Métodos directos de resolución de sistemas de ecuaciones lineales

José Luis Bravo

Índice

Introducción

El objetivo de este tema es obtener la solución del sistema de ecuaciones lineales

$$Ax = b, \quad A \in \mathcal{M}_{n\times n}, \quad b \in \mathbb{R}^n.$$

Nótese que existe una única solución si y sólo si A es no singular.

Además, dicha solución se puede obtener por la regla de Cramer.

El problema que trataremos es cómo aproximar dicha solución de modo eficaz.

Matriz transpuesta conjugada

Dada una matriz $A \in \mathcal{M}_{m\times n}$, $A = (a_{ij})_{i=1\ldots m}^{j=1\ldots n}$ definiremos su transpuesta conjugada como

$$A^* := (\overline{a}_{ji}) \in \mathcal{M}_{n\times m}$$

(La matriz transpuesta conjugada también se denota A^H.)

Sean $A, B \in \mathcal{M}_{m\times n}$, $C \in \mathcal{M}_{n\times p}$. Se verifica:

- $A^{**} = A$.
- $(A+B)^* = A^* + B^*$.
- $(\lambda A)^* = \overline{\lambda} A^*$.
- $(AC)^* = C^* A^*$.

Definiciones

Sean

$$A = (a_{ij})_{1 \le i,j \le n} \in \mathcal{M}_n, \quad I = (\delta_{ij})_{1 \le i,j \le n}.$$

Decimos que A es

- simétrica si real y $A = A^t$.
- ortogonal si real y $AA^t = A^t A = I$.
- hermítica si $A = A^*$.
- unitaria si $AA^* = A^* A = I$.
- normal si $AA^* = A^* A$.

Si A es invertible, entonces A^* es invertible y se verifica

$$(A^*)^{-1} = (A^{-1})^*.$$

Además, $\det A^* = \overline{\det A}$.

Definiciones

Atendiendo a la naturaleza de los elementos:

- A es diagonal si $a_{ij} = 0$, $i \neq j$, $1 \leq i, j \leq n$.
- A es triangular superior si $a_{ij} = 0$, $i > j$, $1 \leq i, j \leq n$.
- A es triangular inferior si $a_{ij} = 0$, $i < j$, $1 \leq i, j \leq n$.
- A es (estrictamente) diagonal dominante si

$$|a_{ii}| \underset{(>)}{\geq} \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad 1 \leq i \leq n.$$

Propiedades

Proposición

Toda matriz estrictamente diagonal dominante es invertible.

Partición de matrices

Dada una matriz $D \in \mathcal{M}_n$, diremos que

$$D = \begin{pmatrix} D_{11} & D_{12} & \cdots & D_{1k_D} \\ D_{21} & D_{22} & \cdots & D_{2k_D} \\ \vdots & \vdots & \ddots & \vdots \\ D_{m_D 1} & D_{m_D 2} & \cdots & D_{m_D k_D} \end{pmatrix},$$

es una partición de D si para cada $1 \leq i, j \leq n$, D_{ij} es una matriz con el mismo número de filas que $D_{i\bar{j}}$ para todo $\bar{j}$ y con el mismo número de columnas que $D_{\bar{i}j}$ para todo $\bar{i}$.

Partición de matrices

Supongamos que $A = (A_{ij})$, $B = (B_{ij})$, $C = (C_{ij})$ son particiones de $A, B, C \in \mathcal{M}_n$.

Teorema

Si cada producto $A_{is}B_{sj}$ se puede formar y

$$C_{ij} = \sum_s A_{is}B_{sj},$$

entonces $C = AB$.

Autovalores y autovectores

Sea $A \in \mathcal{M}_n$. Se llama

- Polinomio característico de A a $p(\lambda) = \det(A - \lambda I)$.
- Autovalores de A a las raíces del polinomio característico.
- Espectro de A, $\mathrm{Sp}(A)$, al conjunto de autovalores de A.
- Radio espectral de A, $\rho(A) := \max\{|\lambda| : \lambda \in \mathrm{Sp}(A)\}$.
- Autovector asociado a un autovalor $\lambda \in \mathrm{Sp}(A)$ a todo vector v que satisfaga $Av = \lambda v$.

