Capítulo 4

Valores e vectores próprios

Os valores e vectores próprios desempenham um papel central em diversas áreas da matemática aplicada, da física, da economia, da engenharia etc.. A diagonalização de matrizes, intimamente relacionada com os conceitos de valores e vectores próprios, desempenha um papel crucial no estudo de sistemas de equações diferenciais, nomeadamente no estudo qualitativo da dinâmica do sistema definido por essas equações.

Na Secção 4.1 estudam-se as propriedades que resultam da definição de valor e vector próprio de uma matriz. A Secção 4.2 é dedicada ao problema da diagonalização de matrizes. Mais tarde, no Capítulo 7, estuda-se a diagonalização (ortogonal) de matrizes simétricas e aplica-se essa diagonalização na identificação de cónicas e superfícies.

A Secção 4.4 apresentamos algumas aplicações dos conceitos anteriormente tratados. Nessa secção, destacamos o papel dos valores e vectores próprios no estudo de sistemas dinâmicos (discretos e contínuos). Em particular, é ilustrado o papel dos valores e vectores próprios no algoritmo PageRank usado pelo motor de busca Google, e na previsão a prazo de sistemas descritos por matrizes de Markov, os quais modelam vários problemas de que a dinâmica de populações é um exemplo. A Secção 4.4.2, é dedicada à resolução de sistemas lineares de equações diferenciais ordinárias descritos por matrizes diagonalizáveis. Aproveitamos ainda para introduzir nessa secção a exponencial de matrizes.

4.1 Valores e vectores próprios de matrizes

Os valores próprios de uma matriz quadrada são escalares que satisfazem uma certa equação matricial. Recorde-se que quando nos referimos a escalares estamos a considerar números reais ou complexos. A cada valor próprio correspondem certos vectores que recebem a designação de vectores próprios. Os números complexos não podem ser evitados quando se lida com valores próprios, uma vez que mesmo uma matriz real pode ter valores próprios complexos. É assim essencial que o leitor possua alguns conhecimentos de números complexos (pelo que deve consultar o Anexo A caso necessite).

Eis a definição de valor e vector próprio de uma matriz.

Definição 4.1. Um escalar λ diz-se um *valor próprio* de uma matriz quadrada A se existe um vector não nulo $\mathbf{x}$, tal que

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{4.1}$$

A um vector não nulo $\mathbf{x}$ que verifica a equação (4.1) chama-se *vector próprio* de A associado ao valor próprio λ.

O par $(\lambda, \mathbf{x})$ diz-se um *par próprio* de A se λ é um valor próprio de A e $\mathbf{x}$ é um vector próprio de A associado a λ.

O conjunto dos valores próprios de uma matriz A designa-se por *espectro* de A e denota-se por $\sigma(A)$.

A terminologia inglesa para valor e vector próprio é respectivamente *"eigenvalue"* e *"eigenvector"*, enquanto que em português do Brasil se usam as designações de autovalor e autovector.

Exemplo 4.1. O vector $\mathbf{x} = (2, 1)$ é um vector próprio da matriz $A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ já que

$$A\mathbf{x} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix} = 5 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5\mathbf{x}.$$

Da igualdade anterior, concluimos que $\mathbf{x}$ é um vector próprio de A associado ao valor próprio $\lambda = 5$. Ou seja, $(5, \mathbf{x})$ é um par próprio de A.

♦

A equação $A\mathbf{x} = \lambda\mathbf{x}$ pode reescrever-se na seguinte forma equivalente

$$A\mathbf{x} = \lambda\mathbf{x} \iff A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \iff (A - \lambda I)\mathbf{x} = \mathbf{0}, \tag{4.2}$$

onde I é a matriz identidade. Assim, a definição de vector próprio é equivalente à existência de uma solução não nula do sistema homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Ora, um sistema homogéneo com matriz dos coeficientes quadrada possui soluções não nulas se e só se o determinante da matriz dos coeficientes é nulo (Proposição 2.2, pág. 95). Podemos portanto enunciar a proposição:

Proposição 4.1. O escalar λ é um valor próprio da matriz quadrada A se e só se satisfaz a equação

$$\det(A - \lambda I) = 0. \tag{4.3}$$

Um vector próprio $\mathbf{x}$ associado ao valor próprio λ é uma solução não nula do sistema homogéneo $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

A equação (4.3) e a solução geral do sistema $(A - \lambda I)\mathbf{x} = \mathbf{0}$, recebem designações que a seguir se especificam.

Definição 4.2. Chama-se *equação característica* da matriz A à equação (na variável λ)

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

O espaço

$$E(\lambda) = \{\mathbf{v} : \ (A - \lambda I)\mathbf{v} = \mathbf{0}\},$$

diz-se o *espaço próprio* do valor próprio λ.

Note-se que o espaço próprio $E(\lambda)$ é a solução geral do sistema $(A - \lambda I)\mathbf{x} = \mathbf{0}$, ou seja, o núcleo de $(A - \lambda I)$. Isto é,

$$E(\lambda) = N(A - \lambda I).$$

É importante observar que resulta imediatamente da igualdade $A\mathbf{x} = \lambda\mathbf{x}$ que, se A é uma matriz real e λ é um valor próprio complexo não real, um vector próprio $\mathbf{x}$ associado a λ é necessariamente um vector complexo. Nesse caso é conveniente encarar A como uma matriz complexa e $E(\lambda) = N(A - \lambda I)$ como subespaço de $\mathbb{C}^n$. Se λ é um valor próprio real da matriz real A, existem vectores próprios em $\mathbb{R}^n$ associados a λ, e nesse caso é usual tomar para $E(\lambda)$ apenas o conjunto dos vectores de $\mathbb{R}^n$ que satisfazem $A\mathbf{x} = \lambda\mathbf{x}$. Por isso, quando uma matriz real A só tem valores próprios reais, consideramos os espaços próprios $E(\lambda) = N(A - \lambda I)$ como subespaços de $\mathbb{R}^n$, caso contrário estes espaços são considerados subespaços de $\mathbb{C}^n$.

Exemplo 4.2. Determinar os valores próprios e os espaços próprios da matriz

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

Os valores próprios são as soluções da equação característica $\det(A-\lambda I) = 0$, isto é,

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -1/2 \\ -1/2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - \frac{1}{4} = 0 \Longleftrightarrow 1 - \lambda = \pm\frac{1}{2}$$

$$\Longleftrightarrow \lambda = \frac{3}{2} \text{ ou } \lambda = \frac{1}{2}.$$

Assim, os valores próprios de A são $\lambda_1 = \frac{3}{2}$ e $\lambda_2 = \frac{1}{2}$.

Os espaços próprios de $\lambda_1 = \frac{3}{2}$ e $\lambda_2 = \frac{1}{2}$ são, respectivamente, as soluções gerais dos sistemas homogéneos $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$ para $i = 1, 2$. Assim,

$$\left(A - \frac{3}{2}I\right)\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} -1/2 & -1/2 \\ -1/2 & -1/2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow -a - b = 0.$$

Logo,

$$E\left(\frac{3}{2}\right) = \{(-b, b);\ b \in \mathbb{R}\} = \mathrm{Span}\{(-1, 1)\}.$$

Para $\lambda_2 = \frac{1}{2}$, tem-se

$$\left(A - \frac{1}{2}I\right)\mathbf{x} = \mathbf{0} \Longleftrightarrow \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow a - b = 0.$$

Portanto,

$$E\left(\frac{1}{2}\right) = \{(b, b);\ b \in \mathbb{R}\} = \mathrm{Span}\{(1, 1)\}.$$

♦

A equação característica de uma matriz A de ordem n é uma equação polinomial uma vez que $\det(A - \lambda I)$ é um polinómio de grau n em λ. É fácil verificar que assim é usando a definição de determinante como a soma de produtos elementares de entradas (multiplicados pelo respectivo sinal). De facto, como a matriz $(A - \lambda I)$ difere da matriz A apenas nas entradas da diagonal principal, todos os produtos elementares de entradas de $(A - \lambda I)$ são polinómios em λ de grau menor ou igual a n e o único produto elementar de grau n é $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$. Assim, $\det(A - \lambda I)$ é a soma de um polinómio de grau n com polinómios de grau inferior, ou seja, $\det(A - \lambda I)$ é um polinómio de grau n.

Definição 4.3. Seja λ um escalar, A uma matriz $n \times n$, e I a matriz identidade de ordem n. O polinómio de grau n em λ, definido por $p(\lambda) = \det(A - \lambda I)$, é denominado *polinómio característico* de A.

A seguir sumarizamos algumas equivalências anteriormente referidas.

Proposição 4.2. Se A é uma matriz quadrada e λ um escalar, são equivalentes as afirmações:

a) λ é um valor próprio de A;

b) O sistema $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tem soluções não nulas;

c) O núcleo de $(A - \lambda I)$ não é trivial, isto é, $N(A - \lambda I) \neq \{\mathbf{0}\}$;

d) Existe um vector $\mathbf{x}$ não nulo tal que $A\mathbf{x} = \lambda\mathbf{x}$;

e) λ é uma raiz do polinómio característico $p(\lambda) = \det(A - \lambda I)$.

O Teorema Fundamental da Álgebra[1] afirma que um polinómio (numa variável), de coeficientes complexos, de grau $n \geq 1$ tem n raízes (contando as raízes repetidas de acordo com a sua multiplicidade). Estas raízes podem ser simples ou múltiplas (com diferentes graus de multiplicidade). As raízes complexas de polinómios de coeficientes reais ocorrem em pares de conjugados. De facto, se p é um polinómio de coeficientes reais e $p(\lambda) = 0$, então $0 = \overline{p(\lambda)} = p(\overline{\lambda})$.

Por conseguinte, o polinómio característico de uma matriz A de ordem n,

$$p(\lambda) = \det(A - \lambda I) = b_0 + b_1\lambda + \cdots + b_{n-1}\lambda^{n-1} + (-1)^n\lambda^n, \tag{4.4}$$

tem n raízes $\lambda_1, \ldots, \lambda_n$, podendo por isso ser escrito como um produto de n factores:

$$p(\lambda) = \det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda). \tag{4.5}$$

Note-se que na factorização (4.5) pode haver factores repetidos.

De seguida introduzimos alguma terminologia usada para caracterizar valores próprios.

[1] Existe um grande número de provas do denominado Teorema Fundamental da Álgebra, algumas de natureza topológica, outras de natureza algébrica ou ainda de natureza analítica. As provas analíticas são do âmbito da Análise complexa e usam nomeadamente o Teorema do integral de Cauchy, ou o Teorema de Liouville ou ainda o designado princípio do argumento. No site `http://www.cut-the-knot.org/do_you_know/fundamental2.shtml`, poderá encontrar várias provas deste teorema bem como várias referências.

Definição 4.4. Seja λ um valor próprio da matriz A.

- A *multiplicidade algébrica* de λ é número de vezes que a raiz λ aparece repetida no polinómio característico de A. Isto é, $mult\,alg(\lambda_i) = k_i$ se e só se $p(\lambda) = (\lambda_1 - \lambda)^{k_1} \cdots (\lambda_s - \lambda)^{k_s}$, onde o espectro de A é $\sigma(A) = \{\lambda_1, \ldots, \lambda_s\}$, com $\lambda_i \neq \lambda_j$.
- λ diz-se um valor próprio *simples* quando $mult\,alg(\lambda) = 1$.
- A *multiplicidade geométrica* de λ é a dimensão do núcleo de $(A - \lambda I)$, isto é, $\dim E(\lambda)$. Dito de outra forma: $mult\,geom(\lambda)$ é o número máximo de vectores próprios linearmente independentes associados a λ.
- λ diz-se um valor próprio *semi-simples* quando $mult\,alg(\lambda) = mult\,geom(\lambda)$.

Na proposição seguinte apresentamos um resultado de utilidade prática, em particular quando se pretende decidir sobre a existência de valores próprios sem os calcular explicitamente. Para tal, é necessário definir o que se entende por traço de uma matriz quadrada.

Definição 4.5. Chama-se *traço* de uma matriz quadrada à soma das entradas da sua diagonal principal, e designamos por $\mathrm{tr}(A)$ o traço de A .

Proposição 4.3. Seja $A = [a_{ij}]_{i,j=1,\ldots,n}$ e $\lambda_1, \lambda_2, \ldots, \lambda_n$ os valores próprios de A. São satisfeitas as igualdades:

$$\begin{aligned} \det(A) &= \lambda_1 \lambda_2 \cdots \lambda_n, \\ \text{e} & \\ \mathrm{tr}(A) &= a_{11} + a_{22} + \cdots + a_{nn} = \lambda_1 + \lambda_2 + \cdots + \lambda_n. \end{aligned}$$

Antes de passarmos à demonstração desta proposição, notemos que no caso particular de uma matriz 2×2 a sua demontração é muito simples. O polinómio característico da matriz $A = [a_{ij}]_{i,j=1,2}$ é

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 - \underbrace{(a_{11} + a_{22})}_{\text{traço de } A}\lambda + \underbrace{a_{11}a_{22} - a_{12}a_{21}}_{\det(A)}. \end{aligned}$$

Por outro lado, se λ_1 e λ_2 são as raízes de $p(\lambda)$, então podemos escrever o polinómio na forma

$$p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2.$$

Comparando as duas expressões obtidas para $p(\lambda)$ segue o resultado enunciado na proposição.

Demonstração. O termo independente do polinómio característico de A, $p(\lambda) = \det(A - \lambda I)$, é $p(0) = \det(A)$. Por outro lado, usando a factorização de p em termos das suas raízes (expressão (4.5)) temos $p(0) = \lambda_1\lambda_2 \cdots \lambda_n$, ficando assim mostrado que o produto dos valores próprios é igual ao determinante da matriz.

Para provar a relação entre o traço da matriz e os seus valores próprios, note-se que usando a factorização (4.5), o coeficiente do termo λ^{n-1} de p é

$$(-1)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n).$$

Se mostrarmos que este coeficiente é igual a $(-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$, provamos que o traço de A é igual à soma dos valores próprios. Para tal, vamos usar indução sobre n. Quando $n = 1$, o resultado é trivialmente satisfeito.

Suponha-se (hipótese de indução) que para qualquer matriz $A_{n-1} = [a_{ij}]$ de ordem $(n-1)$ o coeficiente do termo em λ^{n-2} do seu polinómio característico é $(-1)^{n-2}(a_{11} + a_{22} + \cdots + a_{n-1,n-1})$. Isto é,

$$\det(A_{n-1} - \lambda I) = (-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}(a_{11} + a_{22} + \cdots + a_{n-1,n-1})\lambda^{n-2} + \text{ t.o.i.}, \tag{4.6}$$

onde t.o.i. designa termos de ordem inferior, ou seja, termos que envolvem potências λ^k com $k < n-2$.

Recorrendo ao desenvolvimento de Laplace (ver página 98) segundo a última linha da matriz A (de ordem n), tem-se

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} \\
&= (a_{nn} - \lambda)\det(A_{n-1} - \lambda I) + a_{n1}C_{n1} + \cdots + a_{n,n-1}C_{n,n-1} \\
&= (a_{nn} - \lambda)\det(A_{n-1} - \lambda I) + a_{n1}q_1(\lambda) + \cdots + a_{n,n-1}q_{n-1}(\lambda),
\end{aligned}$$

onde os q_j's designam polinómios em λ de grau menor ou igual a $(n-2)$.

Usando a hipótese de indução, nomeadamente a expressão (4.6), o primeiro termo da soma anterior satisfaz a igualdade

$$(a_{nn} - \lambda)\det(A_{n-1} - \lambda I) = (-1)^n\lambda^n + \\ + (-1)^{n-1}\lambda^{n-1}(a_{11} + a_{22} + \cdots + a_{n-1,n-1} + a_{nn}) + \text{ t.o.i.}.$$

Finalmente, substituindo a expressão anterior na expressão de $\det(A-\lambda I)$, tem-se

$$\det(A-\lambda I) = (-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}(a_{11}+a_{22}+\cdots+a_{n-1,n-1}+a_{nn}) + \text{ t.o.i.},$$

e portanto o enunciado é válido para qualquer n. □

Exemplo 4.3. Consideremos a matriz

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

Como A é triangular superior, a matriz $(A - \lambda I)$ também o é, e portanto o seu determinante é igual ao produto das entradas da sua diagonal principal. Ou seja, o polinómio característico de A é $p(\lambda) = (3-\lambda)^2(2-\lambda)$. Assim, a matriz A tem:

- um valor próprio igual a 3 de multiplicidade algébrica dois.
- um valor próprio simples que é 2.

Refira-se que o polinómio característico é do terceiro grau e as suas três raízes são contadas considerando a raiz 3 duas vezes e a raiz 2 uma vez.

Confirmando os resultados da proposição anterior, tem-se

$$\det(A) = 3 \times 3 \times 2 = 18 \quad \text{e} \quad \operatorname{tr}(A) = 3 + 3 + 2 = 8,$$

onde a raiz repetida do polinómio característico é considerada de acordo com a sua multiplicidade. ♦

Nota 22. *No que se segue abreviamos por vezes o enunciado da proposição anterior dizendo que o determinante (resp. o traço) de uma matriz é igual ao produto (resp. a soma) dos valores próprios da matriz, subentendendo que os valores próprios múltiplos são considerados de acordo com as suas multiplicidades.*

É consequência imediata da proposição anterior o resultado que a seguir se enuncia.

Corolário 4.1. Uma matriz é invertível se e só se zero não é um valor próprio da matriz.

Exercício 4.1. Mostre que se $(\lambda, \mathbf{x})$ é um par próprio de uma matriz invertível A, então $\left(\frac{1}{\lambda}, \mathbf{x}\right)$ é um par próprio de A^{-1}. ▲

As matrizes reais podem ter valores próprios complexos (ver Exemplo 4.4). Sendo o polinómio característico de uma matriz real um polinónio real, as suas raízes complexas ocorrem em pares de conjugados. Isto significa que se $(a+ib)$ é um valor próprio complexo (não real) de uma matriz real, então o seu conjugado $(a-ib)$ também é valor próprio dessa matriz.

Antes de estabelecermos a relação existente entre vectores próprios correspondentes a valores próprios complexos conjugados definimos a matriz conjugada de uma matriz.

Definição 4.6. A matriz conjugada da matriz C é a matriz $\overline{C}$ cujas entradas são os conjugados das entradas de C.

O conjugado de um vector $\mathbf{u}$ é o vector $\overline{\mathbf{u}}$ cujas componentes são os conjugados das componentes de $\mathbf{u}$.

Refira-se que o conjugado de um número real coincide consigo próprio, e portanto se C é uma matriz real tem-se $C = \overline{C}$.

Proposição 4.4. Se $\lambda \in \mathbb{C}$ é um valor próprio de uma matriz real A, então $\overline{\lambda}$ também é um valor próprio de A. Além disso, se $\mathbf{u}$ é um vector próprio de A associado a $\lambda \in \mathbb{C}$, então $\overline{\mathbf{u}}$ é um vector próprio de A associado a $\overline{\lambda}$.

