

# Métodos Numéricos

## Sistemas de Equações Lineares

Carlos J. Luz  
Departamento de Matemática  
Escola Superior de Tecnologia de Setúbal

Ano Lectivo 2003/2004

Os sistemas de equações lineares constituem um instrumento indispensável ao estudo de numerosos problemas oriundos das mais variadas áreas do conhecimento. Se desde há muito encontram incontáveis aplicações em Ciências Exactas e Engenharia, actualmente as Ciências Humanas fazem um uso extensivo dos sistemas de equações lineares através, por exemplo, dos tratamentos estatísticos.

Desde o Ensino Secundário que os sistemas de equações lineares são estudados, sendo aí aprendidos alguns métodos de resolução. Posteriormente, na disciplina de Álgebra Linear, é introduzido o método de eliminação de Gauss e utilizam-se os conceitos de dependência e independência lineares para proceder à classificação dos sistemas de equações lineares quanto à existência de solução.

Numa disciplina de Métodos Numéricos tem-se em vista a descrição de métodos eficientes para a resolução de sistemas de equações de qualquer dimensão. Estes métodos destinam-se a ser implementados num computador, pelo que é necessário atender a algumas técnicas adequadas ao controlo de erros introduzidos pela realização de operações em ponto flutuante.

Os métodos de resolução de sistemas de equações lineares podem ser classificados em **métodos directos** e **métodos iterativos**. Nos métodos directos a solução é obtida após a realização de um número finito de operações aritméticas. Nos métodos iterativos, a partir de uma aproximação inicial, vão-se calculando sucessivamente novas aproximações da solução exacta do sistema. Cada iteração consiste em obter uma nova aproximação (ou iterada) a partir das anteriores, baseando-se a efectividade dum método iterativo na garantia da convergência das sucessivas iteradas para a solução exacta do sistema.

No que se segue, preocupar-nos-emos somente com a resolução de sistemas de equações lineares em que a matriz dos coeficientes é quadrada e não singular<sup>1</sup>. Tratam-se de sistemas com uma única solução, constituídos por  $n$  equações lineares e  $n$  **incógnitas**  $x_1, x_2, \dots, x_n$ , os quais podem ser representados na forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases}, \quad (1)$$

onde  $a_{ij}$  e  $b_i$  ( $i, j = 1, \dots, n$ ) são números reais conhecidos. O elemento  $a_{ij}$  é o **coeficiente da incógnita**  $x_j$  na equação de ordem  $i$  e  $b_i$  é o **termo independente** da mesma equação.

Matricialmente, o sistema anterior pode representar-se por

$$AX = B,$$

em que

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ e } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

são, respectivamente, a **matriz dos coeficientes**, a **matriz das incógnitas** e a **matriz dos termos independentes**. Utilizaremos as igualdades  $A = [a_{ij}]$  e  $X = [x_i]$  para indicar os

---

<sup>1</sup>Recorde-se que uma matriz é *não singular* se é invertível, o que equivale a dizer que o seu determinante é diferente de zero. Caso contrário, a matriz diz-se *singular*.

elementos genéricos duma matriz quadrada  $A$  de ordem  $n$  e duma matriz coluna  $X$ , respectivamente. Faremos também uso da notação vectorial  $\vec{x} = (x_1, x_2, \dots, x_n)$  para representar uma matriz coluna  $X = [x_i]$  com  $n$  elementos. Para efeitos de representação, assumiremos pois que  $X = \vec{x}$ .

Veremos seguidamente alguns métodos directos, nomeadamente o método de eliminação de Gauss e os métodos compactos de Doolittle e Cholesky. Depois, estudaremos os métodos iterativos de Jacobi e de Gauss-Seidel.

## 1 Método de Eliminação de Gauss

Como dissemos, o **método de eliminação de Gauss** é habitualmente estudado na disciplina introdutória de Álgebra Linear. Baseia-se na transformação do sistema dado  $AX = B$  num sistema equivalente  $UX = Y$ , em que  $U$  é uma matriz triangular superior. Sendo  $[A|B]$  a matriz ampliada de  $AX = B$ , a referida transformação é efectuada através da execução repetida de operações elementares sobre as linhas de  $[A|B]$ . Recordam-se estas operações:

**OE1** *Troca entre si de 2 linhas da matriz.*

**OE2** *Multiplicação dos elementos de uma linha por um escalar diferente de zero.*

**OE3** *Subtração a uma linha de outra linha multiplicada por uma constante.*

Qualquer destas operações sobre  $[A|B]$  substitui o sistema inicial por outro equivalente, isto é, com as mesmas soluções. Realizando estas operações de forma sistemática, é possível, a partir do sistema inicial, determinar um sistema equivalente  $UX = Y$ , com  $U$  triangular superior, cujas soluções são facilmente obtidas.

De notar que, durante a resolução do sistema, poderemos, se tal for conveniente, trocar as colunas da matriz ampliada que contêm os coeficientes das incógnitas. Contudo, essa troca implica a troca das correspondentes variáveis na solução  $X$ .

Para ilustrar o método de eliminação de Gauss vejamos um exemplo.

**Exemplo 1.1** Consideremos o sistema

$$\begin{cases} 2x_1 - 6x_2 + 4x_3 = -2 \\ x_1 - 2x_2 + x_3 = 5 \\ 2x_1 - 4x_2 + 6x_3 = 6 \end{cases}.$$

Tem-se então que

$$A = \begin{bmatrix} 2 & -6 & 4 \\ 1 & -2 & 1 \\ 2 & -4 & 6 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} -2 \\ 5 \\ 6 \end{bmatrix},$$

sendo

$$\left[ \begin{array}{ccc|c} 2 & -6 & 4 & -2 \\ 1 & -2 & 1 & 5 \\ 2 & -4 & 6 & 6 \end{array} \right]$$

a respectiva matriz ampliada. Para transformar esta matriz efectuem-se, sucessivamente, as seguintes operações sobre linhas:

1. Dado que  $a_{11} = 2 \neq 0$ , podemos usar este elemento para anular os restantes elementos da 1ª coluna ( $a_{21} = 1$  e  $a_{31} = 2$ ). Para tal, subtraímos aos elementos das 2ª e 3ª linhas o produto da linha 1 pelos multiplicadores

$$m_{21} = \frac{a_{21}}{a_{11}} = \frac{1}{2} \quad \text{e} \quad m_{31} = \frac{a_{31}}{a_{11}} = 1,$$

respectivamente. Este procedimento equivale à execução da operação OE3, podendo ser representado como segue:

$$\left[ \begin{array}{ccc|c} \boxed{2} & -6 & 4 & -2 \\ 1 & -2 & 1 & 5 \\ 2 & -4 & 6 & 6 \end{array} \right] \xrightarrow[\substack{L_2 - \frac{1}{2}L_1 \ (m_{21} = \frac{1}{2}) \\ L_3 - L_1 \ (m_{31} = 1)}]{\text{OE3}} \left[ \begin{array}{ccc|c} 2 & -6 & 4 & -2 \\ 0 & 1 & -1 & 6 \\ 0 & 2 & 2 & 8 \end{array} \right]$$

O elemento  $a_{11} = 2$  que se encontra assinalado por um quadrado, designa-se por **redutor** (em inglês, *pivot*).

2. Continuando a designar por  $a_{ij}$  os elementos da matriz ampliada obtida, tomando  $a_{22} = 1 \neq 0$  como redutor, podemos anular o elemento abaixo ( $a_{32} = 2$ ) utilizando de novo a operação OE3. Bastará, assim, efectuar a subtracção entre a 3ª linha e o produto da 2ª linha pelo multiplicador

$$m_{32} = \frac{a_{32}}{a_{22}} = \frac{2}{1} = 2,$$

a qual se representa por

$$\left[ \begin{array}{ccc|c} 2 & -6 & 4 & -2 \\ 0 & \boxed{1} & -1 & 6 \\ 0 & 2 & 2 & 8 \end{array} \right] \xrightarrow[\substack{L_3 - 2L_2 \ (m_{32} = 2)}]{\text{OE3}} \left[ \begin{array}{ccc|c} 2 & -6 & 4 & -2 \\ 0 & 1 & -1 & 6 \\ 0 & 0 & 4 & -4 \end{array} \right].$$

Designando a matriz ampliada obtida por  $[U|Y]$  verifica-se que

$$U = \begin{bmatrix} 2 & -6 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix}$$

é triangular superior, pelo que o sistema correspondente  $UX = Y$ , isto é,

$$\begin{cases} 2x_1 - 6x_2 + 4x_3 = -2 \\ x_2 - x_3 = 6 \\ 4x_3 = -4, \end{cases}$$

pode ser facilmente resolvido por **rectrosubstituição** ou **substituição ascendente**. De facto, começando pela resolução da 3ª equação e prosseguindo de forma ascendente, obtém-se

$$\begin{cases} x_3 = -4/4 = -1 \\ x_2 = 6 + x_3 = 5 \\ x_1 = \frac{1}{2}(-2 + 6x_2 - 4x_3) = 16. \end{cases}$$

Com vista a descrever no caso geral o método de eliminação de Gauss, designemos por

$$\left[ A^{(1)} | B^{(1)} \right] = \left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{array} \right] \quad (2)$$

a matriz ampliada de um sistema  $AX = B$  com  $n$  equações lineares e  $n$  incógnitas. Podemos então sintetizar aquele método como segue.

Mediante a troca de linhas, suponha-se que é possível colocar um elemento não nulo na posição  $(1, 1)$  (se isto não for possível, terminamos o processo e concluímos que a matriz  $A$  é singular visto que a 1ª coluna é nula). Continuando a usar a representação (2) para a matriz resultante da eventual troca de linhas, podemos tomar o elemento não nulo  $a_{11}^{(1)}$  como redutor para anular todos os restantes elementos da 1ª coluna. Para tal, basta subtrair da linha  $i$  ( $i = 2, \dots, n$ ) a 1ª linha multiplicada por  $m_{i1} = a_{i1}/a_{11}$ , de forma a obter a seguinte matriz:

$$\left[ A^{(2)} | B^{(2)} \right] = \left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right].$$

Tem-se, assim, que

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1}a_{ij}^{(1)}, \quad i = 2, \dots, n, \quad j = 1, \dots, n \quad (3)$$

$$b_i^{(2)} = b_i^{(1)} - m_{i1}b_1^{(1)}, \quad i = 2, \dots, n. \quad (4)$$

A obtenção da matriz  $[A^{(2)} | B^{(2)}]$  culmina o 1º passo do método de eliminação de Gauss.

Seguidamente, se o elemento  $a_{22}^{(2)}$  é não nulo, subtraímos de cada linha  $i$  ( $i = 3, \dots, n$ ) de  $A^{(2)}$  o produto da segunda linha por  $m_{i2} = a_{i2}^{(2)}/a_{22}^{(2)}$ , anulando desta forma os elementos da segunda coluna de  $A^{(2)}$  por baixo do redutor  $a_{22}^{(2)}$ . Obtemos, assim, a matriz

$$\left[ A^{(3)} | B^{(3)} \right] = \left[ \begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{array} \right],$$

em que  $a_{ij}^{(3)}$  e  $b_i^{(3)}$  se obtêm mediante fórmulas análogas a (3) e (4). Caso o elemento  $a_{22}^{(2)}$  seja nulo, por troca de linhas colocamos um elemento não nulo na posição  $(2, 2)$ , o qual pode ser utilizado para anular os elementos que estão abaixo (se não for possível colocar em  $(2, 2)$  um elemento não nulo, paramos o processo concluindo que  $A$  é singular). A obtenção da matriz  $[A^{(3)} | B^{(3)}]$  finaliza o 2º passo do método de eliminação de Gauss.

