

donde $\tilde{A} = (\tilde{a}_{ij})_{1 \leq i, j \leq n}$, con los \tilde{a}_{ij} números complejos, y \tilde{b} es un vector de n componentes que son también números complejos.

En tal caso, la o las soluciones u que se buscan son también vectores cuyas n componentes son números complejos.

El sistema (1.4) puede ser escrito como uno de $2n$ ecuaciones con coeficientes reales y $2n$ incógnitas reales. Para ello, basta tener en cuenta que la matriz \tilde{A} y el vector \tilde{b} pueden ser escritos de manera única en la forma

$$\tilde{A} = A^1 + A^2i, \quad \tilde{b} = b^1 + b^2i,$$

siendo A^1 y A^2 matrices $n \times n$ de coeficientes reales, y b^1 y b^2 vectores de \mathbb{R}^n .

Entonces, tomando $u = u^1 + u^2i$, con u^1 y u^2 vectores incógnitas de \mathbb{R}^n , el sistema (1.4) se escribe

$$(A^1 + A^2i)(u^1 + u^2i) = b^1 + b^2i,$$

es decir,

$$A^1u^1 - A^2u^2 + (A^1u^2 + A^2u^1)i = b^1 + b^2i,$$

sistema que es equivalente a

$$\left(\begin{array}{c|c} A^1 & -A^2 \\ \hline A^2 & A^1 \end{array} \right) \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} = \begin{pmatrix} b^1 \\ b^2 \end{pmatrix}.$$

A partir de ahora, todas las matrices que consideremos en este Tema serán matrices reales, es decir, matrices de números reales. Recordemos que una matriz $n \times n$, A , se dice que es regular si su determinante $|A|$ es distinto de cero, lo cual recordemos que es equivalente a la existencia de la matriz inversa A^{-1} . A este respecto, tenemos el resultado siguiente:

Proposición 1.2 Sea $A = (a_{ij})_{1 \leq i, j \leq n}$ una matriz cuadrada $n \times n$ de coeficientes reales. El sistema de ecuaciones lineales $Au = b$ posee solución para todo $b \in \mathbb{R}^n$ si y sólo si la matriz A es regular, y en tal caso la solución es única y viene dada por $u = A^{-1}b$.

Demostración Evidentemente, si A es regular entonces existe la matriz inversa A^{-1} , y por tanto $Au = b \Leftrightarrow A^{-1}Au = A^{-1}b \Leftrightarrow u = A^{-1}b$.

Recíprocamente, si para todo $b \in \mathbb{R}^n$ existe al menos una solución de $Au = b$, sean e^k , $1 \leq k \leq n$ los n vectores de la base canónica de \mathbb{R}^n . Por hipótesis, para cada k existe al menos una solución $u^k \in \mathbb{R}^n$ del sistema $Au^k = e^k$. Entonces, si consideramos la matriz $n \times n$, que denotaremos por C , cuya k -ésima columna sea u^k , es decir, $C = (u^1 | u^2 | \dots | u^n)$, evidentemente $AC = I_n$, donde por I_n denotamos a la matriz identidad $n \times n$, pero entonces $C = A^{-1}$, y por tanto A es regular. ■

A partir de ahora, **siempre supondremos que A es regular**. Así pues, para cada $b \in \mathbb{R}^n$ dado existirá una y sólo una solución, que se puede hallar, por ejemplo, calculando A^{-1} y multiplicando esta última matriz por b , o bien haciendo uso de la fórmula de Cramer.

Sobre la regla de Cramer, ya hemos demostrado en el tema 1 que el número C_n de operaciones necesarias para resolver $Au = b$ viene dado por

$$C_n = (n + 1)k_n + n,$$

siendo k_n el número de operaciones algebraicas necesarias para calcular el determinante de una matriz $n \times n$ de números reales, y que se obtiene de manera recursiva por

$$\begin{cases} k_2 = 3, \\ k_n = nk_{n-1} + 2n - 1. \end{cases}$$

Recordemos que, con estas fórmulas, en particular se tenía $k_{10} = 9864099$, y $C_{10} = 108505099$, es decir, el número de operaciones necesarias para resolver un sistema de diez ecuaciones y diez incógnitas aplicando la regla de Cramer, resulta ser superior a los cien millones.

Por otra parte, el número d_n de operaciones necesarias para calcular la inversa de una matriz real $n \times n$ es

$$d_n = n^2k_{n-1} + (2n - 1) + n^2,$$

resultado de calcular los n^2 determinantes de los menores de orden $n - 1$, calcular $|A| = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}$, y efectuar las n^2 divisiones por $|A|$.

Además, una vez calculada A^{-1} , para hallar $A^{-1}b$ hace falta efectuar, n veces, n productos y $n - 1$ sumas, es decir, $n(2n - 1)$ operaciones. En consecuencia, para resolver el sistema $Au = b$ por la fórmula $u = A^{-1}b$, hacen falta

$$d_n + n(2n - 1) = n^2k_{n-1} + 3n^2 + n - 1$$

operaciones. En particular, si $n = 10$, obtenemos 98641109 como número de operaciones, que es algo inferior al obtenido por aplicación de la regla de Cramer (téngase en cuenta además que una vez se conoce A^{-1} , la fórmula $u = A^{-1}b$ permite calcular la solución para diferentes valores de b con sólo $n(2n - 1)$ operaciones, cosa que no sucede con la regla de Cramer, en la que al variar b hay que hallar n determinantes nuevos de orden n cada vez).

No obstante lo anterior, hay un tipo particular de sistemas de ecuaciones lineales para los que se rebaja de forma espectacular el número de operaciones necesarias para resolverlos; estos son los denominados sistemas triangulares.

Definición 1.3 a) Se dice que el sistema (1.1) es triangular superior si $a_{ij} = 0$ para todo $1 \leq j < i \leq n$, es decir, si todos los elementos a_{ij} de la matriz A que estén por debajo de la diagonal principal son nulos.

b) Se dice que el sistema (1.1) es triangular inferior si $a_{ij} = 0$ para todo $1 \leq i < j \leq n$, es decir, si todos los elementos a_{ij} de la matriz A que estén por encima de la diagonal principal son nulos.

Así por ejemplo, un sistema triangular superior es de la forma

que se resuelve por un algoritmo de bajada o descenso dado por

$$\left\{ \begin{array}{l} u_1 = \frac{b_1}{a_{11}}, \\ u_2 = \frac{1}{a_{22}}(b_2 - a_{21}u_1), \\ \dots\dots\dots, \\ \dots\dots\dots, \\ u_{n-1} = \frac{1}{a_{n-1,n-1}}(b_{n-1} - a_{n-1,1}u_1 - a_{n-1,2}u_2 - \dots - a_{n-1,n-2}u_{n-2}), \\ u_n = \frac{1}{a_{nn}}(b_n - a_{n,1}u_1 - a_{n,2}u_2 - \dots - a_{n,n-1}u_{n-1}). \end{array} \right. \quad (1.8)$$

Observación 1.4 *Dos clases muy importantes de métodos para la resolución de sistemas de ecuaciones lineales son los métodos directos y los métodos iterados. Los métodos directos, supuesto que no existan errores de redondeo, permiten calcular la solución exacta del sistema. Por el contrario, los métodos iterados en general sólo conducen a una solución aproximada.*

En este tema, vamos a estudiar algunos métodos directos, y todos ellos van a estar basados en la reducción de la matriz de los coeficientes del sistema a una matriz triangular superior (o inferior), para a continuación resolver este último sistema por el algoritmo (1.6) (o (1.8)).

Observación 1.5 *Las matrices de coeficientes que suelen aparecer en las aplicaciones prácticas son, o llenas pero no grandes, o matrices huecas. De manera algo imprecisa, por una matriz llena se entiende una en la que la mayor parte de sus coeficientes son no nulos, y diremos que no es grande si es de orden menor que 30. Una matriz hueca (sparse en inglés) es aquella, siendo grande o no, en que la mayor parte de sus coeficientes son nulos, estando la mayor parte de los no nulos en o por encima de la diagonal principal.*

Las matrices llenas pero no grandes aparecen en una gran variedad de problemas de Estadística, Física, Ingeniería, etc..., mientras que las huecas aparecen generalmente en la resolución numérica de ecuaciones diferenciales.

En general, los métodos directos suelen estar bien adaptados a ambos tipos de matrices, mientras que los métodos iterados están mejor adaptados a las matrices huecas.

2 El método de Gauss

Para el estudio del método de Gauss, procedemos en varias etapas.

Etapá 1.- Descripción del método

El método de Gauss parte de la idea de obtener, mediante cambios en el orden en que se escriben las ecuaciones, y combinaciones lineales de las mismas, un sistema equivalente al de partida que sea triangular superior. Para ello, en una primera

etapa, lo primero que se hace es, mediante cambio de orden si hace falta, escribir el sistema de tal manera que el coeficiente que multiplica a x_1 en la primera ecuación (el elemento a_{11} en la matriz asociada) sea no nulo, y a continuación, efectuando la correspondiente combinación lineal con la primera ecuación, escribir un sistema equivalente al de partida tal que los coeficientes que multiplican a x_1 en la segunda y posteriores ecuaciones (los elementos a_{i1} , $2 \leq i \leq n$ en la nueva matriz asociada) sean todos nulos.

De esta forma, en una segunda etapa, se vuelve cambiar el orden en que se escriben las $n - 1$ últimas ecuaciones, si es preciso, de tal forma que se obtenga que el coeficiente que multiplica a x_2 en la segunda ecuación (el elemento a_{22} en la correspondiente matriz asociada) sea no nulo, y a continuación, efectuando la correspondiente combinación lineal con la segunda ecuación, se obtiene un sistema equivalente al de partida tal que los coeficientes que multiplican a x_2 en la tercera y posteriores ecuaciones (los elementos a_{i2} , $3 \leq i \leq n$ en la nueva matriz asociada) sean todos nulos.