Demonstração. Seja $(\lambda, \mathbf{u})$ um par próprio de A, isto é, $A\mathbf{u} = \lambda\mathbf{u}$. Tomando o conjugado da igualdade $A\mathbf{u} = \lambda\mathbf{u}$, tem-se

$$\overline{A\mathbf{u}} = \overline{(\lambda\mathbf{u})} \Longleftrightarrow \overline{A}\,\overline{\mathbf{u}} = \overline{\lambda}\,\overline{\mathbf{u}} \Longleftrightarrow A\overline{\mathbf{u}} = \overline{\lambda}\,\overline{\mathbf{u}},$$

onde aplicámos a igualdade $\overline{A} = A$ uma vez que, por hipótese, A é real.

Por definição de valor e vector próprio de A, a igualdade $A\overline{\mathbf{u}} = \overline{\lambda}\,\overline{\mathbf{u}}$ significa que $\overline{\lambda}$ é um valor próprio de A e $\overline{\mathbf{u}}$ é um vector próprio associado. □

Exemplo 4.4. A matriz $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ possui os seguintes valores próprios:

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0 \Longleftrightarrow \lambda = i \text{ ou } \lambda = -i.$$

O espaço próprio $E(i)$ é o núcleo da matriz $(A - iI)$, ou seja o conjunto dos vectores $\mathbf{x}$ que verificam

$$(A - iI)\mathbf{x} = 0 \Longleftrightarrow \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow -ia = b.$$

Logo, $E(i) = \{(a, -ia) : a \in \mathbb{C}\} = \text{Span}\{(1, -i)\}$. Como a valores próprios conjugados correspondem vectores próprios conjugados, tem-se

$$E(-i) = \left\{(\bar{a}, \overline{(-ia)}) : a \in \mathbb{C}\right\} = \{\bar{a}\,(1, i) : a \in \mathbb{C}\} = \text{Span}\{(1, i)\}.$$

♦

4.1.1 Valores próprios e comportamento de $f(\mathbf{x}) = A\mathbf{x}$

Nesta secção analisamos algumas relações entre os valores e vectores próprios de uma matriz real A de ordem n e o comportamento da função $f : \mathbb{R}^n \to \mathbb{R}^n$ que aplica um vector $\mathbf{x} \in \mathbb{R}^n$ no vector $A\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \mapsto A\mathbf{x}$.

A função f, definida por $f(\mathbf{x}) = A\mathbf{x}$, é uma função linear, ou seja, uma função que verifica $f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$ para quaisquer $\alpha, \beta \in \mathbb{R}$ e $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Como veremos no Capítulo 6, qualquer função linear f de $\mathbb{R}^n$ em $\mathbb{R}^n$ pode ser escrita na forma $f(\mathbf{x}) = A\mathbf{x}$, com A uma matriz $n \times n$.

Estudaremos aqui com algum detalhe dois casos: 1) A matriz A só tem valores próprios reais; 2) A matriz A é 2×2 e tem um par de valores próprios complexos conjugados. Como veremos ainda neste capítulo, estes dois casos são aqueles que importa estudar se se pretende entender o caso geral de funções f definidas por matrizes A diagonalizáveis.

Caso 1: A matriz A só tem valores próprios reais.

Seja λ um valor próprio real da matriz real A e $f(\mathbf{x}) = A\mathbf{x}$. O espaço próprio $E(\lambda)$ é o espaço gerado pelos vectores próprios associados ao valor próprio λ da matriz A, logo

$$f(\mathbf{x}) = A\mathbf{x} = \lambda\mathbf{x}, \quad \text{para todo} \quad \mathbf{x} \in E(\lambda).$$

A igualdade anterior diz-nos que, se $\mathbf{x} \in E(\lambda)$ então o vector $f(\mathbf{x})$ é um múltiplo de $\mathbf{x}$, e portanto $f(\mathbf{x})$ também pertence a $E(\lambda)$.

Ou seja, se $\lambda \in \mathbb{R}$ qualquer vector do espaço próprio $E(\lambda)$ é aplicado por f num vector do espaço próprio. Na Figura 4.1 ilustramos este facto.

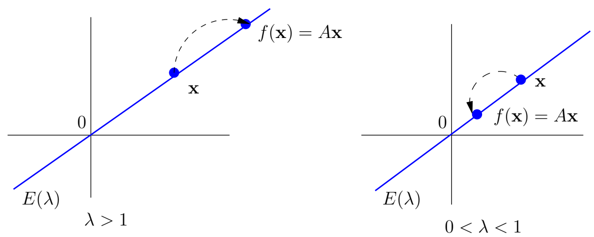

Figura 4.1: Os subespaços próprios de valores próprios reais são invariantes por f.

Um subconjunto S do domínio de uma função g diz-se *invariante* por g se qualquer vector de S é aplicado por g num vector de S .

Se A é uma matriz real de ordem n, os subespaços próprios correspondentes a valores próprios reais são subespaços de $\mathbb{R}^n$ invariantes por $f(\mathbf{x}) = A\mathbf{x}$.

Exemplo 4.5. A matriz A do Exemplo 4.2 (pág. 176) tem espectro $\sigma(A) = \{\frac{3}{2}, \frac{1}{2}\}$. A função $f(\mathbf{x}) = A\mathbf{x}$ é

$$f(\mathbf{x}) = f(x_1, x_2) = \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - \frac{x_2}{2} \\ \\ \frac{-x_1}{2} + x_2 \end{bmatrix}.$$

Ou seja, $f(x_1, x_2) = (x_1 - \frac{x_2}{2}, \frac{-x_1}{2} + x_2)$.

Os espaços próprios de A, obtidos no referido exemplo, são

$$E\left(\frac{3}{2}\right) = \mathrm{Span}\{(-1,1)\} \quad \text{e} \quad E\left(\frac{1}{2}\right) = \mathrm{Span}\{(1,1)\}.$$

Geometricamente, estes espaços próprios são rectas que passam pela origem e têm as direcções dos vectores $(-1,1)$ e $(1,1)$. Assim, a função f: "contrai" vectores na direcção definida por $E\left(\frac{1}{2}\right)$ já que, para $\mathbf{x} \in E\left(\frac{1}{2}\right)$ a respectiva imagem por f é $f(\mathbf{x}) = \frac{\mathbf{x}}{2}$; e "expande" vectores na direcção definida por $E\left(\frac{3}{2}\right)$, visto que $f(\mathbf{x}) = \frac{3\mathbf{x}}{2}$ para $\mathbf{x} \in E\left(\frac{3}{2}\right)$. Na Figura 4.2 ilustramos este facto.

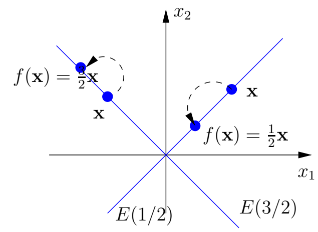

Figura 4.2: A função f "expande" vectores na direcção $E\left(\frac{3}{2}\right)$ e "contrai" na direcção $E\left(\frac{1}{2}\right)$.

Caso 2: A é uma matriz real, 2×2, e tem um par de valores próprios complexos conjugados.

Do caso anterior, sabemos que os subespaços próprios de valores próprios reais, de uma matriz real A, são subespaços invariantes para a função $f : \mathbb{R}^n \to \mathbb{R}^n$, definida por $f(\mathbf{x}) = A\mathbf{x}$. Quando os valores próprios de A são complexos, os respectivos espaços próprios não são subespaços de $\mathbb{R}^n$, apesar da função f estar definida de $\mathbb{R}^n$ em $\mathbb{R}^n$.

Consideremos a matriz A do Exemplo 4.4 (pág. 182) e a função $f(\mathbf{x}) = A\mathbf{x}$. Essa matriz tem valores próprios complexos $\pm i$, e a função $f(\mathbf{x}) = A\mathbf{x}$ é definida por

$$f(\mathbf{x}) = A\mathbf{x} \Longleftrightarrow f(x_1, x_2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$

Geometricamente, f actua no vector $\mathbf{x}$ rodando-o em torno da origem de um ângulo $\dfrac{\pi}{2}$ no sentido directo (ou anti-horário). A Figura 4.3 ilustra este facto. Como se observa neste exemplo, se aplicarmos f, sucessivamente, a um vector $\mathbf{x}$, ao fim de 4 aplicações voltamos a obter o vector $\mathbf{x}$. Na Figura 4.3 denotamos por

$$f^k(\mathbf{x}) = (\underbrace{f \circ f \circ \cdots \circ f}_{\text{k vezes}})(\mathbf{x}),$$

a composição de f consigo própria k vezes (isto é, a transformação obtida por k aplicações sucessivas de f). Note-se que sendo $f(\mathbf{x}) = A\mathbf{x}$ se tem $f^k(\mathbf{x}) = A^k\mathbf{x}$.

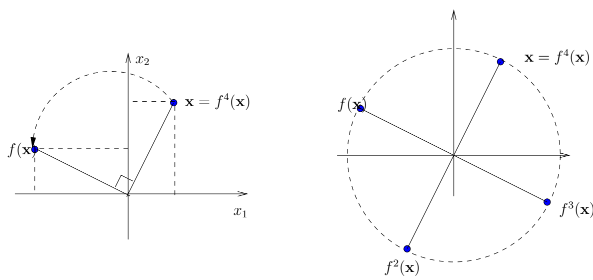

Figura 4.3: A função $f(\mathbf{x}) = A\mathbf{x}$, em que A tem valores próprios $\pm i$ representa uma rotação em $\mathbb{R}^2$.

É óbvio que, exceptuando a origem, nenhum vector de $\mathbb{R}^2$ é aplicado num múltiplo de si próprio. Ou seja, os únicos subespaços de $\mathbb{R}^2$ invariantes por f são $\{(0,0)\}$ e $\mathbb{R}^2$.

Consideremos agora uma matriz A do tipo 2×2 com valores próprios complexos $\lambda = a \pm ib$, com $b \neq 0$,

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Sugere-se como exercício que verifique que esta matriz tem valores próprios $a \pm ib$.

Recorde ainda (Anexo A) que há uma correspondência biunívoca entre pontos do plano de coordenadas (a, b) e números complexos $a+ib$. Usando coordenadas polares, $a = \rho\cos\theta$ e $b = \rho\operatorname{sen}\theta$, o número complexo $\lambda = a+ib$ escreve-se na forma polar: $\lambda = \rho(\cos\theta + i\operatorname{sen}\theta)$. O valor $\rho = |\lambda| = \sqrt{a^2+b^2}$ é a distância de λ à origem, e θ é o ângulo entre a parte positiva do eixo real e o ponto (a, b) (com $-\pi < \theta \leq \pi$). A Figura A.2 da página 512, ilustra a representação polar de um número complexo $a + ib$.

Assim, a matriz A pode escrever-se como um produto de duas matrizes

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} = DR.$$

A matriz $R = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ é uma *matriz de rotação*, visto que para $\mathbf{x} \in \mathbb{R}^2$ o vector $R\mathbf{x}$ é o vector de $\mathbb{R}^2$ que se obtém rodando $\mathbf{x}$ de um ângulo θ, no sentido directo, em torno da origem (como facilmente se verifica).

A matriz $D = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}$ representa uma *expansão* se $\rho = |\lambda| > 1$, e uma *contracção* se $\rho = |\lambda| < 1$ já que, $D\mathbf{x} = \rho\mathbf{x}$.

Assim, a função $f(\mathbf{x}) = A\mathbf{x} = DR\,\mathbf{x}$ actua sobre um vector $\mathbf{x} \in \mathbb{R}^2$ rodando este vector e depois expandindo, contraindo ou mantendo-o, respectivamente nos casos em que $|\lambda| > 1, |\lambda| < 1$ e $|\lambda| = 1$. A Figura 4.4 ilustra esse comportamento de f sobre vectores de $\mathbb{R}^2$.

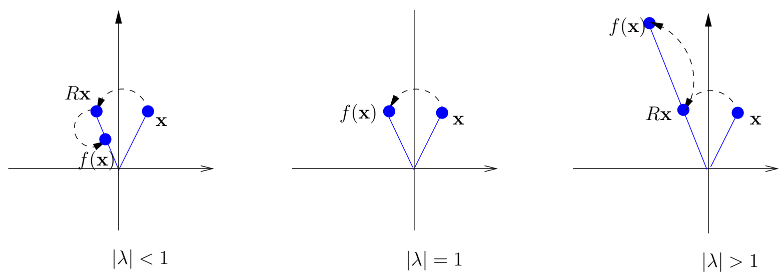

Figura 4.4: A matriz A tem λ como valor próprio complexo e $f(\mathbf{x}) = A\mathbf{x} = DR\mathbf{x}$.

Como veremos na secção seguinte, o comportamento geral de uma função $f : \mathbb{R}^n \to \mathbb{R}^n$ definida por $f(\mathbf{x}) = A\mathbf{x}$, em que A é uma matriz (real) diagonalizável é bem ilustrado pelos dois casos apresentados.

Consideremos agora um exemplo de uma matriz A possuindo valores próprios reais e complexos.

Exemplo 4.6. Consideremos $f(\mathbf{x}) = A\mathbf{x}$ com $A = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1.2 \end{bmatrix}$.

A matriz A tem valores próprios $\lambda_1 = \frac{\sqrt{3}}{2} + i\frac{1}{2}$, $\lambda_2 = \frac{\sqrt{3}}{2} - i\frac{1}{2}$ e $\lambda_3 = 1.2$.

O valor próprio λ_3 é real e o seu espaço próprio é gerado pelo vector $(0, 0, 1)$ (isto é, $E(1.2)$ é o eixo dos zz). Logo, vectores do eixo dos zz são aplicados por f em vectores do eixo dos zz por uma expansão de factor 1.2.

Atendendo à forma da matriz A (diagonal por blocos) é fácil verificar que f aplica vectores do plano xy em vectores deste plano. Como o valor próprio λ_1 (e portanto o seu conjugado λ_2) tem módulo $|\lambda_1| = |\lambda_2| = \sqrt{\frac{3}{4} + \frac{1}{4}} = 1$, a função f aplica um vector $\mathbf{u}$ do plano xy num vector que se obtém de $\mathbf{u}$ por rotação (em torno do eixo dos zz). O ângulo desta rotação é $\pi/6$ (note que $\operatorname{sen}\pi/6 = 1/2$). Por exemplo $f(1,1,0) = (\frac{\sqrt{3}-1}{2}, \frac{\sqrt{3}+1}{2}, 0)$. Para qualquer outro vector $\mathbf{v} = (s_1, s_2, s_3)$, a imagem por f deste vector é o vector $f(\mathbf{v})$ que tem terceira coordenada $1.2s_3$ (expansão na direcção do eixo dos zz), e duas primeiras coordenadas respectivamente, $\cos(\pi/6)s_1 - \operatorname{sen}(\pi/6)s_2$ e $\operatorname{sen}(\pi/6)s_1 + \cos(\pi/6)s_2$ (rotação de (s_1, s_2) de $\pi/6$ em torno da origem).

A Figura 4.5 ilustra aplicações sucessivas de f ao vector $\mathbf{u}$ do plano xy, ao vector $\mathbf{w}$ do eixo dos zz, e ao vector $\mathbf{v}$ que não pertence a estes subespaços. Esta figura ilustra ainda o facto das imagens de aplicações sucessivas de f a vectores $\mathbf{v}$ não pertencentes ao eixo dos zz nem ao plano xy, estarem sobre hélices inscritas num cilindro.

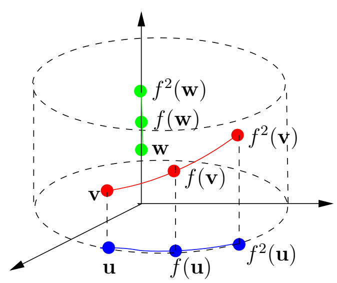

Figura 4.5: Aplicações sucessivas de $f(\mathbf{x}) = A\mathbf{x}$ para a matriz A do Exemplo 4.6.

♦

4.2 Diagonalização de matrizes

O problema central tratado nesta secção é o de saber se dada uma matriz de ordem n existe uma base de $\mathbb{C}^n$ formada por vectores próprios. Quando tal acontece a matriz diz-se diagonalizável, ou ainda, que a matriz é semelhante a uma matriz diagonal. O processo de diagonalização de matrizes desempenha um papel importante em álgebra linear sendo inúmeras as suas aplicações. Por exemplo, a diagonalização de matrizes é utilizada na interpretação da dinâmica de modelos físicos, em computação gráfica, e na identificação de cónicas e de superfícies quadráticas.

Definição 4.7. Duas matrizes quadradas A e B dizem-se *semelhantes* se existe uma matriz invertível P tal que

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

Definição 4.8. Uma matriz quadrada A diz-se *diagonalizável* se é semelhante a uma matriz diagonal D. Isto é, se existe uma matriz invertível P tal que

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

A uma matriz P tal que, $A = PDP^{-1}$ com D diagonal, chama-se *matriz que diagonaliza* A, ou matriz *diagonalizante* de A.

Comecemos por mostrar que os valores próprios de matrizes semelhantes são iguais.

Proposição 4.5. Matrizes semelhantes têm o mesmo polinómio característico. Em particular, os valores próprios são os mesmos e ocorrem com as mesmas multiplicidades.

Demonstração. Sejam A e B matrizes semelhantes. Isto é, existe uma matriz invertível P tal que $A = PBP^{-1}$. Tem-se

$$\begin{aligned}\det(A - \lambda I) &= \det(PBP^{-1} - \lambda I) = \det\left(P(B - \lambda I)P^{-1}\right)\\ &= \det P \det(B - \lambda I)\det(P^{-1}) = \det(B - \lambda I).\end{aligned}$$

Nas igualdades anteriores aplicámos os seguintes factos: o determinante do produto é igual ao produto dos determinantes; o determinante da inversa é o inverso do determinante. Da última igualdade segue que o polinómio característico de A é igual ao de B e portanto A e B possuem os mesmos valores próprios com as mesmas multiplicidades. $\square$

Da Proposição 4.3 sabemos que o traço e o determinante de uma matriz de ordem n são respectivamente iguais à soma e ao produto dos n valores próprios da matriz, por conseguinte, da proposição anterior, segue o corolário que passamos a enunciar.

Corolário 4.2. Matrizes semelhantes têm o mesmo traço e o mesmo determinante.

Se A é uma matriz diagonalizável, isto é, $A = PDP^{-1}$ com D diagonal, pela Proposição 4.5 os valores próprios de D são os valores próprios de A. Como D é diagonal, os seus valores próprios são as entradas da diagonal principal. Logo, $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, onde $\lambda_1, \lambda_2, \ldots, \lambda_n$ são os valores próprios de A.

No caso de A ser diagonalizável, a questão que agora se coloca é a de saber construir uma matriz P que diagonaliza A. O teorema seguinte mostra como construir uma tal matriz, oferecendo simultaneamente uma condição necessária e suficiente para que uma matriz seja diagonalizável.

Teorema 4.1. Uma matriz A do tipo $n \times n$ é diagonalizável se e só se possui n vectores próprios linearmente independentes. Ou seja, se e só se existe uma base de $\mathbb{C}^n$ constituída por vectores próprios de A.

Além disso, se $A = PDP^{-1}$ com $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, então para todo $i = 1, \ldots, n$, a coluna i de P é um vector próprio de A associado ao valor próprio λ_i.

