

Capítulo 2

Métodos de resolución de sistemas lineales

2.1. Comentarios generales

El objetivo de este capítulo es examinar los aspectos numéricos que se presentan al resolver sistemas de ecuaciones lineales de la forma

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

Éste es un sistema de n ecuaciones en las n incógnitas x_1, \dots, x_n . Los elementos a_{ij} y b_i se supone que son números reales fijados. Utilizamos las matrices para representar este tipo de ecuaciones, de forma que el sistema anterior se puede expresar como

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

O de forma compacta, si denotamos las matrices anteriores por A , x y b podemos escribir la ecuación anterior como

$$Ax = b$$

el vector b , es el *vector de términos independientes*.

2.1.1. Sistemas de Ecuaciones

Los métodos de resolución de sistemas lineales se clasifican en métodos directos y en métodos iterativos. Los *métodos directos* permiten obtener la solución del sistema después de un número finito de pasos. Veremos que hay matrices con la propiedad de que es muy fácil la resolución directa de los sistemas correspondientes. Entre ellos cabe destacar el de las matrices triangulares.

El análisis del error en el proceso de resolución de sistemas lineales sirve para obtener acotaciones de los errores en la solución, producidos por los errores de los datos y por los de las operaciones.

Los *métodos iterativos* permiten obtener la solución del sistema a través de aproximaciones sucesivas. En la práctica, estos métodos se suelen utilizar cuando hay pocos elementos no nulos en la matriz.

Operaciones Elementales

Definición 2.1 Decimos que dos sistemas son **equivalentes** si tienen exactamente el mismo conjunto de soluciones.

Las operaciones elementales sobre los sistemas de ecuaciones, tienen como finalidad el transformar sistemas de ecuaciones en sistemas equivalentes que pueden resolverse más rápidamente. El conjunto de esas operaciones elementales es el siguiente

1. Intercambiar dos ecuaciones en el sistema, que equivale al intercambio de dos filas en la matriz del sistema
2. Multiplicar una ecuación por un número distinto de cero, que equivale a multiplicar una fila de la matriz \mathbf{A} por ese número.
3. Añadir a una ecuación un múltiplo de alguna otra ecuación. Es decir sustituimos una fila de la matriz \mathbf{A} por una combinación lineal de ella misma y de las demás.

2.2. Métodos directos

2.2.1. Sistemas fáciles de resolver

Comenzamos con la búsqueda de sistemas que sean fáciles de resolver. Por ejemplo si suponemos que la matriz \mathbf{A} es diagonal

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

la solución del sistema se reduce a n ecuaciones simples y la solución es

$$x = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

Si $a_{ii} = 0$, para algún índice i , pueden darse dos casos

$$\begin{aligned} b_i \neq 0 &\implies x_i \in \mathbb{R}, \text{ existen infinitas soluciones} \\ b_i = 0 &\implies \text{No existe ninguna solución} \end{aligned}$$

Otros sistemas sencillos de resolver son aquellos en los que A es una matriz triangular (superior o inferior). Suponiendo que la matriz A tiene una estructura triangular inferior, el sistema es de la forma

$$\begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

que es fácil de resolver mediante el algoritmo llamado de **sustitución progresiva**

$$\begin{aligned}
 x_1 &= b_1/a_{11} \\
 x_2 &= \frac{b_2 - a_{21}x_1}{a_{22}} \\
 x_3 &= \frac{b_3 - a_{31}x_1 - a_{32}x_2}{a_{33}} \\
 &\vdots \\
 x_k &= \frac{b_k - \sum_{j=1}^{k-1} a_{kj}x_j}{a_{kk}}
 \end{aligned}$$

Análogamente si A es triangular superior

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

se resuelve mediante

$$\begin{aligned}
 x_n &= b_n/a_{nn} \\
 x_{n-1} &= \frac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}} \\
 x_{n-2} &= \frac{b_{n-2} - a_{n-2,n}x_n - a_{n-2,n-1}x_{n-1}}{a_{n-2,n-2}} \\
 &\vdots \\
 x_k &= \frac{b_k - \sum_{j=k+1}^n a_{kj}x_j}{a_{kk}}
 \end{aligned}$$

llamado algoritmo de **sustitución regresiva**.

