

**CAPITULO XVII. TECNICAS ITERATIVAS PARA RESOLVER
SISTEMAS LINEALES**

1. INTRODUCCION Y METODO

Una técnica iterativa para resolver un sistema lineal $A \mathbf{x} = \mathbf{b}$ de $n \times n$ empieza con una aproximación inicial $\mathbf{x}^{(0)}$ a la solución \mathbf{x} , y genera una sucesión de vectores $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ que converge a \mathbf{x} . La mayoría de estas técnicas iterativas involucran un proceso que convierte el sistema $A \mathbf{x} = \mathbf{b}$ en un sistema equivalente de la forma $\mathbf{x} = T \mathbf{x} + \mathbf{c}$ para alguna matriz T de $n \times n$ y un vector \mathbf{c} . Ya seleccionado el vector inicial $\mathbf{x}^{(0)}$ la sucesión de vectores de solución aproximada se genera calculando

$$\mathbf{x}^{(k)} = T \mathbf{x}^{(k-1)} + \mathbf{c} \tag{XVII.1}$$

para cada $k = 1, 2, 3, \dots$. Este tipo de procedimiento nos recuerda a la iteración del punto fijo estudiada en la tercera parte.

Las técnicas iterativas se emplean raras veces para resolver sistemas lineales de dimensión pequeña ya que el tiempo requerido para lograr una precisión suficiente excede al de las técnicas directas como el método de eliminación Gaussiana. Sin embargo, para sistemas grandes con un gran porcentaje de ceros, estas técnicas son eficientes en términos de almacenamiento en la computadora y del tiempo requerido. Los sistemas de este tipo surgen frecuentemente en la solución numérica de problemas de valores en la frontera y de ecuaciones diferenciales parciales.

Ejemplo 1.

El sistema lineal $A \mathbf{x} = \mathbf{b}$ dado por

$$\begin{aligned} E_1 : & 10 x_1 - x_2 + 2 x_3 = 6, \\ E_2 : & -x_1 + 11 x_2 - x_3 + 3 x_4 = 25, \\ E_3 : & 2 x_1 - x_2 + 10 x_3 - x_4 = -11, \\ E_4 : & 3 x_2 - x_3 + 8 x_4 = 15, \end{aligned}$$

tiene por solución a $\mathbf{x} = (1, 2, -1, 1)^t$. Para convertir $A \mathbf{x} = \mathbf{b}$ a la forma $\mathbf{x} = T \mathbf{x} + \mathbf{c}$, resolvemos la ecuación E_i para cada $i = 1, 2, 3, 4$, obteniendo:

$$\begin{aligned} x_1 &= \frac{1}{10}x_2 - \frac{1}{5}x_3 + \frac{3}{5}, \\ x_2 &= \frac{1}{11}x_1 + \frac{1}{11}x_3 - \frac{3}{11}x_4 + \frac{25}{11}, \\ x_3 &= -\frac{1}{5}x_1 + \frac{1}{10}x_2 + \frac{1}{10}x_4 - \frac{11}{10}, \\ x_4 &= -\frac{3}{8}x_2 + \frac{1}{8}x_3 + \frac{15}{8}. \end{aligned}$$

En este ejemplo,

$$T = \begin{pmatrix} 0 & \frac{1}{10} & -\frac{1}{5} & 0 \\ \frac{1}{11} & 0 & \frac{1}{11} & -\frac{3}{11} \\ -\frac{1}{5} & \frac{1}{10} & 0 & \frac{1}{10} \\ 0 & -\frac{3}{8} & \frac{1}{8} & 0 \end{pmatrix} \quad \text{y} \quad \mathbf{c} = \begin{pmatrix} \frac{3}{5} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{pmatrix}.$$

Como una aproximación inicial tomemos a $\mathbf{x}^{(0)} = (0, 0, 0, 0)^t$ y generemos $\mathbf{x}^{(1)}$ mediante:

$$\begin{aligned} x_1^{(1)} &= \frac{1}{10}x_2^{(0)} - \frac{1}{5}x_3^{(0)} + \frac{3}{5} = 0.6000, \\ x_2^{(1)} &= \frac{1}{11}x_1^{(0)} + \frac{1}{11}x_3^{(0)} - \frac{3}{11}x_4^{(0)} + \frac{25}{11} = 2.2727, \\ x_3^{(1)} &= -\frac{1}{5}x_1^{(0)} + \frac{1}{10}x_2^{(0)} + \frac{1}{10}x_4^{(0)} - \frac{11}{10} = -1.1000, \\ x_4^{(1)} &= -\frac{3}{8}x_2^{(0)} + \frac{1}{8}x_3^{(0)} + \frac{15}{8} = 1.8750. \end{aligned}$$

Las iteraciones adicionales $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_4^{(k)})^t$, se generan de manera similar y se presentan en la tabla siguiente.

Tabla 1

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
0	0.0000	0.0000	0.0000	0.0000
1	0.6000	2.2727	-1.1000	1.8750
2	1.0473	1.7159	-0.80523	0.88524
3	0.93264	2.0533	-1.0493	1.1309
4	1.0152	1.9537	-0.96811	0.97385
5	0.98899	2.0114	-1.0103	1.0213
6	1.0032	1.9923	-0.99453	0.99444
7	0.99814	2.0023	-1.0020	1.0036
8	1.0006	1.9987	-0.99904	0.99889
9	0.99968	2.0004	-1.0004	1.0006
10	1.0001	1.9998	-0.99984	0.99980

La decisión de parar después de diez iteraciones está basada en el hecho de que

$$\frac{\|\mathbf{x}^{(10)} - \mathbf{x}^{(9)}\|_\infty}{\|\mathbf{x}^{(10)}\|_\infty} = \frac{8.0 \times 10^{-4}}{1.9998} < 10^{-3}.$$

En realidad, $\|\mathbf{x}^{(10)} - \mathbf{x}\|_\infty = 0.0002$.

El método del ejemplo 1 se llama **método iterativo de Jacobi**. Este consiste en resolver la i -ésima ecuación de $A \mathbf{x} = \mathbf{b}$ para x_i para obtener, siempre y cuando $a_{ii} \neq 0$, que

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

y generar cada $x_i^{(k)}$ de las componentes de $\mathbf{x}^{(k-1)}$ para $k \geq 1$ con

$$x_i^{(k)} = \frac{1}{a_{ii}} \left[\sum_{\substack{j=1 \\ j \neq i}}^n (-a_{ij} x_j^{(k-1)}) + b_i \right] \quad \text{para } i = 1, 2, \dots, n. \quad (XVII.3)$$

El método puede escribirse en la forma $\mathbf{x}^{(k)} = T \mathbf{x}^{(k-1)} + \mathbf{c}$ dividiendo a A en su parte diagonal y no-diagonal. Para ver esto, sean D la matriz diagonal cuya diagonal es la

misma que la diagonal de A , $-L$ la parte triangular estrictamente inferior de A , y $-U$ la parte triangular estrictamente superior de A . Con esta notación, se separa en

$$\begin{aligned}
 A &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} + \\
 &- \begin{pmatrix} 0 & 0 & \dots & 0 \\ -a_{21} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ -a_{n1} & \dots & -a_{n,n-1} & 0 \end{pmatrix} - \begin{pmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & -a_{n-1,n} \\ 0 & 0 & \dots & 0 \end{pmatrix} = \\
 &= D - L - U .
 \end{aligned}$$

La ecuación $A \mathbf{x} = \mathbf{b}$ ó $(D - L - U) \mathbf{x} = \mathbf{b}$ se transforma entonces en $D \mathbf{x} = (L + U) \mathbf{x} + \mathbf{b}$, y finalmente

$$\mathbf{x} = D^{-1} (L + U) \mathbf{x} + D^{-1} \mathbf{b} . \tag{XVII.4}$$

Esto da lugar a la forma matricial de la técnica iterativa de Jacobi:

$$\mathbf{x}^{(k)} = D^{-1} (L + U) \mathbf{x}^{(k-1)} + D^{-1} \mathbf{b} , \quad k = 1, 2, \dots \tag{XVII.5}$$

En la práctica, la ecuación (XVII.3) es la que se usa para los cálculos, reservando a la ecuación (XVII.5) para propósitos teóricos.

