

Clase 18 - Resolución de sistemas lineales (3)

El problema

Dada $A \in \mathbb{R}^{n \times n}$ y $b \in \mathbb{R}^n$ hallar x_* solución de $Ax = b$.

Hasta ahora vimos dos métodos numéricos para resolver este problema: eliminación gaussiana y factorización LU. Ambos métodos son considerados métodos directos y se sabe que luego de un número finito de pasos se obtiene una solución, salvo errores de redondeo. Por otro lado, existe otra familia de métodos para resolver este problema llamados iterativos o indirectos.

Métodos iterativos

Los métodos iterativos para sistemas lineales generan una sucesión de vectores $\{x^{(k)}\}$, a partir de un vector inicial $x^{(0)}$, que convergen a la solución de $Ax = b$, bajo adecuadas hipótesis. La convergencia significa que esta sucesión se detiene cuando se alcanza una precisión determinada luego de un cierto número de iteraciones. El éxito computacional de estos métodos requiere que procedimiento sea simple y de bajo costo computacional. En general estos métodos son adecuados para matrices grandes y que tienen muchos coeficientes iguales a cero (matrices ralas).

Estudiaremos dos métodos iterativos muy conocidos: método de Jacobi y método de Gauss-Seidel. Presentaremos brevemente estos métodos y algunos resultados simples de convergencia. Para fijar ideas veremos inicialmente un ejemplo muy sencillo.

Ejemplo: consideremos el sistema lineal

$$\begin{bmatrix} 7 & -6 \\ -8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix} \longleftrightarrow \begin{cases} 7x_1 - 6x_2 = 3 \\ -8x_1 + 9x_2 = -4 \end{cases}, \quad (1)$$

cuya solución es $x_* = (0.2, -0.2666\dots)$. A partir del sistema lineal (1), podemos despejar las variables:

$$\begin{cases} x_1 = \frac{6}{7}x_2 + \frac{3}{7} \\ x_2 = \frac{8}{9}x_1 - \frac{4}{9} \end{cases} \quad (2)$$

Dada una aproximación $x^{(k-1)} = (x_1^{(k-1)}, x_2^{(k-1)})$, y de (2) se puede generar el siguiente método iterativo:

$$\begin{cases} x_1^{(k)} = \frac{6}{7}x_2^{(k-1)} + \frac{3}{7} \\ x_2^{(k)} = \frac{8}{9}x_1^{(k-1)} - \frac{4}{9} \end{cases},$$

el cual es conocido como método de Jacobi.

También se puede generar otro método iterativo de la siguiente manera:

$$\begin{cases} x_1^{(k)} = \frac{6}{7}x_2^{(k-1)} + \frac{3}{7} \\ x_2^{(k)} = \frac{8}{9}x_1^{(k)} - \frac{4}{9} \end{cases},$$

el cual es conocido como método de Gauss-Seidel.

La tabla siguiente muestra algunas iteraciones de ambos métodos, comenzando en ambos casos con $x^{(0)} = (x_1^{(0)}, x_2^{(0)}) = (0, 0)$. Recordemos que la solución es: $x_* = (0.2, -0.2666\dots)$.

It.	M. de Jacobi		M. de Gauss-Seidel	
k	$x_1^{(k)}$	$x_2^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$
0	0	0	0	0
10	0.14865	-0.19820	0.21978	-0.24909
20	0.18682	-0.24909	0.20130	-0.26531
30	0.19662	-0.26215	0.20009	-0.26659
40	0.19913	-0.26551	0.20001	-0.26666
50	0.19978	-0.26637	0.20000	-0.26667

Los métodos de Jacobi y Gauss-Seidel son conocidos también como métodos de separación (splitting) y parten de una misma idea básica, que consiste en escribir la matriz como $A = M - N$, donde M es una matriz no singular. Así tenemos que

$$Ax = b$$

$$(M - N)x = b$$

$$Mx = Nx + b \quad (3)$$

$$x = M^{-1}(Nx + b) \quad (4)$$

$$x = (M^{-1}N)x + M^{-1}b. \quad (5)$$

Dado $x^{(0)} = (x_1^{(0)}, x_2^{(0)})$, la última ecuación sugiere definir un método iterativo de la forma:

$$x^{(k+1)} = (M^{-1}N)x^{(k)} + M^{-1}b, \quad \text{para } k \geq 0. \quad (6)$$

A continuación veremos algunos resultados de convergencia de estos métodos iterativos.

