

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
FACULTAD DE CIENCIAS EXACTAS, INGENIERÍA Y AGRIMENSURA
(FCEIA-UNR)

MÉTODOS NUMÉRICOS

Prof. Alejandro G. Marchetti

Unidad IV

Sistemas de Ecuaciones Lineales Métodos Directos



Octubre de 2020

Los métodos directos para resolver sistemas de ecuaciones lineales son métodos con un número finito de pasos, y obtienen la solución exacta provisto que todas las operaciones aritméticas sean exactas. El método directo más conocido y utilizado es la eliminación Gaussiana.

1. Eliminación de Gauss

El método de eliminación de Gauss consiste en operar sobre la matriz ampliada del sistema hasta hallar la forma escalonada (una matriz triangular superior). Así, se obtiene un sistema fácil de resolver por sustitución regresiva. Las dos etapas del métodos son:

- 1) Eliminación progresiva de incógnitas. Consiste en transformar el sistema en un sistema triangular superior usando operaciones elementales sobre las filas de la matriz ampliada.
- 2) Resolución del sistema triangular superior mediante sustitución regresiva.

Denotamos el sistema lineal original como $\mathbf{A}^{(1)}\mathbf{x} = \mathbf{b}^{(1)}$, y definimos la matriz ampliada

$$[\mathbf{A}|\mathbf{b}] = [\mathbf{A}^{(1)}|\mathbf{b}^{(1)}] = \left[\begin{array}{ccc|c} a_{11}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ \vdots & & \vdots & \vdots \\ a_{n1}^{(1)} & \dots & a_{nn}^{(1)} & b_n^{(1)} \end{array} \right]$$

Este sistema inicial se reduce en $n - 1$ pasos a la forma:

$$[\mathbf{A}^{(n)}|\mathbf{b}^{(n)}] = \left[\begin{array}{cccc|c} a_{11}^{(1)} & \dots & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & a_{nn}^{(n)} & b_n^{(n)} \end{array} \right]$$

Luego, el sistema triangular superior $\mathbf{A}^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$ se resuelve por sustitución regresiva.

1.1. Algoritmo de Eliminación de Gauss

PASO 1: Supongamos que $a_{11}^{(1)} \neq 0$. Sean los multiplicadores de fila

$$m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \quad i = 2, 3, \dots, n$$

Estos multiplicadores se usan para eliminar x_1 de las ecuaciones 2 a n . Definimos

$$\begin{aligned} a_{ij}^{(2)} &= a_{ij}^{(1)} - m_{i1}a_{1j}^{(1)}, & i, j &= 2, \dots, n \\ b_i^{(2)} &= b_i^{(1)} - m_{i1}b_1^{(1)}, & i &= 2, \dots, n \end{aligned}$$

La primera fila de la matriz ampliada $[\mathbf{A}^{(1)}|\mathbf{b}^{(1)}]$ no se modifica, y la primera columna de $\mathbf{A}^{(1)}$, debajo de la diagonal, se lleva a cero. La matriz ampliada $[\mathbf{A}^{(2)}|\mathbf{b}^{(2)}]$ tiene la forma

$$[\mathbf{A}^{(2)}|\mathbf{b}^{(2)}] = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right] \begin{array}{l} \leftarrow E_1^{(1)} \\ \leftarrow E_2^{(2)} \\ \vdots \\ \leftarrow E_n^{(2)} \end{array}$$

donde empleamos la notación $E_i^{(k)}$ para denotar la fila i de la matriz ampliada en el paso k . Esta fila representa la ecuación i del sistema $A^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$.