2. LOS ALGORITMOS DE JACOBI Y DE GAUSS-SEIDEL

Para resumir el método iterativo de Jacobi, presentamos el siguiente algoritmo:

Algoritmo iterativo de Jacobi.

=====

Para resolver el sistema lineal $A \mathbf{x} = \mathbf{b}$ con una aproximación inicial dada $\mathbf{x}^{(0)}$.

Entrada: número de incógnitas y de ecuaciones n ; las componentes de la matriz $A = (a_{ij})$ donde $1 \leq i, j \leq n$; las componentes b_i , con $1 \leq i \leq n$, del término no homogéneo \mathbf{b} ; las componentes XO_i , con $1 \leq i \leq n$, de la aproximación inicial $\mathbf{XO} = \mathbf{x}^{(0)}$; la tolerancia TOL; el número máximo de iteraciones N_0 .

Salida: solución aproximada x_1, x_2, \dots, x_n ó mensaje de que el número de iteraciones fue excedido.

Paso 1: Tomar $k = 1$.

Paso 2: Mientras que $k \leq N_0$ seguir los pasos 3-6.

Paso 3: Para $i = 1, 2, \dots, n$ tomar

$$x_i = \frac{1}{a_{ii}} \left[- \sum_{\substack{j=1 \\ j \neq i}}^n (a_{ij} XO_j) + b_i \right] .$$

Paso 4: Si $\|\mathbf{x} - \mathbf{XO}\| < TOL$ entonces SALIDA (x_1, x_2, \dots, x_n) ; (procedimiento completado satisfactoriamente) PARAR.

Paso 5: Tomar $k = k + 1$.

Paso 6: Para $i = 1, 2, \dots, n$ tomar $XO_i = x_i$.

Paso 7: SALIDA (*número máximo de iteraciones excedido*);
(*procedimiento completado sin éxito*) PARAR.

El paso 3 del algoritmo requiere que $a_{ii} \neq 0$ para cada $i = 1, 2, \dots, n$. Si éste no es el caso, se puede realizar un reordenamiento de las ecuaciones para que ningún $a_{ii} = 0$, a menos que el sistema sea singular. Se sugiere que las ecuaciones sean arregladas de tal manera que a_{ii} sea lo más grande posible para acelerar la convergencia.

En el paso 4, el criterio de paro ha sido $\|\mathbf{x} - \mathbf{XO}\| < TOL$; otro criterio de paro es iterar hasta que

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|}{\|\mathbf{x}^{(k)}\|}$$

sea menor que alguna tolerancia predeterminada $\varepsilon > 0$. Para este propósito, se puede usar cualquier norma conveniente; la que más se usa es la norma l_∞ .

Un análisis de la ecuación (XVII.3) sugiere una posible mejora en el algoritmo iterativo de Jacobi. Para calcular $x_i^{(k)}$, se usan las componentes de $\mathbf{x}^{(k-1)}$. Como para $i > 1$, $x_1^{(k)}$, $x_2^{(k)}$, \dots , $x_{i-1}^{(k)}$ ya han sido calculadas y supuestamente son mejores aproximaciones a la solución real x_1, x_2, \dots, x_{i-1} que $x_1^{(k)}, x_2^{(k)}, \dots, x_{i-1}^{(k)}$, parece razonable calcular $x_i^{(k)}$ usando los valores calculados más recientemente; es decir,

$$x_i^{(k)} = \frac{1}{a_{ii}} \left[- \sum_{j=1}^{i-1} (a_{ij} x_j^{(k)}) - \sum_{j=i+1}^n (a_{ij} x_j^{(k-1)}) + b_i \right], \quad (XVII.6)$$

para cada $i = 1, 2, \dots, n$ en vez de la ecuación (XVII.3).

Ejemplo 2.

El sistema lineal $A \mathbf{x} = \mathbf{b}$ dado por

$$\begin{aligned} E_1 : & 10 x_1 - x_2 + 2 x_3 = 6, \\ E_2 : & -x_1 + 11 x_2 - x_3 + 3 x_4 = 25, \\ E_3 : & 2 x_1 - x_2 + 10 x_3 - x_4 = -11, \\ E_4 : & 3 x_2 - x_3 + 8 x_4 = 15, \end{aligned}$$

fue resuelto en el ejemplo 1 con el método iterativo de Jacobi. Incorporando la ecuación (XVII.6) en el algoritmo iterativo de Jacobi, se obtienen las ecuaciones que se usarán para cada $k = 1, 2, \dots$:

$$\begin{aligned} x_1^{(k)} &= \frac{1}{10} x_2^{(k-1)} - \frac{1}{5} x_3^{(k-1)} + \frac{3}{5}, \\ x_2^{(k)} &= \frac{1}{11} x_1^{(k)} + \frac{1}{11} x_3^{(k-1)} - \frac{3}{11} x_4^{(k-1)} + \frac{25}{11}, \\ x_3^{(k)} &= -\frac{1}{5} x_1^{(k)} + \frac{1}{10} x_2^{(k)} + \frac{1}{10} x_4^{(k-1)} - \frac{11}{10}, \\ x_4^{(k)} &= -\frac{3}{8} x_2^{(k)} + \frac{1}{8} x_3^{(k)} + \frac{15}{8}. \end{aligned}$$

Tomando $\mathbf{x}^{(0)} = (0, 0, 0, 0)^t$, generamos los vectores iterados de la tabla 2.

Tabla 2

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
0	0.0000	0.0000	0.0000	0.0000
1	0.6000	2.3273	-0.98727	0.87885
2	1.0302	2.0369	-1.0145	0.98435
3	1.0066	2.0035	-1.0025	0.99838
4	1.0009	2.0003	-1.0003	0.99985
5	1.0001	2.0000	-1.0000	1.0000

Ya que

$$\frac{\|\mathbf{x}^{(5)} - \mathbf{x}^{(4)}\|_\infty}{\|\mathbf{x}^{(4)}\|_\infty} = \frac{0.0008}{2.000} = 4 \times 10^{-4} ,$$

se acepta $\mathbf{x}^{(5)}$ como una aproximación razonable a la solución. Es interesante notar que el método de Jacobi en el ejemplo 1 requiere el doble de iteraciones para la misma precisión.

