Apontamentos de Complementos

de

Álgebra Linear e Geometria Analítica

Licenciatura em Tecnologias de Informação Visual
$1^{\underline{o}}$ Ano, $2^{\underline{o}}$ Semestre, Ano Lectivo de 2004/2005

Departamento de Matemática da FCTUC

Joana Teles Correia   Luís Nunes Vicente

13 de Fevereiro de 2005

Conteúdo

Capítulo 1

Complementos sobre valores e vectores próprios

1.1 Valores próprios e vectores próprios de matrizes

Revisões de ALGA: Definição de vector e valor próprios de uma matriz em $M_{n\times n}(\mathbb{C})$. Equação e polinómio característicos. Teorema fundamental da Álgebra e multiplicidade algébrica de valores próprios. Subespaço próprio e multiplicidade geométrica de valores próprios. Valores próprios de matrizes triangulares. Valores próprios de matrizes equivalentes.

1.2 Matrizes diagonalizáveis

Revisões de ALGA: Definição de matriz diagonalizável em $M_{n\times n}(\mathbb{C})$. Independência linear de vectores próprios associados a valores próprios distintos. Matrizes com n valores próprios distintos dois a dois são diagonalizáveis em $M_{n\times n}(\mathbb{C})$. Caracterização necessária e suficiente de diagonalizabilidade de uma matriz em termos das multiplicidades geométricas dos seus valores próprios distintos.

1.3 O caso das matrizes simétricas reais

Revisões de ALGA: Matrizes simétricas em $M_{n\times n}(\mathbb{R})$ têm valores próprios reais e são diagonalizáveis através de matrizes diagonalizantes ortogonais. Vectores próprios de matrizes simétricas em $M_{n\times n}(\mathbb{R})$ associados a valores próprios distintos são ortogonais.

Exemplo de aplicação de diagonalizabilidade em sistemas de equações diferenciais ordinárias.

1.4 Curvas e superfícies do 2º grau

Definição de cónica e sua equação na forma matricial em $\mathbb{R}^2$. Caracterização de uma cónica através de mudanças de coordenadas: rotação (em que a matriz de rotação é dada pela matriz diagonalizante ortogonal da matriz simétrica do termo quadrático da forma matricial) e translacção.
Definição de quadrádica e sua equação na forma matricial em $\mathbb{R}^3$. Caracterização de uma quádrica através de duas mudanças de coordenadas: rotação (em que a matriz de rotação é dada pela matriz diagonalizante ortogonal da matriz simétrica do termo quadrático da forma matricial) e translacção.

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 equivalentes 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 equivalentes 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 equivalentes com elementos reais são ortogonalmente equivalentes. 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 equivalentes A e B têm os mesmos valores próprios. Logo, duas matrizes (ortogonalmente ou) unitariamente equivalentes partilham os mesmos valores próprios.

1.5.1 O Teorema de Schur

Qualquer matriz é unitariamente equivalente 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 equivalente 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 matrix 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

$$\begin{aligned} V_2^*(U_1^*AU_1)V_2 &= \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & r_1^* \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & r_1^*U_2 \\ 0 & A_1U_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & r_1^*U_2 \\ 0 & U_2^*A_1U_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & (r_1^*U_2)_{1,1} & (r_1^*U_2)_{1,2:n} \\ 0 & \lambda_2 & r_2^* \\ 0 & 0 & A_2 \end{bmatrix}. \end{aligned}$$

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

$$U^*AU \;=\; T \quad \text{com} \quad U \;=\; U_1V_2\cdots V_{n-1}$$

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) equivalentes. 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 $U^*T_1U = T_2$ são deixadas como exercício. Deste exemplo, nomeadamente do facto de serem válidas ambas as igualdades

$$U^*T_1U \;=\; T_2 \quad \text{e} \quad I^*T_1I \;=\; T_1,$$

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

$$Q^\top AQ \;=\; T,$$

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

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*

$$A^*A \;=\; AA^*,$$

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

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

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

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

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 equivalentes 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.

