begin{bmatrix} \langle u, u\rangle & \langle u, v\rangle \\ \langle v, u\rangle & \langle v, v\rangle \end{bmatrix} = \begin{bmatrix} \|u\|^2 & 0 \\ 0 & \|v\|^2 \end{bmatrix},$$

logo $\|u\|^2\|v\|^2 = \det\,(A^\top A) = \det\, A^\top \cdot \det\, A = (\det\, A)^2$. Finalmente vem

$$|\det\, A| = \|u\|\|v\| = \text{área de } P[u, v].$$

**Caso 2:** $u$ e $v$ quaisquer com $u \neq 0$
Neste caso, área de $P[u, v] = \|u\|\|h\| = \|u\|\|v - \text{proj}_u v\|$
Mas se $u \perp h$ então, pelo caso 1,




$$\text{área de } P[u, v] = |\det\, B|,$$

onde $B$ é a matriz cujas colunas são $u$ e $v - \text{proj}_u v$.

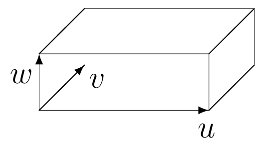
No entanto,

$$\begin{aligned} \det\, \begin{bmatrix} u & v - \text{proj}_u v \end{bmatrix} &= \det\, \begin{bmatrix} u & v \end{bmatrix} - \det\, \begin{bmatrix} u & \text{proj}_u v \end{bmatrix} \\ &= \det\, \begin{bmatrix} u & v \end{bmatrix} - \det\, \begin{bmatrix} u & \frac{\langle u,v\rangle}{\|u\|^2} u \end{bmatrix} \\ &= \det\, \begin{bmatrix} u & v \end{bmatrix} - \frac{\langle u, v\rangle}{\|u\|^2} \det\, \begin{bmatrix} u & u \end{bmatrix} \\ &= \det\, \begin{bmatrix} u & v \end{bmatrix}, \end{aligned}$$

logo $|\det\, B| = |\det\, A|$.

Sejam $u, v, w \in \mathbb{R}^3$, o paralelipípedo definido por $u, v, w$ é o conjunto

$$P[u,v,w] = \{\alpha u + \beta v + \gamma w, 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, 0 \leq \gamma \leq 1\}.$$




**Teorema 2.1.1** *Seja $A$ a matriz cujas colunas são $u, v$ e $w$. O volume de $P[u,v,w]$ é igual a $|det\ A|$.*

**Demonstração.** Exercício (proceder como em $\mathbb{R}^2$).

**Definição 2.1.1** *Sendo $v_1, \ldots, v_n \in \mathbb{R}^n$, o paralelipípedo definido por $v_1, \ldots, v_n$ é o conjunto*

$$P[v_1, \ldots, v_n] = \{\alpha_1 v_1 + \cdots + \alpha_n v_n, 0 \leq \alpha_1 \leq 1, \ldots, 0 \leq \alpha_n \leq 1\}.$$

**Definição 2.1.2** *Sejam $v_1, \ldots, v_n \in \mathbb{R}^n$ $(n \geq 4)$ e $A$ a matriz cujas colunas são $v_1, \ldots, v_n$. Define-se a medida de $P[v_1, \ldots, v_n]$ como sendo $|det\ A|$.*

## 2.2 Produto externo em $\mathbb{R}^3$

Em $\mathbb{R}^3$ é possível determinar um vector ortogonal a dois vectores dados a partir de uma outra operação entre vectores.

Dados $u = (a, b, c)$ e $v = (a', b', c')$ em $\mathbb{R}^3$, determinemos vectores perpendiculares a $u$ e a $v$. Suponhamos que $u$ e $v$ são linearmente independentes. Tem-se dim $\mathcal{L}\{u, v\} = 2$, logo dim $\mathcal{L}\{u, v\}^\perp = 1$, pelo que a direcção dos vectores procurados é univocamente determinada por $u$ e $v$.

Os vectores $x = (x_1, x_2, x_3)$ perpendiculares a $u$ e a $v$ têm de satisfazer o seguinte sistema

$$\begin{cases} \langle u, x \rangle = 0 \\ \langle v, x \rangle = 0 \end{cases} \iff \begin{cases} ax_1 + bx_2 + cx_3 = 0 \\ a'x_1 + b'x_2 + c'x_3 = 0. \end{cases}$$

Como supusemos $u$ e $v$ linearmente independentes, eles não são múltiplos um do outro, logo, pelo menos, um dos três determinantes

$$\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}, \quad \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \quad \begin{vmatrix} b & c \\ b' & c' \end{vmatrix}$$

é diferente de zero. Suponhamos o primeiro não nulo (se for outro o caso, o raciocínio é análogo). Escrevemos o sistema de outra forma

$$\begin{cases} ax_1 + bx_2 = -cx_3 \\ a'x_1 + b'x_2 = -c'x_3 \end{cases}$$

e usando a regra de Cramer

$$x_1 = \frac{\begin{vmatrix} -cx_3 & b \\ -c'x_3 & b' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}} \qquad x_2 = \frac{\begin{vmatrix} a & -cx_3 \\ a' & -c'x_3 \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}}$$

ou

$$x_1 = \frac{x_3\begin{vmatrix} -c & b \\ -c' & b' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}} \qquad x_2 = \frac{x_3\begin{vmatrix} a & -c \\ a' & -c' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}}$$

onde $x_3$ pode tomar qualquer valor. Escolhendo, por exemplo, $x_3 = \begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$, tem-se

$$x_1 = \begin{vmatrix} -c & b \\ -c' & b' \end{vmatrix} \qquad x_2 = \begin{vmatrix} a & -c \\ a' & -c' \end{vmatrix}.$$

Isto motiva a seguinte definição:

**Definição 2.2.1** *Dados $u = (a, b, c)$ e $v = (a', b', c')$ vectores de $\mathbb{R}^3$, chama-se produto externo (ou produto vectorial) de $u$ por $v$ ao vector*

$$u \wedge v = \left( \begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, - \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \right) = (bc' - b'c, -(ac' - a'c), ab' - a'b).$$

**Notação:** $u \wedge v$ ou $u \times v$.

Para recordar a expressão de $u \wedge v$ pode usar-se o seguinte "determinante simbólico"onde $e_1, e_2$ e $e_3$ são os vectores da base canónica de $\mathbb{R}^3$

$$\begin{vmatrix} e_1 & e_2 & e_3 \\ a & b & c \\ a' & b' & c' \end{vmatrix}$$

calculado pelo teorema de Laplace aplicado à primeira linha.

**Exemplo:** Sendo $u = (0, 1, 1)$ e $v = (1, 1, 1)$, tem-se

$$u \wedge v = \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = (0, 1, -1).$$

O vector $(0, 1, -1)$ é um vector ortogonal a $u$ e a $v$.

É fácil obter as principais propriedades do produto externo:

**Teorema 2.2.1** *Sejam $u, v$ e $w$ vectores de $\mathbb{R}^3$ e $\alpha$ um número real. Então:*

1. $u \wedge u = 0$;
2. $u \wedge 0 = 0$;
3. *Se $u$ e $v$ forem linearmente dependentes tem-se $u \wedge v = 0$;*
4. *Se $u \wedge v = 0$ então $u$ e $v$ são linearmente dependentes;*
5. $v \wedge u = -u \wedge v$;
6. $\langle u, u \wedge v \rangle = 0$;
7. $(u+v) \wedge w = u \wedge w + v \wedge w$, $u \wedge (v+w) = u \wedge v + u \wedge w$;
8. $(\alpha u) \wedge v = u \wedge (\alpha v) = \alpha(u \wedge v)$.

**Demonstração.** Exercício.

Vamos agora ver uma fórmula simples para a norma de $u \wedge v$.

**Teorema 2.2.2** *Tem-se*

$$\|u \wedge v\| = \sqrt{\|u\|^2\|v\|^2 - \langle u, v\rangle^2} = \|u\|\|v\| \ sen\ \theta,$$

*onde $\theta$ é o ângulo entre $u$ e $v$.*

**Demonstração.** Sendo $u = (a, b, c)$ e $v = (a', b', c')$, tem-se

$$\|u \wedge v\|^2 = (bc' - b'c)^2 + (ac' - a'c)^2 + (ab' - a'b)^2$$

e

$$\begin{aligned} \|u\|^2\|v\|^2 - \langle u, v\rangle^2 &= (a^2 + b^2 + c^2)(a'^2 + b'^2 + c'^2) - (aa' + bb' + cc')^2 \\ &= \ldots \\ &= (bc' - b'c)^2 + (ac' - a'c)^2 + (ab' - a'b)^2. \end{aligned}$$

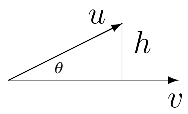
Quanto à segunda igualdade, como $\langle u, v\rangle = \|u\|\|v\| \cos\theta$, tem-se

$$\begin{aligned} \|u\|^2\|v\|^2 - \langle u, v\rangle^2 &= \|u\|^2\|v\|^2 - \|u\|^2\|v\|^2\cos^2\theta \\ &= \|u\|^2\|v\|^2(1 - \cos^2\theta) \\ &= \|u\|^2\|v\|^2\mathrm{sen}^2\theta. \end{aligned}$$

■

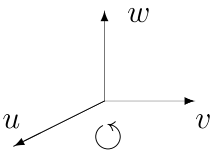
Da igualdade $\|u \wedge v\| = \|u\|\|v\|$ sen $\theta$ concluímos que a norma de $u \wedge v$ é igual à área do paralelogramo definido por $u$ e $v$:

$$\begin{aligned} A &= \|h\|\|v\| \\ &= \|u\|\text{sen } \theta\|v\| \\ &= \|u\|\|v\|\text{sen } \theta. \end{aligned}$$

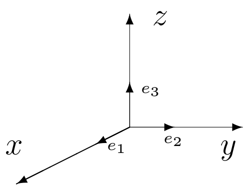



Com base nas propriedades anteriores do produto externo de dois vectores é fácil descrevê-lo geometricamente: é um vector ortogonal ao plano dos dois vectores dados e de comprimento igual à área do paralelogramo definido pelos dois vectores. Resta conhecer o sentido deste vector.

**Definição 2.2.2** *Sejam $u, v, w \in \mathbb{R}^3$ vectores linearmente independentes. Estes três vectores (por esta ordem) formam um triedro directo se a rotação mais curta do vector $u$ que o leva a sobrepor-se ao vector $v$ é feita, para um observador com os pés na origem e a cabeça na extremidade de $w$, no sentido contrário ao dos ponteiros do relógio.*




**Exemplo:** Os vectores $e_1, e_2, e_3$ da base canónica de $\mathbb{R}^3$ formam um triedro directo.




**Teorema 2.2.3** *Sendo $u, v \in \mathbb{R}^3$ linearmente independentes, os vectores $u, v, u \wedge v$ formam um triedro directo.*

**Demonstração.** Ver apontamentos A. P. Santana e J. F. Queiró.

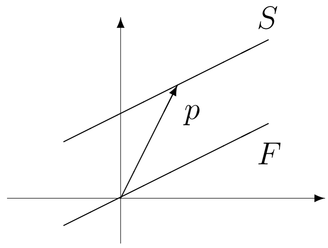
## 2.3 Planos em $\mathbb{R}^n$

Em $\mathbb{R}^2$ – identificado com um plano em que se fixou um sistema de eixos da forma habitual – uma recta que passa pela origem é simplesmente um subespaço de dimensão 1.

$$F = \{\lambda u : \lambda \in \mathbb{R}\}, \quad u \neq 0.$$

Se somarmos qualquer vector de $F$ a p obtemos uma recta que é paralela a $F$ (obtida por translação de $F$):

$$S = \{p + v : v \in F\}.$$




Exactamente da mesma forma podemos descrever uma recta qualquer em $\mathbb{R}^3$ (sendo então $F$ um subespaço de dimensão 1 de $\mathbb{R}^3$) e um plano qualquer de $\mathbb{R}^3$ (sendo $F$ um subespaço de dimensão 2 de $\mathbb{R}^3$).

