11.10 Método de Cholesky

Este método sirve para resolver el sistema $Ax = b$ cuando la matriz A es **definida positiva** (también llamada positivamente definida). Este tipo de matrices se presenta en problemas específicos de ingeniería y física, principalmente.

11.10.1 Matrices definidas positivas

Una matriz simétrica es definida positiva si

$$x^{\mathrm{T}}Ax > 0, \quad \forall\, x \in \mathbb{R}^n,\ x \neq 0. \tag{11.5}$$

Para una matriz cuadrada cualquiera,

$$\begin{aligned}
x^{\mathrm{T}}Ax &= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix} \\
&= \sum_{i=1}^{n} \sum_{j=i}^{n} a_{ij}x_ix_j.
\end{aligned}$$

Si A es simétrica,

$$x^{\mathrm{T}}Ax = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} a_{ij}x_ix_j.$$

Ejemplo 11.7. Sea I la matriz identidad de orden n. Entonces $x^{\mathrm{T}}Ix = x^{\mathrm{T}}x = \|x\|^2$. Luego la matriz I es definida positiva. $\diamond$

Ejemplo 11.8. Sea A la matriz nula de orden n. Entonces $x^{\mathrm{T}}\mathbf{0}\,x = 0$. Luego la matriz nula no es definida positiva. $\diamond$

Ejemplo 11.9. Sea

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}. \\ x^{\mathrm{T}} A x &= x_1^2 + 5x_2^2 + 4x_1x_2 \\ &= x_1^2 + 4x_1x_2 + 4x_2^2 + x_2^2 \\ &= (x_1 + 2x_2)^2 + x_2^2. \end{aligned}$$

Obviamente $x^{\mathrm{T}}Ax \geq 0$. Además $x^{\mathrm{T}}Ax = 0$ si y solamente si los dos sumandos son nulos, es decir, si y solamente si $x_2 = 0$ y $x_1 = 0$, o sea, cuando $x = 0$. Luego A es definida positiva. ◇

Ejemplo 11.10. Sea

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}. \\ x^{\mathrm{T}} A x &= x_1^2 + 4x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2. \end{aligned}$$

Obviamente $x^{\mathrm{T}}Ax \geq 0$. Pero si $x = (6, -3)$, entonces $x^{\mathrm{T}}Ax = 0$. Luego A no es definida positiva. ◇

Ejemplo 11.11. Sea

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}. \\ x^{\mathrm{T}} A x &= x_1^2 + 3x_2^2 + 4x_1x_2 \\ &= (x_1 + 2x_2)^2 - x_2^2. \end{aligned}$$

Si $x = (6, -3)$, entonces $x^{\mathrm{T}}Ax = -9$. Luego A no es definida positiva. ◇

Ejemplo 11.12. Sea

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Como A no es simétrica, entonces no es definida positiva. ◇

Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores propios de A. Si A es simétrica, entonces todos sus valores propios son reales.

Sea δ_i el determinante de la submatriz de A, de tamaño $i\times i$, obtenida al quitar de A las filas $i+1$, $i+2$, ..., n y las columnas $i+1$, $i+2$, ..., n. O sea,

$$\begin{aligned}
\delta_1 &= \det([a_{11}]) = a_{11}, \\
\delta_2 &= \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \\
\delta_3 &= \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \\
&\vdots \\
\delta_n &= \det(A).
\end{aligned}$$

La definición 11.5 tiene relación directa con el nombre matriz definida positiva. Sin embargo, no es una manera fácil o práctica de saber cuándo una matriz simétrica es definida positiva, sobre todo si A es grande. El teorema siguiente presenta algunas de las caracterizaciones de las matrices definidas positivas. Para matrices pequeñas ($n \leq 4$) la caracterización por medio de los δ_i puede ser la de aplicación más sencilla. La última caracterización, llamada factorización de Cholesky, es la más adecuada para matrices grandes. En [Str86], [NoD88] y [Mor01] hay demostraciones y ejemplos.

Teorema 11.1. *Sea A simétrica. Las siguientes afirmaciones son equivalentes.*

- *A es definida positiva.*
- $\lambda_i > 0, \forall i$.
- $\delta_i > 0, \forall i$.
- *Existe U matriz triangular superior e invertible tal que* $A = U^T U$.