O procedimento realizado em cada um dos passos anteriores é seguidamente repetido a partir da posição  $(3, 3)$  e termina quando a matriz dos coeficientes do sistema transformado for

triangular superior. Após  $n - 1$  passos obtém-se a matriz

$$\left[ A^{(n)} | B^{(n)} \right] = \left[ \begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & | & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & | & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & | & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} & | & b_n^{(n)} \end{array} \right],$$

em que  $a_{ii}^{(i)} \neq 0$ ,  $i = 1, \dots, n$ . Considerando  $U = [u_{ij}] = A^{(n)}$  e  $Y = [y_i] = B^{(n)}$ , tem-se que a matriz ampliada  $[A^{(n)} | B^{(n)}]$  corresponde ao sistema  $UX = Y$ , que se encontra na forma triangular. Por retrosubstituição pode-se facilmente resolver este sistema usando as fórmulas

$$\begin{aligned} x_n &= y_n / u_{nn} \\ x_i &= \frac{1}{u_{ii}} \left( y_i - \sum_{j=i+1}^n u_{ij} x_j \right), \quad i = n-1, n-2, \dots, 1. \end{aligned}$$

## 2 Factorização LU

Ao longo da resolução de um sistema  $AX = B$  pelo método de eliminação de Gauss são calculados os multiplicadores  $m_{ij}$ . Por diversos motivos, é de toda a conveniência guardar aqueles multiplicadores. Numa implementação computacional do método de eliminação de Gauss, os locais apropriados para guardar os  $m_{ij}$  são precisamente as entradas da matriz  $A$  que vão sendo sucessivamente anuladas. Assim, admitindo que não há troca de linhas, os multiplicadores  $m_{i1}$ ,  $i = 2, \dots, n$ , devem ser guardados nas posições inicialmente ocupadas pelos elementos  $a_{i1}$  da matriz  $A$ . Seguidamente,  $m_{i2}$ ,  $i = 3, \dots, n$ , devem ser guardados nas posições inicialmente ocupadas pelos elementos  $a_{i2}$  de  $A$  e, assim, sucessivamente. Por fim, o multiplicador  $m_{n,n-1}$  ocupará a posição do elemento  $a_{n,n-1}$  de  $A$ .

Passamos a explicitar as razões que justificam a conservação dos multiplicadores  $m_{ij}$ . Com efeito, consideremos a matriz triangular inferior,

$$L = \left[ \begin{array}{cccc|c} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & m_{n3} & \cdots & 1 \end{array} \right], \quad (5)$$

cujos elementos da diagonal principal são todos iguais a 1 e que, abaixo desta diagonal, apresenta os multiplicadores  $m_{ij}$ . Seja  $A$  a matriz dos coeficientes de um sistema de equações lineares ao qual se aplicou o método de eliminação de Gauss. Admitindo que  $A$  é não singular, tem-se o seguinte resultado cuja demonstração pode ser vista em [1, 2].

**Teorema 2.1** *Suponhamos que ao longo do processo de eliminação de Gauss não houve troca de linhas. Então a matriz  $A$  admite a factorização*

$$A = LU$$

em que  $L$  é a matriz triangular inferior de diagonal unitária e  $U$  é a matriz triangular superior  $A^{(n)}$  obtida no final daquele processo.

**Exemplo 2.1** Para o sistema do exemplo (1.1) tem-se,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 2 & -6 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix}.$$

Facilmente se conclui que

$$LU = \begin{bmatrix} 2 & -6 & 4 \\ 1 & -2 & 1 \\ 2 & -4 & 6 \end{bmatrix} = A. \blacksquare$$

O teorema 2.1 permite escrever o sistema  $AX = B$  na forma

$$(LU)X = B \Leftrightarrow L(UX) = B.$$

Considerando  $UX = Y$ , conclui-se que o método de eliminação de Gauss equivale à factorização  $A = LU$  seguida da resolução de dois sistemas: primeiro,  $LY = B$  para obter  $Y$  e depois,  $UX = Y$  para obter  $X$ . As soluções  $Y$  e  $X$  calculam-se facilmente visto que  $L$  e  $U$  são matrizes triangulares.

Por outro lado, uma vez determinada a factorização  $A = LU$ , havendo necessidade de determinar a solução  $X$  do sistema  $AX = B'$ , em que  $B'$  é diferente de  $B$ , bastará resolver  $LY = B'$  e seguidamente  $UX = Y$ . Trata-se, com efeito, de uma vantagem significativa da factorização  $LU$  visto que, uma vez realizada, pode ser empregue para resolver sistemas que tenham em comum a mesma matriz de coeficientes. Justifica-se, assim, que se guardem os multiplicadores  $m_{ij}$  tal como foi mencionado acima até porque, como veremos, a factorização  $A = LU$  é a fase mais dispendiosa do método de Gauss, em termos computacionais (isto é, em termos de número de operações aritméticas a realizar).

Outra vantagem proporcionada pela factorização  $LU$  é a possibilidade de calcular de uma forma muito simples o determinante de  $A$ . Efectivamente, se  $A = LU$ , resulta das propriedades do determinante que

$$\det A = \det L \times \det U.$$

Como o determinante de uma matriz triangular é o produto dos elementos da diagonal principal, tem-se  $\det L = 1$  e  $\det U = u_{11}u_{22} \cdots u_{nn}$ , pelo que

$$\det A = \det U = u_{11}u_{22} \cdots u_{nn}.$$

**Exemplo 2.2** Aproveitando a factorização  $LU$  da matriz  $A$  obtida no exemplo 2.1, pretende-se calcular  $\det A$  e resolver o sistema  $AX = B'$ , em que

$$B' = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}.$$

Em primeiro lugar, tem-se  $\det A = \det U = 2 \times 1 \times 4 = 8$ .

Para resolver  $AX = B'$  vamos determinar sucessivamente as soluções dos sistemas  $LY = B'$  e  $UX = Y$ . A matriz ampliada de  $LY = B'$  é

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 1/2 & 1 & 0 & 4 \\ 1 & 2 & 1 & 1 \end{array} \right],$$

pelo que, designando por  $y_1, y_2$  e  $y_3$  as componentes de  $Y$ , obtém-se por substituição descendente

$$\begin{cases} y_1 = 2 \\ y_2 = 4 - \frac{1}{2}y_1 = 3 \\ y_3 = 1 - y_1 - 2y_2 = 1 - 2 - 6 = -7 \end{cases}.$$

Por outro lado, a matriz ampliada de  $UX = Y$  é

$$\left[ \begin{array}{ccc|c} 2 & -6 & 4 & 2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 4 & -7 \end{array} \right]$$

pelo que, designando por  $x_1, x_2$  e  $x_3$  as componentes de  $X$ , obtém-se por substituição ascendente

$$\begin{cases} x_1 = \frac{1}{2}(2 + 6x_2 - 4x_3) = 33/4 \\ x_2 = 3 + x_3 = 5/4 \\ x_3 = -7/4 \end{cases}. \blacksquare$$

Como hipótese necessária à validade do teorema 2.1 admitiu-se que, ao longo do método de eliminação de Gauss, não houve troca de linhas. Veremos seguidamente que, no caso em que se efectuem trocas de linhas, o teorema continua válido desde que enunciado de forma ligeiramente diferente.

A troca de linhas na matriz  $A$  equivale à premultiplicação de  $A$  por uma matriz  $P$ , habitualmente designada por **matriz de permutação**. Por exemplo, considerando a matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad (6)$$

a troca das 1ª e 2ª linhas desta matriz equivale à premultiplicação de  $A$  por

$$P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (7)$$

Com efeito,

$$P_{12}A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

é a matriz que resulta de  $A$  pela permutação das referidas linhas.

Uma matriz  $P$  que resulta da troca de duas linhas da matriz identidade diz-se uma **matriz de permutação elementar**. A matriz  $P_{12}$  vista acima é uma matriz de permutação elementar (o índice 12 indica as linhas movimentadas).

Imaginemos mais geralmente que se efectuem na matriz  $A$  dada em (6) duas trocas sucessivas de linhas, a saber: primeiro, trocam-se as 1ª e 2ª linhas e, seguidamente, trocam-se as 2ª e 3ª linhas da matriz  $P_{12}A$ . A realização destas trocas sucessivas é equivalente à premultiplicação de  $A$  pela matriz  $P = P_{23}P_{12}$ , onde

$$P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$



De facto,

$$\begin{aligned}
PA &= P_{23}P_{12}A \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\
&= \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{bmatrix}
\end{aligned}$$

é a matriz que resulta de  $A$  após a realização das trocas mencionadas. A matriz  $P = P_{23}P_{12}$  é uma matriz de permutação obtida a partir do produto de duas matrizes de permutação elementar. Generalizando o raciocínio efectuado, conclui-se que uma sucessão de trocas de linhas efectuadas numa matriz  $A$  equivale à premultiplicação desta matriz por uma matriz de permutação  $P$ , a qual resulta do produto das matrizes de permutação elementar associadas a cada uma das trocas.

Suponha-se agora que se aplica o método de eliminação de Gauss à matriz ampliada  $[A|B]$  de um sistema de equações lineares. Efectuar a totalidade das trocas de linhas ao longo do processo de eliminação é equivalente a realizar todas essas trocas na matriz  $[A|B]$  antes de proceder à eliminação. A realização destas trocas é, como vimos, equivalente à premultiplicação de  $[A|B]$  por uma matriz de permutação  $P$ . Então, aplicando o processo de eliminação de Gauss à matriz ampliada  $[PA|PB]$  não teríamos necessidade de efectuar trocas de linhas, pelo que o teorema 2.1 garante que

$$LU = PA.$$

Tem-se, assim, que nesta igualdade,  $U$  é a matriz triangular superior obtida no final do processo de eliminação e  $L$  é a matriz triangular inferior formada a partir dos multiplicadores utilizados ao longo do mesmo processo. Deste modo, deixando cair a hipótese da ausência de trocas de linhas, o teorema 2.1 mantém-se válido desde que se substitua no seu enunciado a matriz  $A$  pelo produto  $PA$ .

**Exemplo 2.3** Resolver pelo método de eliminação de Gauss o sistema  $AX = B$  em que

$$A = \begin{bmatrix} 0 & -1 & 4 \\ 1 & -2 & 1 \\ 1 & -4 & 6 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} -2 \\ 5 \\ 6 \end{bmatrix}.$$

Obter, em seguida, a factorização  $LU$  da matriz  $PA$ , sendo  $P$  a matriz de permutação correspondente às trocas efectuadas.

Representando a troca de duas linhas  $i$  e  $j$  por  $L_i \leftrightarrow L_j$ , tem-se sucessivamente:

$$\begin{bmatrix} 0 & -1 & 4 & | & -2 \\ 1 & -2 & 1 & | & 5 \\ 1 & -4 & 6 & | & 6 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 1 & -2 & 1 & | & 5 \\ 0 & -1 & 4 & | & -2 \\ 1 & -4 & 6 & | & 6 \end{bmatrix}$$

$$\begin{array}{l} \xrightarrow{\text{OE3}} \\ (m_{21} = 0) \\ L_3 - L_1(m_{31} = 1) \end{array} \left[ \begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & -1 & 4 & -2 \\ 0 & -2 & 5 & 1 \end{array} \right]$$

$$\xrightarrow{\text{OE3}} \left[ \begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & -1 & 4 & -2 \\ 0 & 0 & -3 & 5 \end{array} \right].$$

Por substituição ascendente, obtém-se então

$$\begin{cases} x_3 = -5/3 \\ x_2 = 2 + 4x_3 = -14/3 \\ x_1 = 5 + 2x_2 - x_3 = -8/3 \end{cases}.$$

Conclui-se assim que

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & -3 \end{bmatrix}.$$

Uma vez que só houve a troca das linhas 1 e 2, o produto  $LU$  coincide com  $P_{12}A$ , em que  $P_{12}$  é a matriz dada em (7). Como se pode verificar,

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & 4 \\ 1 & -4 & 6 \end{bmatrix} = P_{12}A. \blacksquare$$

**Exercício 2.1** Utilizando o método de eliminação de Gauss resolva os seguinte sistema de equações:

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 + x_4 = 4 \\ 2x_1 + x_2 + x_3 - x_4 = 7 \\ 2x_1 + x_3 + 2x_4 = -4 \\ x_1 + 2x_2 + 4x_4 = -7 \end{cases}.$$

- (a) Obtenha a factorização  $LU$  da matriz dos coeficientes do sistema e aproveite-a para obter o determinante dessa matriz.  
 (b) Utilize a alínea anterior para resolver o sistema tomando para 2º membro o vector  $(-1, 6, 1, 0)$ .