Propiedades de los autovectores y autovalores

- A es invertible si y sólo si $0 \notin \mathrm{Sp}(A)$.
- Para todo autovalor λ, $\dim(\mathrm{Ker}(A - \lambda I)) > 0$ (todo autovalor tiene un autovector no nulo asociado).
- $\mathrm{Sp}(AB) = \mathrm{Sp}(BA)$. En particular, el espectro es invariante por cambio de base.
- Los autovectores asociados a un autovalor λ constituyen el subespacio vectorial $\mathrm{Ker}(A - \lambda I)$, de dimensión menor o igual que la multiplicidad de λ.
- Si λ_1, λ_2 son dos autovalores distintos de A, entonces

$$\mathrm{Ker}(A - \lambda_1 I) \cap \mathrm{Ker}(A - \lambda_2 I) = \{0\}.$$

Propiedades

Proposición

1. *Los autovalores de una matriz hermítica son siempre reales.*
2. *Los autovalores de una matriz unitaria tienen módulo* 1.

Corolario

Los autovalores de una matriz simétrica son reales.

Matrices semejantes

Se dice que las matrices $A, B \in \mathcal{M}_n$ son semejantes si existe $P \in \mathbb{M}_n$ invertible, tal que

$$B = P^{-1}AP.$$

A y B son semejantes si y sólo si representan la misma aplicación lineal en dos bases distintas.

Sean $A, B \in \mathbb{M}_n$ matrices semejantes. Se verifica:

- $\det A = \det B$.
- $\mathrm{Sp}(A) = \mathrm{Sp}(B)$.

Descomposición de Schur

Teorema (Descomposición de Schur)

Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. *Existe una matriz unitaria* U *tal que la matriz* U^*AU *es triangular superior. Además los elementos de la diagonal son los autovalores de* A.
2. A *es normal si y sólo si existe* U *(unitaria) tal que* U^*AU *es diagonal.*

Matriz definida positiva

Una matriz hermítica $A \in \mathcal{M}_n$ es definida positiva (resp. semidefinida positiva) si

$$v^* A v > 0, \;\; v \in \mathbb{K}^n \backslash \{0\} \quad (\text{resp. } \; v^* A v \geq 0, \;\; v \in \mathbb{K}^n \backslash \{0\}).$$

Proposición

Sea $A \in \mathcal{M}_n$ una matriz hermítica. Se verifica:

1. *A es definida positiva si y sólo si $\mathrm{Sp}(A) \subset \mathbb{R}^+$.*
2. *A es semidefinida positiva si y sólo si $\mathrm{Sp}(A) \subset \mathbb{R}^+ \cup \{0\}$.*

Proposición

Sea $A \in \mathcal{M}_n$. Entonces $A^* A$ es una matriz hermítica y semidefinida positiva.
Si además A es invertible, entonces $A^* A$ es definida positiva.

Norma matricial

Una norma (vectorial) sobre $\mathcal{M}_n$ es una aplicación,

$$\|\cdot\|: \mathcal{M}_n \to \mathbb{R}^+ \cup \{0\}, \quad A \to \|A\|,$$

que verifica:

1. $\|A\| = 0$ si y sólo si $A = 0$.
2. $\|A + B\| \leq \|A\| + \|B\|$, para todo $A, B \in \mathcal{M}_n$.
3. $\|\lambda A\| = |\lambda|\|A\|$, para todo $\lambda \in \mathbb{C}$, $A \in \mathcal{M}_n$.

Decimos que es una norma matricial si además verifica

4. $\|AB\| \leq \|A\|\|B\|$, para todo $A, B \in \mathcal{M}_n$.

Ejemplos

Sea $A \in \mathcal{M}_n$, $A = (a_{ij})_{1 \leq i,j \leq n}$. Podemos definir las siguientes normas:

- Norma no matricial: $\|A\| = \max_{1 \leq i,j \leq n} |a_{ij}|$.
- $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|$.
- $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$.
- $\|A\|_2 = \rho(A^* A)^{1/2}$.