$inserir figura$

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

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

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

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

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

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

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

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*

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

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*

$$H \; = \; h_0 C^0 + h_1 C + \cdots + h_{n-1} C^{n-1},$$

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

$$VHV^{-1} \; = \; V(h_0 C^0 + h_1 C + h_2 C^2)V^{-1} \; = \; h_0 V C^0 V^{-1} + h_1 V C V^{-1} + h_2 V C^2 V^{-1}.$$

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

$$VHV^{-1} \; = \; h_0 D^0 + h_1 D + h_2 D^2.$$

Concluímos que V também diagonaliza H e que os valores próprios de H são os elementos diagonais da matriz diagonal $h_0 I + h_1 D + h_2 D^2$.

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*

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

em que

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

$inserir figura$

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

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

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

Logo,

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

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

$$s_{pq} = \sum_{k=0}^{n-1}(w^{kp})(\overline{w^{kq}}) = \sum_{k=0}^{n-1}(w^{kp})(\bar{w}^{kq}) = \sum_{k=0}^{n-1}(w^{kp})(w^{-kq}) = \sum_{k=0}^{n-1} w^{k(p-q)}.$$

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 se $F\bar{F}$ 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 possivel 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:

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

com

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

Se H for não singular então Λ é não singular e

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

ou, em termos da solução g,

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

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 Λ 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.

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	$\simeq 84000$

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

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 matricias 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 uma valor próprio de H associado ao valor próprio $\lambda = 2 = d$.

Capítulo 2

Geometria analítica

Revisões de ALGA: Definições de produtos internos real e complexo. Norma e distância. Ortogonalidade. Desigualdades triangulares e de Schwartz. Teorema de Pitágoras. Complemento ortogonal de um subespaço. Soma e soma directa de subespaços. Ângulo entre vectores (em $\mathbb{R}^n$ e em $\mathbb{C}^n$).

2.1 Determinantes e medidas de paralelipípedos

Área de paralelipípedos definidos por dois vectores. Volume de paralelipípedos definidos por três vectores. Medida de paralelipípedos definidos por mais de três vectores.

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

Definição e cálculo do produto externo de dois vectores em $\mathbb{R}^3$. Norma do produto externo. Triedos directos usando o produto externo.

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

Planos de dimensão k. Rectas e hiperplanos. Paralelismo. Equação vectorial e equações paramétricas de um plano. Equações normais ou canónicas de uma recta. Equação cartesiana de um hiperplano. Distância de pontos a hiperplanos. Estudo pormenorizado dos casos bidimensional e tridimensional.

Capítulo 3

Complementos sobre problemas de mínimos quadrados

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

Revisões de ALGA: Projecção ortogonal sobre um subespaço. Processo de ortogonalização de Gram-Schmidt. Decomposição QR de uma matriz através do processo de ortogonalização de Gram-Schmidt.

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

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

Veja-se que

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

$$inserir figura$$

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

o que mostra a expressão

$$v_F \;=\; QQ^* v.$$

A esta matriz QQ^* chama-se matriz de projecção ortogonal sobre F.

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

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

Mas, porque $Q^*Q = I$, $A^*A = R^*Q^*QR = R^*R$ e

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

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

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

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

Logo,

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

Desta forma, chamamos à matriz

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

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

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

$$\begin{aligned}\langle f, v - (I - QQ^*)v\rangle &= (v - (I - QQ^*)v)^* f \\ &= (v^* - v^*(I - QQ^*))\, f \\ &= v^*f - v^*f + v^*QQ^*f \\ &= u^*(QQ^*f).\end{aligned}$$

(Note-se que $I - QQ^*$ é hermítica.) Daqui resulta que

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

É fácil de constatar que QQ^* e $I - QQ^*$ 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 QQ^* e $I - QQ^*$ 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 QQ^* ser uma matriz hermítica. O mesmo sucede com $I - QQ^*$ 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, se P é projector ortogonal então P é da forma QQ^*, em que $A = QR$ com $A = [\, v_1 \ \cdots \ v_k \,]$ e $\{v_1, \ldots, v_k\}$ uma base de F_1. Logo, $P^* = P$. ■

Exercícios

1. Mostre que QQ^* e $I - QQ^*$ 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 $I - QQ^*$ (em que Q é a matriz obtida na alínea anterior).

