

TIPOS ESPECIALES DE MATRICES, FACTORACIONES LL^T Y LDL^T

ALAN REYES-FIGUEROA
MÉTODOS NUMÉRICOS II

(AULA 08) 26.JULIO.2022

Estrategias de Pivoteo

Ejemplo:
$$\begin{array}{rclcl} 0.00300x_1 & + & 59.14x_2 & = & 59.17 \\ 5.291x_1 & - & 6.130x_2 & = & 46.78 \end{array} \cdot \text{La solución exacta es } \begin{pmatrix} x_1 = 10.0 \\ x_2 = 1.0 \end{pmatrix}.$$

En aritmética de 4 dígitos, si no usamos pivoteo, el sistema resulta en

$$\begin{pmatrix} 0.00300 & + & 59.14 & \approx & 59.17 \\ & - & 104,300 & \approx & -104,400 \end{pmatrix}$$

y la solución resulta $x_2 = \frac{-104400}{-104300} = 1.001$, y $x_1 = \frac{59.17 - 59.14(1.001)}{0.003} = \frac{59.17 - 59.20}{0.003} = -10.00$.

Usando pivoteo parcial, tendríamos
$$\begin{array}{rclcl} 5.291x_1 & - & 6.130x_2 & = & 46.78 \\ 0.00300x_1 & + & 59.14x_2 & = & 59.17 \end{array} \cdot \text{Luego}$$

$$\begin{pmatrix} 5.291 & - & 6.130 & \approx & 46.78 \\ & + & 59.142 & \approx & 59.14 \end{pmatrix}$$

Así, $x_2 = \frac{59.14}{59.14} = 1.000$ y $x_1 = \frac{46.78 + 6.130(1.000)}{5.291} = 10.00$.

Estrategias de Pivoteo

Ejemplo:
$$\begin{array}{rcl} 30.00x_1 & + & 591400x_2 = 591700 \\ 5.291x_1 & - & 6.130x_2 = 46.78 \end{array} \cdot \text{ La solución exacta es } \begin{pmatrix} x_1 = 10.0 \\ x_2 = 1.0 \end{pmatrix}.$$

En aritmética de 4 dígitos, si usamos pivoteo parcial, el sistema resulta en

$$\begin{pmatrix} 30.00 & + & 591400 & \approx & 591700 \\ & - & 104,300 & \approx & -104,400 \end{pmatrix}$$

y la solución resulta $x_2 = \frac{-104400}{-104300} = 1.001$, $x_1 = \frac{591700 - 591400(1.001)}{30.0} = \frac{591700 - 592000}{30.0} = -10.00$.

Usando pivoteo reescalado de columna, tendríamos

$$s_1 = 591700, \quad s_2 = 46.78, \quad \Rightarrow \quad \frac{|a_{11}|}{s_1} = \frac{30.0}{591700} = 5.070 \times 10^{-5}, \quad \frac{a_{21}}{s_2} = \frac{5.291}{46.78} = 0.1131.$$

De ahí el sistema se escribe
$$\begin{array}{rcl} 5.291x_1 & - & 6.130x_2 = 46.78 \\ 0.00300x_1 & + & 59.14x_2 = 59.17 \end{array}, \text{ y obtenemos la}$$

solución $x_2 = \frac{59.14}{59.14} = 1.000$ y $x_1 = \frac{46.78 + 6.130(1.000)}{5.291} = 10.00$.

Descomposición de Crout

En el caso $A \in \mathbb{R}^{n \times n}$, a la descomposición LU que vimos en el aula anterior, también se le llama la **descomposición de Doolittle**.

$$A = \begin{pmatrix} 1 & & & & \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n1} & \ell_{n2} & \dots & \ell_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & * & * & \dots & * \\ & u_{22} & * & & * \\ & & u_{33} & \ddots & \vdots \\ & & & \ddots & * \\ & & & & u_{nn} \end{pmatrix}.$$

Existe otro tipo de factoración LU , llamada la **descomposición de Crout**. Esta es de la forma

$$A = \begin{pmatrix} \ell_{11} & & & & \\ * & \ell_{22} & & & \\ * & * & \ell_{33} & & \\ \vdots & & \ddots & \ddots & \\ * & * & \dots & * & \ell_{nn} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & u_{13} & \dots & u_{1n} \\ & 1 & u_{23} & \dots & u_{2n} \\ & & 1 & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ & & & & 1 \end{pmatrix}.$$