De manera general, en la etapa $r \leq n - 1$ se vuelve cambiar el orden en que se escriben las $n - (r - 1)$ últimas ecuaciones, si es preciso, de tal forma que se obtenga que el coeficiente que multiplica a x_r en la r -ésima ecuación (el elemento a_{rr} en la correspondiente matriz asociada) sea no nulo, y a continuación, efectuando la correspondiente combinación lineal con la r -ésima ecuación, se escribe un sistema equivalente al de partida tal que los coeficientes que multiplican a x_r en la $(r + 1)$ -ésima y posteriores ecuaciones (los elementos a_{ir} , $r + 1 \leq i \leq n$ en la nueva matriz asociada) sean todos nulos.

Procediendo de esta manera, al cabo de $n - 1$ etapas (como máximo) se obtiene un sistema triangular superior, equivalente al de partida, que podemos resolver mediante un algoritmo de subida.

Así por ejemplo, consideremos el sistema

$$\begin{cases} 3u_2 + 2u_3 = 12 \\ 5u_1 - 6u_2 + 2u_3 = -1 \\ -4u_1 + 2u_2 + u_3 = 3 \end{cases}$$

Intercambiando la primera y la segunda ecuaciones, obtenemos

$$\begin{cases} 5u_1 - 6u_2 + 2u_3 = -1 \\ 3u_2 + 2u_3 = 12 \\ -4u_1 + 2u_2 + u_3 = 3 \end{cases}$$

sistema que ya tiene el primer coeficiente de la segunda ecuación nulo. Para conseguir que el primer coeficiente de la tercera ecuación sea también nulo, basta sumarle a la tercera ecuación la primera multiplicada por $4/5$, obteniéndose como nuevo sistema, equivalente al de partida,

$$\begin{cases} 5u_1 - 6u_2 + 2u_3 = -1 \\ 3u_2 + 2u_3 = 12 \\ -\frac{14}{5}u_2 + \frac{13}{5}u_3 = \frac{11}{5} \end{cases}$$

Ahora, como el coeficiente de x_2 en la segunda ecuación de este último sistema es 3, no nulo, basta sumarle a la tercera ecuación la segunda multiplicada por $(1/3) \cdot (14/5)$, para obtener finalmente el sistema triangular superior

$$\begin{cases} 5u_1 - 6u_2 + 2u_3 = -1 \\ 3u_2 + 2u_3 = 12 \\ \frac{67}{15}u_3 = \frac{67}{5} \end{cases}$$

que, en este sencillo caso, conduce a $u_3 = 3$, $u_2 = 2$ y $u_1 = 1$.

Las operaciones que hemos efectuado en el ejemplo precedente pueden ser escritas, en términos de matrices, de la siguiente manera.

Para comenzar, partimos del sistema $Au = b$, siendo

$$(A|b) = \left(\begin{array}{ccc|c} 0 & 3 & 2 & 12 \\ 5 & -6 & 2 & -1 \\ -4 & 2 & 1 & 3 \end{array} \right)$$

Denotemos $(A^{(1)}|b^{(1)}) = (A|b)$. En una primera etapa, hemos efectuado dos operaciones; una primera operación ha consistido en, dado que $a_{11}^{(1)}$, el elemento de la fila 1 y columna 1 de $A^{(1)}$, es nulo, hemos intercambiado la primera ecuación con otra en la que el primer coeficiente sea no nulo, en concreto, como $a_{22}^{(1)} = 5 \neq 0$, hemos escrito el sistema $\tilde{A}^{(1)}u = \tilde{b}^{(1)}$, donde

$$(\tilde{A}^{(1)}|\tilde{b}^{(1)}) = \left(\begin{array}{ccc|c} 5 & -6 & 2 & -1 \\ 0 & 3 & 2 & 12 \\ -4 & 2 & 1 & 3 \end{array} \right)$$

A continuación, como segunda operación, hemos escrito el sistema $A^{(2)}u = b^{(2)}$, donde

$$(A^{(2)}|b^{(2)}) = \left(\begin{array}{ccc|c} 5 & -6 & 2 & -1 \\ 0 & 3 & 2 & 12 \\ 0 & -\frac{14}{5} & \frac{13}{5} & \frac{11}{5} \end{array} \right)$$

Se observa por tanto que

$$A^{(2)} = E^{(1)}\tilde{A}^{(1)}, \quad b^{(2)} = E^{(1)}\tilde{b}^{(1)},$$

donde si denotamos $\tilde{A}^{(1)} = (\tilde{a}_{ij}^{(1)})$,

$$E^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\tilde{a}_{21}^{(1)}}{\tilde{a}_{11}^{(1)}} & 1 & 0 \\ -\frac{\tilde{a}_{31}^{(1)}}{\tilde{a}_{11}^{(1)}} & 0 & 1 \end{pmatrix}$$

Después, en una segunda etapa, y dado que el sistema ahora resultante tiene en la segunda ecuación como coeficiente de u_2 un número no nulo, es decir, dado que $a_{22}^{(2)}$

es no nulo, no hemos intercambiado la segunda ecuación con la tercera, y hemos procedido directamente a obtener el sistema $A^{(3)}u = b^{(3)}$, donde

$$(A^{(3)}|b^{(3)}) = \left(\begin{array}{ccc|c} 5 & -6 & 2 & -1 \\ 0 & 3 & 2 & 12 \\ 0 & 0 & \frac{67}{15} & \frac{67}{5} \end{array} \right)$$

Se observa que, por analogía con la primera etapa, la matriz $(A^{(3)}|b^{(3)})$ ha sido obtenida como sigue. En primer lugar, se ha tomado $(\tilde{A}^{(2)}|\tilde{b}^{(2)}) = (A^{(2)}|b^{(2)})$, y a continuación

$$A^{(3)} = E^{(2)}\tilde{A}^{(2)}, \quad b^{(3)} = E^{(2)}\tilde{b}^{(2)},$$

donde si denotamos $\tilde{A}^{(2)} = (\tilde{a}_{ij}^{(2)})$,

$$E^{(2)} = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{\tilde{a}_{32}^{(2)}}{\tilde{a}_{22}^{(2)}} & 1 \end{array} \right)$$

Procedamos ahora a describir de manera general el método de Gauss en términos de operaciones matriciales, y a justificar que, bajo la hipótesis de que $|A| \neq 0$, se puede siempre aplicar.

Consideramos por tanto el sistema de n ecuaciones lineales con n incógnitas $Au = b$, con A matriz regular de coeficientes reales.

El método de Gauss para llevar el sistema anterior a uno triangular superior, se lleva a cabo en $n - 1$ etapas.

En una primera etapa, denotemos $(A^{(1)}|b^{(1)}) = (A|b)$. En primer lugar, como $|A| \neq 0$, en la primera columna de A hay al menos un elemento no nulo. Si $a_{11} \neq 0$, denotamos $(\tilde{A}^{(1)}|\tilde{b}^{(1)}) = (A^{(1)}|b^{(1)})$. Por el contrario, si $a_{11} = 0$, y $a_{i1} \neq 0$, con $i > 1$, intercambiamos la fila 1 con la fila i de la matriz $(A^{(1)}|b^{(1)})$, y denotamos en este caso $(\tilde{A}^{(1)}|\tilde{b}^{(1)})$ a la matriz resultante. De esta forma, si denotamos

$$\tilde{A}^{(1)} = (\tilde{a}_{ij}^{(1)}), \quad \tilde{b}^{(1)} = (\tilde{b}_i^{(1)}),$$

tenemos garantizado que $\tilde{a}_{11}^{(1)} \neq 0$.

A continuación, se define

$$(A^{(2)}|b^{(2)}) = E^{(1)}(\tilde{A}^{(1)}|\tilde{b}^{(1)}),$$

donde

$$E^{(1)} = \left(\begin{array}{cccccc} 1 & 0 & 0 & \dots & 0 & 0 \\ -\frac{\tilde{a}_{21}^{(1)}}{\tilde{a}_{11}^{(1)}} & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{\tilde{a}_{n-1,1}^{(1)}}{\tilde{a}_{11}^{(1)}} & 0 & 0 & \dots & 1 & 0 \\ -\frac{\tilde{a}_{n1}^{(1)}}{\tilde{a}_{11}^{(1)}} & 0 & 0 & \dots & 0 & 1 \end{array} \right).$$

De esta forma, se obtiene en particular que la matriz $A^{(2)}$ es de la forma

$$A^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n-1,2}^{(2)} & \cdots & a_{n-1,n}^{(2)} \\ 0 & a_{n,2}^{(2)} & \cdots & a_{n,n}^{(2)} \end{pmatrix},$$

con $a_{11}^{(2)} = \tilde{a}_{11}^{(1)} \neq 0$, y $|A^{(2)}| \neq 0$.

Ahora, de manera general, supongamos que en la etapa $r - 1 \geq 1$ se ha obtenido el sistema $A^{(r)}u = b^{(r)}$, con $A^{(r)}$ matriz $n \times n$ tal que todos los elementos de las $r - 1$ primeras columnas que estén por debajo de la diagonal sean nulos, es decir, de la forma,

$$A^{(r)} = \begin{pmatrix} a_{11}^{(r)} & a_{12}^{(r)} & \cdots & a_{1,r-1}^{(r)} & a_{1r}^{(r)} & \cdots & a_{1n}^{(r)} \\ 0 & a_{22}^{(r)} & \cdots & a_{2,r-1}^{(r)} & a_{2r}^{(r)} & \cdots & a_{2n}^{(r)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{r-1,r-1}^{(r)} & a_{r-1,r}^{(r)} & \cdots & a_{r-1,n}^{(r)} \\ 0 & 0 & \cdots & 0 & a_{r,r}^{(r)} & \cdots & a_{r,n}^{(r)} \\ 0 & 0 & \cdots & 0 & a_{r+1,r}^{(r)} & \cdots & a_{r+1,n}^{(r)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n,r}^{(r)} & \cdots & a_{n,n}^{(r)} \end{pmatrix}$$

Supongamos también que los elementos $a_{ii}^{(r)}$, para $i = 1, \dots, r - 1$, son todos no nulos, y que $|A^{(r)}| \neq 0$.

En tal caso, como evidentemente

$$|A^{(r)}| = a_{11}^{(r)} \cdots a_{r-1,r-1}^{(r)} \begin{vmatrix} a_{r,r}^{(r)} & \cdots & a_{r,n}^{(r)} \\ a_{r+1,r}^{(r)} & \cdots & a_{r+1,n}^{(r)} \\ \vdots & \cdots & \vdots \\ a_{n,r}^{(r)} & \cdots & a_{n,n}^{(r)} \end{vmatrix},$$

está claro que al menos uno de los elementos $a_{i,r}^{(r)}$, con $i \geq r$, es no nulo.