Teorema 1. Sea $b \in \mathbb{R}^n$ y $A = M - N \in \mathbb{R}^{n \times n}$ donde A y M son matrices no singulares. Si $\|M^{-1}N\| < 1$ para alguna norma matricial inducida entonces la sucesión generada por la ecuación (6) converge a la solución de $Ax = b$ para cualquier vector inicial $x^{(0)}$.

Demostración. Restando la ecuación (6) de (5), donde en (5) ponemos la solución x^* se obtiene:

$$x^{(k+1)} - x^* = (M^{-1}N)(x^{(k)} - x^*), \quad \text{para } k \geq 0. \quad (7)$$

Ahora, utilizando alguna norma matricial inducida, se obtiene:

$$\|x^{(k+1)} - x^*\| \leq \|(M^{-1}N)\| \|x^{(k)} - x^*\|, \quad \text{para } k \geq 0. \quad (8)$$

Luego, repitiendo este último paso se tiene:

$$\|x^{(k+1)} - x^*\| \leq \|(M^{-1}N)\|^{k+1} \|x^{(0)} - x^*\|, \quad \text{para } k \geq 0. \quad (9)$$

Usando que $\|(M^{-1}N)\| < 1$, se tiene que

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x^*\| = 0,$$

para cualquier vector inicial $x^{(0)}$. □

La principal diferencial entre el método de Jacobi y el de Gauss-Seidel consiste en cómo es elegida la matriz M . Por simplicidad vamos a descomponer la matriz A de la siguiente forma: $A = L + D + U$, donde L es la parte triangular inferior de A (sin la diagonal), D es la diagonal de A y U es la parte triangular superior de A (sin la diagonal). Es importante no confundir con las matrices L y U de la factorización LU.

Método de Jacobi

En el método de Jacobi se toma $M = D$, y por lo tanto,

$$N = M - A = D - (L + D + U) = -(L + U),$$

esto es,

$$M = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \quad \text{y} \quad N = - \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix}.$$

Si miramos la i -ésima componente en (3) tenemos que

$$\begin{aligned} [Dx^{(k+1)}]_i &= [b - (L + U)x^{(k)}]_i \\ a_{ii}x_i^{(k+1)} &= b_i - [a_{i1} \dots a_{i,i-1} 0 a_{i,i+1} \dots a_{in}] \begin{bmatrix} x_1^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix} \\ a_{ii}x_i^{(k+1)} &= b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \\ x_i^{(k+1)} &= \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)}), \end{aligned}$$

para $i = 1, \dots, n$.

Ahora veremos un pseudocódigo del método de Jacobi. Notar que en el algoritmo usaremos x para representar el vector inicial y sobreescribiremos las sucesivas aproximaciones en el mismo vector. La variable $MaxIt$ representa el número máximo de iteraciones permitidas y $Xtol$ la tolerancia para la norma de la diferencia de dos aproximaciones vectoriales sucesivas.

Algoritmo: método de Jacobi.

input $n, A, b, x, MaxIt, Xtol$

for $k = 1, \dots, MaxIt$ **do**

for $i = 1, \dots, n$ **do**

$$u_i \leftarrow \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j)$$

end for (i)

if $\|u - x\| < Xtol$ **STOP, output:** u es la solución, k iteración.

for $i = 1, \dots, n$ **do**

$x_i \leftarrow u_i$

end for (i)

end for (k)

output x

end

Antes de enunciar un resultado de convergencia, daremos una definición.

Definición 1. Una matriz A , de orden $n \times n$, es **diagonalmente dominante** si

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

Teorema 2. Si A es diagonalmente dominante, entonces la sucesión generada por el método de Jacobi converge a la solución de $Ax = b$, para cualquier vector inicial $x^{(0)} \in \mathbb{R}^n$.

