

Tema 6: Métodos directos para la resolución de S.E.L.

1 Descomposición LR

Sea $A \in \mathcal{M}_n$ tal que $\det(A) \neq 0$, es decir, el S. E. L. $(S) : Ax = \mathbf{b}$ es compatible determinado. Supongamos que **el método de Gauss puede ser aplicado a (S) sin necesidad de intercambiar filas**. En estas condiciones el método de Gauss puede describirse como sigue:

Paso 1: Si escribimos $A = A^{(1)}$ y $\mathbf{b} = \mathbf{b}^{(1)}$ existen $m_2^{(1)}, \dots, m_{n-1}^{(1)}, m_n^{(1)} \in \mathbb{R}$ tales que el sistema $(S) : A^{(1)}\mathbf{x} = \mathbf{b}^{(1)}$ es equivalente a

$$(S_2) : A^{(2)}\mathbf{x} = \mathbf{b}^{(2)}$$

siendo

$$\begin{aligned} A^{(2)} &= F_{n1}(m_n^{(1)}) \cdot F_{(n-1)1}(m_{n-1}^{(1)}) \cdots F_{21}(m_2^{(1)}) \cdot A^{(1)} \\ \mathbf{b}^{(2)} &= F_{n1}(m_n^{(1)}) \cdot F_{(n-1)1}(m_{n-1}^{(1)}) \cdots F_{21}(m_2^{(1)}) \cdot \mathbf{b}^{(1)} \end{aligned}$$

$A^{(2)}$ tiene ceros bajo la diagonal en su primera columna y, si definimos

$$F^{(1)} = F_{n1}(m_n^{(1)}) \cdot F_{(n-1)1}(m_{n-1}^{(1)}) \cdots F_{21}(m_2^{(1)}) \text{ entonces}$$

$$F^{(1)} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ m_2^{(1)} & 1 & 0 & \dots & 0 \\ m_3^{(1)} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_n^{(1)} & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$\text{y } A^{(2)} = F^{(1)}A^{(1)}, \mathbf{b}^{(2)} = F^{(1)}\mathbf{b}^{(1)}.$$

Paso k: Supongamos construido el sistema

$$(S_k) : A^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$$

tal que $A^{(k)}$ tiene ceros bajo la diagonal en sus primeras $k-1$ columnas. Puesto que podemos aplicar el método de Gauss sin intercambiar filas, existen $m_{k+1}^{(k)}, \dots, m_{n-1}^{(k)}, m_n^{(k)} \in \mathbb{R}$ tales que el sistema

$$(S_k) : A^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$$

es equivalente a

$$(S_{k+1}) : A^{(k+1)}\mathbf{x} = \mathbf{b}^{(k+1)}$$

siendo

$$\begin{aligned} A^{(k+1)} &= F_{nk}(m_n^{(k)}) \cdot F_{(n-1)k}(m_{n-1}^{(k)}) \cdots F_{(k+1)k}(m_{k+1}^{(k)}) \cdot A^{(k)} \\ \mathbf{b}^{(k+1)} &= F_{nk}(m_n^{(k)}) \cdot F_{(n-1)k}(m_{n-1}^{(k)}) \cdots F_{(k+1)k}(m_{k+1}^{(k)}) \cdot \mathbf{b}^{(k)} \end{aligned}$$

$A^{(k+1)}$ tiene ceros bajo la diagonal en sus k primeras columnas y si definimos

$$F^{(k)} = F_{nk}(m_n^{(k)}) \cdot F_{(n-1)k}(m_{n-1}^{(k)}) \cdots F_{(k+1)k}(m_{k+1}^{(k)})$$

entonces

$$F^{(k)} = \begin{bmatrix} 1 & \dots & 0 & m_n^{(k)} & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 0 \\ 0 & \dots & 0 & m_{k+1}^{(k)} & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & m_n^{(k)} & \dots & 1 \end{bmatrix}$$

$$\text{y } A^{(k+1)} = F^{(k)} A^{(k)}, \mathbf{b}^{(k+1)} = F^{(k)} \mathbf{b}^{(k)}.$$

Al cabo de $n - 1$ pasos obtenemos un sistema

$$(S_n) : A^{(n)} \mathbf{x} = \mathbf{b}^{(n)}$$

triangular superior tal que $A^{(n)} = F^{(n-1)} F^{(n-2)} \dots F^{(1)} A$. Sea

$$T = F^{(n-1)} F^{(n-2)} \dots F^{(1)}.$$

Claramente, T es triangular inferior con 1's en la diagonal, por serlo cada una de las $F^{(i)}$, luego $L = T^{-1}$ también es triangular inferior con 1's en la diagonal. De hecho,