**Definição 2.3.1** *Um plano de dimensão $k$ em $\mathbb{R}^n$ é um conjunto do tipo*

$$S = \{p + v : v \in F\},$$

*onde $p \in \mathbb{R}^n$ e $F$ é um subespaço de $\mathbb{R}^n$ com dimensão $k$ (notação: $S = \{p\} + F$).*

*Diz-se que $S$ foi obtido de $F$ por uma translação segundo o vector $p$.*

*A $F$ chama-se o subespaço director de $S$ e a $p$ o vector de translação.*

*Um plano de dimensão 1 diz-se uma recta e um plano de dimensão $n-1$ em $\mathbb{R}^n$ diz-se um hiperplano.*

**Exemplos:**

1. Em $\mathbb{R}^2$, os planos de dimensão 0 são os vectores individuais de $\mathbb{R}^2$. Os planos de dimensão 1 (hiperplanos de $\mathbb{R}^2$) são as rectas. O plano de dimensão 2 é o próprio $\mathbb{R}^2$.

2. Em $\mathbb{R}^3$, os planos de dimensão 0 são os vectores individuais de $\mathbb{R}^3$. Os hiperplanos são os planos.

**Definição 2.3.2** *Dois planos da mesma dimensão $S_1 = \{p_1\} + F_1$ e $S_2 = \{p_2\} + F_2$ dizem-se paralelos se $F_1 = F_2$. Se $S_1$ e $S_2$ não tiverem a mesma dimensão a condição de paralelismo é $F_1 \subset F_2$ ou $F_2 \subset F_1$.*

Seja $S = \{p\} + F$ um plano de dimensão $k$ em $\mathbb{R}^n$ e seja $\{v_1, v_2, \ldots, v_k\}$ uma base de $F$. Assim $F = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k : \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}\}$, logo

$$S = \{p + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k : \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}\}.$$

Dado $x \in \mathbb{R}^n$, $x \in S$ se e só se

$$\boxed{x = p + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k,}$$

para alguns $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$. Esta é a **equação vectorial de** $S$. Mas, sendo $x \in \mathbb{R}^n$, a equação vectorial é equivalente a

$$(x_1, x_2, \ldots, x_n) = (p_1, p_2, \ldots, p_n) + \alpha_1(v_{11}, v_{12}, \ldots, v_{1n}) + \cdots + \alpha_k(v_{k1}, \ldots, v_{kn})$$

ou ainda

$$\boxed{\left\{\begin{array}{l} x_1 = p_1 + \alpha_1 v_{11} + \cdots + \alpha_k v_{k1} \\ x_2 = p_2 + \alpha_1 v_{12} + \cdots + \alpha_k v_{k2} \\ \vdots \\ x_n = p_n + \alpha_1 v_{1n} + \cdots + \alpha_k v_{kn} \end{array}\right.}$$

chamadas **equações paramétricas de** $S$.

Diz-se que $S$ passa por $p$ e é paralelo a $F$, ou também que é paralelo aos vectores $v_1, \ldots, v_k$.

**Exemplos:**

1. A recta que passa por $(1, 2, 3)$ e é paralela ao vector $(4, 5, -7)$ tem equação vectorial

   $$(x, y, z) = (1, 2, 3) + \alpha(4, 5, -7), \quad \alpha \in \mathbb{R},$$

   e equações paramétricas

   $$\left\{\begin{array}{l} x = 1 + 4\alpha \\ y = 2 + 5\alpha \\ z = 3 - 7\alpha, \quad \alpha \in \mathbb{R}. \end{array}\right.$$

2. O plano que passa por $(1, 2, 4)$ e é paralelo aos vectores $(1, 2, 1)$ e $(0, 1, 0)$ (é de facto um plano de dimensão 2 porque $(1, 2, 1)$ e $(0, 1, 0)$ são linearmente independentes) tem equação vectorial

   $$(x, y, z) = (1, 2, 4) + \lambda(1, 2, 1) + \mu(0, 1, 0)$$

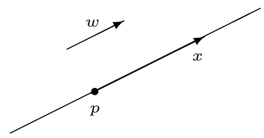
   e equações paramétricas

   $$\left\{\begin{array}{l} x = 1 + \lambda \\ y = 2 + 2\lambda + \mu \\ z = 4 + \lambda. \end{array}\right.$$

**Nota:** Nenhuma destas representações de planos está univocamente determinada, isto é, podemos para o mesmo plano ter várias equações vectoriais assim como vários sistemas de equações paramétricas.

Sejam $p = (p_1, p_2, \ldots, p_n)$ e $w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$. Suponhamos $w_i \neq 0$ para $i = 1, \ldots, n$. Se $x$ pertence à recta em $\mathbb{R}^n$ que passa por $p$ e tem a direcção de $w$ então $x - p$ e $w$ são linearmente dependentes, ou seja, $x - p$ e $w$ são múltiplos um do outro. As equações

$$\boxed{\frac{x_1 - p_1}{w_1} = \frac{x_2 - p_2}{w_2} = \cdots = \frac{x_n - p_n}{w_n}}$$




dizem-se **equações normais ou canónicas** da recta que passa por $p$ e tem a direcção de $w$

**Nota:** As equações paramétricas desta recta são

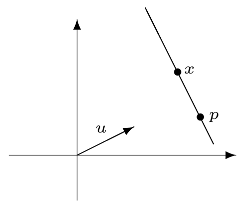
$$\begin{cases} x_1 = p_1 + \lambda w_1 \\ x_2 = p_2 + \lambda w_2 \\ \vdots \\ x_n = p_n + \lambda w_n \end{cases}$$

logo $\lambda = \dfrac{x_1 - p_1}{w_1}$ da primeira igualdade, $\lambda = \dfrac{x_2 - p_2}{w_2}$ da segunda, $\cdots$, $\lambda = \dfrac{x_n - p_n}{w_n}$ da última, ou seja,

$$\frac{x_1 - p_1}{w_1} = \frac{x_2 - p_2}{w_2} = \cdots = \frac{x_n - p_n}{w_n}.$$

Um processo alternativo de descrever planos em $\mathbb{R}^n$ usa o produto interno.

Em $\mathbb{R}^2$, considere-se a recta que passa por $p$ e é perpendicular a $u \neq 0$.
Um ponto $x$ pertence a esta recta se $x - p$ for ortogonal a $u$.




A condição

$$\boxed{\langle u, x - p \rangle = 0}$$

é a equação cartesiana de uma recta (plano de dimensão 1) em $\mathbb{R}^2$.

Mais geralmente vamos ver que em $\mathbb{R}^n$ a condição

$$\langle u, x - p \rangle = 0$$

define um hiperplano.

**Teorema 2.3.1** *Sendo $u, p \in \mathbb{R}^n$, com $u \neq 0$, o conjunto $\{x : \langle u, x - p\rangle = 0\}$ é um hiperplano. Reciprocamente, sendo $S$ um hiperplano, existem $u, p \in \mathbb{R}^n$, com $u \neq 0$, tais que $S = \{x : \langle u, x - p\rangle = 0\}$.*

**Demonstração.**

**1º Caso − $p = 0$.** O conjunto $\{x : \langle u, x\rangle = 0\}$ é $(\mathcal{L}\{u\})^{\perp}$ e como $\mathbb{R}^n = \mathcal{L}\{u\} \oplus (\mathcal{L}\{u\})^{\perp}$, tem-se dim $(\mathcal{L}\{u\})^{\perp} = n - 1$, logo é um subespaço de dimensão $n - 1$, ou seja, é um hiperplano que passa pela origem.

**2º Caso − $p \in \mathbb{R}^n$ qualquer.** $y \in S = \{x : \langle u, x - p\rangle = 0\}$ se e só se $y - p$ pertencer ao subespaço $F = \{x : \langle u, x\rangle = 0\}$. Então, $S = \{p\} + F$ e como $F$ tem dimensão $n - 1$, $S$ é um hiperplano.

Reciprocamente, seja $S$ um hiperplano, $S = \{p\} + F$ onde dim $F = n - 1$. Como $\mathbb{R}^n = F \oplus F^{\perp}$ tem-se dim $F^{\perp} = 1$. Seja $u \in F^{\perp}$ não nulo, então $F^{\perp} = \mathcal{L}\{u\}$ e vem $F = (F^{\perp})^{\perp} = \{x : \langle u, x\rangle = 0\}$. Como $S = \{p\} + F$, vem que

$$S = \{x : \langle u, x - p\rangle = 0\}.$$

■

**Definição 2.3.3** *Chama-se equação cartesiana do hiperplano à condição*

$$\langle u, x - p\rangle = 0.$$

*O vector $u$ diz-se um vector ortogonal ao hiperplano.*

**Exemplos:**

1. Em $\mathbb{R}^2$ os hiperplanos são as rectas.

   A equação cartesiana da recta que passa por $(p_1, p_2)$ e é ortogonal a $(u_1, u_2)$ é

$$\begin{aligned} \langle (u_1, u_2), (x, y) - (p_1, p_2)\rangle &= 0 \\ u_1(x - p_1) + u_2(y - p_2) &= 0 \\ u_1 x + u_2 y &= u_1 p_1 + u_2 p_2. \end{aligned}$$

   Fazendo $\langle u, p\rangle = b$ vem

$$u_1 x + u_2 y = b.$$

2. Em $\mathbb{R}^3$ os hiperplanos são os planos.

   A equação cartesiana do plano que passa por $(p_1, p_2, p_3)$ e é ortogonal a $(u_1, u_2, u_3)$ é

$$\begin{aligned} \langle (u_1, u_2, u_3), (x, y, z) - (p_1, p_2, p_3)\rangle &= 0 \\ u_1(x - p_1) + u_2(y - p_2) + u_3(z - p_3) &= 0 \\ u_1 x + u_2 y + u_3 z &= u_1 p_1 + u_2 p_2 + u_3 p_3. \end{aligned}$$

   Fazendo $\langle u, p\rangle = b$ vem

$$u_1 x + u_2 y + u_3 z = b.$$

**Nota:** O vector ortogonal a um hiperplano é essencialmente único, no sentido de que dois vectores ortogonais a um mesmo hiperplano são múltiplos um do outro.

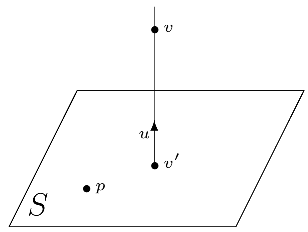
## 2.4 Problemas métricos

### 2.4.1 Distância de um ponto a um hiperplano

Em $\mathbb{R}^3$, como determinar a distância de um ponto a um hiperplano?

Seja $S$ o plano de $\mathbb{R}^3$ que contém $p$ e é perpendicular a $u \neq 0$, logo a equação cartesiana deste plano é

$$\langle x - p, u \rangle = 0.$$




Queremos determinar a distância de $v$ a $S$. Consideramos a recta que passa por $v$ e é perpendicular a $S$, isto é, a recta que passa por $v$ e é paralela a $u$. Sendo $v'$ a intersecção desta recta com $S$, a distância pretendida é

$$\|v - v'\|.$$

Encontramos assim o seguinte sistema

$$\begin{cases} v' \in S \\ v' = v + \lambda u \end{cases} \iff \begin{cases} \langle v' - p, u \rangle = 0 \\ v' - v = \lambda u. \end{cases}$$

Da primeira equação vem que $\lambda = -\dfrac{\langle v - p, u \rangle}{\langle u, u \rangle}$ e substituindo na segunda obtém-se

$$v - v' = \frac{\langle v - p, u \rangle}{\langle u, u \rangle}\, u.$$

Logo,

$$\|v - v'\| = \frac{|\langle v - p, u \rangle|}{|\langle u, u \rangle|}\, \|u\| = \frac{|\langle v - p, u \rangle|}{\|u\|}.$$

Em $\mathbb{R}^n$, a distância de $v$ ao hiperplano de equação cartesiana $\langle u, x - p \rangle = 0$ é igual à norma da projecção ortogonal de $v - p$ sobre $u$, isto é, é igual a

$$\frac{|\langle v - p, u \rangle|}{\|u\|}.$$

Em $\mathbb{R}^2$, a distância do ponto $(q_1, q_2)$ à recta de equação cartesiana $a_1x_1 + a_2x_2 = b$ é

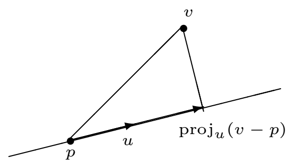
$$\frac{|a_1q_1 + a_2q_2 - b|}{\sqrt{a_1^2 + a_2^2}}.$$

Em $\mathbb{R}^3$, a distância do ponto $(q_1, q_2, q_3)$ ao plano de equação cartesiana $a_1x_1 + a_2x_2 + a_3x_3 = b$ é
$$\frac{|a_1q_1 + a_2q_2 + a_3q_3 - b|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

### 2.4.2 Distância de um ponto a uma recta

Em $\mathbb{R}^n$, como determinar a distância de um ponto $v$ a uma recta $x = p + \alpha u$, $\alpha \in \mathbb{R}$?