Teorema (Teorema 22.1, Libro de Trefethen)

Sea $A = LU$ la factoración de una matriz no singular $A \in \mathbb{R}^{n \times n}$ se calcula por eliminación gaussiana sin pivoteo en una computadora que satisfaga los axiomas de la aritmética de punto flotante (13.5) y (13.7). Si A tiene una factorización LU , entonces, para todos los ε_{maq} suficientemente pequeños, la factorización completa con éxito en aritmética de punto flotante (no se encuentran pivotes cero), y las matrices calculadas \tilde{L} y \tilde{U} satisfacen

$$\tilde{L}\tilde{U} = A + \delta A, \quad \frac{\|\delta A\|}{\|\tilde{L}\| \|\tilde{U}\|} = O(\varepsilon_{maq}),$$

para alguna $\delta A \in \mathbb{R}^{n \times n}$.

Teorema (Teorema 22.2, Libro de Trefethen)

Sea $PA = LU$ la factoración de una matriz $A \in \mathbb{R}^{n \times n}$ calculada por eliminación gaussiana con pivoteo parcial (algoritmo 21.1) en un computador que satisfaga los axiomas (13.5) y (13.7). Entonces las matrices calculadas \tilde{P} , \tilde{L} y \tilde{U} satisfacen

$$\tilde{L}\tilde{U} = \tilde{P}A + \delta A, \quad \frac{\|\delta A\|}{\|A\|} = O(\rho \varepsilon_{maq}),$$

para alguna $\delta A \in \mathbb{R}^{n \times n}$, donde ρ es el factor de crecimiento para A . Si $|\ell_{jk}| < 1$ para cada $j > k$, lo que implica que no hay empates en la selección de pivotes, entonces $\tilde{P} = P$ para toda ε_{maq} suficientemente pequeña.

En otras palabras, la eliminación gaussiana es estable hacia atrás.

Diagonal Dominancia

Definición

Una matriz $A \in \mathbb{R}^{n \times n}$ es **diagonalmente dominante** cuando cada entrada de la diagonal principal (en módulo) es mayor o igual que la suma del resto de entradas en la misma fila

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|, \quad \text{para todo } i = 1, 2, \dots, n.$$

En caso las desigualdes sean todas estrictas, decimos que A es **estrictamente diagonal dominante**.

Ejemplo:

$$A = \begin{pmatrix} 4 & 2 & 0 \\ 3 & 6 & -1 \\ 0 & 3 & -5 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 6 & -2 \\ 3 & -1 & 0 \\ -4 & 2 & 1 \end{pmatrix}$$

A es diagonalmente dominante, pero B no. De hecho, A es estrictamente diagonal dominante.

Diagonal Dominancia

Teorema

Una matriz estrictamente diagonal dominante A es no singular. En este caso, la eliminación gaussiana se puede realizar en cualquier sistema lineal de la forma $A\mathbf{x} = \mathbf{b}$ sin intercambios de fila o columna, y los cálculos serán estables respecto al crecimiento del error de redondeo.

Prueba: Considere el sistema $A\mathbf{x} = \mathbf{0}$ y suponga que existe una solución no trivial $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ para este sistema. Sea k un índice para el que $0 < |x_k| = \max_{1 \leq j \leq n} |x_j|$. Como $\sum_{j=1}^n a_{ij}x_j = 0$, para cada $i = 1, 2, \dots, n$, entonces cuando $i = k$ se tiene $a_{kk}x_k = -\sum_{j \neq k} a_{kj}x_j$. De la desigualdad triangular

$$|a_{kk}| |x_k| = \left| \sum_{j \neq k} a_{kj}x_j \right| \leq \sum_{j \neq k} |a_{kj}| |x_j| \quad \Rightarrow \quad |a_{kk}| \leq \sum_{j \neq k} |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{j \neq k} |a_{kj}|.$$

Esta desigualdad contradice la dominancia diagonal estricta de A . Por consiguiente, la única solución para $A\mathbf{x} = \mathbf{0}$ es $\mathbf{x} = \mathbf{0}$. Esto muestra que A es no singular.