Algoritmo para la resolución de un sistema triangular

$$x_n = b_n/a_n$$

Para $i = n - 1, \dots, 1$

$$x_i = b_i$$

Para $j = i + 1, \dots, n$

$$x_i = x_i - a_{ij}x_j$$

$$x_i = x_i/a_{ii}$$

Coste computacional

k	Sumas	Productos	Divisiones
n	—	—	1
$n - 1$	1	1	1
\vdots	\vdots	\vdots	\vdots
1	$n - 1$	$n - 1$	1
Total	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$	n

Sumando todas las operaciones tenemos un total de

$$\text{TOTAL OPERACIONES} = \frac{n(n-1)}{2} + \frac{n(n-1)}{2} + n = n^2$$

2.2.2. Método de Gauss

Sea un sistema

$$Ax = b$$

con $|A| \neq 0$. El método de Gauss consiste en llevar el sistema a la forma triangular produciendo ceros por debajo de la diagonal principal, realizando siempre operaciones elementales.

Partiendo del sistema inicial, vamos a construir una sucesión de sistemas intermedios hasta llegar a un sistema triangular. Es decir si llamamos $A^{(1)} = A = (a_{ij}^{(1)})$ y $b^{(1)} = b$, el sistema inicial es $A^{(1)}x = b$. Realizando operaciones elementales convenientemente vamos obteniendo sistemas equivalentes ($A^{(k)} = b^{(k)}$), donde la matriz $A^{(k)} = (a_{ij}^{(k)})$, tiene nulos los elementos por debajo de la diagonal en las primeras $k - 1$ columnas

$$A^{(k)} = \begin{bmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \dots & a_{1k}^{(k)} & \dots & a_{1n}^{(k)} \\ 0 & a_{22}^{(k)} & \dots & a_{2k}^{(k)} & \dots & a_{2n}^{(k)} \\ \vdots & 0 & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{bmatrix} \quad k = 2, \dots, n$$

De esta forma obtenemos un sistema

$$A^{(n)}x = b^{(n)}$$

donde $A^{(n)}$ es una matriz triangular superior. Paralelamente a estas operaciones sobre la matriz A , se realizan las transformaciones convenientes sobre el vector b de forma que no se altere el sistema.

El paso k -ésimo se realiza cambiando sólo los valores de los elementos de las filas que van desde la $(k + 1)$ -ésima hasta la última y los elementos correspondientes del vector b de términos independientes según el siguiente esquema

$$A^{(k+1)} = M^{(k)}A^{(k)}$$

$$b^{(k+1)} = M^{(k)}b^{(k)}$$

donde

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

y

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

$M^{(k)}$ es la **matriz de Frobenius**, que tiene como objetivo sacar 0 debajo de la diagonal en la columna k . El elemento $a_{kk}^{(k)}$ se llama **pivote**. La nueva matriz $A^{(k+1)} = (a_{ij}^{(k+1)})$, tendrá los siguientes coeficientes

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} + m_{ik}a_{kj}^{(k)}$$

$$m_{ik} = -\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$$

$$b_i^{(k+1)} = b_i^{(k)} + m_{ik}b_k^{(k)}$$

En la segunda fase se resuelve el sistema triangular

$$A^{(n)}x = b^{(n)}$$

por el procedimiento descrito en la sección anterior.

Notar que en la etapa k -ésima del proceso de reducción tenemos que dividir por el elemento $a_{kk}^{(k)}$. Necesitamos hacer las siguientes observaciones:

1. Si el elemento pivote $a_{kk}^{(k)} = 0$, entonces se puede permutar esta fila por la primera fila cuyo elemento de la columna k no sea 0 ($a_{ik}^{(k)} \neq 0, i > k$), que existe siempre ya que $|A| \neq 0$, obtenemos así una matriz equivalente

$$\tilde{A}^{(k)} = P_{k,t_k}A^{(k)}$$

$$\tilde{b}^k = P_{k,t_k}b^{(k)}$$

$$t_k = \min \{t \geq k : a_{tk}^{(k)} \neq 0\}$$

donde P_{k,t_k} es la matriz de permutación, que intercambia las filas k y t_k , esta matriz es la identidad en la que se han intercambiado las filas k y t_k