Demonstração. A igualdade $A = PDP^{-1}$ é equivalente a $AP = PD$. O produto AP é a matriz cujas colunas são o produto de A pelas colunas de P (ver Definição 1.12, pág. 39). Assim, designando as colunas de P por $\mathbf{c}_i$, tem-se

$$AP = A\begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A\mathbf{c}_1 & A\mathbf{c}_2 & \cdots & A\mathbf{c}_n \\ | & | & & | \end{bmatrix}.$$

Por outro lado,

$$PD = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{c}_1 & \lambda_2\mathbf{c}_2 & \cdots & \lambda_n\mathbf{c}_n \\ | & | & & | \end{bmatrix}.$$

Logo, $AP = PD$ se e só se a i-ésima coluna de P verifica $A\mathbf{c}_i = \lambda_i\mathbf{c_i}$, para todo $i = 1, \ldots, n$. Ou seja, se e só se $\mathbf{c}_i$ é um vector próprio de A associado ao valor próprio λ_i.

A matriz P é invertível se e só se tem n colunas linearmente independentes (cf. Proposição 3.15, pág. 153). Conclui-se portanto que é necessário e suficiente para que A seja diagonalizável que existam n vectores próprios de A linearmente independentes. $\square$

Note-se que nem todas as matrizes são diagonalizáveis como se verifica no exemplo seguinte.

Exemplo 4.7. Considere a matriz $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Verifiquemos se esta matriz é ou não diagonalizável.

Uma vez que a matriz A é triangular, os valores próprios de A são $\lambda_1 = 2$ e $\lambda_2 = 1$ (as entradas da diagonal principal). O valor próprio 2 é simples e o valor próprio 1 tem multiplicidade algébrica 2, visto que o polinómio característico de A é $p(\lambda) = (2-\lambda)(1-\lambda)^2$.

Para que A seja diagonalizável têm de existir 3 vectores próprios linearmente independentes. Determinemos os espaços próprios.

$$(A - 2I)\mathbf{x} = \mathbf{0} \iff \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} -b + c = 0 \\ c = 0 \end{cases}$$

Logo,

$$E(2) = \left\{(a,b,c) \in \mathbb{R}^3 : \; b = c = 0\right\} = \{(a,0,0) : \; a \in \mathbb{R}\} = \text{Span}\{(1,0,0)\}.$$

Para $\lambda_2 = 1$:

$$(A - I)\mathbf{x} = \mathbf{0} \iff \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} a = 0 \\ c = 0 \end{cases}$$

$$E(1) = \left\{(a,b,c) \in \mathbb{R}^3 : \; a = c = 0\right\} = \{(0,b,0) : \; b \in \mathbb{R}\} = \text{Span}\{(0,1,0)\}.$$

Como a dimensão de cada espaço próprio é igual a 1, existem no máximo dois vectores próprios linearmente independentes (um vector retirado de cada espaço próprio). Ou seja, não existe um número suficiente de vectores próprios (que seria 3) para a matriz ser diagonalizável. Portanto, a matriz A não é diagonalizável. ♦

Pelo teorema anterior sabemos que é condição necessária e suficiente para uma matriz A, de ordem n, ser diagonalizável que possua n vectores próprios linearmente independentes, ou seja, que exista uma base do espaço linear complexo $\mathbb{C}^n$ formada por vectores próprios de A. Como veremos adiante (Proposição 4.6

e Corolário 4.4), esta condição é equivalente à soma das dimensões dos espaços próprios ser igual a n.

Na proposição seguinte mostramos que vectores próprios associados a valores próprios distintos são necessariamente linearmente independentes.

Proposição 4.6. Seja $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ um conjunto de valores próprios distintos da matriz A de ordem n.

1) Se $\{(\lambda_1, \mathbf{x}_1), (\lambda_2, \mathbf{x}_2), \ldots, (\lambda_k, \mathbf{x}_k)\}$ é um conjunto de pares próprios de A, então o conjunto $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ é linearmente independente.

2) Se B_i é uma base de $E(\lambda_i)$, então o conjunto $B = B_1 \cup B_2 \cup \cdots \cup B_k$ é linearmente independente.

Demonstração. 1): Suponhamos, por absurdo, que S é linearmente dependente, e que $S' = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$ é um subconjunto de S com o maior número possível de vectores linearmente independentes (isto é, S' é uma base de $\operatorname{Span} S$ cuja existência é assegurada pela Proposição 3.8, pág. 133).

Qualquer vector $\mathbf{x}_j$ de $S \setminus S'$ escreve-se de forma única como combinação linear dos vectores de S' (cf. Teorema 3.1). Ou seja, existem escalares α_i únicos tais que

$$\mathbf{x}_j = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_p \mathbf{x}_p. \tag{4.7}$$

Multiplicando (à esquerda) a igualdade anterior pela matriz $(A - \lambda_j I)$, obtemos

$$\begin{aligned} (A - \lambda_j I)\mathbf{x}_j = \alpha_1 (A - \lambda_j I)\mathbf{x}_1 + \cdots + \alpha_p (A - \lambda_j I)\mathbf{x}_p \quad &\iff \\ \mathbf{0} = \alpha_1 (\lambda_1 - \lambda_j)\mathbf{x}_1 + \cdots + \alpha_p (\lambda_p - \lambda_j)\mathbf{x}_p, & \end{aligned}$$

já que $\mathbf{x}_i$ é um vector próprio de A associado a λ_i. Como S' é linearmente independente, segue da última igualdade que $\alpha_i(\lambda_i - \lambda_j) = 0$, para todo o $i = 1, \ldots, p$. Sendo distintos os valores próprios de A, obtemos de $\alpha_i(\lambda_i - \lambda_j) = 0$ que $\alpha_i = 0$ para todo $i = 1, \ldots, p$. Por conseguinte, de (4.7) resulta que $\mathbf{x}_j$ é o vector zero, o que contraria a hipótese de $\mathbf{x}_j$ ser vector próprio. Logo, S é linearmente independente.

2): Para mostrar que $B = B_1 \cup B_2 \cup \cdots \cup B_k$ é linearmente independente, basta provar que B é uma base de $E(\lambda_1) + \cdots + E(\lambda_k)$ (ver demostração da Proposição 3.10, pág. 141). Ou equivalentemente, mostrar que para $j = 1, \ldots, k$ se tem

$$X_j = E(\lambda_j) \cap \left(E(\lambda_1) + \cdots + E(\lambda_{j-1}) + E(\lambda_{j+1}) + \cdots + E(\lambda_k)\right) = \{\mathbf{0}\}\,.$$

Suponhamos, por absurdo, que existe um vector $\mathbf{x} \in X_j$ tal que $\mathbf{x} \neq \mathbf{0}$. Como o vector $\mathbf{x}$ pertence a $E(\lambda_j)$, tem-se

$$A\mathbf{x} = \lambda_j \mathbf{x}. \tag{4.8}$$

Uma vez que $\mathbf{x}$ também pertence a $E(\lambda_1)+\cdots+E(\lambda_{j-1})+E(\lambda_{j+1})+\cdots+E(\lambda_k)$, podemos escrever

$$\mathbf{x} = \mathbf{v}_1 + \cdots + \mathbf{v}_{j-1} + \mathbf{v}_{j+1} + \cdots + \mathbf{v}_k, \tag{4.9}$$

para certos vectores $\mathbf{v}_i \in E(\lambda_i)$. Multiplicando (à esquerda) a igualdade (4.9), respectivamente, por A e por λ_j, obtemos

$$\begin{aligned} A\mathbf{x} &= A\mathbf{v}_1 + \cdots + A\mathbf{v}_{j-1} + A\mathbf{v}_{j+1} + \cdots + A\mathbf{v}_k \\ &= \lambda_1\mathbf{v}_1 + \cdots + \lambda_{j-1}\mathbf{v}_{j-1} + \lambda_{j+1}\mathbf{v}_{j+1} + \cdots + \lambda_k\mathbf{v}_k \\ \lambda_j\mathbf{x} &= \lambda_j\mathbf{v}_1 + \cdots + \lambda_j\mathbf{v}_{j-1} + \lambda_j\mathbf{v}_{j+1} + \cdots + \lambda_j\mathbf{v}_k. \end{aligned}$$

De (4.8) segue que as últimas expressões são iguais, pelo que

$$(\lambda_1 - \lambda_j)\mathbf{v}_1 + + \cdots + (\lambda_{j-1} - \lambda_j)\mathbf{v}_{j-1} + (\lambda_{j+1} - \lambda_j)\mathbf{v}_j + \cdots + (\lambda_k - \lambda_j)\mathbf{v}_k = \mathbf{0}.$$

Como pelo item 1) os vectores $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_k$ são linearmente independentes, da última igualdade resulta que $(\lambda_i - \lambda_j) = 0$ para todo $i = 1, \ldots, k$ e $i \neq j$, o que contraria a hipótese dos valores próprios serem distintos. Logo, $X_j = \{\mathbf{0}\}$. Como j é qualquer, temos $X_j = \{\mathbf{0}\}$ para todo $j = 1, \ldots, k$, o que prova que B é uma base de $E(\lambda_1) + \cdots + E(\lambda_k)$. $\square$

Como corolário da proposição anterior e do Teorema 4.1 podemos enunciar o seguinte resultado.

Corolário 4.3. Uma matriz $n \times n$ com n valores próprios distintos é diagonalizável.

Exemplo 4.8. Verifiquemos que $A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$ é diagonalizável, e determinemos uma matriz P que diagonaliza A.

O polinómio característico de A é

$$\det(A-\lambda I) = \begin{vmatrix} 4-\lambda & 0 & 1 \\ 2 & 3-\lambda & 2 \\ 1 & 0 & 4-\lambda \end{vmatrix} = (3-\lambda)\begin{vmatrix} 4-\lambda & 1 \\ 1 & 4-\lambda \end{vmatrix}$$

$$= (3-\lambda)\left(\lambda^2 - 8\lambda + 15\right) = (3-\lambda)^2(5-\lambda).$$

Assim, a matriz A tem valores próprios $\lambda_1 = 3$ e $\lambda_2 = 5$, com multiplicidades algébricas 2 e 1, respectivamente. Saliente-se ainda que no cálculo de $\det(A - \lambda I)$ aplicámos o desenvolvimento de Laplace utilizando a segunda coluna, o que produziu imediatamente uma factorização do polinómio (e portanto uma raiz).

Calculemos bases para os espaços próprios de A.

$$(A-3I)\mathbf{x} = \mathbf{0} \iff \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff a + c = 0$$

$$\begin{aligned} E(3) = \left\{(a,b,c) \in \mathbb{R}^3 :\ a = -c\right\} &= \{(-c,b,c) :\ b,c \in \mathbb{R}\} \\ &= \operatorname{Span}\{(-1,0,1),(0,1,0)\}. \end{aligned}$$

$$\begin{aligned} (A-5I)\mathbf{x} = \mathbf{0} \iff \begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} &\iff \begin{cases} -a \phantom{{}-2b} + c = 0 \\ 2a - 2b + 2c = 0 \end{cases} \\ &\iff a = c \text{ e } b = 2a. \end{aligned}$$

$$\begin{aligned} E(5) = \left\{(a,b,c) \in \mathbb{R}^3 :\ a = c \text{ e } b = 2a\right\} &= \{(a,2a,a) :\ a \in \mathbb{R}\} \\ &= \operatorname{Span}\{(1,2,1)\}. \end{aligned}$$

Como $\{(-1,0,1),(0,1,0)\}$ e $\{(1,2,1)\}$ são bases, respectivamente, dos subespaços $E(3)$ e $E(5)$, tem-se

$$\dim E(3) = 2 \quad \text{e} \quad \dim E(5) = 1.$$

Conclui-se portanto que a multiplicidade geométrica de $\lambda = 3$ é dois, e a de $\lambda = 5$ é um. Por conseguinte, a matriz A só tem valores próprios semi-simples.

Existem três vectores próprios de A linearmente independentes, e portanto A é diagonalizável.

Uma matriz P que diagonaliza A (isto é, tal que $A = PDP^{-1}$ com D diagonal), é uma matriz cujas colunas formam uma base constituída por vectores próprios de A. É claro que P depende da forma como se constrói a matriz D. Assim, se escolhermos $D = \operatorname{diag}(3, 5, 3)$, a matriz P possui, na 1ª e 3ª colunas vectores próprios associados ao valor próprio $\lambda_1 = 3$, e na segunda coluna um vector próprio associado a $\lambda_2 = 5$. Para que P seja invertível, temos de escolher vectores próprios linearmente independentes. Por exemplo,

$$P = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Também podemos considerar, por exemplo,

$$P = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$

correspondendo a uma outra colocação dos valores próprios de A na diagonal principal de D. ♦

Exemplo 4.9. Seja $A = \begin{bmatrix} 2 & 5 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

Usando o desenvolvimento de Laplace ao longo da terceira coluna de $(A-\lambda I)$, temos

$$\det(A-\lambda I) = \begin{vmatrix} 2-\lambda & 5 & 0 \\ -1 & -\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} = (5-\lambda)\left[-\lambda(2-\lambda)+5\right] = (5-\lambda)(\lambda^2-2\lambda+5).$$

Logo, 5 e $1 \pm 2i$ são valores próprios de A, já que

$$(\lambda^2 - 2\lambda + 5) = 0 \Longleftrightarrow \lambda = \frac{2 \pm \sqrt{-16}}{2} = 1 \pm 2i.$$

A matriz A é diagonalizável uma vez que tem três valores próprios distintos (cf. Corolário 4.3). Para factorizar A na forma $A = PDP^{-1}$ vamos calcular os espaços próprios considerando-os como subespaços de $\mathbb{C}^n$ (a matriz é real mas tem valores próprios complexos).

$$(A-5I)\mathbf{x} = \begin{bmatrix} -3 & 5 & 0 \\ -1 & -5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{cases} -3a+5b & = 0 \\ -a-5b & = 0 \end{cases} \iff a = b = 0.$$

Logo,

$$E(5) = \{(0,0,c) : c \in \mathbb{C}\} = \text{Span}\{(0,0,1)\}.$$

$$(A-(1+2i)I)\mathbf{x} = \begin{bmatrix} 1-2i & 5 & 0 \\ -1 & -1-2i & 0 \\ 0 & 0 & 4-2i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\iff \begin{cases} -x-(1+2i)y & = 0 \\ (4-2i)z & = 0 \end{cases} \iff x = -(1+2i)y \text{ e } z = 0.$$

Note-se que a matriz $A-(1+2i)I$ deverá ter determinante igual a zero uma vez que $1+2i$ é valor próprio de A. Ou seja, as linhas de $A-(1+2i)I$ são linearmente dependentes. Desta observação podemos concluir (sem verificação adicional) que as duas primeiras linhas da matriz são linearmente dependentes, e consequentemente o sistema $(A-(1+2i)I)\mathbf{x} = \mathbf{0}$ reduz-se às duas equações indicadas acima. Logo,

$$E(1+2i) = \{(-(1+2i)y, y, 0) :\ y \in \mathbb{C}\} = \text{Span}\{(1+2i,-1,0)\}.$$

Como a valores próprios complexos conjugados correspondem vectores próprios conjugados, tem-se

$$E(1-2i) = \{(-(1-2i)\bar{y}, \bar{y}, 0) :\ y \in \mathbb{C}\} = \text{Span}\{(1-2i,-1,0)\}.$$

Assim, uma matriz P que diagonaliza A e a correspondente matriz diagonal D podem ser

$$P = \begin{bmatrix} 1+2i & 1-2i & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1+2i & 0 & 0 \\ 0 & 1-2i & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

Sugere-se que confirme a igualdade $A = PDP^{-1}$. ♦

Como vimos, uma matriz A de ordem n é diagonalizável se e só se a soma das dimensões dos espaços próprios de A (ou seja, a soma das multiplicidades geométricas dos valores próprios de A) for igual a n. Já se observou no Exemplo 4.7 que existem matrizes cuja soma das multiplicidades geométricas dos seus valores próprios é inferior à ordem da matriz. Coloca-se naturalmente a questão de saber se essa soma pode ser superior a n. A resposta a esta questão é negativa como se deduz da proposição seguinte.

Proposição 4.7. Se λ é um valor próprio da matriz A, então

$$mult\,geom(\lambda) \leq mult\,alg(\lambda).$$

Demonstração. Seja A uma matriz de ordem n e μ um valor próprio de A com multiplicidade algébrica igual a k e multiplicidade geométrica igual a r.

É óbvio que $r \leq n$, caso contrário $r = \dim N(A - \mu I)$ seria maior do que n, o que é impossível já que a ordem de A é n. De igual modo, $k \leq n$ já que $\det(A - \lambda I)$ é um polinómio de grau n e portanto não tem raízes de multiplicidade superior a n.

Provemos agora que $r \leq k \leq n$. Suponhamos, por absurdo, que $r > k$. Ou seja, que existem r vectores próprios $\mathbf{u}_1, \ldots, \mathbf{u}_r$ linearmente independentes associados ao valor próprio μ. Podemos completar o conjunto $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ por forma a obter uma base de $\mathbb{C}^n$. Seja $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots, \mathbf{v}_{n-r}\}$ uma tal base, e P a matriz cujas colunas são os vectores de B, colocados segundo a ordem pela qual aparecem em B. As primeiras r colunas $\mathbf{u}_i$ de P verificam $A\mathbf{u}_i = \mu\mathbf{u}_i$, e portanto PAP^{-1} é uma matriz em blocos da forma

$$PAP^{-1} = T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix},$$

onde T_{11} é uma matriz diagonal do tipo $r \times r$ com todas as entradas na diagonal principal iguais a μ, isto é, $T_{11} = \mu I_r$. A matriz A e a matriz $T = PAP^{-1}$ têm o mesmo polinómio característico (cf. Proposição 4.5), ou seja

$$p(\lambda) = \det(A - \lambda I) = \det(T - \lambda I) = \begin{vmatrix} T_{11} - \lambda I_r & T_{12} \\ 0 & T_{22} - \lambda I_{n-r} \end{vmatrix}.$$

Efectuando r aplicações sucessivas do desenvolvimento de Laplace segundo a primeira coluna, tem-se

$$p(\lambda) = \det(T - \lambda I) = \det(T_{11} - \lambda I_r)\det(T_{22} - \lambda I_{n-r}) = (\mu - \lambda)^r q(\lambda),$$

onde q é um polinómio de grau $n-r$. Da expressão anterior conclui-se que μ é uma raiz de p com multiplicidade algébrica pelo menos $r > k$, o que é uma contradição.

□

A soma das multiplicidades algébricas dos valores próprios de uma matriz de ordem n é exactamente n, consequentemente segue como corolário da proposição anterior, da Proposição 4.6 e do Teorema 4.1, o seguinte:

Corolário 4.4. Uma matriz A é diagonalizável se e só se todo o valor próprio λ de A satisfaz a igualdade

$$mult\, alg(\lambda) = mult\, geom(\lambda).$$

Ou seja, uma matriz é diagonalizável se e só todos os valores próprios são semi-simples.

Terminamos esta secção referindo alguns resultados sobre diagonalização de matrizes reais com valores próprios complexos.