   (d) Verifique que $I - QQ^*$ é um projector ortogonal.

   (e) Sobre que subespaço é que a matriz $I - QQ^*$ 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 QQ^* (em que Q é a matriz obtida na alínea anterior).

   (d) Verifique que QQ^* é um projector ortogonal.

   (e) Sobre que subespaço é que a matriz QQ^* 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

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

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 Q à custa do produto de k matrizes unitárias $n \times n$:

$$\hat{Q}_k \cdots \hat{Q}_2 \hat{Q}_1 A \; = \; \hat{R},$$

o que é equivalente a

$$A \; = \; \hat{Q}\hat{R},$$

com $\hat{Q} = \hat{Q}_1^* \hat{Q}_2^* \cdots \hat{Q}_k^*$. A matriz produto $\hat{Q}$ é unitária (ver exercícios da Secção 1.5). As dimensões são as seguintes:

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

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

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

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

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

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

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

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

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

$$\underset{A}{\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}} \xrightarrow{\hat{Q}_1} \underset{\hat{Q}_1 A}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \\ 0 & \bar{\times} & \bar{\times} \end{bmatrix}} \xrightarrow{\hat{Q}_2} \underset{\hat{Q}_2\hat{Q}_1 A}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\times}} \end{bmatrix}} \xrightarrow{\hat{Q}_3} \underset{\hat{R} = \hat{Q}_3\hat{Q}_2\hat{Q}_1 A}{\begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & \bar{\bar{\times}} & \bar{\bar{\times}} \\ 0 & 0 & \bar{\bar{\bar{\times}}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}.$$

A matriz R é definida pelas três primeiras linhas de $\hat{R} = \hat{Q}_3\hat{Q}_2\hat{Q}_1 A$, ou seja, por

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

A matriz Q é dada pelas três primeiras colunas de $\hat{Q} = \hat{Q}_1^*\hat{Q}_2^*\hat{Q}_3^*$.

Regressemos ao caso geral e começemos por analisar a matriz $\hat{Q}_1 \in M_{n\times n}(\mathbb{C})$. 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 $\hat{Q}_1$ seja tal que

$$\hat{Q}_1 v_1 = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

O escalar α, que no esquema corresponde ao símbolo $\bar{\times}$ na posição $1,1$ de $\hat{Q}_1 A$, vai ser escolhido de forma a valer $\|v_1\|$, o que está de acordo como a decomposição QR calculada através do processo de ortogonalização de Gram-Schmidt. Assim,

$$\hat{Q}_1 v_1 \;=\; \|v_1\| e_1,$$

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 $\hat{Q}_1$ com as propriedades desejadas. No caso em causa, o reflector de Householder procurado é dado por

$$\hat{Q}_1 \;=\; I - \frac{2}{\|h_1\|^2} h_1 h_1^* \;\in\; M_{n\times n}(\mathbb{C}) \qquad \text{com} \qquad h_1 \;=\; \|v_1\| e_1 - v_1.$$

Confirmaremos na próxima secção que $\hat{Q}_1 v_1 = \|v_1\| e_1$.

O passo seguinte consiste em identificar a matriz $\hat{Q}_2$. Observando o esquema anterior, constata-se que $\hat{Q}_2$, ao multiplicar, à esquerda, $\hat{Q}_1 A$, deverá deixar intacta a primeira linha desta última matriz. Tal efeito é conseguido ao exigir-se que

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

Se $\bar{Q}_2 \in M_{n-1\times n-1}(\mathbb{C})$ for unitária $\hat{Q}_2$ 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 $\hat{Q}_1 A$ abaixo da posição $2,2$. Recorremos, novamente, aos reflectores de Householder para conseguir este último objectivo. A matriz $\hat{Q}_2$ é, assim, dada por

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

e $\bar{v}_2$ o vector em $\mathbb{C}^{n-1}$ formado pelos elementos da segunda coluna de $\hat{Q}_1 A$, 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:

$$\hat{Q}_1 A \;=\; \begin{bmatrix} \bar{\times} & \bar{\times} & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \\ 0 & [\bar{\times}] & \bar{\times} \end{bmatrix}.$$