La técnica presentada en el ejemplo 2 se llama **método iterativo de Gauss-Seidel**. Para escribir este método en la forma matricial (XVII.1) se multiplican ambos lados de la ecuación (XVII.6) por a_{ii} y se recolectan todos los k -ésimos términos iterados para dar

$$a_{i1} x_1^{(k)} + a_{i2} x_2^{(k)} + \dots + a_{ii} x_i^{(k)} = -a_{i,i+1} x_{i+1}^{(k-1)} - \dots - a_{in} x_n^{(k-1)} + b_i ,$$

para cada $i = 1, 2, \dots, n$. Escribiendo las n ecuaciones tenemos:

$$a_{11} x_1^{(k)} = -a_{12} x_2^{(k-1)} - a_{13} x_3^{(k-1)} - \dots - a_{1n} x_n^{(k-1)} + b_1 ,$$

$$a_{21} x_1^{(k)} + a_{22} x_2^{(k)} = -a_{23} x_3^{(k-1)} \dots - a_{2n} x_n^{(k-1)} + b_2 ,$$

... ..

$$a_{n1} x_1^{(k)} + a_{n2} x_2^{(k)} + \dots + a_{nn} x_n^{(k)} = b_n ,$$

y se sigue que, en forma matricial, el método de Gauss-Seidel puede ser representado como $(D - L) \mathbf{x}^{(k)} = U \mathbf{x}^{(k-1)} + \mathbf{b}$, ó

$$\mathbf{x}^{(k)} = (D - L)^{-1} U \mathbf{x}^{(k-1)} + (D - L)^{-1} \mathbf{b} . \tag{XVII.7}$$

Para que la matriz triangular inferior $(D - L)$ sea no singular, es necesario y suficiente que $a_{ii} \neq 0$ para cada $i = 1, 2, \dots, n$.

Para resumir el método iterativo de Gauss-Seidel, presentamos el siguiente algoritmo:

Algoritmo iterativo de Gauss-Seidel.

=====
 Para resolver el sistema lineal $A \mathbf{x} = \mathbf{b}$ con una aproximación inicial dada $\mathbf{x}^{(0)}$.

Entrada: número de incógnitas y de ecuaciones n ; las componentes de la matriz $A = (a_{ij})$ donde $1 \leq i, j \leq n$; las componentes b_i , con $1 \leq i \leq n$, del término no homogéneo \mathbf{b} ; las

componentes XO_i , con $1 \leq i \leq n$, de la aproximación inicial $\mathbf{XO} = \mathbf{x}^{(0)}$; la tolerancia TOL; el número máximo de iteraciones N_0 .

Salida: solución aproximada x_1, x_2, \dots, x_n ó mensaje de que el número de iteraciones fue excedido.

Paso 1: Tomar $k = 1$.

Paso 2: Mientras que $k \leq N_0$ seguir los pasos 3–6.

Paso 3: Para $i = 1, 2, \dots, n$ tomar

$$x_i = \frac{1}{a_{ii}} \left[- \sum_{j=1}^{i-1} (a_{ij} x_j) - \sum_{j=i+1}^n (a_{ij} XO_j) + b_i \right].$$

Paso 4: Si $\|\mathbf{x} - \mathbf{XO}\| < TOL$ entonces SALIDA (x_1, x_2, \dots, x_n) ; *(procedimiento completado satisfactoriamente)* PARAR.

Paso 5: Tomar $k = k + 1$.

Paso 6: Para $i = 1, 2, \dots, n$ tomar $XO_i = x_i$.

Paso 7: SALIDA *(número máximo de iteraciones excedido)*; *(procedimiento completado sin éxito)* PARAR.

=====

Los resultados de los ejemplos 1 y 2 parecen implicar que el método de Gauss-Seidel es superior al método de Jacobi. Este es generalmente cierto, pero no siempre. En realidad, hay sistemas lineales para los cuales el método de Jacobi converge y el método de Gauss-Seidel no, y viceversa.

3. CONVERGENCIA DE LOS PROCESOS ITERATIVOS

Para estudiar la convergencia de las técnicas generales de iteración, consideramos la fórmula (XVII.1)

$$\mathbf{x}^{(k)} = T \mathbf{x}^{(k-1)} + \mathbf{c}$$

para cada $k = 1, 2, \dots$, donde $\mathbf{x}^{(0)}$ es arbitrario. Este estudio requerirá del siguiente lema:

Lema XVII.1

Si el radio espectral $\rho(T)$ satisface que $\rho(T) < 1$, ó si la norma de la matriz T satisface que $\|T\| < 1$, entonces $(I - T)^{-1}$ existe y

$$(I - T)^{-1} = I + T + T^2 + \dots$$

Teorema XVII.2

Para cualquier $\mathbf{x}^{(0)} \in \mathcal{R}^n$, la sucesión $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ definida por (XVII.1)

$$\mathbf{x}^{(k)} = T \mathbf{x}^{(k-1)} + \mathbf{c}$$

para cada $k \geq 1$ y $\mathbf{c} \neq \mathbf{0}$, converge a la solución única de $\mathbf{x} = T \mathbf{x} + \mathbf{c}$ si y sólo si $\rho(T) < 1$.

Demostración: de la ecuación (XVII.1), se tiene que

$$\begin{aligned}
 \mathbf{x}^{(k)} &= T \mathbf{x}^{(k-1)} + \mathbf{c} = \\
 &= T (T \mathbf{x}^{(k-2)} + \mathbf{c}) + \mathbf{c} = \\
 &= T^2 \mathbf{x}^{(k-2)} + (T + I) \mathbf{c} = \\
 &\quad \dots \\
 &= T^k \mathbf{x}^{(0)} + (T^{k-1} + \dots + T + I) \mathbf{c} .
 \end{aligned}$$

Suponiendo que $\rho(T) < 1$, podemos usar el Teorema XIII.15 y el Lema XVII.1 para obtener

$$\begin{aligned}
 \lim_{k \rightarrow \infty} \mathbf{x}^{(k)} &= \lim_{k \rightarrow \infty} T^k \mathbf{x}^{(0)} + \lim_{k \rightarrow \infty} \left(\sum_{j=0}^{k-1} T^j \right) \mathbf{c} \\
 &= 0 \cdot \mathbf{x}^{(0)} + (I - T)^{-1} \mathbf{c} = (I - T)^{-1} \mathbf{c} .
 \end{aligned}$$

De (XVII.1) $\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = (I - T)^{-1} \mathbf{c}$ será la solución única de $\mathbf{x} = T \mathbf{x} + \mathbf{c}$.

Para probar el recíproco, sea $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ convergente a \mathbf{x} para cualquier $\mathbf{x}^{(0)}$. De la ecuación (XVII.1) sigue que $\mathbf{x} = T \mathbf{x} + \mathbf{c}$, así que para cada k ,

$$\mathbf{x} - \mathbf{x}^{(k)} = T (\mathbf{x} - \mathbf{x}^{(k-1)}) = \dots = T^k (\mathbf{x} - \mathbf{x}^{(0)}) .$$

Por lo tanto, para cualquier vector $\mathbf{x}^{(0)}$,

$$\lim_{k \rightarrow \infty} T^k (\mathbf{x} - \mathbf{x}^{(0)}) = \lim_{k \rightarrow \infty} \mathbf{x} - \mathbf{x}^{(k)} = \mathbf{0} .$$

Consecuentemente, si \mathbf{z} es un vector arbitrario y $\mathbf{x}^{(0)} = \mathbf{x} - \mathbf{z}$, entonces

$$\lim_{k \rightarrow \infty} T^k \mathbf{z} = \lim_{k \rightarrow \infty} T^k [\mathbf{x} - (\mathbf{x} - \mathbf{z})] = \mathbf{0} ,$$

lo cual, por el Teorema XIII.15, implica que $\rho(T) < 1$.

c.q.d.

