

1.II.3. Sistemas de ecuaciones lineales: métodos iterativos.

Manuel Palacios

Departamento de Matemática Aplicada

Centro Politécnico Superior

Universidad de Zaragoza

Otoño 2003

Referencias

- [1] Burden, R. L. and Faires, J. D.: Análisis Numérico. *Grupo Editorial Iberoamerica*, 1985.
- [2] Gasca, M.: Cálculo numérico: resolución de ecuaciones y sistemas. *Librería Central*, 1987.
- [3] Hairer, E.: Introduction à l'Analyse Numérique. *Université de Genève, Dept. de Mathématiques*, 1993.
- [4] Conde, C. y Winter, G.: Métodos y algoritmos del álgebra numérica. *Editorial Reverté*, 1990.
- [5] Griffel, D. H.: Linear Algebra and its applications. *Ellis Horwood*, 1989.
- [6] Strang, G.: Linear Algebra and its Applications. 3th ed. *Hardcourt Brace Jovanovich, Inc.*, 1988.

1 Métodos iterativos lineales

Consideraremos un sistema lineal

$$Ax = b, \quad (1)$$

siendo A una matriz inversible, es decir, el sistema es compatible determinado.

Un método iterativo para la resolución de sistemas lineales permite construir una sucesión $\{x^{(m)}\}_{m=0}^{\infty}$ de vectores que “aproximan” la solución exacta $x = A^{-1}b$ del sistema en el sentido de que

$$\lim_{m \rightarrow \infty} x^{(m)} = x$$

Los métodos iterativos más usados son los que construyen la solución en la forma

$$x^{(m+1)} = Bx^{(m)} + c, \quad B \in M_{\mathbb{R}}(n), \quad c \in \mathbb{R}^n, \quad (2)$$

a partir de $x^{(0)} \in \mathbb{R}^n$ arbitrario.

Definición 1.1 *Un método (2) se dice convergente, si para cualquier $x^{(0)} \in \mathbb{R}^n$, se verifica:*

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

Una condición mínima que se ha de cumplir es que, en el límite, la ecuación (2) se transforme en

$$x = Bx + c, \quad (3)$$

siendo $x = A^{-1}b$ la solución exacta, para lo cual se ha de cumplir equivalentemente que

$$c = (I - B)A^{-1}b \quad (4)$$

Definición 1.2 *Un método (2) se dice consistente si se cumple la condición (3) o, equivalentemente, la (4)*

Teorema 1.3 *Un método (2) es convergente para el sistema (1) si y solo se verifican las dos condiciones siguientes:*

- i) el método es consistente con el sistema
- ii) $\rho(B) < 1$

Demostr.: $\boxed{\Rightarrow}$ Si el método es convergente, pasando al límite en (2) resulta la consistencia.

Por otro lado, de (2) y (3), resulta:

$$x^{(m+1)} - x = B(x^{(m)} - x) = B^2(x^{(m-1)} - x) = \dots = B^{m+1}(x^{(0)} - x)$$

Como esta diferencia debe tender a 0 por la convergencia, se debe cumplir

$$\lim_{m \rightarrow \infty} B^m v = 0, \quad \forall v \in \mathbb{R}^n,$$

lo cual, como ya sabemos, se cumple si $\rho(B) < 1$.

☞ Si se cumplen ii), debe existir alguna norma tal que $\|B\| < 1$. Consideremos una norma vectorial compatible con ella. Como más arriba, se tiene:

$$x^{(m)} - x = B^m (x^{(0)} - x)$$

Tomando normas resulta:

$$\|x^{(m)} - x\| = \|B\|^m \|x^{(0)} - x\|$$

Pasando al límite, se tiene la convergencia. ■

1.1 Construcción de métodos iterativos

Consideremos la descomposición

$$A = M - N, \quad M \text{ regular}$$

El sistema (1) se puede escribir en la forma

$$Mx = Nx + b, \quad \text{es decir, } x = M^{-1}Nx + M^{-1}b$$

lo que sugiere estudiar el método

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

Evidentemente, $B = M^{-1}N$ y $c = M^{-1}b = (I - B)A^{-1}b$, por lo que el método así construido es consistente con el sistema. Para que el método sea convergente es suficiente que para alguna norma se cumpla $\|M^{-1}N\| < 1$