Valores próprios complexos

Recordemos que se λ é um valor próprio complexo não real de uma matriz real A de ordem n, e $\mathbf{x}$ um vector próprio associado a λ, então $\mathbf{x}$ não é um vector de $\mathbb{R}^n$. Assim, se a matriz real A é diagonalizável e tem valores próprios complexos, então a matriz P na factorização $A = PDP^{-1}$ possui entradas complexas (ver Exemplo 4.9). É habitual designar-se este facto dizendo que a matriz A é *diagonalizável em* $\mathbb{C}^n$.

Quando uma matriz diagonalizável A tem valores próprios complexos a factorização $A = PDP^{-1}$ (com D diagonal), não é a factorização mais conveniente para, por exemplo, estudar o comportamento geométrico da função $f : \mathbb{R}^n \to \mathbb{R}^n$, definida por $\mathbf{x} \mapsto A\mathbf{x} = PDP^{-1}\mathbf{x}$, uma vez que $P^{-1}\mathbf{x}$ não pertence a $\mathbb{R}^n$. No sentido de esclarecer esta questão, iremos mostrar que se A é diagonalizável em $\mathbb{C}^n$, então existem matrizes <u>reais</u> M e Σ tais que $A = M\Sigma M^{-1}$, com Σ uma matriz diagonal por blocos, com um bloco diagonal (correspondente aos valores próprios reais), e blocos 2×2 da forma

$$S = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}. \tag{4.10}$$

(correspondentes a cada par de valores próprios complexos $\lambda = a \pm bi$). Ou seja,

$$\Sigma = \begin{bmatrix} D & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & S_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & S_k \end{bmatrix}, \tag{4.11}$$

onde as matrizes S_i são da forma (4.10), a matriz D é uma matriz diagonal tendo na diagonal principal os valores próprios reais de A e $\mathbf{0}$ designa matrizes nulas.

Antes de procedermos à demonstração deste teorema estabelecem-se alguns resultados preliminares.

Define-se a parte real e a parte imaginária de um vector $\mathbf{u} \in \mathbb{C}^n$ como sendo os vectores de $\mathbb{R}^n$ cujas componentes são, respectivamente, a parte real e a parte imaginária das componentes correspondentes de $\mathbf{u}$. Por exemplo, se $\mathbf{u} = (5 - i, -2 + 3i)$, então $\operatorname{Re}\mathbf{u} = (5, -2)$ e $\operatorname{Im}\mathbf{u} = (-1, 3)$.

Lema 4.1. Os vectores $\mathbf{u} \in \mathbb{C}^n$ e $\overline{\mathbf{u}} \in \mathbb{C}^n$ são linearmente independentes sobre $\mathbb{C}$ se e só se $\operatorname{Re}\mathbf{u}$ e $\operatorname{Im}\mathbf{u}$ são vectores (de $\mathbb{R}^n$) linearmente independentes.

Demonstração. Qualquer vector $\mathbf{u}$ de $\mathbb{C}^n$ escreve-se na forma $\mathbf{u} = \operatorname{Re}\mathbf{u} + i\operatorname{Im}\mathbf{u}$. Como $\operatorname{Re}\mathbf{u} = \operatorname{Re}\overline{\mathbf{u}}$ e $\operatorname{Im}\overline{\mathbf{u}} = -\operatorname{Im}\mathbf{u}$, tem-se

$$\begin{aligned}(\alpha + i\beta)\mathbf{u} + (\gamma + i\delta)\overline{\mathbf{u}} &= (\alpha + i\beta)\left(\operatorname{Re}\mathbf{u} + i\operatorname{Im}\mathbf{u}\right) + (\gamma + i\delta)\left(\operatorname{Re}\mathbf{u} - i\operatorname{Im}\mathbf{u}\right) \\ &= \left[(\alpha + \gamma)\operatorname{Re}\mathbf{u} + (\delta - \beta)\operatorname{Im}\mathbf{u}\right] + i\left[(\beta + \delta)\operatorname{Re}\mathbf{u} + (\alpha - \gamma)\operatorname{Im}\mathbf{u}\right].\end{aligned}$$

Logo, $(\alpha + i\beta)\mathbf{u} + (\gamma + i\delta)\overline{\mathbf{u}} = 0 + 0i$ é equivalente ao sistema

$$\begin{cases} (\alpha + \gamma)\operatorname{Re}\mathbf{u} + (\delta - \beta)\operatorname{Im}\mathbf{u} = 0 \\ (\beta + \delta)\operatorname{Re}\mathbf{u} + (\alpha - \gamma)\operatorname{Im}\mathbf{u} = 0. \end{cases} \tag{4.12}$$

Uma vez que se verifica a equivalência

$$\alpha = \beta = \gamma = \delta = 0 \iff \begin{cases} \alpha + \gamma = 0 \\ \alpha - \gamma = 0 \\ \delta + \beta = 0 \\ \delta - \beta = 0, \end{cases}$$

os vectores $\mathbf{u}$ e $\overline{\mathbf{u}}$ são linearmente independentes se e só se $\operatorname{Re}\mathbf{u}$ e $\operatorname{Im}\mathbf{u}$ são linearmente independentes. $\square$

Proposição 4.8. Seja A uma matriz real 2×2 com valores próprios $a\pm bi$ ($b \neq 0$) e $\mathbf{v} \in \mathbb{C}^2$ um vector próprio associado a $\lambda = a - bi$. Então,

$$A = MSM^{-1}, \quad \text{com} \quad M = \begin{bmatrix} | & | \\ \operatorname{Re}\mathbf{v} & \operatorname{Im}\mathbf{v} \\ | & | \end{bmatrix} \quad \text{e} \quad S = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

onde $\operatorname{Re}\mathbf{v}$ e $\operatorname{Im}\mathbf{v}$ designam, respectivamente, a parte real e a parte imaginária do vector $\mathbf{v}$.

Demonstração. O Lema 4.1 garante que as colunas de M são linearmente independentes, uma vez que o vectores $\mathbf{v}$ e $\overline{\mathbf{v}}$ são vectores próprios associados a valores próprios distintos, e portanto linearmente independentes (cf. Proposição 4.6). Assim, M é invertível e $A = MSM^{-1}$ é equivalente a $AM = MS$. Necessitamos pois de mostrar que $AM = MS$, com M e S da forma indicada no enunciado. Ora,

$$\begin{aligned} MS &= \begin{bmatrix} | & | \\ \operatorname{Re}\mathbf{v} & \operatorname{Im}\mathbf{v} \\ | & | \end{bmatrix} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \\ &= \begin{bmatrix} M\begin{bmatrix} a \\ b \end{bmatrix} & M\begin{bmatrix} -b \\ a \end{bmatrix} \end{bmatrix} && \text{(pela Definição 1.12)} \\ &= \begin{bmatrix} | & | \\ a\operatorname{Re}\mathbf{v} + b\operatorname{Im}\mathbf{v} & -b\operatorname{Re}\mathbf{v} + a\operatorname{Im}\mathbf{v} \\ | & | \end{bmatrix} && \text{(pela Definição 1.11)} \end{aligned}$$

e

$$AM = A\begin{bmatrix} | & | \\ \operatorname{Re}\mathbf{v} & \operatorname{Im}\mathbf{v} \\ | & | \end{bmatrix} = \begin{bmatrix} | & | \\ A\operatorname{Re}\mathbf{v} & A\operatorname{Im}\mathbf{v} \\ | & | \end{bmatrix} \qquad \text{(pela Definição 1.12).}$$

Da definição de valor e vector próprio, temos

$$\begin{aligned} A\mathbf{v} = \lambda\mathbf{v} &\iff A\left(\operatorname{Re}\mathbf{v} + i\operatorname{Im}\mathbf{v}\right) = (a - ib)\left(\operatorname{Re}\mathbf{v} + i\operatorname{Im}\mathbf{v}\right) \\ &\iff A\operatorname{Re}\mathbf{v} + iA\operatorname{Im}\mathbf{v} = \left(a\operatorname{Re}\mathbf{v} + b\operatorname{Im}\mathbf{v}\right) + i\left(-b\operatorname{Re}\mathbf{v} + a\operatorname{Im}\mathbf{v}\right) \\ &\iff A\operatorname{Re}\mathbf{v} = a\operatorname{Re}\mathbf{v} + b\operatorname{Im}\mathbf{v} \quad \text{e} \quad A\operatorname{Im}\mathbf{v} = -b\operatorname{Re}\mathbf{v} + a\operatorname{Im}\mathbf{v}, \end{aligned}$$

onde na última equivalência aplicámos o facto de dois vectores complexos serem iguais se e só se as respectivas partes reais e imaginárias forem iguais. Por conseguinte, a igualdade $AM = MS$ é satisfeita. ☐

Enunciemos agora o teorema já referido.

Teorema 4.2. Seja A uma matriz real, $n \times n$, diagonalizável, e com p valores próprios reais e k pares de valores próprios complexos conjugados ($p + 2k = n$). Existem matrizes reais M e Σ tais que $A = M\Sigma M^{-1}$. A matriz Σ é uma matriz diagonal por blocos da forma

$$\Sigma = \begin{bmatrix} D & 0 & \cdots & 0 \\ 0 & S_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_k \end{bmatrix}.$$

Os blocos (na diagonal) de Σ têm as seguintes propriedades:

- O bloco D é uma matriz diagonal de ordem p com entradas na diagonal principal iguais aos valores próprios reais de A, repetidos de acordo com as suas multiplicidades.

- Cada bloco S_j é um bloco 2×2 da forma $\begin{bmatrix} a_j & -b_j \\ b_j & a_j \end{bmatrix}$, com $a_j \pm ib_j$ um par de valores próprios complexos conjugados de A.

As colunas de M têm as seguintes propriedades::

- Para $i = 1, \ldots, p$, a coluna i de M é um vector próprio $\mathbf{v}_i$ associados ao valor próprio real λ_i de A.

- As colunas de $p + 1$ a n são, respectivamente, os pares de vectores $\operatorname{Re} \mathbf{v}_j$ e $\operatorname{Im} \mathbf{v}_j$, ($j = 1, \ldots, k$), onde $\mathbf{v}_j$ é um vector próprio associado ao valor próprio (complexo) $\lambda_j = a_j - ib_j$.

Demonstração. A matriz A é diagonalizável e portanto existe uma base de $\mathbb{C}^n$ constituída por n vectores próprios de A. Seja $\{\mathbf{u}_1, \ldots, \mathbf{u}_p, \mathbf{v}_1, \overline{\mathbf{v}}_1, \ldots, \mathbf{v}_k, \overline{\mathbf{v}}_k\}$ uma base de $\mathbb{C}^n$, em que $\mathbf{u}_i$ é um vector próprio associado a um valor próprio real e $\mathbf{v}_j, \overline{\mathbf{v}}_j$ são vectores próprios associados, respectivamente, ao par de valores próprios complexos $a_j - ib_j, a_j + ib_j$. Do Lema 4.1 e da Proposição 4.6, segue que

$$B = (\mathbf{u}_1, \ldots, \mathbf{u}_p, \operatorname{Re} \mathbf{v}_1, \operatorname{Im} \mathbf{v}_1, \ldots, \operatorname{Re} \mathbf{v}_k, \operatorname{Im} \mathbf{v}_k)$$

é uma base ordenada de $\mathbb{R}^n$.

Coloquem-se os vectores da base B numa matriz M por colunas, respeitando a ordem de B. Efectuando os produtos MA e ΣM, obtém-se $MA = \Sigma M$ (cf. Proposição 4.8). Como as colunas de M formam uma base, a matriz M é invertível e portanto $MA = \Sigma M$ é equivalente a $MAM^{-1} = \Sigma$. $\square$

Ilustremos a aplicabilidade deste teorema à matriz do Exemplo 4.9. Nesse exemplo, verificámos que o espectro de A é $\sigma(A) = \{5, 1+2i, 1-2i\}$, e consequentemente A é diagonalizável. Os espaços próprios são

$$E(5) = \mathrm{Span}\{(0,0,1)\}, \quad E(1-2i) = \mathrm{Span}\{(1-2i, -1, 0)\}.$$

Seja $\mathbf{v} = (1-2i, -1, 0)$ um vector próprio associado a $\lambda = 1-2i$. Os vectores $\mathrm{Re}\,\mathbf{v}$ e $\mathrm{Im}\,\mathbf{v}$ são

$$\mathrm{Re}\,\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathrm{Im}\,\mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}.$$

Por conseguinte, o Teorema 4.2 diz-nos que podemos tomar para M e Σ as matrizes

$$M = \begin{bmatrix} 0 & 1 & -2 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{e} \quad \Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 2 & 1 \end{bmatrix}.$$

Ou seja, $A = M\Sigma M^{-1}$ com M e Σ reais. Sugerimos que compare esta factorização com a factorização $A = PDP^{-1}$ obtida no Exemplo 4.9.

Finalizamos esta secção fazendo uma referência breve ao comportamento da função $f : \mathbb{R}^n \to \mathbb{R}^n$, definida por $f(\mathbf{x}) = A\mathbf{x}$, no caso em que A é diagonalizável e possui valores próprios complexos.

Seja A uma matriz real, de ordem n, diagonalizável. Então, $A = M\Sigma M^{-1}$ com M real e Σ uma matriz real, diagonal por blocos, da forma (4.11). A matriz M é invertível, e pelo Teorema 3.8 a matriz M realiza a mudança da base constituída pelos vectores coluna de M (que é uma base de $\mathbb{R}^n$ constituída por vectores próprios de A correspondentes a valores próprios reais, e pelos vectores das partes reais e imaginárias de vectores próprios associados aos valores próprios complexos) para a base canónica de $\mathbb{R}^n$. A matriz M determina uma mudança de variáveis de $\mathbf{x}$ para $\mathbf{y}$, mediante a igualdade $\mathbf{x} = M\mathbf{y}$, ou seja, $f(\mathbf{x}) = A\mathbf{x} = M\Sigma M^{-1}\mathbf{x} = M\Sigma\mathbf{y}$.

A acção de f sobre um vector $\mathbf{x}$ (ou equivalentemente a acção de A sobre $\mathbf{x}$) pode traduzir-se do seguinte modo: (i) fazer a mudança de variáveis de $\mathbf{x}$ para

$\mathbf{y}$; (ii) fazer actuar a matriz Σ em $\mathbf{y}$; (iii) seguidamente, sobre o vector obtido, efectuar a mudança de variáveis (inversa) para a variável inicial. No diagrama seguinte ilustra-se este processo.

$$
\begin{array}{ccc}
\mathbb{R}^n \ni \mathbf{x} & \xrightarrow{A} & A\mathbf{x} \in \mathbb{R}^n \\
{\scriptstyle M^{-1}}\Big\downarrow & & \Big\uparrow{\scriptstyle M} \\
\mathbb{R}^n \ni \mathbf{y} & \xrightarrow{\Sigma} & \Sigma\mathbf{y} \in \mathbb{R}^n
\end{array}
$$

A matriz Σ é constituída por blocos diagonais e por blocos associados a pares de valores próprios complexos conjugados, os quais são matrizes do tipo estudado no Caso 2 da Secção 4.1.1 (ver página 184). Do estudo efectuado nessa secção, sabemos como actuam os blocos de Σ em vectores de $\mathbb{R}^n$.

Exemplo 4.10. Considere-se a matriz $A = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} \\ \frac{1}{2\sqrt{3}} & \frac{\sqrt{3}}{2} \end{bmatrix}$.

Os valores próprios de A são $\lambda = \frac{1}{2}(\sqrt{3} \pm i)$. Podemos verificar que $\mathbf{v} = \begin{bmatrix} -i\sqrt{3} \\ 1 \end{bmatrix}$ é um vector próprio de A associado ao valor próprio $\lambda = \frac{1}{2}(\sqrt{3} - i)$, já que

$$A\mathbf{v} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} \\ \frac{1}{2\sqrt{3}} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} -i\sqrt{3} \\ 1 \end{bmatrix} = \frac{\sqrt{3} - i}{2} \begin{bmatrix} -i\sqrt{3} \\ 1 \end{bmatrix} = \lambda\mathbf{v}.$$

A Proposição 4.8 garante que a matriz A é da forma $A = MSM^{-1}$, com

$$S = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \quad \text{e} \quad M = \begin{bmatrix} 0 & -\sqrt{3} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} | & | \\ \operatorname{Re}\mathbf{v} & \operatorname{Im}\mathbf{v} \\ | & | \end{bmatrix}.$$

Atendendo ao estudo realizado na Secção 4.1.1-Caso 2, sabemos que a matriz S é uma matriz de rotação (os valores próprios têm módulo igual a 1), e que S actua em vectores de $\mathbb{R}^2$ rodando-os (no sentido directo) em torno da origem de um ângulo $\frac{\pi}{6}$ (note que $\cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$ e $\operatorname{sen}\frac{\pi}{6} = \frac{1}{2}$) .

Na Figura 4.6 encontram-se representadas sucessivas aplicações de S ao vector $\mathbf{x}_0 = (1, 3)$ através de pontos a cor azul. Cada um destes pontos é obtido do anterior por uma rotação de $\pi/6$, ou seja, os pontos correspondentes a aplicações sucessivas de S situam-se sobre uma circunferência de centro na origem e raio igual à distância de $\mathbf{x}_0$ à origem, isto é, de raio igual a $\sqrt{10}$.

As imagens de aplicações sucessivas de A ao mesmo ponto $\mathbf{x}_0$ (representadas na Figura 4.6 a vermelho) estão por sua vez sobre uma elipse. Note-se que $\mathbf{x} = M\mathbf{y}$ é da forma

$$\mathbf{x} = M\mathbf{y} \Longleftrightarrow \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} -\sqrt{3}y_1 \\ y_0 \end{bmatrix},$$

e portanto se $\mathbf{x} = (x_0, x_1)$ pertence à circunferência de equação $x_0^2 + x_1^2 = 10$, então o ponto $\mathbf{y} = (y_0, y_1)$ pertence à elipse definida por $y_0^2 + 3y_1^2 = 10$.

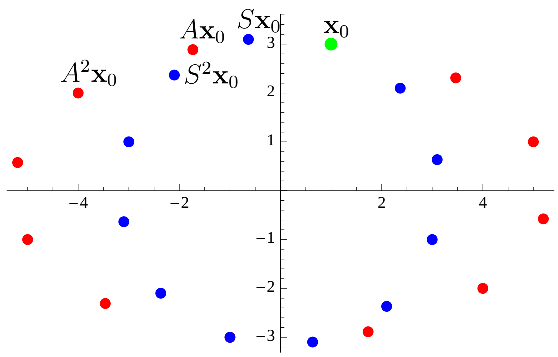

Figura 4.6: Sucessivas aplicações de $f(\mathbf{x}) = A\mathbf{x}$ e de $g(\mathbf{w}) = S\mathbf{w}$ ao ponto $\mathbf{x}_0$, onde $A = MSM^{-1}$. A matriz A tem valores próprios complexos de módulo 1.

4.3 Potências de uma matriz e valores próprios

Certas propriedades dos valores e vectores próprios de potências de uma matriz desempenham um papel fundamental em álgebra linear e nas aplicações. Neste texto apresentam-se alguns exemplos ilustrativos da relevância dos valores e vectores próprios de potências de matrizes, nomeadamente no estudo do comportamento a longo prazo de cadeias de Markov (estudadas na próxima secção), ou na determinação da forma canónica de Jordan de uma matriz (tratada no Capítulo 8).