2. Al estar el elemento $a_{kk}^{(k)}$ en el denominador puede crear errores de truncamiento grandes al tratar la ecuación por ordenador. Si $a_{kk}^{(k)}$ es pequeño el método de Gauss es muy inestable y los errores en los datos pueden aumentar considerablemente. Modificamos el algoritmo de Gauss para evitar esta inestabilidad. Tenemos dos opciones

- a) **Pivote parcial (por columnas)**: Se elige como nuevo elemento pivote el elemento de máximo valor absoluto entre los elementos $a_{ik}^{(k)}$ $i = k, \dots, n$

$$|a_{t_k k}^{(k)}| = \max \{|a_{tk}^{(k)}| : t \geq k\}$$

- b) **Pivote total o completo**: El nuevo elemento pivote pasa a ser el elemento de máximo valor absoluto entre los elementos $a_{ij}^{(k)}$ $i, j = k, \dots, n$

$$|a_{t_k s_k}^{(k)}| = \max \{|a_{ts}^{(k)}| : k \leq t \leq n; k \leq s \leq n\}$$

El problema es que se pueden intercambiar columnas y por tanto, las incógnitas y esto hay que tenerlo en cuenta.

Ejemplo 2.1 Resolver por el método de Gauss el siguiente sistema

$$\left. \begin{array}{cccccc} x_1 & + & x_2 & & + & 3x_4 & = & 4 \\ 2x_1 & + & x_2 & - & x_3 & + & x_4 & = & 1 \\ 3x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & -3 \\ -x_1 & + & 2x_2 & + & 3x_3 & - & x_4 & = & 4 \end{array} \right\}$$

Solución: Tomemos la matriz del sistema y apliquemos el método de Gauss

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 4 \\ 2 & 1 & -1 & 1 & 1 \\ 3 & -1 & -1 & 2 & -3 \\ -1 & 2 & 3 & -1 & 4 \end{array} \right) \rightsquigarrow$$

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 4 \\ 0 & -1 & -1 & -5 & -7 \\ 0 & -4 & -1 & -7 & -15 \\ 0 & 3 & 3 & 2 & 8 \end{array} \right) \rightsquigarrow$$

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 4 \\ 0 & -1 & -1 & -5 & -7 \\ 0 & 0 & 3 & 13 & 13 \\ 0 & 0 & 0 & -13 & -13 \end{array} \right) \rightsquigarrow$$

$$\left. \begin{array}{cccccc} x_1 & + & x_2 & & + & 3x_4 & = & 4 \\ & - & x_2 & - & x_3 & - & 5x_4 & = & -7 \\ & & & & 3x_3 & + & 13x_4 & = & 13 \\ & & & & & - & 13x_4 & = & -13 \end{array} \right\}$$

resolviendo ahora el sistema triangular obtenemos las soluciones:

$$x_4 = 1; x_3 = 0; x_2 = 2; x_1 = -1$$

Corrección de errores

A partir del sistema

$$\left. \begin{array}{cccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \dots & + & a_{nn}x_n & = & b_n \end{array} \right\}$$

obtenemos mediante el método de Gauss, un valor aproximado de la solución

$$(\tilde{x}_1, \dots, \tilde{x}_n)$$

al substituir esta solución aproximada en el sistema, obtendremos un valor aproximado para los coeficientes independientes

$$\left. \begin{array}{cccc} a_{11}\tilde{x}_1 & + & \dots & + & a_{1n}\tilde{x}_n & = & \tilde{b}_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}\tilde{x}_1 & + & \dots & + & a_{nn}\tilde{x}_n & = & \tilde{b}_n \end{array} \right\}$$

si definimos ahora $\Delta x_j = x_j - \tilde{x}_j \Rightarrow x_j = \tilde{x}_j + \Delta x_j$ y substituímos en el sistema obtenemos

$$\left. \begin{array}{cccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \dots & + & a_{nn}x_n & = & b_n \end{array} \right\} \rightsquigarrow$$

$$\left. \begin{array}{r} a_{11}(\tilde{x}_1 + \Delta x_1) + \dots + a_{1n}(\tilde{x}_n + \Delta x_n) = b_1 \\ \vdots \\ a_{n1}(\tilde{x}_1 + \Delta x_1) + \dots + a_{nn}(\tilde{x}_n + \Delta x_n) = b_n \end{array} \right\} \rightsquigarrow$$