Un Teorema parecido nos dará condiciones de suficiencia para la convergencia de los procesos de iteración usando las normas en lugar del radio espectral.

Teorema XVII.3

Si $\|T\| < 1$, para cualquier norma matricial natural, entonces la sucesión definida en la ecuación (XVII.1), $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$, converge para cualquier $\mathbf{x}^{(0)} \in \mathcal{R}^n$, a un vector $\mathbf{x} \in \mathcal{R}^n$, y se satisfacen las siguientes cotas de error:

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \|T\|^k \|\mathbf{x}^{(0)} - \mathbf{x}\| , \quad (\text{XVII.8})$$

y

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{\|T\|^k}{1 - \|T\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| . \quad (\text{XVII.9})$$

Demostración: comenzando con un vector arbitrario $\mathbf{x}^{(0)}$, formaremos una secuencia de aproximaciones

$$\begin{aligned} \mathbf{x}^{(1)} &= T \mathbf{x}^{(0)} + \mathbf{c} , \\ \mathbf{x}^{(2)} &= T \mathbf{x}^{(1)} + \mathbf{c} , \\ &\dots \quad \dots \quad \dots \quad \dots \\ \mathbf{x}^{(k)} &= T \mathbf{x}^{(k-1)} + \mathbf{c} , \end{aligned}$$

de donde

$$\mathbf{x}^{(k)} = T^k \mathbf{x}^{(0)} + (T^{k-1} + \dots + T + I) \mathbf{c} .$$

Como para $\|T\| < 1$ tenemos $\|T^k\| \rightarrow 0$ cuando $k \rightarrow \infty$, se deduce que

$$\lim_{k \rightarrow \infty} T^k = 0 \quad \text{y} \quad \lim_{k \rightarrow \infty} (I + T + T^2 + \dots + T^{k-1}) = \sum_{k=0}^{\infty} T^k = (I - T)^{-1} .$$

Y por tanto, pasando al límite cuando $k \rightarrow \infty$, tenemos

$$\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = (I - T)^{-1} \mathbf{c} .$$

Esto prueba la convergencia del proceso iterativo. Además, tenemos $(I - T) \mathbf{x} = \mathbf{c}$ ó $\mathbf{x} = T \mathbf{x} + \mathbf{c}$, lo cual quiere decir que el vector \mathbf{x} en el límite es una solución del sistema. Como la matriz $(I - T)$ no es singular, la solución \mathbf{x} es única. Hemos así demostrado la primera parte del Teorema.

Demostramos ahora la cota de error (XVII.8). Supongamos que $\mathbf{x}^{(k+p)}$ y $\mathbf{x}^{(k)}$ son dos aproximaciones de la solución del sistema lineal $\mathbf{x} = T \mathbf{x} + \mathbf{c}$; de la ecuación (XVII.1), tenemos:

$$\begin{aligned} \|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}\| &= \|T \mathbf{x}^{(k+p-1)} - T \mathbf{x}^{(k-1)}\| = \|T (\mathbf{x}^{(k+p-1)} - \mathbf{x}^{(k-1)})\| = \dots \\ &= \|T^k (\mathbf{x}^{(p)} - \mathbf{x}^{(0)})\| \leq \|T\|^k \|\mathbf{x}^{(p)} - \mathbf{x}^{(0)}\| . \end{aligned}$$

Ahora pasando al límite cuando $p \rightarrow \infty$, obtenemos

$$\lim_{p \rightarrow \infty} \|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}\| \leq \lim_{p \rightarrow \infty} \|T\|^k \|\mathbf{x}^{(p)} - \mathbf{x}^{(0)}\| = \|T\|^k \lim_{p \rightarrow \infty} \|\mathbf{x}^{(p)} - \mathbf{x}^{(0)}\|$$

y entonces

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \|T\|^k \|\mathbf{x} - \mathbf{x}^{(0)}\| ,$$

que es la cota de error (XVII.8)

Finalmente demostramos la cota de error (XVII.9). Como antes, supongamos que $\mathbf{x}^{(k+p)}$ y $\mathbf{x}^{(k)}$ son dos aproximaciones de la solución del sistema lineal $\mathbf{x} = T \mathbf{x} + \mathbf{c}$. Tenemos

$$\|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}\| \leq \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| + \|\mathbf{x}^{(k+2)} - \mathbf{x}^{(k+1)}\| + \dots + \|\mathbf{x}^{(k+p)} - \mathbf{x}^{(k+p-1)}\| .$$

Por lo visto antes:

$$\|\mathbf{x}^{(m+1)} - \mathbf{x}^{(m)}\| \leq \|T\| \|\mathbf{x}^{(m)} - \mathbf{x}^{(m-1)}\| \leq \|T\|^{m-k} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| ,$$

para $m > k \geq 1$. Entonces tenemos:

$$\begin{aligned} \|\mathbf{x}^{(p+k)} - \mathbf{x}^{(k)}\| &\leq \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| + \|T\| \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| + \dots + \\ &\quad + \|T\|^{p-1} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq \frac{1}{1 - \|T\|} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq \\ &\leq \frac{\|T\|}{1 - \|T\|} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \dots \leq \frac{\|T\|^k}{1 - \|T\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| , \end{aligned}$$

de donde se deduce la cota de error (XVII.9).

c.q.d.

Notése que si en particular elegimos $\mathbf{x}^{(0)} = \mathbf{c}$, entonces $\mathbf{x}^{(1)} = T \mathbf{c} + \mathbf{c}$ y

$$\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| = \|T \mathbf{c}\| \leq \|T\| \|\mathbf{c}\| ,$$

y la cota (XVII.9) nos da:

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \frac{\|T\|^{k+1}}{1 - \|T\|} \|\mathbf{c}\| . \quad (\text{XVII.9}')$$

Ejemplo 3.

Demostrar que el proceso de iteración de Jacobi es convergente para el sistema lineal siguiente:

$$\begin{aligned} E_1 : & 10 x_1 - x_2 + 2 x_3 - 3 x_4 = 0 , \\ E_2 : & x_1 + 10 x_2 - x_3 + 2 x_4 = 5 , \\ E_3 : & 2 x_1 + 3 x_2 + 20 x_3 - x_4 = -10 , \\ E_4 : & 3 x_1 + 2 x_2 + x_3 + 20 x_4 = 15 . \end{aligned}$$

¿Cuántas iteraciones han de efectuarse para hallar las raíces del sistema con un error menor de 10^{-4} ?