PASO k: Suponga que para $i = 1, \dots, k-1$ los x_i han sido eliminados de las ecuaciones $i+1, \dots, n$. Tenemos:

$$[A^{(k)}|\mathbf{b}^{(k)}] = \left[\begin{array}{cccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \dots & \dots & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \ddots & \ddots & & & \vdots & \vdots \\ \vdots & & 0 & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} & b_k^{(k)} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} & b_n^{(k)} \end{array} \right] \begin{array}{l} \leftarrow E_1^{(1)} \\ \leftarrow E_2^{(2)} \\ \vdots \\ \leftarrow E_k^{(k)} \\ \vdots \\ \leftarrow E_n^{(k)} \end{array}$$

Supongamos que el elemento pivote $a_{kk}^{(k)} \neq 0$. Definimos los multiplicadores de fila

$$m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k+1, \dots, n$$

Estos multiplicadores se usan para eliminar x_k de las ecuaciones $k+1, \dots, n$:

$$E_i^{(k+1)} = E_i^{(k)} - m_{ik} E_k^{(k)}, \quad i = k+1, \dots, n$$

Continuando de esta manera, después de $n-1$ pasos, obtenemos el sistema escalonado (triangular superior) $A^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$.

Sustitución regresiva: Primero obtenemos x_n de la última ecuación del sistema escalonado $A^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$,

$$x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}}$$

Este resultado se puede sustituir hacia atrás en la $(n-1)$ -ésima ecuación y despejar x_{n-1} , y así para las incógnitas restantes.

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

Esto completa el algoritmo de eliminación de Gauss.

Ejemplo 1 Resolver el sistema lineal

$$\begin{array}{rrcr} x_1 + 2x_2 + & x_3 & = & 0 \\ 2x_1 + 2x_2 + 3x_3 & = & 3 \\ -x_1 - 3x_2 & = & 2 \end{array}$$

Representamos este sistema lineal con la matriz ampliada

$$[A|\mathbf{b}] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 2 & 3 & 3 \\ -1 & -3 & 0 & 2 \end{array} \right]$$

En el siguiente diagrama, las flechas indican los pasos de la eliminación, y los multiplicadores utilizados se indican al costado de las flechas:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 2 & 3 & 3 \\ -1 & -3 & 0 & 2 \end{array} \right] \xrightarrow[m_{31}=-1]{m_{21}=2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -2 & 1 & 3 \\ 0 & -1 & 1 & 2 \end{array} \right] \xrightarrow{m_{32}=\frac{1}{2}} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{array} \right]$$

Resolviendo el sistema $A^{(3)}\mathbf{x} = \mathbf{b}^{(3)}$,

$$x_3 = 1, \quad x_2 = -1, \quad x_1 = 1$$

1.2. Pivoteo Parcial

En cada paso del proceso de eliminación visto en la sección anterior, supusimos que los elementos pivote eran distintos de cero, $a_{kk}^{(k)} \neq 0$. Para eliminar esta hipótesis, podemos emplear pivoteo parcial. En el caso de que $a_{kk}^{(k)} = 0$, examinamos los elementos $a_{ik}^{(k)}$ en las filas $E_i^{(k)}$ para $i = k+1, \dots, n$. Siendo A no singular, se puede demostrar que al menos uno de dichos elementos es distinto de cero. La ecuación $E_i^{(k)}$ con $a_{ik}^{(k)} \neq 0$ se intercambia con $E_k^{(k)}$ y luego se continua el proceso de eliminación.

Ejemplo 2 Resolver el sistema lineal

$$6x_1 + 2x_2 + 2x_3 = -2$$

$$2x_1 + \frac{2}{3}x_2 + \frac{1}{3}x_3 = 1$$

$$x_1 + 2x_2 - x_3 = 0$$

utilizando una calculadora decimal con una mantisa de 4 dígitos.