Este procedimento é aplicado, sucessivamente, até ser concluída a triangularização. No passo j, com $2 \leq j \leq k$, a matriz $\hat{Q}_j$ tem a forma

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

em que $\bar{Q}_j$ é um reflector de Householder de ordem $n-j+1$. A matriz $\hat{Q}_j$ 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*

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

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 $\mathcal{L}\{h\}$, 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.

$$inserirfigura$$

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

$$Hy' = \|y'\|e_1.$$

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 $s = \text{sen}(\theta)$, $c = \cos(\theta)$ e θ 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 $\hat{Q}\hat{R}$ em que $\hat{Q} \in M_{4\times 4}(\mathbb{R})$ é ortogonal e $\hat{R} \in M_{4\times 2}(\mathbb{R})$ tem zeros nas linhas 3 e 4.
   (d) Indique, sem efectuar quaisquer contas, como decomporia B na forma QR (com $Q \in M_{4\times 2}(\mathbb{R})$ e $R \in M_{2\times 2}(\mathbb{R})$) a partir da decomposição $\hat{Q}\hat{R}$ 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 $z \in \mathbb{C}$, 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 $s = \operatorname{sen}(\theta)$, $c = \cos(\theta)$ e θ 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 matricias intimamente relacionada com os valores singulares é a norma Euclideana. 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 Euclideana 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,

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

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 Euclideana, definida por intermédio da norma Euclideana vectorial. A norma Euclideana 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 Euclideana 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 Euclideana 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

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

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

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

$inserir figura$

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

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 Euclideana.

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 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 σ_1 e σ_2. Os vectores u_1 e u_2 são os vectores singulares de A à esquerda associados, respectivamente, aos valores singulares σ_1 e σ_2. De igual forma, os vectores v_1 e v_2 são os vectores singulares de A à direita associados, respectivamente, aos valores singulares σ_1 e σ_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.

$inserir figura$

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 *Sejam m e n números inteiros positivos tais que $m \geq n$. Seja $A \in M_{m\times n}(\mathbb{C})$. 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 *Sejam m e n números inteiros positivos tais que $m \geq n$. Seja $A \in M_{m\times n}(\mathbb{C})$. Um decomposição em valores singulares (DVS) de A (na forma completa) é um produto da forma*

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

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

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

com $\hat{\Sigma}$ uma matriz diagonal n-por-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 Σ. 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

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

Transconjugando, obtemos a seguinte decomposição em valores singulares para A na forma reduzida:

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

com $U = \bar{V}$, $\hat{\Sigma} = \hat{\hat{\Sigma}}$ e $\hat{V} = \hat{\hat{U}}$. 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 Σ 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 Euclideana, 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 existências destas bases?) Ponha-se

$$U_1 = \left[\begin{array}{cccc} \frac{1}{\|u_1\|}u_1 & u_2 & \cdots & u_m \end{array}\right] \quad \text{e} \quad V_1 = \left[\begin{array}{cccc} v_1 & v_2 & \cdots & v_m \end{array}\right].$$

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 ortogalidade entre os vectores u_i, $i = 1, \ldots, m$, obtiveram-se zeros abaixo de σ_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 $w = 0$.

Começemos 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 Euclideana $\|Cx\| \leq \|C\|\|x\|$ (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 Σ é uma matriz em escada com r *pivots* (os σ's positivos), o que mostra o pretendido.
■

Os vectores singulares à esquerda contém uma base para o espaço das colunas de A. Os vectores singulares à direita contém 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 Σ, 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 Σ^2 são unitariamente diagonalizáveis 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 Euclideana 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 σ_1) ao quadrado.

Corolário 3.3.1

$$\|A\| \;=\; \sqrt{\rho(A^*A)}.$$

O cálculo da norma Euclideana 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 $m = n$ e A for hermítica então os seus valores singulares são os valores absolutos dos seus valores próprios.*

Demonstração. Como $A^*A = V\Sigma^2V^*$ e $A = A^*$, vem que $A^2 = V\Sigma^2V^*$. 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 $\lambda_1 = -1$ e $\lambda_2 = 3$ (confirme). Consequentemente, os seus valores singulares são $\sigma_1 = 3$ e $\sigma_2 = 1$. A norma Euclideana de A é igual a 3.