$$\left. \begin{array}{r} a_{11}\tilde{x}_1 + \dots + a_{1n}\tilde{x}_n + a_{11}\Delta x_1 + \dots + a_{1n}\Delta x_n = b_1 \\ \vdots \\ a_{n1}\tilde{x}_1 + \dots + a_{nn}\tilde{x}_n + a_{n1}\Delta x_1 + \dots + a_{nn}\Delta x_n = b_n \end{array} \right\} \rightsquigarrow$$

$$\left. \begin{array}{r} a_{11}\Delta x_1 + \dots + a_{1n}\Delta x_n = b_1 - \tilde{b}_1 \\ \vdots \\ a_{n1}\Delta x_1 + \dots + a_{nn}\Delta x_n = b_n - \tilde{b}_n \end{array} \right\}$$

este sistema es así mismo un sistema de ecuaciones que se puede utilizar para calcular el incremento Δx_j .

Ejemplo 2.2 Sea el sistema de ecuaciones siguiente:

$$\begin{array}{rclcl} 3x_1 & -0,1x_2 & -0,2x_3 & = & 7,85 \\ 0,1x_1 & +7x_2 & -0,3x_3 & = & -19,3 \\ 0,3x_1 & -0,2x_2 & +10x_3 & = & 71,4 \end{array}$$

utilizando la eliminación gaussiana obtenemos la siguiente solución

$$\begin{array}{rcl} x_1 & = & 3,17 \\ x_2 & = & -2,51 \\ x_3 & = & 7,02 \end{array}$$

sustituyendo estos valores en las ecuaciones originales del sistema, obtenemos un valor aproximado de los términos independientes

$$\tilde{b} = (8,36, -19,4, 71,7)$$

para calcular una aproximación del error cometido construimos el siguiente sistema

$$\begin{array}{rclclcl} 3\Delta x_1 & -0,1\Delta x_2 & -0,2\Delta x_3 & = & (7,85 - 8,36) & = & -0,51 \\ 0,1\Delta x_1 & +7\Delta x_2 & -0,3\Delta x_3 & = & (-19,3 - (-19,4)) & = & 0,1 \\ 0,3\Delta x_1 & -0,2\Delta x_2 & +10\Delta x_3 & = & (71,4 - 71,7) & = & -0,3 \end{array}$$

Que si lo resolvemos (con 3 cifras decimales) obtenemos

$$\Delta x = (-0,171, 0,157 \times 10^{-1}, -0,246 \times 10^{-1})$$

y las soluciones serán

$$\begin{array}{rclcl} x_1 & = & 3,17 - 0,1710 & = & 3,00 \\ x_2 & = & -2,51 + 0,0157 & = & -2,49 \\ x_3 & = & 7,02 - 0,0246 & = & 7,00 \end{array}$$

2.2.3. Método de Gauss-Jordan

Consiste en transformar A en una matriz diagonal; en vez de eliminar x_k solamente de la fila i ; con $i > k$, se elimina de todas las filas salvo para $i = k$.

Ejemplo 2.3 Vamos a resolver mediante el método de Gauss-Jordan el sistema

$$\left. \begin{array}{r} x + y - z = 1 \\ 3x + 2y + z = 1 \\ 5x + 3y + 4z = 2 \end{array} \right\}$$

Operamos primero sobre las filas, para construir una matriz triangular superior. A continuación operamos sobre las columnas y construimos la matriz diagonal necesaria (también se pueden realizar las operaciones conjuntamente, haciendo ceros en filas y columnas).

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 3 & 2 & 1 & 1 \\ 5 & 3 & 4 & 2 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -1 & 4 & -2 \\ 0 & -2 & 9 & -3 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -1 & 4 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & -1 & 0 & -6 \\ 0 & 0 & 1 & 1 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & -1 & 0 & -6 \\ 0 & 0 & 1 & 1 \end{array} \right)$$

con solución

$$x = -4; y = -6; z = 1$$

Nota 1 El método de Gauss puede llevarse a cabo de modo estable sin intercambio de filas para

- Matrices estrictamente diagonal dominantes

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

- Matrices simétricas definidas positivas.