Reduciendo el sistema a la forma especial para la iteración de Jacobi, tenemos

$$\begin{aligned} x_1 &= 0.1 x_2 - 0.2 x_3 + 0.3 x_4 , \\ x_2 &= -0.1 x_1 + 0.1 x_3 - 0.2 x_4 + 0.5 , \\ x_3 &= -0.1 x_1 - 0.15 x_2 + 0.05 x_4 - 0.5 , \\ x_4 &= -0.15 x_1 - 0.1 x_2 - 0.05 x_3 + 0.75 . \end{aligned}$$

Entonces la matriz del sistema es:

$$T = \begin{pmatrix} 0 & 0.1 & -0.2 & 0.3 \\ -0.1 & 0 & 0.1 & -0.2 \\ -0.1 & -0.15 & 0 & 0.05 \\ -0.15 & -0.1 & -0.05 & 0 \end{pmatrix} .$$

Utilizando, por ejemplo, la norma l_1 , tenemos:

$$\|T\|_1 = \max\{0.35, 0.35, 0.35, 0.55\} = 0.55 < 1 .$$

En consecuencia el proceso de iteración para el sistema dado es convergente. Si consideramos como aproximación inicial de la raíz \mathbf{x} el vector

$$\mathbf{x}^{(0)} = \mathbf{c} = (0.0, 0.5, -0.5, 0.75)^t ,$$

entonces

$$\|\mathbf{c}\|_1 = 0.0 + 0.5 + 0.5 + 0.75 = 1.75 .$$

Sea ahora k el número de iteraciones requeridas para conseguir la exactitud especificada. Utilizando la fórmula (XVII.9'), tenemos:

$$\|\mathbf{x} - \mathbf{x}^{(k)}\|_1 \leq \frac{\|T\|_1^{k+1}}{1 - \|T\|_1} \|\mathbf{c}\|_1 = \frac{0.55^{k+1} \times 1.75}{0.45} < 10^{-4} .$$

De aquí,

$$0.55^{k+1} < \frac{45}{175} 10^{-4}$$

o sea

$$\begin{aligned} (k+1) \log_{10} 0.55 &< \log_{10} 45 - \log_{10} 175 - 4 \\ -(k+1) 0.25964 &< 1.65321 - 2.24304 - 4 = -4.58983 \end{aligned}$$

y consecuentemente

$$k+1 > \frac{4.58983}{0.25964} \approx 17.7 \quad \implies \quad k > 16.7 .$$

Podemos tomar $k = 17$. Nótese que la estimación teórica del número de iteraciones necesarias para asegurar la exactitud especificada es excesivamente alto. A menudo se obtiene la exactitud deseada en un número menor de iteraciones.

Para aplicar los resultados de arriba a las técnicas iterativas de Jacobi o Gauss-Seidel, necesitamos escribir las matrices de iteración del método de Jacobi, T_J , dadas en (XVII.5) y del método de Gauss-Seidel, T_{GS} , dadas en (XVII.7), como

$$T_J = D^{-1} (L + U) \quad \text{y} \quad T_{GS} = (D - L)^{-1} U .$$

De ser $\rho(T_J)$ ó $\rho(T_{GS})$ menores que uno, es claro que la sucesión $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ converge a la solución \mathbf{x} de $A \mathbf{x} = \mathbf{b}$. Por ejemplo, el esquema de Jacobi (ver ecuación (XVII.5)) tiene:

$$\mathbf{x}^{(k)} = D^{-1} (L + U) \mathbf{x}^{(k-1)} + D^{-1} \mathbf{b} ,$$

y si $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ converge a \mathbf{x} , entonces

$$\mathbf{x} = D^{-1} (L + U) \mathbf{x} + D^{-1} \mathbf{b} .$$

Esto implica que

$$D \mathbf{x} = (L + U) \mathbf{x} + \mathbf{b} \quad \text{y} \quad (D - L - U) \mathbf{x} = \mathbf{b} .$$

Ya que $D - L - U = A$, luego \mathbf{x} satisface $A \mathbf{x} = \mathbf{b}$. De manera parecida se procede con el esquema de Gauss-Seidel dado por la ecuación (XVII.7).

Podemos dar ahora condiciones de suficiencia fáciles de verificar para la convergencia de los métodos de Jacobi y de Gauss-Seidel.

Teorema XVII.4

Si A es una matriz estrictamente dominante diagonalmente, entonces, para cualquier elección de $\mathbf{x}^{(0)} \in \mathcal{R}^n$ ambos métodos, el de Jacobi o el de Gauss-Seidel, dan lugar a sucesiones $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ que convergen a la solución de $A \mathbf{x} = \mathbf{b}$.

La relación entre la rapidez de convergencia y el radio espectral de la matriz de iteración T se puede ver de la desigualdad (XVII.8). Como (XVII.8) se satisface para cualquier norma matricial natural se sigue, de la afirmación que siguió al Teorema XIII.14, que

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| \approx \rho(T)^k \|\mathbf{x}^{(0)} - \mathbf{x}\| . \quad (\text{XVII.10})$$

Supongamos que $\rho(T) < 1$ y que se va a usar $\mathbf{x}^{(0)} = \mathbf{0}$ en una técnica iterativa para aproximar \mathbf{x} con un error relativo máximo de 10^{-t} . Por la estimación (XVII.10), el error relativo después de k iteraciones es aproximadamente $\rho(T)^k$, así que se espera una precisión de 10^{-t} si

$$\rho(T)^k \leq 10^{-t} ,$$

esto es, si

$$k \geq \frac{t}{-\log_{10} \rho(T)} .$$

Por lo tanto, es deseable escoger la técnica iterativa con el menor $\rho(T) < 1$ para el sistema particular $A \mathbf{x} = \mathbf{b}$.

En general no se conoce cuál de las dos técnicas, la de Jacobi o la de Gauss-Seidel, debe usarse. Sin embargo, en un caso especial, sí se conoce la respuesta.

Teorema XVII.5 (Stein-Rosenberg)

Si $a_{ij} \leq 0$ para cada $i \neq j$ y $a_{ii} > 0$ para cada $i = 1, 2, \dots, n$, entonces se satisface una y solamente una de las siguientes afirmaciones:

- a) $0 < \rho(T_{GS}) < \rho(T_J) < 1$;
- b) $1 < \rho(T_J) < \rho(T_{GS})$;
- c) $\rho(T_{GS}) = \rho(T_J) = 0$;
- d) $\rho(T_J) = \rho(T_{GS}) = 1$;

Para el caso especial descrito en el Teorema XVII.5, vemos que cuando un método converge, entonces ambos convergen, siendo el método de Gauss-Seidel más rápido que el método de Jacobi.

4. LOS METODOS DE RELAJACION

Como la razón de convergencia de un procedimiento depende del radio espectral de la matriz asociada con el método, una manera de seleccionar un procedimiento que nos lleve a una convergencia acelerada consiste en escoger un método cuya matriz asociada tenga un radio espectral mínimo. Estos procedimientos nos llevan a los métodos de relajación. Pero antes de formular la teoría de los métodos de relajación, veamos las ideas fundamentales de la forma más simple. Supongamos que se dispone de un sistema de ecuaciones lineales

$$\begin{aligned}
 E_1 : & \quad a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1 , \\
 E_2 : & \quad a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n = b_2 , \\
 & \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\
 E_n : & \quad a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n = b_n .
 \end{aligned}
 \tag{XVII.11}$$

Transformaremos este sistema de la manera siguiente: pondremos los términos constantes a la izquierda y dividiremos la primera ecuación por $-a_{11}$, la segunda por $-a_{22}$, etc. Obtendremos entonces un sistema que está listo para la relajación:

$$\begin{aligned}
 E_1 : & \quad - x_1 + b_{12} x_2 + \dots + b_{1n} x_n + c_1 = 0 , \\
 E_2 : & \quad b_{21} x_1 - x_2 + \dots + b_{2n} x_n + c_2 = 0 , \\
 & \quad \dots \\
 E_n : & \quad b_{n1} x_1 + b_{n2} x_2 + \dots - x_n + c_n = 0 ,
 \end{aligned}
 \tag{XVII.12}$$

donde

$$b_{ij} = -\frac{a_{ij}}{a_{ii}} \quad (i \neq j) \quad \text{y} \quad c_i = \frac{b_i}{a_{ii}} .
 \tag{XVII.13}$$