Comecemos por observar que sendo A uma matriz diagonalizável de ordem n, isto é, $A = PDP^{-1}$ com $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$, qualquer potência positiva de A também é uma matriz diagonalizável, visto que

$$A^k = \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{k \text{ factores}} = PD^kP^{-1}, \tag{4.13}$$

onde $D^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$. A igualdade $P^{-1}A^kP = D^k$ implica que os valores próprios de A^k são $\lambda_1^k, \ldots, \lambda_n^k$ (a matriz A^k é semelhante à matriz D^k). Além disso, as colunas de P, que são vectores próprios de A, também são vectores próprios de A^k já que: se $(\lambda, \mathbf{u})$ é um par próprio de A, então $(\lambda^k, \mathbf{u})$ é um par próprio de A^k (ver demonstração da próxima proposição).

Proposição 4.9. Seja $p(\lambda) = (-1)^n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$ o polinómio característico da matriz A e $\mathbf{u}$ um vector próprio de A. Então

$$p(A)\mathbf{u} = \left[(-1)^nA^n + b_{n-1}A^{n-1} + \cdots + b_1A + b_0\right]\mathbf{u} = \mathbf{0}.$$

Demonstração. Comecemos por mostrar que se $(\mu, \mathbf{u})$ é um par próprio de A, então $(\mu^k, \mathbf{u})$ é um par próprio de A^k, com k um inteiro positivo. De facto, se $A\mathbf{u} = \mu\mathbf{u}$, resulta

$$\begin{aligned} A^k\mathbf{u} &= A^{k-1}(A\mathbf{u}) = A^{k-1}(\mu\mathbf{u}) = \mu A^{k-1}\mathbf{u} = \mu A^{k-2}(A\mathbf{u}) = \mu^2A^{k-2}\mathbf{u} \\ &= \cdots = \mu^k\mathbf{u}. \end{aligned}$$

Por definição de valor e vector próprio, a igualdade $A^k\mathbf{u} = \mu^k\mathbf{u}$ significa que μ^k é um valor próprio de A e $\mathbf{u}$ é um vector próprio associado.

Para mostrar que $p(A)\mathbf{u} = \mathbf{0}$ basta mostrar que $p(A)\mathbf{u} = p(\mu)\mathbf{u}$, onde $(\mu, \mathbf{u})$ é um par próprio de A. Aplicando o facto de $(\mu^k, \mathbf{u})$ ser um par próprio de A^k no cálculo de $p(A)\mathbf{u}$, tem-se

$$\begin{aligned} p(A)\mathbf{u} &= (-1)^nA^n\mathbf{u} + b_{n-1}A^{n-1}\mathbf{u} + \cdots + b_1A\mathbf{u} + b_0\mathbf{u} \\ &= (-1)^n\mu^n\mathbf{u} + b_{n-1}\mu^{n-1}\mathbf{u} + \cdots + b_1\mu\mathbf{u} + b_0\mathbf{u} \\ &= ((-1)^n\mu^n + b_{n-1}\mu^{n-1} + \cdots + b_1\mu + b_0)\mathbf{u} = p(\mu)\mathbf{u} = 0 \times \mathbf{u} = \mathbf{0}, \end{aligned}$$

onde a penúltima igualdade resulta do facto de μ ser valor próprio de A, e portanto raiz de p. $\square$

Corolário 4.5. Seja $p(\lambda) = (-1)^n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0 = 0$ a equação característica da matriz A. Se A tem n vectores próprios linearmente independentes, a matriz A satisfaz a sua equação característica. Isto é,

$$p(A) = (-1)^nA^n + b_{n-1}A^{n-1} + \cdots + b_1A + b_0 = \mathbf{O},$$

onde $\mathbf{O}$ designa a matriz nula.

Demonstração. Sejam $\mathbf{u}_1, \ldots, \mathbf{u}_n$ vectores próprios linearmente independentes de A e X a matriz cujas colunas são estes vectores. Da proposição anterior tem-se $p(A)\mathbf{u}_i = \mathbf{0}$ para $i = 1, \ldots, n$. Assim, usando a Definição 1.12 de produto de matrizes, obtemos

$$p(A)X = p(A)\begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ p(A)\mathbf{u}_1 & p(A)\mathbf{u}_2 & \ldots & p(A)\mathbf{u}_n \\ | & | & & | \end{bmatrix} = \mathbf{O}.$$

Como as colunas de X são linearmente independentes, a matriz X é invertível (Proposição 3.15). Logo, multiplicando (à direita) a equação matricial $p(A)X = \mathbf{O}$ por X^{-1}, obtém-se $p(A) = \mathbf{O}$. □

O resultado do corolário anterior é igualmente válido no caso da matriz não admitir n vectores próprios linearmente independentes. Esta generalização constitui o famoso Teorema de Cayley-Hamilton[2] que passamos a enunciar.

Teorema 4.3. Cayley-Hamilton
Toda a matriz quadrada verifica a sua equação característica. Ou seja, se $(-1)^n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$ é o polinómio característico de A, então

$$(-1)^n A^n + b_{n-1}A^{n-1} + \cdots + b_1 A + b_0 I = \mathbf{O}.$$

O exercício guiado seguinte apresenta uma demonstração do Teorema de Cayley-Hamilton.

Exercício 4.2. Mostre o Teorema de Cayley-Hamilton.
Comece por justificar por que razão a matriz adjunta $\operatorname{adj}(A - \lambda I)$ pode ser escrita na forma $\operatorname{adj}(A-\lambda I) = B_{n-1}\lambda^{n-1} + \cdots + B_1\lambda + B_0$, onde B_i são matrizes $n \times n$. Aplique a fórmula (2.14), na página 102, à matriz $(A - \lambda I)$, e conclua que

$$\begin{aligned} -B_{n-1} &= (-1)^n I \\ AB_{n-k} - B_{n-k-1} &= b_{n-k} I \quad \text{para } k = 1, \ldots, n-1 \\ AB_0 &= b_0 I. \end{aligned}$$

Multiplique as igualdades anteriores respectivamente por A^n, A^{n-k} e I, adicione, e obterá $p(A) = \mathbf{O}$.

▲

[2] Arthur Cayley (1821 – 1895), matemático inglês. Sir William Rowan Hamilton (1805 – 1865), físico, astrónomo e matemático irlandês.

4.4 Aplicações: Sistemas dinâmicos

Apresentamos nesta secção alguns exemplos ilustrativos da importância dos valores e vectores próprios de uma matriz no estudo de certos modelos matemáticos. Em primeiro lugar trataremos sistemas dinâmicos discretos e posteriormente equações diferenciais (sistemas dinâmicos contínuos).

Um sistema dinâmico é um modelo matemático que descreve a evolução no tempo do estado de um sistema. O modelo matemático procura descrever o resultado de uma determinada experiência, efectuada repetidas vezes. Nos modelos mais simples, o resultado de cada experiência depende apenas do resultado da experiência anterior. Consideremos o exemplo seguinte.

Exemplo 4.11. Uma universidade tem na totalidade dos seus cursos de licenciatura (com a duração de três anos) um *numerus clausus* de 850 alunos. Em cada ano lectivo, 80% dos estudantes dos cursos de licenciatura transitam de ano (ou terminam, caso estejam no terceiro ano) e 20% ficam retidos no mesmo ano. O número de alunos que frequentam as licenciaturas dessa universidade no ano lectivo k representa-se pelo *vector de estado* $\mathbf{x}_k \in \mathbb{R}^3$ cujas componentes x_{k1}, x_{k2} e x_{k3} são, respectivamente, o número de alunos no primeiro, segundo e terceiro ano das licenciaturas. Suponha-se que número de alunos de licenciatura no ano lectivo 2010/11, era 1600 no primeiro ano, 950 no segundo ano e 1100 no terceiro. Representamos o número de alunos de licenciatura no ano lectivo 2010/11 pelo vector de estado $\mathbf{x}_0 = (1600, 950, 1100)$. O número de alunos de licenciatura no ano lectivo seguinte é representado pelo vector $\mathbf{x}_1$. De acordo com os dados do problema o vector $\mathbf{x}_1$ é

$$\mathbf{x}_1 = \begin{bmatrix} 850 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2 & 0 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0.8 & 0.2 \end{bmatrix} \mathbf{x}_0 = \mathbf{b} + A\mathbf{x_0}.$$

O número de alunos de licenciatura nos anos lectivos subsequentes é representado pelos vectores de estado $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \ldots$. O vector de estado $\mathbf{x}_{k+1}$ é dado por

$$\mathbf{x}_{k+1} = \mathbf{b} + A\mathbf{x}_k \quad k = 0, 1, 2, \ldots \tag{4.14}$$

A equação (4.14) é uma fórmula de recorrência que permite obter o vector de estado de um determinado ano lectivo à custa dos vectores de estado de anos lectivos anteriores. ♦

Estudamos a seguir sistemas dinâmicos discretos lineares do tipo

$$\mathbf{x}_{k+1} = A\mathbf{x}_k, \quad \text{com} \quad \mathbf{x}_k \in \mathbb{R}^n \text{ e } k = 0, 1, 2, 3, \ldots$$

onde A é uma matriz de ordem n. Este tipo de sistemas é também designado por *equação às diferenças*, de 1ª ordem homogénea. Note-se que (4.14) é uma equação às diferenças de 1ª ordem não homogénea.

O análogo contínuo de uma equação às diferenças é uma *equação diferencial*. Uma equação diferencial modela sistemas físicos cujos estados são observados de forma contínua. No final deste capítulo estudaremos este tipo de equações.

4.4.1 Sistemas dinâmicos discretos

Consideremos o sistema dinâmico discreto

$$\mathbf{x}_k = A\mathbf{x}_{k-1}, \quad k = 0, 1, 2, \ldots$$

onde A é uma matriz real do tipo $n \times n$.

Chamamos *órbita* do ponto $\mathbf{x}_0$ ao conjunto de pontos da sucessão $\{\mathbf{x}_s\}$, isto é, $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$. A órbita do ponto inicial $\mathbf{x}_0$ é determinada pelas potências A^k e por $\mathbf{x}_0$, uma vez que

$$\begin{aligned}
&\mathbf{x}_1 = A\mathbf{x}_0 \\
&\mathbf{x}_2 = A\mathbf{x}_1 = A^2\mathbf{x}_0 \\
&\vdots \\
&\mathbf{x}_k = A\mathbf{x}_{k-1} = A^k\mathbf{x}_0, \quad k \geq 1.
\end{aligned}$$

Suponhamos que A é diagonalizável e que existe uma base $B = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ de $\mathbb{R}^n$ formada por vectores próprios de A. Nesta base, o ponto inicial $\mathbf{x}_0$ escreve-se como combinação linear (única) dos vectores de B, seja

$$\mathbf{x}_0 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n.$$

Como vimos na demonstração da Proposição 4.9, se $(\lambda_i, \mathbf{u}_i)$ é um par próprio de A, então $(\lambda_i^k, \mathbf{u}_i)$ é um par próprio de A^k. Por conseguinte, o estado do sistema no instante k é dado por

$$\begin{aligned}
\mathbf{x}_k = A^k\mathbf{x}_0 &= A^k(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n) \\
&= c_1A^k\mathbf{u}_1 + c_2A^k\mathbf{u}_2 + \cdots + c_nA^k\mathbf{u}_n \qquad (4.15) \\
&= c_1\lambda_1^k\mathbf{u}_1 + c_2\lambda_2^k\mathbf{u}_2 + \cdots + c_n\lambda_n^k\mathbf{u}_n,
\end{aligned}$$

onde $(\lambda_i, \mathbf{u}_i)$ é um par próprio de A, para $i = 1, \ldots, n$.

A expressão (4.15) permite determinar o *comportamento a longo prazo* do sistema. Este comportamento é dado pelo vector $\mathbf{x}_\infty$ definido por

$$\mathbf{x}_\infty = \lim_{k\to\infty} \mathbf{x}_k = \lim_{k\to\infty} \left[c_1\lambda_1^k\mathbf{u}_1 + c_2\lambda_2^k\mathbf{u}_2 + \cdots + c_n\lambda_n^k\mathbf{u}_n\right]. \tag{4.16}$$

Apresentamos a seguir alguns exemplos de sistemas dinâmicos discretos.

Números de Fibonacci e o número de Ouro

É no mínimo surpreendente como, quer na Natureza quer em certas criações artísticas em arquitectura e pintura, se podem encontrar os chamados números de Fibonacci. Um exemplo é a flor do girassol que tem 233 sementes em 144 espirais, números que correspondem aos 12º e 13º termos da sucessão de Fibonacci. Aconselhamos uma visita ao site "goldennumber.net" [3], ou ao site de Ron Knott[4], onde pode encontrar vários exemplos de como as sucessões de Fibonacci aparecem na Arte e na Natureza, bem como outras curiosidades relacionadas com o "número de ouro".

A sucessão de números de Fibonacci foi usada pelo seu criador, Leonardo de Pisa[5] (mais tarde conhecido por Fibonacci), como um modelo matemático simples para descrever o crescimento de uma população de coelhos, nas seguintes condições:

1) Admite-se que os coelhos não morrem;

2) Supõe-se que um casal de coelhos demora dois meses a atingir a maturidade, altura em que se passa a reproduzir mensalmente dando origem a um novo casal de coelhos.

Denotemos por F_k o número de casais de coelhos no mês k. A evolução da população de coelhos pode modelar-se da seguinte forma:

- O processo inicia-se no mês $k = 1$ com um casal de coelhos, isto é, $F_1 = 1$.
- No mês seguinte o número de casais é ainda igual a 1, ou seja, $F_2 = 1$, visto que o casal original está ainda imaturo.

[3] http://goldennumber.net/
[4] http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fibnat.html
[5] Leonardo de Pisa (1170 — 1250), matemático italiano considerado um dos mais talentosos matemático da Idade Média.

- Decorridos dois meses, tem-se $F_3 = 2$, correspondente ao casal original e a um casal recém-nascido.

A sucessão $\{F_k\}$ obtida pelo processo anterior é designada por *sucessão de Fibonacci*. O termo de ordem k da sucessão de Fibonacci, verifica

$$F_{k+2} = F_{k+1} + F_k \quad \text{com } F_1 = 1 \text{ e } F_2 = 1.$$

Os primeiros termos desta sucessão são

$$1, 1, 2, 3, 5, 8, 13, \dots$$

A sucessão de Fibonacci pode ser escrita na forma matricial da maneira que indicaremos a seguir. Usaremos os valores próprios e os vectores próprios da matriz que define a sucessão para obter uma expressão explícita (não recursiva) de F_k e provar que o crescimento da sucessão de Fibonacci é do tipo exponencial.

Escrevendo $\mathbf{x}_{k-1} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$, a equação $F_{k+2} = F_{k+1} + F_k$ é equivalente a

$$\begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} \Longleftrightarrow \mathbf{x}_k = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{x}_{k-1} = A\mathbf{x}_{k-1}, \quad \text{para } k = 1, 2, \dots \tag{4.17}$$

e $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} F_2 \\ F_1 \end{bmatrix}$.

Atendendo à expressão obtida em (4.15), se a matriz A for diagonalizável (e com valores próprios reais), tem-se

$$\mathbf{x}_{k-1} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = c_1 \lambda_1^{k-1} \mathbf{u}_1 + c_2 \lambda_2^{k-1} \mathbf{u}_2, \tag{4.18}$$

onde $(\lambda_1, \mathbf{u}_1)$ e $(\lambda_2, \mathbf{u}_2)$ são pares próprios A, e (c_1, c_2) é o vector das coordenadas de $\mathbf{x}_0$ na base ordenada $(\mathbf{u}_1, \mathbf{u}_2)$.

Comecemos por calcular os valores próprios de A.

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - \lambda - 1.$$

As raízes de $p(\lambda) = \lambda^2 - \lambda - 1$ são

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \approx 1.618\dots \quad \text{e} \quad \lambda_2 = \frac{1-\sqrt{5}}{2} \approx -0.618\dots.$$

O valor próprio λ_1 é conhecido por *número de ouro* sendo habitualmente designado pela letra ϕ.

A matriz A tem valores próprios reais e distintos, logo é diagonalizável (cf. Corolário 4.3). Determinemos uma base de $\mathbb{R}^2$ constituída por vectores próprios de A. Para $i = 1, 2$ temos

$$\begin{bmatrix} 1-\lambda_i & 1 \\ 1 & -\lambda_i \end{bmatrix} \mathbf{u}_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{cases} (1-\lambda_i)a + \quad b = 0 \\ \qquad\qquad a - \lambda_i b = 0 \end{cases} \iff a = \lambda_i b.$$

Logo, podemos tomar $\mathbf{u}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$ e $\mathbf{u}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$ como vectores próprios associados, respectivamente, a λ_1 e λ_2. As coordenadas de $\mathbf{x}_0$ na base ordenada $(\mathbf{u}_1, \mathbf{u}_2)$ são:

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} \iff \begin{cases} c_1\lambda_1 + c_2\lambda_2 = 1 \\ \quad c_1 + \quad c_2 = 1 \end{cases}$$
$$\iff c_1 = 1 - c_2, c_2 = \frac{\lambda_1 - 1}{\lambda_1 - \lambda_2} \iff c_1 = \frac{1}{\sqrt{5}}\phi, c_2 = -\frac{1}{\sqrt{5}}\lambda_2.$$

Assim, a expressão (4.18) toma a forma

$$\mathbf{x}_{k-1} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^k \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix} - \left(\frac{1-\sqrt{5}}{2} \right)^k \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix} \right].$$

Logo, os termos da sucessão de Fibonacci são dados por

$$F_k = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right], \quad k = 1, 2, \ldots. \tag{4.19}$$

A fórmula (4.19) é conhecida como a *fórmula de Binet* para os números de Fibonacci. Note que nesta fórmula temos que $\left(\frac{1-\sqrt{5}}{2}\right)^k \to 0$ quando $k \to \infty$, já que $0 < \lambda_2 < 1$. Por conseguinte, para k suficientemente grande, o valor de F_k pode aproximar-se por $\frac{1}{\sqrt{5}}\phi^k$, onde $\lambda_1 = \phi > 1$ é o maior valor próprio de A.

Matrizes de Markov

As chamadas cadeias de Markov[6] aparecem naturalmente na modelação matemática de problemas de biologia, química, economia, etc.. Trata-se de sistemas dinâmicos

[6] Andrey Markov (1856-1922), matemático russo.

discretos, $\mathbf{x}_{k+1} = M\mathbf{x}_k$, em que a matriz M é uma matriz cujos vectores coluna são *vectores de probabilidades*, isto é, vectores de componentes não negativas e tais que a soma das componentes é igual a 1.