Definición 2.2 Una matriz $A \in M_{n \times n}(\mathbb{R})$, es **definida positiva**, si

$$\forall x \neq 0 \implies x^T \mathbf{A} x > 0$$

o equivalentemente que todos sus valores propios son positivos.

Algoritmo de Gauss

En el proceso de triangulación del método de Gauss descrito partimos de

$$\left. \begin{array}{l} A^{(1)}x = b^{(1)} \\ |A| \neq 0 \end{array} \right\}$$

y llegamos a la matriz $A^{(n)}$, que es triangular superior. Suponiendo que no hay intercambio de filas, el algoritmo sería el siguiente

Algoritmo del método de Gauss $a_{ij}^{(1)} = a_{ij}$

$b_i^{(1)} = b_i$
 Para $k = 1, \dots, n$
 Para $i = k + 1, \dots, n$
 $m_{ik} = -a_{ik}^{(k)} / a_{kk}^{(k)}$
 Para $j = k + 1, \dots, n$
 $a_{ij}^{(k+1)} = a_{ij}^{(k)} + m_{ik} a_{kj}^{(k)}$
 $b_i^{(k+1)} = b_i^{(k)} + m_{ik} b_k^{(k)}$

Coste computacional

k	Sumas	Productos	Divisiones
1	$(n-1)^2 + (n-1)$	$(n-1)^2 + (n-1)$	$n-1$
2	$(n-2)^2 + (n-2)$	$(n-2)^2 + (n-2)$	$n-2$
\vdots	\vdots	\vdots	\vdots
$n-1$	$1^2 + 1$	$1^2 + 1$	1
Total	$S_{n-1} + \frac{n(n-1)}{2}$	$S_{n-1} + \frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$

Siendo S_{n-1} la suma de los cuadrados de los $(n-1)$ primeros números naturales

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

Sumando todas las operaciones tenemos un total de

$$\text{TOTAL OPERACIONES} = \frac{4n^3 + 9n^2 - 7n}{6}$$

que cuando n es grande se aproxima a

$$\text{TOTAL OPERACIONES} \approx \frac{2}{3}n^3$$

Nota 2 Si no hay intercambio de filas, es decir

$$A^{(n)} = M^{(n-1)} \dots M^{(2)} M^{(1)} A^{(1)} \implies A^{(1)} = M^{-1} A^{(n)}$$

siendo

$$\left(M^{(k)}\right)^{-1} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \ddots & \dots & 0 & \dots & 0 \\ \vdots & 0 & 1 & \vdots & \dots & \vdots \\ 0 & \dots & -m_{k+1,k} & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -m_{n,k} & 0 & \dots & 1 \end{bmatrix}$$

Si $L = M^{-1}; U = A^{(n)} \implies A^{(1)} = LU$ con L una matriz triangular inferior y U una matriz triangular superior.

2.2.4. Factorización LU

Dado el sistema

$$Ax = b; \quad |A| \neq 0$$

si se produce la descomposición

$$A = LU$$

con L triangular inferior y U triangular superior, resolver el sistema inicial es equivalente a resolver los siguientes sistemas

$$\left. \begin{array}{l} Ly = b \\ Ux = y \end{array} \right\}$$

$$b = Ly = L(Ux) = LUx = Ax$$

Teorema 2.1 Sea $A \in M_{n \times n}(\mathbb{R})$; $|A_k| \neq 0$ $k = 1, \dots, n$. Siendo

$$A_k = (a_{ij})_{1 \leq i, j \leq k}$$

Entonces

$$A = LU$$

con L matriz triangular inferior con elementos diagonales $l_{ii} = 1, \forall i$ y U matriz triangular superior con elementos diagonales $u_{ii} \neq 0, \forall i$. Esta descomposición es única.

Cálculo directo de la factorización LU

Si

$$\left. \begin{array}{l} Ax = b \\ |A| \neq 0 \\ |A_k| \neq 0; k = 1, \dots, n \end{array} \right\} \Rightarrow \exists A = LU$$

donde

$$L = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \quad U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & u_{nn} \end{bmatrix}$$

con $l_{ij} = 0 \forall j > i$; $u_{ij} = 0 \forall j < i$.

