Descomposiciones matriciales LU y Choleski

Una matriz cuadrada $A \in M_{n\times n}(\mathbb{R})$ posee descomposición LU cuando existen matrices L, $U \in M_{n\times n}(\mathbb{R})$ triangular inferior (lower) y triangular superior (upper) respectivamente, tales que $A = LU$. Tal descomposición permite resolver de forma muy rápida cualquier sistema compatible determinado $Ax = b$ a través de dos algoritmos de bajada y subida:

$$LUx = b \;\Leftrightarrow\; [Lz = b \;\&\; Ux = z].$$

Los resultados conocidos de CNI respecto a la descomposición de tipo LU de una matriz cuadrada son:

Descomposición LU

Teorema. Si los menores principales de una matriz A de dimensión n son no nulos entonces A admite una descomposición LU. Esta descomposición es única si los elementos de la diagonal principal de L son todos unos.

Una condición suficiente para aplicar el resultado anterior es la siguiente:

Definición. Sea $A = (a_{ij})$ una matriz cuadrada $n \times n$. Se dice que A es diagonalmente dominante en sentido estricto si

$$|a_{ii}| > \sum_{j=1,\, j\neq i}^{n} |a_{ij}| \quad \text{para toda fila } i = 1, ..., n. \tag{1}$$

Teorema. Sea $A = (a_{ij})$ una matriz cuadrada $n \times n$ diagonalmente dominante en sentido estricto. Entonces,

a) A es regular,

b) A admite una única factorización LU de Doolittle.

Cálculo efectivo LU

Para el cálculo efectivo de la factorización LU de una matriz dada, en la práctica se obtienen los elementos de L y de U por identificación de forma recursiva, admitiendo que existe la factorización. Así, si

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{n,n} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ l_{n-1,1} & l_{n-1,2} & \dots & 1 & 0 \\ l_{n,1} & l_{n,2} & \dots & l_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1,n-1} & u_{1n} \\ 0 & u_{22} & \dots & u_{2,n-1} & u_{2n} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & u_{n-1,n-1} & u_{n-1,n} \\ 0 & 0 & \dots & 0 & u_{n,n} \end{pmatrix},$$

identificando los elementos de la primera fila de A con los correspondientes de LU, y los de la primera columna de A con los correspondientes de LU, obtenemos

$$\begin{cases} a_{1j} = u_{1j}, \quad j = 1, ..., n, \\ a_{i1} = l_{i1}u_{11} \Rightarrow l_{i1} = \dfrac{a_{i1}}{u_{11}}, \quad i = 2, ..., n. \end{cases}$$

Identificando a continuación los elementos de la segunda fila de A con los correspondientes de LU, y los de la segunda columna de A con los correspondientes de LU, obtenemos

$$\begin{cases} a_{2j} = l_{21}u_{1j} + u_{2j} \Rightarrow u_{2j} = a_{2j} - l_{21}u_{1j}, \quad j = 2, ..., n, \\ a_{i2} = l_{i1}u_{12} + l_{i2}u_{22} \Rightarrow l_{i2} = \dfrac{1}{u_{22}}(a_{i2} - l_{i1}u_{12}), \quad i = 3, ..., n. \end{cases}$$

Y en general, si suponemos conocidas las $k-1$ primeras filas de U y las $k-1$ primeras columnas de L, entonces, identificando los elementos de la k-ésima fila de A con los correspondientes de LU, y los de la k-ésima columna de A con los correspondientes de LU, obtenemos

$$\begin{cases} a_{kj} = l_{k1}u_{1j} + ... + l_{k,k-1}u_{k-1,j} + u_{kj} \\ \Rightarrow u_{kj} = a_{kj} - \displaystyle\sum_{r=1}^{k-1} l_{kr}u_{rj}, \quad j = k, ..., n, \\ a_{ik} = l_{i1}u_{1k} + ... + l_{i,k-1}u_{k-1,k} + l_{ik}u_{kk} \\ \Rightarrow l_{ik} = \dfrac{1}{u_{kk}}\left(a_{ik} - \displaystyle\sum_{r=1}^{k-1} l_{ir}u_{rk}\right), \quad i = k+1, ..., n. \end{cases}$$

Descomposición de Choleski

Forzando un poco más la factorización LU de una matriz, en casos especiales encontramos que pueden elegirse $L = U^* = B$, es lo que se conoce como la factorización de Choleski para una matriz cuadrada simétrica: $A = BB^*$. Que exista, a partir del Teorema de Doolittle, está intimamente relacionada con que en la expresión $A = LU = LDL^*$ los elementos de la matriz diagonal D sean positivos, lo que conduce a la siguiente definición y siguientes resultados en CNI.