Diagonal Dominancia

Mostramos ahora que las matrices U_1, U_2, \dots, U_n generadas por el proceso de eliminación gaussiana son estrictamente diagonal dominantes. Eso garantiza que en cada etapa el elemento pivote es distinto a cero.

Como $A = U_1$ es estrictamente diagonal dominante, $a_{11} \neq 0$ y U_2 se puede calcular.

Además, para $i = 2, 3, \dots, n$ y $j = 2, 3, \dots, n$

$$u_{ij}^{(2)} = u_{ij}^{(1)} - \frac{u_{i1}^{(1)}}{u_{ii}^{(2)}} u_{1j}^{(2)}$$

Primero, $u_{i1}^{(2)} = 0$. La desigualdad triangular implica que

$$\sum_{j=2, j \neq i}^n |u_{ij}^{(2)}| = \sum_{j=2, j \neq i}^n \left| u_{ij}^{(1)} - \frac{u_{i1}^{(1)}}{u_{ii}^{(2)}} u_{1j}^{(2)} \right| \leq \sum_{j=2, j \neq i}^n |u_{ij}^{(1)}| + \sum_{j=2, j \neq i}^n \frac{|u_{i1}^{(1)}|}{|u_{ii}^{(2)}|} |u_{1j}^{(2)}|$$

Pero, siendo A es estrictamente diagonalmente dominante, sabemos

$$\sum_{j=2, j \neq i}^n |u_{ij}^{(1)}| \leq |u_{ii}^{(1)}| - |u_{i1}^{(1)}| \quad \text{y} \quad \sum_{j=2, j \neq i}^n |u_{1j}^{(1)}| \leq |u_{11}^{(1)}| - |u_{1i}^{(1)}|,$$

Diagonal Dominancia

por lo que

$$\begin{aligned}\sum_{j=2, j \neq i}^n |u_{ij}^{(2)}| &< |u_{ii}^{(1)}| - |u_{i1}^{(1)}| + \frac{|u_{i1}^{(1)}|}{|u_{11}^{(2)}|} (|u_{11}^{(1)}| - |u_{1i}^{(1)}|) = |u_{11}^{(1)}| + \frac{|u_{i1}^{(1)}| |u_{1i}^{(1)}|}{|u_{11}^{(1)}|} \\ &< \left| |u_{11}^{(1)}| + \frac{|u_{i1}^{(1)}| |u_{1i}^{(1)}|}{|u_{11}^{(1)}|} \right| = |u_{ii}^{(2)}|\end{aligned}$$

Esto establece la diagonal dominancia para las filas $2, 3, \dots, n$. La primera fila de U_2 y de $U_1 = A$ son la misma, por lo que U_2 es estrictamente diagonal dominante.

Este proceso continúa de manera inductiva hasta que se obtiene U_n triangular superior y estrictamente diagonal dominante. Esto implica que todos los elementos diagonales son no-nulos, y se puede realizar la eliminación gaussiana sin intercambios de fila.

La demostración de estabilidad para este procedimiento se puede encontrar en el libro de WENDROFF, *Theoretical Numerical Analysis*.

Matrices Positivo Definidas

Definición

Una matriz simétrica $A \in \mathbb{R}^{n \times n}$ es **positiva definida** ($A \succ 0$), si $\mathbf{x}^T A \mathbf{x} > 0$, $\forall \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

Observaciones:

- No todos los autores requieren la simetría. En el libro de GOLUB y VAN LOAN, para que A sea positiva definida se requiere únicamente que $\mathbf{x}^T A \mathbf{x} > 0$, para todo $\mathbf{x} \neq \mathbf{0}$.
- Existen definiciones similares:
 - A es **positiva semi-definida** ($A \succeq 0$) si $\mathbf{x}^T A \mathbf{x} \geq 0$, $\forall \mathbf{x} \neq \mathbf{0}$,
 - A es **negativa definida** ($A \prec 0$) si $\mathbf{x}^T A \mathbf{x} < 0$, $\forall \mathbf{x} \neq \mathbf{0}$,
 - A es **negativa semi-definida** ($A \preceq 0$) si $\mathbf{x}^T A \mathbf{x} \leq 0$, $\forall \mathbf{x} \neq \mathbf{0}$,
 - A es **no definida** si no cumple ninguna de las anteriores.
- El signo de $\mathbf{x}^T A \mathbf{x}$ clasifica las formas cuadráticas (recordar $\mathbf{x}^T A \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$).