Utilizando la definición de producto de dos matrices

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

donde hemos utilizado la propiedad de que hay componentes que son cero.

Se puede comprobar perfectamente cuales son los elementos de la primera fila de U

$$u_{1j} = a_{1j} \quad j = 1, \dots, n$$

Si suponemos que estamos en el paso k obtendremos

$$a_{kj} = l_{k1} u_{1j} + l_{k2} u_{2j} + \dots + l_{kk} u_{kj} \quad j \geq k$$

$$a_{ik} = l_{i1} u_{1k} + l_{i2} u_{2k} + \dots + l_{ik} u_{kk} \quad i \geq k$$

El método viene dado por las siguientes ecuaciones, para $k = 1, \dots, n$

$$u_{kj} = a_{kj} - \sum_{r=1}^{k-1} l_{kr} u_{rj} \quad (j = k, \dots, n)$$

$$l_{kk} = 1, \quad l_{ik} = \frac{1}{u_{kk}} \left(a_{ik} - \sum_{r=1}^{k-1} l_{ir} u_{rk} \right) \quad (i = k+1, \dots, n)$$

Una vez tenemos hecha la descomposición, se resuelven los sistemas triangulares que se obtienen. El primer sistema se resuelve como

$$b_k = l_{k1}y_1 + l_{k2}y_2 + \dots + l_{k,k-1}y_{k-1} + l_{kk}y_k \quad k = 1, \dots, n$$

Algoritmo para la descomposición LU

Para $k = 1, \dots, n$

$$\begin{aligned} &\text{Para } j = k, k+1, \dots, n \\ &\quad u_{kj} = a_{kj} - \sum_{p=1}^{k-1} l_{kp}u_{pj} \\ &\quad y_k = b_k - \sum_{p=1}^{k-1} l_{kp}y_p \\ &\text{Para } i = k+1, \dots, n \\ &\quad l_{ik} = \left(a_{ik} - \sum_{p=1}^{k-1} l_{ip}u_{pk} \right) / u_{kk} \end{aligned}$$

Este algoritmo realiza la descomposición y a la vez calcula $y = L^{-1}b$. Se calcula como

- Primera fila de U e y_1 .

$$u_{1j} = a_{1j} \quad (j = 1, \dots, n)$$

$$y_1 = b_1$$

- Primera columna de L

$$l_{11} = 1$$

$$l_{i1} = a_{i1}/u_{11} \quad (i = 2, \dots, n)$$

- Segunda fila de U e y_2

$$u_{2j} = a_{2j} - (l_{21}u_{1j}) \quad (j = 2, \dots, n)$$

$$y_2 = b_2 - l_{21}y_1$$

- Segunda columna de L

$$l_{i2} = (a_{i2} - l_{i1}u_{12}) / u_{22} \quad (i = 3, \dots, n)$$

⋮

El orden de cálculo es importante: $u_{1j} (j = 1, \dots, n)$; $l_{i1} (i = 2, \dots, n)$; $u_{2j} (j = 2, \dots, n)$; $l_{i2} (i = 3, \dots, n)$; \dots , $u_{n-1,n-1}$, $u_{n-1,n}$, $l_{n,n-1}$ y u_{nn} .

Este método se denomina de *Doolittle*. Existe otra descomposición LU , pero en la que es la matriz U la que tiene los elementos de la diagonal principal iguales a 1, dicho método se llama de *Crout* y utiliza las siguientes matrices

$$L = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix} \quad U = \begin{bmatrix} 1 & u_{12} & \dots & u_{1n} \\ 0 & 1 & \dots & u_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

donde ahora $l_{ij} \neq 0$; $i \geq j$. El método sería el siguiente, para $k = 1, \dots, n$

$$\begin{aligned} l_{ik} &= a_{ik} - \sum_{r=1}^{k-1} l_{ir}u_{rk} \quad (i = k, \dots, n) \\ u_{kk} &= 1; \quad u_{kj} = \frac{1}{l_{kk}} \left(a_{kj} - \sum_{r=1}^{k-1} l_{kr}u_{rj} \right) \quad (j = k+1, \dots, n) \end{aligned}$$

2.2.5. Método de Cholesky

Si $A \in M_n(\mathbb{R})$, entonces:

Teorema 2.2 A es simétrica y definida positiva $\iff \exists L$ matriz triangular inferior con $l_{kk} > 0$, tal que $A = LL^T$.