**Resolução:** (a) Neste caso tem-se

$$A = \begin{bmatrix} 3 & 2 & 4 & 1 \\ 2 & 1 & 1 & -1 \\ 2 & 0 & 1 & 2 \\ 1 & 2 & 0 & 4 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 4 \\ 7 \\ -4 \\ -7 \end{bmatrix},$$

pelo que

$$\left[ \begin{array}{cccc|c} 3 & 2 & 4 & 1 & 4 \\ 2 & 1 & 1 & -1 & 7 \\ 2 & 0 & 1 & 2 & -4 \\ 1 & 2 & 0 & 4 & -7 \end{array} \right]$$

é a matriz ampliada do sistema. Aplicando o método de eliminação de Gauss a esta matriz vem:

$$\begin{aligned}
& \left[ \begin{array}{cccc|c} 3 & 2 & 4 & 1 & 4 \\ 2 & 1 & 1 & -1 & 7 \\ 2 & 0 & 1 & 2 & -4 \\ 1 & 2 & 0 & 4 & -7 \end{array} \right] \xrightarrow{L_1 \leftrightarrow L_4} \left[ \begin{array}{cccc|c} \boxed{1} & 2 & 0 & 4 & -7 \\ 2 & 1 & 1 & -1 & 7 \\ 2 & 0 & 1 & 2 & -4 \\ 3 & 2 & 4 & 1 & 4 \end{array} \right] \\
& \xrightarrow{OE3} \left[ \begin{array}{cccc|c} 1 & 2 & 0 & 4 & -7 \\ 0 & \boxed{-3} & 1 & -9 & 21 \\ 0 & -4 & 1 & -6 & 10 \\ 0 & -4 & 4 & -11 & 25 \end{array} \right] \xrightarrow{OE3} \left[ \begin{array}{cccc|c} 1 & 2 & 0 & 4 & -7 \\ 0 & -3 & 1 & -9 & 21 \\ 0 & 0 & \boxed{-1/3} & 6 & -18 \\ 0 & 0 & 8/3 & 1 & -3 \end{array} \right] \\
& \begin{matrix} m_{21} = 2 \\ m_{31} = 2 \\ m_{41} = 3 \end{matrix} \quad \begin{matrix} m_{32} = 4/3 \\ m_{42} = 4/3 \end{matrix} \\
& \xrightarrow{OE3} \left[ \begin{array}{cccc|c} 1 & 2 & 0 & 4 & -7 \\ 0 & -3 & 1 & -9 & 21 \\ 0 & 0 & -1/3 & 6 & -18 \\ 0 & 0 & 0 & \boxed{49} & -147 \end{array} \right] \\
& \begin{matrix} m_{43} = -8 \end{matrix}
\end{aligned}$$

Por substituição ascendente obtém-se finalmente

$$\begin{cases} x_4 = -147/49 = -3 \\ x_3 = -3(-18 - 6x_4) = 0 \\ x_2 = -\frac{1}{3}(21 - x_3 + 9x_4) = 2 \\ x_1 = -7 - 2x_2 - 4x_4 = 1. \end{cases}$$

(b) Atendendo aos multiplicadores utilizados no processo de eliminação de Gauss temos que

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 \\ m_{31} & m_{32} & 1 & 0 \\ m_{41} & m_{42} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 4/3 & 1 & 0 \\ 3 & 4/3 & -8 & 1 \end{bmatrix};$$

por outro lado,

$$U = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & -3 & 1 & -9 \\ 0 & 0 & -1/3 & 6 \\ 0 & 0 & 0 & 49 \end{bmatrix}.$$

Uma vez que só houve a troca das linhas 1 e 4, o produto  $LU$  coincide com  $P_{14}A$ , em que  $P_{14}$  é a matriz

$$P_{14} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

De facto, como se pode verificar, tem-se

$$LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 4/3 & 1 & 0 \\ 3 & 4/3 & -8 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 4 \\ 0 & -3 & 1 & -9 \\ 0 & 0 & -1/3 & 6 \\ 0 & 0 & 0 & 49 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 1 & 1 & -1 \\ 2 & 0 & 1 & 2 \\ 3 & 2 & 4 & 1 \end{bmatrix}$$

e

$$P_{14}A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 & 4 & 1 \\ 2 & 1 & 1 & -1 \\ 2 & 0 & 1 & 2 \\ 1 & 2 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 4 \\ 2 & 1 & 1 & -1 \\ 2 & 0 & 1 & 2 \\ 3 & 2 & 4 & 1 \end{bmatrix}.$$

Para obter o determinante de  $A$  observe-se que a igualdade  $P_{14}A = LU$  implica que

$$\begin{aligned} \det P_{14}A &= \det LU \Leftrightarrow \det P_{14} \times \det A = \det L \times \det U \\ &\Leftrightarrow \det P_{14} \times \det A = 1 \times [1 \cdot (-3) \cdot (-1/3) \cdot 49] \\ &\Leftrightarrow \det P_{14} \times \det A = 49. \end{aligned}$$

Como  $\det P_{14} = -1$  (recorde-se que a troca de duas linhas ou colunas de uma matriz obriga a trocar o sinal ao determinante) conclui-se que  $\det A = -49$ .

(c) Para resolver

$$AX = B' = \begin{bmatrix} -1 \\ 6 \\ 1 \\ 0 \end{bmatrix}$$

utilizando a factorização  $LU$  de  $P_{14}A$  teremos primeiro que premultiplicar  $B'$  por  $P_{14}$  como forma de efectuarmos em  $B'$  a troca de linhas realizada em  $A$ . Este procedimento conduz à resolução do sistema equivalente  $P_{14}AX = P_{14}B'$ , para o que é suficiente determinar sucessivamente as soluções dos sistemas  $LY = P_{14}B'$  e  $UX = Y$ . A matriz ampliada de  $LY = P_{14}B'$  é

$$\left[ \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 6 \\ 2 & 4/3 & 1 & 0 & 1 \\ 3 & 4/3 & -8 & 1 & -1 \end{array} \right],$$

pelo que, designando por  $y_1, y_2, y_3$  e  $y_4$  as componentes de  $Y$ , obtém-se por substituição descendente

$$\begin{cases} y_1 = 0 \\ y_2 = 6 - 2y_1 = 6 \\ y_3 = 1 - 2y_1 - \frac{4}{3}y_2 = -7 \\ y_4 = -1 - 3y_1 - \frac{4}{3}y_2 + 8y_3 = -65 \end{cases}.$$

Por outro lado, a matriz ampliada de  $UX = Y$  é

$$\left[ \begin{array}{cccc|c} 1 & 2 & 0 & 4 & 0 \\ 0 & -3 & 1 & -9 & 6 \\ 0 & 0 & -1/3 & 6 & -7 \\ 0 & 0 & 0 & 49 & -65 \end{array} \right]$$

pelo que, designando por  $x_1, x_2, x_3$  e  $x_4$  as componentes de  $X$ , obtém-se por substituição ascendente

$$\begin{cases} x_4 = -65/49 \\ x_3 = -3 \cdot (-7 - 6x_4) = -141/49 \\ x_2 = -\frac{1}{3}(6 - x_3 + 9x_4) = 50/49 \\ x_1 = 0 - 2x_2 - 4x_4 = 160/49 \end{cases} \quad \blacksquare$$

## 2.1 Avaliação do esforço computacional

O esforço computacional realizado por um algoritmo é habitualmente medido pelo número de operações aritméticas elementares (adições, subtracções, multiplicações ou divisões) efectuadas. Neste contexto, é costume usar-se a terminologia inglesa pelo que a locução “número de operações aritméticas efectuadas” é vulgarmente substituída por “numero de *flops* efectuados” (o termo *flops* abrevia a expressão inglesa *floating point operations*).

O objectivo que temos agora em vista é a contagem do número de *flops* necessários para resolver um sistema  $AX = B$  pelo método de eliminação de Gauss. Vamos considerar separadamente a criação das matrizes  $L$  e  $U$  a partir de  $A$ , depois a transformação de  $B$  no vector  $Y$  e finalmente a resolução de  $UX = Y$  utilizando a substituição ascendente.

### 1. Cálculo de $L$ e $U$

Considerando a descrição do método de eliminação de Gauss feita na secção 1, para obter a matriz  $A^{(2)}$  são necessárias  $n - 1$  divisões para calcular os multiplicadores  $m_{i1}$ ,  $i = 2, \dots, n$ . A obtenção dos elementos  $a_{ij}^{(2)}$  requer, por outro lado, a realização de  $(n - 1)^2$  multiplicações e  $(n - 1)^2$  subtracções. Utilizando o mesmo raciocínio podemos proceder à contagem das operações realizadas na obtenção de cada matriz  $A^{(i)}$ ,  $i = 1, \dots, n$ ; a tabela seguinte resume estas contagens.

Obtenção de $A^{(i)}$	Adições/Subtracções	Multiplicações	Divisões
$A^{(2)}$	$(n - 1)^2$	$(n - 1)^2$	$n - 1$
$A^{(3)}$	$(n - 2)^2$	$(n - 2)^2$	$n - 2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$A^{(n)} = U$	1	1	1
Total	$\frac{n(n-1)(2n-1)}{6}$	$\frac{n(n-1)(2n-1)}{6}$	$\frac{n(n-1)}{2}$

O valor total em cada coluna foi obtido usando as igualdades

$$\sum_{j=1}^p j = \frac{p(p-1)}{2} \quad \text{e} \quad \sum_{j=1}^p j^2 = \frac{p(p+1)(2p+1)}{6}, \quad p \geq 1.$$

Para a factorização  $LU$ , temos assim que o número total de *flops* é

$$\frac{n(n-1)(2n-1)}{6} + \frac{n(n-1)(2n-1)}{6} + \frac{n(n-1)}{2} = \frac{2n^3}{3} - \frac{n^2}{2} - \frac{n}{6}$$

### 2. Transformação de $B$ no vector $Y$

Neste caso temos o seguinte quadro:

Obtenção de $B^{(i)}$	Adições/Subtracções	Multiplicações
$B^{(2)}$	$n - 1$	$n - 1$
$B^{(3)}$	$n - 2$	$n - 2$
$\vdots$	$\vdots$	$\vdots$
$B^{(n)} = Y$	1	1
Total	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$

O número total de *flops* é

$$\frac{n(n-1)}{2} + \frac{n(n-1)}{2} = n^2 - n.$$

### 3. Resolução de $UX = Y$

Neste caso tem-se

Obtenção de $x_i$	Adições/Subtrações	Multiplicações	Divisões
$x_n$	0	0	1
$x_{n-1}$	1	1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_1$	$n-1$	$n-1$	1
Total	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$	$n$

Assim, o número total de *flops* é

$$\frac{n(n-1)}{2} + \frac{n(n-1)}{2} + n = n^2.$$

Dos pontos 1., 2. e 3. conclui-se que o número total de *flops* necessários para resolver  $AX = B$  é

$$\left(\frac{2n^3}{3} - \frac{n^2}{2} - \frac{n}{6}\right) + (n^2 - n) + n^2 = \frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n \approx \frac{2}{3}n^3.$$

A estimativa de  $2n^3/3$  apontada para o número total de *flops* é válida para grandes valores de  $n$ . De notar que a potência  $n^3$  só aparece no cálculo do número de *flops* da factorização  $LU$ . Deste modo, o cálculo de  $L$  e  $U$  constitui o maior esforço computacional do processo de eliminação de Gauss. Uma vez obtidas estas matrizes, são somente necessários  $2n^2 - n \approx 2n^2$  *flops* para resolver  $AX = B$ . Conclui-se, portanto, que é relativamente económico resolver outros sistemas com a mesma matriz dos coeficientes, desde que a factorização  $LU$  tenha sido guardada.

Convém referir que, do ponto de vista computacional, o processo de eliminação de Gauss é radicalmente mais económico do que a regra de Cramer. Efectivamente, se nesta regra os determinantes forem calculados através da regra de Laplace, o número de *flops* efectuados é de  $(n+1)!$  para um sistema com  $n$  equações e  $n$  incógnitas. Assim, para  $n = 10$ , o método de eliminação de Gauss efectua cerca de 700 *flops* enquanto a regra de Cramer realiza  $11! = 39\,916\,800$  *flops*. Por esta razão não é recomendável a utilização da regra de Cramer na resolução de sistemas de equações lineares de grandes ou mesmo moderadas dimensões.

## 3 Pesquisa de redutor

Por definição, os redutores  $a_{ii}^{(i)}$ ,  $i = 1, \dots, n$ , utilizados no método de eliminação de Gauss são não nulos. Contudo, numa implementação computacional do método, não é suficiente que os redutores sejam diferentes de zero. Na realidade, os erros de arredondamento das operações em vírgula flutuante, podem tornar não nulo um redutor que seria nulo caso os cálculos fossem realizados em aritmética exacta. Compreensivelmente, a utilização deste redutor pode introduzir erros grosseiros nos cálculos seguintes, os quais se podem repercutir indesejavelmente

na solução final do sistema. Para minorar este tipo de problemas, recorre-se a determinados métodos que visam estabelecer critérios a usar na escolha dos redutores. Veremos seguidamente os métodos designados por **pesquisa parcial de redutor** e **pesquisa total de redutor**.

