

Métodos Iterativos para Sistemas de Ecuaciones

MA-1006

UCR

Temas de sección

- a) Normas, condición y estabilidad.
- b) Método de Jacobi, método de Gauss-Seidel, métodos SOR.
- c) Métodos para sistema de ecuaciones no lineales.

¿Cuál es la idea?

¿Cuál es la idea de esta sección?

Definición

Una norma en \mathbb{R}^n es una función $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ que satisface las siguientes propiedades

- a) $\|\mathbf{x}\| \geq 0$, para todo $\mathbf{x} \in \mathbb{R}^n$.
- b) $\|\mathbf{x}\| = 0$ si y solo si $\mathbf{x} = 0$.
- c) $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$, para todo $\alpha \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$.
- d) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

Definición (Normas de vectores)

Considere el espacio vectorial \mathbb{R}^n (el espacio de vectores columna). Sea \mathbf{x} un vector $\mathbf{x} \in \mathbb{R}^n$. Se definen las siguientes normas:

- $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$
- $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$
- $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$

Ejercicio (Normas)

Sea $\mathbf{x} = (-1, 1, -2)^t$. Determine $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$ y $\|\mathbf{x}\|_\infty$.

Solución:

$$\|\mathbf{x}\|_1 = |-1| + |1| + |-2| = 4$$

$$\|\mathbf{x}\|_2 = (|-1|^2 + |1|^2 + |-2|^2)^{1/2} = \sqrt{6}$$

$$\|\mathbf{x}\|_\infty = |-2| = 2$$

Definición (Normas de matrices)

Una norma (matricial) es una función $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ que satisface las siguientes propiedades para $A, B \in \mathbb{R}^{m \times n}$:

- a) $\|A\| \geq 0$.
- b) $\|A\| = 0$ si y solo si $A = 0$.
- c) $\|\alpha A\| = |\alpha| \|A\|$, para todo $\alpha \in \mathbb{R}$.
- d) $\|A + B\| \leq \|A\| + \|B\|$.
- e) $\|AB\| \leq \|A\| \|B\|$ (consistente).

Teorema

Si $\|\cdot\|$ es una norma en \mathbb{R}^n , entonces

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

es una norma matricial.

Comentario: a una norma matricial definida a partir de una norma de vectores se le denomina norma *inducida* o norma *natural*.

Normas de matrices

Sea $A \in \mathbb{R}^{n \times n}$. Tenemos los siguientes resultados:

- a) Sean $\{a_j\}_{j=1}^n \in \mathbb{R}^n$ los vectores **columna** de A .

$$\|A\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1.$$

Esto es, la 1-norma de una matriz corresponde a la máxima suma de los valores absolutos de las entradas de cada columna.

- b) Si denotamos a_i^* a las n **filas** de A , entonces:

$$\|A\|_\infty = \max_{1 \leq i \leq n} \|a_i^*\|_1.$$

Esto es, la ∞ -norma de una matriz corresponde al máximo de la suma de los valores absolutos de cada fila.

Ejercicio (Normas de matrices)

Determine $\|A\|_1$ y $\|A\|_\infty$ donde

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{pmatrix}$$

Solución:

$$\|A\|_1 = 6$$

$$\|A\|_\infty = 7$$

Ejercicio (Normas de matrices en Matlab)

El comando *norm* de *Matlab* calcula la 2-norma de un vector (o matriz). Además,

- a) *norm(v,p)* devuelve la *p*-norma del vector (o matriz).
- b) Note que si $p = 2$, entonces $\text{norm}(v) = \text{norm}(v, p)$.
- c) En el comando *norm(v,p)*, *p* también puede tomar el valor *Inf* (norma infinito).

Verifique que para

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{pmatrix}$$

Se tiene que $\|A\|_1 = 6$ y $\|A\|_\infty = 7$.

Definición (Radio espectral)

Sean $\lambda_1, \dots, \lambda_n$ valores propios de una matriz A . Entonces su radio espectral $\rho(A)$ se define como

$$\rho(A) = \max_{i=1, \dots, n} \{|\lambda_i|\}$$

Teorema

Sea $\|\cdot\|$ una norma matricial en $\mathbb{R}^{n \times n}$. Entonces para todo $A \in \mathbb{R}^{n \times n}$ se tiene

- $\|A^k\| \leq \|A\|^k$
- $\rho(A) \leq \|A\|$

Número de condición

Definición

El número de condición de una matriz A está dado por

$$\kappa(A) = \|A\| \|A^{-1}\|$$

Se dice que la matriz A es mal condicionada si $\kappa(A) \gg 1$. En caso contrario se dice que es bien condicionada. Además, el sistema $Ax = b$ es mal condicionado si la matriz A es mal condicionada.