Será a distância de $v-p$ à sua projecção ortogonal sobre $u$, ou seja,
$$\|v - p - \mathrm{proj}_u(v-p)\|.$$




Fazendo alguns cálculos,
$$\begin{aligned}\|v - p - \mathrm{proj}_u(v-p)\|^2 &= \langle v - p - \mathrm{proj}_u(v-p), v - p - \mathrm{proj}_u(v-p)\rangle \\ &= \cdots \\ &= \frac{\|v-p\|^2\|u\|^2 - \langle v-p, u\rangle^2}{\|u\|^2}.\end{aligned}$$

Logo, a distância pretendida é
$$\frac{\sqrt{\|v-p\|^2\|u\|^2 - \langle v-p, u\rangle^2}}{\|u\|}.$$

### 2.4.3 Outras distâncias

Em $\mathbb{R}^3$, tem-se

**Distância entre dois planos $P$ e $P'$**

- Se $P$ e $P'$ forem coincidentes, $d(P, P') = 0$;
- Se $P$ e $P'$ se intersectam (segundo uma recta), $d(P, P') = 0$;
- Se $P$ e $P'$ forem estritamente paralelos, $d(P, P') = d(x, P')$, com $x \in P$.

**Distância de uma recta $r$ a um plano $P$**

- Se $r$ for concorrente com $P$, $d(r, P) = 0$;
- Se $r$ for paralela a $P$, $d(r, P) = d(p, P)$, com $p \in r$.

**Distância entre duas rectas $r$ e $r'$**

- Se as rectas forem concorrentes, $d(r, r') = 0$;
- Se as rectas forem paralelas, $d(r, r') = d(p, r')$, com $p \in r$;
- Se as rectas forem enviesadas, $d(r, r') = d(r, P)$, onde $P$ é o plano que contém $r'$ e é paralelo a $r$.

### 2.4.4 Ângulos

O **ângulo entre duas rectas** $x = p + \alpha v$ e $x = q + \alpha w$ é o ângulo entre $v$ e $w$ (ou o suplementar desse, se ele não pertencer a $[0, \pi/2]$).

O **ângulo entre dois planos** $\langle u, x - p \rangle = 0$ e $\langle v, x - q \rangle = 0$ é o ângulo entre $v$ e $w$ (ou o suplementar desse, se ele não pertencer a $[0, \pi/2]$).

# Capítulo 3

# Complementos sobre problemas de mínimos quadrados

## 3.1 Decomposição QR de uma matriz — processo de ortogonalização de Gram-Schmidt

**Revisões de ALGA:**

Se $x \in V$, um vector $x_F$ diz-se a **projecção ortogonal de** $x$ **sobre o subespaço** $F$ de $V$ se

$$x - x_F \in F^{\perp}.$$

Relativamente a uma base ortogonal $\{v_1, \ldots, v_k\}$ (vectores ortogonais dois a dois) de $F$ tem-se

$$x_F = \mathrm{proj}_F x = \sum_{i=1}^{k} \frac{\langle x, v_i \rangle}{\|v_i\|^2} v_i = \sum_{i=1}^{k} \mathrm{proj}_{v_i} x.$$

**Ortogonalização de Gram-Schmidt:** processo iterativo para obter uma base **ortogonal** $\{u_1, u_2, \ldots, u_k\}$ de um subespaço vectorial $F$ a partir de uma base qualquer $\{v_1, v_2, \ldots, v_k\}$ de $F$.

$$\begin{array}{rcl}
u_1 & = & v_1 \\
u_2 & = & v_2 - \mathrm{proj}_{u_1} v_2 \\
u_3 & = & v_3 - \mathrm{proj}_{u_1} v_3 - \mathrm{proj}_{u_2} v_3 \\
 & \vdots & \\
u_k & = & v_k - \mathrm{proj}_{u_1} v_k - \mathrm{proj}_{u_2} v_k - \cdots - \mathrm{proj}_{u_{k-1}} v_k
\end{array}$$

Seja $A \in M_{n\times k}(\mathbb{C})$ uma matriz com as colunas linearmente independentes, isto é, $\mathrm{car}(A) = k$. Designem-se as colunas de $A$ por $v_1, \ldots, v_k$ (são vectores de $\mathbb{C}^n$) e por $u_1, \ldots, u_k$ as colunas duas a duas ortogonais que se obtêm das de $A$ aplicando-lhes o processo de ortogonalização de Gram-Schmidt.

$$\begin{aligned}
u_1 &= v_1 \\
u_2 &= v_2 - \alpha_{12}u_1 \\
u_3 &= v_3 - \alpha_{13}u_1 - \alpha_{23}u_2 \\
&\vdots \\
u_k &= v_k - \alpha_{1k}u_1 - \alpha_{2k}u_2 - \cdots - \alpha_{k-1,k}u_{k-1}
\end{aligned}$$

com $\alpha_{ij} \in \mathbb{C}$. Estas relações podem escrever-se

$$\begin{aligned}
v_1 &= u_1 \\
v_2 &= \alpha_{12}u_1 + u_2 \\
v_3 &= \alpha_{13}u_1 + \alpha_{23}u_2 + u_3 \\
&\vdots \\
v_k &= \alpha_{1k}u_1 + \alpha_{2k}u_2 + \cdots + \alpha_{k-1,k}u_{k-1} + u_k.
\end{aligned}$$

Sendo $U$ a matriz cujas colunas são $u_1, \ldots, u_k$, estas igualdades podem resumir-se pela igualdade matricial $A = UT$, onde

$$T = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1k} \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2k} \\ 0 & 0 & 1 & \cdots & \alpha_{3k} \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Assim,

$$A = \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} T = \begin{bmatrix} \frac{u_1}{\|u_1\|} & \cdots & \frac{u_k}{\|u_k\|} \end{bmatrix} \begin{bmatrix} \|u_1\| & & \\ & \ddots & \\ & & \|u_k\| \end{bmatrix} T.$$

Faça-se $\hat{Q} = \begin{bmatrix} \frac{u_1}{\|u_1\|} & \cdots & \frac{u_k}{\|u_k\|} \end{bmatrix}$ e $\hat{R} = \begin{bmatrix} \|u_1\| & & \\ & \ddots & \\ & & \|u_k\| \end{bmatrix} T$. Tem-se

$$\boxed{A = \hat{Q}\hat{R}}.$$

$\hat{Q}$ é uma matriz $n \times k$ e tem as colunas ortonormadas. $\hat{R}$ é uma matriz quadrada de ordem $k$, triangular superior e não singular.

Seja $F$ um subespaço de dimensão $k$ e $\{u_1, \ldots, u_k\}$ uma sua base ortonormada. Recorde-se que a projecção ortogonal de $v$ sobre o subespaço $F$ pode ser expressa como

$$v_F = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_k \rangle u_k.$$

Ponha-se

$$\hat{Q} = [\, u_1 \ \cdots \ u_k \,] \in M_{n\times k}(\mathbb{C}).$$

Veja-se que

$$\begin{aligned}
\hat{Q}\hat{Q}^* v &= [\, u_1 \ \cdots \ u_k \,] \begin{bmatrix} u_1^* \\ \vdots \\ u_k^* \end{bmatrix} v \\
&= [\, u_1 \ \cdots \ u_k \,] \begin{bmatrix} u_1^* v \\ \vdots \\ u_k^* v \end{bmatrix} \\
&= [\, u_1 \ \cdots \ u_k \,] \begin{bmatrix} \langle v, u_1 \rangle \\ \vdots \\ \langle v, u_k \rangle \end{bmatrix} \\
&= \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_k \rangle u_k \\
&= v_F,
\end{aligned}$$

o que mostra a expressão

$$v_F = \hat{Q}\hat{Q}^* v.$$

A esta matriz $\hat{Q}\hat{Q}^*$ chama-se matriz de projecção ortogonal sobre $F$.

Seja $\hat{Q}$ a matriz da decomposição QR de uma matriz $A \in M_{n\times k}(\mathbb{C})$, ou seja, considere-se, agora, $A = \hat{Q}\hat{R}$ com $\hat{R} \in M_{k\times k}(\mathbb{C})$ uma matriz triangular superior não singular. Desta forma, como $A = \hat{Q}\hat{R}$ é equivalente a $\hat{Q} = A\hat{R}^{-1}$, vem que

$$\begin{aligned}
\hat{Q}\hat{Q}^* &= (A\hat{R}^{-1})(A\hat{R}^{-1})^* &&= (A\hat{R}^{-1})((\hat{R}^{-1})^* A^*) \\
&= A(\hat{R}^{-1}(\hat{R}^{-1})^*)A^* &&= A(\hat{R}^{-1}(\hat{R}^*)^{-1})A^* \\
&= A(\hat{R}^*\hat{R})^{-1}A^*.
\end{aligned}$$

Mas, porque $\hat{Q}^*\hat{Q} = I$, $A^*A = \hat{R}^*\hat{Q}^*\hat{Q}\hat{R} = \hat{R}^*\hat{R}$ e

$$\hat{Q}\hat{Q}^* = A(A^*A)^{-1}A^* \in M_{n\times n}(\mathbb{C}).$$

Vamos, agora, tentar identificar uma matriz de projecção ortogonal sobre $F^\perp$, o complemento ortogonal do subespaço $F$. Dado o vector $v$, sabemos que

$$v = v_F + v_{F^\perp},$$

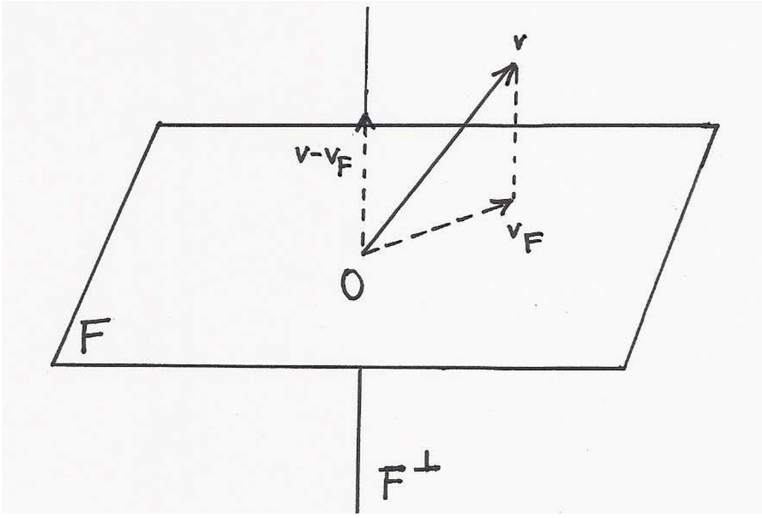



Figura 3.1.1: Projecção ortogonal sobre $F$ e $F^{\perp}$.

em que $v_{F^\perp}$ designa a projecção ortogonal de $v$ sobre $F^{\perp}$.

Logo,

$$\begin{aligned} v_{F^\perp} &= v - v_F \\ &= v - \hat{Q}\hat{Q}^* v \\ &= (I - \hat{Q}\hat{Q}^*)v. \end{aligned}$$

Desta forma, chamamos à matriz

$$I - \hat{Q}\hat{Q}^* = I - A(A^*A)^{-1}A^* \in M_{n\times n}(\mathbb{C}),$$

matriz da projecção ortogonal sobre $F^{\perp}$.

Verifiquemos que $v_{F^\perp} = (I - \hat{Q}\hat{Q}^*)v$ é, de facto, a projecção ortogonal de $v$ sobre $F^{\perp}$:

$$\begin{aligned} \langle f, v - (I - \hat{Q}\hat{Q}^*)v \rangle &= \left(v - (I - \hat{Q}\hat{Q}^*)v\right)^* f \\ &= \left(v^* - v^*(I - \hat{Q}\hat{Q}^*)\right) f \\ &= v^* f - v^* f + v^*\hat{Q}\hat{Q}^* f \\ &= v^*(\hat{Q}\hat{Q}^* f). \end{aligned}$$

(Note-se que $I - \hat{Q}\hat{Q}^*$ é hermítica.) Daqui resulta que

$$\langle f, v - (I - \hat{Q}\hat{Q}^*)v \rangle = 0 \quad \text{se} \quad f \in F^{\perp}.$$

É fácil de constatar que $\hat{Q}\hat{Q}^*$ e $I - \hat{Q}\hat{Q}^*$ são matrizes que verificam $P^2 = P$.