La solución exacta es

$$x_1 = 2,6 \quad x_2 = -3,8 \quad x_3 = -5,0$$

Veamos la solución que se obtiene aplicando el proceso de eliminación de gauss con la aritmética de la calculadora.

$$\left[\begin{array}{ccc|c} 6,000 & 2,000 & 2,000 & -2,000 \\ 2,000 & 0,6667 & 0,3333 & 1,000 \\ 1,000 & 2,000 & -1,000 & 0,000 \end{array} \right] \xrightarrow[m_{31}=0,1667]{m_{21}=0,3333} \left[\begin{array}{ccc|c} 6,000 & 2,000 & 2,000 & -2,000 \\ 0,000 & 0,0001000 & -0,3333 & 1,667 \\ 0,000 & 1,667 & -1,333 & 0,3334 \end{array} \right]$$

$$\xrightarrow{m_{32}=16670} \left[\begin{array}{ccc|c} 6,000 & 2,000 & 2,000 & -2,000 \\ 0,000 & 0,0001000 & -0,3333 & 1,667 \\ 0,000 & 0,000 & 5555 & -27790 \end{array} \right]$$

Resolviendo por sustitución regresiva obtenemos

$$x_1 = 1,335 \quad x_2 = -0,000 \quad x_3 = -5,003$$

Vemos que la solución obtenida difiere mucho de la solución verdadera. Este error se debe a que el elemento pivote $a_{22}^{(2)} = 0,0001000$ debió haber sido igual a cero, pero no lo es debido a los errores de redondeo. Este elemento pivote tiene un error relativo infinito.

Para evitar este problema podemos intercambiar las ecuaciones $E_2^{(2)}$ y $E_3^{(2)}$, y luego continuar la eliminación:

$$\left[\begin{array}{ccc|c} 6,000 & 2,000 & 2,000 & -2,000 \\ 0,000 & 1,667 & -1,333 & 0,3334 \\ 0,000 & 0,0001000 & -0,3333 & 1,667 \end{array} \right] \xrightarrow{m_{32}=0,00005999} \left[\begin{array}{ccc|c} 6,000 & 2,000 & 2,000 & -2,000 \\ 0,000 & 1,667 & -1,333 & 0,3334 \\ 0,000 & 0,000 & -0,3332 & 1,667 \end{array} \right]$$

Mediante sustitución regresiva obtenemos en este caso

$$x_1 = -2,602 \quad x_2 = -3,801 \quad x_3 = -5,003$$

que se aproxima mucho más a la solución real del sistema.

Vemos en el Ejemplo 2 que no es suficiente simplemente pedir que los elementos pivote sean distintos de cero. Puede ocurrir que un elemento sea distinto de cero debido a errores de redondeo, lo cual conduce a errores gruesos en los cálculos subsiguientes. Para evitar esto, introducimos el siguiente procedimiento de pivoteo parcial:

En el paso k , calcular

$$c = \max_{k \leq i \leq n} |a_{ik}^{(k)}|$$

Si el elemento $|a_{kk}^{(k)}| < c$, luego intercambiar la ecuación $E_k^{(k)}$ con la correspondiente ecuación tal que $|a_{kk}^{(k)}| = c$.

Con el pivoteo parcial, los multiplicadores m_{ik} satisfacen:

$$|m_{ik}| \leq 1, \quad i = k+1, \dots, n \\ k = 1, \dots, n-1$$

Además del pivoteo parcial, existe también el pivoteo completo, en el cual, tanto en las columnas como en las filas se busca el elemento de mayor valor absoluto y luego se intercambian. El pivoteo completo agrega complejidad al cambiar el orden de las variables x .

1.3. Número de Operaciones del Método de Gauss

Para analizar el número de operaciones necesarias para resolver el sistema $\mathbf{Ax} = \mathbf{b}$ por eliminación de Gauss, consideraremos separadamente la generación de $\mathbf{A}^{(n)}$ a partir de $\mathbf{A}^{(1)}$, la modificación de $\mathbf{b}^{(1)}$ a $\mathbf{b}^{(n)}$, y finalmente la obtención de la solución \mathbf{x} por sustitución regresiva.