Demostración. En el método de Jacobi, la matriz M es la diagonal de A y debe ser inversible para que el método esté bien definido, por lo que $a_{ii} \neq 0, i = 1, \dots, n$. La matriz de iteración está dada por

$$\begin{aligned} M^{-1}N &= - \begin{bmatrix} 1/a_{11} & 0 & \dots & 0 \\ 0 & 1/a_{22} & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1/a_{nn} \end{bmatrix} \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{21} & 0 & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix} \\ &= - \begin{bmatrix} 0 & a_{12}/a_{11} & \dots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0 & \dots & a_{2n}/a_{22} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & \dots & 0 \end{bmatrix} \end{aligned} \quad (10)$$

Luego,

$$\|M^{-1}N\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n \frac{|a_{ij}|}{|a_{ii}|} < 1,$$

pues A es diagonalmente dominante. Finalmente, la convergencia es una consecuencia directa del teorema anterior.

□

Método de Gauss-Seidel

En el método de Gauss-Seidel se toma $M = L + D$, y por lo tanto,

$$N = M - A = L + D - (L + D + U) = -U,$$

esto es,

$$M = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \quad \text{y} \quad N = - \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Luego por (3) tenemos que

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

Así, analizando la i -ésima componente, obtenemos

$$\begin{aligned}[Dx^{(k+1)}]_i &= [b - Lx^{(k+1)} - Ux^{(k)}]_i \\ a_{ii}x_i^{(k+1)} &= b_i - [a_{i1} \dots a_{i,i-1} 0 \dots 0] \begin{bmatrix} x_1^{(k+1)} \\ \vdots \\ x_n^{(k+1)} \end{bmatrix} - [0 \dots 0 a_{i,i+1} \dots a_{in}] \begin{bmatrix} x_1^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix} \\ a_{ii}x_i^{(k+1)} &= b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \\ x_i^{(k+1)} &= \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)}),\end{aligned}$$

para $i = 1, \dots, n$.

Ahora veremos un pseudocódigo del método de Gauss-Seidel. Notar que en el algoritmo usaremos x para representar el vector inicial y sobreescibiremos las sucesivas aproximaciones en el mismo vector. La variable *MaxIt* representa el número máximo de iteraciones permitidas y *Xtol* la tolerancia para la norma de la diferencia de dos aproximaciones vectoriales sucesivas.

Algoritmo: método de Gauss-Seidel.

input $n, A, b, x, MaxIt, Xtol$

for $k = 1, \dots, MaxIt$ **do**

for $i = 1, \dots, n$ **do**

$u_i \leftarrow 0$

end for (i)

for $i = 1, \dots, n$ **do**

$$u_i \leftarrow \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}u_j - \sum_{j=i+1}^n a_{ij}x_j)$$

end for (i)

if $\|u - x\| < Xtol$ **STOP, output:** u es la solución, k iteración.

for $i = 1, \dots, n$ **do**

$x_i \leftarrow u_i$

end for (i)

end for (k)

output x

end

El resultado de convergencia que enunciamos para este método es similar al del método de Jacobi, aunque su demostración no es tan directa.

Teorema 3. Si A es diagonalmente dominante, entonces la sucesión generada por el método de Gauss-Seidel converge a la solución de $Ax = b$, para cualquier vector inicial $x^{(0)} \in \mathbb{R}^n$.

Observación: existen resultados que muestran que si ambos métodos convergen, el método de Gauss-Seidel lo hace más rápido que el método de Jacobi. Esto es razonable, pues al calcular la i -ésima componente, en el método de Gauss-Seidel, se utiliza información recientemente calculada en la misma iteración. Por otro lado, si ambos métodos divergen, el método de Gauss-Seidel lo hace más rápido que el método de Jacobi. Por esta razón el método era más atractivo computacionalmente que el de Jacobi. Sin embargo, el método de Jacobi ha vuelto a cobrar popularidad en las últimas décadas, pues es paralelizable mientras que el método de Gauss-Seidel no lo es.

Por último vamos a enunciar un resultado auxiliar y un teorema que, bajo ciertas hipótesis, asegura la convergencia de los métodos como Jacobi o Gauss-Seidel.

Teorema 4. Para cada matriz $A \in \mathbb{R}^{n \times n}$, se cumple que $\rho(A)$ es igual al $\inf\{\|A\|\}$ sobre todas las normas matriciales inducidas.

Teorema 5. Una condición necesaria y suficiente para que la sucesión generada por el método iterativo $x^{(k+1)} = (M^{-1}N)x^{(k)} + M^{-1}b$, para $k \geq 0$, converja a la única solución de $Ax = b$ para todo $x^{(0)}$ inicial, es que $\rho(M^{-1}N) < 1$.