Teorema 1.4 *Todo método iterativo (2) convergente es de la forma (5) con $A = M - N$ y M regular.*

Demostr.: Suponiendo conocidos A, b, B y c , vamos a calcular M y N . Como $\rho(B) < 1$, luego $I - B$ es regular, y $c = M^{-1}b = (I - B)A^{-1}b$, basta tomar

$$M = A(I - B)^{-1} \quad \text{y} \quad N = MB = A(I - B)^{-1}B,$$

que, evidentemente, verifican $M - N = A$. ■

Los métodos (5) suelen presentarse en la forma:

$$Mx^{(m+1)} = Nx^{(m)} + b, \quad (6)$$

ya que, habitualmente, la matriz M suele ser diagonal o triangular inferior, por lo que la resolución de este sistema es de bajo coste computacional.

1.2 Estabilidad

Para estudiar la estabilidad, hay que analizar la influencia de perturbaciones. Así que, supongamos que los cálculos se realizan cometiendo un pequeño error, es decir, en vez de

$$x^{(m)} = B x^{(m-1)} + c,$$

se calcula

$$\tilde{x}^{(m)} = B \tilde{x}^{(m-1)} + c + \rho^{(m)}$$

En consecuencia, se tendrá

$$\begin{aligned}\tilde{x}^{(m)} - x^{(m)} &= B (\tilde{x}^{(m-1)} - x^{(m-1)}) + \rho^{(m)} \\ \tilde{x}^{(0)} - x^{(0)} &= \rho^{(0)};\end{aligned}$$

de donde

$$\tilde{x}^{(m)} - x^{(m)} = B^m \rho^{(0)} + \sum_{j=1}^m B^{m-j} \rho^{(j)},$$

y tomando normas compatibles:

$$\begin{aligned}\|\tilde{x}^{(m)} - x^{(m)}\| &\leq \sum_{j=1}^m \|B\|^{m-j} \|\rho^{(j)}\| \leq \sup\{\|\rho^{(j)}\|\} \sum_{j=1}^m \|B\|^j \leq \\ &\leq \sup\{\|\rho^{(j)}\|\} \sum_{j=1}^{\infty} \|B\|^j = \frac{1}{1 - \|B\|} \sup\{\|\rho^{(j)}\|\}\end{aligned}$$

Por lo tanto, pequeños errores de redondeo producirán errores pequeños en el resultado en la iteración m y el método será estable, siempre que sea convergente.

Notemos que un valor de $\|B\| < 1$, pero próximo a 1 producirá una cota muy grande. Interesa que $\|B\|$ sea pequeña.

1.3 Velocidad de convergencia

Más arriba hemos obtenido

$$\|x^{(m)} - x\| = \|B^m\| \|x^{(0)} - x\|$$

Esto puede ser interpretado como que en cada una de las m primeras iteraciones el error se ha reducido en un factor de $\|B^m\|^{1/m}$ y, en consecuencia, se puede estimar que para que el error se reduzca en un factor de 10^{-t} se deben realizar N iteraciones, cumpliéndose:

$$(\|B^m\|^{1/m})^N \leq 10^{-t}, \quad \text{o bien} \quad N \geq \frac{t}{-\log_{10}(\|B^m\|^{1/m})}$$

El número

$$R(B^m) = -\log_{10} \|B^m\|^{1/m} \tag{7}$$

es denominado **velocidad media de convergencia** en m iteraciones.

Se puede demostrar que

$$\rho(B) = \lim_{m \rightarrow \infty} \|B^m\|^{1/m}$$

Al número $-\log_{10}(1/\rho(B))$ se le denomina **velocidad asintótica** de convergencia.

