

Métodos iterativos para sistemas lineales

José Luis Bravo

Índice

- 1 Métodos iterativos para sistemas lineales
 - Método del punto fijo
 - Métodos de Jacobi y Gauss-Seidel
 - Criterio de parada

- 2 Autovalores y autovectores
 - Métodos para calcular el valor propio dominante
 - Métodos basados en transformaciones matriciales
 - Método QR de Francis-Kublanovskaya
 - Factorización QR Householder

Métodos Iterativos

Los métodos directos que hemos estudiado tienen inconvenientes si se aplican a sistemas de ecuaciones de grandes dimensiones, porque requieren muchas operaciones y son sensibles a errores de redondeo.

Los métodos iterativos están especialmente indicados en la resolución de este tipo de sistemas, o en aquellos en las que las matrices son dispersas (poseen muchos ceros).

Convergencia de matrices

Una sucesión de matrices $\{A_r\}_{r=1,2,\dots}$, $A_r \in \mathcal{M}_n$ se dice que converge a la matriz $A \in \mathcal{M}_n$ si para una (cualquier) norma matricial

$$\lim_{r \rightarrow \infty} \|A_r - A\| = 0.$$

Proposición

Sea $A \in \mathcal{M}_n$. Entonces se verifica

- 1 $\rho(A) \leq \|A\|$ para toda norma matricial.
- 2 Para todo $\epsilon > 0$, existe una norma matricial inducida $\|\cdot\|$ tal que

$$\|A\| \leq \rho(A) + \epsilon.$$

Sucesiones geométricas

Teorema

Sea $A \in \mathcal{M}_n$. Son equivalentes:

- 1 $\lim_{k \rightarrow \infty} A^k = 0$
- 2 $\lim_{k \rightarrow \infty} A^k v = 0$, para todo $v \in \mathbb{C}^n$.
- 3 $\rho(A) < 1$.
- 4 Existe una norma matricial (inducida) tal que $\|A\| < 1$.

Teorema

Sea $A \in \mathcal{M}_n$ y $\|\cdot\|$ una norma matricial. Entonces

$$\lim_{k \rightarrow \infty} \|A^k\|^{1/k} = \rho(A).$$

Teorema del punto fijo

Teorema (Teorema del punto fijo de Banach)

Sea $T: \mathbb{K}^n \rightarrow \mathbb{K}^n$ una aplicación contractiva (es decir, existe $K < 1$ tal que $\|T(y) - T(x)\| \leq K\|y - x\|$ para todo $x, y \in \mathbb{K}^n$). Entonces existe un único punto fijo de T , $z \in \mathbb{K}^n$.

Es más, para cualquier $x_0 \in \mathbb{K}^n$, sea $\{x_k\}$ la sucesión definida por

$$x_{k+1} = T(x_k), \quad n > 0.$$

Entonces se verifica

$$\lim_{k \rightarrow \infty} x_k = z.$$

Método del punto fijo para sistemas lineales

Consideremos el sistema $Ax = b$. Vamos a transformarlo para aplicar el método del punto fijo. Para ello descomponemos $A = M - N$, donde M es una matriz invertible. Entonces x es solución de

$$Mx = Nx + b.$$

Tomamos un vector inicial x_0 y definimos la sucesión

$$Mx_{n+1} = Nx_n + b.$$

Si la sucesión converge, converge a la solución de $Ax = b$.

Método del punto fijo para sistemas lineales

Dado un sistema $Ax = b$ y una descomposición $A = M - N$, definimos la aplicación afín

$$G(x) = M^{-1}(Nx + b).$$

Como

$$\|G(y) - G(x)\| \leq \|M^{-1}N\| \|y - x\|, \quad \text{para todo } x, y \in \mathbb{K}^n,$$

si $\|M^{-1}N\| < 1$, entonces la función G es contractiva.

Convergencia del método del punto fijo

Teorema

Sean $A, M, N \in \mathcal{M}_n$ tales que M es invertible y $A = M - N$. Dado $x_0 \in \mathbb{K}^n$, definimos la sucesión $x_{k+1} = G(x_k)$, donde $G(x) = M^{-1}(Nx + b)$.