**Pesquisa parcial de redutor.** Seja

$$\left[ A^{(k)} | B^{(k)} \right] = \left[ \begin{array}{cccccc|c} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1k}^{(k)} & \cdots & a_{1n}^{(k)} & b_1^{(k)} \\ 0 & a_{22}^{(k)} & \cdots & a_{2k}^{(k)} & \cdots & a_{2n}^{(k)} & b_2^{(k)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & \vdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(1)} \end{array} \right]$$

a matriz ampliada no início do passo  $k \in \{1, \dots, n-1\}$  do método de eliminação de Gauss. Considere-se

$$c_k = \max_{k \leq i \leq n} |a_{ik}^{(k)}|, \quad (8)$$

isto é, maior dos valores absolutos dos elementos  $a_{kk}^{(k)}, \dots, a_{nk}^{(k)}$  da coluna  $k$  abaixo e sobre a diagonal principal de  $A^{(k)}$ . Suponhamos que este máximo é atingido para o elemento  $a_{rk}^{(k)}$  que se encontra na linha  $r \in \{k, \dots, n\}$ , isto é,  $c_k = |a_{rk}^{(k)}|$ . Então, se  $r > k$ , trocam-se as linhas  $r$  e  $k$  da matriz  $[A^{(k)} | B^{(k)}]$  e prossegue-se com o passo  $k$  do processo de eliminação de Gauss. De notar que este procedimento garante que os multiplicadores  $m_{ik}$ ,  $i = k+1, \dots, n$ , verificam  $|m_{ik}| \leq 1$  o que evita uma grande variação dos elementos de  $[A^{(k)} | B^{(k)}]$  e, possivelmente, erros propagados de grande magnitude.

**Exemplo 3.1** Arredondando os resultados dos cálculos para 4 dígitos, vamos resolver o sistema

$$\begin{cases} 0.005x + 0.81y = 0.7842 \\ 2.333x + 1.52y = 1.225 \end{cases}$$

sem pesquisa de redutor e com pesquisa parcial de redutor. A solução do sistema com 4 algarismos exactos é

$$x_1 = -0.1061 \quad \text{e} \quad x_2 = 0.9688.$$

#### 1. Resolução sem pesquisa de redutor

A transformação da matriz ampliada numa matriz triangular representa-se por

$$\begin{aligned} & \left[ \begin{array}{cc|c} 0.005 & 0.81 & 0.7842 \\ 2.333 & 1.52 & 1.225 \end{array} \right] \\ & \quad \downarrow m_{21} = \frac{2.333}{0.005} = 466 \\ & \left[ \begin{array}{cc|c} 0.005 & 0.81 & 0.7842 \\ 0 & -375.9 & -364.2 \end{array} \right] \end{aligned}$$

Logo, a solução obtida por substituição ascendente é

$$\begin{cases} x_2 = \frac{-364.2}{-375.9} = 0.9689 \\ x_1 = \frac{0.7842 - 0.81 \cdot 0.9689}{0.005} = -0.1218 \end{cases} \quad (9)$$

## 2. Resolução com pesquisa de redutor

Neste caso teremos de começar por trocar as 1ª e 2ª linhas da matriz ampliada visto que

$$c_1 = \max \{|0.005|, |2.333|\} = 2.333.$$

A transformação da matriz ampliada numa matriz triangular representa-se por

$$\begin{aligned} & \left[ \begin{array}{cc|c} 0.005 & 0.81 & 0.7842 \\ 2.333 & 1.52 & 1.225 \end{array} \right] \\ & \quad \downarrow L_1 \leftrightarrow L_2 \\ & \left[ \begin{array}{cc|c} 2.333 & 1.52 & 1.225 \\ 0.005 & 0.81 & 0.7842 \end{array} \right] \\ & \quad \downarrow m_{21} = \frac{0.005}{2.333} = 0.0021 \\ & \left[ \begin{array}{cc|c} 2.333 & 1.52 & 1.225 \\ 0 & 0.8068 & 0.7816 \end{array} \right] \end{aligned}$$

Logo, a solução obtida por substituição ascendente é

$$\begin{cases} x_2 = \frac{0.7816}{0.8068} = 0.9688 \\ x_1 = \frac{1.225 - 1.52 \cdot 0.9688}{2.333} = -0.1061 \end{cases} \quad (10)$$

Vê-se assim que, em (9), a percentagem de erro de  $x_2$  é de cerca de 0.01% mas a percentagem de erro de  $x_1$  é de cerca de 14.8%. Por outro lado, em (10), foi obtida a solução do sistema com 4 algarismos exactos. Isto ilustra o efeito positivo da pesquisa parcial de redutor no método de eliminação de Gauss. ■

**Pesquisa total de redutor.** Este procedimento consiste em calcular, no passo  $k$  da eliminação de Gauss, a quantidade

$$c_k = \max_{i \leq i, j \leq n} |a_{ij}^{(k)}|.$$

Trata-se do maior dos valores absolutos dos elementos da submatriz

$$\left[ \begin{array}{ccc} a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ \vdots & \cdots & \vdots \\ a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{array} \right]$$

de  $A^{(k)}$ . Uma vez calculado  $c_k$  procede-se à troca das linhas de  $[A^{(k)}|B^{(k)}]$  e das colunas de  $A^{(k)}$  de modo a trazer para a posição do redutor o elemento para o qual se obteve o valor de  $c_k$ . De notar que a troca de colunas deve levar a que se troquem igualmente as correspondentes variáveis na matriz das incógnitas.

A pesquisa total de redutor é o procedimento que em geral melhor permite minorar os efeitos decorrentes da propagação dos erros de arredondamento. Ao mesmo tempo, o número de *flops* que requer é muito superior ao número exigido pela pesquisa parcial de redutor. Acresce ainda que, na generalidade dos casos, a precisão permitida pela pesquisa parcial de redutor é semelhante à fornecida pela pesquisa total. Por estes motivos, a pesquisa parcial de redutor é o procedimento mais utilizado no método de eliminação de Gauss para controlar os erros de arredondamento.



## 4 Métodos Compactos

Vimos na secção 2 que o método de eliminação de Gauss permite obter a factorização  $LU$  da matriz dos coeficientes dum sistema  $AX = B$ . E, tal como aí foi referido, uma vez obtidas as matrizes  $L$  e  $U$ , a solução do sistema  $AX = B$  pode ser determinada resolvendo sucessivamente os sistemas triangulares  $LY = B$  e  $UX = Y$ . Acontece, no entanto, que as matrizes  $L$  e  $U$  podem ser calculadas directamente sem recorrer ao método de eliminação de Gauss. Somos então conduzidos aos chamados **métodos compactos**, designação que se deve a que os elementos de  $L$  e  $U$  são agora calculados de uma só vez através de fórmulas explícitas, ao contrário do que acontece na eliminação de Gauss onde aqueles elementos são obtidos por etapas à medida que o processo avança. Veremos seguidamente o **método de Doolittle** que procede à factorização  $LU$  da matriz  $A$  de uma forma compacta. Apresentaremos ainda nesta secção um outro método compacto – o **método de Cholesky** –, que se baseia numa factorização da matriz  $A$  distinta da factorização  $LU$ .

### 4.1 Método de Doolittle

Este método não é mais do que a expressão do método de eliminação de Gauss na forma matricial. Consiste no cálculo directo da factorização  $LU$  da matriz  $A$ , sendo  $L$  a matriz (5) e  $U$  a matriz  $A^{(n)}$  obtida no final da eliminação de Gauss. Antes da apresentação das fórmulas para o cálculo directo de  $L$  e  $U$ , vejamos um exemplo.

**Exemplo 4.1** Vamos resolver pelo método de Doolittle o sistema do exemplo 1.1. A factorização  $LU$  da matriz  $A$  pode obter-se a partir de

$$A = [a_{ij}] = LU \Leftrightarrow \begin{bmatrix} 2 & -6 & 4 \\ 1 & -2 & 1 \\ 2 & -4 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}.$$

Multiplicando as matrizes que se encontram à direita da última igualdade obtém-se

$$\begin{bmatrix} 2 & -6 & 4 \\ 1 & -2 & 1 \\ 2 & -4 & 6 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ m_{21}u_{11} & u_{22} + m_{21}u_{12} & u_{23} + m_{21}u_{13} \\ m_{31}u_{11} & m_{31}u_{12} + m_{32}u_{22} & u_{33} + m_{31}u_{13} + m_{32}u_{23} \end{bmatrix}.$$

Igualando os elementos homólogos da 1ª coluna e da 1ª linha das duas matrizes, vem:

$$\begin{aligned} 1^\text{ª} \text{ coluna} &\rightarrow \begin{cases} u_{11} = a_{11} = 2 \\ m_{21}u_{11} = a_{21} = 1 \Rightarrow m_{21} = \frac{1}{2} \\ m_{31}u_{11} = a_{31} = 2 \Rightarrow m_{31} = \frac{2}{2} = 1 \end{cases} \\ 1^\text{ª} \text{ linha} &\rightarrow \begin{cases} u_{12} = a_{12} = -6 \\ u_{13} = a_{13} = 4 \end{cases} \end{aligned}$$

Prosseguindo com os elementos que restam da 2ª coluna e da 2ª linha das duas matrizes, obtém-se

$$\begin{aligned} 2^\text{ª} \text{ coluna} &\rightarrow \begin{cases} u_{22} + m_{21}u_{12} = a_{22} = -2 \Rightarrow u_{22} = -2 - \frac{1}{2}(-6) = 1 \\ m_{31}u_{12} + m_{32}u_{22} = a_{32} = -4 \Rightarrow m_{32} = \frac{-4 - 1(-6)}{1} = 2 \end{cases} \\ 2^\text{ª} \text{ linha} &\rightarrow \begin{cases} u_{23} + m_{21}u_{13} = a_{23} = 1 \Rightarrow u_{23} = 1 - \frac{1}{2} \cdot 4 = -1 \end{cases} \end{aligned}$$

Finalmente, os elementos da 3ª coluna e da 3ª linha das duas matrizes conduzem a

$$\left. \begin{array}{l} 3^{\text{a}} \text{ coluna} \\ 3^{\text{a}} \text{ linha} \end{array} \right\} \rightarrow \{u_{33} + m_{31}u_{13} + m_{32}u_{23} = a_{33} = 6 \Rightarrow u_{33} = 6 - 1 \cdot 4 - 2(-1) = 4\}.$$

A factorização  $A = LU$  é então

$$\begin{bmatrix} 2 & -6 & 4 \\ 1 & -2 & 1 \\ 2 & -4 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -6 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix},$$

como já se tinha visto no exemplo 2.1. A partir daqui resolve-se  $LY = B$ , isto é,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 6 \end{bmatrix} \Leftrightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \\ -4 \end{bmatrix}$$

e, finalmente, determina-se a solução de  $UX = Y$ :

$$\begin{bmatrix} 2 & -6 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \\ -4 \end{bmatrix} \Leftrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 16 \\ 5 \\ -1 \end{bmatrix}. \blacksquare$$

Observando a determinação das matrizes  $L$  e  $U$  no exemplo anterior, constata-se que o procedimento efectuado é equivalente ao cálculo dos elementos da matriz

$$L + U - I = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ m_{21} & u_{22} & u_{23} \\ m_{31} & m_{32} & u_{33} \end{bmatrix},$$

onde  $I$  representa a matriz identidade. A ordem pela qual os elementos desta matriz podem ser calculados (outras sequências de cálculos existem) também é sugerida no exemplo: primeiro são calculados os elementos da 1ª coluna e da 1ª linha de  $L + U - I$ ; depois são calculados os restantes elementos da 2ª coluna e da 2ª linha de  $L + U - I$ ; finalmente, é obtido o elemento restante que se encontra na 3ª coluna e 3ª linha de  $L + U - I$ .

A extensão deste procedimento ao caso geral não apresenta dificuldades. Sendo  $A$  uma matriz não singular de ordem de ordem  $n$ , o cálculo dos elementos  $m_{ij}$  e  $u_{ij}$  da matriz  $L + U - I$  pode fazer-se mediante as seguintes fórmulas, as quais podem ser facilmente deduzidas usando a técnica apresentada no exemplo anterior:

$$\left\{ \begin{array}{ll} u_{1j} = a_{1j} & j = 1, \dots, n \\ u_{ij} = a_{ij} - \sum_{s=1}^{i-1} m_{is}u_{sj} & j = i, \dots, n; \quad i \geq 2 \\ m_{i1} = \frac{a_{i1}}{u_{11}} & i = 2, \dots, n \\ m_{ij} = \frac{1}{u_{jj}} \left( a_{ij} - \sum_{s=1}^{j-1} m_{is}u_{sj} \right) & i = j+1, \dots, n; \quad j \geq 2 \end{array} \right. . \quad (11)$$

A ordem pela qual os elementos de  $L + U - I$  devem ser calculados está esquematizada a seguir:

$$L + U - I = \left[ \begin{array}{c|c|c|c} \boxed{1} & \boxed{2} & & \\ \boxed{3} & \boxed{4} & & \\ \boxed{5} & \boxed{6} & & \\ & \text{etc.} & & \end{array} \right]. \quad (12)$$

De notar que os elementos  $u_{jj}$  intervenientes em (11) têm de ser não nulos, o que pode sempre ser conseguido por troca de linhas, caso a matriz  $A$  seja não singular.