11.10.2 Factorización de Cholesky

Antes de estudiar el caso general, veamos la posible factorización para los ejemplos de la sección anterior.

La matriz identidad se puede escribir como $I = I^{\mathrm{T}} I$, siendo I triangular superior invertible. Luego existe la factorización de Cholesky para la matriz identidad.

Si existe la factorización de Cholesky de una matriz, al ser U y U^{T} invertibles, entonces A debe ser invertible. Luego la matriz nula no tiene factorización de Cholesky.

Sea

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}.$$

Entonces

$$\begin{aligned} \begin{bmatrix} u_{11} & 0 \\ u_{12} & u_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} &= \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \\ u_{11}^2 &= 1 \\ u_{11}u_{12} &= 2, \\ u_{12}^2 + u_{22}^2 &= 5 \end{aligned}$$

Se deduce que

$$\begin{aligned} u_{11} &= 1 \\ u_{12} &= 2, \\ u_{22} &= 1, \\ U &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Entonces existe la factorización de Cholesky de A.

Cuando se calculó u_{11} se hubiera podido tomar $u_{11} = -1$ y se hubiera podido obtener otra matriz U. Se puede demostrar que si se escogen los elementos diagonales u_{ii} positivos, entonces la factorización, cuando existe, es única.

Sea

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

Entonces

$$
\begin{aligned}
\begin{bmatrix} u_{11} & 0 \\ u_{12} & u_{22} \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} &= \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \\
u_{11}^2 &= 1 \\
u_{11}u_{12} &= 2, \\
u_{12}^2 + u_{22}^2 &= 4
\end{aligned}
$$

Se deduce que

$$
\begin{aligned}
u_{11} &= 1 \\
u_{12} &= 2, \\
u_{22} &= 0, \\
U &= \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Entonces, aunque existe U tal que $A = U^{\mathrm{T}}U$, sin embargo no existe la factorización de Cholesky de A ya que U no es invertible.

Sea

$$
A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.
$$

Entonces

$$
\begin{aligned}
\begin{bmatrix} u_{11} & 0 \\ u_{12} & u_{22} \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix} &= \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \\
u_{11}^2 &= 1 \\
u_{11}u_{12} &= 2, \\
u_{12}^2 + u_{22}^2 &= 3
\end{aligned}
$$

Se deduce que

$$
\begin{aligned}
u_{11} &= 1 \\
u_{12} &= 2, \\
u_{22}^2 &= -1.
\end{aligned}
$$

Entonces no existe la factorización de Cholesky de A.

En el caso general,

$$
\begin{bmatrix}
u_{11} & & & & & & \\
\vdots & & & & & & \\
u_{1k} & \cdots & u_{kk} & & & & \\
\vdots & & & & & & \\
u_{1j} & \cdots & u_{kj} & \cdots & u_{jj} & & \\
\vdots & & & & & & \\
u_{1n} & \cdots & u_{kn} & \cdots & u_{jn} & \cdots & u_{nn}
\end{bmatrix}
\begin{bmatrix}
u_{11} & \cdots & u_{1k} & \cdots & u_{1j} & \cdots & u_{1n} \\
 & & & & & & \vdots \\
 & & u_{kk} & \cdots & u_{kj} & \cdots & u_{kn} \\
 & & & & & & \vdots \\
 & & & & u_{jj} & \cdots & u_{jn} \\
 & & & & & & \vdots \\
 & & & & & & u_{nn}
\end{bmatrix}
$$

El producto de la fila 1 de U^{T} por la columna 1 de U da:

$$u_{11}^2 = a_{11}.$$

Luego

$$u_{11} = \sqrt{a_{11}}. \tag{11.6}$$

El producto de la fila 1 de U^{T} por la columna j de U da:

$$u_{11}u_{1j} = a_{1j}.$$

Luego

$$u_{1j} = \frac{a_{1j}}{u_{11}}, \qquad j = 2, ..., n. \tag{11.7}$$

Al hacer el producto de la fila 2 de U^{T} por la columna 2 de U, se puede calcular u_{22}. Al hacer el producto de la fila 2 de U^{T} por la columna j de U, se puede calcular u_{2j}. Se observa que el cálculo de los elementos de U se hace fila por fila. Supongamos ahora que se conocen los elementos de las filas 1, 2, ..., $k-1$ de U y se desea calcular los elementos de la fila k de U. El producto de la fila k de U^{T} por la columna k de U da:

$$
\begin{aligned}
\sum_{i=1}^{k} u_{ik}^2 &= a_{kk} \\
\sum_{i=1}^{k-1} u_{ik}^2 + u_{kk}^2 &= a_{kk}.
\end{aligned}
$$