Métodos: Jacobi y Gauss

Dado un sistema lineal de tamaño $n \times n$ dado por $A\mathbf{x} = \mathbf{b}$, esto es,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n = b_3 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Queremos generar un método iterativo para aproximar la solución \mathbf{x} a dicho sistema.

El método debe iniciar con una aproximación inicial $\mathbf{x}^{(0)}$ para la solución \mathbf{x} y debe generar una sucesión $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ que sea **convergente** a \mathbf{x} .

Método de Jacobi

El método de Jacobi se obtiene primeramente despejando cada variable x_i de la ecuación i de la forma

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

Note que aquí se requiere imponer la restricción $a_{ii} \neq 0$. Luego, para cada $k \geq 1$, se obtienen las componentes $x_i^{(k)}$ de $\mathbf{x}^{(k)}$ a partir de $\mathbf{x}^{(k-1)}$

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

Veamos esto con un ejemplo.

Método de Jacobi

Ejemplo

Considere el sistema

$$\begin{cases} 5x - 2y = -2 \\ 2x - 8y = 0 \end{cases} \quad (1)$$

La solución exacta a este sistema está dada por

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -4/9 \\ -1/9 \end{pmatrix} = \begin{pmatrix} -0.\bar{4} \\ -0.\bar{1} \end{pmatrix}.$$

Aplique el método de Jacobi para aproximar la solución utilizando como vector inicial $\mathbf{x}^{(0)} = (0, 0)^t$.

Solución:

Solución:

A partir del sistema, se despeja la primera variable de la primera ecuación y la segunda variable de la segunda ecuación:

$$\begin{aligned}x &= \frac{-2}{5} + \frac{2y}{5} &= 0x + \frac{2y}{5} - \frac{2}{5} \\y &= \frac{0}{(-8)} - \frac{2x}{(-8)} &= \frac{x}{4} + 0y + 0,\end{aligned}$$

este sistema se puede escribir como sigue:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 2/5 \\ 1/4 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -2/5 \\ 0 \end{pmatrix}. \quad (2)$$

Defina $T = \begin{pmatrix} 0 & 2/5 \\ 1/4 & 0 \end{pmatrix}$ y $\mathbf{c} = (-2/5, 0)^t$. La ecuación (2) sugiere que podemos utilizar la iteración

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{c} \quad (3)$$

para aproximar la solución al sistema (4).

- En la primera iteración, utilizamos $\mathbf{x}^{(0)} = (0, 0)^t$ en la ecuación (3) y obtenemos

$$\mathbf{x}^{(1)} = \begin{pmatrix} 0 & 2/5 \\ 1/4 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -2/5 \\ 0 \end{pmatrix} = \begin{pmatrix} -2/5 \\ 0 \end{pmatrix}.$$

- En la segunda iteración, utilizamos $\mathbf{x}^{(1)} = (-2/5, 0)^t$ en la ecuación (3) y obtenemos

$$\mathbf{x}^{(2)} = \begin{pmatrix} 0 & 2/5 \\ 1/4 & 0 \end{pmatrix} \begin{pmatrix} -2/5 \\ 0 \end{pmatrix} + \begin{pmatrix} -2/5 \\ 0 \end{pmatrix} = \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix} = \begin{pmatrix} -0.4 \\ -0.1 \end{pmatrix}.$$

- En la tercera iteración:

$$\mathbf{x}^{(3)} = \begin{pmatrix} 0 & 2/5 \\ 1/4 & 0 \end{pmatrix} \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix} + \begin{pmatrix} -2/5 \\ 0 \end{pmatrix} = \begin{pmatrix} -11/25 \\ -1/10 \end{pmatrix} = \begin{pmatrix} -0.44 \\ -0.1 \end{pmatrix}$$

De esta forma se puede continuar el procedimiento hasta que la aproximación sea lo suficientemente buena. Se puede utilizar como condición de parada utilizando una tolerancia ε de la siguiente forma

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_{\infty}}{\|\mathbf{x}^{(k)}\|_{\infty}} < \varepsilon.$$

Método de Jacobi

El método anterior también se puede escribir como sigue: podemos la matriz A en tres matrices D, L, U , donde la matriz D tiene las entradas diagonales de A , la matriz L es la parte (estrictamente) triangular inferior de A y la matriz U contiene la parte estrictamente triangular superior de A , esto es, $A = (D - L - U)$. Entonces, el procedimiento del ejemplo (4) es equivalente a

$$A\mathbf{x} = \mathbf{b}$$

$$(D - L - U)\mathbf{x} = \mathbf{b}$$

$$D\mathbf{x} = (L + U)\mathbf{x} + \mathbf{b}$$

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

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

entonces, la iteración de Jacobi puede ser escrita de la forma

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

Ejercicio

Para el sistema

$$\begin{cases} 5x - 2y = -2 \\ 2x - 8y = 0 \end{cases} \quad (4)$$

Determine T_J y c_J .