Tem-se, portanto, que a aplicação do método de Doolittle à resolução de um sistema  $AX = B$  consiste na factorização de  $A$  mediante (11), seguida da determinação das soluções dos sistemas triangulares  $LY = B$  e  $UX = Y$  por substituição descendente e ascendente, respectivamente.

**Exercício 4.1** Resolver pelo método de Doolittle o sistema

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

Vamos aplicar à matriz dos coeficientes

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

as fórmulas (11) e a ordem de cálculo esquematizada em (12). Temos então:

$$\begin{aligned} \textcircled{1} & \rightarrow \begin{cases} u_{11} = a_{11} = 2 \\ m_{21} = \frac{a_{21}}{u_{11}} = \frac{3}{2} \\ m_{31} = \frac{a_{31}}{u_{11}} = \frac{2}{2} = 1 \end{cases} \\ \textcircled{2} & \rightarrow \begin{cases} u_{12} = a_{12} = 1 \\ u_{13} = a_{13} = 3 \end{cases} \\ \textcircled{3} & \rightarrow \begin{cases} u_{22} = a_{22} - \sum_{s=1}^1 m_{2s}u_{s2} = a_{22} - m_{21}u_{12} = 2 - \frac{3}{2} \cdot 1 = \frac{1}{2} \\ m_{32} = \frac{1}{u_{22}} \left( a_{32} - \sum_{s=1}^1 m_{3s}u_{s2} \right) = \frac{a_{32} - m_{31}u_{12}}{u_{22}} = \frac{2 - 1 \cdot 1}{1/2} = 2 \end{cases} \\ \textcircled{4} & \rightarrow \begin{cases} u_{23} = a_{23} - \sum_{s=1}^1 m_{2s}u_{s3} = a_{23} - m_{21}u_{13} = 4 - \frac{3}{2} \cdot 3 = -\frac{1}{2} \end{cases} \\ \textcircled{5} & \rightarrow \begin{cases} u_{33} = a_{33} - \sum_{s=1}^2 m_{3s}u_{s3} = a_{33} - m_{31}u_{13} - m_{32}u_{23} = 1 - 3 + 1 = -1. \end{cases} \end{aligned}$$

Assim, tem-se que

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 3/2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{e} \quad U = \begin{bmatrix} 2 & 1 & 3 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & -1 \end{bmatrix}.$$

Resolve-se agora  $LY = B$ , em que  $B$  é o 2º membro do sistema dado, isto é,

$$\begin{bmatrix} 1 & 0 & 0 \\ 3/2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

pelo que

$$\begin{cases} y_1 = 1 \\ y_2 = 2 - \frac{3}{2}y_1 = 2 - \frac{3}{2} = \frac{1}{2} \\ y_3 = 3 - y_1 - 2y_2 = 3 - 1 - 1 = 1 \end{cases}.$$

Finalmente, determina-se a solução de  $UX = Y$ , isto é,

$$\begin{bmatrix} 2 & 1 & 3 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \\ 1 \end{bmatrix},$$

donde sai que

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

Portanto, a solução do sistema dado é  $x_1 = 2$ ,  $x_2 = 0$  e  $x_3 = -1$ . ■

Tal como no método de eliminação de Gauss, é possível implementar no método de Doolittle a pesquisa parcial de redutor. No entanto, existe um outro método compacto – o **método de Crout** – em que a implementação da pesquisa parcial de redutor se torna mais eficiente. Este procedimento realiza igualmente uma factorização  $LU$  da matriz  $A$ , mas agora os elementos da diagonal principal da matriz triangular superior  $U$  são todos iguais a um, enquanto na diagonal principal da matriz triangular inferior  $L$  podem figurar quaisquer valores. A descrição detalhada deste método pode ser vista em [2].

## 4.2 Método de Cholesky

O **método de Cholesky** é aplicável a sistemas cuja matriz dos coeficientes é simétrica e definida positiva. Uma matriz  $A = [a_{ij}]$  de ordem  $n$  diz-se **definida positiva** se

$$X^T A X = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j > 0,$$

para toda a matriz coluna não nula  $X = [x_i]$  com  $n$  elementos. Faremos uso adiante da seguinte caracterização das matrizes simétricas definidas positivas cuja demonstração pode ser vista em [5, 6].

**Teorema 4.1** *Seja  $A$  uma matriz de ordem  $n$  simétrica. Então  $A$  é definida positiva se e só se todas as submatrizes  $A_i$  formadas pelos elementos das primeiras  $i$  linhas e  $i$  colunas de  $A$  têm determinante positivo, isto é,  $\det A_i > 0$ , para  $i = 1, \dots, n$ .*

**Exemplo 4.2** A matriz simétrica  $A = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 10 \end{bmatrix}$  é definida positiva pois

$$\det A_1 = \det [2] = 2 > 0 \quad \det A_2 = \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = 2 > 0 \quad \text{e}$$

$$\det A_3 = \det A = \det \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 10 \end{bmatrix} = 8 > 0. \blacksquare$$

Na resolução de sistemas  $AX = B$  pelo método de Cholesky, a matriz  $A$  é factorizada em duas matrizes  $L$  e  $L^T$ , tais que  $A = LL^T$ . A matriz  $L$  é triangular inferior (não necessariamente de diagonal unitária) sendo  $L^T$  a sua matriz transposta, logo triangular superior. Assim, o sistema original  $AX = B$  pode escrever-se na forma  $(LL^T)X = B$  ou, equivalentemente,  $L(L^T X) = B$ . Portanto, à semelhança do que foi visto para a decomposição  $LU$ , a solução de  $AX = B$  obtém-se resolvendo sucessivamente os sistemas triangulares

$$LY = B \text{ e } L^T X = Y. \quad (13)$$

Vejamos como realizar a factorização  $LL^T$  de uma matriz  $A$  de ordem  $n$ , suposta simétrica e definida positiva:

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

Pretendemos que  $LL^T = A$ , isto é,

$$\begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Efectuando o produto da 1ª linha de  $L$  pela 1ª coluna de  $L^T$  obtém-se

$$l_{11}^2 = a_{11} \Rightarrow l_{11} = \sqrt{a_{11}}.$$

De notar que  $l_{11} > 0$  pois, como  $A$  definida positiva, vem pelo teorema 4.1,  $\det(A_1) = a_{11} > 0$ .

Multiplicando seguidamente a 2ª linha de  $L$  pelas 1ª e 2ª colunas de  $L^T$  obtém-se, respectivamente,

$$l_{21}l_{11} = a_{21} \quad \text{e} \quad l_{21}^2 + l_{22}^2 = a_{22},$$

pelo que

$$l_{21} = \frac{a_{21}}{l_{11}} \quad \text{e} \quad l_{22} = \sqrt{a_{22} - l_{21}^2}.$$

Tem-se igualmente que  $l_{22} > 0$  pois,  $\det A_2 = l_{11}^2 (a_{22} - l_{21}^2)$  e este determinante é positivo, atendendo a que  $A$  é definida positiva.

Efectuando sucessivamente o procedimento anterior chegaríamos à etapa final que consiste em multiplicar a última linha de  $L$  por todas as colunas de  $L^T$ . Obter-se-ia então

$$l_{n1}l_{11} = a_{n1}, \quad l_{n1}l_{21} + l_{n2}l_{22} = a_{n2}, \dots, l_{n1}^2 + l_{n2}^2 + \dots + l_{nn}^2 = a_{nn},$$

pelo que

$$l_{n1} = \frac{a_{n1}}{l_{11}}, \quad l_{n2} = \frac{1}{l_{22}} \left( a_{n2} - \sum_{k=1}^1 l_{nk}l_{2k} \right), \dots, l_{nn} = \sqrt{a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2}.$$

O facto de  $A$  ser definida positiva garante que  $l_{ii} > 0$ ,  $i = 1, 2, \dots, n$ .

As fórmulas anteriores sugerem que os elementos  $l_{ij}$  da matriz triangular inferior  $L$  podem ser obtidos, para cada  $i = 1, 2, \dots, n$ , mediante as igualdades

$$\begin{cases} l_{ij} = \frac{1}{l_{jj}} \left( a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk} \right) & j = 1, \dots, i-1 \\ l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2} \end{cases} \quad (14)$$

A ordem pela qual os elementos de  $L = [l_{ij}]$  (com  $l_{ij} = 0$  se  $i < j$ ) devem ser calculados está esquematizada a seguir:

$$L = \begin{bmatrix} \boxed{1} & & & & \\ \boxed{2} & \boxed{3} & & & \\ \boxed{4} & & \boxed{5} & & \\ \boxed{6} & & & \boxed{7} & \\ \text{etc.} & & & & \end{bmatrix}$$

**Exemplo 4.3** Vamos resolver pelo método de Cholesky o sistema de equações lineares  $AX = B$ , onde

$$A = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 10 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 6 \\ 14 \\ 42 \end{bmatrix}.$$

Vimos no exemplo anterior que  $A$  é definida positiva. Para obter a factorização  $LL^T$  da matriz  $A$ , apliquemos as fórmulas (14):

$$\begin{aligned} l_{11} &= \sqrt{a_{11}} = \sqrt{2}, \\ l_{21} &= \frac{a_{21}}{l_{11}} = \frac{0}{\sqrt{2}} = 0, \\ l_{22} &= \sqrt{a_{22} - \sum_{k=1}^1 l_{2k}^2} = \sqrt{a_{22} - l_{21}^2} = \sqrt{1 - 0} = 1 \end{aligned}$$

$$\begin{aligned}
l_{31} &= \frac{a_{31}}{l_{11}} = \frac{2}{\sqrt{2}} = \sqrt{2}, \\
l_{32} &= \frac{a_{32} - \sum_{k=1}^1 l_{3k}l_{2k}}{l_{22}} = \frac{a_{32} - l_{31}l_{21}}{l_{22}} = \frac{2 - \sqrt{2} \times 0}{1} = 2 \\
l_{33} &= \sqrt{\left(a_{33} - \sum_{k=1}^2 l_{3k}^2\right)} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{10 - 2 - 4} = 2.
\end{aligned}$$

A factorização  $A = LL^T$  é então

$$\begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 10 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ \sqrt{2} & 2 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & \sqrt{2} \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}.$$

A partir daqui resolve-se  $LY = B$ , isto é,

$$\begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ \sqrt{2} & 2 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ 42 \end{bmatrix}$$

$$\Downarrow$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} \\ 14 \\ 4 \end{bmatrix}$$

e, finalmente, obtém-se a solução de  $L^T X = Y$ :

$$\begin{bmatrix} \sqrt{2} & 0 & \sqrt{2} \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} \\ 14 \\ 4 \end{bmatrix}$$

$$\Downarrow$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \\ 2 \end{bmatrix}. \blacksquare$$

## 5 Métodos Iterativos

Faremos seguidamente um estudo sumário dos métodos iterativos para a resolução de um sistema de equações lineares. Concretamente, apresentaremos os métodos de Jacobi e Gauss-Seidel, cada um dos quais pode ser visto como um conjunto de procedimentos que permite obter uma solução aproximada do sistema a partir de outra solução aproximada, procedimentos esses que podem ser indefinidamente repetidos.

Ao aplicar um método directo para resolver um sistema de equações lineares o utilizador conhece com antecedência o esforço computacional que vai ser dispendido. Pelo contrário, num método iterativo, não se pode à partida prever o número de soluções aproximadas que devem ser calculadas para se obter uma dada precisão, visto que esse número pode variar de sistema de equações para sistema de equações. Porquê então usar métodos iterativos? De

facto, em muitas situações reais é necessário resolver sistemas de grande dimensão (isto é, com 1000 ou mais equações e variáveis) cuja matriz dos coeficientes é **esparsa** (isto é, o número de elementos não nulos é relativamente pequeno). Torna-se então impraticável utilizar um método directo para resolver estes sistemas, atendendo ao espaço de memória necessário para guardar todos coeficientes nulos. Por isso, a utilização de métodos iterativos nestes casos tem todo o cabimento, sendo mesmo em muitos deles a única possibilidade que existe para obter a solução do sistema.