Definición. Sea $A = (a_{ij})$ una matriz $n \times n$ de coeficientes reales. Diremos que A es definida positiva si

$$\sum_{i,j=1}^{n} a_{ij}u_iu_j > 0, \quad \text{para todo } u \in \mathbb{R}^n \setminus \{0\}. \tag{2}$$

Teorema. (**Factorización de Cholesky**) Si A es una matriz $n \times n$ simétrica definida positiva, entonces existe al menos una matriz $n \times n$ triangular inferior B tal que $A = BB^*$. Además, se puede imponer que $b_{ii} > 0$ para todo $i = 1, ..., n$, y en tal caso la factorización anterior es única.

El recíproco del Teorema de factorización de Cholesky es también cierto.

Proposición. Si B es una matriz $n \times n$ triangular inferior regular, y definimos $A = BB^*$, entonces A es simétrica y definida positiva.

Proposición. (**Criterio de Sylvester de matriz definida positiva**) Sea A una matriz real simétrica $n \times n$. Dicha matriz es definida positiva si y sólo si todos sus menores principales son positivos.

Cálculo efectivo factorización de Choleski

Para el cálculo efectivo de la factorización de Cholesky, se parte de la igualdad

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{n,n} \end{pmatrix}$$

$$= \begin{pmatrix} b_{11} & 0 & \dots & 0 & 0 \\ b_{21} & b_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n-1} & 0 \\ b_{n,1} & b_{n,2} & \dots & b_{n,n-1} & b_{n,n} \end{pmatrix} \begin{pmatrix} b_{11} & b_{21} & \dots & b_{n-1,1} & b_{n1} \\ 0 & b_{22} & \dots & b_{n-1,2} & b_{n2} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & b_{n-1,n-1} & b_{n,n-1} \\ 0 & 0 & \dots & 0 & b_{n,n} \end{pmatrix}$$

con $b_{ii} > 0$.

De esta forma, identificando los elementos de la primera fila de A con los correspondientes de BB^*, obtenemos

$$\begin{cases} a_{11} = b_{11}^2 \Rightarrow b_{11} = \sqrt{a_{11}} \\ a_{1j} = b_{11} b_{j1} \Rightarrow b_{j1} = \dfrac{a_{1j}}{b_{11}}, \quad j = 2, ..., n, \end{cases}$$

identificando a continuación los elementos de la segunda fila de A con los correspondientes de BB^*, obtenemos

$$a_{2j} = b_{21} b_{j1} + b_{22} b_{j2}, \quad j = 2, ..., n,$$

y en consecuencia,

$$\begin{cases} b_{22} = \sqrt{a_{22} - b_{21}^2}, \\ b_{j2} = \dfrac{1}{b_{22}}(a_{2j} - b_{21} b_{j1}), \quad j = 3, ..., n, \end{cases}$$

y en general si suponemos conocidas las $k-1$ primeras columnas de B entonces, identificando los elementos de la k-ésima fila de A con los correspondientes de BB^*, obtenemos

$$a_{kj} = b_{k1} b_{j1} + b_{k2} b_{j2} + ... + b_{kk} b_{jk}, \quad j = k, ..., n,$$

y en consecuencia,

$$\begin{cases} b_{kk} = \sqrt{a_{kk} - \displaystyle\sum_{r=1}^{k-1} b_{kr}^2}, \\ b_{jk} = \dfrac{1}{b_{kk}} \left(a_{kj} - \displaystyle\sum_{r=1}^{k-1} b_{kr} b_{jr} \right), \quad j = k+1, ..., n. \end{cases}$$

Se puede demostrar que el número de operaciones necesarias para llevar a cabo el método de cálculo anterior resulta ser aproximadamente la mitad de sumas y productos que en el método de Gauss o el LU, a cambio de calcular además n raíces cuadradas. Cuando n es grande, el método de Cholesky es el más favorable para el caso de matrices simétricas y definidas positivas.