Matrices Positivo Definidas

Ejemplo: La matriz $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ es definida positiva.

Tomemos $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$, $\mathbf{x} \neq \mathbf{0}$. Entonces

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= 2x_1^2 - x_1x_2 - x_2x_1 + 2x_2^2 - x_2x_3 - x_3x_2 + 2x_3^2 \\ &= x_1^2 + (x_1^2 - 2x_1x_2 + x_2^2) + (x_2^2 - 2x_2x_3 + x_3^2) + x_3^2 \\ &= x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 \geq 0. \end{aligned}$$

Mas aún, esta suma es cero, únicamente si todos los términos se anulan, y esto se cumple si, y sólo si, $x_1 = x_2 = x_3 = 0$.

Entonces $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$. Esto muestra que A es definida positiva.

Matrices Positivo Definidas

Teorema

Una matriz simétrica A es definida positiva \Leftrightarrow todos sus autovalores son positivos.

Prueba: (\Rightarrow). Supongamos que $A \succ 0$, y sea λ un autovalor de A . Observe primero que, como A es simétrica, $\lambda \in \mathbb{R}$. Por otro lado, existe $\mathbf{x} \neq \mathbf{0}$ autovector asociado a $\lambda \Rightarrow A\mathbf{x} = \lambda\mathbf{x}$. Entonces $0 < \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T (\lambda \mathbf{x}) = \lambda \mathbf{x}^T \mathbf{x} = \lambda \|\mathbf{x}\|_2^2$. Como $\|\mathbf{x}\|_2^2 > 0$, entonces $\lambda > 0$.

(\Leftarrow) Suponga ahora que $\lambda_i > 0$ para todo autovalor de A . Por el Teorema Espectral, A posee una base ortonormal de autovectores $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$, con autovalores asociados $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$. Sea $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$. Escribimos $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{q}_i$, con al menos una $c_i \neq 0$. Entonces

$$\begin{aligned} \mathbf{x}^T A \mathbf{x} &= \left(\sum_{i=1}^n c_i \mathbf{q}_i \right)^T A \left(\sum_{i=1}^n c_i \mathbf{q}_i \right) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \mathbf{q}_i^T A \mathbf{q}_j = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \mathbf{q}_i^T (\lambda_j \mathbf{q}_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n c_i c_j \lambda_j (\mathbf{q}_i^T \mathbf{q}_j) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \lambda_j \delta_{ij} = \sum_{i=1}^n c_i^2 \lambda_i > 0. \quad \square \end{aligned}$$

Matrices Positivo Definidas

Propiedades

Si $A \in \mathbb{R}^{n \times n}$ es una matriz simétrica y positiva definida, entonces

- a) A es no singular,
- b) $\max_{1 \leq j, k \leq n} |a_{jk}| \leq \max_{1 \leq i \leq n} |a_{ii}|$,
- c) $a_{ii} > 0$, para todo $i = 1, 2, \dots, n$,
- d) $(a_{ij})^2 < a_{ii}a_{jj}$, para cada $i \neq j$.
- e) (Criterio de SYLVESTER) Todos los determinantes principales menores de $M_i = A[1 : i, 1 : i]$ y $N_i = A[i : n, i : n]$, para $i = 1, 2, \dots, n$ son positivos.

Prueba: (a) Como todos los autovalores λ_i de A son positivos, entonces $\text{rank } A = n$ y $\text{Ker } A = \{0\}$. Esto muestra que A es no singular.

(b) Para cada $i = 1, 2, \dots, n$, tome $\mathbf{x} = \mathbf{e}_i$. Entonces $0 < \mathbf{x}^T A \mathbf{x} = \sum_i \sum_j a_{ij} x_i x_j = a_{ii}$.