Para a resolver iterativamente  $AX = B$  podemos começar por decompor a matriz  $A$  (que continuaremos a supor não singular) na soma de duas matrizes  $M$  e  $N$ , isto é,

$$A = M + N. \quad (15)$$

O sistema original escreve-se então na forma

$$(M + N)X = B$$

ou, equivalentemente,

$$MX = B - NX.$$

Esta expressão sugere o processo iterativo

$$MX^{(k+1)} = B - NX^{(k)}, \quad k = 0, 1, \dots, \quad (16)$$

em que  $X^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$  designa a solução aproximada da solução do sistema  $AX = B$  obtida ao fim de  $k$  iterações. Escolhendo para  $M$  uma matriz invertível, a igualdade anterior escreve-se na forma

$$X^{(k+1)} = M^{-1}B - M^{-1}NX^{(k)}, \quad k = 0, 1, \dots$$

Assim, a partir de uma solução aproximada inicial  $X^{(0)}$ , é formada a sucessão de iteradas

$$\begin{aligned} X^{(1)} &= M^{-1}B - M^{-1}NX^{(0)} \\ X^{(2)} &= M^{-1}B - M^{-1}NX^{(1)} \\ &\vdots \\ X^{(k+1)} &= M^{-1}B - M^{-1}NX^{(k)} \\ &\vdots \end{aligned}$$

que desejavelmente deve convergir para a solução de  $AX = B$ . Veremos adiante condições que garantem a convergência desta sucessão. Antes, porém, vamos descrever os métodos de Jacobi e Gauss-Seidel que correspondem precisamente a duas diferentes escolhas da matriz  $M$ .

## 5.1 Método de Jacobi

Começemos por considerar a seguinte decomposição da matriz  $A = [a_{ij}]$  ( $i, j = 1, \dots, n$ ) dos coeficientes do sistema  $AX = B$ :

$$A = L + D + U, \quad (17)$$



onde

$$L = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ a_{21} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-2} & 0 & \\ a_{n1} & \cdots & \cdots & a_{n,n-1} & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ 0 & a_{22} & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

e

$$U = \begin{bmatrix} 0 & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

são, respectivamente, as matrizes estritamente diagonal inferior, diagonal e estritamente diagonal superior em que podemos decompor a matriz  $A$ .

O método de Jacobi baseia-se na seguinte escolha das matrizes  $M$  e  $N$  da decomposição  $A = M + N$ :

$$M = D \quad \text{e} \quad N = L + U.$$

O processo iterativo (16) assume então a forma

$$DX^{(k+1)} = B - (L + U) X^{(k)}, \quad k = 0, 1, \dots$$

Admitindo que todos os  $a_{ii} \neq 0$  (o que é sempre possível por meio de trocas de linhas e colunas desde que a matriz  $A$  seja não singular), esta igualdade é equivalente a

$$\boxed{X^{(k+1)} = D^{-1}B - D^{-1}(L + U) X^{(k)}, \quad k = 0, 1, \dots} \quad (18)$$

ou ainda a

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right), \quad i = 1, 2, \dots, n; \quad k = 0, 1, \dots,$$

considerando  $X^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ .

**Exemplo 5.1** Vamos aplicar o método de Jacobi à resolução do sistema  $AX = B$ , com

$$A = \begin{bmatrix} 5 & 1 & 1 \\ 3 & 10 & 2 \\ 1 & -1 & 3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

Começamos por deduzir a fórmula iterativa do método de Jacobi e, tomando  $X^{(0)} = (0, 0, 0)$ , calcularemos três iteradas.

Tendo em conta (18), uma vez que

$$M = D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad N = L + U = \begin{bmatrix} 0 & 1 & 1 \\ 3 & 0 & 2 \\ 1 & -1 & 0 \end{bmatrix}$$

e

$$M^{-1} = D^{-1} = \begin{bmatrix} \frac{1}{5} & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix},$$

vem

$$\begin{aligned} X^{(k+1)} &= D^{-1}B - D^{-1}(L + U)X^{(k)} \\ &= \begin{bmatrix} 0 \\ 1/10 \\ 2/3 \end{bmatrix} - \begin{bmatrix} 0 & 1/5 & 1/5 \\ 3/10 & 0 & 1/5 \\ 1/3 & -1/3 & 0 \end{bmatrix} X^{(k)}, \quad k = 0, 1, \dots \end{aligned}$$

Considerando agora  $X^{(0)} = (0, 0, 0)$ , obtém-se

$$\begin{aligned} X^{(1)} &= \begin{bmatrix} 0 \\ 1/10 \\ 2/3 \end{bmatrix} - \begin{bmatrix} 0 & 1/5 & 1/5 \\ 3/10 & 0 & 1/5 \\ 1/3 & -1/3 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{2}{3} \end{bmatrix} \\ X^{(2)} &= \begin{bmatrix} 0 \\ 1/10 \\ 2/3 \end{bmatrix} - \begin{bmatrix} 0 & 1/5 & 1/5 \\ 3/10 & 0 & 1/5 \\ 1/3 & -1/3 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} -\frac{23}{150} \\ -\frac{1}{30} \\ \frac{7}{10} \end{bmatrix} \quad \text{e} \\ X^{(3)} &= \begin{bmatrix} 0 \\ 1/10 \\ 2/3 \end{bmatrix} - \begin{bmatrix} 0 & 1/5 & 1/5 \\ 3/10 & 0 & 1/5 \\ 1/3 & -1/3 & 0 \end{bmatrix} \begin{bmatrix} -\frac{23}{150} \\ -\frac{1}{30} \\ \frac{7}{10} \end{bmatrix} = \begin{bmatrix} -\frac{2}{15} \\ \frac{3}{500} \\ \frac{53}{75} \end{bmatrix} \end{aligned}$$

isto é,

$$X^{(3)} \approx \begin{bmatrix} -0.1333 \\ 0.006 \\ 0.7067 \end{bmatrix}.$$

De notar que a solução do sistema com quatro algarismos exactos é  $X = (-0.1429, 0, 0.7143)$ . ■

## 5.2 Método de Gauss-Seidel

Considerando de novo a igualdade (17), o método de Gauss-Seidel baseia-se na seguinte escolha:

$$M = L + D \quad \text{e} \quad N = U.$$

O processo iterativo (16) escreve-se então na forma

$$(L + D) X^{(k+1)} = B - U X^{(k)}, \quad k = 0, 1, \dots, \quad (19)$$

ou, admitindo que  $M = L + D$  é invertível (condição equivalente a  $a_{ii} \neq 0, \forall i$ ),

$$\boxed{X^{(k+1)} = (L + D)^{-1} B - (L + D)^{-1} U X^{(k)}, \quad k = 0, 1, \dots} \quad (20)$$

**Exemplo 5.2** Vamos deduzir a fórmula iterativa do método de Gauss-Seidel quando aplicado à resolução do sistema de equações do exemplo 5.1. Calcularemos seguidamente três iteradas tomando  $X^{(0)} = (0, 0, 0)$ .

Tendo em conta (20), uma vez que

$$M = L + D = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 10 & 0 \\ 1 & -1 & 3 \end{bmatrix}, \quad N = U = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

e

$$M^{-1} = (L + D)^{-1} = \begin{bmatrix} \frac{1}{5} & 0 & 0 \\ -\frac{3}{50} & \frac{1}{10} & 0 \\ -\frac{13}{150} & \frac{1}{30} & \frac{1}{3} \end{bmatrix}$$

vem

$$\begin{aligned} X^{(k+1)} &= (L + D)^{-1} B - (L + D)^{-1} U X^{(k)} \\ &= \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{7}{10} \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & -\frac{3}{50} & \frac{7}{50} \\ 0 & -\frac{13}{150} & -\frac{1}{50} \end{bmatrix} X^{(k)}, \quad k = 0, 1, \dots \end{aligned}$$

Considerando agora  $X^{(0)} = (0, 0, 0)$ , obtém-se

$$\begin{aligned} X^{(1)} &= \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{7}{10} \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & -\frac{3}{50} & \frac{7}{50} \\ 0 & -\frac{13}{150} & -\frac{1}{50} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{7}{10} \end{bmatrix} \\ X^{(2)} &= \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{7}{10} \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & -\frac{3}{50} & \frac{7}{50} \\ 0 & -\frac{13}{150} & -\frac{1}{50} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{7}{10} \end{bmatrix} = \begin{bmatrix} -\frac{4}{25} \\ \frac{1}{125} \\ \frac{271}{375} \end{bmatrix} \quad \text{e} \\ X^{(3)} &= \begin{bmatrix} 0 \\ \frac{1}{10} \\ \frac{7}{10} \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & -\frac{3}{50} & \frac{7}{50} \\ 0 & -\frac{13}{150} & -\frac{1}{50} \end{bmatrix} \begin{bmatrix} -\frac{4}{25} \\ \frac{1}{125} \\ \frac{271}{375} \end{bmatrix} = \begin{bmatrix} -\frac{274}{1875} \\ -\frac{13}{18750} \\ \frac{13409}{18750} \end{bmatrix}, \end{aligned}$$

ou seja,

$$X^{(3)} \simeq \begin{bmatrix} -0.1461 \\ -0.0007 \\ 0.7152 \end{bmatrix}.$$

Vê-se assim que  $X^{(3)} = (-0.1461, -0.0007, 0.7152)$  está mais próxima da solução do sistema  $X = (-0.1429, 0, 0.7143)$  do que a solução aproximada  $X^{(3)} = (-0.1333, 0.006, 0.7067)$  obtida pelo método de Jacobi. ■

De (19) resulta também que

$$DX^{(k+1)} = B - LX^{(k+1)} - UX^{(k)}, \quad k = 0, 1, \dots$$

Supondo que  $a_{ii} \neq 0$ , deduz-se que a fórmula iterativa do método de Gauss-Seidel pode ser escrita na forma

$$X^{(k+1)} = D^{-1}(B - LX^{(k+1)} - UX^{(k)}), \quad k = 0, 1, \dots \quad (21)$$

Alternativamente, considerando  $X^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ , (21) exprime-se na forma

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad i = 1, 2, \dots, n; \quad k = 0, 1, \dots,$$

a qual evidencia uma importante característica do método de Gauss-Seidel: *cada componente  $x_i^{(k+1)}$  da aproximação  $X^{(k+1)}$  é calculada recorrendo às componentes  $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$  obtidas anteriormente na mesma iterada.* Este facto não ocorre no método de Jacobi em que o cálculo das componentes de  $X^{(k+1)}$  é feito unicamente com base nas componentes de  $X^{(k)}$ . Por esta razão o método de Jacobi é por vezes designado por método das *substituições simultâneas* enquanto o método de Gauss-Seidel é conhecido por método das *substituições sucessivas*.

O aproveitamento imediato da última componente calculada de  $X^{(k+1)}$  no cálculo da componente seguinte, propicia, como ilustra o exemplo anterior, uma convergência mais rápida do método de Gauss-Seidel quando comparado com o método de Jacobi.

### 5.3 Normas vectoriais e matriciais

Dissemos no exemplo anterior que a 3ª iterada obtida pelo método de Gauss-Seidel está mais próxima da solução do sistema do que a 3ª iterada obtida pelo método de Jacobi. A medição desta proximidade e mais geralmente dos erros das iteradas de um qualquer método iterativo requer o uso de normas. Apresentaremos seguidamente algumas das normas vectoriais e matriciais mais utilizadas no actual contexto bem como as suas principais propriedades.

Geometricamente, a norma de um vector de  $\mathbb{R}^2$  ou  $\mathbb{R}^3$  pode ser interpretada como o comprimento desse vector. Em geral, uma **norma vectorial** definida em  $\mathbb{R}^n$  é uma aplicação deste espaço em  $\mathbb{R}$  que verifica determinado conjunto de propriedades. Mais precisamente:

**Definição 5.1** *Uma norma de  $\mathbb{R}^n$  é uma aplicação que a cada vector  $\vec{x} \in \mathbb{R}^n$  faz corresponder o número real  $\|\vec{x}\|$  e que verifica os seguintes axiomas:*

1.  $\|\vec{x}\| > 0$  se  $\vec{x} \neq \vec{0}$  e  $\|\vec{0}\| = 0$ . (Positividade)
2.  $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ , para  $\vec{x}, \vec{y} \in \mathbb{R}^n$ . (Desigualdade triangular)
3.  $\|\lambda \vec{x}\| = |\lambda| \|\vec{x}\|$ , para  $\vec{x} \in \mathbb{R}^n$  e  $\lambda \in \mathbb{R}$ . (Homogeneidade)

O espaço vectorial  $\mathbb{R}^n$  munido de uma norma diz-se um **espaço vectorial normado**. Um vector de  $\mathbb{R}^n$  diz-se **unitário** ou **normalizado** quando tem norma igual a 1.