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 σ_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
   $$\|AV\|_F \;=\; \|V\|_F \quad \text{e} \quad \|AV\| \;=\; \|A\|.$$
   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 Euclideanas 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 Euclideana?

(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-columa (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 $A \in M_{m\times n}(C)$ 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.

$inserirfigura$

Figura 3.4.5: Projecção ortogonal de b sobre $C(A)$.

3.4 Resolução de problemas de mínimos quadrados

Consideremos um sistema de equações lineares impossível, escrito na forma

$$Ax \ = \ b,$$

com $A \in M_{m\times n}(\mathbb{C})$ e $b \in \mathbb{C}^m$. As letras m e n respresentam números inteiros positivos a satisfazer $m > n$. Uma solução x_{mn} no sentido dos mínimos quadrados deste sistema satisfaz

$$\|Ax_{mn} - b\| \ = \ \min\left\{\|Ax - b\| \, , \, x \in \mathbb{C}^n\right\}.$$

Sabemos dois factos importantes.

1. O vector Ax_{mn} é a projecção ortogonal sobre o subespaço da colunas de A. O vector x_{mn} é solução de
$$Ax \ = \ \mathrm{proj}_{C(A)}b.$$
Ver Figura 3.4.5.

2. O vector x_{mn} é solução das chamadas equações normais:
$$A^*Ax \ = \ A^*b.$$

Ambos estes sistemas são possíveis (determinados ou indeterminados consoante a característica de A). Tem-se, ainda, que

$$Ax_{mn} \ = \ \mathrm{proj}_{C(A)}b \ = \ \langle b, v_1\rangle v_1 + \cdots + \langle b, v_r\rangle v_r,$$

em que $\mathrm{car}(A) = r$ e $\{v_1 \ldots, v_r\}$ é uma base ortonormada de $C(A)$.

Se $\mathrm{car}(A) = n$ $(r = n)$ então x_{mn} é definido, unicamente, por

$$Ax \ = \ \mathrm{proj}_{C(A)}b \quad \text{ou} \quad A^*Ax \ = \ A^*b.$$

Ambos estes sistemas são, neste caso, possíveis determinados. Estes sistemas são possíveis indeterminados se $\mathrm{car}(A) < n$.

Vamos sugerir dois processos de cálculo para x_{mn}.

3.4.1 Através da decomposição QR

Substituindo A nas equações normais pela sua decomposição QR obtém-se um sistema equivalente, envolvendo apenas a matriz R:

$$\begin{aligned} A^*Ax &= A^*b \\ &\Updownarrow \\ (QR)^*(QR)x &= (QR)^*b \\ &\Updownarrow \\ R^*(Q^*Q)Rx &= R^*Q^*b \\ &\Updownarrow \\ R^*Rx &= R^*\bar{y} \quad \text{com} \quad \bar{y} = Q^*b. \end{aligned}$$

Se a característica de A for igual a n então a matriz triangular R é não singular (os seus elementos diagonais são diferentes de zero). Nesta situação, multiplicando, à esquerda, por R^* e pela sua inversa, proporciona as duas implicações da equivalência:

$$A^*Ax = A^*b \quad \Longleftrightarrow \quad Rx = \bar{y}.$$

Algoritmo 3.4.1 [Determinar x_{mn} pela decomposição QR.]

1. *Factorizar $A = QR$, pondo $\bar{y} = Q^*b$.*

2. *Se car$(R) = n$, então resolver o sistema possível determinado:*

   $$Rx = \bar{y}.$$

   Caso contrário (car$(R) < n$), determinar uma solução do sistema possível indeterminado:

   $$R^*Rx = R^*\bar{y}.$$

Note-se que a matriz do sistema $Rx = \bar{y}$ é triangular superior, mas o mesmo já não acontece com a matriz do sistema $R^*Rx = R^*\bar{y}$.