Matrices Positivo Definidas

(c) Para cada $k \neq j$, definamos $\mathbf{x} \in \mathbb{R}^n$ por $\mathbf{x} = \mathbf{e}_j - \mathbf{e}_k = (0, \dots, 0, 1, 0, \dots, 0, -1, 0, \dots, 0)$, y $\mathbf{y} = \mathbf{e}_j + \mathbf{e}_k = (0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots, 0)$. Resulta

$$0 < \mathbf{x}^T \mathbf{A} \mathbf{x} = a_{jj} - a_{jk} - a_{kj} + a_{kk} = a_{jj} - 2a_{jk} + a_{kk},$$

$$0 < \mathbf{y}^T \mathbf{A} \mathbf{y} = a_{jj} + a_{jk} + a_{kj} + a_{kk} = a_{jj} + 2a_{jk} + a_{kk} < \max_{1 \leq i \leq n} |a_{ii}|$$

Entonces $2a_{jk} < a_{jj} + a_{kk}$ y $-2a_{jk} < a_{jj} + a_{kk}$. Luego

$$|a_{jk}| < \frac{a_{jj} + a_{kk}}{2} \leq \max_{1 \leq i \leq n} |a_{ii}|, \quad \Rightarrow \quad \max_{1 \leq k, j \leq n} |a_{jk}| < \max_{1 \leq i \leq n} |a_{ii}|.$$

(d) Para $i \neq j$, definimos $\mathbf{x} = \alpha \mathbf{e}_i + \mathbf{e}_j = (0, \dots, 0, \alpha, 0, \dots, 0, 1, 0, \dots, 0)$, con $\alpha \in \mathbb{R}$ arbitrario. Entonces $0 < \mathbf{x}^T \mathbf{A} \mathbf{x} = a_{ii}\alpha^2 + 2a_{ij}\alpha + a_{jj}$. Este polinomio cuadrático en α es siempre positivo, de modo que no posee raíces reales. Entonces su discriminante es negativo. Así, $4a_{ij}^2 - 4a_{ii}a_{jj} < 0 \Rightarrow a_{ij}^2 < a_{ii}a_{jj}$, para todo $i \neq j$.

(e) Pendiente. Se deduce de la descomposición de Cholesky $\mathbf{A} = \mathbf{R}^T \mathbf{R}$. \square

Matrices Positivo Definidas

Teorema

Una matriz simétrica $A \in \mathbb{R}^{n \times n}$ es positiva definida si, y sólo si, la eliminación gaussiana se puede realizar sin intercambios de fila, y todos los elementos pivote son positivos. Además, en este caso, los cálculos son estables respecto al crecimiento del error de redondeo.

Prueba: Ver libro de WENDROFF, *Theoretical Numerical Analysis*.

Eliminación Gaussiana Simétrica

Sea $A \in \mathbb{R}^{n \times n}$ positiva definida. Nos interesa descomponer A en factores triangulares LU . Si aplicamos un solo paso de la eliminación gaussiana a la matriz A , con un 1 en la primera entrada ($a_{11} = 1$) obtenemos

$$\begin{pmatrix} 1 & \mathbf{w}^T \\ \mathbf{w} & K \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{w} & I \end{pmatrix} \begin{pmatrix} 1 & \mathbf{w}^T \\ \mathbf{0} & K - \mathbf{w}\mathbf{w}^T \end{pmatrix}.$$

Ahora introducimos ceros en la segunda columna. Sin embargo, para mantener la simetría, se hace una variante de la descomposición LU , llamada la **factoración de Cholesky**, que primero introduce ceros en la primera fila para coincidir con los ceros recién introducidos en la primera columna de U . Podemos hacer esto por una operación triangular superior derecha que resta múltiplos de la primera columna de los siguientes:

$$\begin{pmatrix} 1 & \mathbf{w}^T \\ \mathbf{0} & K - \mathbf{w}\mathbf{w}^T \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & K - \mathbf{w}\mathbf{w}^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{w}^T \\ \mathbf{0} & I \end{pmatrix}.$$

Esta operación es la transpuesta de la triangular inferior arriba.

Descomposición de Cholesky

Tenemos

$$A = \begin{pmatrix} 1 & \mathbf{w}^T \\ \mathbf{w} & K \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{w} & I \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & K - \mathbf{w}\mathbf{w}^T \end{pmatrix} \begin{pmatrix} 1 & \mathbf{w}^T \\ \mathbf{0} & I \end{pmatrix}.$$

La idea de la factorización de Cholesky es continuar este proceso, haciendo cero una columna y una fila de A simétricamente, hasta que se reduce a la identidad.