Por exemplo, suponha-se que o administrador de uma firma de aluguer de viaturas, com três agências localizadas em cidades distintas, pretende estimar o número de carros em cada agência num dado instante. Admita-se que um cliente pode alugar uma viatura numa agência e entregá-la noutra. É claro que o administrador da firma não pode saber de antemão qual o número exacto de viaturas que será entregue numa dada agência, mas pode calcular a percentagem de carros que se encontram numa dada agência. Designe-se por $\mathbf{x}_k = (x_{k1}, x_{k2}, x_{k3})$ o vector de estado do mês k, isto é, o vector em que a componente x_{ki} representa a probabilidade de um carro da frota se encontrar na agência i ($i = 1, 2, 3$) no mês k. Calculando as probabilidades de um carro que está na agência i se encontrar na agência j no mês seguinte, o administrador da firma concluiu que $\mathbf{x}_{k+1} = M\mathbf{x}_k$, onde

$$M = \begin{bmatrix} 1/2 & 0 & 1/10 \\ 1/10 & 7/10 & 1/10 \\ 2/5 & 3/10 & 4/5 \end{bmatrix}.$$

Este é um problema típico que é modelado matematicamente por uma cadeia de Markov com três estados (as agências). A entrada p_{ij} da matriz $M = [p_{ij}]$ designa a probabilidade do sistema se encontrar no estado i quando na observação anterior se encontrava no estado j. A matriz M é designada por *matriz de transição* da cadeia de Markov. Como cada vector coluna de M é um vector de probabilidade, a matriz M goza da propriedade de ter todas as entradas não negativas e da soma das entradas de cada coluna ser constante e igual a 1. Estas matrizes são designadas por *matrizes de Markov* ou *matrizes estocásticas*.

Seja M uma matriz de Markov. Uma *cadeia de Markov* associada a M é uma sucessão de vectores de probabilidade, $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \ldots$ satisfazendo

$$\mathbf{x}_{k+1} = M\mathbf{x}_k, \quad k = 0, 1, 2, \ldots \tag{4.20}$$

Como se disse, as matrizes de Markov caracterizam-se por terem entradas não negativas (dizendo-se matrizes não negativas) e a soma das entradas de cada coluna ser igual a 1. Vemos a seguir que estas duas propriedades têm fortes implicações no tipo de valores próprios destas matrizes e no comportamento a longo prazo das cadeias de Markov.

Se A é uma matriz tal que a soma das entradas de cada linha é constante igual a s, então o vector $\mathbf{u} = (1, 1, \ldots, 1)$ satisfaz a igualdade

$$A\mathbf{u} = s\mathbf{u}.$$

A igualdade anterior significa que s é um valor próprio de A e $\mathbf{u}$ é um vector próprio associado. Como o determinante de uma matriz é igual ao determinante da sua transposta, tem-se $\det(A - \lambda I) = \det(A - \lambda I)^T = \det(A^T - \lambda I)$. Ou seja, as matrizes A e A^T têm os mesmos valores próprios. Por conseguinte, se uma matriz tem a soma das entradas de cada coluna constante, esta constante é um valor próprio da matriz.

Proposição 4.10. Se uma matriz A tem a soma das entradas de cada coluna (ou de cada linha) constante, então esta constante é um valor próprio de A.

Em particular, $\lambda = 1$ é um valor próprio de uma matriz de Markov.

Nota 23. *Apesar de uma matriz e a sua transposta terem os mesmos valores próprios, isso não significa que pares próprios de A sejam também pares próprios de A^T. Ou seja, se $\mathbf{v}$ é um vector próprio de A associado ao valor próprio λ, não significa que $\mathbf{v}$ é um vector próprio de A^T associado ao valor próprio λ. Deixamos como exercício encontrar um contra-exemplo.*

Dada uma cadeia de Markov definida pela matriz M, um *vector de equilíbrio*, ou *vector estacionário*, é um vector de probabilidades $\mathbf{q}$ que satisfaz

$$M\mathbf{q} = \mathbf{q}. \tag{4.21}$$

Ou seja, um vector de equilíbrio é um vector próprio de M associado ao valor próprio 1 que também é um vector de probabilidades. Saliente-se que, se $\mathbf{x}_k$ é um vector de equilíbrio, então $\mathbf{x}_j = \mathbf{x}_k$ para $j \geq k$, o que justifica a designação de vector de equilíbrio.

Para uma matriz de Markov, a existência de vectores próprios associados ao valor próprio $\lambda = 1$ que sejam vectores de probabilidades decorre da teoria geral das matrizes não negativas. A teoria das matrizes não negativas (conhecida pela designação de Teoria de Perron-Frobenius[7]) está fora do âmbito deste texto. Utilizamos no entanto alguns resultados fundamentais desta teoria, convidando o leitor interessado a consultar obras especializadas. Para os resultados que aqui utilizamos sugere-se a leitura de Meyer [9].

[7]Oskar Perron (1880 – 1975) e Ferdinand Georg Frobenius (1849 – 1917).

Define-se *raio espectral* de uma matriz como sendo o máximo dos módulos dos seus valores próprios. Isto é,

$$\rho(A) = \max_{i=1,\ldots,n} \{|\lambda_i|\},$$

onde λ_i é um valor próprio da matriz A (de ordem n). Resume-se no quadro seguinte alguns resultados da Teoria de Perron-Frobenius cuja prova pode encontrar em Meyer [9].

A Teoria de Perron-Frobenius para matrizes não negativas garante que uma matriz de Markov admite um vector próprio $\mathbf{u}$, associado ao valor próprio $\lambda = 1$, com todas as componentes positivas. Além disso, o raio espectral de uma matriz de Markov é exactamente 1.

No caso da matriz de Markov M ser positiva (isto é, com todas as entradas positivas), tem-se:

i) O valor próprio $\lambda = 1$ tem multiplicidade algébrica igual a 1 e os restantes valores próprios de M têm módulo inferior a 1.

ii) Se $\mathbf{y}$ é um qualquer vector de probabilidades, então

$$\lim_{k\to\infty} M^k\mathbf{y} = \mathbf{q}, \tag{4.22}$$

onde $\mathbf{q}$ é um vector de probabilidades cujas componentes são todas positivas.

O facto de uma matriz de Markov M possuir um par próprio $(1, \mathbf{u})$ em que o vector próprio $\mathbf{u} = (u_1, \ldots, u_n)$ tem todas as componentes positivas vai implicar que qualquer cadeia de Markov admita pelo menos um vector de equilíbrio. Com efeito, o vector $\mathbf{q} = \frac{1}{\sum_{i=1}^n u_i}\mathbf{u}$ ainda é um vector próprio de M associado ao valor próprio $\lambda = 1$, e além disso é um vector de probabilidades. Logo, $\mathbf{q}$ é um vector de equilíbrio.

Proposição 4.11. Qualquer cadeia de Markov admite pelo menos um vector de equilíbrio.

Saliente-se que o valor próprio $\lambda = 1$ de matrizes de Markov não negativas pode ter multiplicidade geométrica superior a um, e portanto haver mais do que um vector de equilíbrio. No entanto, quando a matriz de Markov é positiva, a unicidade do vector de equilíbrio está garantida, uma vez que o valor próprio $\lambda = 1$ é um valor próprio simples. Além disso, a longo prazo o sistema tende para o vector de equilíbrio independentemente do vector de estado inicial (pelo item (ii) no quadro acima).

No caso da matriz de Markov M ser positiva e diagonalizável, é fácil ver que a expressão (4.16) converge. De facto, como $\lambda_1 = 1$ é um valor próprio de M de multiplicidade algébrica 1, e os restantes valores próprios de M têm módulo inferior a 1, a expressão (4.16) reduz-se a

$$\mathbf{x}_\infty = \lim_{k\to\infty} M^k \mathbf{x}_0 = \lim_{k\to\infty} \left[c_1\lambda_1^k \mathbf{u}_1 + c_2\lambda_2^k \mathbf{u}_2 + \cdots + c_n\lambda_n^k \mathbf{u}_n\right] = c_1\mathbf{u}_1.$$

Como $\mathbf{x}_\infty$ é um vector de probabilidades (de facto é o vector de equilíbrio), a constante c_1 é igual ao inverso da soma das componentes de um vector próprio (positivo) $\mathbf{u}_1$ associado a $\lambda_1 = 1$.

Exemplo 4.12. Suponha-se que anualmente 1.5% da população que vive na área metropolitana de Lisboa (AML) muda-se para outras regiões do país, e 9% da população portuguesa muda-se para AML. Sabendo que no ano de 1970, 18% da população de Portugal vivia na AML, pertende-se determinar qual a distribuição da população portuguesa a longo prazo.

Tomando para vector de estado inicial $\mathbf{x}_0 = (0.18, 0.82)$, correspondendo a ter (em 1970) 18% da população na AML (e portanto 82% fora desta região), a evolução no tempo da percentagem da população portuguesa vivendo na AML é descrita pelo sistema $\mathbf{x}_k = M\mathbf{x}_{k-1}$, $k = 1, 2, \ldots$ onde M é a matriz de Markov

$$M = \begin{bmatrix} 0.985 & 0.09 \\ 0.015 & 0.91 \end{bmatrix}.$$

A matriz M corresponde aos movimentos transcritos na tabela seguinte (onde FAML designa não residentes na AML).

		De	
		AML	FAML
Para	AML	0.985	0.09
	FAML	0.015	0.91

A matriz M é positiva e tem 1 como valor próprio visto que a soma das entradas de cada coluna é igual a 1 (cf. Proposição 4.10). Como a soma dos valores próprios é igual ao traço da matriz (Proposição 4.3), o outro valor próprio é

$\lambda_2 = 0.895$. Uma vez que M é uma matriz positiva já sabíamos que $\lambda_1 = 1$ seria o maior valor próprio e que a sua multiplicidade algébrica seria igual a um. São vectores próprios associados, respectivamente a λ_1 e a λ_2, os vectores $\mathbf{u}_1 = (0.09, 0.015)$ e $\mathbf{u}_2 = (-1, 1)$, como pode confirmar calculando $M\mathbf{u}_1$ e $M\mathbf{u}_2$.

Usando a equação (4.15), no ano k, a percentagem da população portuguesa na AML e fora desta área é dada por

$$\mathbf{x}_k = c_1 \begin{bmatrix} 0.09 \\ 0.015 \end{bmatrix} + c_2(0.895)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \tag{4.23}$$

onde c_1 e c_2 são as coordenadas do vector inicial $\mathbf{x}_0 = (0.18, 0.82)$ na base ordenada $(\mathbf{u}_1, \mathbf{u}_2)$. A longo prazo, a distribuição da população portuguesa tende para

$$\mathbf{x}_\infty = \lim_{k\to\infty} \mathbf{x}_k = c_1 \begin{bmatrix} 0.09 \\ 0.015 \end{bmatrix},$$

uma vez que $(0.895)^k \to 0$ quando $k \to \infty$. Este vector é de facto, o vector de equilíbrio. Uma vez que $\mathbf{x}_\infty$ é um vector de probabilidades tem-se $c_1 = \frac{1}{0.09+0.015} = 0.105$. Pode confirmar-se este resultado calculando as coordenadas c_1 e c_2 de $\mathbf{x}_0$ na base $(\mathbf{u}_1, \mathbf{u}_2)$.

Assim, a longo prazo temos $\mathbf{x}_\infty = \dfrac{1}{0.105}\mathbf{u}_1 \approx (0.86, 0.14)$, ou seja, a longo prazo 86% da população portuguesa viverá na área metropolitana de Lisboa. ⧫

Exemplo 4.13. Este exemplo baseia-se no artigo de Kurt Bryan e Tanya Leise intitulado: ”The $\$25000000000$[8] eigenvector. The linear Algebra behind Google”, publicado em 2006 pela SIAM Review [3].

Nos finais dos anos 90 a empresa fundadora do motor de busca Google[9] apresentou um processo de pesquisa na *net* de palavras chave que listava os resultados segundo a sua relevância. Tal não acontecia com os motores de busca existentes à época, nos quais o utilizador era obrigado a percorrer várias páginas de listagem de sites irrelevantes até encontrar a informação desejada.

Um dos algoritmos usados pelo Google para seriar os sites por ordem decrescente de importância é o denominado algoritmo PageRank.[10]

[8] O valor estimado da empresa Google quando em 2004 se tornou uma empresa pública.

[9] Google é um trocadilho da palavra anglo-saxónica ”googol” a qual significa 10^{100}. O termo reflecte o número enorme e sempre crescente de utilizadores da net.

[10] PageRank foi desenvolvido na Universidade de Stanford (USA) por Larry Page e posteriormente por Sergey Brin como parte de um projecto de investigação. Page e Brin fundaram a empresa Google em 1998.

Apresentamos aqui um exemplo muito simples que ilustra a importância da álgebra linear na quantificação da relevância dos sites da net. A relevância de uma dada página é quantificada atribuindo-lhe uma classificação (um número real não negativo). Esta classificação depende do número de links (ligações ou citações) que essa página faz para outras páginas, bem como do número de citações que as outras páginas lhe fazem.

Suponha que o número de páginas (interligadas) numa rede é $n > 1$ e que cada página é designada por um inteiro k. Cada link vai representar-se por uma seta. Uma seta com origem em A e ponto final B, indica um link da página A para a página B. Um exemplo é a rede com cinco páginas representada pelo *grafo direccionado*[11] da Figura 4.7.

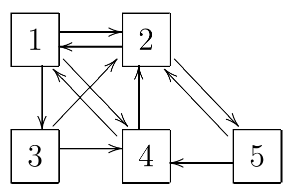

Figura 4.7: Uma rede com cinco páginas. Uma seta de A para B indica um link da página A para a B.

Designemos por x_k o valor da relevância (ou importância) da página k da rede. O valor de x_k é não negativo e $x_j > x_k$ significa que a página j tem mais importância que a página k. Uma forma simples de atribuir a importância a uma dada página k seria considerar x_k igual ao número de setas que entram na página k, ou seja, o número de citações que as outras páginas da rede fazem à página k. Por exemplo, para a rede da Figura 4.7 teríamos $x_1 = 2, x_2 = 4, x_3 = 1, x_4 = 3$ e $x_5 = 1$, significando que a página mais relevante seria a página 2 e as menos importantes a 3 e a 5.

A caracterização anterior é insuficiente visto que uma citação proveniente de uma página pouco importante não deve ter o mesmo valor que uma citação proveniente de outra mais importante, e as autocitações não devem ser consideradas. Uma outra forma de atribuir o valor da importância da página k seria considerar x_k igual à soma do valores das importâncias das páginas que a citam. Isso daria, $x_1 = x_2 + x_4$, $x_2 = x_1 + x_3 + x_4 + x_5$, $x_3 = x_1$, $x_4 = x_1 + x_3 + x_5$ e $x_5 = x_2$.

[11] Um grafo consiste num conjunto de vértices e arestas. Cada aresta liga um par de vértices. Um grafo diz-se direccionado se está atribuído um sentido a cada aresta.

Há contudo uma outra característica a levar em conta neste modelo, nomeadamente o facto de uma página não dever ganhar uma relevância superior só pelo simples facto de fazer muitas citações a outras páginas. Neste sentido, o valor da importância de uma dada página deve ser dividido pelo número de citações que faz a outras páginas da rede, ou seja, se a página j faz um total de s_j links para as outras páginas da rede deve considerar-se que a sua relevância é x_j/s_j (note que s_j é o número de setas que saem do vértice j do grafo). Ou seja, uma página que faz s citações confere a cada página citada o valor $1/s$ da sua importância.

Desta forma, para a rede apresentada na Figura 4.7 teríamos a seguinte modificação nas relações obtidas anteriormente:

$$\begin{array}{ll} x_1 & = \frac{1}{2}x_2 + \frac{1}{2}x_4 \\ x_2 & = \frac{1}{3}x_1 + \frac{1}{2}x_3 + \frac{1}{2}x_4 + \frac{1}{2}x_5 \\ x_3 & = \frac{1}{3}x_1 \\ x_4 & = \frac{1}{3}x_1 + \frac{1}{2}x_3 + \frac{1}{2}x_5 \\ x_5 & = \frac{1}{2}x_2 \end{array} \tag{4.24}$$

As equações anteriores podem reescrever-se na forma matricial $\mathbf{x} = M\mathbf{x}$, onde M é a matriz de Markov

$$M = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{x} = (x_1, x_2, x_3, x_4, x_5).$$

Para determinar o valor das relevâncias das cinco páginas web representadas no grafo da Figura 4.7 há que determinar o vector $\mathbf{x}$ que verifica as igualdades (4.24), ou seja, um vector de equilíbrio da cadeia de Markov definida por M. Temos assim de determinar um vector próprio associado ao valor próprio 1 de M que seja um vector de probabilidades. O vector $(15, 18, 5, 12, 9)$ é um vector próprio associado ao valor próprio 1, e portanto um vector de equilíbrio é $\frac{1}{59}(15, 18, 5, 12, 9) \approx (0.25, 0.31, 0.09, 0.20, 0.15)$. Logo, a página mais importante será a página 2, seguida de 1, 4, 5 e 3.

Evidentemente que no caso concreto da seriação realizada pelo Google a matriz M terá uma grandeza da ordem dos biliões, pelo que no tratamento computacional deste modelo assumem especial relevância os métodos numéricos para cálculo de valores e vectores próprios de matrizes de grandes dimensões.

Convém referir que o vector de equilíbrio pode não ser único, uma vez que, como se observa neste exemplo, a matriz que modela o funcionamento do Google não é necessariamente positiva. Ou seja, o subespaço próprio associado ao valor próprio 1 pode ter dimensão superior a 1. Este caso, é tratado no artigo [3] anteriormente referido, sendo aí apresentado um algoritmo que permite que o Google produza sempre uma listagem de sites ordenados por ordem decrescente de relevância.

Aconselha-se ao leitor interessado em aprofundar os detalhes da implementação do algoritmo PageRank e de outros motores de busca, a leitura de [7]. ⧫

4.4.2 Equações diferenciais ordinárias

O análogo contínuo de um sistema dinâmico discreto são os sistemas modelados por equações diferenciais, ou seja, por equações que envolvem uma função e as suas derivadas. Nesta secção abordaremos alguns aspectos da resolução de equações diferenciais ordinárias (EDO[12]) lineares, de primeira ordem. Trataremos equações do tipo $\mathbf{x}'(t) = A\mathbf{x}(t)+\mathbf{b}(t)$, onde A é uma matriz real (constante), $\mathbf{x}(t)$ e $\mathbf{b}(t)$ são vectores $n \times 1$ cujas componentes são funções reais de variável real, e $\mathbf{x}'(t)$ designa a derivada de $\mathbf{x}(t)$, isto é, a função com valor em vectores cujas componentes são as derivadas em ordem a $t \in \mathbb{R}$ das componentes de $\mathbf{x}(t)$. Eis a expressão matricial de $\mathbf{x}$ e $\mathbf{x}'$:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \qquad \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}.$$

Em particular, iremos determinar o conjunto de todas as soluções de um sistema de equações diferenciais (lineares) da forma $\mathbf{x}'(t) = A\mathbf{x}(t)$, em que A é uma matriz diagonalizável. O caso em que A não é diagonalizável é tratado no Capítulo 8. Para além do estudo da solução geral do sistema referido abordaremos outros tópicos relacionados, nomeadamente a exponencial de matrizes e a redução de uma equação diferencial de ordem n, homogénea e de coeficientes constantes, a um sistema do tipo $\mathbf{x}'(t) = A\mathbf{x}(t)$.