Ejercicio

Encuentre las primeras dos iteraciones del método de Jacobi para el sistema

$$\begin{cases} 3x - y + z = 1 \\ 3x + 6y + 2z = 0 \\ 3x + 3y + 7z = 4 \end{cases} \quad (5)$$

utilizando el vector inicial $\mathbf{x}^{(0)} = (0, 0, 0)^t$.

Método de Gauss-Seidel

El método se basa en la igualdad

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

En forma matricial, podemos descomponer $A = D - L - U$ y realizar un despeje

$$(D - L)\mathbf{x} = U\mathbf{x} + \mathbf{b}, \quad \text{'hacia adelante'}$$

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

La ecuación anterior sugiere plantear la iteración

$$\mathbf{x}^{(k)} = (D - L)^{-1}U\mathbf{x}^{(k-1)} + (D - L)^{-1}\mathbf{b}. \quad (6)$$

Si definimos $T_g = (D - L)^{-1}U$ y $\mathbf{c}_g = (D - L)^{-1}\mathbf{b}$, la ecuación (6) también se puede escribir de la forma $\mathbf{x}^{(k)} = T_g \mathbf{x}^{(k-1)} + \mathbf{c}_g$.

Ejemplo

Aproximar la solución al sistema

$$\begin{cases} 5x - 2y = -2 \\ 2x - 8y = 0 \end{cases} \quad (7)$$

utilizando el vector inicial $\mathbf{x}^{(0)} = (0, 0)^t$.

Solución:

Ahora vamos a utilizar el método de Gauss-Seidel. Separamos la matriz A de la siguiente manera

$$A = \begin{pmatrix} 5 & -2 \\ 2 & -8 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -8 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = D - L - U,$$

luego debemos encontrar explícitamente la matriz $D - L$ y $(D - L)^{-1}$

$$(D - L) = \begin{pmatrix} 5 & 0 \\ 2 & -8 \end{pmatrix} \quad \text{y} \quad (D - L)^{-1} = \frac{1}{40} \begin{pmatrix} 8 & 0 \\ 2 & -5 \end{pmatrix}.$$

De lo anterior tenemos que

$$T_g = (D - L)^{-1}U = \frac{1}{40} \begin{pmatrix} 8 & 0 \\ 2 & -5 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2/5 \\ 0 & 1/10 \end{pmatrix}$$

$$\mathbf{c} = (D - L)^{-1}\mathbf{b} = \frac{1}{40} \begin{pmatrix} 8 & 0 \\ 2 & -5 \end{pmatrix} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix}$$

- En la primera iteración, utilizamos $\mathbf{x}^{(0)} = (0, 0)^t$ y obtenemos

$$\mathbf{x}^{(1)} = \begin{pmatrix} 0 & 2/5 \\ 0 & 1/10 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix} = \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix}.$$

- En la segunda iteración, utilizamos $\mathbf{x}^{(1)} = (-2/5, -1/10)^t$ en la ecuación (3) y obtenemos

$$\begin{aligned} \mathbf{x}^{(2)} &= \begin{pmatrix} 0 & 2/5 \\ 0 & 1/10 \end{pmatrix} \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix} + \begin{pmatrix} -2/5 \\ -1/10 \end{pmatrix} \\ &= \begin{pmatrix} -11/25 \\ -11/100 \end{pmatrix} = \begin{pmatrix} -0.44 \\ -0.11 \end{pmatrix} \end{aligned}$$

Ejercicio

Determine las matrices T_g y \mathbf{c}_g del método de Gauss-Seidel para el sistema

$$\begin{cases} 3x - y + z = 1 \\ 3x + 6y + 2z = 0 \\ 3x + 3y + 7z = 4 \end{cases} \quad (8)$$

Convergencia

Teorema

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

$$\mathbf{x}^{(k)} = T\mathbf{x}^{(k-1)} + \mathbf{c}, \quad \text{para } k \geq 1$$

converge a la solución única de $\mathbf{x} = T\mathbf{x} + \mathbf{c}$ si y solo si $\rho(T) < 1$.

Convergencia

Definición

Una matriz $A \in \mathbb{R}^{n \times n}$ se llama estrictamente diagonal dominante por filas si

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

Otros resultados de convergencia son los siguientes:

- 1 Si A es estrictamente diagonal dominante, entonces, las sucesiones obtenidas con el método de Jacobi y Gauss-Seidel convergen.
- 2 Si A es simétrica y definida positiva, entonces el método de Gauss-Seidel converge.
- 3 Si tanto A como $2D - A$ son simétricas y definidas positivas, entonces el método de Jacobi converge.

Método de Sobrerrelajación (SOR)

Los métodos de sobrerrelajación (de Gauss-Seidel) son utilizados para acelerar convergencia. En resumen, los métodos utilizan un peso $0 < \omega < 1$ para *promediar* el valor de $x_i^{(k+1)}$ en cada iteración.