- 1) Cálculo de $\mathbf{A}^{(n)}$. En el paso 1, se necesitan $n-1$ divisiones para calcular los multiplicadores m_{i1} , $2 \leq i \leq n$. Luego, se usan $(n-1)^2$ multiplicaciones y $(n-1)^2$ sumas para crear los nuevos elementos $a_{ij}^{(2)}$. Continuando de esta manera en cada paso, obtenemos el conteo de operaciones indicado en la Tabla 1.

Tabla 1: Conteo de operaciones en la eliminación gaussiana.

Paso	Sumas	Multiplicaciones	Divisiones
1	$(n-1)^2$	$(n-1)^2$	$n-1$
2	$(n-2)^2$	$(n-2)^2$	$n-2$
\vdots	\vdots	\vdots	\vdots
$n-1$	1	1	1
Total	$\frac{n(n-1)(2n-1)}{6}$	$\frac{n(n-1)(2n-1)}{6}$	$\frac{n(n-1)}{2}$

Denotamos por $SR(\cdot)$ el número de sumas y restas, y por $MD(\cdot)$ el número de multiplicaciones y divisiones. Luego,

$$SR(\mathbf{A}^{(1)} \rightarrow \mathbf{A}^{(n)}) = \frac{n(n-1)(2n-1)}{6} \cong \frac{n^3}{3} \\ MD(\mathbf{A}^{(1)} \rightarrow \mathbf{A}^{(n)}) = \frac{n(n-1)(2n-1)}{6} + \frac{n(n-1)}{2} = \frac{n(n^2-1)}{3} \cong \frac{n^3}{3}$$

siendo las estimaciones finales válidas para valores grandes de n .

2) Modificación de $\mathbf{b}^{(1)}$ a $\mathbf{b}^{(n)}$.

$$SR(\mathbf{b}^{(1)} \rightarrow \mathbf{b}^{(n)}) = (n-1) + (n-2) + \cdots + 1 = \frac{n(n-1)}{2}$$

$$MD(\mathbf{b}^{(1)} \rightarrow \mathbf{b}^{(n)}) = \frac{n(n-1)}{2}$$

3) Solución de $\mathbf{A}^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$.

$$SR(\mathbf{x}) = \frac{n(n-1)}{2}$$

$$MD(\mathbf{x}) = \frac{n(n+1)}{2}$$

Vemos que, para valores de n grandes, el principal costo computacional de la eliminación de Gauss se da en la generación de $\mathbf{A}^{(n)}$ a partir de $\mathbf{A}^{(1)}$. Puede entonces afirmarse que para sistemas de grandes dimensiones el número de operaciones de la eliminación de Gauss es del orden de $\frac{2n^3}{3}$.

1.4. Casos Especiales

Definición 1 Decimos que una matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ es estrictamente diagonal dominante si

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

Teorema 1 Toda matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ estrictamente diagonal dominante es no singular. Para estas matrices, el sistema $\mathbf{Ax} = \mathbf{b}$ se puede resolver por eliminación de Gauss sin necesidad de pivoteo.

Teorema 2 Para toda matriz \mathbf{A} simétrica y definida positiva, el sistema $\mathbf{Ax} = \mathbf{b}$ se puede resolver por eliminación de Gauss sin necesidad de pivoteo, siendo todos los elementos pivotes positivos.

1.5. Método de Gauss-Jordan

El método de Gauss-Jordan es una variante de la eliminación de Gauss en el que se eliminan las incógnitas tanto por encima como por debajo de la diagonal. En este método, la matriz ampliada $[\mathbf{A}|\mathbf{b}]$ se convierte luego de n pasos en la matriz $[\mathbf{I}|\mathbf{x}]$, obteniéndose de este modo la solución \mathbf{x} .