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -m_2^{(1)} & 1 & 0 & \dots & 0 \\ -m_3^{(1)} & -m_3^{(2)} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_n^{(1)} & -m_n^{(2)} & -m_n^{(3)} & \dots & 1 \end{bmatrix}$$

Sea ahora, $R = A^{(n)}$. Entonces,

$$R = TA \quad \Longleftrightarrow \quad A = LR,$$

siendo R regular y triangular superior y L triangular inferior con unos en la diagonal. Hemos probado así, el siguiente

Teorema. 1.1 Dado un S. E. L. $(S) : A\mathbf{x} = \mathbf{b}$ tal que $A \in \mathcal{M}_n$ y $\det(A) \neq 0$ si podemos aplicar el método de Gauss a (S) sin necesidad de intercambiar filas, existe una matriz triangular inferior con diagonal unidad, L , y una matriz regular y triangular superior, R , tal que

$$A = LR.$$

Decimos en tal caso que A posee factorización (o descomposición) LR .

Teorema. 1.2 Sea A una matriz cuadrada regular. Si existe la descomposición LR de A entonces es única.

Nota. 1.3 Dado un sistema de ecuaciones lineales $(S) : A\mathbf{x} = \mathbf{b}$, si la matriz A posee factorización LR entonces podemos resolverlo del siguiente modo:

- 1) Calculamos la factorización LR de A .
- 2) Resolvemos el sistema $L\mathbf{y} = \mathbf{b}$.
- 3) Una vez obtenida la solución $\mathbf{y}$, resolvemos el sistema $R\mathbf{x} = \mathbf{y}$.

Esto proporciona una solución de (S) , ya que tenemos

$$A\mathbf{x} = \mathbf{b} \quad \Longleftrightarrow \quad L \underbrace{R\mathbf{x}}_{=\mathbf{y}} = \mathbf{b} \quad \Longleftrightarrow \quad L\mathbf{y} = \mathbf{b}.$$

La ventaja de este método, a la hora de resolver varios S. E. L.,

$$(SL_1) : A\mathbf{x} = \mathbf{b}_1, \dots, (SL_r) : A\mathbf{x} = \mathbf{b}_r$$

con la misma matriz de coeficientes, A , es evidente, puesto que, una vez calculada la factorización LR de A , lo que puede hacerse de una vez por todas, resolver cada uno de los sistemas (SL_i) , sólo requiere una aplicación del método de bajada (para resolver $L\mathbf{y} = \mathbf{b}_i$) y una aplicación del método de subida (para resolver $R\mathbf{x} = \mathbf{y}$), es decir, sólo del orden de $2n^2$ operaciones para cada uno de ellos, una vez factorizada A .

Número de operaciones.- Veamos el número de operaciones necesarias para resolver $L\mathbf{y} = \mathbf{b}$.

Para despejar y_i de $a_{i1}y_1 + \dots + a_{ii}y_i = b_i$ hacen falta $i-1$ multiplicaciones, $i-2$ sumas, 1 resta y 1 división, luego en total tenemos $2(i-1)+1$. Por tanto $\sum_{i=1}^n (2(i-1)+1) = 2(n-1)n/2 + n = n^2$. Ahora bien, como los a_{ii} son 1, nos ahorramos n divisiones, luego necesitamos $n^2 - n$.

Para resolver $R\mathbf{x} = \mathbf{y}$ ya se vió en el tema 4 que hacían falta n^2 operaciones.

Por lo tanto para resolver cada nuevo sistema necesitamos $2n^2 - n$ operaciones en lugar de las $(4n^3 + 9n^2 - 7n)/6$ necesarias por el método de Gauss (ver tema 4).

A continuación veremos condiciones suficientes para garantizar la existencia de la descomposición LR .

Definición. 1.4 Sea $A = (a_{ij}) \in \mathcal{M}_n$, llamaremos submatriz fundamental de A de orden k , ($1 \leq k \leq n$), a la matriz

$$A_k = \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix}$$

Teorema. 1.5 Sea A una matriz regular de orden n . Son equivalentes:

- 1) Existe la descomposición LR de A .
- 2) $\det(A_k) \neq 0$, $\forall k = 1, \dots, n$.