Una condición necesaria y suficiente para que la sucesión anterior converja para todo valor inicial (a un punto fijo de G) es

$$\rho(M^{-1}N) < 1.$$

Método de Jacobi

El **método de iteración de Jacobi** consiste en definir M como una matriz diagonal con la misma diagonal que A y

$$\mathbf{N} = M - A$$

Así, la sucesión se construye partiendo de un valor inicial \mathbf{x}_0 y definiendo

$$\mathbf{M}\mathbf{x}_{k+1} = \mathbf{N}\mathbf{x}_k + \mathbf{b}, \quad k \geq 0$$

Nótese que en este caso, M^{-1} es una matriz diagonal tal que el i -ésimo elemento de su diagonal es a_{ii}^{-1} , $1 \leq i \leq n$.

Método de Gauss-Seidel

El **método de iteración de Gauss-Seidel** se obtiene al tomar M como una matriz triangular inferior cuyos elementos no nulos coinciden con los de A , es decir, si $A = (a_{ij})$ y $M = (m_{ij})$, entonces $m_{ij} = a_{ij}$ si $i \geq j$ y $m_{ij} = 0$ si $i < j$.

La matriz $N = (n_{ij}) = M - A$ verifica $n_{ij} = -a_{ij}$ si $i < j$ y $n_{ij} = 0$ si $i \geq j$.

La sucesión se construye partiendo de un valor inicial \mathbf{x}_0 y definiendo

$$\mathbf{M}\mathbf{x}_{k+1} = N\mathbf{x}_k + \mathbf{b}, \quad k \geq 0$$

Ejemplo de Métodos iterativos

Ejemplo: Dado el sistema de ecuaciones

$$\begin{aligned} 2x - y &= 9 \\ x + 6y - 2z &= 15 \\ 4x - 3y + 8z &= 1 \end{aligned}, \text{ con } \mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 6 & -2 \\ 4 & -3 & 8 \end{pmatrix}$$

tenemos:

$$\mathbf{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 4 & -3 & 0 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Partimos de un valor inicial

$$\mathbf{x}_0 = (0, 0, 0)$$

Ejemplo de Métodos iterativos

El método de Jacobi consiste en definir

$$\mathbf{D}\mathbf{x}_{k+1} = -(\mathbf{L} + \mathbf{U})\mathbf{x}_k + \mathbf{b}, \quad k \geq 0$$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{pmatrix} = - \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -2 \\ 4 & -3 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \begin{pmatrix} 9 \\ 15 \\ 1 \end{pmatrix}$$

Por tanto,

$$\begin{aligned} 2x_{k+1} &= y_k + 9 \\ 6y_{k+1} &= -x_k + 2z_k + 15 \\ 8z_{k+1} &= -4x_k + 3y_k + 1 \end{aligned}$$

$$\begin{aligned} \text{Partiendo de } \mathbf{x}_0 = (0, 0, 0), \quad 2x_1 &= y_0 + 9 = 9 \\ 6y_1 &= -x_0 + 2z_0 + 15 = 15 \\ 8z_1 &= -4x_0 + 3y_0 + 1 = 1 \end{aligned}$$

Ejemplo de Métodos iterativos

Para el método de Jacobi, tenemos:

- $\mathbf{x}_0 = (0, 0, 0)$
- $\mathbf{x}_1 = (4.5, 2.5, 0.125)$
- $\mathbf{x}_2 = (5.75, 1.7916667, -1.1875)$
- $\mathbf{x}_3 = (5.3958333, 1.1458333, -2.078125)$
- $\mathbf{x}_4 = (5.07291667, 0.90798611, -2.14322917)$

Nótese que, en este caso, el sistema es resoluble, y la solución real es $(5, 1, -2)$, hacia la cual converge la sucesión creada por el método.