Supongamos que $\mathbf{x}^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ es la aproximación inicial a la solución del sistema dado. Sustituyendo estos valores en el sistema tendremos los **restos**

$$\begin{aligned}
 R_1^{(0)} &= c_1 - x_1^{(0)} + \sum_{j=2}^n b_{1j} x_j^{(0)} = x_1^{(1)} - x_1^{(0)} , \\
 & \quad \dots \quad \dots \quad \dots \quad \dots \\
 R_k^{(0)} &= c_k - x_k^{(0)} + \sum_{\substack{j=1 \\ j \neq k}}^n b_{kj} x_j^{(0)} = x_k^{(1)} - x_k^{(0)} , \\
 & \quad \dots \quad \dots \quad \dots \quad \dots \\
 R_n^{(0)} &= c_n - x_n^{(0)} + \sum_{j=1}^{n-1} b_{nj} x_j^{(0)} = x_n^{(1)} - x_n^{(0)} .
 \end{aligned}
 \tag{XVII.14}$$

Si damos un incremento $\delta x_s^{(0)}$ a una de las incógnitas $x_s^{(0)}$, el resto correspondiente $R_s^{(0)}$ quedará disminuido en $\delta x_s^{(0)}$ y todos los otros restos $R_i^{(0)}$ ($i \neq s$) quedarán aumentados en

$b_{is} \delta x_s^{(0)}$. De este modo, para hacer que desaparezca el resto siguiente $R_i^{(1)}$ es suficiente dar a $x_s^{(1)}$ un incremento $\delta x_s^{(1)} = R_s^{(0)}$ y tendremos

$$R_s^{(1)} = 0 \quad \text{y} \quad R_i^{(1)} = R_i^{(0)} + b_{is} \delta x_s^{(0)} \quad \text{para} \quad i \neq s . \quad (\text{XVII.15})$$

Así el **método de relajación**, en su forma más simple, consiste en reducir el resto numéricamente más elevado a cero, en cada etapa, cambiando el valor del componente apropiado de la aproximación. El proceso acaba cuando todos los restos del último sistema transformado son iguales a cero con la exactitud requerida.

Ejemplo 4.

Vamos a resolver el sistema lineal $A \mathbf{x} = \mathbf{b}$ dado por

$$\begin{aligned} E_1 : & 10 x_1 - 2 x_2 - 2 x_3 = 6 , \\ E_2 : & - x_1 + 10 x_2 - 2 x_3 = 7 , \\ E_3 : & - x_1 - x_2 + 10 x_3 = 8 , \end{aligned}$$

con el método de relajación y aritmética de dos dígitos.

Reduzcamos el sistema a una forma conveniente para la relajación:

$$\begin{aligned} E_1 : & - x_1 + 0.2 x_2 + 0.2 x_3 + 0.6 = 0 , \\ E_2 : & 0.1 x_1 - x_2 + 0.2 x_3 + 0.7 = 0 , \\ E_3 : & 0.1 x_1 + 0.1 x_2 - x_3 + 0.8 = 0 . \end{aligned}$$

Eligiendo los valores $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$ como aproximaciones iniciales de las raíces, obtenemos los restos:

$$R_1^{(0)} = 0.60, \quad R_2^{(0)} = 0.70, \quad R_3^{(0)} = 0.80.$$

Siendo $R_3^{(0)} = 0.80$ el resto más grande, consideramos

$$\delta x_3^{(0)} = 0.80$$

de donde obtendremos los restos

$$\begin{aligned} R_1^{(1)} &= R_1^{(0)} + 0.2 \times 0.8 = 0.60 + 0.16 = 0.76 , \\ R_2^{(1)} &= R_2^{(0)} + 0.2 \times 0.8 = 0.70 + 0.16 = 0.86 , \\ R_3^{(1)} &= R_3^{(0)} - 0.8 = 0.80 - 0.80 = 0.0 . \end{aligned}$$

Establezcamos ahora

$$\delta x_2^{(1)} = 0.86$$

y así sucesivamente. Los resultados de los cálculos se dan en la tabla 3.

Sumando todos los incrementos $\delta x_i^{(k)}$ ($i = 1, 2, 3; k = 0, 1, \dots$), tendremos los valores de las raíces:

$$\begin{aligned} x_1 &= 0.0 + 0.93 + 0.062 + 0.007 = 0.999 = 1.0 , \\ x_2 &= 0.0 + 0.86 + 0.13 + 0.016 = 1.006 = 1.0 , \\ x_3 &= 0.0 + 0.80 + 0.18 + 0.019 = 0.999 = 1.0 . \end{aligned}$$

En este caso el sistema se ha resuelto exactamente.

Tabla 3

k	x_1	R_1	x_2	R_2	x_3	R_3
0	0.0	0.60 0.16 0.76	0.0	0.70 0.16 0.86	0.0 0.80	0.80 -0.80 0.0
1		0.17 0.93	0.86	-0.86 0.0		0.086 0.086
2	0.93	-0.93 0.0		0.093 0.093		0.093 0.18
3		0.036 0.036		0.036 0.13	0.18	-0.18 0.0
4		0.026 0.062	0.13	-0.13 0.0		0.013 0.013
5	0.062	-0.062 0.0		0.012 0.012		0.0062 0.019
6		0.0038 0.0038		0.0038 0.016	0.019	-0.019 0.0
7		0.0032 0.007	0.016	-0.016 0.0		0.0016 0.0016
8	0.007	-0.007 0.0		0.0007 0.0007		0.0007 0.0023
Σ	1.0		1.0		1.0	

Vamos ahora a describir los métodos de relajación. Antes de describir un procedimiento para seleccionar tales métodos, necesitamos introducir una manera nueva de medir la cantidad por la cual una aproximación a la solución de un sistema lineal difiere de la solución real del sistema. El método hace uso del denominado **vector residual**.

Definición. Si $\tilde{\mathbf{x}} \in \mathcal{R}^n$ es una aproximación a la solución del sistema lineal definido por $A \mathbf{x} = \mathbf{b}$, el **vector residual** de $\tilde{\mathbf{x}}$ con respecto a este sistema se define como $\mathbf{r} = \mathbf{b} - A \tilde{\mathbf{x}}$.

En procedimientos como los métodos de Jacobi o de Gauss-Seidel se asocia un vector residual con cada cálculo de una componente aproximada del vector solución. El objetivo del método consiste en generar una sucesión de aproximaciones que hagan que los vectores residuales asociados converjan a cero. Supongamos que tomamos

$$\mathbf{r}_i^{(k)} = (r_{1i}^{(k)}, r_{2i}^{(k)}, \dots, r_{ni}^{(k)})^t$$

para denotar al vector residual para el método de Gauss-Seidel correspondiente al vector solución aproximado

$$(x_1^{(k)}, x_2^{(k)}, \dots, x_{i-1}^{(k)}, x_i^{(k-1)}, \dots, x_n^{(k-1)})^t .$$

La m -ésima componente de $\mathbf{r}_i^{(k)}$ es

$$r_{mi}^{(k)} = b_m - \sum_{j=1}^{i-1} a_{mj} x_j^{(k)} - \sum_{j=i}^n a_{mj} x_j^{(k-1)} \quad (XVII.16)$$