Definición. 1.6 Una matriz $A \in \mathcal{M}_n$ simétrica se denomina definida positiva (resp. semidefinida positiva) si:

$$\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}, \quad \mathbf{x}^t A \mathbf{x} > 0 \text{ (resp. } \mathbf{x}^t A \mathbf{x} \geq 0 \text{)}.$$

Teorema. 1.7 Sea $A \in \mathcal{M}_n$ simétrica. Son equivalentes:

- 1) A es definida positiva.
- 2) $\det(A_k) > 0$, $\forall k = 1, \dots, n$.

Nota. 1.8 De paso hemos demostrado que A simétrica es definida positiva si y solo si todos los pivotes (sin intercambio de filas) son mayores que 0.

Del teorema anterior se sigue:

Teorema. 1.9 Si A es simétrica definida positiva entonces existe la descomposición LR de A .

A continuación veremos como obtener la descomposición LR de una matriz de manera directa:
Cálculo directo de la factorización LR :

Sea $A \in \mathcal{M}_n$ tal que existen $L = (l_{ij}) \in \mathcal{M}_n$ triangular inferior con diagonal unidad y $R = (r_{ij}) \in \mathcal{M}_n$ regular y triangular superior tales que $A = LR$, es decir,

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

Multiplicando la primera fila de L por R obtenemos:

$$r_{1k} = a_{1k}, \quad \forall k = 1, \dots, n.$$

Además, $\forall i = 1, \dots, n$, $l_{i1}r_{11} = a_{i1}$, luego

$$\forall i = 1, \dots, n, \quad l_{i1} = \frac{a_{i1}}{a_{11}}.$$

Multiplicando la segunda fila de L por R :

$$\forall i = 2, \dots, n, \quad l_{21}r_{1i} + r_{2i} = a_{2i}$$

lo que nos permite obtener la segunda fila de R

$$\forall i = 2, \dots, n, \quad r_{2i} = a_{2i} - l_{21}r_{1i}.$$

Además, $\forall i = 3, \dots, n$, $l_{i1}r_{12} + l_{i2}r_{22} = a_{i2}$, luego

$$\forall i = 3, \dots, n, \quad l_{i2} = (a_{i2} - l_{i1}r_{12}) \frac{1}{r_{22}}.$$

En general,

$$\forall p = 2, \dots, n, \quad \forall k = p, \dots, n, \quad r_{pk} = a_{pk} - \sum_{i=1}^{p-1} l_{pi}r_{ik}.$$

lo que nos da la p -ésima fila de R , y la p -ésima columna de L :

$$\forall p = 2, \dots, n, \quad \forall k = p+1, \dots, n, \quad l_{kp} = \frac{1}{r_{pp}}(a_{kp} - \sum_{i=1}^{p-1} l_{ki}r_{ip}).$$

2 Factorización de Cholesky

Definición. 2.1 Diremos que una matriz $A \in \mathcal{M}_n$, simétrica, posee factorización de Cholesky si existe una matriz $B \in \mathcal{M}_n$, triangular inferior con elementos positivos en la diagonal, tal que

$$A = BB^t.$$

Proposición. 2.2 La factorización de Cholesky, si existe, es única.

Teorema. 2.3 Si $A \in \mathcal{M}_n$ es simétrica definida positiva, entonces posee descomposición de Cholesky.

Cálculo directo de la factorización de Cholesky:

Sea $A = (a_{ij}) \in \mathcal{M}_n$ simétrica definida positiva y $B = (b_{ij}) \in \mathcal{M}_n$ triangular inferior con diagonal positiva tal que $A = BB^t$. Entonces,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{21} & \dots & b_{n1} \\ 0 & b_{22} & \dots & b_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_{nn} \end{bmatrix}$$

Por tanto, $\forall p = 2, \dots, n$, $\forall k = p+1, \dots, n$, tenemos

$$b_{pp} = (a_{pp} - \sum_{i=1}^{p-1} b_{pi}^2)^{1/2} \quad \text{y} \quad b_{kp} = \frac{1}{b_{pp}}(a_{kp} - \sum_{i=1}^{p-1} b_{pi}b_{ki})$$

3 Aritmética de punto flotante. Errores de redondeo

El punto flotante es un método para representar números en la memoria de un ordenador. Básicamente consiste en lo siguiente:

Un número $N \in \mathbb{R}$ se representa por

$$\pm f \cdot 10^e$$

donde:

f se denomina **mantisa** y, normalmente, aunque dependiendo del ordenador, está formado por 16 dígitos, siendo $0 \leq f < 1$.

e se denomina **característica** y está comprendido, dependiendo de cada ordenador, entre -308 y 308 .