La construcción se hace de forma análoga a los métodos anteriores, partiendo de que A puede descomponerse en la forma anterior

$$A = LL^T$$

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

haciendo el producto e identificado, podemos obtener las ecuaciones que deben cumplir los coeficientes de L .

Algoritmo del método de Cholesky

Para $k = 1, \dots, n$

$$l_{kk} = \sqrt{a_{kk} - \sum_{p=1}^{k-1} l_{kp}^2}$$

Para $j = k + 1, \dots, n$

$$l_{jk} = \left(a_{kj} - \sum_{p=1}^{k-1} l_{jp} l_{kp} \right) / l_{kk}$$

2.3. Métodos iterativos

En esta sección construiremos a partir del sistema de ecuaciones inicial $Ax = b$, otro sistema lineal equivalente de la forma

$$x = Mx + N$$

que resolveremos de forma iterativa. Partiendo de una aproximación inicial $x^{(0)}$, generamos una sucesión de vectores $\{x^{(k)}\}_k$ que converja a x , la solución exacta del sistema de partida. Consideraremos sólo métodos que definan la sucesión de la forma

$$x^{(k+1)} = Mx^{(k)} + N$$

$$M \in M_{n \times n}(\mathbb{R})$$

$$N \in M_{n \times 1}(\mathbb{R})$$

siendo M es la matriz **iteración** del método.

Definición 2.3 Un método construido según el esquema anterior se dice **convergente** al sistema $Ax = b$, si $\forall x^{(0)} \in \mathbb{R}^n$, se verifica

$$\lim_{k \rightarrow \infty} x^{(k)} = x = A^{-1}b$$

Si el método es convergente, entonces al tomar límites obtendremos

$$x = Mx + N$$

y como $x = A^{-1}b \Rightarrow$

$$A^{-1}b = MA^{-1}b + N \Rightarrow$$

$$N = A^{-1}b - MA^{-1}b = (I - M)A^{-1}b$$

Definición 2.4 Un método es **consistente** con el sistema $Ax = b$ si se verifica

$$\begin{aligned} x &= Mx + N \\ & \quad y \\ N &= (I - M)A^{-1}b \end{aligned}$$

Definición 2.5 Sea $A \in M_{n \times n}(\mathbb{R})$, definimos su radio espectral $\rho(A)$ como

$$\rho(A) = \max_{i=1, \dots, n} \{|\lambda_i| : \lambda_i \text{ es valor propio de } A\}$$

Teorema 2.3 El método es convergente respecto a $Ax = b$, independientemente de la elección del punto inicial $x^0 \iff$

1. Es consistente
2. El radio espectral de la matriz de iteración es menor que 1:

$$\rho(M) < 1$$

Construcción de métodos iterativos

Sea $Ax = b$ con A invertible; entonces $x = A^{-1}b$. Supongamos que ponemos A de la siguiente forma

$$A = R - S$$

con R invertible (la descomposición existe siempre, por ejemplo con $R = A, S = 0$). Tenemos por tanto

$$Ax = b \Rightarrow (R - S)x = Rx - Sx = b \Rightarrow Rx = Sx + b$$

multiplicando por R^{-1}

$$R^{-1}Rx = R^{-1}Sx + R^{-1}b$$

$$x = R^{-1}Sx + R^{-1}b$$

que sugiere por tanto el método

$$x^{(k+1)} = R^{-1}Sx^{(k)} + R^{-1}b$$

que es del tipo general con

$$M = R^{-1}S$$

y

$$N = R^{-1}b$$

por tanto se cumple el siguiente teorema.

Teorema 2.4 El método

$$x^{(k+1)} = R^{-1}Sx^{(k)} + R^{-1}b$$

con $A = R - S$ y R invertible es consistente y será convergente si $(R^{-1}S) < 1$.