Comecemos por precisar a nomenclatura usada na classificação de equações diferenciais.

- <u>Equação diferencial</u>: uma equação que envolve uma função, por exemplo, $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$, e as suas derivadas.

- <u>Equação diferencial ordinária (EDO)</u>: uma equação que envolve uma função de uma única variável real t, e as suas derivadas.

[12] A abreviatura EDO em língua inglesa é ODE, de "ordinary differential equation".

- Ordem de uma EDO: ordem da derivada de maior ordem que aparece na equação.

Alguns exemplos de equações diferenciais do tipo $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$:

(a) $x' - 2x = 5t$ (onde $\mathbf{x}(t) = [x(t)], \mathbf{b} = [5t]$ e $A = [-2]$)

(b) $\begin{cases} x_1'(t) = 2x_1(t) - x_2(t) \\ x_2'(t) = 4x_1(t) + 10x_2(t) \end{cases} \iff \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 4 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \iff \mathbf{x}'(t) = A\mathbf{x}(t).$

Quando a matriz A é de ordem superior a 1, como no exemplo (b) anterior, a equação $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$ é um sistema de EDOs lineares de 1ª ordem. Neste texto usaremos indistintamente a designação EDO para uma equação diferencial ou para um sistema de EDOs.

A equação
$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$$
diz-se homogénea se $\mathbf{b}(t) = \mathbf{0}$. Em geral chama-se a
$$\mathbf{x}'(t) = A\mathbf{x}(t)$$
a equação *homogénea associada* à equação $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$.

Uma *solução* da equação diferencial $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$ é uma função diferenciável $\mathbf{u}$ que verifica a equação, ou seja, tal que $\mathbf{u}'(t) = A\mathbf{u}(t) + \mathbf{b}(t)$. Chama-se *solução geral* de uma equação diferencial ao conjunto de todas as soluções da equação.

Estamos interessados em obter o conjunto de todas as soluções da equação diferencial $\mathbf{x}'(t) = A\mathbf{x}(t)$, onde A é uma matriz (constante) do tipo $n \times n$. Para tal comecemos por abordar o caso mais simples em que $n = 1$, ou seja, de uma equação do tipo $x'(t) = kx(t)$, com k uma constante real.

Exemplo 4.14. Considere-se a equação $x'(t) = 3x(t)$. A solução geral desta equação é o conjunto de todas as funções reais x de variável real cuja derivada é o triplo da função. Esse conjunto solução é constituído por todas as funções da forma $x(t) = ce^{3t}$, onde c designa uma constante real arbitrária. Não existe outro tipo de funções que verifiquem a equação, como pode confirmar usando o Exercício 4.3 adiante. Assim, a solução geral da equação dada, é o subespaço

$$\left\{x(t) = ce^{3t} : \quad c \in \mathbb{R}\right\} = \text{Span}\{e^{3t}\} \tag{4.25}$$

do espaço linear $\mathcal{C}$ das funções reais de variável real, contínuas com derivada contínua. As operações de adição e multiplicação por escalares para as quais $\mathcal{C}$ é um espaço linear são as operações definidas na página 167.

O conjunto (4.25) é uma família de funções parametrizadas por $c \in \mathbb{R}$. Na Figura 4.8 encontram-se representados alguns elementos desta família.

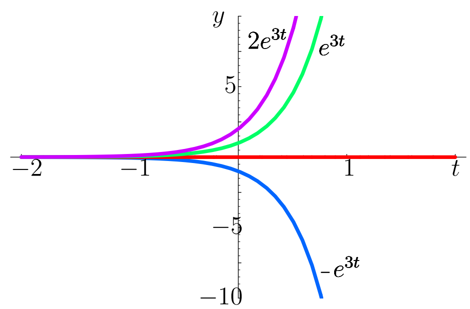

Figura 4.8: A solução geral da equação $x' = 3x$ é $x(t) = ce^{3t}$. A vermelho a função nula, correspondente a $c = 0$; a verde a solução correspondente a $c = x(0) = 1$; a azul a solução correspondente a $c = x(0) = -1$.

Obtivemos uma infinidade de soluções para a equação $x' = 3x$. Porém, se considerarmos o problema de saber quantas soluções da equação tomam um certo valor num dado ponto, a resposta é: uma única solução. Este tipo de problema é designado por *problema de valor inicial* (ou abreviadamente p.v.i.). Por exemplo, o problema de valor inicial

$$\begin{cases} x' = 3x \\ x(0) = 1, \end{cases}$$

apenas possui a solução $x(t) = e^{3t}$, já que impondo a condição inicial $x(0) = 1$ na solução geral da equação resulta

$$x(t) = ce^{3t} \Rightarrow x(0) = 1 \Rightarrow c = 1.$$

♦

Exercício 4.3. Mostre que $u(t)$ é uma solução de $x' = kx$ (com k uma constante real) se e só se o produto $u(t)e^{-kt}$ é uma constante. ▲

Sistemas de EDOs lineares, de primeira ordem, homogéneos

Vimos que no caso de uma EDO homogénea do tipo $x' = kx$ (com $x : \mathbb{R} \to \mathbb{R}$) o conjunto solução geral é o espaço linear gerado por e^{kt}. A solução geral de um sistema de EDOs homogéneo do tipo $\mathbf{x}' = A\mathbf{x}$ é igualmente um subespaço do espaço linear das funções $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$, contínuas com derivada contínua, com as operações usuais de adição de funções e multiplicação de uma função por um escalar. De facto, é fácil mostrar que:

- A soma de duas soluções de $\mathbf{x}' = A\mathbf{x}$ ainda é uma solução desta equação.
- O produto de uma solução de $\mathbf{x}' = A\mathbf{x}$ por um escalar ainda é uma solução desta equação.

Refira-se que a solução (constante) nula $\mathbf{x}(t) = (0, 0, \ldots, 0)$ é sempre solução do sistema $\mathbf{x}' = A\mathbf{x}$, como não podia deixar de ser uma vez que a solução geral é um subespaço.

Como veremos, a solução geral de $\mathbf{x}' = A\mathbf{x}$, em que A é uma matriz de ordem n, é um (sub)espaço linear de dimensão n. Assim, a determinação da solução geral passa pela obtenção de uma base deste espaço. Uma base para a solução geral recebe a designação de *conjunto fundamental de soluções,* atendendo a que qualquer solução de $\mathbf{x}' = A\mathbf{x}$ se obtém como combinação linear dos elementos do referido conjunto.

Para obter uma base para a solução geral, temos de saber determinar soluções linearmente independentes da equação $\mathbf{x}' = A\mathbf{x}$. A Proposição 4.12 adiante, consequência imediata do *Teorema de existência e unicidade de soluções* de problemas de valor inicial, fornece-nos um teste para a independência linear de soluções da equação $\mathbf{x}' = A\mathbf{x}$. Enunciamos em seguida uma versão do teorema referido cuja prova pode ser encontrada em qualquer livro dedicado ao estudo de equações diferenciais como, por exemplo, Braun [2].

Teorema 4.4. Existência e unicidade de soluções

Seja A uma matriz real de ordem n, $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ uma função contínua e $\mathbf{x}'$ a sua derivada.

Existe uma e uma só solução do problema de valor inicial

$$\mathbf{x}' = A\mathbf{x} \quad \text{e} \quad \mathbf{x}(t_0) = (x_1^0, x_2^0, \ldots, x_n^0).$$

Além disso, esta solução existe para todo $t \in \mathbb{R}$.

Deixamos como exercício a demonstração da unicidade de soluções de um problema de valor inicial.

O teorema anterior permite-nos provar que o conjunto solução geral da equação $\mathbf{x}' = A\mathbf{x}$ é um subespaço de dimensão n, onde n é a ordem da matriz A. No exercício a seguir sugere-se uma demonstração.

Exercício 4.4. Mostre que o conjunto solução geral da equação $\mathbf{x}' = A\mathbf{x}$ é um subespaço de dimensão n, onde n é a ordem da matriz A.

Sugestão: Considere o conjunto $B = \{\phi_1(t), \dots, \phi_n(t)\}$, onde $\phi_j(t)$ é a solução do problema de valor inicial $\mathbf{x}' = A\mathbf{x}$ e $\mathbf{x}(0) = \mathbf{e}_j$, com $\mathbf{e}_j$ o vector da base canónica de $\mathbb{R}^n$, $\mathbf{e}_j = (0, 0, \dots, 1, \dots, 0)$. Mostre que B é linearmente independente e que qualquer solução da equação $\mathbf{x}' = A\mathbf{x}$ é uma combinação linear dos elementos de B.

▲

A proposição que enunciamos a seguir fornece um teste de grande utilidade prática na verificação da independência linear de soluções de $\mathbf{x}' = A\mathbf{x}$. Nomeadamente, essa proposição garante que as soluções $\mathbf{u}_1(t), \dots, \mathbf{u}_k(t)$ de $\mathbf{x}' = A\mathbf{x}$ são linearmente independentes se e só se os vectores de $\mathbb{R}^n$, $\mathbf{u}_1(t_0), \dots, \mathbf{u}_k(t_0)$ são linearmente independentes, onde t_0 é um valor de t que podemos escolher da forma mais conveniente.

Proposição 4.12. Seja $B = \{\mathbf{u}_1(t), \dots, \mathbf{u}_k(t)\}$ um conjunto de soluções de $\mathbf{x}' = A\mathbf{x}$ e t_0 um valor fixo de t. O conjunto B é linearmente independente se e só se $\{\mathbf{u}_1(t_0), \dots, \mathbf{u}_k(t_0)\}$ é linearmente independente em $\mathbb{R}$.

Demonstração. Suponha-se que B é linearmente dependente. Ou seja, existem constantes $c_1, \dots, c_k$ não todas nulas, tais que $c_1\mathbf{u}_1(t) + \cdots + c_k\mathbf{u}_k(t) = \mathbf{0}$. Calculando esta igualdade em $t = t_0$, tem-se

$$c_1\mathbf{u}_1(t_0) + \cdots + c_k\mathbf{u}_k(t_0) = \mathbf{0},$$

com pelo menos um dos c_i's é não nulo. Logo, $\{\mathbf{u}_1(t_0), \dots, \mathbf{u}_k(t_0)\}$ é linearmente dependente.

Para a implicação recíproca, suponha-se que existem constantes $c_1, \dots, c_k$, não todas nulas, tais que $c_1\mathbf{u}_1(t_0) + \cdots + c_k\mathbf{u}_k(t_0) = \mathbf{0}$. Considere-se a função

$$\phi(t) = c_1\mathbf{u}_1(t) + \cdots + c_k\mathbf{u}_k(t).$$

A função ϕ é uma solução da equação diferencial $\mathbf{x}' = A\mathbf{x}$ já que, ϕ é uma combinação linear de soluções. Além disso, $\phi(t_0) = \mathbf{0}$. Pelo Teorema de existência e unicidade de soluções de um problema de valor inicial (Teorema 4.4), o problema de valor inicial $\mathbf{x}' = A\mathbf{x}$ com $\mathbf{x}(t_0) = \mathbf{0}$ só admite a solução nula. Logo, $\phi(t) = \mathbf{0}$ para todo o t, e portanto $c_1\mathbf{u}_1(t) + \cdots + c_k\mathbf{u}_k(t) = \mathbf{0}$ com pelo menos um dos c_i's não nulo. Ou seja, B é linearmente dependente. $\square$

Pretendemos agora determinar a solução geral de $\mathbf{x}' = A\mathbf{x}$ em que $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$ é uma função contínua e A uma matriz real $n \times n$. Para tal, comecemos por abordar o caso em que A é a matriz diagonal $A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. O sistema $\mathbf{x}' = A\mathbf{x}$ reduz-se a n equações do tipo já estudado. Nomeadamente,

$$\mathbf{x}' = A\mathbf{x} \iff \begin{cases} x_1' = \lambda_1 x_1 \\ x_2' = \lambda_2 x_2 \\ \vdots \\ x_n' = \lambda_n x_n. \end{cases}$$

Como vimos anteriormente, no Exercício 4.3, a solução geral de cada equação $x_i' = \lambda_i x_i$ é dada por $x_i(t) = c_i e^{\lambda_i t}$ onde c_i é uma constante real arbitrária. Logo, a solução geral deste sistema é

$$\begin{aligned}\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t}. \end{bmatrix} &= c_1 e^{\lambda_1 t}\mathbf{e}_1 + c_2 e^{\lambda_2 t}\mathbf{e}_2 + \cdots + c_n e^{\lambda_n t}\mathbf{e}_n \\ &= \begin{bmatrix} | & | & & | \\ e^{\lambda_1 t}\mathbf{e}_1 & e^{\lambda_2 t}\mathbf{e}_2 & \cdots & e^{\lambda_n t}\mathbf{e}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = X(t)\mathbf{c},\end{aligned}$$

onde $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ são os vectores da base canónica de $\mathbb{R}^n$. Assim, um conjunto gerador da solução geral é $\left\{e^{\lambda_1 t}\mathbf{e}_1, e^{\lambda_2 t}\mathbf{e}_2, \ldots, e^{\lambda_n t}\mathbf{e}_n\right\}$. Este conjunto é uma base para a solução geral já que, avaliando os elementos deste conjunto em $t_0 = 0$ se obtém a base canónica de $\mathbb{R}^n$, e portanto a Proposição 4.12 garante que $\left\{e^{\lambda_1 t}\mathbf{e}_1, \ldots, e^{\lambda_n t}\mathbf{e}_n\right\}$ é linearmente independente.

Podemos tirar as conclusões que se seguem relativas à solução geral do sistema de EDOs $\mathbf{x}' = A\mathbf{x}$ em que A é uma matriz diagonal do tipo $n \times n$.

- Existem n soluções linearmente independentes da forma $e^{\lambda_i t}\mathbf{e}_i$ para $\mathbf{x}' = A\mathbf{x}$.

- A solução geral do sistema é da forma $\mathbf{x}(t) = X(t)\mathbf{c}$, onde X é uma matriz que tem para colunas os n vectores $e^{\lambda_i t}\mathbf{e}_i$, e $\mathbf{c}$ é um vector coluna constante arbitrário. Como se verifica facilmente, os pares $(\lambda_i, \mathbf{e}_i)$ são pares próprios de A.

- A matriz $X(t)$ é invertível, visto que as suas colunas são linearmente independentes. Consequentemente $\det X(t) \neq 0$ para todo o t.

- A derivada da matriz $X(t)$ satisfaz a equação $\mathbf{X}' = A\mathbf{X}$, onde a derivada de uma matriz é a matriz se que obtém derivando entrada a entrada. É fácil verificar que de facto $X'(t) = AX(t)$.

As conclusões anteriores, válidas para uma matriz diagonal, permanecem válidas para qualquer matriz quadrada, como veremos a seguir.

Definição 4.9. Seja A uma matriz do tipo $n \times n$. Chama-se matriz *solução fundamental* do sistema $\mathbf{x}' = A\mathbf{x}$ a qualquer matriz $X(t)$ cujas colunas sejam n soluções linearmente independentes de $\mathbf{x}' = A\mathbf{x}$.

Visto que a solução geral de um sistema homogéneo de n equações diferenciais lineares de 1ª ordem é um espaço linear de dimensão n, as colunas de uma matriz solução fundamental X formam uma base para este espaço. Assim, a solução geral de $\mathbf{x}' = A\mathbf{x}$ é o conjunto das combinações lineares das colunas de X.

A solução geral de $\mathbf{x}' = A\mathbf{x}$ é dada por

$$\mathbf{x}(t) = X(t)\mathbf{c},$$

onde $X(t)$ é uma matriz solução fundamental do sistema e $\mathbf{c}$ é um vector coluna constante arbitrário.

Note-se que, por definição de produto de uma matriz por um vector, a expressão $X(t)\mathbf{c}$ designa precisamente uma combinação linear (arbitrária) das colunas de X.

Exercício 4.5. Mostre que $X(t)$ é uma matriz solução fundamental do sistema $\mathbf{x}' = A\mathbf{x}$ se e só se $X'(t) = AX(t)$ e $\det(X(0)) \neq 0$.

Relembre que a notação $X'(t)$ designa a matriz cujas entradas são as derivadas das entradas de $X(t)$. ▲

Como determinar n soluções linearmente independentes para $\mathbf{x}' = A\mathbf{x}$? A proposição seguinte responde (parcialmente) a esta questão.

Proposição 4.13. Seja A uma matriz real $n \times n$, e $\mathbf{u}$ um vector constante não nulo.

(1) A função $e^{\lambda t}\mathbf{u}$ é solução de $\mathbf{x}' = A\mathbf{x}$ se e só se λ é valor próprio de A com vector próprio associado $\mathbf{u}$.

(2) Se $\mathbf{x}(t) = \mathbf{x}_1(t) + i\mathbf{x}_2(t)$ é uma solução (complexa) de $\mathbf{x}' = A\mathbf{x}$, então $\operatorname{Re}\mathbf{x}(t) = \mathbf{x}_1(t)$ e $\operatorname{Im}\mathbf{x}(t) = \mathbf{x}_2(t)$ são duas soluções reais de $\mathbf{x}' = A\mathbf{x}$.

(3) Se $\lambda = a + ib$, com $b \neq 0$, é um valor próprio (complexo) de A e $\mathbf{u}$ um vector próprio associado, então $\operatorname{Re}(e^{\lambda t}\mathbf{u})$ e $\operatorname{Im}(e^{\lambda t}\mathbf{u})$ são duas soluções reais linearmente independentes de $\mathbf{x}' = A\mathbf{x}$. Ou seja,

$$e^{at}\left(\cos(bt)\operatorname{Re}\mathbf{u} - \operatorname{sen}(bt)\operatorname{Im}\mathbf{u}\right) \quad \text{e} \quad e^{at}\left(\operatorname{sen}(bt)\operatorname{Re}\mathbf{u} + \cos(bt)\operatorname{Im}\mathbf{u}\right),$$

são duas soluções reais linearmente independentes de $\mathbf{x}' = A\mathbf{x}$.

Demonstração. (1) Como $(e^{\lambda t}\mathbf{u})' = \lambda e^{\lambda t}\mathbf{u}$, tem-se que $\mathbf{z}' = A\mathbf{z}$ para $\mathbf{z}(t) = e^{\lambda t}\mathbf{u}$ se e só se

$$\lambda e^{\lambda t}\mathbf{u} = A(e^{\lambda t}\mathbf{u}) \iff \lambda\mathbf{u} = A\mathbf{u},$$

onde na equivalência anterior se aplicou o facto de $e^{\lambda t}$ nunca se anular.

Ou seja, $e^{\lambda t}\mathbf{u}$ é solução de $\mathbf{x}' = A\mathbf{x}$ se e só se $\mathbf{u}$ é um vector próprio de A associado ao valor próprio λ.