PASO k:

1) Normalizar $E_k^{(k)}$ dividiéndolo por $a_{kk}^{(k)}$.

$$a_{kj}^{k+1} = \frac{a_{kj}^{(k)}}{a_{kk}^{(k)}}, \quad j = k, \dots, n$$

$$b_k^{(k+1)} = \frac{b_k^{(k)}}{a_{kk}^{(k)}}$$

2) Eliminación de x_k en las ecuaciones por encima y por debajo de la ecuación k .

$$\begin{aligned} a_{ij}^{k+1} &= a_{ij}^{(k)} - a_{ik}^{(k)} a_{kj}^{(k+1)} \\ b_i^{(k+1)} &= b_i^{(k)} - a_{ik}^{(k)} b_k^{(k+1)} \end{aligned}$$

para $j = k, \dots, n, i = 1, \dots, n, i \neq k$.

El método de Gauss-Jordan usa del orden de $\frac{n^3}{2}$ multiplicaciones y divisiones y $\frac{n^3}{2}$ sumas y restas, por lo que requiere un mayor número de operaciones que la eliminación de Gauss.

2. Factorización LU

2.1. Factorización LU a partir de la Eliminación Gaussiana

Mediante la eliminación de Gauss, supuesta sin pivoteo, obtenemos el sistema triangular superior $\mathbf{U}\mathbf{x} = \mathbf{g}$, con $\mathbf{U} = \mathbf{A}^{(n)}$, y $\mathbf{g} = \mathbf{b}^{(n)}$,

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & u_{nn} \end{bmatrix} \quad \text{con } u_{ij} = a_{ij}^{(i)}$$

Definimos la matriz triangular inferior \mathbf{L} , basada en los multiplicadores m_{ik} de la eliminación gaussiana,

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ m_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & 1 \end{bmatrix}$$

Teorema 3 (Factorización LU) Sea $\mathbf{A} \in \mathbb{R}^{n \times n}$ una matriz no singular, y sean \mathbf{L} y \mathbf{U} las matrices triangular inferior y triangular superior, definidas anteriormente usando la eliminación de Gauss. Luego, si \mathbf{U} es generada sin pivoteo se tiene

$$\mathbf{A} = \mathbf{LU}$$

Resolver el sistema $\mathbf{Ax} = \mathbf{b}$ es equivalente a resolver $\mathbf{LUx} = \mathbf{b}$, lo cual es equivalente a resolver dos sistemas triangulares:

$$\begin{aligned} \mathbf{Lg} &= \mathbf{b} && \text{Sistema triangular inferior. Se resuelve por sustitución progresiva.} \\ \mathbf{Ux} &= \mathbf{g} && \text{Sistema triangular superior. Se resuelve por sustitución regresiva.} \end{aligned}$$

Para resolver la sustitución progresiva y regresiva se requieren en total alrededor de n^2 multiplicaciones y divisiones y n^2 sumas y restas. Para determinar las matrices \mathbf{L} y \mathbf{U} se requiere el mismo número de operaciones que las que se necesitan para generar la matriz $\mathbf{A}^{(n)} = \mathbf{U}$ en la eliminación de Gauss. Sin embargo, una vez que se tiene la factorización, los sistemas en los que intervenga la matriz \mathbf{A} pueden resolverse fácilmente para cualquier número de vectores \mathbf{b} .

En la programación de la factorización LU a partir de la eliminación gaussiana, los elementos $a_{ij}^{(k+1)}$, $j \geq i$, siempre se almacenan reemplazando los elementos $a_{ij}^{(k)}$. A medida que los elementos debajo de la diagonal se hacen iguales a cero, es conveniente almacenar en dicho espacio los multiplicadores m_{ij} , ocupando el espacio utilizado originalmente para almacenar los elementos a_{ij} , $i > j$.

2.2. Unicidad de la factorización LU

Teorema 4 (Unicidad de la factorización LU) Si $A \in \mathbb{R}^{n \times n}$ es tal que la eliminación de Gauss puede realizarse sin pivoteo, luego A puede factorizarse como $A = LU$, donde $U = A^{(n)}$ es el resultado final de la eliminación de Gauss aplicada a A , y L es una matriz triangular inferior con $l_{ii} = 1$, $i = 1, \dots, n$, y $l_{ij} = m_{ij}$, $i > j$. Dicha factorización es única.