Ejemplo de Métodos iterativos

El método de Gauss-Seidel consiste en definir

$$(\mathbf{D} + \mathbf{L})\mathbf{x}_{k+1} = -\mathbf{U}\mathbf{x}_k + \mathbf{b}, \quad k \geq 0$$

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 6 & 0 \\ 4 & -3 & 8 \end{pmatrix} \begin{pmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{pmatrix} = - \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} + \begin{pmatrix} 9 \\ 15 \\ 1 \end{pmatrix}$$

Por tanto,

$$\begin{aligned} 2x_{k+1} &= y_k + 9 \\ x_{k+1} + 6y_{k+1} &= 2z_k + 15 \\ 4x_{k+1} - 3y_{k+1} + 8z_{k+1} &= 1 \end{aligned}$$

Partiendo de $\mathbf{x}_0 = (0, 0, 0)$,

$$\begin{aligned} 2x_1 &= y_0 + 9 \\ 6y_1 &= -x_1 + 2z_0 + 15 \\ 8z_1 &= -4x_1 + 3y_1 + 1 \end{aligned}$$

Ejemplo de Métodos iterativos

Para el método de Gauss-Seidel, tenemos:

- $\mathbf{x}_0 = (0, 0, 0)$
- $\mathbf{x}_1 = (4.5, 1.75, -1.46875)$
- $\mathbf{x}_2 = (5.375, 1.11458333, -2.14453125)$
- $\mathbf{x}_3 = (5.05729167, 0.94227430, -2.05029297)$
- $\mathbf{x}_4 = (4.97113715, 0.98804615, -1.99005127)$

Convergencia de los métodos iterativos

Teorema

Dado el sistema $\mathbf{Ax} = \mathbf{b}$, si \mathbf{A} es estrictamente diagonal dominante, entonces existe una única solución del sistema y los métodos iterativos de Jacobi y Gauss-Seidel producen una sucesión de vectores que converge a dicha solución.

Criterio de parada

Consideremos un método iterativo de la forma

$$Mx^{(k+1)} = Nx^{(k)} + b,$$

donde $x^{(0)}$ es el punto inicial escogido, $B = M^{-1}N$ la matriz del método y b un vector constante.

Denotemos $\delta_k = \|x^{(k)} - x^{(k-1)}\|$. Si $\epsilon^{(k+1)} = \|x - x^{(k+1)}\|$, entonces una estimación para dicho valor es

$$\epsilon^{(k+1)} \approx \frac{\delta_{k+1}^2}{\delta_k - \delta_{k+1}}.$$

Introducción

Sea $A \in \mathcal{M}_n$ una matriz. Recordemos que sus autovalores pueden ser calculados como las raíces del polinomio

$$p(\lambda) = \det(A - \lambda I),$$

y, dado un autovalor λ , podemos calcular su autovector asociado como la solución del sistema lineal homogéneo (singular)

$$(A - \lambda I)x = 0.$$

Sin embargo estos problemas (en especial el primero) están mal condicionados. Por ello, estudiaremos métodos para determinar directamente los autovalores.

Localización de los autovalores

Teorema (Teorema de Gershgorin)

Sea $A \in \mathcal{M}_n(\mathbb{C})$. La unión de todos los discos

$$K_i = \left\{ \mu \in \mathbb{C} : |\mu - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}| \right\}$$

contiene todos los autovalores de la matriz $A = (a_{ij})_{1 \leq i, j \leq n}$.

Corolario

Si la unión $M_1 = \cup_{j=1}^m K_j$ de m discos K_j , $j = 1, \dots, m$ y la unión M_2 de los discos restantes son disjuntas, entonces M_1 contiene exactamente m autovalores de A .

Método de la potencia

Sea $A \in \mathcal{M}_n$ y supongamos que sus autovalores, $\lambda_1, \dots, \lambda_n$ verifican

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

y que existen n autovectores linealmente independientes, u_1, u_2, \dots, u_n (tal que $Au_i = \lambda_i u_i$).