Em  $\mathbb{R}^n$  podem definir-se diversas normas. Entre estas contam-se as **normas  $l_p$**  ( $p \geq 1$ ) definidas, para  $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ , por

$$\|\vec{x}\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{\frac{1}{p}}.$$

Para  $p = 1$  obtém-se a **norma da soma** (ou **norma  $l_1$** ) em  $\mathbb{R}^n$ , dada por

$$\|\vec{x}\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|.$$

Para  $p = 2$  obtém-se a **norma euclidiana** (ou **norma  $l_2$** ) em  $\mathbb{R}^n$  definida por

$$\|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Em particular, se  $n = 1$  vem  $\vec{x} = x_1$  e a norma deste vector coincide com o módulo ou valor absoluto de  $x_1$  pois

$$\|x\|_2 = \sqrt{x_1^2} = |x_1|.$$

Caso  $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$  ou  $\vec{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$  tem-se

$$\|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2} \quad \text{ou} \quad \|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Finalmente, define-se também em  $\mathbb{R}^n$  a **norma do máximo** (ou **norma  $l_\infty$** ) por

$$\|\vec{x}\|_\infty = \max \{|x_1|, |x_2|, \dots, |x_n|\}.$$

Cada matriz quadrada de ordem  $n$  de elementos reais tem  $n^2$  elementos pelo que pode ser vista como um vector de  $\mathbb{R}^{n^2}$ . Deste modo, a norma de uma matriz quadrada poderá ser definida através de uma norma vectorial definida em  $\mathbb{R}^{n^2}$ . Contudo, o conjunto das matrizes de ordem  $n$  tem uma estrutura mas rica que a do espaço  $\mathbb{R}^{n^2}$ , dado que aí está definido uma operação de multiplicação; acontece que é muitas vezes necessário relacionar o “tamanho” do produto matricial  $AB$  com os “tamanhos” das matrizes  $A$  e  $B$ . Justifica-se, assim, a seguinte definição de **norma matricial**:

**Definição 5.2** *Uma norma no conjunto das matrizes quadradas de ordem  $n$  de elementos reais é uma aplicação que a cada matriz  $A$  de ordem  $n$  faz corresponder o número real  $\|A\|$  que verifica os seguintes axiomas:*

1.  $\|A\| > 0$  se  $A \neq \mathbf{O}$  e  $\|\mathbf{O}\| = 0$  ( $\mathbf{O}$  é a matriz nula de ordem  $n$ ). (Positividade)
2.  $\|A + B\| \leq \|A\| + \|B\|$ , para quaisquer matrizes  $A$  e  $B$  de ordem  $n$ . (Desigualdade triangular)
3.  $\|\lambda A\| = |\lambda| \|A\|$ , para toda a matriz  $A$  de ordem  $n$  e  $\lambda \in \mathbb{R}$ . (Homogeneidade)
4.  $\|AB\| \leq \|A\| \|B\|$ , para quaisquer matrizes  $A$  e  $B$  de ordem  $n$ . (Submultiplicatividade)

Observa-se que a definição de norma de uma matriz é análoga à de norma de um vector só que é obrigatória a verificação da condição 4., a qual garante que a norma matricial é **submultiplicativa**.

Vejamos alguns exemplos de normas matriciais. Uma das normas mais frequentemente usadas é a **norma de Frobenius** definida, para uma matriz  $A = [a_{ij}]$  de ordem  $n$  por

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

Cada norma vectorial definida em  $\mathbb{R}^n$  “induz” uma norma matricial no conjunto das matrizes reais de ordem  $n$ .

**Teorema 5.1** Consideremos uma norma vectorial definida em  $\mathbb{R}^n$  e seja  $A$  uma matriz real de ordem  $n$ . Então,

$$\|A\| = \max \{ \|A\vec{x}\| : \|\vec{x}\| = 1 \}$$

é uma norma matricial, que se diz **induzida** pela norma vectorial de  $\mathbb{R}^n$ . Além disto, verifica-se a desigualdade

$$\|A\vec{x}\| \leq \|A\| \|\vec{x}\|, \quad (22)$$

para qualquer  $\vec{x} \in \mathbb{R}^n$ .

A demonstração deste teorema pode ver-se em [6]. De notar que desigualdade (22) compatibiliza a utilização da norma vectorial com a norma matricial induzida. Por outro lado, a norma induzida pode ser alternativamente definida por

$$\|A\| = \max \left\{ \frac{\|A\vec{x}\|}{\|\vec{x}\|} : \vec{x} \neq \vec{0} \right\}.$$

Com efeito,

$$\|A\| = \max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|}{\|\vec{x}\|} = \max_{\vec{x} \neq \vec{0}} \left\| \frac{1}{\|\vec{x}\|} A\vec{x} \right\| = \max_{\vec{x} \neq \vec{0}} \left\| A \frac{\vec{x}}{\|\vec{x}\|} \right\| = \sup_{\|\vec{y}\|=1} \|A\vec{y}\|,$$

onde  $\vec{y} = \vec{x}/\|\vec{x}\|$  é um vector unitário. De referir ainda que da definição de norma induzida resulta imediatamente que  $\|I\| = 1$ , onde  $I$  representa a matriz identidade.

As normas matriciais induzidas pelas normas vectoriais  $l_p$  ( $p \geq 1$ ) definidas em  $\mathbb{R}^n$  são as chamadas **normas matriciais- $p$** , as quais se costumam representar por

$$\|A\|_p = \max \left\{ \|A\vec{x}\|_p : \|\vec{x}\|_p = 1 \right\}.$$

Sendo  $A = [a_{ij}]$  uma matriz de ordem  $n$  obtêm-se as seguintes concretizações da norma- $p$  (ver demonstração em [6]):

$$\text{Se } p = 1 \Rightarrow \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (\text{norma do máximo das somas por coluna}) \quad (23)$$

$$\text{Se } p = \infty \Rightarrow \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (\text{norma do máximo das somas por linha}) \quad (24)$$

Tem-se assim que  $\|A\|_1$  é a maior das somas dos valores absolutos dos elementos de cada coluna de  $A$ , enquanto  $\|A\|_\infty$  é a maior das somas dos valores absolutos dos elementos de cada linha de  $A$ .

**Exemplo 5.3** Calcular as normas  $\|A\|_1$  e  $\|A\|_\infty$  da matriz

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

De (23) e (24) deduz-se que

$$\|A\|_1 = \max (|a_{11}| + |a_{21}| + |a_{31}|, |a_{12}| + |a_{22}| + |a_{32}|, |a_{13}| + |a_{23}| + |a_{33}|)$$

e

$$\|A\|_{\infty} = \max(|a_{11}| + |a_{12}| + |a_{13}|, |a_{21}| + |a_{22}| + |a_{23}|, |a_{31}| + |a_{32}| + |a_{33}|),$$

respectivamente.

Por outro lado, sendo

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

tem-se

$$\begin{aligned} \|A\|_1 &= \max(2 + 1 + 2 + 0, 1 + 2 + 1 + 2, 2 + 1 + 2 + 0, 2 + 2 + 1 + 1) \\ &= \max(5, 6, 5, 6) = 6 \end{aligned}$$

e

$$\begin{aligned} \|A\|_{\infty} &= \max(2 + 1 + 2 + 2, 1 + 2 + 1 + 2, 2 + 1 + 2 + 1, 0 + 2 + 0 + 1) \\ &= \max(7, 6, 6, 3) = 7. \blacksquare \end{aligned}$$

## 5.4 Convergência dos métodos iterativos

Consideremos de novo a forma geral dos métodos iterativos dada em (16),

$$MX^{(k+1)} = B - NX^{(k)}, \quad k = 0, 1, \dots,$$

ou

$$X^{(k+1)} = M^{-1}B - M^{-1}NX^{(k)}, \quad k = 0, 1, \dots,$$

supondo a invertibilidade de  $M$ . Assim, considerando  $G = -M^{-1}N$  e  $C = M^{-1}B$ , o processo iterativo anterior pode escrever-se na forma

$$X^{(k+1)} = GX^{(k)} + C, \quad k = 0, 1, \dots \quad (25)$$

A matriz  $G$  diz-se a **matriz de iteração** do processo iterativo.

Enunciamos seguidamente uma condição suficiente de convergência do processo iterativo (25) bem como uma estimativa da norma do erro de cada iterada.

**Teorema 5.2** *Suponhamos que  $\|G\| < 1$  para alguma norma matricial induzida por uma norma vectorial de  $\mathbb{R}^n$ . Então, qualquer que seja a aproximação inicial  $X^{(0)}$ , o método iterativo (25) converge para a solução  $X$  do sistema  $AX = B$ .*

*Além disto, para  $k = 0, 1, \dots$ , é válida a seguinte estimativa da norma do erro da iterada  $X^{(k+1)}$ :*

$$\|X - X^{(k+1)}\| \leq \frac{\|G\|}{1 - \|G\|} \|X^{(k+1)} - X^{(k)}\|.$$

**Demonstração:** Visto que,  $X^{(k+1)} = GX^{(k)} + C$  para  $k = 0, 1, \dots$ , tem-se

$$X^{(k+1)} - X^{(k)} = GX^{(k)} + C - (GX^{(k-1)} + C) = G(X^{(k)} - X^{(k-1)}).$$

Por indução facilmente se conclui que, para  $k = 1, 2, \dots$ ,

$$X^{(k+1)} - X^{(k)} = G^k(X^{(1)} - X^{(0)})$$

pelo que, para qualquer inteiro  $m > 0$ ,

$$\begin{aligned} X^{(k+m)} - X^{(k)} &= \left( X^{(k+m)} - X^{(k+m-1)} \right) + \dots + \left( X^{(k+1)} - X^{(k)} \right) \\ &= \left( G^{k+m-1} + \dots + G^k \right) \left( X^{(1)} - X^{(0)} \right). \end{aligned}$$

Considerando uma norma vectorial de  $\mathbb{R}^n$  e a respectiva norma matricial induzida, tem-se então

$$\begin{aligned} \left\| X^{(k+m)} - X^{(k)} \right\| &\leq \left( \|G\|^{k+m-1} + \dots + \|G\|^k \right) \left\| X^{(1)} - X^{(0)} \right\| \\ &\leq \frac{\|G\|^k}{1 - \|G\|} \left\| X^{(1)} - X^{(0)} \right\| \end{aligned} \quad (26)$$

atendendo a que  $\sum_{t=0}^{m-1} \|G\|^t \leq \frac{1}{1 - \|G\|}$ , pois  $\|G\| < 1$ . Esta última desigualdade e (26) garantem que, para qualquer  $\delta > 0$ ,

$$\left\| X^{(k+m)} - X^{(k)} \right\| < \delta$$

para  $k$  suficientemente grande, pelo que a sucessão  $X^{(k)}$ ,  $k \geq 0$ , é de Cauchy e, portanto, convergente. Sendo  $X = \lim_{k \rightarrow \infty} X^{(k)}$ , tem-se

$$GX^{(k)} + C = G \lim_{k \rightarrow \infty} X^{(k)} + C,$$

ou seja,

$$X = GX + C. \quad (27)$$

Fica pois provado que processo iterativo converge para a solução de  $AX = B$  dado que este sistema e (27) são equivalentes.

Finalmente, subtraindo a (27) a equação (25) obtém-se

$$X - X^{(k+1)} = G \left( X - X^{(k)} \right), \quad k = 0, 1, \dots,$$

donde sai, considerando uma norma vectorial e a respectiva norma matricial induzida,

$$\begin{aligned} \left\| X - X^{(k+1)} \right\| &\leq \|G\| \left\| X - X^{(k)} \right\| = \|G\| \left\| X - X^{(k+1)} + X^{(k+1)} - X^{(k)} \right\| \\ &\leq \|G\| \left\| X - X^{(k+1)} \right\| + \|G\| \left\| X^{(k+1)} - X^{(k)} \right\|, \end{aligned}$$

ou seja,

$$\left\| X - X^{(k+1)} \right\| \leq \frac{\|G\|}{1 - \|G\|} \left\| X^{(k+1)} - X^{(k)} \right\|,$$

o que prova o pretendido. ■

#### 5.4.1 Aplicação ao método de Jacobi

No método de Jacobi considera-se  $M = D$ , pelo que atendendo a (??) e a (25) resulta que a matriz de iteração e o vector  $C$  (designados agora respectivamente por  $G_J$  e  $C_J$ ) são dados por

$$G_J = -M^{-1}N = D^{-1}(L + U) \quad \text{e} \quad C_J = M^{-1}B = D^{-1}B.$$



Portanto, se  $\|G_J\| < 1$  para alguma norma matricial induzida por uma norma vectorial, o método é convergente.