Si $a_{rr}^{(r)} \neq 0$, denotamos $(\tilde{A}^{(r)}|\tilde{b}^{(r)}) = (A^{(r)}|b^{(r)})$. Por el contrario, si $a_{rr}^{(r)} = 0$, y $a_{ir}^{(r)} \neq 0$, con $i > r$, entonces intercambiamos la fila r con la fila i de la matriz $(A^{(r)}|b^{(r)})$, y denotamos en este caso $(\tilde{A}^{(r)}|\tilde{b}^{(r)})$ a la matriz resultante. De esta forma, de nuevo, si denotamos

$$\tilde{A}^{(r)} = (\tilde{a}_{ij}^{(r)}), \quad \tilde{b}^{(r)} = (\tilde{b}_i^{(r)}),$$

tenemos garantizado que $\tilde{a}_{rr}^{(r)} \neq 0$.

A continuación, se define

$$(A^{(r+1)}|b^{(r+1)}) = E^{(r)}(\tilde{A}^{(r)}|\tilde{b}^{(r)}), \quad (2.9)$$

siendo

$$E^{(r)} = \left(\begin{array}{ccc|ccc} & & & I_r & & \Theta_{r \times (n-r)} \\ & & & \hline 0 & \dots & -\frac{\tilde{a}_{r+1,r}^{(r)}}{\tilde{a}_{rr}^{(r)}} & & & \\ \vdots & \ddots & \vdots & & & \\ 0 & \dots & -\frac{\tilde{a}_{n,r}^{(r)}}{\tilde{a}_{rr}^{(r)}} & & I_{n-r} & \end{array} \right), \quad (2.10)$$

donde, en general, denotamos por I_k a la matriz identidad $k \times k$, y por $\Theta_{r \times (n-r)}$ a la matriz nula $r \times (n-r)$.

De esta forma, se obtiene en particular que la matriz $A^{(r+1)}$ es de la forma

$$A^{(r+1)} = \begin{pmatrix} a_{11}^{(r+1)} & a_{12}^{(r+1)} & \dots & a_{1,r-1}^{(r+1)} & a_{1r}^{(r+1)} & \dots & a_{1n}^{(r+1)} \\ 0 & a_{22}^{(r+1)} & \dots & a_{2,r-1}^{(r+1)} & a_{2r}^{(r+1)} & \dots & a_{2n}^{(r+1)} \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{r-1,r-1}^{(r+1)} & a_{r-1,r}^{(r+1)} & \dots & a_{r-1,n}^{(r+1)} \\ 0 & 0 & \dots & 0 & a_{r,r}^{(r+1)} & \dots & a_{r,n}^{(r+1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & a_{r+1,n}^{(r+1)} \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & a_{n,n}^{(r+1)} \end{pmatrix},$$

siendo $a_{ii}^{(r+1)} \neq 0$ para $i = 1, \dots, r$, $|A^{(r+1)}| \neq 0$.

Se observa, por tanto, que con este procedimiento la matriz $A^{(n)}$ es triangular superior con determinante no nulo (de hecho, el determinante de $A^{(n)}$ es igual en valor absoluto al determinante de A), con lo que en particular los $a_{ii}^{(n)}$ son todos no nulos, y se puede resolver el sistema $A^{(n)}u = b^{(n)}$ por el algoritmo de subida.

Obsérvese también que la matriz $A^{(r+1)}$ se construye mediante las fórmulas

$$a_{ij}^{(r+1)} = \tilde{a}_{ij}^{(r)}, \quad \forall 1 \leq i \leq r, \quad \forall 1 \leq j \leq n, \quad (2.11)$$

$$a_{ij}^{(r+1)} = 0, \quad \forall r+1 \leq i \leq n, \quad \forall 1 \leq j \leq r, \quad (2.12)$$

$$a_{i,j}^{(r+1)} = \tilde{a}_{i,j}^{(r)} - \frac{\tilde{a}_{i,r}^{(r)}}{\tilde{a}_{r,r}^{(r)}} \tilde{a}_{r,j}^{(r)}, \quad \forall r+1 \leq i \leq n, \quad \forall r+1 \leq j \leq n. \quad (2.13)$$

Análogamente, el vector $b^{(r+1)}$ se construye mediante las fórmulas

$$b_i^{(r+1)} = \tilde{b}_i^{(r)}, \quad \forall 1 \leq i \leq r, \quad (2.14)$$

$$b_i^{(r+1)} = \tilde{b}_i^{(r)} - \frac{\tilde{a}_{i,r}^{(r)}}{\tilde{a}_{r,r}^{(r)}} \tilde{b}_r^{(r)}, \quad \forall r+1 \leq i \leq n. \quad (2.15)$$

Etapa 2.- Cómputo del número de operaciones

Para el cómputo del número de operaciones que hay que realizar como máximo en el método de Gauss, no vamos a contabilizar los cambios de filas que eventualmente haya que realizar en cada etapa. Teniendo en cuenta esto, está claro que las únicas operaciones que hay que realizar en cada etapa vienen dadas por las fórmulas (2.13) y (2.15).

Así, en la etapa r , para el cálculo de los $a_{i,j}^{(r+1)}$, $r+1 \leq i, j \leq n$, hace falta efectuar $n-r$ divisiones, $(n-r)^2$ productos, y $(n-r)^2$ sumas. En la misma etapa, para el cálculo de los $b_i^{(r+1)}$, $r+1 \leq i \leq n$, y si descontamos las operaciones ya realizadas para obtener las $a_{i,j}^{(r+1)}$, hace falta efectuar $(n-r)$ productos, y $(n-r)$ sumas.

Por tanto, en la etapa r , para obtener $(A^{(r+1)}|b^{(r+1)})$, hace falta realizar como máximo $2(n-r)^2 + 3(n-r)$ operaciones, y en consecuencia, para llegar al sistema triangular superior $A^{(n)}u = b^{(n)}$, hay que realizar a lo más una cantidad de operaciones igual a $\sum_{r=1}^{n-1} [2(n-r)^2 + 3(n-r)]$.

Ahora bien, teniendo en cuenta que se satisfacen las siguiente igualdades

$$\sum_{k=1}^m k = \frac{m(m+1)}{2}, \quad y \quad \sum_{k=1}^m k^2 = \frac{m(m+1)(2m+1)}{6}, \quad \forall m \geq 1,$$

cuya comprobación por inducción se deja como ejercicio, obtenemos

$$\sum_{r=1}^{n-1} (n-r) = (n-1) + (n-2) + \dots + 2 + 1 = \frac{(n-1)n}{2},$$

y

$$\sum_{r=1}^{n-1} (n-r)^2 = (n-1)^2 + (n-2)^2 + \dots + 2^2 + 1^2 = \frac{(n-1)n(2n-1)}{6}.$$

Por tanto, el número de operaciones que como máximo hay que realizar para llegar al sistema triangular superior $A^{(n)}u = b^{(n)}$, es igual a

$$\frac{3(n-1)n}{2} + \frac{(n-1)n(2n-1)}{3},$$

con lo que, sumando a esta cantidad las n^2 operaciones que a lo más hay que efectuar para resolver un sistema triangular, obtenemos que el número máximo de operaciones necesarias para resolver el sistema de orden n , $Au = b$, por el método de Gauss, si no se computan los eventuales cambios de filas, viene dado por

$$\frac{3(n-1)n}{2} + \frac{(n-1)n(2n-1)}{3} + n^2 = \frac{4n^3 + 9n^2 - 7n}{6}.$$

En consecuencia, para resolver por el método de Gauss un sistema de 10 ecuaciones lineales con 10 incógnitas, hacen falta a lo más 805 operaciones (compárese esta cantidad con la estimada para resolver un sistema de orden 10 por la regla de Cramer o por cálculo de la inversa de A).

Etapa 3.- El problema de la elección de pivote

Los elementos $\tilde{a}_{rr}^{(r)} \neq 0$ que se utilizan para dividir en las fórmulas (2.13) y (2.15), se denominan pivotes.

Teniendo en cuenta que, en la práctica, cuando se efectúan los cálculos en el método de Gauss se producen errores de redondeo, la elección del pivote en cada etapa r es una cuestión de importancia. En concreto, si por ejemplo en la etapa $r-1$ se ha obtenido que $a_{rr}^{(r)} \neq 0$, pero su valor absoluto es muy pequeño en comparación con el de otro elemento $a_{ir}^{(r)} \neq 0$ con $i > r$, entonces la división por el primero producirá errores de redondeo mucho mayores que la división por este último, y en consecuencia, aún siendo $a_{rr}^{(r)} \neq 0$, es mejor utilizar el elemento $a_{ir}^{(r)}$ como pivote r -ésimo $\tilde{a}_{rr}^{(r)}$.

Por ejemplo, consideremos el sistema

$$\begin{cases} \varepsilon u_1 + u_2 = 1, \\ u_1 + u_2 = 2, \end{cases}$$

cuya solución exacta es

$$u_1 = \frac{1}{1-\varepsilon}, \quad u_2 = \frac{1-2\varepsilon}{1-\varepsilon}$$

Es claro que para valores positivos de ε muy próximos a cero, los valores de las soluciones u_1 y u_2 estarán muy cercanos a 1.

Si resolvemos el sistema por el método de Gauss, tomando como pivote ε , obtendremos como sistema triangular equivalente

$$\begin{cases} \varepsilon u_1 + u_2 = 1, \\ (1 - \frac{1}{\varepsilon})u_2 = 2 - \frac{1}{\varepsilon}, \end{cases}$$

sistema al que al aplicarle el algoritmo de subida produce como solución

$$u_2 = \frac{2 - \frac{1}{\varepsilon}}{1 - \frac{1}{\varepsilon}}, \quad u_1 = \frac{1 - u_2}{\varepsilon}. \quad (2.16)$$

En la computadora, si ε es suficientemente pequeño, los números $2 - 1/\varepsilon$ y $1 - 1/\varepsilon$ se comportan como iguales. Así por ejemplo, si $\varepsilon = 10^{-8}$, entonces $1/\varepsilon = 10^8$, con lo que

$$\begin{cases} 2 - 1/\varepsilon = -99999998 = -0.99999998 \times 10^8, \\ 1 - 1/\varepsilon = -99999999 = -0.99999999 \times 10^8, \end{cases}$$

con lo que si suponemos que se manejan tan sólo siete cifras decimales y que se utiliza aproximación por redondeo, tanto $2 - 1/\varepsilon$ como $1 - 1/\varepsilon$ producen el número -0.1×10^9 .