ó

$$r_{mi}^{(k)} = b_m - \sum_{j=1}^{i-1} a_{mj} x_j^{(k)} - \sum_{j=i+1}^n a_{mj} x_j^{(k-1)} - a_{mi} x_i^{(k-1)}$$

para cada $m = 1, 2, \dots, n$. En particular, la i -ésima componente de $\mathbf{r}_i^{(k)}$ es

$$r_{ii}^{(k)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} - a_{ii} x_i^{(k-1)} ;$$

así que

$$a_{ii} x_i^{(k-1)} + r_{ii}^{(k)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} . \quad (XVII.17)$$

Recuérdese, sin embargo, que en el método de Gauss-Seidel $x_i^{(k)}$ se escoge como

$$x_i^{(k)} = \frac{- \sum_{j=1}^{i-1} (a_{ij} x_j^{(k)}) - \sum_{j=i+1}^n (a_{ij} x_j^{(k-1)}) + b_i}{a_{ii}} , \quad (XVII.6)$$

así que la ecuación (XVII.17) puede escribirse como $a_{ii} x_i^{(k-1)} + r_{ii}^{(k)} = a_{ii} x_i^{(k)}$ ó

$$x_i^{(k)} = x_i^{(k-1)} + \frac{r_{ii}^{(k)}}{a_{ii}} . \quad (XVII.18)$$

Podemos derivar otra conexión entre los vectores residuales y la técnica de Gauss-Seidel. De (XVII.16), la i -ésima componente de $\mathbf{r}_{i+1}^{(k)}$ es

$$\begin{aligned} r_{i,i+1}^{(k)} &= b_i - \sum_{j=1}^i a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} \\ &= b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} - a_{ii} x_i^{(k)} . \end{aligned} \quad (XVII.19)$$

La ecuación (XVII.6) implica que $r_{i,i+1}^{(k)} = 0$. Entonces, en cierto sentido, la técnica de Gauss-Seidel está ideada para requerir que la i -ésima componente de $\mathbf{r}_{i+1}^{(k)}$ sea cero.

Reducir una coordenada del vector residual a cero, sin embargo, no es necesariamente la manera más eficiente de reducir la norma del vector $\mathbf{r}_{i+1}^{(k)}$. En realidad, modificando el procedimiento de Gauss-Seidel en la forma de la ecuación (XVII.18) a:

$$x_i^{(k)} = x_i^{(k-1)} + \omega \frac{r_{ii}^{(k)}}{a_{ii}} \tag{XVII.20}$$

para ciertas elecciones de ω positivo nos llevará a una convergencia significativamente más rápida.

Los métodos que emplean la ecuación (XVII.20) se conocen como **métodos de relajación**. Para $0 < \omega < 1$, los procedimientos se llaman **métodos de sub-relajación** y se pueden emplear para obtener la convergencia de algunos sistemas que no son convergentes por el método de Gauss-Seidel. Para $\omega > 1$, los procedimientos se llaman **métodos de sobre-relajación** y se pueden usar para acelerar la convergencia de sistemas que son convergentes por el método de Gauss-Seidel. Estos métodos se abrevian frecuentemente como **SOR** (de **S**uccessive **O**ver-**R**elaxation) y son particularmente útiles para resolver los sistemas lineales que aparecen en la solución numérica de ciertas ecuaciones diferenciales parciales.

Antes de ilustrar las ventajas del método SOR notamos que usando la ecuación (XVII.17), la ecuación (XVII.20) se puede reformular para propósitos de cómputo como

$$x_i^{(k)} = (1 - \omega) x_i^{(k-1)} + \frac{\omega}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} \right]. \tag{XVII.21}$$

Para determinar la forma matricial del método SOR reescribimos (XVII.21) como

$$a_{ii} x_i^{(k)} + \omega \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} = (1 - \omega) a_{ii} x_i^{(k-1)} - \omega \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} + \omega b_i$$

así que

$$(D - \omega L) \mathbf{x}^{(k)} = [(1 - \omega) D + \omega U] \mathbf{x}^{(k-1)} + \omega \mathbf{b}$$

ó

$$\mathbf{x}^{(k)} = (D - \omega L)^{-1} [(1 - \omega) D + \omega U] \mathbf{x}^{(k-1)} + \omega (D - \omega L)^{-1} \mathbf{b} .$$

Algoritmo iterativo Successive Over-Relaxation (SOR).

=====

Para resolver el sistema lineal $A \mathbf{x} = \mathbf{b}$ dados el parámetro ω y una aproximación inicial $\mathbf{x}^{(0)}$.

Entrada: número de incógnitas y de ecuaciones n ; las componentes de la matriz $A = (a_{ij})$ donde $1 \leq i, j \leq n$; las componentes b_i , con $1 \leq i \leq n$, del término no homogéneo \mathbf{b} ; las componentes XO_i , con $1 \leq i \leq n$, de la aproximación inicial $\mathbf{XO} = \mathbf{x}^{(0)}$; el parámetro ω ; la tolerancia TOL; el número máximo de iteraciones N_0 .

Salida: solución aproximada x_1, x_2, \dots, x_n ó mensaje de que el número de iteraciones fue excedido.

Paso 1: Tomar $k = 1$.

Paso 2: Mientras que $k \leq N_0$ seguir los pasos 3–6.

Paso 3: Para $i = 1, 2, \dots, n$ tomar

$$x_i = (1 - \omega) XO_i + \frac{\omega}{a_{ii}} \left[- \sum_{j=1}^{i-1} (a_{ij} x_j) - \sum_{j=i+1}^n (a_{ij} XO_j) + b_i \right].$$

Paso 4: Si $\|\mathbf{x} - \mathbf{XO}\| < TOL$ entonces SALIDA (x_1, x_2, \dots, x_n) ;
(procedimiento completado satisfactoriamente) PARAR.

Paso 5: Tomar $k = k + 1$.

Paso 6: Para $i = 1, 2, \dots, n$ tomar $XO_i = x_i$.

Paso 7: SALIDA (número máximo de iteraciones excedido);
(procedimiento completado sin éxito) PARAR.

Ejemplo 5.

El sistema lineal $A \mathbf{x} = \mathbf{b}$ dado por

$$\begin{aligned} E_1 : & 4 x_1 + 3 x_2 & = & 24 , \\ E_2 : & 3 x_1 + 4 x_2 - x_3 & = & 30 , \\ E_3 : & - x_2 + 4 x_3 & = & -24 , \end{aligned}$$

tiene por solución $\mathbf{x} = (3, 4, -5)^t$. Se usarán los métodos de Gauss-Seidel y el SOR con $\omega = 1.25$ para resolver este sistema usando $\mathbf{x}^{(0)} = (1, 1, 1)^t$ para ambos métodos. Las ecuaciones para el método de Gauss-Seidel son

$$\begin{aligned} x_1^{(k)} &= -0.75 x_2^{(k-1)} + 6 , \\ x_2^{(k)} &= -0.75 x_1^{(k)} + 0.25 x_3^{(k-1)} + 7.5 , \\ x_3^{(k)} &= 0.25 x_2^{(k)} - 6 , \end{aligned}$$

para cada $k = 1, 2, \dots$, y las ecuaciones para el método SOR con $\omega = 1.25$ son

$$\begin{aligned} x_1^{(k)} &= -0.25 x_1^{(k-1)} - 0.9375 x_2^{(k-1)} + 7.5 , \\ x_2^{(k)} &= -0.9375 x_1^{(k)} - 0.25 x_2^{(k-1)} + 0.3125 x_3^{(k-1)} + 9.375 , \\ x_3^{(k)} &= 0.3125 x_2^{(k)} - 0.25 x_3^{(k-1)} - 7.5 . \end{aligned}$$

Las primeras siete iteraciones de cada método se muestran en las tablas 4 y 5.