Recordemos que un método es consistente si

$$\begin{aligned} x &= Mx + N \\ & \quad y \\ N &= (I - M)A^{-1}b \end{aligned}$$

En este caso $M = R^{-1}S$ y $N = R^{-1}b$, podemos comprobar por tanto que

$$x = Mx + N \Leftrightarrow x = R^{-1}Sx + R^{-1}b$$

y multiplicando por R

$$Rx = Sx + b \Leftrightarrow Rx - Sx = b \Leftrightarrow (R - S)x = b$$

pero $A = R - S$, por tanto

$$Ax = b.$$

Además

$$(I - M)A^{-1}b = (I - R^{-1}S)A^{-1}b$$

como $S = R - A$

$$\begin{aligned} (I - R^{-1}S)A^{-1}b &= (I - R^{-1}(R - A))A^{-1}b = (I - R^{-1}R + R^{-1}A)A^{-1}b \\ &= (I - I + R^{-1}A)A^{-1}b = R^{-1}AA^{-1}b = R^{-1}b = N \end{aligned}$$

por tanto

$$(I - M)A^{-1}b = N$$

y el método es consistente. Para que sea convergente $\rho(M) < 1$, es decir, $\rho(R^{-1}S) < 1$.

Podemos por tanto deducir el siguiente resultado

Teorema 2.5 *Todo método*

$$x^{(k+1)} = Mx^{(k)} + N$$

convergente es de la forma

$$x^{(k+1)} = R^{-1}Sx^{(k)} + R^{-1}b$$

con $A = R - S$ y R invertible.

Vamos a deducir diversos métodos a partir de la ecuación anterior. Para ello descomponemos la matriz A como sigue

$$A = D - L - U$$

con

$$D = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & a_{nn} \end{bmatrix}$$

$$L = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ -a_{21} & 0 & 0 & \dots & 0 \\ -a_{31} & -a_{32} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -a_{n1} & -a_{n2} & \dots & -a_{n,n-1} & 0 \end{bmatrix}$$

$$U = \begin{bmatrix} 0 & -a_{12} & -a_{13} & \dots & -a_{1n} \\ 0 & 0 & -a_{23} & \dots & -a_{2n} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -a_{n-1,n} \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

2.3.1. Método de Jacobi

Consiste en elegir

$$\begin{aligned} R &= D \\ S &= L + U \end{aligned}$$

tendremos por tanto la ley de recurrencia siguiente

$$\begin{aligned} Dx^{(k+1)} &= (L + U)x^{(k)} + b \\ x^{(k+1)} &= D^{-1}(L + U)x^{(k)} + D^{-1}b \end{aligned}$$

que en componentes toma la forma siguiente

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

expresión general del método de Jacobi; para obtener $x_i^{(k+1)}$, hacen falta todos los valores de $x^{(k)}$, que se utilizan simultáneamente, por eso a este método también se le suele denominar *método de aproximaciones simultáneas*.

Notar que teniendo en cuenta que

$$A = D - L - U \Rightarrow L + U = D - A$$

el método también puede ponerse como

$$Dx^{(k+1)} = (D - A)x^{(k)} + b$$

o multiplicando por la inversa D^{-1} , como

$$x^{(k+1)} = (I - D^{-1}A)x^{(k)} + D^{-1}b$$

siendo la matriz de iteración

$$M = I - D^{-1}A.$$

2.3.2. Método de Gauss-Seidel

En este caso elegimos

$$\begin{aligned} R &= D - L \\ S &= U \end{aligned}$$

tendremos por tanto la ley de recurrencia siguiente

$$\begin{aligned} (D - L)x^{(k+1)} &= Ux^{(k)} + b \\ x^{(k+1)} &= (D - L)^{-1}Ux^{(k)} + (D - L)^{-1}b \end{aligned}$$

que es un sistema triangular que se resuelve de arriba hacia abajo, para calcular $x_i^{(k+1)}$ se utilizan los valores previamente calculados de $x_j^{(k+1)}$ para $j = 1, \dots, i - 1$, y los valores de $x_j^{(k)}$. Este método también se denomina *método de aproximaciones sucesivas* y en general converge más rápidamente que el anterior.

Notar que M , la matriz de iteración, es en este caso

$$M = (D - L)^{-1}U$$

El sistema expresado en componentes toma la forma siguiente

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

Teorema 2.6 *Si la matriz A del sistema de ecuaciones $Ax = b$ es estrictamente diagonal dominante entonces los métodos de Jacobi y Gauss-Seidel convergen.*