La iteración

$$\left\{ \begin{array}{l} \mathbf{x}^{(0)}, \omega \in \mathbb{R} - \{0\} \text{ dados,} \\ \text{para } k = 0, 1, 2, \dots, \text{ hacer:} \\ \hat{x}_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right) \\ x_i^{(k+1)} = (1 - \omega) x_i^{(k)} + \omega \hat{x}_i^{(k+1)}, \quad \forall i = 1, 2, \dots, n \end{array} \right.$$

se conoce como **iteración de sobrerrelajación hacia adelante**.

La iteración

$$\left\{ \begin{array}{l} \mathbf{x}^{(0)}, \omega \in \mathbb{R} - \{0\} \text{ dados,} \\ \text{para } k = 0, 1, 2, \dots, \text{ hacer:} \\ \hat{x}_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^n a_{ij} x_j^{(k+1)} \right) \\ x_i^{(k+1)} = (1 - \omega) x_i^{(k)} + \omega \hat{x}_i^{(k+1)}, \quad \forall i = n, n-1, \dots, 1 \end{array} \right.$$

se conoce como **iteración de sobrerrelajación hacia atrás**.

Teorema

Si $A \in \mathbb{R}^{n \times n}$ es simétrica y definida positiva y $\omega \in]0, 2[$, entonces la iteración SOR converge hacia la solución de $A\mathbf{x} = \mathbf{b}$, para todo $\mathbf{x}^{(0)}, \mathbf{b} \in \mathbb{R}^n$.

Sistemas no lineales: Método de Newton

Considere el sistema de ecuaciones (no lineales)

$$\begin{aligned}x_1^2 - x_2^2 &= -2x_2 \\ 2x_1 + x_2^2 &= 6\end{aligned}$$

Dos soluciones (aproximadas) a este sistema son
(0.625204094, 2.179355825) y (2.109511920, -1.334532188).

Este sistema se puede expresar también de la forma

$$F(\mathbf{x}) = \begin{cases} f_1(x_1, x_2) &= x_1^2 - x_2^2 + 2x_2 = 0 \\ f_2(x_1, x_2) &= 2x_1 + x_2^2 - 6 = 0 \end{cases}$$

Nos referimos al sistema de ecuaciones como $F(\mathbf{x}) = \mathbf{0}$, donde $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Podemos aproximar la solución de este sistema con el método de Newton multivariable.

Definición

Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. La recursión definida por

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - [J_f(\mathbf{x}^{(k)})]^{-1} F(\mathbf{x}^{(k)}), \quad k = 0, 1, 2, \dots$$

donde $\mathbf{x}^{(0)} \in \mathbb{R}^n$ es el método de Newton para el sistema $F(\mathbf{x}) = 0$.

Note que se asume que $J_f(\mathbf{x}^{(k)})$ existe y es no singular para cada $k = 0, 1, 2, \dots$. La matriz jacobiana en el punto \mathbf{x} está dada por

$$J_f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_n(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

Teorema

Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ y suponga que

- 1 $F(\mathbf{c}) = 0$, F está definida y es continua en un vecindario de \mathbf{c} ,
- 2 las derivadas parciales de F están definidas y son continuas,
- 3 la matriz jacobiana $J_f(\mathbf{c})$ es no singular,

entonces la sucesión $\mathbf{x}^{(k)}$ definida por el método de Newton converge a la solución \mathbf{c} para $\mathbf{x}^{(0)}$ suficientemente cerca de \mathbf{c} y la convergencia es cuadrática.

Ejercicio

Considere el sistema

$$x_1^2 - x_2^2 = -2x_2$$

$$2x_1 + x_2^2 = 6$$

Utilice el método de Newton con $\mathbf{x}^{(0)} = (0, 0)^t$ para calcular $\mathbf{x}^{(1)}$ y $\mathbf{x}^{(2)}$.

% Ejemplo: Newton en dos variables: se necesita F, JF, v (aprox. i

clearvars

F = @(x1,x2) [x1^2 - x2^2 + 2*x2; 2*x1 + x2^2 - 6];

J = @(x1,x2) [2*x1 -2*x2 + 2; 2 2*x2]; % Jacobiano

v = [1; 1]; % [-4;-5]; % Condicion Inicial

tol = 10^-6;

iterMax = 10;

error = 1; % error en la aprox.

k = 1; % contador

while (error > tol) && (k <= iterMax)

vaux = v - J(v(1),v(2))\F(v(1),v(2));

error = norm(vaux-v, Inf); % norma infinito

v = vaux;

if (k==iterMax)

fprintf("Máximo de iteraciones excedido %3d \n", k)

end

k= k+1;

end

disp(v)

disp(k)