Ejercicio 1.5 Encontrar la velocidad de convergencia en 5 iteraciones del método iterativo que toma $M = D$ (diagonal principal de A) al aplicarlo a la resolución del sistema cuya matriz ampliada es la siguiente:

$$A = \begin{pmatrix} 800.2669 & 0.8648 & 0.8642 \\ 0.2161 & 800.1441 & 0.1440 \end{pmatrix}$$

Solución: La matriz de iteración B y su potencia serán:

$$B = M^{-1}N = \begin{pmatrix} 0 & -0.0010906 \\ -0.000270076 & 0 \end{pmatrix}, \quad B^5 = \begin{pmatrix} 0 & -9.203 * 10^{-17} \\ -2.30032 * 10^{-17} & 0 \end{pmatrix}$$

Así que

$$\|B^5\|_{\infty} = 9.203 * 10^{-17}$$

Y por lo tanto,

$$R(B^5) = 3.20721$$

El valor propio de B de módulo mayor es $\lambda = 0.000540226$, por lo que la velocidad asintótica será

$$\log_{10}(1/\rho(B)) = 3.26742$$

Como se puede observar ambas son muy cercanas. ■

1.4 Cota del error

Se puede demostrar el siguiente resultado

Teorema 1.6

$$\|x^{(m)} - x\| \leq \frac{\|B\|}{1 - \|B\|} \|x^{(m)} - x^{(m-1)}\|$$

Demotr.: Tomando normas en la siguiente expresión

$$x^{(m)} - x = x^{(m)} - x^{(m+1)} + x^{(m+1)} - x,$$

resulta

$$\|x^{(m)} - x\| \leq \|x^{(m)} - x^{(m+1)}\| + \|x^{(m+1)} - x\| \leq \|x^{(m+1)} - x^{(m)}\| + \|B\| \|x^{(m)} - x\|,$$

es decir,

$$(1 - \|B\|) \|x^{(m)} - x\| \leq \|x^{(m+1)} - x^{(m)}\| \leq \|B\| \|x^{(m)} - x^{(m-1)}\|,$$

de donde se deduce la desigualdad buscada. ■

Esta acotación suele utilizarse como test de parada del algoritmo.

2 Métodos iterativos más usuales

Consideraremos los métodos de Jacobi, Gauss-Seidel y relajación, que son los más habitualmente utilizados. Todos ellos se construyen mediante el proceso descrito en el apartado 1.1 de la sección anterior. Para su definición, descomponemos la matriz A en la forma

$$A = D - L - U,$$

donde D es la matriz diagonal principal de A , L es una matriz triangular inferior estricta cuyos elementos no nulos coinciden con los de A y U es una matriz triangular superior estricta cuyos elementos no nulos coinciden con los de A .

Entonces, el método de **Jacobi** se define considerando la matriz $M = D$, por lo que $N = L + U$. Por lo tanto, la matriz de iteración, que denotaremos J , será

$$J = M^{-1} N = D^{-1} (L + U) = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & & \dots & -\frac{a_{n-1n}}{a_{nn}} & 0 \end{pmatrix}$$

Ya que la matriz A es regular, caso de que algún a_{jj} fuese nulo, haría falta permutar previamente las filas.

El método de Jacobi se puede escribir en componentes en la forma:

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

El método de **Gauss-Seidel** se define considerando la matriz $M = D - L$, por lo que $N = U$. Por lo tanto, la matriz de iteración, que denotaremos G , será

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

aunque habitualmente se expresa en la forma (6), es decir,

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

por lo que se puede escribir en componentes en la forma:

$$\sum_{j=1}^i a_{ij} x_j^{(m+1)} = -\sum_{j=i+1}^n a_{ij} x_j^{(m)} + b_i, \quad i = 1, 2, \dots, n$$

o bien,

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

Ya que la matriz A es regular, caso de que algún a_{jj} fuese nulo, haría falta permutar previamente las filas.

Observemos que para determinar la aproximación $x^{(m+1)}$, en el método de Jacobi, se utilizan los $x_i^{(m)}$ de la iteración anterior, mientras que, en el método de Gauss-Seidel, se utilizan los últimos $x^{(m+1)}$ calculados. Por eso, en el primero se requiere mantener en memoria dos vectores, el $x^{(m+1)}$ y el $x^{(m)}$, mientras que en el de Gauss-Seidel solo se necesita uno de ellos.