Sea $x_0 = \alpha_1 u_1 + \dots + \alpha_n u_n$, con $\alpha_1 > 0$ y consideremos la sucesión $\{x_k\}$ definida por

$$x_k = A^k x_0.$$

Entonces

$$x_k = \lambda_1^k \left(\alpha_1 u_1 + \left(\frac{\lambda_2}{\lambda_1} \right)^k \alpha_2 u_2 + \dots + \left(\frac{\lambda_n}{\lambda_1} \right)^k \alpha_n u_n \right).$$

Método de la potencia

Es decir, para k suficientemente grande, se verifica

$$x_{k+1} \approx \lambda_1 x_k.$$

Para obtener el valor de λ_1 , basta aplicar a ambos términos una función lineal $\phi: \mathbb{K}^n \rightarrow \mathbb{K}$, por ejemplo:

- 1 Método de Rayleigh

$$\lambda_1 \approx \frac{x_k^* x_{k+1}}{x_k^* x_k}.$$

- 2 Sea j la posición de la mayor coordenada en módulo de x_k , entonces

$$\lambda_1 \approx \frac{e_j^* x_{k+1}}{e_j^* x_k}.$$

Cálculo del autovector

El método de la potencia también permite calcular un vector propio asociado a λ_1 . Conocido λ_1 , se verifica que

$$\lim_{k \rightarrow \infty} \frac{x_k}{\lambda_1^k} = \alpha_1 u_1$$

Nota: Si el autovalor de módulo máximo no es único, existen modificaciones del método que permiten estimarlo. En particular, si λ_1 es un autovalor múltiple (y verifica que su módulo es estrictamente mayor que el del resto de autovalores), entonces el método de la potencia también estima λ_1 .

Método de la potencia inversa

El **método de la potencia inversa** permite calcular el autovalor de módulo mínimo. Sea $A \in \mathcal{M}_n$ invertible y diagonalizable, con autovalores $\lambda_1, \dots, \lambda_n$ tal que

$$|\lambda_1| \geq |\lambda_2| \geq \dots > |\lambda_n| (> 0).$$

Entonces los autovalores de A^{-1} son $\lambda_1^{-1}, \dots, \lambda_n^{-1}$ y verifican

$$|\lambda_1^{-1}| \leq |\lambda_2^{-1}| \leq \dots < |\lambda_n^{-1}|.$$

Por tanto, podemos aproximar λ_n^{-1} mediante el método de la potencia, aplicado a la matriz A^{-1}

Método de la potencia desplazado

Consideremos la matriz $A - \mu I$ que supondremos diagonalizable.

Si λ es un autovalor de A , entonces $\lambda - \mu$ es un autovalor de $A - \mu I$. El parámetro μ se denomina desplazamiento.

El desplazamiento permite calcular el autovalor principal cuando hay varios autovalores con el mismo módulo; basta desplazar en cualquier dirección tal que los autovalores resultantes ya no tengan el mismo módulo.

Método de la potencia (inversa) desplazado

Una variante del método anterior permite obtener el autovalor más próximo a un número complejo μ dado.

Consideremos la matriz $A - \mu I$ que supondremos diagonalizable. Sean $\bar{\lambda}_1, \dots, \bar{\lambda}_n$ los autovalores de $A - \mu I$ y supongamos que

$$|\bar{\lambda}_1| \geq |\bar{\lambda}_2| \geq \dots > |\bar{\lambda}_n| > 0.$$

Entonces los autovalores de $(A - \mu I)^{-1}$ son $\bar{\lambda}_1^{-1}, \dots, \bar{\lambda}_n^{-1}$ y verifican

$$|\bar{\lambda}_1^{-1}| \leq |\bar{\lambda}_2^{-1}| \leq \dots < |\bar{\lambda}_n^{-1}|.$$

Por tanto, podemos aproximar $(\lambda_n - \mu)^{-1}$ mediante el método de la potencia, aplicado a la matriz $(A - \mu I)^{-1}$.

Factorización QR

El Método QR de Francis-Kublanovskaya es un método iterativo que, bajo ciertas condiciones permite obtener todos los autovalores de una matriz.