Normas matriciales compatibles e inducidas

Sea $\|\cdot\|_v$ una norma vectorial en $\mathbb{K}^n$ y $\|\cdot\|_M$ una norma matricial en $\mathcal{M}_n$.

Decimos que $\|\cdot\|_M$ es una norma matricial compatible si

$$\|Av\|_v \leq \|A\|_M \|v\|_v, \quad \forall v \in \mathbb{K}^n,\ A \in \mathcal{M}_n.$$

Decimos que $\|\cdot\|_M$ es una norma matricial inducida si

$$\|A\|_M = \text{máx}_{v\neq 0} \frac{\|Av\|_v}{\|v\|_v} = \text{máx}_{\|v\|_v=1} \|Av\|_v, \quad \forall A \in \mathcal{M}_n.$$

Si $\|\cdot\|_M$ es una norma matricial inducida entonces $\|I\|_M = 1$.

Ejemplos

- $\|A\|_1$ inducida por $\|v\|_1 = \sum_{i=1}^n |v_i|$, $v \in \mathbb{K}^n$.
- $\|A\|_\infty$ inducida por $\|v\|_\infty = \text{máx}_{i=1}^n |v_i|$, $v \in \mathbb{K}^n$.
- $\|A\|_2$ inducida por $\|v\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2}$, $v \in \mathbb{K}^n$.
- La norma de Frobenius:

$$\|A\| = \left(\sum_{1 \leq i,j \leq n} |a_{ij}|^2 \right)^{1/2} = \operatorname{tr}(A^* A)^{1/2}$$

  no es una norma inducida.
- Toda norma vectorial es compatible con la matricial inducida, pero dada una norma matricial, existen infinitas normas compatibles con ella.

La norma inducida es una norma matricial

Proposición

Si $\|\cdot\|_v$ *es una norma vectorial sobre* $\mathbb{K}^n$*, entonces la norma inducida*

$$\|A\|_M = \max_{\|v\|_v=1} \|Av\|_v$$

es una norma matricial, compatible con $\|\cdot\|_v$*.*

Normas matriciales inducidas por las vectoriales

Proposición

1. *La norma inducida por* $\|v\|_1 = \sum_{i=1}^n |v_i|$, $v \in \mathbb{K}^n$, *es*
$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{ij}|.$$
2. *La norma inducida por* $\|v\|_\infty = \max_{1\le i\le n} |v_i|$, $v \in \mathbb{K}^n$, *es*
$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|.$$
3. *La norma inducida por* $\|v\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2}$, $v \in \mathbb{K}^n$, *es*
$$\|A\|_2 = \rho(A^* A)^{1/2}.$$

Error y error residual

Sea $\tilde{x}$ una solución aproximada de $Ax = b$. Definimos el *vector de error* como

$$x_\delta = \tilde{x} - x,$$

donde x es la solución exacta del problema.

Definimos el *vector de error residual* como

$$b_\delta = A\tilde{x} - b.$$

Nótese que $Ax_\delta = b_\delta$.

Fijada una norma, definimos los errores relativos asociados a los errores anteriores como

$$E = \frac{\|x_\delta\|}{\|x\|}, \quad R = \frac{\|b_\delta\|}{\|b\|}.$$

Condicionamiento

Consideremos una norma matricial inducida por una vectorial (denotaremos ambas como $\|\cdot\|$).

Sea $A \in \mathcal{M}_n$. Denominamos *condicionamiento* de A respecto a la norma $\|\cdot\|$ a

$$\text{cond}(A) := \|A\|\|A^{-1}\|.$$

Proposición

Sea $\|\cdot\|$ una norma matricial inducida y $A \in \mathcal{M}_n$ una matriz invertible. Se verifican las siguientes propiedades:

1. $\text{cond}(A) \geq 1$.
2. $\text{cond}(A) = \text{cond}(A^{-1})$.
3. $\text{cond}(\lambda A) = \text{cond}(A)$ *para todo* $\lambda \in \mathbb{R}\setminus\{0\}$.