Para que la reducción triangular simétrica funcione en general, necesitamos una factorización que funcione para cualquier todo $a_{11} > 0$, no sólo el caso $a_{11} = 1$. La generalización se logra ajustando algunos de los elementos de la fila 1, por un factor de $\alpha = \sqrt{a_{11}}$:

$$A = \begin{pmatrix} a_{11} & \mathbf{w}^T \\ \mathbf{w} & K \end{pmatrix} = \begin{pmatrix} \alpha & \mathbf{0} \\ \frac{\mathbf{w}}{\alpha} & I \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & K - \frac{\mathbf{w}\mathbf{w}^T}{a_{11}} \end{pmatrix} \begin{pmatrix} \alpha & \frac{\mathbf{w}^T}{\alpha} \\ \mathbf{0} & I \end{pmatrix} = R_1^T A_1 R_1.$$

Si la entrada superior izquierda de la submatriz $K - \frac{\mathbf{w}\mathbf{w}^T}{a_{11}}$ es positiva, el proceso puede repetirse de la misma forma, para factorizarla. Así, $A_1 = R_2^T A_2 R_2$ y $A = R_1^T R_2^T A_2 R_2 R_1$. Continuando este proceso, obtenemos eventualmente

Descomposición de Cholesky

$$A = \underbrace{R_1^T R_2^T \cdots R_n^T}_{R^T} \underbrace{R_n \cdots R_2 R_1}_R = R^T R,$$

donde R es triangular superior, y $r_{jj} > 0, \forall j$. Esta factoración se conoce como la **descomposición LL^T** o **factoración de CHOLESKY**.

La descripción anterior deja una pregunta. ¿Cómo sabemos que la entrada superior izquierda de la submatriz $K - \frac{\mathbf{w}\mathbf{w}^T}{a_{11}}$ es positiva? La respuesta es que debe ser positiva porque $K - \frac{\mathbf{w}\mathbf{w}^T}{a_{11}}$ es positiva definida, ya que es el $(n-1) \times (n-1)$ submatriz principal inferior derecha de la matriz definida positiva $R_1^{-T} A R_1^{-1}$. Por inducción, el mismo argumento muestra que todas las submatrices subsiguientes son definidas positivas, y por lo tanto el proceso no concluye con éxito.

Teorema

$A \in \mathbb{R}^{n \times n}$ es simétrica y positiva definida \Leftrightarrow admite una factoración LL^T . \square

Descomposición de Cholesky

Algoritmo: (Factoración de Cholesky ó LL^T).

Inputs: $A \in \mathbb{R}^{n \times n}$ simétrica y positiva definida, *Outputs:* $R \in \mathbb{R}^{n \times m}$ tal que $A = R^T R$.

Initialize $R = A$.

for $k = 1$ to n :

 for $j = k + 1$ to n :

$$R_{j,k:n} = R_{j,k:n} - (R_{kj}/R_{kk}) R_{k,k:n},$$

$$R_{k,k:n} = R_{k,k:n} / \sqrt{R_{kk}}.$$

El número de operaciones aritméticas es $O(\frac{1}{3}n^3)$.

Descomposición de Cholesky

Teorema

Sea $A \in \mathbb{R}^{n \times n}$, simétrica y positiva definida. Si la descomposición de Cholesky de A se calcula mediante el algoritmo anterior en un computador que satisface los axiomas de la aritmética de punto flotante, entonces para todo ε_{maq} suficientemente pequeño, este proceso está garantizado de ejecutarse hasta el final (es decir, no surgirán entradas pivote cero o negativas r_{kk}), generando una matriz \tilde{R} que satisface

$$\tilde{R}^T \tilde{R} = A + \delta A, \quad \frac{\|\delta A\|}{\|A\|} \leq O(\varepsilon_{maq}),$$

para alguna $\delta A \in \mathbb{R}^{n \times n}$.