Demostración. Demostraremos la unicidad de la factorización LU. Notar que los factores L y U son no singulares ya que son matrices triangulares con elementos diagonales distintos de cero. Supongamos que $A = L_1 U_1 = L_2 U_2$ son dos factorizaciones LU de A , luego

$$\begin{aligned} L_2^{-1} L_1 U_1 &= U_2 \\ L_2^{-1} L_1 &= U_2 U_1^{-1} \end{aligned} \quad (1)$$

Sabemos que:

- La inversa de una matriz triangular inferior (superior) es una matriz triangular inferior (superior).
- El producto de dos matrices triangulares inferiores (superiores) es una matriz triangular inferior (superior).

Por lo tanto, $L_2^{-1} L_1$ es triangular inferior, mientras que $U_2 U_1^{-1}$ es triangular superior. Luego, la ecuación (1) implica $L_2^{-1} L_1 = U_2 U_1^{-1} = D$, siendo D una matriz diagonal. Como $[L_2]_{ii} = [L_2^{-1}]_{ii} = [L_1]_{ii} = 1$, tenemos $D = I$, y por ende $L_1 = L_2$ y $U_1 = U_2$. \square

2.3. Factorización LU con Matriz de Permutación

En el razonamiento previo hemos supuesto que A es tal que el sistema $Ax = b$ puede resolverse mediante el método de eliminación de Gauss sin intercambios de filas (sin pivoteo). Cuando se requieren intercambios de filas se puede usar una matriz de permutación.

Definición 2 (Matriz de permutación) Una matriz de permutación $P \in \mathbb{R}^{n \times n}$, es una matriz en la que hay exactamente una entrada cuyo valor es 1 en cada fila y en cada columna, siendo todas las demás entradas iguales a 0.

Ejemplo:

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

Teorema 5 Para toda matriz $A \in \mathbb{R}^{n \times n}$ no singular, existe una matriz de permutación P tal que PA posee una factorización LU, es decir, $PA = LU$.

Las matrices P , L y U se pueden generar al programar la eliminación de Gauss con pivoteo parcial, teniendo en cuenta los intercambios de filas requeridos.

Una vez obtenida la factorización $PA = LU$, el sistema $Ax = b$ se resuelve permutando primero los elementos en b para construir $\tilde{b} = Pb$. Luego se resuelve los sistemas triangulares $Ly = \tilde{b}$ y $Ux = y$ por sustitución progresiva y regresiva, respectivamente.

2.4. Método de Doolittle

El método de Doolittle es un método de factorización LU denominado compacto, porque es sencillo de programar y requiere poca memoria una vez implementado en el ordenador. Existen muchos métodos compactos de factorización de matrices, y suelen diferir en el tratamiento que hacen de los elementos de la diagonal de las matrices L y/o U . En concreto, el método de Doolittle genera una matriz L que tiene 1 en todos los elementos de la diagonal. Este método proporciona una ventaja de cálculo en ordenadores modernos, que cuentan con memoria caché además de la memoria principal. En este tipo de arquitecturas de ordenador, la secuencia en la que se realizan los cálculos puede ser más importante que la cantidad de cálculos que se realizan. **El método de Doolittle tiene una secuencia de operaciones óptima para ejecutarse en un ordenador con memoria caché.**

El resultado del método serán dos matrices triangulares tales que $A = LU$. La matriz L será triangular inferior y de diagonal unitaria. La matriz U será triangular superior. Al ser las matrices triangulares, podemos calcular cada elemento de la matriz A mediante

$$a_{ij} = \sum_{k=1}^{\min(i,j)} l_{ik}u_{kj}$$

donde a , l y u se refieren a los elementos de las matrices A , L y U , respectivamente. El límite superior de la suma se debe a que las matrices triangulares tienen ceros en su mitad inferior o superior, y por tanto, el producto matricial sumaría cero.