Condicionamiento

Teorema

Sean $b, b_\delta \in \mathbb{R}^n$ *no idénticamente nulos. Denotemos* x *y* $x + x_\delta$ *a las soluciones respectivas de los sistemas lineales*

$$Ax = b \quad A(x + x_\delta) = b + b_\delta.$$

Entonces se verifica

$$\frac{1}{\operatorname{cond}(A)} \frac{\|b_\delta\|}{\|b\|} \leq \frac{\|x_\delta\|}{\|x\|} \leq \operatorname{cond}(A) \frac{\|b_\delta\|}{\|b\|}$$

Además, para toda matriz A *invertible, existen* $b, b_\delta \in \mathbb{R}^n$ *no idénticamente nulos tal que las desigualdades se alcanzan.*

Condicionamiento

Teorema

Sean $b, b_\delta \in \mathbb{R}^n$ no identicamente nulos, $A, A_\delta \in \mathcal{M}_n$ tales que A y $A + A_\delta$ son matrices invertibles y denotemos x y $x + x_\delta$ a las soluciones respectivas de

$$Ax = b, \quad (A + A_\delta)(x + x_\delta) = b + b_\delta.$$

Entonces se verifica

$$\frac{\|x_\delta\|}{\|x\|} \leq \frac{\operatorname{cond}(A)}{1 - \operatorname{cond}(A)\|A_\delta\|/\|A\|} \left(\frac{\|b_\delta\|}{\|b\|} + \frac{\|A_\delta\|}{\|A\|} \right)$$

Cota superior asintótica (notación de Landau)

Decimos que una sucesión $\{x_n\}$ es del orden menor o igual que otra sucesión $\{y_n\}$ si existen constantes C, N tales que $|x_n| < C|y_n|$ para $n > N$ y lo denotamos

$$\{x_n\} = \mathcal{O}\left(\{y_n\}\right).$$

Decimos que dos sucesiones $\{x_n\}, \{y_n\}$ son del mismo orden, y lo denotamos $\{x_n\} = \Theta\left(\{y_n\}\right)$ si

$$\{x_n\} = \mathcal{O}\left(\{y_n\}\right), \quad \{y_n\} = \mathcal{O}\left(\{x_n\}\right).$$

Decimos que el orden de $\{x_n\}$ es estrictamente menor que el de $\{y_n\}$ si $\lim_{n\to\infty} x_n/y_n = 0$ y lo denotamos $\{x_n\} = o\left(\{y_n\}\right)$.

Sistemas fáciles de resolver

Consideremos que tenemos el sistema de ecuaciones lineales

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

donde $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$, $\mathbf{b} \in \mathbb{K}^n$ son conocidos.
Para ciertas matrices $\mathbf{A}$ el sistema es fácil de resolver:

1. $\mathbf{A}$ diagonal.
2. $\mathbf{A}$ triangular superior.
3. $\mathbf{A}$ triangular inferior.
4. $\mathbf{A}$ ortogonal o unitaria.

Factorización de matrices

Supongamos que la matriz $\mathbf{A}$ factoriza como producto de varias matrices

$$\mathbf{A} = \mathbf{M_1}\mathbf{M_2}\ldots\mathbf{M_k},$$

de modo que las matrices $\mathbf{M_i}$ se correspondan con matrices de sistemas "fáciles de resolver".
Entonces, podemos resolver el sistema recursivamente:

$$\mathbf{M_1}\mathbf{y_1} = \mathbf{b}, \mathbf{M_2}\mathbf{y_2} = \mathbf{y_1}, \ldots \mathbf{M_k}\mathbf{y_k} = \mathbf{y_{k-1}},$$

y tendremos que la solución es $\mathbf{x} = \mathbf{y_k}$.
Por ejemplo, si $\mathbf{A} = \mathbf{M_1}\mathbf{M_2}$, basta resolver $\mathbf{M_1}\mathbf{y} = \mathbf{b}$ y $\mathbf{M_2}\mathbf{x} = \mathbf{y}$.

Factorización LU

Decimos que una matriz invertible **A** admite una factorización LU si se puede escribir en la forma:

$$\mathbf{A} = \mathbf{LU}$$

donde

- **L** es una matriz triangular inferior ($l_{ij} = 0$ si $i < j$).
- **U** es una matriz triangular superior ($u_{ij} = 0$ si $i > j$).