**Definição 3.1.1** *Uma matriz* $P \in M_{n\times n}(\mathbb{C})$ *que verifica*

$$P^2 = P$$

*chama-se um projector.*

Vimos, também, que as matrizes $\hat{Q}\hat{Q}^*$ e $I - \hat{Q}\hat{Q}^*$ projectam ortogonalmente sobre $F$ e $F^\perp$, respectivamente, o que motiva a seguinte definição.

**Definição 3.1.2** *Um projector ortogonal* $P \in M_{n\times n}(\mathbb{C})$ *é um projector tal que*

$$F_1 = \{Px,\ x \in \mathbb{C}^n\} \quad e \quad F_2 = \{(I-P)x,\ x \in \mathbb{C}^n\}$$

*são subespaços ortogonais.*

Existe uma caracterização necessária e suficiente para um projector ser ortogonal. Fizemos menção, anteriormente, ao facto de $\hat{Q}\hat{Q}^*$ ser uma matriz hermítica. O mesmo sucede com $I - \hat{Q}\hat{Q}^*$ e, como se prova de seguida, com qualquer projector ortogonal.

**Teorema 3.1.1** *Um projector* $P \in M_{n\times n}(\mathbb{C})$ *é um projector ortogonal se e só se* $P = P^*$.

**Demonstração.** Se $P = P^*$ então

$$\begin{aligned}\langle Px, (I-P)y\rangle &= ((I-P)y)^*(Px) &&= y^*(I-P)^*(Px)\\ &= y^*((I-P)P)x &&= y^*(P-P^2)x\\ &= 0,\end{aligned}$$

o que mostra a ortogonalidade entre os subespaços $F_1$ e $F_2$ da definição de projector ortogonal.

No outro sentido, suponhamos que $P$ é um projector ortogonal. Se dim $F_1 = k$, considere-se uma base ortonormada de $\mathbb{C}^n$ $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ onde $\{v_1, \ldots, v_k\}$ é uma base de $F_1$ e $\{v_{k+1}, \ldots, v_n\}$ é uma base de $F_2$. Então $Pv_j = v_j$ para $j = 1, \ldots, k$ e $Pv_j = 0$ para $j \geq k+1$. Sendo $Q$ a matriz unitária cujas colunas são $v_1, \ldots, v_n$, tem-se

$$PQ = \begin{bmatrix} | & & | & | & & | \\ v_1 & \ldots & v_k & 0 & \ldots & 0 \\ | & & | & | & & | \end{bmatrix}$$

e

$$Q^*PQ = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} = D.$$

Logo $P = QDQ^*$ e $P^* = (QDQ^*)^* = QD^*Q^* = QDQ^* = P$.

Nota: A matriz $P$ é também da forma $\hat{Q}\hat{Q}^*$, onde $\hat{Q}$ é a matriz $n \times k$ cujas colunas são $v_1, \ldots, v_k$. ■

### Exercícios

1. Mostre que $\hat{Q}\hat{Q}^*$ e $I - \hat{Q}\hat{Q}^*$ são projectores.

2. Prove que se $P$ é um projector ortogonal então $I - 2P$ é uma matriz unitária. Dê uma interpretação geométrica deste facto.

3. Considere os vectores

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{e} \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

   (a) Determine uma base ortonormada para o subespaço gerado por $v_1$ e $v_2$ através do processo de ortogonalização de Gram-Schmidt.

   (b) Seja $A$ a matriz 4-por-2 cujas colunas são $v_1$ e $v_2$. Com base na alínea anterior, escreva a decomposição QR da matriz $A$.

   (c) Calcule $I - \hat{Q}\hat{Q}^*$ (em que $\hat{Q}$ é a matriz obtida na alínea anterior).

   (d) Verifique que $I - \hat{Q}\hat{Q}^*$ é um projector ortogonal.

   (e) Sobre que subespaço é que a matriz $I - \hat{Q}\hat{Q}^*$ projecta ortogonalmente?

4. Considere os vectores

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{e} \quad v_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

   (a) Determine uma base ortonormada para o subespaço gerado por $v_1$, $v_2$ e $v_3$ através do processo de ortogonalização de Gram-Schmidt.

   (b) Seja $A$ a matriz 3-por-3 cujas colunas são $v_1$, $v_2$ e $v_3$. Com base na alínea anterior, escreva a decomposição QR da matriz $A$.

   (c) Calcule $\hat{Q}\hat{Q}^*$ (em que $\hat{Q}$ é a matriz obtida na alínea anterior).

   (d) Verifique que $\hat{Q}\hat{Q}^*$ é um projector ortogonal.

   (e) Sobre que subespaço é que a matriz $\hat{Q}\hat{Q}^*$ projecta ortogonalmente?

## 3.2 Decomposição QR de uma matriz — triangularização de Householder

O processo de ortogonalização de Gram-Schmidt pode ser resumido na forma matricial

$$A = \hat{Q}\hat{R} = UT, \quad \hat{Q} = UD^{-1} \quad \text{e} \quad \hat{R} = DT,$$

com as mesmas notações e hipóteses da secção anterior. Repare-se que $A = UT$ é equivalente a

$$AT^{-1} \;=\; U$$

e que $T^{-1}$ pode ser expressa como o produto de $k-1$ matrizes,

$$T^{-1} \;=\; E_2 \cdots E_k,$$

com

$$E_j \;=\; \begin{bmatrix} 1 & & & -\frac{\langle v_j, u_1\rangle}{\|u_1\|^2} & & \\ & \ddots & & \vdots & & \\ & & 1 & -\frac{\langle v_j, u_{j-1}\rangle}{\|u_{j-1}\|^2} & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}, \quad j = 2, \ldots, k.$$

Cada uma destas matrizes $E_j$ resulta, por sua vez, do produto de $j-1$ matrizes elementares.

O processo de ortogonalização de Gram-Schmidt pode ser encarado, deste modo, como uma *ortogonalização triangular.*

### 3.2.1 Triangularização ortogonal

Existe uma outra forma de calcular a decomposição QR de uma matriz. Este processo, descrito nesta secção, pode ser visto como uma *triangularização ortogonal.* A ideia é formar $\hat{Q}$ à custa do produto de $k$ matrizes unitárias $n \times n$:

$$Q_k \cdots Q_2 Q_1 A \;=\; R,$$

o que é equivalente a

$$A \;=\; QR,$$

com $Q = Q_1^* Q_2^* \cdots Q_k^*$. A matriz produto $Q$ é unitária (ver exercícios da Secção 1.5). As dimensões são as seguintes:

$$A \in M_{n\times k}(\mathbb{C}), \quad Q \in M_{n\times n}(\mathbb{C}) \quad \text{e} \quad R \in M_{n\times k}(\mathbb{C}).$$

Esquematicamente, temos, quando $n = 5$ e $k = 3$, que

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}} = \underset{Q}{\begin{bmatrix} * & * & * & \star & \star \\ * & * & * & \star & \star \\ * & * & * & \star & \star \\ * & * & * & \star & \star \\ * & * & * & \star & \star \end{bmatrix}} \underset{R}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\bar{\times}}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}.$$

Daqui, recupera-se a forma reduzida da decomposição QR de $A$:

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}} = \underset{\hat{Q}}{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}} \underset{\hat{R}}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\bar{\times}}} \end{bmatrix}}.$$

Em termos gerais, a forma reduzida $A = \hat{Q}\hat{R}$ da decomposição QR de $A$ é obtida através da sua forma completa $A = QR$, atribuindo

$$\hat{Q} \;=\; Q_{1:n,1:k} \quad \text{e} \quad \hat{R} \;=\; R_{1:k,1:k},$$

ou seja, retirando a $Q$ as suas últimas $n-k$ colunas e a $R$ as suas últimas $n-k$ linhas.

Passamos, então, à descrição do processo de triangularização ortogonal de $A$, baseado na sua decomposição $A = QR$. Este processo (quando $n = 5$ e $k = 3$) desenrola-se de acordo com o seguinte esquema:

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}} \xrightarrow{Q_1} \underset{Q_1A}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \end{bmatrix}} \xrightarrow{Q_2} \underset{Q_2Q_1A}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\times}} \end{bmatrix}} \xrightarrow{Q_3} \underset{R = Q_3Q_2Q_1A}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\bar{\times}}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}.$$

A matriz $\hat{R}$ é definida pelas três primeiras linhas de $R = Q_3Q_2Q_1A$, ou seja, por

$$\hat{R} \;=\; \begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\bar{\times}}} \end{bmatrix}.$$

A matriz $\hat{Q}$ é dada pelas três primeiras colunas de $Q = Q_1^*Q_2^*Q_3^*$.

Regressemos ao caso geral e comecemos por analisar a matriz $Q_1 \in M_{n\times n}(\mathbb{C})$. Esta matriz terá de ser unitária. Além disso, e de acordo com o esquema apresentado, terá de anular os elementos da primeira coluna de $A$ abaixo do elemento na posição $1,1$. Seja $v_1$ a primeira coluna de $A$. Pretende-se que a matriz unitária $Q_1$ seja tal que

$$Q_1v_1 \;=\; \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

O escalar $\alpha$, que no esquema corresponde ao símbolo $\bar{\times}$ na posição $1,1$ de $Q_1A$, vai ser escolhido de forma a valer $\|v_1\|$, o que está de acordo com a decomposição QR calculada através do processo de ortogonalização de Gram-Schmidt. Assim,

$$Q_1 v_1 \;=\; \|v_1\| e_1,$$

em que $e_1$ representa a primeira coluna da matriz identidade de ordem $n$.

Como iremos ver mais à frente, existe uma classe de matrizes, conhecida por reflectores de Householder, que nos dá uma solução para a matriz $Q_1$ com as propriedades desejadas. No caso em causa, o reflector de Householder procurado é dado por

$$Q_1 \;=\; I - \frac{2}{\|h_1\|^2} h_1 h_1^* \;\in\; M_{n\times n}(\mathbb{C}) \qquad \text{com} \qquad h_1 \;=\; \|v_1\| e_1 - v_1.$$

Confirmaremos na próxima secção que $Q_1 v_1 = \|v_1\| e_1$.

O passo seguinte consiste em identificar a matriz $Q_2$. Observando o esquema anterior, constata-se que $Q_2$, ao multiplicar, à esquerda, $Q_1A$, deverá deixar intacta a primeira linha desta última matriz. Tal efeito é conseguido ao exigir-se que

$$Q_2 \;=\; \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \bar{Q}_2 & \\ 0 & & & \end{bmatrix} \;\in\; M_{n\times n}(\mathbb{C}).$$

Se $\bar{Q}_2 \in M_{n-1\times n-1}(\mathbb{C})$ for unitária então $Q_2$ também o será (ver exercícios da Secção 1.5). Além disto, a matriz $\bar{Q}_2$ terá que anular os elementos da segunda coluna de $Q_1A$ abaixo da posição $2,2$. Recorremos, novamente, aos reflectores de Householder para conseguir este último objectivo. A matriz $\bar{Q}_2$ é, assim, dada por

$$\bar{Q}_2 \;=\; I - \frac{2}{\|h_2\|^2} h_2 h_2^* \;\in\; M_{n-1\times n-1}(\mathbb{C}) \qquad \text{com} \qquad h_2 \;=\; \|\bar{v}_2\| \bar{e}_1 - \bar{v}_2$$

e $\bar{v}_2$ o vector em $\mathbb{C}^{n-1}$ formado pelos elementos da segunda coluna de $Q_1A$, da posição $2,2$ até à posição $n,2$. O vector $\bar{e}_1$ é a primeira coluna da matriz identidade de ordem $n-1$. No esquema apresentado para $n=5$ e $k=2$, as componentes do vector $\bar{v}_2$ aparecem, de seguida, entre parêntesis rectos:

$$Q_1A \;=\; \begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \end{bmatrix}.$$

Este procedimento é aplicado, sucessivamente, até ser concluída a triangularização. No passo $j$, com $2 \le j \le k$, a matriz $Q_j$ tem a forma

$$Q_j \;=\; \begin{bmatrix} I_{j-1\times j-1} & 0 \\ 0 & \bar{Q}_j \end{bmatrix} \;\in\; M_{n\times n}(\mathbb{C}),$$

em que $\bar{Q}_j$ é um reflector de Householder de ordem $n-j+1$. A matriz $Q_j$ é unitária por $\bar{Q}_j$ também o ser (ver exercícios da Secção 1.5).