Luego

$$u_{kk} = \sqrt{a_{kk} - \sum_{i=1}^{k-1} u_{ik}^2}, \qquad k = 2, ..., n. \tag{11.8}$$

El producto de la fila k de U^{T} por la columna j de U da:

$$\sum_{i=1}^{k} u_{ik}u_{ij} = a_{kj}.$$

Luego

$$u_{kj} = \frac{a_{kj} - \sum_{i=1}^{k-1} u_{ik}u_{ij}}{u_{kk}}, \qquad k = 2, ..., n, \;\; j = k+1, ..., n. \tag{11.9}$$

Si consideramos que el valor de la sumatoria es 0 cuando el límite inferior es más grande que el límite superior, entonces las fórmulas 11.8 y 11.9 pueden ser usadas para $k = 1, ..., n$.

Ejemplo 11.13. Sea

$$A = \begin{bmatrix} 16 & -12 & 8 & -16 \\ -12 & 18 & -6 & 9 \\ 8 & -6 & 5 & -10 \\ -16 & 9 & -10 & 46 \end{bmatrix}.$$

$$u_{11} = \sqrt{16} = 4$$

$$u_{12} = \frac{-12}{4} = -3$$

$$u_{13} = \frac{8}{4} = 2$$

$$u_{14} = \frac{-16}{4} = -4$$

$$u_{22} = \sqrt{18 - (-3)^2} = 3$$

$$u_{23} = \frac{-6 - (-3)(2)}{3} = 0$$

$$u_{24} = \frac{9 - (-3)(-4)}{3} = -1$$

$$u_{33} = \sqrt{5 - (2^2 + 0^2)} = 1$$

$$u_{34} = \frac{-10 - (\ 2(-4) + 0(-1)\ )}{1} = -2$$

$$u_{44} = \sqrt{46 - (\ (-4)^2 + (-1)^2 + (-2)^2\ )} = 5\,.$$

Entonces,

$$U \;=\; \begin{bmatrix} 4 & -3 & 2 & -4 \\ 0 & 3 & 0 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}. \ \diamond$$

La factorización de Cholesky no existe cuando en la fórmula 11.8 la cantidad dentro del radical es negativa o nula. Utilizando el producto escalar, las fórmulas 11.8 y 11.9 se pueden reescribir así:

$$\begin{aligned} t &= a_{kk} - \text{prodEsc}(\ U(1:k-1,k)\ ,\ U(1:k-1,k)\ ), \\ u_{kk} &= \sqrt{t}, \\ u_{kj} &= \frac{a_{kj} - \text{prodEsc}(\ U(1:k-1,k)\ ,\ U(1:k-1,j)\ )}{u_{kk}} \end{aligned}$$

Para ahorrar espacio de memoria, los valores u_{kk} y u_{kj} se pueden almacenar sobre los antiguos valores de a_{kk} y a_{kj}. O sea, al empezar el algoritmo se tiene la matriz A. Al finalizar, en la parte triangular superior del espacio ocupado por A estará U.

$$t = a_{kk} - \text{prodEsc}(\ A(1:k-1,k)\ ,\ A(1:k-1,k)\ ), \tag{11.10}$$

$$a_{kk} = \sqrt{t}, \tag{11.11}$$

$$a_{kj} = \frac{a_{kj} - \text{prodEsc}(\ A(1:k-1,k)\ ,\ A(1:k-1,j)\ )}{a_{kk}} \tag{11.12}$$

El siguiente es el esquema del algoritmo para la factorización de Cholesky. Si acaba normalmente, la matriz A es definida positiva. Si en algún momento $t \le \varepsilon$, entonces A no es definida positiva.

datos: A, ε
para $k = 1, ..., n$
    cálculo de t según (11.10)
    si $t \le \varepsilon$ **ent salir**
    $a_{kk} = \sqrt{t}$
    para $j = k+1, ..., n$
        cálculo de a_{kj} según (11.12)
    fin-para j
fin-para k