En tales circunstancias, el algoritmo (2.16) produciría $u_2 = 1$, que es un valor aproximado aceptable de dicha incógnita, pero entonces se obtendría $u_1 = 0/\varepsilon = 0$, que es un valor muy alejado de la verdadera solución.

En este ejemplo, las dificultades desaparecen si se cambia de pivote, es decir, si se intercambia primero el orden de las ecuaciones escribiendo el sistema

$$\begin{cases} u_1 + u_2 = 2, \\ \varepsilon u_1 + u_2 = 1, \end{cases}$$

y a continuación se aplica el método de Gauss, tomando como pivote 1, con lo que obtenemos

$$\begin{cases} u_1 + u_2 = 2, \\ (1 - \varepsilon)u_2 = 1 - 2\varepsilon, \end{cases}$$

y en consecuencia como algoritmo de subida

$$u_2 = \frac{1 - 2\varepsilon}{1 - \varepsilon}, \quad u_1 = 2 - u_2,$$

que cuando ε es pequeño, por ejemplo como antes, produce $u_2 = 1$, $u_1 = 1$, que son ambos valores aproximados aceptables de las verdaderas soluciones.

Para evitar situaciones como la del ejemplo precedente, es usual seguir alguna estrategia en la elección del pivote.

Una de tales estrategias es la denominada de pivote parcial, que consiste en, en cada etapa r , determinar $i_r \geq r$ tal que

$$|a_{i_r r}^{(r)}| = \max_{r \leq i \leq n} |a_{ir}^{(r)}|,$$

e intercambiar (si $i_r > r$) las filas r e i_r , es decir, tomar como $\tilde{a}_{rr}^{(r)}$ el elemento $a_{i_r r}^{(r)}$.

Para la mayor parte de los sistemas de ecuaciones lineales, la estrategia de pivote parcial resulta ser suficiente.

Otra estrategia que se usa a veces es la de pivote total. Dicha estrategia consiste en determinar en cada etapa r el elemento de mayor valor absoluto entre todos los de la matriz $A^{(r)}$ que forman la submatriz formada por las filas y columnas entre la r y la n . Es decir, se determinan $r \leq i_r, j_r \leq n$ tales que

$$|a_{i_r j_r}^{(r)}| = \max_{r \leq i, j \leq n} |a_{ij}^{(r)}|,$$

y, si $i_r > r$, se intercambian las filas r e i_r , para a continuación, si $j_r > r$, intercambiar las columnas r y j_r . Naturalmente, en tal caso, si hay intercambio de columnas, hay también que efectuar el mismo intercambio en las incógnitas del sistema.

En general, las estrategias de cambio de pivote tienen la ventaja de producir menores errores de redondeo cuando se aplica el método de Gauss, pero naturalmente, al exigir efectuar comparaciones en cada etapa para determinar el nuevo pivote, aumentan los tiempos de cálculo.

Etapas 4.- Descomposición $A = LU$ cuando no hay que hacer cambios de filas

Supongamos que la matriz A es tal que al aplicarle el método de Gauss, en cada etapa r el término $a_{rr}^{(r)}$ es distinto de cero, y no efectuamos cambios de filas, es decir, $\tilde{A}^{(r)} = A^{(r)}$. En tal caso, para cada r , de acuerdo con (2.9),

$$A^{(r+1)} = E^{(r)} A^{(r)},$$

con $E^{(r)}$ la matriz definida por (2.10).

En consecuencia, teniendo en cuenta que $A^{(1)} = A$, tenemos

$$A^{(n)} = E^{(n-1)} \dots E^{(2)} E^{(1)} A \tag{2.17}$$

Ahora bien, la matriz $A^{(n)}$ es triangular superior, y la vamos a denotar U (del inglés upper).

Por otra parte, cada una de las matrices $E^{(r)}$ es, en particular, triangular inferior con la diagonal principal formada por unos. El producto de dos matrices triangulares inferior con las diagonales principales respectivas formadas por unos, es también una matriz triangular inferior con la misma propiedad (hágase como ejercicio). Asimismo, toda matriz triangular inferior con la diagonal principal formada por unos es invertible, y su inversa es de nuevo triangular inferior con la diagonal principal formada por unos (hágase también como ejercicio).

Teniendo en cuenta lo anterior, la matriz $(E^{(n-1)} \dots E^{(1)})^{-1}$ es triangular inferior con la diagonal principal formada por unos. Si denotamos a esta matriz por L (del inglés lower), podemos afirmar el siguiente resultado:

Proposición 2.1 *Si A es una matriz regular $n \times n$ de coeficientes reales tal que al aplicarle el método de Gauss, en cada etapa r el término $a_{rr}^{(r)}$ es distinto de cero, entonces existe una matriz triangular inferior L , con la diagonal principal formada por unos, y una matriz triangular superior U , de tal manera que*

$$A = LU \tag{2.18}$$

3 El método de Gauss-Jordan

El método de Gauss-Jordan es una variante del método de Gauss, en el que se busca llegar a un sistema de ecuaciones lineales, equivalente al de partida, tal que la matriz de coeficientes sea diagonal, es decir tal que los únicos elementos no nulos de la matriz sean los de la diagonal principal. Para ello, se procede de la manera que describimos a continuación.

Al igual que en el método de Gauss, suponemos que queremos resolver el sistema de n ecuaciones con n incógnitas $Au = b$, siendo A una matriz regular $n \times n$ de coeficientes reales.

El método de Gauss-Jordan para llevar el sistema anterior a uno triangular superior, se lleva a cabo en n etapas.

En una primera etapa, se actúa igual que en el método de Gauss, es decir, se denota $(A^{(1)}|b^{(1)}) = (A|b)$, se elige $i \geq 1$ tal que $a_{i1} \neq 0$, y si $i \neq 1$, se intercambian la fila 1 con la fila i de la matriz $(A^{(1)}|b^{(1)})$, y se denota $(\tilde{A}^{(1)}|\tilde{b}^{(1)})$ a la matriz resultante.

A continuación, se define

$$(A^{(2)}|b^{(2)}) = E^{(1)}(\tilde{A}^{(1)}|\tilde{b}^{(1)}),$$

donde $E^{(1)}$ está definida como en el método de Gauss.

De esta forma, se obtiene en particular que la matriz $A^{(2)}$ es de la forma

$$A^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \dots & a_{1n}^{(r)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n-1,2}^{(2)} & \dots & a_{n-1,n}^{(2)} \\ 0 & a_{n,2}^{(2)} & \dots & a_{n,n}^{(2)} \end{pmatrix},$$

con $a_{11}^{(2)} = \tilde{a}_{11}^{(1)} \neq 0$, y $|A^{(2)}| \neq 0$.

Ahora, de manera general, supongamos que en la etapa $r - 1 \geq 1$ se ha obtenido el sistema $A^{(r)}u = b^{(r)}$, con $A^{(r)}$ matriz $n \times n$ tal que todos los elementos de las $r - 1$ primeras columnas que no estén en la diagonal sean nulos, es decir, de la forma,

$$A^{(r)} = \begin{pmatrix} a_{11}^{(r)} & 0 & \cdots & 0 & a_{1r}^{(r)} & \cdots & a_{1n}^{(r)} \\ 0 & a_{22}^{(r)} & \cdots & 0 & a_{2r}^{(r)} & \cdots & a_{2n}^{(r)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{r-1,r-1}^{(r)} & a_{r-1,r}^{(r)} & \cdots & a_{r-1,n}^{(r)} \\ 0 & 0 & \cdots & 0 & a_{r,r}^{(r)} & \cdots & a_{r,n}^{(r)} \\ 0 & 0 & \cdots & 0 & a_{r+1,r}^{(r)} & \cdots & a_{r+1,n}^{(r)} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n,r}^{(r)} & \cdots & a_{n,n}^{(r)} \end{pmatrix}$$

Supongamos también que los elementos $a_{ii}^{(r)}$, para $i = 1, \dots, r - 1$, son todos no nulos, y que $|A^{(r)}| \neq 0$.

En tal caso, como evidentemente

$$|A^{(r)}| = a_{11}^{(r)} \cdots a_{r-1,r-1}^{(r)} \begin{vmatrix} a_{r,r}^{(r)} & \cdots & a_{r,n}^{(r)} \\ a_{r+1,r}^{(r)} & \cdots & a_{r+1,n}^{(r)} \\ \vdots & \cdots & \vdots \\ a_{n,r}^{(r)} & \cdots & a_{n,n}^{(r)} \end{vmatrix},$$

está claro que al menos uno de los elementos $a_{i,r}^{(r)}$, con $i \geq r$, es no nulo.

Se elige $i \geq r$ tal que $a_{i,r}^{(r)} \neq 0$, se intercambian, si es preciso, la fila r con la fila i de la matriz $(A^{(r)}|b^{(r)})$, y se denota $(\tilde{A}^{(r)}|\tilde{b}^{(r)})$ a la matriz resultante. De esta forma, de nuevo, si denotamos

$$\tilde{A}^{(r)} = (\tilde{a}_{ij}^{(r)}), \quad \tilde{b}^{(r)} = (\tilde{b}_i^{(r)}),$$

tenemos garantizado que $\tilde{a}_{rr}^{(r)} \neq 0$.