Obviamente, este método tiene varias limitaciones:

- 1) Sólo pueden representarse una cantidad finita de números reales (de hecho, sólo podemos representar números racionales).
- 2) Hay números que no pueden ser representados de manera exacta, con los consiguientes problemas a la hora de realizar cálculos de manera efectiva.

Para evitar este tipo de problemas suele aumentarse, si es posible, el tamaño de la mantisa, o emplear el redondeo a una cantidad fija de dígitos en la mantisa, a la hora de almacenar números en la memoria. Una forma de redondear a t dígitos es la siguiente:

Dado un número real N con más de t dígitos, si el dígito $t + 1$ de N es 5, 6, 7, 8 ó 9, se aumenta el dígito t en una unidad y se desprecian los restantes, a partir de t . Si el dígito $t + 1$ es 0, 1, 2, 3 ó 4 despreciamos todos los dígitos a partir de t .

Este modo de proceder provoca algunos resultados indeseados cuando se realiza una gran cantidad de operaciones, ya que los errores de redondeo se acumulan alejando mucho el resultado final del que se habría obtenido si los cálculos hubiesen sido exactos. Como ejemplo consideremos el sistema:

$$(S) : \begin{cases} x_1 - x_2 = 0 \\ 0.01x_1 + x_2 = 1 \end{cases}$$

(S) es compatible determinado y su solución exacta es

$$x_1 = x_2 = \frac{1}{1.01} \approx 0.9$$

Si al resolverlo utilizamos un ordenador que trabaja en punto flotante redondeando a 2 dígitos, obtendremos como solución, aplicando el método de Gauss,

$$x_1 = x_2 = 1$$

la cual es una solución relativamente aceptable. Sin embargo, intercambiando las ecuaciones de (S) , obtendríamos el sistema:

$$\begin{cases} 0.01x_1 + x_2 = 1 \\ x_1 - x_2 = 0 \end{cases}$$

y, al resolverlo utilizando de nuevo el método de Gauss, hallaríamos como solución

$$x_1 = 0 \quad x_2 = 1$$

que está bastante lejos de la solución exacta. La clave de esta fuerte diferencia se encuentra en que al resolver el segundo sistema hemos utilizado un pivote, 0.01, muy pequeño.

4 Método de Gauss con pivote parcial y total

En muchas ocasiones, al aplicar el método de Gauss a la resolución de un sistema de ecuaciones resulta necesario el intercambio de filas. Además, como hemos visto en la sección anterior, una elección de pivotes adecuada puede reducir notablemente los errores de redondeo. Por ello, en la práctica se utilizan varias estrategias a la hora de elegir los pivotes en el método de Gauss. Básicamente, tenemos dos posibilidades:

- a) **Pivote parcial:** Se trata de aplicar el método de Gauss eligiendo como pivote, en cada etapa, el elemento que posee mayor valor absoluto de entre los que se encuentran bajo la diagonal en la correspondiente columna. Esto obliga en general a realizar intercambios de filas.
- b) **Pivote total:** Se trata de elegir como pivote, en cada etapa k , el elemento de mayor valor absoluto de entre los que se encuentran en la submatriz determinada por la intersección de las k últimas filas y las k últimas columnas. En consecuencia, debemos realizar intercambios de filas y de columnas.

Los siguientes resultados nos proporcionan la base teórica necesaria para estos métodos:

Teorema. 4.1 (del pivote parcial)

Sea $A \in \mathcal{M}_n$ regular. Existe una matriz de permutación $P \in \mathcal{M}_n$ tal que PA posee factorización LR .

Nota. 4.2 En general, dada $A \in \mathcal{M}_n$, si tenemos $PA = LR$, resolver el sistema $(S) : Ax = b$ es equivalente a resolver

$$(S') : PAx = Pb,$$

lo cual puede hacerse mediante la descomposición LR que posee PA ,

$$LRx = Pb.$$

Teorema. 4.3 (del pivote total)

Dada $A \in \mathcal{M}_n$ regular, existen $P, Q \in \mathcal{M}_n$ matrices de permutación tales que PAQ posee factorización LR .

Nota. 4.4 Dado el sistema $(S) : Ax = b$, si $PAQ = LR$ podemos resolverlo del siguiente modo:

- 1) Resolvemos $LRy = Pb$.
- 2) Una vez obtenida y calcular $x = Qy$.

Esto nos proporciona la solución de (S) ya que:

$$Ax = b \iff PAx = Pb \iff PAQy = Pb.$$