Para obtener una precisión de siete lugares decimales el método de Gauss-Seidel requiere de 34 iteraciones en contra de las 14 que se necesitan en el método de sobre-relajación con $\omega = 1.25$.

Tabla 4

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1.000000	1.000000	1.000000
1	5.250000	3.812500	-5.046875
2	3.1406250	3.8828125	-5.0292969
3	3.0878906	3.9267578	-5.0183105
4	3.0549317	3.9542236	-5.0114441
5	3.0343323	3.9713898	-5.0071526
6	3.0214577	3.9821186	-5.0044703
7	3.0134111	3.9888241	-5.0027940

Tabla 5

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	1.000000	1.000000	1.000000
1	6.312500	3.5195313	-6.6501465
2	2.6223144	3.9585266	-4.6004238
3	3.1333027	4.0102646	-5.0966864
4	2.9570513	4.0074838	-4.9734897
5	3.0037211	4.0029250	-5.0057135
6	2.9963275	4.0009263	-4.9982822
7	3.0000498	4.0002586	-5.0003486

Un problema que se presenta al usar el método SOR, es cómo escoger el valor apropiado de ω . Aún cuando no se conoce una respuesta completa a esta pregunta para un sistema lineal general $n \times n$, los siguientes resultados pueden usarse en ciertas situaciones.

Teorema XVII.6 (Kahan)

Si $a_{ii} \neq 0$ para cada $i = 1, 2, \dots, n$, entonces $\rho(T_\omega) \geq |\omega - 1|$. Esto implica que $\rho(T_\omega) < 1$ sólo si $0 < \omega < 2$, donde $T_\omega = (D - \omega L)^{-1} [(1 - \omega) D + \omega U]$ es la matriz de iteración del método SOR.

Teorema XVII.7 (Ostrowski-Reich)

Si A es una matriz positiva definida y $0 < \omega < 2$, entonces el método SOR converge para cualquier elección de la aproximación inicial $\mathbf{x}^{(0)}$ del vector solución.

Teorema XVII.8

Si A es una matriz positiva definida y tridiagonal, entonces $\rho(T_{GS}) = [\rho(T_J)]^2 < 1$, la elección óptima de ω para el método SOR es

$$\omega = \frac{2}{1 + \sqrt{1 - [\rho(T_J)]^2}}, \tag{XVII.22}$$

y con este valor de ω , $\rho(T_\omega) = \omega - 1$.

Ejemplo 6.

En el ejemplo 5 la matriz A estaba dada por

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

Esta matriz es positiva definida y tridiagonal, así que se aplica el Teorema XVII.8. Como

$$\begin{aligned} T_J &= D^{-1} (L + U) = \\ &= \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} 0 & -3 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & -0.75 & 0 \\ -0.75 & 0 & 0.25 \\ 0 & 0.25 & 0 \end{pmatrix}, \end{aligned}$$

tenemos que

$$T_J - \lambda I = \begin{pmatrix} -\lambda & -0.75 & 0 \\ -0.75 & -\lambda & 0.25 \\ 0 & 0.25 & -\lambda \end{pmatrix},$$

con lo que

$$\det(T_J - \lambda I) = -\lambda(\lambda^2 - 0.625).$$

Por lo tanto,

$$\rho(T_J) = \sqrt{0.625} \approx 0.79057$$

y la ecuación (XVII.22) nos da

$$\omega = \frac{2}{1 + \sqrt{1 - \rho(T_{GS})}} = \frac{2}{1 + \sqrt{1 - [\rho(T_J)]^2}} = \frac{2}{1 + \sqrt{1 - 0.625}} \approx 1.24.$$

Esto explica la rápida convergencia obtenida usando $\omega = 1.25$ en el ejemplo 5.

5. ELECCION DEL METODO PARA RESOLVER SISTEMAS LINEALES

Cuando el sistema lineal es lo suficientemente pequeño para que sea fácilmente acomodado en la memoria principal de un ordenador, es en general más eficaz usar una técnica directa que minimice el efecto del error de redondeo. Específicamente, es adecuado el algoritmo de eliminación Gaussiana con pivoteo escalado de columna.

Los sistemas lineales grandes cuyos coeficientes son entradas básicamente de ceros y que aparecen en patrones regulares se pueden resolver generalmente de una manera eficiente usando un procedimiento iterativo como el discutido en este capítulo. Los sistemas de este tipo aparecen naturalmente, por ejemplo, cuando se usan técnicas de diferencias finitas para resolver problemas de valor en la frontera, una aplicación común en la solución numérica de ecuaciones diferenciales parciales.

EJERCICIOS.

1. Encontrar las dos primeras iteraciones del método de Jacobi, del método de Gauss-Seidel y del método SOR para los siguientes sistemas lineales, usando $\mathbf{x}^{(0)} = \mathbf{0}$.

$$\begin{array}{rcl}
 a) & 2 x_1 & - x_2 + x_3 = -1, \\
 & 3 x_1 & + 3 x_2 + 9 x_3 = 0, \\
 & 3 x_1 & + 3 x_2 + 5 x_3 = 4.
 \end{array}$$

$$\begin{array}{rcl}
 b) & & 2 x_2 + 4 x_3 = 0, \\
 & x_1 & - x_2 - x_3 = 0.375, \\
 & x_1 & - x_2 + 2 x_3 = 0.
 \end{array}$$

$$\begin{array}{rcl}
 c) & 10 x_1 & - x_2 = 9, \\
 & - x_1 & + 10 x_2 - 2 x_3 = 7, \\
 & & - 2 x_2 + 10 x_3 = 6.
 \end{array}$$

$$\begin{array}{rcl}
 d) & 2 x_1 & = 3, \\
 & x_1 & + 1.5 x_2 = 4.5, \\
 & & - 3 x_2 + 0.5 x_3 = -6.6, \\
 & 2 x_1 & - 2 x_2 + x_3 + x_4 = 0.8.
 \end{array}$$

$$\begin{array}{rcl}
 e) & 2 x_1 & - x_2 + 10 x_3 = -11, \\
 & & 3 x_2 - x_3 + 8 x_4 = -11, \\
 & 10 x_1 & - x_2 + 2 x_3 = 6, \\
 & - x_1 & + 11 x_2 - x_3 + 3 x_4 = 25.
 \end{array}$$

$$\begin{array}{rcl}
 f) & 10 x_1 & - x_2 + 2 x_3 = 6, \\
 & - x_1 & + 11 x_2 - x_3 + 3 x_4 = 25, \\
 & 2 x_1 & - x_2 + 10 x_3 = -11, \\
 & & 3 x_2 - x_3 + 8 x_4 = -11.
 \end{array}$$

$$\begin{array}{rcl}
 g) & 4 x_1 & - 2 x_2 = 0, \\
 & -2 x_1 & + 5 x_2 - x_3 = 2, \\
 & & - x_2 + 4 x_3 + 2 x_4 = 3, \\
 & & & 2 x_3 + 3 x_4 = -2.
 \end{array}$$

2. Aplicar, si es posible, los algoritmos iterativos de Jacobi, de Gauss-Seidel y el algoritmo SOR, con $\omega = 1.2$, para resolver los sistemas lineales del ejercicio 1. Usar $TOL = 10^{-2}$ y el número máximo de iteraciones $N = 25$.