Visto que

$$G_J = \begin{bmatrix} 0 & -\frac{a_{21}}{a_{11}} & \dots & \dots & -\frac{a_{n1}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & -\frac{a_{n-1,n}}{a_{n-1,n-1}} \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & -\frac{a_{n,n-1}}{a_{nn}} & 0 \end{bmatrix},$$

conclui-se que a norma  $\|G_J\|_\infty$  induzida pela norma do máximo vem dada por

$$\|G_J\|_\infty = \max_{1 \leq i \leq n} \left( \sum_{j=1, j \neq i}^n \left| \frac{a_{ij}}{a_{ii}} \right| \right).$$

Assim, ter-se-á  $\|G_J\|_\infty < 1$  desde que

$$\sum_{j=1, j \neq i}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1, \quad \forall i = 1, \dots, n,$$

o que equivale a

$$\sum_{j=1, j \neq i}^n |a_{ij}| < |a_{ii}|, \quad \forall i = 1, \dots, n,$$

isto é, os elementos da diagonal principal são, em valor absoluto, superiores à soma dos valores absolutos dos elementos da respectiva linha. Uma matriz  $A = [a_{ij}]$  que satisfaça estas condições diz-se de **diagonal estritamente dominante por linhas**. Podemos pois enunciar:

**Teorema 5.3** *Se  $A$  for de diagonal estritamente dominante por linhas, o método de Jacobi converge, independentemente da aproximação inicial  $X^{(0)}$  escolhida.*

**Exemplo 5.4** Mostremos que o método de Jacobi aplicado ao sistema do exemplo 5.1 (pág. 24) converge e calculemos uma estimativa da norma do erro da iterada  $X^{(3)}$  aí obtida.

Verifiquemos que a matriz

$$A = \begin{bmatrix} 5 & 1 & 1 \\ 3 & 10 & 2 \\ 1 & -1 & 3 \end{bmatrix}$$

é de diagonal estritamente dominante por linhas. Com efeito, tem-se

$$i = 1 \Rightarrow \sum_{j=1, j \neq 1}^3 |a_{1j}| = |1| + |1| = 2 < |5| = a_{11},$$

$$i = 2 \Rightarrow \sum_{j=1, j \neq 2}^3 |a_{2j}| = |3| + |2| = 5 < |10| = a_{22} \quad \text{e}$$

$$i = 3 \Rightarrow \sum_{j=1, j \neq 3}^3 |a_{3j}| = |1| + |-1| = 2 < |3| = a_{33},$$

pelo que

$$\|G_J\|_\infty = \max \left\{ \frac{2}{5}, \frac{5}{10}, \frac{2}{3} \right\} = \frac{2}{3} < 1.$$

Portanto, o método de Jacobi converge.

Para calcular uma estimativa da norma do erro de  $X^{(3)}$  vamos usar a fórmula dada no teorema 5.2 com a norma vectorial do máximo. Assim, tem-se

$$\begin{aligned} \|X - X^{(3)}\|_\infty &\leq \frac{\|G_J\|_\infty}{1 - \|G_J\|_\infty} \|X^{(3)} - X^{(2)}\|_\infty \\ &= \frac{2/3}{1 - 2/3} \max \left\{ \left| \frac{1}{50} \right|, \left| \frac{59}{1500} \right|, \left| \frac{1}{150} \right| \right\} \approx 0.0787. \blacksquare \end{aligned}$$

#### 5.4.2 Aplicação ao método de Gauss-Seidel

No método de Gauss-Seidel a matriz de iteração (que designaremos por  $G_{GS}$ ) é dada por

$$G_{GS} = -M^{-1}N = -(L + D)^{-1}U.$$

Não é imediato indicar a forma dos elementos da matriz de iteração com vista a calcular  $\|G_{GS}\|_\infty$ . No entanto, prova-se um resultado análogo ao do teorema 5.3, podendo a demonstração ser vista em [1, 8].

**Teorema 5.4** *Se  $A$  for de diagonal estritamente dominante por linhas, o método de Gauss-Seidel converge, independentemente da aproximação inicial  $X^{(0)}$  escolhida.*

**Exemplo 5.5** Voltando ao sistema do exemplo 5.4, temos que o método de Gauss-Seidel é igualmente convergente pois, como aí vimos, a matriz  $A$  é de diagonal estritamente dominante por linhas. Calculemos agora uma estimativa da norma do erro da iterada  $X^{(3)}$  obtida no exemplo 5.2 (pág. 26) por aplicação do método de Gauss-Seidel ao mesmo sistema.

Com efeito, tem-se

$$G_{SD} = -M^{-1}N = \begin{bmatrix} 0 & -\frac{1}{5} & -\frac{1}{5} \\ 0 & \frac{3}{50} & -\frac{7}{50} \\ 0 & \frac{13}{150} & \frac{1}{50} \end{bmatrix}$$

pelo que

$$\|G_{GS}\|_\infty = \max \left\{ \frac{2}{5}, \frac{1}{5}, \frac{8}{75} \right\} = \frac{2}{5} < 1$$

e

$$\begin{aligned} \|X - X^{(3)}\|_\infty &\leq \frac{\|G_{GS}\|_\infty}{1 - \|G_{GS}\|_\infty} \|X^{(3)} - X^{(2)}\|_\infty \\ &\lesssim \frac{2/5}{1 - 2/5} \max \{ |0.0139|, |-0.0087|, |-0.0075| \} \approx 0.0092. \end{aligned}$$

Verifica-se assim que a estimativa da norma do erro da iterada  $X^{(3)}$  obtida pelo método de Gauss-Seidel é cerca de oito vezes e meia inferior à estimativa homóloga calculada no método de Jacobi. Isto ilustra a maior velocidade de convergência que em geral o método de Gauss-Seidel revela quando comparado com o método de Jacobi.

**Exercício 5.1** Considere o sistema de equações lineares  $AX = B$  com

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & a & 1/4 \\ 1/3 & 1/4 & a \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

em que  $a$  é um parâmetro real.

1. (a) Enuncie, justificando, uma condição sobre os elementos de  $A$  que garanta a convergência dos métodos iterativos de Jacobi e Gauss-Seidel;
- (b) Para  $a = 1$ , determine duas iterações do método de Gauss-Seidel tomando  $X^{(0)} = M^{-1}B$  para aproximação inicial. Indique uma estimativa da norma do erro da última iterada obtida.

**Resolução:** (a) Para assegurar a convergência dos métodos iterativos indicados basta que a matriz  $A$  seja de diagonal estritamente dominante por linhas, isto é,

$$\sum_{j=1, j \neq i}^3 |a_{ij}| < |a_{ii}|, \quad i = 1, \dots, 3.$$

Então, dado que a primeira linha já satisfaz este critério pois,

$$\sum_{j=1, j \neq 1}^3 |a_{1j}| = \left| \frac{1}{2} \right| + \left| \frac{1}{3} \right| = \frac{5}{6} < 1 = |a_{11}|,$$

deverá ter-se adicionalmente

$$\begin{aligned} |a_{22}| &= |a| > \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \quad \text{e} \\ |a_{33}| &= |a| > \frac{1}{3} + \frac{1}{4} = \frac{7}{12}. \end{aligned}$$

Assim, aquele critério fica satisfeito desde que

$$|a| > \frac{3}{4}.$$

(b) Começemos por determinar as matrizes necessárias à implementação do método de Gauss-Seidel. Sucessivamente, tem-se

$$M = L + D = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/3 & 1/4 & 1 \end{bmatrix}, \quad N = U = \begin{bmatrix} 0 & 1/2 & 1/3 \\ 0 & 0 & 1/4 \\ 0 & 0 & 0 \end{bmatrix}$$

e

$$M^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -5/24 & -1/4 & 1 \end{bmatrix}.$$

Assim,

$$\begin{aligned}
X^{(k+1)} &= M^{-1}B - M^{-1}NX^{(k)} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -5/24 & -1/4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -5/24 & -1/4 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1/2 & 1/3 \\ 0 & 0 & 1/4 \\ 0 & 0 & 0 \end{bmatrix} X^{(k)} \\
&= \begin{bmatrix} 1 \\ -1/2 \\ 19/24 \end{bmatrix} - \begin{bmatrix} 0 & 1/2 & 1/3 \\ 0 & -1/4 & 1/12 \\ 0 & -5/48 & -19/144 \end{bmatrix} X^{(k)}.
\end{aligned}$$

Calculemos agora duas iterações supondo  $X^{(0)} = M^{-1}B = \begin{bmatrix} 1 \\ -1/2 \\ 19/24 \end{bmatrix}$ :

$$\begin{aligned}
X^{(1)} &= \begin{bmatrix} 1 \\ -1/2 \\ 19/24 \end{bmatrix} - \begin{bmatrix} 0 & 1/2 & 1/3 \\ 0 & -1/4 & 1/12 \\ 0 & -5/48 & -19/144 \end{bmatrix} \begin{bmatrix} 1 \\ -1/2 \\ 19/24 \end{bmatrix} \\
&= \begin{bmatrix} 71/72 \\ -199/288 \\ 2917/3456 \end{bmatrix}
\end{aligned}$$

e

$$\begin{aligned}
X^{(2)} &= \begin{bmatrix} 1 \\ -1/2 \\ 19/24 \end{bmatrix} - \begin{bmatrix} 0 & 1/2 & 1/3 \\ 0 & -1/4 & 1/12 \\ 0 & -5/48 & -19/144 \end{bmatrix} \begin{bmatrix} 71/72 \\ -199/288 \\ 2917/3456 \end{bmatrix} \\
&= \begin{bmatrix} \frac{11033}{10368} \\ -\frac{30817}{41472} \\ \frac{413587}{497664} \end{bmatrix},
\end{aligned}$$

ou seja,  $X^{(2)} \approx (1.0641, -0.7431, 0.8311)$ .

Para calcular uma estimativa da norma do erro da iterada  $X^{(2)}$ , como

$$G_{GS} = -M^{-1}N = \begin{bmatrix} 0 & -1/2 & -1/3 \\ 0 & 1/4 & -1/12 \\ 0 & 5/48 & 19/144 \end{bmatrix},$$

a norma do máximo das somas por linhas de  $G_{GS}$  é dada por

$$\|G_{GS}\|_{\infty} = \max \left\{ \frac{1}{2} + \frac{1}{3}, \frac{1}{4} + \frac{1}{12}, \frac{5}{48} + \frac{19}{144} \right\} = \frac{5}{6} < 1.$$

Assim, utilizando a fórmula do teorema 5.2,

$$\begin{aligned}
\|X - X^{(2)}\|_{\infty} &\leq \frac{\|G_{GS}\|_{\infty}}{1 - \|G_{GS}\|_{\infty}} \|X^{(2)} - X^{(1)}\|_{\infty} \\
&= \frac{5/6}{1 - 5/6} \max \{ |0.0780|, |-0.0521|, |-0.0130| \} \approx 0.39.
\end{aligned}$$

## Referências

- [1] Atkinson, K. E., *An Introduction to Numerical Analysis*, John Wiley and Sons, 1978.
- [2] Carpentier, M. P. J., *Análise Numérica - Teoria*, Associação dos Estudantes do IST, 1993.
- [3] Conte, S. D. e De Boor, C., *Elementary Numerical Analysis*, McGraw-Hill, 1983.
- [4] Dahlquist e Bjork, *Numerical Methods*, Prentice Hall, 1974.
- [5] Dias Agudo, F. R., *Introdução à Álgebra Linear e Geometria Analítica*, 4<sup>a</sup> edição, Escolar Editora, 1996,
- [6] Horn, R. A. e Johnson, C. R., *Matrix Analysis*, Cambridge University Press, 1985
- [7] Kreyzig, E., *Advanced Engineering Mathematics* (Caps. 17 e 18), Wiley, 1999.
- [8] Pina, H., *Métodos Numéricos*, McGraw-Hill, 1995.
- [9] Valença, M. R., *Métodos Numéricos*, 3<sup>a</sup> edição, Livraria Minho, 1993.