3.4.2 Através da decomposição em valores singulares

Substituindo A pela sua decomposição em valores singulares na forma reduzida obtém-se um sistema equivalente, envolvendo apenas a matriz $\hat{\Sigma}$:

$$\begin{aligned} A^*Ax &= A^*b \\ &\Updownarrow \\ (\hat{U}\hat{\Sigma}V^*)^*(\hat{U}\hat{\Sigma}V^*)x &= (\hat{U}\hat{\Sigma}V^*)^*b \\ &\Updownarrow \\ V\hat{\Sigma}^*(\hat{U}^*\hat{U})\hat{\Sigma}V^*x &= V\hat{\Sigma}^*\hat{U}^*b \\ &\Updownarrow \\ \hat{\Sigma}^2V^*x &= \hat{\Sigma}\hat{U}^*b \\ &\Updownarrow \\ \hat{\Sigma}^2\bar{x} &= \hat{\Sigma}\bar{y} \quad \text{com} \quad \bar{x} = V^*x \quad \text{e} \quad \bar{y} = \hat{U}^*b. \end{aligned}$$

A matriz do sistema $\hat{\Sigma}^2\bar{x} = \hat{\Sigma}\bar{y}$ é diagonal. Por extenso este sistema tem a forma

$$\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_n \end{bmatrix} = \begin{bmatrix} \sigma_1\bar{y}_1 \\ \vdots \\ \sigma_n\bar{y}_n \end{bmatrix}.$$

Seja r a característica de A (igual à característica de $\hat{\Sigma}$).

Vejamos, primeiro, o caso $r < n$. Os últimos $n-r$ valores singulares de A são nulos e obtemos o sistema possível indeterminado

$$\begin{bmatrix} \sigma_1^2 & & & & & \\ & \ddots & & & & \\ & & \sigma_r^2 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_r \\ \bar{x}_{r+1} \\ \vdots \\ \bar{x}_n \end{bmatrix} = \begin{bmatrix} \sigma_1\bar{y}_1 \\ \vdots \\ \sigma_r\bar{y}_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

As incógnitas $\bar{x}_{r+1}, \ldots, \bar{x}_n$ são livres. As incógnitas básicas tomam os valores $\bar{x}_i = \bar{y}_i/\sigma_i$, $i = 1, \ldots, r$.

Se a característica de A for igual a n então a matriz diagonal $\hat{\Sigma}$ é não singular (os seus elementos diagonais são diferentes de zero). Nesta situação, multiplicando, à esquerda, por $\hat{\Sigma}$ e pela sua inversa, proporciona as duas implicações da equivalência:

$$A^*Ax \; = \; A^*b \quad \Longleftrightarrow \quad \hat{\Sigma}\bar{x} \; = \; \bar{y}.$$

O sistema $\hat{\Sigma}\bar{x} = \bar{y}$ tem a solução $\hat{\Sigma}^{-1}\bar{y}$ dada por $\bar{x}_i = \bar{y}_i/\sigma_i$, $i = 1, \ldots, n$.

O algoritmo para a resolução de um problema de mínimos quadrados usando a decomposições em valores singulares é descrito de seguida. Depois de determinar $\bar{x}$, da forma acabada de explicar, há que recuperar a solução nas variáveis originais através da mudança de variáveis $\bar{x} = V^*x \Leftrightarrow x = V\bar{x}$.

Algoritmo 3.4.2 [Determinar x_{mn} pela DVS.]

1. *Factorizar $A = \hat{U}\hat{\Sigma}V^*$, pondo $\bar{y} = \hat{U}^*b$.*

2. *Se car($\hat{\Sigma}$) $= n$, então:*
$$\bar{x} \; = \; \hat{\Sigma}^{-1}\bar{y}.$$
Caso contrário (car($\hat{\Sigma}$) $< n$), determinar uma solução para o sistema possível indeterminado:
$$\hat{\Sigma}^2\bar{x} \; = \; \hat{\Sigma}\bar{y}.$$