Práctica 1bis: Descomposiciones matriciales LU y Choleski

Existen funciones ya programadas en MATLAB para calcular descomposiciones LU y Choleski (ver ayuda sobre los comandos `lu` y `chol`). Nuestro objetivo es programar nosotros mismos dichas descomposiciones.

1. Programar la descomposición LU de una matriz dada A, supuesto que se cumplen las condiciones del Teorema de Doolittle.

   Comprobar el funcionamiento del programa con el siguiente ejemplo:

$$\begin{pmatrix} 5 & 6 & 7 \\ 10 & 20 & 23 \\ 15 & 50 & 67 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 5 & 6 & 7 \\ 0 & 8 & 9 \\ 0 & 0 & 10 \end{pmatrix}.$$

2. La resolución programada (bucle) de un sistema triangular superior por algoritmo de subida es $u_n = b_n/a_{n,n}$; $u_{n-1} = (b_{n-1} - a_{n-1,n}u_n)/a_{n-1,n-1}$; ..., $u_1 = (b_1 - a_{1,2}u_2 - \ldots - a_{1,n}u_n)/a_{1,1}$.

   La resolución programada (bucle) de un sistema triangular inferior por algoritmo de bajada es $u_1 = b_1/a_{1,1}$; $u_2 = (b_2 - a_{2,1}u_1)/a_{2,2}$; ..., $u_n = (b_n - a_{n,1}u_1 - \ldots - a_{n,n-1}u_{n-1})/a_{n,n}$.

   Realizar dos programas que resuelvan sistemas triangulares por algoritmos de subida y bajada respectivamente.

3. Usar la matriz dada en el ejercicio 1 para resolver el sistema $Ax = (18, 53, 132)^t$utilizando la descomposición LU obtenida entonces y los algoritmos de subida y bajada del ejercicio 2. La solución es $(1, 1, 1)^t$.

4. Programar la descomposición de Choleski de una matriz supuesto que satisfaga las condiciones del teorema de descomposición de Choleski.

   Aplicarlo para obtener B tal que $A = BB^*$. La matriz A dato y la solución B son:

$$A = \begin{pmatrix} 1 & 2 & 4 & 7 \\ 2 & 13 & 23 & 38 \\ 4 & 23 & 77 & 122 \\ 7 & 38 & 122 & 294 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{pmatrix}.$$

5. Resolver el sistema $Ax = (14, 76, 226, 461)^t$ donde A es la matriz del ejercicio anterior. Para resolverlo, aplica la descomposición de Choleski

   y resuelve dos algoritmos, uno de subida y otro de bajada. La solución del sistema es $(1, 1, 1, 1)^t$.

6. Completar los ejercicios 1 y 4 para comprobar si se cumplen las hipótesis del teorema de Doolittle y del teorema de Choleski y criterio de Sylvester respectivamente sobre una matriz A dada.

7. **(Estructura tridiagonal de la factorización de Cholesky).** Consideremos la matriz tridiagonal simétrica de dimensión $n \times n$:

$$A = \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 & 0 & 0 \\ -1 & 2 & -1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 & 2 & -1 \\ 0 & 0 & 0 & \ldots & 0 & -1 & 2 \end{pmatrix}.$$

   Comprobar con distintos valores de n que la factorización de Cholesky de A mantiene la estructura tridiagonal.

   Indicación: Construcción de matrices "grandes". (Para más información, usar `help` con los comandos anteriores.)

   - `zeros(n,m)` matriz $n \times m$ llena de ceros.
   - `ones(n,m)` matriz $n \times m$ llena de unos.
   - `diag` aplicado a matriz, devuelve su diagonal; aplicado a un vector, genera matriz diagonal con ese vector en ella y resto de elementos nulos. También válida para generar diagonales superiores o inferiores (p.ej. `diag(v,1)` o `diag(v,-1)`).
   - `sparse` (avanzado: agiliza cálculos en una matriz hueca eliminando las entradas nulas).