A continuación, se define

$$(A^{(r+1)}|b^{(r+1)}) = J^{(r)}(\tilde{A}^{(r)}|\tilde{b}^{(r)}), \quad (3.19)$$

siendo $J^{(r)} = (c_{ij}^{(r)})$, la matriz $n \times n$ definida por

$$J^{(r)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & -\frac{\tilde{a}_{1r}^{(r)}}{\tilde{a}_{rr}^{(r)}} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & -\frac{\tilde{a}_{2r}^{(r)}}{\tilde{a}_{rr}^{(r)}} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & -\frac{\tilde{a}_{r-1r}^{(r)}}{\tilde{a}_{rr}^{(r)}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \frac{1}{\tilde{a}_{rr}^{(r)}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -\frac{\tilde{a}_{r+1r}^{(r)}}{\tilde{a}_{rr}^{(r)}} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & -\frac{\tilde{a}_{nr}^{(r)}}{\tilde{a}_{rr}^{(r)}} & 0 & \cdots & 1 \end{pmatrix},$$

De esta manera se obtiene que la matriz $A^{(r+1)}$ es de la forma

$$A^{(r+1)} = \begin{pmatrix} a_{11}^{(r+1)} & 0 & \dots & 0 & 0 & a_{1,r+1}^{(r+1)} & \dots & a_{1n}^{(r+1)} \\ 0 & a_{22}^{(r+1)} & \dots & 0 & 0 & a_{2,r+1}^{(r+1)} & \dots & a_{2n}^{(r+1)} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{r-1,r-1}^{(r+1)} & 0 & a_{r-1,r+1}^{(r+1)} & \dots & a_{r-1,n}^{(r+1)} \\ 0 & 0 & \dots & 0 & a_{r,r}^{(r+1)} & a_{r,r+1}^{(r+1)} & \dots & a_{r,n}^{(r+1)} \\ 0 & 0 & \dots & 0 & 0 & a_{r+1,r+1}^{(r+1)} & \dots & a_{r+1,n}^{(r+1)} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & a_{n,r+1}^{(r+1)} & \dots & a_{n,n}^{(r+1)} \end{pmatrix},$$

siendo $a_{ii}^{(r+1)} \neq 0$ para $i = 1, \dots, r$, $|A^{(r+1)}| \neq 0$.

Se observa por tanto, que con este procedimiento, la matriz $A^{(n+1)}$ es diagonal, con todos los $a_{ii}^{(n+1)}$ no nulos, y se puede resolver el sistema $A^{(n+1)}u = b^{(n+1)}$ de manera inmediata.

Obsérvese que si en una posible etapa $n+1$ del método de Gauss-Jordan se multiplica cada fila i de $(A^{(n+1)}|b^{(n+1)})$ por $1/a_{ii}^{(n+1)}$, entonces el vector $\tilde{b}^{(n+1)}$ resultante es la solución del sistema $Au = b$.

Como un ejemplo consideremos el sistema

$$\begin{cases} 2u_1 + u_2 - 3u_3 = -1, \\ -u_1 + 3u_2 + 2u_3 = 12, \\ 3u_1 + u_2 - 3u_3 = 0. \end{cases}$$

Denotemos

$$(A|b) = \left(\begin{array}{ccc|c} 2 & 1 & -3 & -1 \\ -1 & 3 & 2 & 12 \\ 3 & 1 & -3 & 0 \end{array} \right).$$

Denotemos $(A^{(1)}|b^{(1)}) = (A|b)$, matriz que ya tiene el primer coeficiente de la primera ecuación igual a 2, no nulo.

Sumando a la segunda fila la primera multiplicada por $1/2$, y a la tercera fila la primera multiplicada por $-3/2$, obtenemos la matriz

$$(A^{(2)}|b^{(2)}) = \left(\begin{array}{ccc|c} 2 & 1 & -3 & -1 \\ 0 & 7/2 & 1/2 & 23/2 \\ 0 & -1/2 & 3/2 & 3/2 \end{array} \right).$$

Ahora, como el segundo coeficiente de la segunda fila de esta última matriz es $7/2$, no nulo, le sumamos a la primera fila de esta matriz la segunda multiplicada por $-2/7$, y a la tercera fila la segunda multiplicada por $1/7$, con lo que obtenemos la matriz

$$(A^{(3)}|b^{(3)}) = \left(\begin{array}{ccc|c} 2 & 0 & -22/7 & -30/7 \\ 0 & 7/2 & 1/2 & 23/2 \\ 0 & 0 & 11/7 & 22/7 \end{array} \right).$$

Finalmente, como el tercer coeficiente de la tercera fila de la matriz $(A^{(3)}|b^{(3)})$ es $11/7$, no nulo, le sumamos a la primera fila de esta matriz la tercera multiplicada

por $(22/7)/(11/7) = 2$, y a la segunda fila la tercera multiplicada por $-(1/2)/(11/7) = -7/22$, con lo que obtenemos la matriz

$$(A^{(4)}|b^{(4)}) = \left(\begin{array}{ccc|c} 2 & 0 & 0 & 2 \\ 0 & 7/2 & 0 & 21/2 \\ 0 & 0 & 11/7 & 22/7 \end{array} \right),$$

que en este sencillo caso conduce inmediatamente a la solución del sistema de partida, $u_1 = 1$, $u_2 = 3$ y $u_3 = 2$.

Dicha solución también se obtiene como la cuarta columna de la matriz $(A^{(5)}|b^{(5)})$ obtenida dividiendo por 2 la primera fila de $(A^{(4)}|b^{(4)})$, y dividiendo por $7/2$ la segunda y por $11/7$ la tercera fila de la misma matriz, que conduce a

$$(A^{(5)}|b^{(5)}) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

Observación 3.1 *El método de Gauss-Jordan no aporta ninguna ventaja sobre el de Gauss para resolver un sistema lineal concreto, pero en cambio sí tiene interés cuando se trata de efectuar el cálculo de la inversa de una matriz. En concreto, si A es una matriz regular $n \times n$ de coeficientes reales, y queremos hallar A^{-1} , lo que podemos hacer es escribir esta última matriz por columnas,*

$$A^{-1} = (u^{(1)}|u^{(2)}|\dots|u^{(n)})$$

De esta forma, como $AA^{-1} = I_n$, hemos de resolver n sistemas lineales

$$Au^{(k)} = e^{(k)} \quad 1 \leq k \leq n,$$

donde por $e^{(k)}$ denotamos al k -ésimo vector canónico de \mathbb{R}^n .

Para la resolución conjunta de estos n sistemas lineales, se escribe $(A|I)$, y aplicando Gauss-Jordan, en $n+1$ etapas se llega a $(I|B)$, siendo entonces $B = A^{-1}$.

4 El método LU

Sabemos por la Proposición 2.1 que si A es una matriz regular $n \times n$ de coeficientes reales tal que al aplicarle el método de Gauss, en cada etapa r el término $a_{rr}^{(r)}$ es distinto de cero, entonces existe una matriz triangular inferior L , con la diagonal principal formada por unos, y una matriz triangular superior U , de tal manera que que se puede escribir como el producto de L por U , es decir, $A = LU$. En tal caso, para resolver el sistema de ecuaciones lineales $Au = b$, se escribe $LUu = b$, y denotando $v = Uu$, se observa que lo que hay que hacer es resolver en primer lugar por el algoritmo de descenso el sistema

$$Lv = b,$$

para a continuación, una vez conocida v , resolver por el algoritmo de ascenso el sistema

$$Uu = v$$

Así por ejemplo, para resolver el sistema

$$\begin{cases} 5u_1 + 2u_2 + u_3 = 12, \\ 5u_1 - 6u_2 + 2u_3 = -1, \\ -4u_1 + 2u_2 + u_3 = 3, \end{cases}$$

si sabemos que su matriz A de coeficientes admite la descomposición

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -4/5 & -9/20 & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & 9/4 \end{pmatrix},$$

descomposición que se puede calcular aplicando el método de Gauss, para hallar la solución del sistema original, resolvemos en primer lugar el sistema

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -4/5 & -9/20 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 12 \\ -1 \\ 3 \end{pmatrix},$$

que tiene como solución

$$v_1 = 12, \quad v_2 = -13, \quad v_3 = 27/4.$$

A continuación, resolvemos el sistema

$$\begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & 9/4 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 12 \\ -13 \\ 27/4 \end{pmatrix},$$

y obtenemos la solución

$$u_3 = 3, \quad u_2 = 2, \quad u_1 = 1.$$

Resulta pues de interés el disponer de criterios para saber si dada una matriz cuadrada, ésta puede ser escrita como el producto de una matriz triangular inferior L , con la diagonal principal formada por unos, y una matriz triangular superior U . En caso afirmativo, se dice que dicha matriz admite una factorización LU . Sobre este particular, lo primero que se tiene es un resultado de unicidad.

Proposición 4.1 *Si A es una matriz $n \times n$ tal que su determinante $|A|$ es distinto de cero, entonces A puede ser escrita a lo más de una única manera como el producto de una matriz triangular inferior L , con la diagonal principal formada por unos, y una matriz triangular superior U .*

Demostración.- Supongamos que

$$A = L_1 U_1 = L_2 U_2$$

con L_1 y L_2 matrices triangular inferior con la diagonal principal formada por unos, y U_1 y U_2 matrices triangular superior.

En tal caso, como $0 \neq |A| = |L_1||U_1| = |L_2||U_2|$, en particular las matrices L_2 y U_1 son invertibles, y por tanto, de la igualdad $L_1U_1 = L_2U_2$ obtenemos

$$L_2^{-1}L_1 = U_2U_1^{-1}.$$

Pero $L_2^{-1}L_1$ es una matriz triangular inferior con la diagonal principal formada por unos, y $U_2U_1^{-1}$ es una matriz triangular superior, en consecuencia, de la última igualdad, deducimos que

$$L_2^{-1}L_1 = U_2U_1^{-1} = I_n,$$

y por tanto $L_1 = L_2$ y $U_1 = U_2$. ■

Nos preguntamos a continuación por algún criterio que proporcione existencia de descomposición LU para una matriz dada.

Definición 4.2 *Sea A una matriz cuadrada $n \times n$. Para cada entero $1 \leq r \leq n$, se denomina submatriz principal de orden r de A , y la denotaremos A_r , a la matriz $r \times r$ formada por las r primeras filas y las r primeras columnas de A . Al determinante $|A_r|$ lo denominaremos el determinante principal de orden r de A .*

Teorema 4.3 (de Doolittle) *Sea A una matriz cuadrada $n \times n$ tal que todos los determinantes principales $|A_r|$, para $1 \leq r \leq n$, sea no nulos. Entonces, A admite una y sólo una factorización LU .*

Demostración.- La unicidad de factorización LU es consecuencia de la Proposición 4.1. Para la existencia, por la Proposición 2.1, basta ver que al aplicarle a la matriz A el método de Gauss, en cada etapa r el término $a_{rr}^{(r)}$ es distinto de cero. Para ello, vamos a proceder por inducción.