En el caso de que $i \leq j$, es decir, el número de fila es más pequeño que el número de la columna, en otras palabras, estamos por encima de la diagonal (o en la diagonal), entonces $\min(i, j) = i$ y tendremos

$$l_{ii}u_{ij} = u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}$$

ya que $l_{ii} = 1$. Es decir, hemos despejado del sumatorio el sumando cuando $k = i$, y eso nos permite obtener una expresión genérica para calcular u_{ij} .

Si empezamos en $i = 1$ la ecuación anterior nos dice que $u_{1j} = a_{1j}$, por lo que ya hemos calculado la primera fila de la matriz U . Para seguir calculando filas de U es necesario conocer los valores de L . Si nos movemos por debajo de la diagonal, $j \leq i$, tendremos que $\min(i, j) = j$, y obtenemos

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}}{u_{jj}}$$

Por tanto, podemos calcular $l_{i1} = \frac{a_{i1}}{u_{11}}$, es decir, hemos calculado ya la primera columna de L . Al conocer la primera columna de L podemos intentar calcular la siguiente fila de U . Al terminar una fila de U , ya conocemos los elementos para calcular la siguiente columna de L . Aplicando sucesivamente este método, obtendremos de manera completa las matrices L y U .

Ejemplificamos para un sistema de 3×3 .

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

Luego

$$\begin{aligned} a_{11} &= u_{11}, & a_{12} &= u_{12}, & a_{13} &= u_{13} \\ a_{21} &= m_{21}u_{11}, & a_{31} &= m_{31}u_{11} \end{aligned}$$

con lo cual obtenemos la primera columna de \mathbf{L} y la primera fila de \mathbf{U} (suponiendo $u_{11} \neq 0$).

Multiplicamos la fila 2 de \mathbf{L} por las columnas 2 y 3 de \mathbf{U} :

$$\begin{aligned} a_{22} &= m_{21}u_{12} + u_{22} \\ a_{23} &= m_{21}u_{13} + u_{23} \end{aligned}$$

De estas ecuaciones obtenemos u_{22} y u_{23} .

Multiplicamos la fila 3 de \mathbf{L} por las columnas 2 y 3 de \mathbf{U} :

$$\begin{aligned} m_{31}u_{12} + m_{32}u_{22} &= a_{32} \\ m_{31}u_{13} + m_{32}u_{23} + u_{33} &= a_{33} \end{aligned}$$

De estas ecuaciones obtenemos m_{32} y u_{33} (suponiendo que $u_{22} \neq 0$).

2.5. Descomposición de Crout

Es una descomposición LU alternativa que usa una matriz \mathbf{U} con números 1 en la diagonal. Se resuelve de manera similar al método de Doolittle. Ejemplificamos para un sistema de 3×3 .

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

$$\begin{aligned} l_{11} &= a_{11}, & l_{21} &= a_{21}, & l_{31} &= a_{31} \\ u_{12} &= \frac{a_{12}}{l_{11}}, & u_{13} &= \frac{a_{13}}{l_{11}} \\ l_{22} &= a_{22} - l_{21}u_{12}, & l_{32} &= a_{32} - l_{31}u_{12} \\ u_{23} &= \frac{a_{23} - l_{21}u_{13}}{l_{22}}, & l_{33} &= a_{33} - l_{31}u_{13} - l_{32}u_{23} \end{aligned}$$

3. Factorización de Cholesky

Teorema 6 La matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ es definida positiva si y solo si existe una única matriz triangular superior \mathbf{R} con elementos diagonales positivos tal que $\mathbf{A} = \mathbf{R}^T \mathbf{R}$. Esta es la factorización de Cholesky de \mathbf{A} .