(2) Dizer que $\mathbf{x}(t) = \mathbf{x}_1(t) + i\mathbf{x}_2(t)$ é uma solução complexa de $\mathbf{x}' = A\mathbf{x}$ é equivalente a

$$\begin{aligned}
\mathbf{x}'(t) = A\mathbf{x} &\iff \mathbf{x}_1'(t) + i\mathbf{x}_2'(t) = A\left(\mathbf{x}_1(t) + i\mathbf{x}_2(t)\right) \\
&\iff \mathbf{x}_1'(t) + i\mathbf{x}_2'(t) = A\mathbf{x}_1(t) + iA\mathbf{x}_2(t) \\
&\iff \mathbf{x}_1'(t) = A\mathbf{x}_1(t) \quad \text{e} \quad \mathbf{x}_2'(t) = A\mathbf{x}_2(t).
\end{aligned}$$

Ou seja, $\operatorname{Re}\mathbf{x}(t) = \mathbf{x}_1(t)$ e $\operatorname{Im}\mathbf{x}(t) = \mathbf{x}_2(t)$ são duas soluções reais de $\mathbf{x}' = A\mathbf{x}$.

(3) Do item (1) tem-se que $e^{\lambda t}\mathbf{u}$ é uma solução (complexa) da equação diferencial, e pelo item (2) resulta que $\operatorname{Re}(e^{\lambda t}\mathbf{u})$ e $\operatorname{Im}(e^{\lambda t}\mathbf{u})$ são duas soluções reais da respectiva equação diferencial. A independência linear destas soluções segue do Lema 4.1 (na página 198) e da Proposição 4.12 considerando $t_0 = 0$.

Calculemos a parte real e imaginária de $e^{\lambda t}\mathbf{u}$. Para tal, relembremos que $e^{ibt} = \cos(bt) + i\,\mathrm{sen}(bt)$ (ver (A.1) no Apêndice A).

$$\begin{aligned} e^{(a+ib)t}(\operatorname{Re}\mathbf{u} + i\operatorname{Im}\mathbf{u}) &= e^{at}e^{ibt}(\operatorname{Re}\mathbf{u} + i\operatorname{Im}\mathbf{u}) \\ &= e^{at}(\cos(bt) + i\,\mathrm{sen}(bt))(\operatorname{Re}\mathbf{u} + i\operatorname{Im}\mathbf{u}) \\ &= e^{at}\left[(\cos(bt)\operatorname{Re}\mathbf{u} - \mathrm{sen}(bt)\operatorname{Im}\mathbf{u})\right. \\ &\quad \left.+ i\,(\mathrm{sen}(bt)\operatorname{Re}\mathbf{u} + \cos(bt)\operatorname{Im}\mathbf{u})\right]. \end{aligned}$$

Assim, $\operatorname{Re}(e^{\lambda t}\mathbf{u})$ e $\operatorname{Im}(e^{\lambda t}\mathbf{u})$ são dadas pelas expressões no enunciado.

□

Podemos agora determinar a solução geral de $\mathbf{x}' = A\mathbf{x}$ para matrizes diagonalizáveis A. O caso em que A não é diagonalizável é tratado no Capítulo 8.

Recorde-se que uma matriz A, de ordem n, é diagonalizável se e só se admite n vectores próprios linearmente independentes. Neste caso, pela Proposição 4.13, o sistema $\mathbf{x}' = A\mathbf{x}$ tem n soluções reais linearmente independentes da seguinte forma: (i) $e^{\lambda t}\mathbf{u}$ com λ um valor próprio real de A e $\mathbf{u}$ um vector próprio associado; (ii) para cada par de valores próprios complexos conjugados λ e $\overline{\lambda}$, existem duas soluções reais (linearmente independentes) da forma $\operatorname{Re}(e^{\lambda t}\mathbf{u})$ e $\operatorname{Im}(e^{\lambda t}\mathbf{u})$, onde $\mathbf{u}$ é um vector próprio associado a λ. A solução geral de $\mathbf{x}' = A\mathbf{x}$ é uma combinação linear destas n soluções linearmente independentes.

Apresentamos a seguir um exemplo da determinação da solução geral de um sistema $\mathbf{x}' = A\mathbf{x}$ em que A possui valores próprios reais e complexos.

Exemplo 4.15. Determine-se a solução geral do sistema $\mathbf{x}' = A\mathbf{x}$, em que A é a matriz do Exemplo 4.9 (pág. 194). O sistema correspondente é

$$\begin{cases} x_1' &= 2x_1 + 5x_2 \\ x_2' &= -x_1 \\ x_3' &= 5x_3. \end{cases}$$

A matriz A possui valores próprios 5 e $1 \pm 2i$ e os espaços próprios são:

$$E(5) = \operatorname{Span}\{(0,0,1)\}, \quad E(1+2i) = \operatorname{Span}\{(1+2i,-1,0)\}$$
$$E(1-2i) = \operatorname{Span}\{(1-2i,-1,0)\}.$$

Usando a Proposição 4.13, são soluções (reais) linearmente independentes do sistema:

$$e^{5t}\begin{bmatrix}0\\0\\1\end{bmatrix}, \quad \operatorname{Re}\left(e^{(1+2i)t}\begin{bmatrix}1+2i\\-1\\0\end{bmatrix}\right) \quad \text{e} \quad \operatorname{Im}\left(e^{(1+2i)t}\begin{bmatrix}1+2i\\-1\\0\end{bmatrix}\right).$$

Como $e^{(1+2i)t} = e^t e^{2it} = e^t(\cos(2t) + i\,\mathrm{sen}(2t))$, resulta

$$\mathrm{Re}\left(e^{(1+2i)t}\begin{bmatrix}1+2i\\-1\\0\end{bmatrix}\right) = e^t\left(\cos(2t)\begin{bmatrix}1\\-1\\0\end{bmatrix} - \mathrm{sen}(2t)\begin{bmatrix}2\\0\\0\end{bmatrix}\right)$$
$$\mathrm{Im}\left(e^{(1+2i)t}\begin{bmatrix}1+2i\\-1\\0\end{bmatrix}\right) = e^t\left(\mathrm{sen}(2t)\begin{bmatrix}1\\-1\\0\end{bmatrix} + \cos(2t)\begin{bmatrix}2\\0\\0\end{bmatrix}\right).$$

Conclui-se que a solução geral do sistema é o conjunto das funções

$$\mathbf{x}(t) = c_1 e^{5t}\begin{bmatrix}0\\0\\1\end{bmatrix} + c_2 e^t\left(\cos(2t)\begin{bmatrix}1\\-1\\0\end{bmatrix} - \mathrm{sen}(2t)\begin{bmatrix}2\\0\\0\end{bmatrix}\right) +$$
$$+ c_3 e^t\left(\mathrm{sen}(2t)\begin{bmatrix}1\\-1\\0\end{bmatrix} + \cos(2t)\begin{bmatrix}2\\0\\0\end{bmatrix}\right),$$

com c_1, c_2 e c_3 constantes reais arbitrárias.

Note-se que uma matriz solução fundamental para este sistema é

$$X(t) = \begin{bmatrix}0 & e^t(\cos(2t) - 2\,\mathrm{sen}(2t)) & e^t(\mathrm{sen}(2t) + 2\cos(2t))\\ 0 & -e^t\cos(2t) & -e^t\,\mathrm{sen}(2t)\\ e^{5t} & 0 & 0\end{bmatrix}.$$

♦

Exercício 4.6. Para a matriz A do Exemplo 4.8 (pág. 194), mostre que uma matriz solução fundamental para o sistema $\mathbf{x}' = A\mathbf{x}$ é

$$X(t) = \begin{bmatrix}-e^{3t} & 0 & e^{5t}\\ 0 & e^{3t} & 2e^{5t}\\ e^{3t} & 0 & e^{5t}\end{bmatrix}.$$

Calcule ainda $X(t)X(0)^{-1}$.

▲

Da definição de matriz solução fundamental, vê-se facilmente que não existe uma única matriz solução fundamental. Não é difícil mostrar que quaisquer duas matrizes $X(t)$, $Y(t)$, solução fundamental de um sistema, satisfazem uma relação do tipo $Y(t) = X(t)K$ onde K é uma matriz (constante) invertível.

Exercício 4.7. Mostre que se $X(t)$ e $Y(t)$ são duas quaisquer matrizes solução fundamental do sistema $\mathbf{x}' = A\mathbf{x}$, então existe uma matriz real invertível K tal que $Y(t) = X(t)K$.

Sugestão: Escrever as colunas de Y como combinação linear das colunas de X e usar os factos $\det Y(0) \neq 0$ e $\det X(0) \neq 0$ para mostrar que K é invertível.
▲

A Proposição 4.13 diz-nos como determinar uma base do conjunto das soluções de $\mathbf{x}' = A\mathbf{x}$ no caso em que A é diagonalizável. Conhecida a solução geral da equação homogénea $\mathbf{x}' = A\mathbf{x}$ podemos determinar a solução geral da equação não homogénea $\mathbf{x}' = A\mathbf{x} + \mathbf{b}(t)$ desde que se conheça uma solução desta equação. De facto, à semelhança do que acontece para sistemas de equações lineares (não diferenciais), tem-se:

A solução geral da equação $\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$ é a soma de uma solução particular $\mathbf{x}_p$ da equação com a solução geral $\mathbf{x}_h$ da equação homogénea associada, $\mathbf{x}'(t) = A\mathbf{x}(t)$. Ou seja, a solução geral é

$$\{\mathbf{x}(t) : \mathbf{x}(t) = \mathbf{x}_p(t) + \mathbf{x}_h(t)\} .$$

Deixamos como exercício a demonstração deste resultado que se reduz a um mero decalque da demonstração apresentada para o Teorema 3.7 da página 153.

No exemplo que se segue aplicamos o resultado anterior para determinar a solução geral de uma equação diferencial não homogénea.

Exemplo 4.16. Considere-se a seguinte EDO não homogénea $x' - 3x = 2t - 3t^2$.

A função $v(t) = t^2$ é uma solução desta equação como facilmente se verifica. A equação homogénea associada, $x' - 3x = 0$, tem solução geral $x(t) = ce^{3t}$, com $c \in \mathbb{R}$. Logo, a solução geral da equação não homogénea é: $\{ce^{3t} + t^2 : c \in \mathbb{R}\}$.
♦

Exponencial de matrizes

Como se viu, dada uma equação diferencial, $\mathbf{x}' = A\mathbf{x}$, existem várias matrizes solução fundamental dessa equação. A exponencial e^{At} vai ser definida como sendo uma matriz solução fundamental particular do sistema $\mathbf{x}' = A\mathbf{x}$. É possível definir a matriz exponencial e^{At} como uma série de potências da matriz At, semelhante à série de potências que define a função exponencial real. No entanto essa via sai do âmbito deste texto, o leitor interessado poderá consultar, por exemplo, Braun [2].

Relembremos (Exercício 4.5) que uma matriz solução fundamental do sistema $\mathbf{x}' = A\mathbf{x}$ é uma matriz que verifica a equação diferencial para matrizes $X'(t) = AX(t)$.

Definição 4.10. Seja A uma matriz quadrada. A exponencial e^{At} é a matriz solução fundamental do sistema $\mathbf{x}' = A\mathbf{x}$ cujo valor em $t = 0$ é a matriz identidade. Ou seja, e^{At} é a solução do problema de valor inicial

$$\begin{cases} X' = AX \\ X(0) = I, \end{cases}$$

onde I designa a matriz identidade e X uma matriz da mesma ordem de A.

Nota 24. *O problema de valor inicial na definição anterior é um problema de valor inicial para matrizes. Este problema pode ser visto como um conjunto de n problemas de valor inicial para o sistema $\mathbf{x}' = A\mathbf{x}$ (correspondentes a problemas de valor inicial para as colunas de X).*

Tendo em conta a definição anterior e o resultado do Exercício 4.7 (pág. 228), podemos enunciar a proposição seguinte.

Proposição 4.14. Seja A uma matriz real $n \times n$ e $X(t)$ uma qualquer matriz solução fundamental do sistema $\mathbf{x}' = A\mathbf{x}$. A exponencial e^{At} é dada por

$$e^{At} = X(t)X(0)^{-1}.$$

Demonstração. Se e^{At} e $X(t)$ são duas matrizes solução fundamental do sistema $\mathbf{x}' = A\mathbf{x}$, do Exercício 4.7, segue que

$$e^{At} = X(t)K,$$

onde K é uma matriz (constante) invertível. Como por definição $e^{A0} = I$, resulta

$$e^{A0} = I = X(0)K \iff K = X(0)^{-1}.$$

Por conseguinte, $e^{At} = X(t)X(0)^{-1}$. □

Nota 25. *Uma vez que e^{At} é solução de um problema de valor inicial, segue da unicidade de soluções deste tipo de problemas que a matriz e^{At} é única. Além disso, como e^{At} é uma matriz solução fundamental de $\mathbf{x}' = A\mathbf{x}$, é satisfeita a igualdade $\frac{d}{dt}(e^{At}) = Ae^{At}$ (ver Exercício 4.5).*

Exercício 4.8. a) Mostre que $e^{(A+B)t} = e^{At}e^{Bt}$ se e só se A e B comutam.

b) Use o resultado anterior para mostrar que a matriz inversa de e^{At} é e^{-At}.

Sugestão: Para a alínea a) mostre os resultados:

- Se $AB = BA$, então $X(t) = Be^{At}$ e $Y(t) = e^{At}B$ são matrizes solução fundamental do mesmo p.v.i.. Use a unicidade de soluções de problemas de valor inicial para mostrar que $X(t) = Y(t)$.
- Mostre que $e^{At}e^{Bt}$ e $e^{t(A+B)}$ resolvem o mesmo problema de valor inicial.
- Finalmente, mostre que se $X(t) = Y(t)$, o que necessariamente implica $X'(t) = Y'(t)$, se tem $AB = BA$.

▲

Exemplo 4.17. Calculemos e^{At} para $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

No Exemplo 4.4 (pág. 182) calculámos os valores próprios e os espaços próprios desta matriz. Os valores próprios de A são $\lambda = \pm i$ e $E(-i) = \text{Span}\{(1, i)\}$. Logo, pelo item (2) da Proposição 4.13, são soluções (reais) linearmente independentes do sistema $\mathbf{x}' = A\mathbf{x}$, os vectores

$$\text{Re}\left(e^{-it}\begin{bmatrix} 1 \\ i \end{bmatrix}\right) \quad \text{e} \quad \text{Im}\left(e^{-it}\begin{bmatrix} 1 \\ i \end{bmatrix}\right),$$

ou, equivalentemente,

$$\cos(-t)\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \text{sen}(-t)\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos t \\ \text{sen}\, t \end{bmatrix}, \qquad \text{sen}(-t)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \cos(-t)\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\text{sen}\, t \\ \cos t \end{bmatrix}.$$

Nas igualdades anteriores aplicámos $\cos(-t) = \cos t$ e $\text{sen}(-t) = -\text{sen}\, t$.

Uma matriz solução fundamental para o sistema $\mathbf{x}' = A\mathbf{x}$ é uma matriz cujas colunas são soluções linearmente independentes do sistema. Neste caso, uma matriz solução fundamental $X(t)$ é, por exemplo,

$$X(t) = \begin{bmatrix} \cos t & -\text{sen}\, t \\ \text{sen}\, t & \cos t \end{bmatrix}.$$

Como $X(0)$ é a matriz identidade, temos $e^{At} = X(t)X(0)^{-1} = X(t)$. Verifique ainda que se tivéssemos considerado para $X(t)$ a matriz que se obtém trocando as colunas da matriz acima a expressão $X(t)X(0)^{-1}$ produziria o mesmo resultado (e^{At} é única). ⧫

Exercício 4.9. Considere o sistema de equações diferenciais $\mathbf{x}' = A\mathbf{x}$, onde A é uma matriz diagonalizável, isto é, $A = PDP^{-1}$ com D diagonal.

a) Mostre que o sistema dado é equivalente ao sistema $\mathbf{y}' = D\mathbf{y}$, onde $\mathbf{y} = P^{-1}\mathbf{x}$.

b) Use a alínea anterior, para mostrar a igualdade $e^{At} = Pe^{Dt}P^{-1}$.

▲

Equações de ordem n e redução de ordem

Para finalizar esta secção vamos verificar que podemos obter a solução geral de equações diferenciais ordinárias lineares homogéneas, de coeficientes constantes e de ordem superior à primeira, resolvendo um sistema da forma $\mathbf{x}' = A\mathbf{x}$.

Uma *equação diferencial ordinária, linear, homogénea, de coeficientes constantes, e de ordem* n, é uma equação diferencial da forma

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0, \tag{4.26}$$

onde (os coeficientes) $a_i \in \mathbb{R}$ para $i = 1, \ldots, n$, a função $y : \mathbb{R} \to \mathbb{R}$ é real de variável real, e $y^{(k)}$ designa a derivada de ordem k de y.

Introduzindo novas variáveis $x_1 = y, x_2 = y', \ldots, x_n = y^{(n-1)}$, a equação (4.26) é equivalente a um sistema da forma $\mathbf{x}' = A\mathbf{x}$. Nomeadamente

$$\begin{array}{l} x_1 = y \\ x_2 = y' \\ \vdots \\ x_n = y^{(n-1)} \end{array} \quad \Longrightarrow \begin{cases} x_1' = x_2 \\ x_2' = x_3 \\ \vdots \\ x_n' = -a_0x_1 - a_1x_2 - \cdots - a_{n-1}x_n. \end{cases} \tag{4.27}$$

A matriz A do sistema obtido é conhecida por matriz *companheira* da equação (4.26) e tem a forma

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots\cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots\cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots\cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1}. \end{bmatrix}.$$

Exercício 4.10. Mostre que o polinómio característico da matriz companheira da equação (4.26) é $p(\lambda) = (-1)^n[\lambda^n + a_{n-1}\lambda^{(n-1)} + \cdots + a_1\lambda + a_0]$.
Sugestão: Calcule $\det(A - \lambda I)^T$ usando o método de eliminação de Gauss. ▲

Atendendo ao último exercício, o polinómio característico da matriz companheira de uma equação de ordem n pode ser obtido directamente a partir da equação (4.26).

Após reduzir a equação (4.26) a um sistema de EDOs homogéneo de primeira ordem, $\mathbf{x}' = A\mathbf{x}$, a solução geral da equação (4.26) obtém-se da solução geral do sistema (4.27) considerando apenas a primeira componente da solução $\mathbf{x}$ do sistema, visto que fizemos $y = x_1$.

Exemplo 4.18. Determinemos a solução geral da equação diferencial de terceira ordem, $y''' - 6y'' + 11y' - 6y = 0$.

Fazendo $y = x_1, y' = x_2$ e $y'' = x_3$, a equação diferencial reduz-se ao sistema

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -11 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Pelo Exercício 4.10, a equação característica da matriz companheira é $\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$. É fácil verificar que esta equação tem raízes $\lambda = 3, \lambda = 2$ e $\lambda = 1$. Como os valores próprios são distintos, a matriz companheira é diagonalizável.

Uma base de vectores próprios é, por exemplo, $\{(1, 3, 9), (1, 2, 4), (1, 1, 1)\}$. Esta base é constituída por vectores próprios associados respectivamente a $3, 2$ e 1. A solução geral do sistema é assim

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = c_1 e^{3t} \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + c_3 e^{t} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Uma vez que fizemos $y = x_1$, a solução geral da equação diferencial de terceira ordem é

$$y(t) = x_1(t) = c_1 e^{3t} + c_2 e^{2t} + c_3 e^{t},$$

onde c_1, c_2 e c_3 são constantes reais arbitrárias. ♦