Desde luego, por hipótesis, $a_{11}^{(1)} = a_{11} \neq 0$. Supongamos por tanto que

$$a_{rr}^{(r)} \neq 0, \quad \text{para } r = 1, \dots, k-1, \quad (4.20)$$

y comprobemos que entonces también $a_{kk}^{(k)} \neq 0$.

Desde luego, por (4.20), $\tilde{A}^{(r)} = A^{(r)}$ para $r = 1, \dots, k-1$, y en tal caso, de acuerdo con (2.9),

$$A^{(r+1)} = E^{(r)}A^{(r)}, \quad \text{para } r = 1, \dots, k-1,$$

con $E^{(r)}$ la matriz definida por (2.10).

En consecuencia, teniendo en cuenta que $A^{(1)} = A$, tenemos

$$A^{(k)} = E^{(k-1)} \dots E^{(1)}A. \quad (4.21)$$

Ahora bien, como ya se dijo anteriormente, la matriz producto $E^{(k-1)} \dots E^{(1)}$ es triangular inferior con la diagonal principal formada por unos, con lo que de (4.21), efectuando el producto de las matrices por cajas, se obtiene en particular que la matriz

$$\begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \dots & a_{1,k-1}^{(k)} & a_{1k}^{(k)} \\ 0 & a_{22}^{(k)} & \dots & a_{2,k-1}^{(k)} & a_{2k}^{(k)} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & a_{k-1,k-1}^{(k)} & a_{k-1,k}^{(k)} \\ 0 & 0 & \dots & 0 & a_{k,k}^{(k)} \end{pmatrix},$$

es igual al producto de una matriz $k \times k$ de determinante igual a 1, por la matriz principal A_k , y por tanto,

$$0 \neq |A_k| = a_{11}^{(k)} \dots a_{kk}^{(k)},$$

con lo que $a_{kk}^{(k)} \neq 0$, como queríamos demostrar. ■

Observación 4.4 *Si A es una matriz $n \times n$ no singular, pero que no satisface las condiciones del Teorema 4.3, puede suceder que sí se satisfagan dichas condiciones si se efectúan intercambios de filas en la citada matriz.*

Por otra parte, si A es singular, puede suceder que admita factorización LU , en este caso no única. Un ejemplo será visto en clase de problemas.

A continuación, exponemos otro criterio para saber si una matriz admite factorización LU de Doolittle.

Definición 4.5 *Sea $A = (a_{ij})$ una matriz cuadrada $n \times n$. Se dice que A es diagonalmente dominante en sentido estricto si*

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \quad \text{para toda fila } i = 1, \dots, n. \quad (4.22)$$

Teorema 4.6 *Sea $A = (a_{ij})$ una matriz cuadrada $n \times n$ diagonalmente dominante en sentido estricto. Entonces,*

- a) *A es regular,*
- b) *A admite una única factorización LU de Doolittle.*

Demostración.-

a) Supongamos que $|A| = 0$. En tal caso, existe $w \in \mathbb{R}^n$, $w \neq 0$, tal que $Aw = 0$. Sea $1 \leq k \leq n$ tal que $|w_k| = \max_{1 \leq j \leq n} |w_j| \neq 0$.

Como $\sum_{j=1}^n a_{kj}w_j = 0$, despejando tenemos

$$a_{kk}w_k = - \sum_{j=1, j \neq k}^n a_{kj}w_j,$$

y por lo tanto

$$|a_{kk}||w_k| \leq \sum_{j=1, j \neq k}^n |a_{kj}||w_j|,$$

con lo que

$$|a_{kk}| \leq \sum_{j=1, j \neq k}^n |a_{kj}| \frac{|w_j|}{|w_k|} \leq \sum_{j=1, j \neq k}^n |a_{kj}|,$$

en contradicción con el hecho de que A es diagonalmente dominante en sentido estricto.

b) Puesto que A es diagonalmente dominante en sentido estricto, es inmediato que cada submatriz principal A_r es también diagonalmente dominante en sentido

estricto. Gracias entonces al apartado anterior, los determinantes principales son todos no nulos, por lo que basta tener en cuenta el Teorema 4.3 para concluir con este apartado. ■

Para el cálculo efectivo de la factorización LU de una matriz dada, en la práctica se obtienen los elementos de L y de U por identificación de forma recursiva, admitiendo que existe la factorización. Así, si

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{n,n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ l_{21} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ l_{n-1,1} & l_{n-1,2} & \cdots & 1 & 0 \\ l_{n,1} & l_{n,2} & \cdots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1,n-1} & u_{1n} \\ 0 & u_{22} & \cdots & u_{2,n-1} & u_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & \cdots & 0 & u_{n,n} \end{pmatrix},$$

identificando los elementos de la primera fila de A con los correspondientes de LU , y los de la primera columna de A con los correspondientes de LU , obtenemos

$$\begin{cases} a_{1j} = u_{1j}, & j = 1, \dots, n, \\ a_{i1} = l_{i1}u_{11} \Rightarrow l_{i1} = \frac{a_{i1}}{u_{11}}, & i = 2, \dots, n. \end{cases}$$

Identificando a continuación los elementos de la segunda fila de A con los correspondientes de LU , y los de la segunda columna de A con los correspondientes de LU , obtenemos

$$\begin{cases} a_{2j} = l_{21}u_{1j} + u_{2j} \Rightarrow u_{2j} = a_{2j} - l_{21}u_{1j}, & j = 2, \dots, n, \\ a_{i2} = l_{i1}u_{12} + l_{i2}u_{22} \Rightarrow l_{i2} = \frac{1}{u_{22}}(a_{i2} - l_{i1}u_{12}), & i = 3, \dots, n. \end{cases}$$

Y en general, si suponemos conocidas las $k-1$ primeras filas de U y las $k-1$ primeras columnas de L , entonces, identificando los elementos de la k -ésima fila de A con los correspondientes de LU , y los de la k -ésima columna de A con los correspondientes de LU , obtenemos

$$\begin{cases} a_{kj} = l_{k1}u_{1j} + \cdots + l_{k,k-1}u_{k-1,j} + u_{kj} \\ \Rightarrow u_{kj} = a_{kj} - \sum_{r=1}^{k-1} l_{kr}u_{rj}, & j = k, \dots, n, \\ a_{ik} = l_{i1}u_{1k} + \cdots + l_{i,k-1}u_{k-1,k} + l_{ik}u_{kk} \\ \Rightarrow l_{ik} = \frac{1}{u_{kk}} \left(a_{ik} - \sum_{r=1}^{k-1} l_{ir}u_{rk} \right), & i = k+1, \dots, n. \end{cases}$$

Obsérvese que el cálculo anterior puede ser hecho en paralelo, lo cual redunda en un ahorro considerable de tiempo. El número de operaciones necesarias para

llevar a cabo el método de cálculo expuesto es del mismo orden que en el de Gauss. El método de Gauss suele ser el mejor para resolver un sistema concreto, pero el método LU es preferible si se trata de resolver varios sistemas lineales con la misma matriz A y distintos b , ya que una vez se factoriza A , todos los sistemas se resuelven con un ahorro considerable de tiempo en los cálculos.

Como ejercicio, se propone que se encuentre la factorización LU de la matriz

$$A = \begin{pmatrix} 6 & 2 & 1 & -1 \\ 2 & 4 & 1 & 0 \\ 1 & 1 & 4 & -1 \\ -1 & 0 & -1 & 3 \end{pmatrix}.$$

Observación 4.7 (factorización de Crout) *A la factorización LU de una matriz que acabamos de exponer se la denomina también factorización de Doolittle de dicha matriz. Otra posibilidad es descomponer una matriz A como el producto de una triangular inferior por otra triangular superior con la diagonal principal formada por unos, a dicha descomposición se la denomina factorización de Crout de A . Dicha factorización puede ser obtenida fácilmente a partir de la de Doolittle; en efecto, sea $A = LU$, con A matriz no singular,*

$$L = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ l_{n-1,1} & l_{n-1,2} & \dots & 1 & 0 \\ l_{n,1} & l_{n,2} & \dots & l_{n,n-1} & 1 \end{pmatrix},$$

$$U = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1,n-1} & u_{1n} \\ 0 & u_{22} & \dots & u_{2,n-1} & u_{2n} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & \dots & 0 & u_{n,n} \end{pmatrix}.$$

En tal caso, si denotamos

$$\tilde{D} = \begin{pmatrix} 1/u_{11} & 0 & \dots & 0 & 0 \\ 0 & 1/u_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1/u_{n-1,n-1} & 0 \\ 0 & 0 & \dots & 0 & 1/u_{n,n} \end{pmatrix},$$

entonces la matriz $\tilde{U} = \tilde{D}U$ es de la forma

$$\tilde{U} = \tilde{D}U = \begin{pmatrix} 1 & u_{12}/u_{11} & \dots & u_{1,n-1}/u_{11} & u_{1n}/u_{11} \\ 0 & 1 & \dots & u_{2,n-1}/u_{22} & u_{2n}/u_{22} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & u_{n-1,n}/u_{n-1,n-1} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

es decir, es triangular superior con la diagonal principal formada por unos, y

$$A = LU = L\tilde{D}^{-1}\tilde{U} = \tilde{L}\tilde{U},$$

con

$$\begin{aligned}
\tilde{L} &= L\tilde{D}^{-1} \\
&= \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ l_{n-1,1} & l_{n-1,2} & \dots & 1 & 0 \\ l_{n,1} & l_{n,2} & \dots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & 0 & \dots & 0 & 0 \\ 0 & u_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & u_{n-1,n-1} & 0 \\ 0 & 0 & \dots & 0 & u_{n,n} \end{pmatrix} \\
&= \begin{pmatrix} u_{11} & 0 & \dots & 0 & 0 \\ l_{21}u_{11} & u_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ l_{n-1,1}u_{11} & l_{n-1,2}u_{22} & \dots & u_{n-1,n-1} & 0 \\ l_{n,1}u_{11} & l_{n,2}u_{22} & \dots & l_{n,n-1}u_{n-1,n-1} & u_{nn} \end{pmatrix},
\end{aligned}$$

matriz triangular inferior.