Factorización LU

Si conocemos una factorización LU de una matriz $\mathbf{A}$, $\mathbf{A} = \mathbf{LU}$, podemos resolver el sistema $\mathbf{Ax} = \mathbf{b}$ con el siguiente procedimiento:

- Resolvermos mediante sustitución progresiva el sistema

$$\mathbf{Ly} = \mathbf{b}.$$

- Obtenemos la solución $\mathbf{x}$ resolviendo por sustitución regresiva el sistema

$$\mathbf{Ux} = \mathbf{y}.$$

También es útil para calcular el determinante, inversas, etc.

Factorización LU

Obtener una factorización LU de una matriz $\mathbf{A}$ es equivalente a resolver el sistema

$$\mathbf{A} = \mathbf{LU}$$

donde los elementos de $\mathbf{L}$ y $\mathbf{U}$ son las incógnitas.

Tenemos n^2 ecuaciones y $n^2 + n$ incógnitas.

Debemos fijar n condiciones. Algunas posibilidades:

- Factorización de Doolittle, si L es triangular inferior unitaria.
- Factorización de Crout, si U es triangular superior unitaria.
- Factorización de Choleski, si $U = L^t$.

Por omisión, se considerará que la factorización LU es la de Doolittle.

Existencia de la factorización

Teorema

Si los n menores principales de la matriz $A \in \mathcal{M}_n$ son no singulares, entonces la matriz A admite una factorización LU.

Matriz de permutación

Dada $\sigma \in S_n$, denominamos **matriz de permutaciones asociada a** σ a la matriz

$$P = \begin{pmatrix} e_{\sigma(1)} \\ \dots \\ e_{\sigma(n)} \end{pmatrix}.$$

Para toda matriz $A \in \mathcal{M}_n$, **PA** permuta por σ las filas de **A**.

Lema

Sea P una matriz de permutaciones. Entonces en cada fila y en cada columna de P hay un único 1 y el resto de posiciones contienen ceros. Además, $\det(P) = \operatorname{sig}(\sigma)$ *y*

$$PP^t = P^t P = I.$$

Factorización LU con pivoteo

Decimos que $A \in \mathcal{M}_n$ admite una factorización LU con pivoteo si existen

1. una matriz de permutación P,
2. una matriz triangular inferior $L = (l_{ij})_{1 \leq i,j \leq n}$, con $l_{ii} = 1$, $1 \leq i \leq n$ y $|l_{ij}| \leq 1$ $1 \leq i,j \leq n$,
3. una matriz triangular superior U invertible

tales que $PA = LU$.

Si A admite una factorización LU con pivoteo entonces A factoriza como $A = P^t LU$.

Factorización LU con pivoteo

Para calcular la factorización con pivoteo, partimos de $P = L = Id$, $U = A$.

Para cada columna k, comenzando por la primera:

1. Permutamos las filas de U de modo que en la columa k el elemento de la diagonal sea mayor en módulo que los elementos que están debajo.
2. Esa misma permutación se la aplicamos a P y a L, pero en L no modificamos los elementos de la diagonal.
3. Hacemos ceros en la columna k de U por debajo de la diagonal.
4. Los factores que hemos utilizado en el paso anterior, los guardamos en la posición correspondiente de L.

Factorización LU con pivoteo

Teorema

Para toda matriz $A \in \mathcal{M}_n$ invertible, existe una matriz de permutación P, una matriz triangular inferior $L = (l_{ij})_{1 \leq i,j \leq n}$, con $l_{ii} = 1$, $1 \leq i \leq n$ y $|l_{ij}| \leq 1$ $1 \leq i,j \leq n$, y una matriz triangular superior U invertible tales que

$$PA = LU.$$

Factorización LU de Cholesky

Sea $A \in \mathcal{M}_n(\mathbb{R})$. Decimos que A admite una *factorización LU de Cholesky* si existe una matriz **real** triangular inferior **L** tal que los elementos de su diagonal son positivos y

$$\mathbf{A} = \mathbf{L}\mathbf{L}^\top.$$

Teorema

Una matriz **A** *real es simétrica y definida positiva si y sólo si admite una factorización LU de Choleski.*