### 3.2.2 Reflectores de Householder

Ficou por mostrar que um reflector de Householder é uma matriz unitária com as propriedades de anulação referidas. Apresenta-se, primeiro, a definição de um reflector de Householder.

**Definição 3.2.1** *Dado um vector $y \in \mathbb{C}^n$, o correspondente reflector de Householder $H \in M_{n\times n}(\mathbb{C})$ é dado por*

$$H = I - \frac{2}{\|h\|^2}hh^* \qquad \textit{com} \qquad h = \|y\|e_1 - y,$$

*em que $e_1$ é a primeira coluna da matriz identidade de ordem $n$.*

A interpretação geométrica do efeito de multiplicar à esquerda um reflector de Householder pelo vector que o define permite conhecer melhor a expressão para $H$. Como o próprio nome indica, $H$ vai *reflectir* $y$ em torno de um determinado subespaço, de acordo com a Figura 3.2.2. A matriz

$$\frac{1}{\|h\|^2}hh^* = \frac{1}{\langle h, h\rangle}hh^*$$

é uma matriz de projecção ortogonal sobre o subespaço $\mathcal{L}\{h\}$ gerado pelo vector $h$. Logo,

$$I - \frac{1}{\|h\|^2}hh^*$$

projecta ortogonalmente sobre o complemento ortogonal $\mathcal{L}\{h\}^\perp$. Assim, se o Ponto $A$ (da Figura 3.2.2) representar as coordenadas do vector $y$, então o vector

$$\left(I - \frac{1}{\|h\|^2}hh^*\right)y$$

corresponderá às coordenadas do Ponto B.

Assim sendo, a reflexão do Ponto A em torno da recta $\mathcal{L}\{h\}^\perp$, assinalado na figura por C, tem coordenadas dadas por

$$\left(I - \frac{2}{\|h\|^2}hh^*\right)y.$$

Da figura depreende-se que as coordenadas do Ponto C são as do vector $\|y\|e_1$.

A demonstração de que $Hy = \|y\|e_1$ pressupõe, por hipótese, que a primeira coordenada de $y$ é um número real.

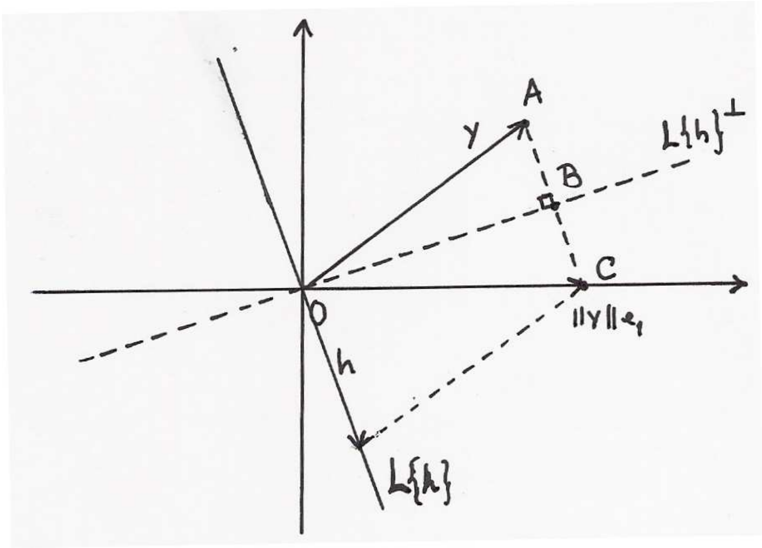



Figura 3.2.2: A interpretação geométrica de um reflector de Householder.

**Proposição 3.2.1** *Seja $y$ um vector em $\mathbb{C}^n$ com $y_1 \in \mathbb{R}$. Se $H \in M_{n\times n}(\mathbb{C})$ for o reflector de Householder associado ao vector $y$ então*

$$Hy = \|y\|e_1,$$

*em que $e_1$ representa a primeira coluna da matriz identidade de ordem $n$.*

**Demonstração.** Em primeiro lugar, calculemos o quadrado da norma do vector $h = \|y\|e_1 - y$:

$$\begin{aligned}(\|y\|e_1 - y)^* (\|y\|e_1 - y) &= \|y\|^2 - \|y\|e_1^* y - \|y\|y^* e_1 + \|y\|^2 \\ &= 2\left(-\|y\|e_1^* y + \|y\|^2\right).\end{aligned}$$

A segunda igualdade só foi possível porque $y_1$ é real.

Efectuando os restantes cálculos, obtém-se

$$\begin{aligned}\left(I - \tfrac{2}{\|h\|^2}hh^*\right)y &= \left(I - \tfrac{2}{\|\|y\|e_1-y\|^2}(\|y\|e_1 - y)(\|y\|e_1 - y)^*\right)y \\ &= y - \tfrac{2(\|y\|e_1^* y - y^* y)}{\|\|y\|e_1-y\|^2}(\|y\|e_1 - y) \\ &= y + \|y\|e_1 - y \\ &= \|y\|e_1. \blacksquare\end{aligned}$$

A restrição de $y_1$ ser real não perturba o que foi desenvolvido anteriormente na Subsecção 3.2.1. Se $y_1$ não for um número real, multiplica-se, primeiro, o vector $y$ por $y_1^*$, obtendo-se um vector $y'$ em que a primeira componente é, de certeza, um escalar real. Depois, determinar-se-ia um reflector de Householder $H'$ tal que

$$H'y' \;=\; \|y'\|e_1.$$

Dividindo esta expressão por $y_1^*$, obtém-se

$$H'y \;=\; H'\frac{1}{y_1^*}y' \;=\; \frac{\|y'\|}{y_1^*}e_1.$$

Identificou-se, assim, um reflector de Householder que anula as componentes do vector $y$ abaixo da primeira posição.

## Exercícios

1. Considere a matriz

$$H \;=\; \begin{bmatrix} -c & s \\ s & c \end{bmatrix} \;\in\; M_{2\times 2}(\mathbb{R}),$$

   com $s = \operatorname{sen}(\theta)$, $c = \cos(\theta)$ e $\theta$ um número real.

   (a) Prove que $\det(H) = -1$.

   (b) Mostre que $H$ é um reflector de Householder, identificando o vector $h$ a ele associado.

2. Sejam $A$ e $y$ a matriz e o vector dados, respectivamente, por

$$A \;=\; \begin{bmatrix} 1 & 3 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & -\sqrt{5} & 1 \end{bmatrix} \qquad \text{e} \qquad y \;=\; \begin{bmatrix} 2 \\ 0 \\ -\sqrt{5} \end{bmatrix}.$$

   (a) Calcule uma matriz ortogonal $H$ em $M_{3\times 3}(\mathbb{R})$ tal que $Hy$ tenha os elementos nas posições $2,1$ e $3,1$ nulos.

   (b) A partir do resultado da alínea (a), calcule uma matriz ortogonal $U$ em $M_{4\times 4}(\mathbb{R})$ tal que $UA$ tenha zeros na coluna 1 a seguir ao elemento na posição $1,1$ e zeros na coluna 2 a seguir ao elemento na posição $2,2$.

   (c) Seja $B$ a matriz constituída pelas duas primeiras colunas de $A$. Com base na alínea (b), indique, sem efectuar quaisquer contas, como poderia escrever $B$ na forma $QR$ em que $Q \in M_{4\times 4}(\mathbb{R})$ é ortogonal e $R \in M_{4\times 2}(\mathbb{R})$ tem zeros nas linhas 3 e 4.

   (d) Indique, sem efectuar quaisquer contas, como decomporia $B$ na forma $\hat{Q}\hat{R}$ (com $\hat{Q} \in M_{4\times 2}(\mathbb{R})$ e $\hat{R} \in M_{2\times 2}(\mathbb{R})$) a partir da decomposição $QR$ da alínea anterior.

3. Calcule os valores próprios e vectores próprios de um reflector de Householder.

4. Seja $x$ um vector de $\mathbb{C}^n$. Considere-se a matriz em $M_{n\times n}(\mathbb{C})$ que se obtém da matriz identidade alterando, apenas, os seus elementos nas posições $i,i,\ i,j,\ j,i$ e $j,j$,

$$G = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \bar{c} & \cdots & \bar{s} & & & \\ & & & \vdots & \ddots & \vdots & & & \\ & & & -s & \cdots & c & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix},$$

em que

$$c = \frac{x_i}{\sqrt{|x_i|^2 + |x_j|^2}} \quad \text{e} \quad s = \frac{x_j}{\sqrt{|x_i|^2 + |x_j|^2}}.$$

Os índices $i$ e $j$ verificam $i \neq j$ e $i,j \in \{1,\ldots,n\}$. Esta matriz é conhecida por *rotacional* de Givens (associado ao vector $x$).

   (a) Demonstre que a $j$-ésima componente de $Gx$ é nula.

   (b) Mostre que $G$ é uma matriz unitária.

   (c) Dado um vector $z \in \mathbb{C}^n$, explique como poderia determinar, utilizando $n-1$ rotacionais de Givens, uma matriz unitária $U$ tal que $Uz$ fosse um múltiplo de $e_1$ (a primeira coluna da matriz identidade de ordem $n$).

5. Considere a matriz

$$G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \in M_{2\times 2}(\mathbb{R}),$$

com $s = \mathrm{sen}(\theta)$, $c = \cos(\theta)$ e $\theta$ um número real.

   (a) Confirme que $\det(G) = 1$.

   (b) Mostre que a transformação linear associada a $G$ corresponde a uma *rotação* (em vez de uma reflexão). Em que sentido é feita a rotação?

   (c) Identifique $G$ como um rotacional de Givens.

## 3.3 Decomposição em valores singulares de uma matriz

A decomposição em valores singulares generaliza a diagonalização através de valores próprios, proporcionando uma diagonalização (num sentido mais lato) de qualquer matriz, quadrada ou rectangular.

Os valores e vectores singulares de uma matriz transmitem, fielmente, várias propriedades matriciais, constituindo um instrumento precioso em problemas envolvendo sistemas de equações lineares, mínimos quadrados e outras manipulações matriciais.

### 3.3.1 Normas matriciais

Uma das propriedades matriciais intimamente relacionada com os valores singulares é a norma euclidiana. A definição de norma faz-se em qualquer espaço vectorial. Existem diversas normas no espaço vectorial $M_{m\times n}(\mathbb{C})$. Vamos apresentar apenas duas: a norma euclidiana e a norma de Frobenius (ou do traço). Um estudo detalhado sobre normas vectoriais e matriciais (que incluiria a demonstração dos factos referidos seguidamente) está fora do contexto deste curso.

A norma de Frobenius em $M_{m\times n}(\mathbb{C})$ é definida através de um produto interno neste espaço vectorial. É possível provar que é um produto interno complexo a aplicação que, a cada duas matrizes $A$ e $B$ em $M_{m\times n}(\mathbb{C})$, faz corresponder o número complexo

$$\langle A, B\rangle \;=\; \mathrm{tr}(A^*B),$$

em que $\mathrm{tr}(A^*B)$ designa o traço de $A^*B$. Repare-se que $A^*B$ é uma matriz quadrada $n$-por-$n$. O traço de uma matriz quadrada $C \in M_{n\times n}(\mathbb{C})$ é definido pela soma dos seus elementos diagonais:

$$\mathrm{tr}(C) \;=\; \sum_{i=1}^{n} c_{ii}.$$

A norma de Frobenius é a norma (matricial) associada a este produto interno, ou seja, é a aplicação que, a cada matriz em $M_{m\times n}(\mathbb{C})$, faz corresponder o número real

$$\|A\|_F \;=\; \sqrt{\langle A, A\rangle} \;=\; \sqrt{\mathrm{tr}(A^*A)}.$$

É fácil provar que

$$\|A\|_F \;=\; \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}.$$

A norma de Frobenius de uma matriz calcula-se, deste modo, tomando a raiz quadrada da soma dos quadrados dos módulos de todos os elementos da matriz. A ordem pela qual esta soma é feita é arbitrária. Por exemplo,

$$\left\| \begin{bmatrix} 1 & i & 2 \\ 1 & 1 & 1 \end{bmatrix} \right\|_F \;=\; \sqrt{9} \;=\; 3.$$

A norma de Frobenius de uma matriz mantém-se igual se conjugarmos os seus elementos, pois todos são tomados em módulo. Destes dois factos simples resulta imediatamente que

$$\|A^*\|_F \;=\; \sqrt{\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|^2} \;=\; \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} \;=\; \|A\|_F.$$

A simplicidade do processo de cálculo da norma de Frobenius e o facto de ser definida por um produto interno tornam esta norma atraente em vários contextos.