Obsérvese que en el caso particular en que A es simétrica,

$$A = LU \Rightarrow A = A^* = (LU)^* = U^*L^*,$$

y por tanto

$$\tilde{L} = U^*, \quad \tilde{U} = L^*.$$

5 El método de Cholesky

En esta sección vamos a estudiar un caso particular de factorización LU , en el que no vamos a pedir que los elementos de la diagonal principal de L sean unos, y U va a coincidir con L^* , la traspuesta de L .

A este respecto, observemos que se tiene el siguiente resultado.

Proposición 5.1 *Sea A una matriz $n \times n$ simétrica y regular que admita factorización LU . Entonces existe una matriz diagonal D tal que*

$$A = LDL^*. \quad (5.23)$$

Si además $d_{ii} > 0$ para todo $i = 1, \dots, n$, entonces, denotando por $D^{1/2}$ a la matriz diagonal cuyos elementos de la diagonal principal sean las raíces $(d_{ii})^{1/2}$, y definiendo

$$B = LD^{1/2}, \quad (5.24)$$

se tiene que B es triangular inferior, y

$$A = BB^*. \quad (5.25)$$

Demostración.- De la igualdad $A = LU$, obtenemos también $A = A^* = U^*L^*$, con lo que

$$LU = U^*L^*,$$

y por tanto

$$U(L^*)^{-1} = L^{-1}U^*.$$

Como $U(L^*)^{-1}$ es triangular superior, y $L^{-1}U^*$ es triangular inferior, de la última igualdad se deduce que ambos productos han de ser iguales a una matriz diagonal D , de modo que en particular $U(L^*)^{-1} = D$, y por tanto $U = DL^*$, con lo que A se escribe en la forma (5.23).

Si $d_{ii} > 0$ para todo $i = 1, \dots, n$, tenemos que $D = D^{1/2}D^{1/2}$, y en consecuencia

$$A = LD^{1/2}D^{1/2}L^* = (LD^{1/2})(LD^{1/2})^*,$$

con lo que si definimos B por (5.24), esta matriz es triangular inferior y satisface (5.25). ■

Definición 5.2 Sea $A = (a_{ij})$ una matriz $n \times n$ de coeficientes reales. Diremos que A es definida positiva si

$$\sum_{i,j=1}^n a_{ij}u_iu_j > 0, \quad \text{para todo } u \in \mathbb{R}^n \setminus \{0\}. \quad (5.26)$$

Observación 5.3 Obsérvese que

$$\sum_{i,j=1}^n a_{ij}u_iu_j = (u_1, \dots, u_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{n,n} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = u^*Au,$$

y por tanto (5.26) puede ser escrita de manera más simple

$$u^*Au > 0, \quad \text{para todo } u \in \mathbb{R}^n \setminus \{0\}. \quad (5.27)$$

Así por ejemplo, la matriz $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$, es definida positiva, ya que

$$\begin{aligned} & (u_1, u_2, u_3) \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \\ &= 2u_1^2 + 2u_2^2 + 2u_3^2 - 2u_1u_2 - 2u_2u_3 \\ &= u_1^2 - 2u_1u_2 + u_2^2 + u_2^2 - 2u_2u_3 + u_3^2 + u_1^2 + u_3^2 \\ &= (u_1 - u_2)^2 + (u_2 - u_3)^2 + u_1^2 + u_3^2 > 0, \quad \text{si } (u_1, u_2, u_3) \neq (0, 0, 0). \end{aligned}$$

Observación 5.4 En la literatura matemática a veces se exige en la definición de matriz definida positiva el que ésta además sea simétrica. Nosotros hemos prescindido de esta exigencia. Por ejemplo, como se puede comprobar fácilmente, la matriz $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ no es simétrica, pero satisface (5.26).

Observación 5.5 Una matriz diagonal es definida positiva si y sólo si todos los elementos de la diagonal son estrictamente positivos (compruébese).

Observación 5.6 *Demuéstrese como ejercicio que una matriz $n \times n$ es definida positiva si y sólo si existe un $\alpha > 0$ tal que $u^*Au \geq \alpha\|u\|^2$, para todo $u \in \mathbb{R}^n$, donde $\|u\|^2 = \sum_{k=1}^n u_k^2$ es el cuadrado de la norma euclidiana de u .*

Proposición 5.7 *Si A es una matriz $n \times n$ definida positiva, entonces*

- a) *A es regular y, más exactamente, el determinante de A es positivo.*
- b) *Las submatrices principales A_r , $r = 1, \dots, n$, son todas definidas positivas.*

Demostración.-

a) Si A es singular, entonces existe $u \in \mathbb{R}^n \setminus \{0\}$ tal que $Au = 0$, y en consecuencia $u^*Au = 0$, contra el hecho de que A es definida positiva. Así pues, podemos afirmar por ahora que toda matriz definida positiva tiene determinante no nulo.

Consideremos ahora la familia A_λ de matrices definida por

$$A_\lambda = \lambda A + (1 - \lambda)I_n, \quad \lambda \in [0, 1].$$

Es inmediato comprobar que la matriz A_λ es definida positiva, y por tanto $|A_\lambda| \neq 0$, para todo $\lambda \in [0, 1]$. Observemos que la función $g(\lambda) = |A_\lambda|$ es un polinomio en la variable λ , y por tanto es continua en el intervalo $[0, 1]$. Basta ahora tener en cuenta que según lo anterior $g(\lambda)$ no se anula en dicho intervalo, y que $g(0) = |I_n| = 1 > 0$, para concluir que en particular $0 < g(1) = |A|$.

b) Si fijado $r = 1, \dots, n$, tomamos $u = (u_1, \dots, u_r, 0, \dots, 0) \neq 0$, obtenemos

$$0 < u^*Au = (u_1, \dots, u_r)A_r \begin{pmatrix} u_1 \\ \vdots \\ u_r \end{pmatrix},$$

y en consecuencia A_r es definida positiva. ■

Corolario 5.8 *Si A es una matriz $n \times n$ definida positiva, entonces A admite factorización LU.*

Demostración.- El resultado es consecuencia inmediata del Teorema 4.3 y la Proposición 5.7. ■

Observación 5.9 *Como una consecuencia evidente de la Proposición 5.7, podemos afirmar que si una matriz A es definida positiva, entonces todos sus menores principales son positivos. El recíproco no es cierto en general; así por ejemplo, la matriz $A = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$, tiene todos sus menores principales positivos, pero no es definida positiva. Para que el recíproco sea cierto, es necesario exigir que A sea simétrica. En concreto, se satisface el resultado siguiente*

Proposición 5.10 (Criterio de Sylvester de matriz definida positiva) *Sea A una matriz real simétrica $n \times n$. Dicha matriz es definida positiva si y sólo si todos sus menores principales son positivos.*

Vamos a comprobar la Proposición precedente en los casos $n = 2$ y $n = 3$, y se propone como ejercicio que se compruebe el caso $n = 4$.

Si $n = 2$ y A es simétrica con los menores principales positivos, entonces A es de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix},$$

satisfaciéndose

$$a_{11} > 0, \quad a_{11}a_{22} - a_{12}^2 > 0. \quad (5.28)$$

Pero, dado un $x \in \mathbb{R}^2$, con $x^* = (x_1, x_2)$, se tiene

$$\begin{aligned} x^*Ax &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \\ &= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2, \end{aligned}$$

y de esta última igualdad resulta inmediato que (5.28) implica que A es definida positiva.

Ahora, si suponemos $n = 3$ y A es simétrica con los menores principales positivos, entonces A es de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix},$$

satisfaciéndose

$$a_{11} > 0, \quad a_{11}a_{22} - a_{12}^2 > 0, \quad |A| > 0. \quad (5.29)$$

Ahora bien, si $x \in \mathbb{R}^3 \setminus \{0\}$, con $x^* = (x_1, x_2, x_3)$, se tiene

$$\begin{aligned} x^*Ax &= a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 + a_{22}x_2^2 + a_{33}x_3^2 \\ &= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 \\ &\quad + \left(a_{33} - \frac{a_{13}^2}{a_{11}}\right)x_3^2 + 2\left(a_{23} - \frac{a_{12}a_{13}}{a_{11}}\right)x_2x_3. \end{aligned} \quad (5.30)$$

Si $(x_2, x_3) = (0, 0)$, entonces, $x^*Ax = a_{11}x_1^2 > 0$, por ser $x_1 \neq 0$ y $a_{11} > 0$.

Si $(x_2, x_3) \neq (0, 0)$, entonces observando que por (5.30),

$$x^*Ax \geq \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 + \left(a_{33} - \frac{a_{13}^2}{a_{11}}\right)x_3^2 + 2\left(a_{23} - \frac{a_{12}a_{13}}{a_{11}}\right)x_2x_3,$$

concluimos de lo ya demostrado para $n = 2$, que una condición suficiente para que x^*Ax sea positivo en este caso es que se satisfagan

$$a_{22} - \frac{a_{12}^2}{a_{11}} > 0, \quad \begin{vmatrix} a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ a_{23} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{vmatrix} > 0. \quad (5.31)$$

Ahora bien, teniendo en cuenta que

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ 0 & a_{23} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{vmatrix},$$

igualdad que se obtiene de restar en el determinante de la izquierda la primera fila multiplicada por a_{12}/a_{11} de la segunda, y la primera fila multiplicada por a_{13}/a_{11} de la tercera, es inmediato comprobar que

$$\begin{vmatrix} a_{22} - \frac{a_{12}^2}{a_{11}} & a_{23} - \frac{a_{12}a_{13}}{a_{11}} \\ a_{23} - \frac{a_{12}a_{13}}{a_{11}} & a_{33} - \frac{a_{13}^2}{a_{11}} \end{vmatrix} = \frac{1}{a_{11}} |A|,$$

y por tanto, (5.31) es consecuencia inmediata de (5.29).