3. *Efectuar o produto matriz-vector $x = V\bar{x}$.*

3.4.3 Pseudo-inversa de uma matriz

Quando $A \in M_{m\times n}(\mathbb{C})$ tem característica n, a matriz $A^*A \in M_{n\times n}(\mathbb{C})$ tem característica n (logo é invertível) e

$$A^+ \;=\; (A^*A)^{-1}\,A^* \;\in\; M_{n\times m}(\mathbb{C})$$

é a chamada *pseudo-inversa* de A. A pseudo-inversa de A é uma inversa *à esquerda*, no sentido em que

$$A^+A \;=\; I, \qquad I \in M_{n\times n}(\mathbb{C}).$$

A pseudo inversa de uma matriz pode ser calculada através das suas decomposições $A = QR$ e $A = \hat{U}\hat{\Sigma}V^*$:

$$A^+ \;=\; R^{-1}Q^* \quad \text{e} \quad A^+ \;=\; V\hat{\Sigma}^{-1}\hat{U}^*.$$

Exercícios

1. Seja Q a matriz dada por

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

   (a) Calcule $P = I - QQ^*$ e mostre que se trata de um projector ortogonal.

   (b) Sobre que subespaço projecta P ortogonalmente?

   (c) Calcule a solução no sentido dos mínimos quadrados de

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;=\; \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

   utilizando, apenas, a matriz Q.

   (d) Calcule a decomposição em valores singulares de Q.

2. Considere um sistema de equações lineares da forma

$$Ax \;=\; b, \qquad A \in M_{m\times n}(\mathbb{C}) \quad \text{e} \quad b \in \mathbb{C}^m,$$

   com m e n números inteiros positivos a verificar $m < n$. Assuma que a característica da matriz A é igual a m.

   (a) Quantos valores singulares não nulos tem A?

   (b) Substitua A em $Ax = b$ pela transconjugada de QR, em que Q e R são os factores da decomposição $A^* = QR$.

   (c) Multiplique, depois, ambos os membros do sistema por $(R^*)^{-1}$, indicando o resultado obtido. Classifique o sistema resultante desta operação.

3. Prove que a pseudo-inversa de uma matriz $A \in M_{m\times n}(\mathbb{C})$ com característica n pode ser escrita nas formas

$$A^+ = R^{-1}Q^* \quad \text{e} \quad A^+ = V\hat{\Sigma}^{-1}\hat{U}^*,$$

dadas as respectivas decomposições matriciais (QR e em valores singulares na forma reduzida).

4. Quando $A \in M_{m\times n}(\mathbb{C})$ tem característica m, é possível definir uma inversa *à direita* de A, dada por

$$A^{\oplus} = A^* (AA^*)^{-1}.$$

   (a) Prove que $AA^{\oplus} = I \in M_{m\times m}$.

   (b) Escreva $A^{\oplus}$ em função da decomposição QR de A^*.

   (c) Escreva $A^{\oplus}$ em função da DVS de A dada na forma reduzida.

5. Considere um sistema de equações lineares da forma

$$Ax = b, \qquad A \in M_{m\times n}(\mathbb{C}) \quad \text{e} \quad b \in \mathbb{C}^m,$$

com m e n números inteiros positivos a verificar $m < n$. Assuma que a característica da matriz A é igual a m.

   (a) Classifique o sistema e substitua A em $Ax = b$ pela sua decomposição em valores singulares (na forma reduzida $U\hat{\Sigma}\hat{V}^*$).

   (b) Multiplique, depois, ambos os membros do sistema por U^* e $\hat{\Sigma}^{-1}$, por esta ordem, indicando o resultado obtido. Classifique, novamente, o sistema resultante destas operações.

   (c) Mostre que $\hat{V}\hat{\Sigma}^{-1}U^*b$ é uma solução do sistema. Identifique esta solução como sendo $A^{\oplus}b$.

Bibliografia

[1] G. H. Golub e C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1989.

[2] R. A. Horn e C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985 (edição de 1993).

[3] A. P. Santana e J. F. Queiró, *Álgebra Linear e Geometria Analítica*, Departamento de Matemática, Universidade de Coimbra, 2003.

[4] L. N. Trefethen e D. Bau III, *Numerical Linear Algebra*, SIAM, Filadélfia, 1997.

[5] C. F. Van Loan, *Computational Frameworks for the Fast Fourier Transform*, SIAM, Filadélfia, 1992.