Dada una matriz $A \in \mathcal{M}_n(\mathbb{C})$, supondremos que podemos factorizarla como

$$A = QR,$$

donde Q es una matriz unitaria ($Q^*Q = I$) y R es una matriz triangular superior.

Esta factorización se denomina **factorización QR**.

Método QR de Francis-Kublanovskaya

Sea $A_1 = A$ y Q_1, R_1 una factorización QR de A_1 . Definimos la matriz

$$A_2 = R_1 Q_1 = Q_1^t A_1 Q_1.$$

En particular A_1 y A_2 tienen los mismos autovalores. El método QR de Francis-Kublanovskaya consiste en ir calculando recursivamente

$$A_k = R_{k-1} Q_{k-1},$$

donde $Q_{k-1} R_{k-1} = A_{k-1}$, R_{k-1} es triangular superior y $Q_{k-1}^t Q_{k-1} = I$.

Teorema

Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que sus autovalores λ_i , $1 \leq i \leq n$ verifican

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|.$$

Entonces

$$\lim_{k \rightarrow \infty} A_k = \begin{pmatrix} \lambda_1 & c_{12} & \dots & c_{1n} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{n-1,n} \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Es más,

$$(A_k)_{i,i-1} = \mathcal{O} \left(\left| \frac{\lambda_i}{\lambda_{i-1}} \right|^k \right), \quad i = 2, \dots, n.$$

Si además A es simétrica, entonces $c_{ij} = 0$, $1 \leq i \neq j \leq n$.

Reflexión de Householder

Dado un vector v unitario (con norma euclídea) de \mathbb{R}^n , la **reflexión de Householder** es la matriz

$$P = I - 2vv^t.$$

Proposición

La reflexión de Householder es una matriz simétrica y ortogonal. Además

- 1 $PP = I_d$.
- 2 $\|Pc\|_2 = \|c\|_2$ para todo $c \in \mathbb{R}^n$.
- 3 *Geoméricamente, se corresponde con una reflexión respecto al hiperplano ortogonal a v .*

Reflexión de Householder

Proposición

Dado $c \in \mathbb{R}^n$, existe una reflexión tal que $Pc = \pm \|c\| e_1$.

Basta tomar

$$v = \frac{u}{\|u\|_2}, \quad u = c \pm \|c\|_2 e_1.$$

Factorización QR de Householder

Sea $A \in \mathcal{M}_n$ y sean c la primera columna. Definimos la matriz

$$Q^{(1)} = I - 2vv^t, \quad v = \frac{u}{\|u\|}, \quad u = c \pm \|c\|e_1.$$

Entonces $Q^{(1)}$ es una matriz unitaria tal que la primera columna de $R^{(1)} = Q^{(1)}A$ tiene ceros debajo de la diagonal.

Factorización QR de Householder

Definimos $Q^{(k)}$, $R^{(k)}$ por recursivamente. Sea \tilde{R} la submatriz de $R^{(k-1)}$ formada por las últimas $n - k + 1$ filas y columnas.

Procediendo como antes, existe una reflexión de Householder \tilde{Q} tal que $\tilde{Q}\tilde{R}$ tiene ceros debajo de la diagonal de la primera columna. Entonces definimos

$$Q^{(k)} = \begin{pmatrix} Id & 0 \\ 0 & \tilde{Q} \end{pmatrix}, \quad R^{(k)} = Q^{(k)}R^{(k-1)}.$$

Finalmente $Q = Q^{(1)}Q^{(2)} \dots Q^{(n-1)}$ y $R = Q^{(n-1)} \dots Q^{(2)}Q^{(1)}A$ producen la factorización QR buscada.

Ejemplo

Ejemplo. Obtener una factorización QR de

$$A = \begin{pmatrix} \frac{16}{25} & -\frac{14}{25} & -2 \\ -\frac{12}{25} & -\frac{52}{25} & -1 \\ -\frac{3}{5} & -\frac{3}{5} & -3 \end{pmatrix}.$$