Observación 5.11 Sea A una matriz real simétrica $n \times n$. Recordemos que en tal caso, todos los autovalores de A son números reales, y que existe una matriz $n \times n$ P que es ortogonal (es decir tal que $PP^* = I_n$) y $A = PDP^*$, siendo D la matriz diagonal formada por los autovalores de A (contados cada uno como indique su multiplicidad). Teniendo en cuenta lo anterior, es inmediato comprobar que se satisface el siguiente resultado (demuéstrese como ejercicio):

Proposición 5.12 Una matriz real simétrica es definida positiva si y sólo si todos sus autovalores son positivos.

Proposición 5.13 Sean A y B matrices $n \times n$. Entonces, la matriz producto BAB^* es definida positiva si y sólo si A es definida positiva y B es regular.

Demostración.-

a) Supongamos que A es definida positiva y B es regular. En tal caso, B^* también es regular, y en consecuencia, si $u \in \mathbb{R}^n$ no es el cero, $B^*u \neq 0$, y por tanto

$$u^*(BAB^*)u = (u^*B)A(B^*u) = (B^*u)^*A(B^*u) > 0,$$

con lo que BAB^* es definida positiva.

b) Supongamos que BAB^* es definida positiva.

Si B no fuese regular, existiría $v \in \mathbb{R}^n \setminus \{0\}$, tal que $B^*v = 0$. Pero entonces, para ese v ,

$$0 < v^*(BAB^*)v = (v^*B)A(B^*v) = 0^*A0 = 0,$$

lo cual es una contradicción. En consecuencia, B es regular.

Si A no fuese definida positiva, existiría $w \in \mathbb{R}^n \setminus \{0\}$, tal que $w^*Aw \leq 0$. En tal caso, como B es regular, existe un $v \in \mathbb{R}^n \setminus \{0\}$, tal que $B^*v = w$, y entonces

$$0 < v^*(BAB^*)v = (v^*B)A(B^*v) = (B^*v)^*A(B^*v) = w^*Aw \leq 0,$$

lo cual es una contradicción. En consecuencia, A es definida positiva. ■

Teorema 5.14 (de factorización de Cholesky) *Si A es una matriz $n \times n$ simétrica definida positiva, entonces existe al menos una matriz $n \times n$ triangular inferior B tal que $A = BB^*$. Además, se puede imponer que $b_{ii} > 0$ para todo $i = 1, \dots, n$, y en tal caso la factorización anterior es única.*

Demostración.-

a) Existencia de la factorización $A = BB^*$.

Por el Corolario 5.8, la matriz A posee factorización LU , y por ser A simétrica, por la Proposición 5.1 la matriz A se puede escribir en la forma $A = LDL^*$, con D diagonal. Pero entonces, por la Proposición 5.13 la matriz D es definida positiva, con lo que $d_{ii} > 0$ para todo $i = 1, \dots, n$, y por tanto, nuevamente por la Proposición 5.1, si definimos $B = LD^{1/2}$, entonces B es triangular inferior y tal que $A = BB^*$. Obsérvese finalmente que $b_{ii} = l_{ii}\sqrt{d_{ii}}$, con lo que si tomamos para cada $i = 1, \dots, n$ la raíz $\sqrt{d_{ii}}$ con el mismo signo que el de l_{ii} , tendremos garantizado que $b_{ii} > 0$.

b) Unicidad de la factorización $A = BB^*$ con $b_{ii} > 0$.

Por la libertad en la elección del signo para las $\sqrt{d_{ii}}$, está claro la no unicidad de la factorización $A = BB^*$, pero vamos ahora a obtener dicha unicidad con el requerimiento adicional de que todos los b_{ii} sean positivos.

Sean $B^{(1)}$ y $B^{(2)}$ dos matrices $n \times n$ triangulares inferior tales que los elementos de la diagonal principal de ambas sean positivos, es decir, $b_{ii}^{(1)} > 0$, $b_{ii}^{(2)} > 0$, para todo $i = 1, \dots, n$, y satisfagan

$$A = B^{(1)}(B^{(1)})^* = B^{(2)}(B^{(2)})^*.$$

En tal caso,

$$(B^{(2)})^{-1}B^{(1)} = (B^{(2)})^*((B^{(1)})^*)^{-1} = (B^{(2)})^*((B^{(1)})^{-1})^* = ((B^{(1)})^{-1}B^{(2)})^*,$$

con lo que al ser $(B^{(2)})^{-1}B^{(1)}$ triangular inferior y $((B^{(1)})^{-1}B^{(2)})^*$ triangular superior,

$$(B^{(2)})^{-1}B^{(1)} = ((B^{(1)})^{-1}B^{(2)})^* = \widetilde{D}, \quad (5.32)$$

con \widetilde{D} matriz diagonal

Pero $(B^{(2)})^{-1}$ es una matriz triangular inferior cuyos elementos de la diagonal principal son de la forma $1/b_{ii}^{(2)}$, y en consecuencia los elementos de la diagonal principal de $(B^{(2)})^{-1}B^{(1)}$ son de la forma $b_{ii}^{(1)}/b_{ii}^{(2)}$.

De manera análoga, se obtiene que los elementos de la diagonal principal de $(B^{(1)})^{-1}B^{(2)}$ son de la forma $b_{ii}^{(2)}/b_{ii}^{(1)}$.

En consecuencia, teniendo en cuenta la igualdad (5.32), tenemos que

$$\frac{b_{ii}^{(1)}}{b_{ii}^{(2)}} = \frac{b_{ii}^{(2)}}{b_{ii}^{(1)}}, \quad \text{para todo } i = 1, \dots, n,$$

es decir,

$$(b_{ii}^{(1)})^2 = (b_{ii}^{(2)})^2, \quad \text{para todo } i = 1, \dots, n,$$

con lo que, teniendo en cuenta que los $b_{ii}^{(1)}$ y los $b_{ii}^{(2)}$ son positivos, podemos afirmar que

$$b_{ii}^{(1)} = b_{ii}^{(2)}, \quad \text{para todo } i = 1, \dots, n,$$

y en consecuencia $\widetilde{D} = I_n$, la matriz identidad $n \times n$, y nuevamente por (5.32), obtenemos $B^{(1)} = B^{(2)}$. ■

El recíproco del Teorema de factorización de Cholesky es también cierto.

Proposición 5.15 *Si B es una matriz $n \times n$ triangular inferior regular, y definimos $A = BB^*$, entonces A es simétrica y definida positiva.*

Demostración.- En primer lugar,

$$A^* = (BB^*)^* = BB^* = A,$$

y por tanto A es simétrica.

Además, si $u \in \mathbb{R}^n \setminus \{0\}$, entonces por ser B regular, también $B^*u \in \mathbb{R}^n \setminus \{0\}$. Pero entonces,

$$u^*Au = u^*BB^*u = (B^*u)^*B^*u = \sum_{j=1}^n (B^*u)_j^2 > 0,$$

y por tanto A es definida positiva. ■

Para el cálculo efectivo de la factorización de Cholesky, se parte de la igualdad

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{n,n} \end{pmatrix} = \begin{pmatrix} b_{11} & 0 & \dots & 0 & 0 \\ b_{21} & b_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n-1} & 0 \\ b_{n,1} & b_{n,2} & \dots & b_{n,n-1} & b_{n,n} \end{pmatrix} \begin{pmatrix} b_{11} & b_{21} & \dots & b_{n-1,1} & b_{n1} \\ 0 & b_{22} & \dots & b_{n-1,2} & b_{n2} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & b_{n-1,n-1} & b_{n,n-1} \\ 0 & 0 & \dots & 0 & b_{n,n} \end{pmatrix}$$

con $b_{ii} > 0$.

De esta forma, identificando los elementos de la primera fila de A con los correspondientes de BB^* , obtenemos

$$\begin{cases} a_{11} = b_{11}^2 \Rightarrow b_{11} = \sqrt{a_{11}} \\ a_{1j} = b_{11}b_{j1} \Rightarrow b_{j1} = \frac{a_{1j}}{b_{11}}, \quad j = 2, \dots, n, \end{cases}$$

identificando a continuación los elementos de la segunda fila de A con los correspondientes de BB^* , obtenemos

$$a_{2j} = b_{21}b_{j1} + b_{22}b_{j2}, \quad j = 2, \dots, n,$$

y en consecuencia,

$$\begin{cases} b_{22} = \sqrt{a_{22} - b_{21}^2}, \\ b_{j2} = \frac{1}{b_{22}}(a_{2j} - b_{21}b_{j1}), \quad j = 3, \dots, n, \end{cases}$$

y en general si suponemos conocidas las $k - 1$ primeras columnas de B entonces, identificando los elementos de la k -ésima fila de A con los correspondientes de BB^* , obtenemos

$$a_{kj} = b_{k1}b_{j1} + b_{k2}b_{j2} + \dots + b_{kk}b_{jk}, \quad j = k, \dots, n,$$

y en consecuencia,

$$\begin{cases} b_{kk} = \sqrt{a_{kk} - \sum_{r=1}^{k-1} b_{kr}^2}, \\ b_{jk} = \frac{1}{b_{kk}} \left(a_{kj} - \sum_{r=1}^{k-1} b_{kr}b_{jr} \right), \quad j = k + 1, \dots, n. \end{cases}$$

Se puede demostrar que el número de operaciones necesarias para llevar a cabo el método de cálculo anterior resulta ser aproximadamente la mitad de sumas y productos que en el método de Gauss o el LU , a cambio de calcular además n raíces cuadradas. Cuando n es grande, el método de Cholesky es el más favorable para el caso de matrices simétricas y definidas positivas.

Como ejercicio, se propone que se encuentre la factorización de Cholesky de la matriz

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix},$$

y se resuelva el sistema $Au = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix}$.

Referencias

- [1] R.L. Burden & J.D. Faires, *Análisis Numérico*, Grupo Editorial Iberoamérica, México 1985.
- [2] D. Kincaid & W. Cheney, *Análisis Numérico*, Addison-Wesley Iberoamericana, Wilmington, 1994.

Como bibliografía complementaria se pueden consultar:

- [3] P.G. Ciarlet, *Introduction à l'Analyse Numérique Matricielle et à l'Optimization*, Masson, Paris 1982.
- [4] J. Stoer & R. Bulirsch, *Introduction to Numerical Analysis*, W.H. Freeman and Co., San Francisco, 1975.
- [5] R. Théodor, *Initiation à l'Analyse Numérique*, Masson, Paris, 1989.