Outra norma matricial relevante é a norma euclidiana, definida por intermédio da norma euclidiana vectorial. A norma euclidiana faz corresponder, a cada matriz em $M_{m\times n}(\mathbb{C})$, o número real

$$\|A\| \;=\; \max_{x\in\mathbb{C}^n:\, x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

Recorde-se que

$$\|y\| \;=\; \sqrt{\langle y, y\rangle} \;=\; \sqrt{\sum_{i=1}^{n}|y_i|^2} \qquad (y \;\in\; \mathbb{C}^n).$$

A demonstração de que estamos na presença de uma norma é omitida. Vamos, apenas, indicar por que motivo está esta aplicação bem definida. Veja-se que

$$\|A\| \;=\; \max_{x\in\mathbb{C}^n:\, x\neq 0} \frac{\left\|\frac{1}{\|x\|}Ax\right\|}{\left\|\frac{1}{\|x\|}x\right\|}.$$

Logo, efectuando a mudança de variável $z = (1/\|x\|)x$, obtemos

$$\|A\| \;=\; \max_{z\in\mathbb{C}^n:\, \|z\|=1} \|Az\|.$$

Observa-se, desta forma, que a norma euclidiana de $A$ é o máximo de $\|Az\|$ (uma função contínua de $z$) em $\{z \in \mathbb{C}^n :\; \|z\| = 1\}$ (um conjunto limitado e fechado em $\mathbb{C}^n$). Pelo Teorema de Weirstrass existe (pelo menos) um $w \in \mathbb{C}^n$ tal que

$$\|A\| \;=\; \|Aw\| \quad \text{e} \quad \|w\| = 1.$$

A matriz $A \in M_{m\times n}(\mathbb{C})$ define uma transformação linear $A : \mathbb{C}^n \to \mathbb{C}^m$ que, a cada $x \in \mathbb{C}^n$ faz corresponder o vector $Ax \in \mathbb{C}^m$. A norma euclidiana está relacionada com a forma como a transformação linear $A$ distorce a esfera de raio 1 centrada na origem 0,

$$E(0;1) \;=\; \{x \in \mathbb{C}^n :\; \|x\| = 1\}\,.$$

Este efeito fica registado pela imagem, por $A$, da esfera unitária $E(0;1)$, dada por

$$A(E(0;1)) \;=\; \{Ax \in \mathbb{C}^n :\; \|x\| = 1\}\,.$$

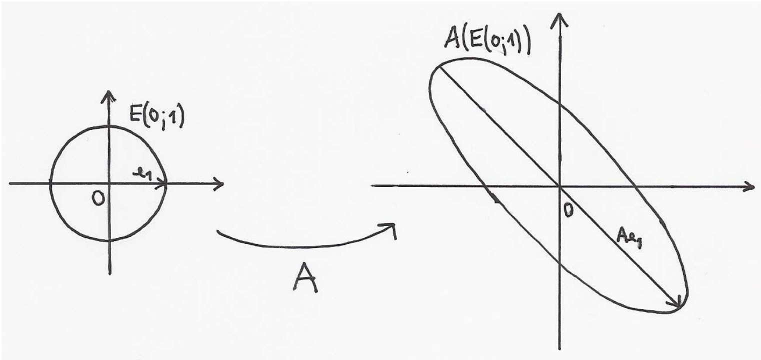



Figura 3.3.3: O efeito de $A$ sobre a circunferência $E(0;1)$ e a interpretação geométrica da norma de $A$.

Vejamos o exemplo em $M_{m\times n}(\mathbb{R})$ com $m = n = 2$ e

$$A = \begin{bmatrix} \frac{3\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{3\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

A circunferência unitária $E(0;1)$ em $\mathbb{R}^2$ e a sua imagem $A(E(0;1))$ por $A$ estão desenhadas na Figura 3.3.3. Neste exemplo tem-se que $\|A\| = 3$. Um dos vectores a fazer o papel do vector $w$ é a primeira coluna da matriz identidade de ordem 2:

$$\|Ae_1\| = \left\| \begin{bmatrix} \frac{3\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{3\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\| = 3 \left\| \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} \right\| = 3.$$

O outro vector nestas circunstâncias é $w = -e_1$.

A imagem $A(E(0;1))$ é uma elipse em $\mathbb{R}^2$. (Em $\mathbb{R}^n$, obter-se-ia uma *hiperelipse*). O eixo de maior comprimento desta elipse vai, da esquerda para a direita na figura, de $A(-e_1)$ a $Ae_1$. A norma de $A$ é o factor de distorção máxima da bola unitária pela transformação linear $A$.

### 3.3.2 Interpretação geométrica

Os valores singulares de uma matriz têm uma interpretação geométrica no âmbito da que foi dada para a norma euclidiana.

Sejam $u_1$ e $u_2$ e $v_1$ e $v_2$ dois pares de vectores ortonormados, para os quais existem escalares reais $\sigma_1 \geq \sigma_2 \geq 0$ tais que

$$Av_1 = \sigma_1 u_1 \quad \text{e} \quad Av_2 = \sigma_2 u_2,$$

ou, de forma equivalente,

$$A[\, v_1 \; v_2 \,] = [\, u_1 \; u_2 \,] \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}.$$

Utilizando as designações

$$\hat{U} \;=\; [\,u_1\, u_2\,], \quad V \;=\; [\,v_1\, v_2\,] \quad \text{e} \quad \hat{\Sigma} \;=\; \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix},$$

vem que

$$AV \;=\; \hat{U}\hat{\Sigma}.$$

As matrizes $\hat{U}$ e $V$ são unitárias. Logo, $AV = \hat{U}\hat{\Sigma}$ é equivalente a

$$A \;=\; \hat{U}\hat{\Sigma}V^*.$$

Esta expressão para $A$ é a sua decomposição em valores singulares. Os valores singulares de $A$ são os reais $\sigma_1$ e $\sigma_2$. Os vectores $u_1$ e $u_2$ são os vectores singulares de $A$ à esquerda associados, respectivamente, aos valores singulares $\sigma_1$ e $\sigma_2$. De igual forma, os vectores $v_1$ e $v_2$ são os vectores singulares de $A$ à direita associados, respectivamente, aos valores singulares $\sigma_1$ e $\sigma_2$.

Regressando ao exemplo anterior, identificamos

$$v_1 \;=\; -e_1 \;=\; \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad \text{e} \quad v_2 \;=\; e_2 \;=\; \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Assim,

$$Av_1 \;=\; 3\begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \;=\; 3u_1 \quad \text{com} \quad u_1 \;=\; \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

e

$$Av_2 \;=\; \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \;=\; 1u_2 \quad \text{com} \quad u_2 \;=\; \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}.$$

Calcularam-se, assim, os valores singulares

$$\sigma_1 \;=\; 3 \quad \text{e} \quad \sigma_2 \;=\; 1.$$

Os vectores $v_1$ e $v_2$ são os vectores na circunferência $E(0;1)$ da Figura 3.3.4. Representam-se, na elipse $A(E(0,1))$, os vectores $\sigma_1 u_1$ e $\sigma_2 u_2$. O maior valor singular de $A$ coincide com a sua norma, uma coincidência que será enunciada e provada rigorosamente mais adiante. Os valores singulares de $A$ são os factores de distorção máxima e mínima de $E(0;1)$ pela transformação linear $A$.

Reunindo toda a informação, apresentamos a decomposição em valores singulares de $A$ determinada:

$$A \;=\; \hat{U}\hat{\Sigma}V^* \;=\; \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^*.$$

Aproveita-se esta ocasião para indicar que a decomposição em valores singulares de uma matriz não é única. Se tivéssemos escolhido $v_1 = e_1$ e $v_2 = e_2$, teríamos chegado a

$$A \;=\; \hat{U}\hat{\Sigma}V^* \;=\; \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^*.$$

Existe, no entanto, um certo tipo de unicidade na decomposição em valores singulares, como este exemplo deixa antever e como veremos mais à frente.

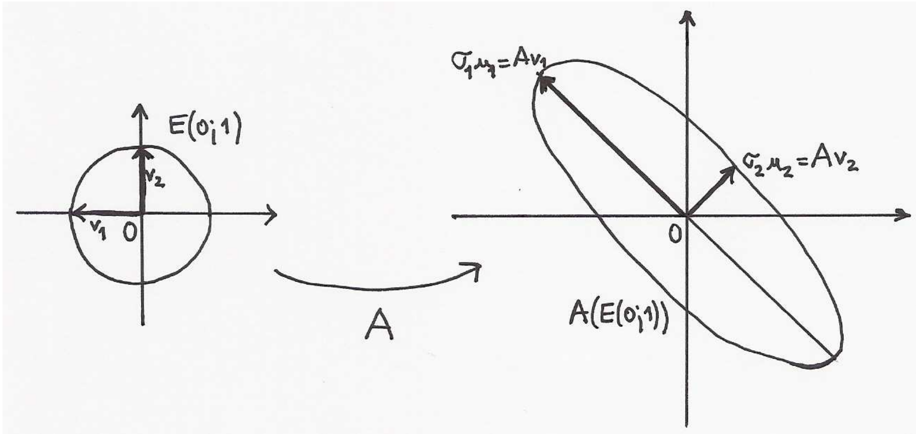



Figura 3.3.4: O efeito de $A$ sobre a circunferência $E(0;1)$ e a interpretação geométrica dos valores e vectores singulares de $A$.

### 3.3.3 Formas reduzida e completa

A definição seguinte enquadra a matriz do exemplo anterior, assim como qualquer matriz quadrada ($m = n$), e generaliza o conceito da decomposição em valores singulares (DVS) a matrizes rectangulares. Nesta primeira fase, consideram-se matrizes rectangulares em que $m > n$ ou, por outras palavras, estudam-se matrizes com mais linhas do que colunas.

**Definição 3.3.1** *Seja $A \in M_{m\times n}(\mathbb{C})$ com $m \geq n$. Um decomposição em valores singulares (DVS) de $A$ (na forma reduzida) é um produto da forma*

$$A \;=\; \hat{U}\hat{\Sigma}V^*,$$

*em que $\hat{U} \in M_{m\times n}(\mathbb{C})$ tem colunas ortonormadas, $\hat{\Sigma}$ é uma matriz diagonal $n$-por-$n$ com elementos diagonais reais satisfazendo*

$$\sigma_1 \;\geq\; \sigma_2 \;\geq\; \cdots \;\geq\; \sigma_n \;\geq\; 0$$

*e $V$ é uma matriz unitária em $M_{n\times n}(\mathbb{C})$.*

Esquematicamente, quando $m = 5$ e $n = 3$, a forma reduzida da DVS de $A$ apresenta as seguintes dimensões:

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}} \;=\; \underset{\hat{U}}{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}} \;\; \underset{\hat{\Sigma}}{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}} \;\; \underset{V^*}{\begin{bmatrix} \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond \end{bmatrix}}.$$

O exemplo visto na subsecção anterior era para o caso $m = n = 2$. Listamos outros exemplos, não necessariamente com matrizes quadradas:

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

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

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

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -5i & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & 1 \end{bmatrix},$$

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

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

Está fora do âmbito deste curso o processo de cálculo da DVS. Espera-se, porém, que os exemplos apresentados tenham deixado pistas para a determinação da DVS em casos simples.

A forma completa da decomposição em valores singulares no caso $m \geq n$ é formalizada na seguinte definição.