Resolver el sistema $\mathbf{Ax} = \mathbf{b}$ es equivalente a resolver $\mathbf{R}^T \mathbf{Rx} = \mathbf{b}$, lo cual es equivalente a resolver dos sistemas triangulares:

$$\begin{aligned} \mathbf{R}^T \mathbf{g} &= \mathbf{b} && \text{Sistema triangular inferior. Se resuelve por sustitución progresiva.} \\ \mathbf{Rx} &= \mathbf{g} && \text{Sistema triangular superior. Se resuelve por sustitución regresiva.} \end{aligned}$$

4. Factorización QR

Sea $A \in \mathbb{R}^{m \times n}$, $A = [\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n]$ una matriz con columnas $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ linealmente independientes (esto implica $m \geq n$). Aplicando Gram-Schmidt a las columnas de A , resulta una base ortogonal normalizada $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ del espacio columna de A , donde

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{\nu_1} \quad \text{y} \quad \mathbf{q}_k = \frac{\mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{a}_k^\top \mathbf{q}_i) \mathbf{q}_i}{\nu_k}, \quad k = 2, \dots, n$$

donde

$$\nu_1 = \|\mathbf{a}_1\| \quad \text{y} \quad \nu_k = \left\| \mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{a}_k^\top \mathbf{q}_i) \mathbf{q}_i \right\|, \quad k = 2, \dots, n$$

Reordenando estas ecuaciones, tenemos

$$\begin{aligned} \mathbf{a}_1 &= \nu_1 \mathbf{q}_1 \\ \mathbf{a}_k &= (\mathbf{a}_k^\top \mathbf{q}_1) \mathbf{q}_1 + \dots + (\mathbf{a}_k^\top \mathbf{q}_{k-1}) \mathbf{q}_{k-1} + \nu_k \mathbf{q}_k, \quad k = 2, \dots, n \end{aligned}$$

En forma matricial:

$$[\mathbf{a}_1 | \mathbf{a}_2 | \dots | \mathbf{a}_n] = [\mathbf{q}_1 | \mathbf{q}_2 | \dots | \mathbf{q}_n] \begin{bmatrix} \nu_1 & \mathbf{a}_2^\top \mathbf{q}_1 & \mathbf{a}_3^\top \mathbf{q}_1 & \dots & \mathbf{a}_n^\top \mathbf{q}_1 \\ 0 & \nu_2 & \mathbf{a}_3^\top \mathbf{q}_2 & \dots & \mathbf{a}_n^\top \mathbf{q}_2 \\ 0 & 0 & \nu_3 & \dots & \mathbf{a}_n^\top \mathbf{q}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \nu_n \end{bmatrix}$$

o bien,

$$A = QR$$

donde

- $A \in \mathbb{R}^{m \times n}$: matriz con columnas linealmente independientes
- $Q \in \mathbb{R}^{m \times n}$: base ortogonal normalizada del espacio columna de A
- $R \in \mathbb{R}^{n \times n}$: matriz triangular superior con elementos diagonales positivos

Teorema 7 (Factorización QR) Toda matriz $A \in \mathbb{R}^{m \times n}$ con columnas linealmente independientes puede factorizarse de manera única como $A = QR$ con Q y R definidas anteriormente.

Si $A \in \mathbb{R}^{n \times n}$ es no singular, tenemos $Q^\top = Q^{-1}$, ya que Q tiene columnas ortogonales normalizadas. El sistema $A\mathbf{x} = Q\mathbf{R}\mathbf{x} = \mathbf{b}$ es equivalente al sistema

$$R\mathbf{x} = Q^\top \mathbf{b}$$

el cual es un sistema triangular superior que se resuelve por sustitución regresiva.

Bibliografía

1. Kendall E. Atkinson, *An Introduction to Numerical Analysis*, Second edition, John Wiley & Sons, 1989.
2. Richard L. Burden y J. Douglas Faires, *Análisis Numérico*, Séptima edición, Cengage Learning, México D.F., 2009.
3. https://mat.caminos.upm.es/wiki/Factorización_de_Doolittle.