Ejemplo 2.1 Resolver mediante los métodos de Jacobi y Gauss-Seidel el sistema $Ax = b$ siguiente:

$$\bar{A} = \begin{pmatrix} 10 & -1 & 2 & 0 & 6 \\ -1 & 11 & -1 & 3 & 25 \\ 2 & -1 & 10 & -1 & -11 \\ 0 & 3 & -1 & 8 & 15 \end{pmatrix}$$

Solución:

La matriz de iteración del método de Jacobi es:

$$J = D^{-1}(L + U) = \begin{pmatrix} 0 & 1/10 & -2/10 & 0 & 0 \\ 1/11 & 0 & 1/11 & -3/11 & 0 \\ -2/10 & 1/10 & 0 & 1/10 & 0 \\ 0 & -3/8 & 1/8 & 0 & 0 \end{pmatrix},$$

por lo que las ecuaciones del método son:

$$\begin{aligned} x_1^{(k+1)} &= 1/10 x_2^{(k)} - 2/10 x_3^{(k)} + 3/5 \\ x_2^{(k+1)} &= 1/11 x_1^{(k)} + 1/11 x_3^{(k)} - 3/11 x_4^{(k)} + 25/11 \\ x_3^{(k+1)} &= -2/10 x_1^{(k)} + 1/10 x_2^{(k)} + 1/10 x_4^{(k)} - 11/10 \\ x_4^{(k+1)} &= -3/8 x_2^{(k)} + 1/8 x_3^{(k)} + 15/8 \end{aligned}$$

Comenzando con el vector inicial $x_0 = (0, 0, 0, 0)$, se obtienen los resultados de la tabla adjunta:

0	1	2	3	9	10
0.0000	0.6000	1.0473	0.9326	0.9997	1.0001
0.0000	2.2727	1.7159	2.0533	2.0004	1.9998
0.0000	-1.1000	-0.8052	-1.0493	-1.0004	-0.9998
0.0000	1.8750	0.8852	1.1309	1.0006	0.9998

Mientras que para el método de Gauss-Seidel,

$$G = (D - L)^{-1}U = \begin{pmatrix} 0 & 1/10 & -2/10 & 0 \\ 0 & 1/110 & 4/55 & -3/11 \\ 0 & -21/1100 & 13/275 & 4/55 \\ 0 & -51/8800 & -47/2200 & 49/440 \end{pmatrix},$$

y las ecuaciones del método se escriben mejor en la forma

$$\begin{aligned} x_1^{(k+1)} &= && 1/10 x_2^{(k)} & -2/10 x_3^{(k)} & & +3/5 \\ x_2^{(k+1)} &= 1/11 x_1^{(k+1)} && & +1/11 x_3^{(k)} & -3/11 x_4^{(k)} & +25/11 \\ x_3^{(k+1)} &= -2/10 x_1^{(k+1)} & +1/10 x_2^{(k+1)} && & +1/10 x_4^{(k)} & -11/10 \\ x_4^{(k+1)} &= && -3/8 x_2^{(k+1)} & +1/8 x_3^{(k+1)} & & +15/8 \end{aligned}$$

Comenzando con el mismo vector inicial, se obtienen los resultados de la tabla adjunta:

0	1	2	3	4	5
0.0000	0.6000	1.0300	1.0065	1.0009	1.0001
0.0000	2.3272	2.037	2.0036	2.0003	2.0000
0.0000	-0.9873	-1.014	-1.0025	-1.0003	-1.0000
0.0000	0.8789	0.9844	0.9983	0.9999	1.0000

En este ejemplo, el método de Gauss-Seidel converge más rápidamente que el de Jacobi, aunque esto no es así en general. ■

La familia de **métodos de relajación** (abreviadamente, SOR) se consigue tomando $M = \frac{1}{\omega} D - L$, resultando $N = \frac{1-\omega}{\omega} D + U$, siendo $\omega \neq 0$. La matriz de iteración, que denotaremos G_ω , es

$$G_\omega = \left(\frac{1}{\omega} D - L\right)^{-1} \left(\frac{1-\omega}{\omega} D + U\right)$$

Observemos que esta familia se reduce al método de Gauss-Seidel en el caso particular $\omega = 1$.