**Definição 3.3.2** *Seja $A \in M_{m\times n}(\mathbb{C})$ com $m \geq n$. Um decomposição em valores singulares (DVS) de $A$ (na forma completa) é um produto da forma*

$$A = U\Sigma V^*,$$

*em que $U \in M_{m\times m}(\mathbb{C})$ e $V \in M_{n\times n}(\mathbb{C})$ são matrizes unitárias e $\Sigma$ é da forma*

$$\begin{bmatrix} \hat{\Sigma} \\ 0 \end{bmatrix} \in M_{m\times n}(\mathbb{C}),$$

*com $\hat{\Sigma}$ uma matriz diagonal $n\times n$ cujos elementos diagonais são números reais a satisfazer*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Esquematicamente, quando $m = 5$ e $n = 3$, a forma completa da DVS de $A$ apresenta as seguintes dimensões:

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}} = \underset{U}{\begin{bmatrix} * & * & * & \star & \star \\ * & * & * & \star & \star \\ * & * & * & \star & \star \\ * & * & * & \star & \star \\ * & * & * & \star & \star \end{bmatrix}} \underset{\Sigma}{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}} \underset{V^*}{\begin{bmatrix} \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond \end{bmatrix}}.$$

Quando $m = n$ as formas reduzida e completa coincidem. Se $m > n$, passa-se da forma completa para a forma reduzida eliminando as últimas $m - n$ colunas de $U$ e as últimas $m - n$ linhas (de zeros) de $\Sigma$. Para passar da forma reduzida para a completa é necessário acrescentar $m - n$ linhas de zeros a $\hat{\Sigma}$ e completar as colunas de $\hat{U}$ de forma a obter uma base ortonormada para $\mathbb{C}^m$. Escolha-se num dos exemplos anteriores ($m = 3$ e $n = 2$) e escreva-se uma forma completa para a DVS, assinalando entre parêntesis curvos os elementos acrescentados:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -5i & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & (0) \\ 0 & 0 & (i) \\ 1 & 0 & (0) \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 0 \\ (0) & (0) \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & 1 \end{bmatrix},$$

Note-se que a forma de completar $\hat{U}$ não é única.

A situação rectangular em que existem menos linhas do que colunas ($m < n$) resulta da aplicação da DVS à matriz adjunta $A^*$. Como $A^*$ tem mais linhas do que colunas, esta matriz admite uma DVS na forma reduzida

$$A^* = \hat{\bar{U}}\hat{\bar{\Sigma}}\bar{V}^*.$$

Transconjugando, obtemos a seguinte decomposição em valores singulares para $A$ na forma reduzida:

$$A = \bar{V}\hat{\bar{\Sigma}}\hat{\bar{U}}^* = U\hat{\Sigma}\hat{V}^*,$$

com $U = \bar{V}$, $\hat{\Sigma} = \hat{\bar{\Sigma}}$ e $\hat{V} = \hat{\bar{U}}$. Esquematicamente, quando $m = 3$ e $n = 5$, a forma reduzida da DVS de $A$ apresenta as seguintes dimensões:

$$\underset{A}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} = \underset{U}{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}} \underset{\hat{\Sigma}}{\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}} \underset{\hat{V}^*}{\begin{bmatrix} \diamond & \diamond & \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond & \diamond & \diamond \end{bmatrix}}.$$

A forma completa no caso $m < n$ consistiria em completar as colunas de $\hat{V}$ (ou as linhas de $\hat{V}^*$) de forma a obter uma base ortonormada para $\mathbb{C}^n$ e a acrescentar $n - m$ colunas de zeros a $\hat{\Sigma}$. Esquematicamente, quando $m = 3$ e $n = 5$, a forma completa da DVS de $A$ apresenta as seguintes dimensões:

$$\underset{A}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} = \underset{U}{\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}} \underset{\Sigma}{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 \end{bmatrix}} \underset{V^*}{\begin{bmatrix} \diamond & \diamond & \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond & \diamond & \diamond \\ \star & \star & \star & \star & \star \\ \star & \star & \star & \star & \star \end{bmatrix}}.$$

A passagem da forma completa para a forma reduzida no caso $m < n$ é igualmente simples e passaria por eliminar as últimas $n-m$ colunas (de zeros) de $\Sigma$ e as últimas $n-m$ colunas de $V$ (ou as últimas $n - m$ linhas de $V^*$).

### 3.3.4 Existência e unicidade

Uma das questões à qual é preciso dar resposta é a existência de decomposições em valores singulares. O teorema seguinte garante a existência de uma DVS para qualquer matriz. A demonstração é feita recorrendo à forma completa no caso $m \geq n$. Vimos anteriormente, como passar da forma completa à reduzida e como converter o formato rectangular vertical no formato rectangular horizontal. Uma vez garantida a existência para a forma reduzida com as dimensões $m \geq n$, todos os outros casos aparecem como simples corolários.

**Teorema 3.3.1** *Toda a matriz $A$ em $M_{m\times n}(\mathbb{C})$ admite uma decomposição em valores singulares.*

**Demonstração.** Sem perda de generalidade concentrar-nos-emos no formato quadrangular ($m = n$) ou rectangular vertical ($m > n$). O caso $n = 1$ é trivial e deixado como exercício. Vamos considerar, assim, que $n > 1$, o que implica $m > 1$. A demonstração é feita por indução matemática sobre $m$ e $n$. Prova-se que a tese é válida para $m$ e $n$ partindo do princípio que ela é verdadeira para $m - 1$ e $n - 1$.

Tome-se

$$\sigma_1 = \|A\|.$$

Pela definição de norma euclidiana, sabe-se que existe um vector $v_1 \in \mathbb{C}^n$ tal que

$$\|A\| = \|Av_1\| \quad \text{e} \quad \|v_1\| = 1.$$

Faça-se

$$u_1 = Av_1 \in \mathbb{C}^m.$$

Sejam, ainda,

$$\left\{\frac{1}{\|u_1\|}u_1, u_2, \ldots, u_m\right\} \quad \text{e} \quad \{v_1, v_2, \ldots, v_n\}$$

bases ortonormadas de, respectivamente, $\mathbb{C}^m$ e $\mathbb{C}^n$. (Como é que se garante a existência destas bases?) Ponha-se

$$U_1 \;=\; \left[\begin{array}{cccc} \frac{1}{\|u_1\|}u_1 & u_2 & \cdots & u_m \end{array}\right] \quad \text{e} \quad V_1 \;=\; \left[\begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array}\right].$$

Vem, então, que

$$\begin{aligned}
S \;=\; U_1^* A V_1 \;&=\; \left[\begin{array}{c} \frac{1}{\|u_1\|}u_1^* \\ u_2^* \\ \vdots \\ u_m^* \end{array}\right] \left[\begin{array}{cccc} Av_1 & Av_2 & \cdots & Av_n \end{array}\right] \\
&=\; \left[\begin{array}{c} \frac{1}{\|u_1\|}u_1^* \\ u_2^* \\ \vdots \\ u_m^* \end{array}\right] \left[\begin{array}{cccc} u_1 & Av_2 & \cdots & Av_n \end{array}\right] \\
&=\; \left[\begin{array}{cc} \sigma_1 & w^* \\ 0 & B \end{array}\right].
\end{aligned}$$

Dada a ortogonalidade entre os vectores $u_i$, $i = 1, \ldots, m$, obtiveram-se zeros abaixo de $\sigma_1$. Utilizou-se, também, o facto de $\sigma_1 = \|Av_1\| = \|u_1\|$. O vector $w$ está em $\mathbb{C}^{n-1}$ e a matriz $B$ em $M_{m-1\times n-1}(\mathbb{C})$.

Para que $S$ fosse adquirindo a forma de uma matriz diagonal seria conveniente que $w$ fosse nulo. O passo seguinte da demonstração consiste, precisamente, em provar que $w = 0$.

Comecemos por limitar, inferiormente, a norma de $S$. Efectuando cálculos com a norma vectorial em $\mathbb{C}^m$, vem que

$$\left\|\left[\begin{array}{cc} \sigma_1 & w^* \\ 0 & B \end{array}\right]\left[\begin{array}{c} \sigma_1 \\ w \end{array}\right]\right\|^2 \;=\; \left\|\left[\begin{array}{c} \sigma_1^2 + w^*w \\ Bw \end{array}\right]\right\|^2 \;=\; \left(\sigma_1^2 + w^*w\right)^2 + \|Bw\|^2 \;\geq\; \left(\sigma_1^2 + w^*w\right)^2.$$

Logo, tomando raízes quadradas e fazendo mais cálculos vectoriais,

$$\left\|S\left[\begin{array}{c} \sigma_1 \\ w \end{array}\right]\right\| \;\geq\; \sigma_1^2 + w^*w \;=\; \sqrt{\sigma_1^2 + w^*w}\left\|\left[\begin{array}{c} \sigma_1 \\ w \end{array}\right]\right\|.$$

Combinando esta desigualdade com a propriedade da norma matricial euclidiana $\|Cx\| \leq \|C\|\|x\|$ (ver exercício nesta secção), resulta que

$$\sqrt{\sigma_1^2 + w^*w}\left\|\left[\begin{array}{c} \sigma_1 \\ w \end{array}\right]\right\| \;\leq\; \left\|S\left[\begin{array}{c} \sigma_1 \\ w \end{array}\right]\right\| \;\leq\; \|S\|\left\|\left[\begin{array}{c} \sigma_1 \\ w \end{array}\right]\right\|.$$

Provou-se, deste modo, que

$$\|S\| \;\geq\; \sqrt{\sigma_1^2 + w^* w}.$$

Por outro lado, como $U_1$ e $V_1$ são matrizes unitárias, as normas de $A$ e de $S$ coincidem (ver exercícios), o que permite afirmar que

$$\|S\| \;=\; \|A\| \;=\; \sigma_1.$$

Combinando estas duas relações envolvendo $\|S\|$, escrevemos, depois de tomar os quadrados,

$$\sigma_1^2 \;\geq\; \sigma_1^2 + w^* w,$$

o que implica que $\|w\|^2 = w^* w = 0$, ou seja, que $w = 0$.

Vamos, neste momento da demonstração, aplicar a hipótese de indução à matriz $B \in M_{m-1 \times n-1}(\mathbb{C})$, garantindo a existência de matrizes unitárias $U_2 \in M_{m-1\times m-1}(\mathbb{C})$ e $V_2 \in M_{n-1\times n-1}(\mathbb{C})$ tais que

$$U_2^* B V_2 \;=\; \Sigma_2 \;=\; \begin{bmatrix} \sigma_2 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{bmatrix}.$$

Os valores singulares de $B$ aparecem por ordem decrescente ($\sigma_2 \geq \cdots \geq \sigma_n \geq 0$).

O mesmo tipo de cálculos utilizado na demonstração do Teorema de Schur conduzir-nos-ia a

$$\begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} U_1^* A V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} \;=\cdots=\; \begin{bmatrix} \sigma_1 & 0 \\ 0 & U_2^* B V_2 \end{bmatrix} \;=\; \begin{bmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}.$$

Para terminar, basta identificar a DVS de $A$:

$$U^* A V \;=\; \Sigma$$

com

$$U \;=\; U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}, \quad V \;=\; V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix} \quad \text{e} \quad \Sigma \;=\; \begin{bmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}.$$

As matrizes $U$ e $V$ assim definidas são unitárias (um tipo de factos também necessário na demonstração do Teorema de Schur).

Observe-se que os valores singulares de $A$ respeitam a ordem decrescente da DVS. Sabemos que $\sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Além disso, as normas de $A$ e de $B$ satisfazem $\|A\| \geq \|B\|$ (porquê?). Logo,

$$\sigma_1 \;=\; \|A\| \;\geq\; \|B\| \;=\; \sigma_2,$$

e não restam quaisquer dúvidas que construímos uma decomposição em valores singulares para a matriz $A$. ■

A decomposição em valores singulares de uma matriz apresenta propriedades de unicidade semelhantes às da diagonalização por valores e vectores próprios. Entende-se por *sinal complexo* um número complexo cujo módulo vale 1.

**Teorema 3.3.2** *Os valores singulares de uma matriz $A$ em $M_{m\times n}(\mathbb{C})$ são únicos. Se os valores singulares forem distintos dois a dois então os vectores singulares (à esquerda e à direita) são únicos a menos de um sinal complexo.*

A demonstração da unicidade da DVS é omitida. Dada a unicidade dos valores singulares e quando estes são distintos dois a dois, é fácil observar geometricamente (nos casos reais bi ou tridimensionais) que os vectores singulares só podem ser modificados por uma multiplicação por $-1$. Foi esta a forma como, para a matriz do exemplo correspondente à Figura 3.3.4, apresentámos uma DVS alternativa.

### 3.3.5 Propriedades

Em todos os exemplos apresentados o número de valores singulares não nulos coincidiu com a característica da respectiva matriz.