Teorema

Sea $A \in \mathbb{R}^{n \times n}$, simétrica y positiva definida. La solución del sistema $A\mathbf{x} = \mathbf{b}$ a través de la factorización Cholesky es estable hacia atrás, lo que genera una solución calculada que satisface

$$(A + \delta A)\tilde{\mathbf{x}} = \mathbf{b}, \quad \frac{\|\delta A\|}{\|A\|} = O(\varepsilon_{maq}),$$

para alguna $\delta A \in \mathbb{R}^{n \times n}$.

Descomposición LDL^T

Una variante de la descomposición de Cholesky es la **descomposición LDL^T** . Esta tiene la particularidad que evita calcular las raíces cuadradas $\sqrt{r_{kk}}$ necesarias en el algoritmo de Cholesky.

Para ello, en el primer paso, se factora la matriz como un producto de la forma

$$A = \begin{pmatrix} a_{11} & \mathbf{w}^T \\ \mathbf{w} & K \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \frac{\mathbf{w}}{a_{11}} & I \end{pmatrix} \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & K - \frac{\mathbf{w}\mathbf{w}^T}{a_{11}} \end{pmatrix} \begin{pmatrix} 1 & \frac{\mathbf{w}^T}{a_{11}} \\ \mathbf{0} & I \end{pmatrix} = L_1 D_1 L_1^T,$$

donde L_1 es triangular inferior con 1s en la diagonal, y D_1 ahora tiene en si primera entrada el valor a_{11} en lugar de 1. Continuando este proceso, obtenemos eventualmente

$$A = \underbrace{L_1 L_2 \cdots L_n}_L D \underbrace{L_n^T \cdots L_2^T L_1^T}_{L^T} = LDL^T,$$

donde L es triangular inferior con 1s en la diagonal, D es diagonal, y $d_{jj} > 0, \forall j$. Esta factoración se conoce como la **descomposición LDL^T** .

Descomposición LDL^T

Algoritmo: (Factoración LDL^T).

Inputs: $A \in \mathbb{R}^{n \times n}$ simétrica y positiva definida,

Outputs: $R, D \in \mathbb{R}^{n \times m}$, tales que $A = R^T D R$.

Initialize $R = A$, $D = I$.

for $k = 1$ to n :

$$D_{kk} = R_{kk},$$

for $j = k + 1$ to n :

$$R_{j,k:n} = R_{j,k:n} - (R_{kj}/D_{kk}) R_{k,k:n},$$

$$R_{k,k:n} = R_{k,k:n}/D_{kk}.$$

Teorema

$A \in \mathbb{R}^{n \times n}$ es simétrica y positiva definida \Leftrightarrow admite una factoración LDL^T . \square

Relaciones entre Descomposiciones

Tenemos varias relaciones entre las factoraciones discutidas. Suponemos aquí que $A \in \mathbb{R}^{n \times n}$ es simétrica y positiva definida.

- Si $A = LDL^T$, entonces $R = D^{1/2}L^T$, resulta en una factoración de Cholesky para A :
$$R^T R = (LD^{1/2})(D^{1/2}L^T) = LDL^T = A.$$

- Si $A = LDL^T$, entonces $U = DL^T$ es triangular superior, y LU resulta en una factoración de Doolittle para A :

$$LU = L(DL^T) = LDL^T = A.$$

- Si $A = LDL^T$, entonces $Z = LD$ es triangular inferior, y ZL^T resulta en una factoración de Crout para A :

$$ZL^T = (LD)L^T = LDL^T = A.$$

- Si $A = R^T R$, entonces haciendo la factoración LU de R^T , obtenemos $R^T = LU$, con L triangular inferior con 1's en la diagonal, y U es diagonal. Luego, $D = UU^T$ es digagonal y resulta en una factoración LDL^T para A :

$$LDL^T = L(UU^T)L^T = (LU)(U^T L^T) = R^T R = A.$$

La factoración de Cholesky es útil en muchas aplicaciones:

- mínimos cuadrados
- optimización no-lineal (método de NEWTON, DFP, BGFS)
- simulación Monte Carlo
- generación de matrices de covarianza
- filtros de KALMAN

Implementaciones computacionales:

- LAPACK (*Linear Algebra Package*, 1970's), Fortran 77.
- LINPACK (*Linear Package*, 1976, Argonne Labs.) Fortran, C.
- BLAS (*Basic Linear Algebra Subprograms*).