En componentes, se puede expresar en la forma:

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

o bien,

$$x_i^{(m+1)} = \omega \tilde{x}_i^{(m+1)} + (1-\omega) x_i^{(m)}, \quad i = 1, 2, \dots, n, \quad (8)$$

siendo

$$\tilde{x}_i^{(m+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)} \right), \quad i = 1, 2, \dots, n,$$

es decir, cada componente de la iteración $(m+1)$ -ésima se obtiene como cierto promedio entre la solución del método de Gauss-Seidel y la solución de la iteración anterior.

3 Resultados de convergencia

3.1 Matrices estrictamente diagonal dominantes

Teorema 3.1 Para esta matrices, se cumple que los métodos de Gauss-Seidel y de Jacobi son convergentes y $\rho(J) \leq K$, y $\rho(G) \leq K$, siendo

$$K = \max\left\{ \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} \right\} < 1$$

Demostr.: Consultar [2]. ■

En el ejemplo (2.1) se tiene:

$$\rho(J) = 0.4264, \quad \rho(G) = 0.0898 \text{ y } K = 0.5$$

3.2 Matrices simétricas definidas positivas

Teorema 3.2 Si A es simétrica y $a_{ii} > 0$, $\forall i$, entonces, el método de Gauss-Seidel converge $\iff A$ definida positiva

Demostr.: Consultar [2]. ■

Teorema 3.3 Si A es simétrica y $D + L + U$ es definida positiva, entonces, el método de Jacobi converge $\iff A$ definida positiva

Demostr.: Consultar [2]. ■

3.3 Matrices tridiagonales

Teorema 3.4 En este caso, los radios espectrales de las matrices de iteración de los métodos de Jacobi y Gauss-Seidel verifican: $\rho(G) = \rho(J)^2$ y, por tanto, ambos convergen o ambos divergen. En el caso convergente, el de Gauss-Seidel es, asintóticamente, más rápido que el de Jacobi.

Demostr.: Consultar [2]. ■

3.4 Convergencia de los métodos de relajación

Teorema 3.5 (Kahan) Si $a_{ii} \neq 0$, $\forall i$, $\rho(G_\omega) \geq |\omega - 1|$

Demostr.: Basta observar que

$$\det G_\omega = \det D^{-1} (1 - \omega)^n \det D = (1 - \omega)^n = \prod \lambda_i \Rightarrow \rho(G_\omega) \geq |\omega - 1|.$$

Teorema 3.6 El método SOR solo puede converge si $0 < \omega < 2$

Demotr.: Es consecuencia inmediata del teorema anterior. ■

Teorema 3.7 *Si A es estrictamente diagonal dominante y $0 < \omega \leq 1 \implies$ SOR converge*

Demotr.: Basta calcular la norma de Chebyshev del vector $y = Gu$, con $\|u\| = 1$. Teniendo en cuenta (8), resulta

$$\|y\|_\infty = \|Gu\|_\infty \leq \omega \|G\|_\infty + (1 - \omega) < 1, \text{ si } 0 < \omega \leq 1,$$

luego

$$\|G\|_\infty = \max_{\|u\|_\infty=1} \|Gu\|_\infty < 1. \blacksquare$$

Teorema 3.8 (*Ostrowski-Reich*)

Si A es simétrica, real, con $a_{ii} > 0, \forall i$, entonces, SOR converge \iff A definida positiva y $0 < \omega < 2$

Teorema 3.9 *Si A es tridiagonal simétrica definida positiva, entonces: los métodos de Jacobi, Gauss-Seidel y SOR convergen y el valor óptimo del parámetro ω es*

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(B_J)^2}}$$