**Teorema 3.3.3** *Se $A$ em $M_{m\times n}(\mathbb{C})$ tiver característica $r$ então $A$ tem $r$ valores singulares positivos.*

**Demonstração.** Uma vez que $U$ é quadrada e com característica $m$ e que $V^*$ é quadrada e com característica $n$, sabemos, pelas propriedades da característica de uma matriz, que

$$\mathrm{car}(A) \;=\; \mathrm{car}(U\Sigma V^*) \;=\; \mathrm{car}(\Sigma).$$

Mas $\Sigma$ é uma matriz em escada com $r$ *pivots* (os $\sigma$'s positivos), o que mostra o pretendido.
■

Os vectores singulares à esquerda contêm uma base para o espaço das colunas de $A$. Os vectores singulares à direita contêm uma base para o espaço nulo de $A$. Estes factos são enunciados e provados de seguida.

**Teorema 3.3.4** *Seja $A \in M_{m\times n}(\mathbb{C})$ uma matriz com característica $r$. Seja $A = U\Sigma V^*$ a sua DVS na forma completa. Então:*

1. *Uma base para $C(A)$ é dada por $\{u_1, \ldots, u_r\}$.*

2. *Uma base para $N(A)$ é dada por $\{v_{r+1}, \ldots, v_n\}$.*

**Demonstração.** Como $\dim(C(A)) = r$ e $\dim(N(A)) = n - r$ e quer as colunas de $U$ quer as de $V$ são linearmente independentes, basta provar que $u_i \in C(A)$, $i = 1, \ldots, r$, e que $v_i \in N(A)$, $i = r + 1, \ldots, n$.

É fácil verificar que $u_i \in C(A)$, para um dado $i$ em $\{1, \ldots, r\}$, multiplicando $A$ por $v_i$:

$$Av_i \;=\; U\Sigma V^* v_i \;=\; U\Sigma e_i \;=\; \sigma_i U e_i \;=\; \sigma_i u_i,$$

em que $e_i$ designa a $i$-ésima coluna da matriz identidade de ordem $n$. Como $Av_i \in C(A)$, tem-se que $u_i = (1/\sigma_i)Av_i \in C(A)$.

Para mostrar que $v_i \in N(A)$, com $i \in \{r+1, \ldots, n\}$, calcula-se, novamente, o mesmo produto $Av_i$, mas desta vez com índices $i$ correspondentes às colunas nulas de $\Sigma$, obtendo-se

$$Av_i \;=\; U\Sigma V^* v_i \;=\; U\Sigma e_i \;=\; U0 \;=\; 0. \;\blacksquare$$

A matriz $A^*A$ é hermítica e os seus valores próprios são escalares reais. O produto $A^*A$, com $A$ dada pela forma completa da sua DVS, escreve-se como

$$\begin{aligned} A^*A &= (U\Sigma V^*)^*(U\Sigma V^*) \\ &= V\Sigma^*(U^*U)\Sigma V^* \\ &= V\Sigma^2 V^*. \end{aligned}$$

A matriz $\Sigma^2 = \Sigma^*\Sigma$ é uma matriz diagonal $n \times n$. Observamos, assim, que $A^*A$ e $\Sigma^2$ são unitariamente semelhantes e, como tal, têm os mesmos valores próprios. Logo, os valores próprios não nulos de $A^*A$ são $\sigma_1^2, \ldots, \sigma_r^2$. Reunimos esta informação no seguinte enunciado.

**Teorema 3.3.5** *Seja $A \in M_{m\times n}(\mathbb{C})$ uma matriz com característica $r$. Então os $r$ valores singulares não nulos de $A$ são as raízes quadradas dos $r$ valores próprios não nulos de $A^*A$.*

Um primeiro corolário deste teorema indica-nos um processo para o cálculo da norma euclidiana de uma matriz. Do enunciado do teorema concluímos que o maior valor próprio de $A^*A$ (que designaremos por $\rho(A^*A)$) coincide com o maior valor singular de $A$ (conhecido por $\sigma_1$) ao quadrado.

**Corolário 3.3.1**

$$\|A\| \;=\; \sqrt{\rho(A^*A)}.$$

O cálculo da norma euclidiana de matrizes diagonais, por exemplo, passa a ser trivial. Vejamos o seguinte caso concreto.

$$A \;=\; \begin{bmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \;\Longrightarrow\; A^*A \;=\; A^2 \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 4 \end{bmatrix} \;\Longrightarrow\; \|A\| \;=\; \sqrt{9} \;=\; 3.$$

Outra consequência imediata prende-se com a relação entre valores próprios e valores singulares de matrizes hermíticas.

**Corolário 3.3.2** *Se $A$ for hermítica então os seus valores singulares são os valores absolutos dos seus valores próprios.*

**Demonstração.** Como $A^*A = V\Sigma^2 V^*$ e $A = A^*$, vem que $A^2 = V\Sigma^2 V^*$. Logo, os valores próprios de $A^2$ são os valores singulares de $A$ ao quadrado.

Por outro lado, se $A = RDR^*$ for uma diagonalização unitária da matriz hermítica $A$ definida pelos seus valores e vectores próprios, tem-se que $A^2 = RD^2R^*$. Logo, os valores próprios de $A^2$ são os valores próprios de $A$ ao quadrado.

Destas duas observações conclui-se o que se pretende provar. ■

A matriz hermítica

$$A = \begin{bmatrix} 1 & 4i \\ -4i & 1 \end{bmatrix}$$

tem valores próprios $\lambda_1 = -3$ e $\lambda_2 = 5$ (confirme). Consequentemente, os seus valores singulares são $\sigma_1 = 5$ e $\sigma_2 = 3$. A norma euclidiana de $A$ é igual a 5.

Terminamos esta subsecção dedicada às propriedades da decomposição em valores singulares com uma fórmula para a forma reduzida da DVS muito utilizada em aplicações.

**Teorema 3.3.6** *Uma matriz $A \in M_{m\times n}(\mathbb{C})$ com característica $r$ pode ser escrita na forma*

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^*,$$

*em que $u_i$, $i = 1, \ldots, r$, são os primeiros $r$ vectores singulares à esquerda de $A$, $v_i$, $i = 1, \ldots, r$, são os primeiros $r$ vectores singulares à direita de $A$ e $\sigma_i$, $i = 1, \ldots, r$, são os valores singulares positivos de $A$.*

A demonstração deste facto resume-se a multiplicação por blocos. O caso geral é deixado como exercício. Exemplifiquemos a demonstração com o caso $m = 2$ e $n = 3$ (com $\sigma_1 \geq \sigma_2 > 0$):

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \end{bmatrix}} = \underset{U}{\begin{bmatrix} * & * \\ * & * \end{bmatrix}} \underset{\hat{\Sigma}}{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}} \underset{\hat{V}^*}{\begin{bmatrix} \diamond & \diamond & \diamond \\ \diamond & \diamond & \diamond \end{bmatrix}}.$$

Designamos por $u_1$ e $u_2$ as colunas de $U$ e por $v_1^*$ e $v_2^*$ as linhas de $\hat{V}^*$ (ou seja, as colunas de $\hat{V}$ conjugadas). Multiplicando por blocos, conseguimos exprimir $A$ da forma desejada:

$$A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 v_1^* \\ \sigma_2 v_2^* \end{bmatrix} = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^*.$$

Cada matriz da forma $u_i v_i^*$ tem característica igual a 1. Um produto da forma $uv^*$, com $u \in \mathbb{C}^m$ e $v \in \mathbb{C}^n$ pode ser escrito na forma

$$uv^* = u \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}^* = u \begin{bmatrix} \overline{v_1} & \cdots & \overline{v_n} \end{bmatrix} = \begin{bmatrix} \overline{v_1}\, u & \cdots & \overline{v_n}\, u \end{bmatrix}.$$

Observa-se que todas as colunas de $uv^*$ são múltiplos do vector $u$. A característica de $uv^*$ é, mesmo, igual a 1. A partir desta observação e do último teorema podemos afirmar que qualquer matriz (com característica $r$) é uma combinação linear de $r$ matrizes de característica 1. Os coeficientes desta combinação linear são os valores singulares $\sigma_1, \ldots, \sigma_r$ da matriz. Em determinadas aplicações, aparecem matrizes cujos valores singulares mais pequenos deveriam ser nulos, mas não o são por determinados motivos (por exemplo, ruído em experiências). É frequente, nestes casos, substituir esses valores singulares por zero, desprezando as suas contribuições, e considerar uma matriz aproximada com menos termos na combinação linear das matrizes de característica 1.

## Exercícios

1. Mostre que
$$\|A\|_F \;=\; \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} \qquad (A \;\in\; M_{m\times n}(\mathbb{C}))\,.$$

2. Calcule a norma de Frobenius das seguintes matrizes:
$$\begin{bmatrix} i & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 & 0 \\ 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 & i \end{bmatrix}, \quad \begin{bmatrix} \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} \\ \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} \\ \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} \\ \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} \\ \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} & \sqrt{5} \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & 5 & 4 & 3 & 2 \\ 2 & 1 & 5 & 4 & 3 \\ 3 & 2 & 1 & 5 & 4 \\ 4 & 3 & 2 & 1 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{bmatrix}.$$

3. Mostre que $\|Ax\| \leq \|A\|\|x\|$, em que $A \in M_{m\times n}(\mathbb{C})$ e $x \in \mathbb{C}^n$.

4. Considere uma matriz $A \in M_{m\times n}(\mathbb{C})$.

   (a) Seja $U \in M_{m\times m}(\mathbb{C})$ uma matriz unitária. Prove que
   $$\|UA\|_F \;=\; \|A\|_F \quad \text{e} \quad \|UA\| \;=\; \|A\|.$$
   Sugestão: Utilize as fórmulas $\|C\|_F = \sqrt{\operatorname{tr}(C^*C)}$ e $\|C\| = \sqrt{\rho(C^*C)}$.

   (b) Considere, agora, $V \in M_{n\times n}(\mathbb{C})$ uma matriz unitária. Mostre que
   $$\|AV\|_F \;=\; \|A\|_F \quad \text{e} \quad \|AV\| \;=\; \|A\|.$$
   Sugestão: No caso da norma de Frobenius recorra a $\|C^*\|_F = \|C\|_F$ e ao resultado da alínea anterior. Para a norma euclidiana use $\|C\| = \max_{\|z\|=1}\|Cz\|$.

5. Seja $D$ uma matriz diagonal $n$-por-$n$ com elementos complexos. Calcule $\|D\|$ e $\|D\|_F$ em função dos elementos diagonais de $D$.

6. Sejam $u \in \mathbb{C}^m$ e $v \in \mathbb{C}^n$. Considere a matriz definida por
$$A \;=\; uv^* \;\in\; M_{m\times n}(\mathbb{C}).$$

(a) Mostre que

$$\|Ax\| \;\leq\; \|u\|\|v\|\|x\| \qquad \forall x \in \mathbb{C}^n.$$

Conclua que $\|A\| \leq \|u\|\|v\|$.

(b) Faça $x = v$ e conclua, daqui, que $\|A\| = \|u\|\|v\|$.

(c) O que pode dizer sobre $\|A\|_F$?

7. Calcule decomposições em valores singulares (nas formas reduzida e completa) das seguintes matrizes:

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} i & i \\ i & i \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 0 & 0 \\ -2 & 0 \\ 0 & 0 \end{bmatrix}.$$

8. Calcule as normas euclidianas das matrizes dadas no exercício anterior.

9. Considere uma matriz $A$, 4-por-2, cuja decomposição em valores singulares (na forma reduzida) é dada por:

$$A \;=\; \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \\ \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

(a) Qual é a característica da matriz $A$? E a sua norma euclidiana?

(b) Escreva a matriz $A$ como uma soma de duas matrizes de característica 1.

(c) Indique os valores próprios de $A^*A$.

(d) Escreva uma forma completa para esta decomposição.

10. Escreva a DVS (formas reduzida e completa) de uma matriz-coluna (ou vector) $u \in M_{m\times 1}(\mathbb{C})$.

11. Escreva a DVS (formas reduzida e completa) de uma matriz-linha $v \in M_{1\times n}(\mathbb{C})$.

12. Seja $A \in M_{m\times n}(\mathbb{C})$ uma matriz com característica $r$ e valores singulares positivos $\sigma_1, \ldots, \sigma_r$. Mostre, usando propriedades da norma de Frobenius enunciadas em exercícios anteriores, que

$$\|A\|_F \;=\; \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}.$$

13. Utilizando as propriedades dos determinantes, mostre que o módulo do determinante de uma matriz quadrada é igual ao produto dos seus valores singulares.