Cholesky-felbontás

Sáfár Orsolya

Cholesky-felbontás

Definíció: Az $\mathbf{A} \in \mathbb{R}^{n \times n}$ mátrix Cholesky-felbontásának nevezzük az $\mathbf{A} = \mathbf{U}^\mathbf{T}\mathbf{U}$ szorzatot, ahol $\mathbf{U}$ egy olyan felső háromszög-mátrix, amelynek minden diagonálisbeli eleme pozitív.

Tétel: Egy mátrixnak pontosan akkor létezik Cholesky-felbontása, ha szimmetrikus és pozitív definit. Ha létezik a Cholesky-felbontás, akkor egyértelmű és az alábbi képlettel lehet kiszámítani $\mathbf{U}$ elemeit:

$$\begin{aligned} u_{11} &= \sqrt{a_{11}}, \text{aztán} \quad i = 2, \ldots, n\text{-re:} \\ u_{ij} &= \left(a_{ij} - \sum_{k=1}^{j-1} u_{ik} u_{jk} \right) / u_{jj}, \quad j = 1, \ldots, i-1 \\ u_{ii} &= \sqrt{a_{ii} - \sum_{k=1}^{i-1} u_{ik}^2} \end{aligned} \tag{1}$$

Bizonyítás

A létezés igazolásához teljes indukcióval dolgozunk a mátrix mértére.

- $i = 1$-re $\mathbf{A} = [a_{11}]$, ahol $a_{11} > 0$ ha $\mathbf{A}$ pozitív definit, és ekkor $[\sqrt{a_{11}}] = [u_{11}]$ tényleg a Cholesky-felbontást adja.
- Tegyük fel tehát, hogy i-ig igaz az állítás.
- Ekkor $i + 1$-re:

$$\mathbf{A_{i+1}} = \begin{bmatrix} \mathbf{A_i} & \mathbf{v} \\ \mathbf{v^T} & \alpha \end{bmatrix}$$

alakba írható a mátrix, ahol $\mathbf{A_i}$ pozitív definit, ugyanis $\mathbf{A_{i+1}}$ főminorja. A felírásban kihasználtuk $\mathbf{A_{i+1}}$ szimmetriáját is, továbbá $\alpha > 0$ szintén a pozitív definitségből következik a $(0, \dots, 0, 1)$ vektorral. Másrészt az indukciós feltétel miatt $\mathbf{A_i} = \mathbf{U_i^T U_i}$, ahol $\mathbf{U_i}$ felső háromszög mátrix pozitív diagonálissal.

Bizonyítás - folytatás

Ezek után keressük $\mathbf{A_{i+1}}$ felbontását

$$\mathbf{U_{i+1}} = \begin{bmatrix} \mathbf{U_i} & \mathbf{u} \\ \mathbf{0^T} & \beta \end{bmatrix}$$

alakban, ekkor az

$$\begin{bmatrix} \mathbf{U_i^T} & \mathbf{0} \\ \mathbf{u^T} & \beta \end{bmatrix} \begin{bmatrix} \mathbf{U_i} & \mathbf{u} \\ \mathbf{0^T} & \beta \end{bmatrix} = \begin{bmatrix} \mathbf{A_i} & \mathbf{v} \\ \mathbf{v^T} & \alpha \end{bmatrix}$$

egyenlőségnek kell teljesülnie, azaz $\mathbf{U_i}^T \mathbf{u} = \mathbf{v}$ és $\mathbf{u}^T \mathbf{u} + \beta^2 = \alpha$ adódik. Innen $\mathbf{u}$ egyértelmű, mert $\mathbf{U_i}$ reguláris (hiszen pozitív definit).

Bizonyítás - folytatás

Másrészt

$$0 < \det(\mathbf{A_{i+1}}) = \det(\mathbf{U_{i+1}})^2 = \beta^2 \det(\mathbf{U_i})^2$$

mutatja, hogy β valós, és így $\beta = \sqrt{\alpha - \mathbf{u^T u}}$. Ekkor a tételben szereplő formula a $\mathbf{v} = \mathbf{U_i u} = (a_{1,i}, \ldots, a_{i-1,i})^T$ egyenlet megoldásából adódik egyszerű visszahelyettesítéssel.

Továbbá tetszőleges reguláris $\mathbf{B} \in \mathbb{R}^{n \times n}$-es mátrixra $\mathbf{B^T B}$ pozitív definit, hiszen $\mathbf{x^T B^T B x} = ||\mathbf{Bx}||_2^2 > 0$ ha $\mathbf{x} \neq \mathbf{0}$, így csakis pozitív definit mátrixnak létezhet Cholesky-felbontása.

Cholesky-felbontás futásideje

Tétel: A Cholesky-felbontás $\frac{n^3}{3} + \mathbf{O}(n^2)$ lebegőpontos művelet kiszámításával készíthető el.

Bizonyítás: Használjuk fel a kiszámításra vonatkozó képleteket!

$$u_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} u_{ik} u_{jk}\right) / u_{jj}, \quad i = 2, \ldots n, \quad j = 1, \ldots, i-1$$

$$u_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} u_{ik}^2} \quad i = 2, \ldots n$$

u_{11} meghatározása egy művelet, u_{ij}-é $1 + 2(j-1) + 1$, végül u_{ii}-é $1 + 1 + 2(i-1)$. Azaz:

$$1 + \sum_{i=2}^{n} \sum_{j=1}^{i} 2 + 2j = 1 + \sum_{i=2}^{n} (i+2)(i+1) = \frac{n^3}{3} + \mathbf{O}(n^2)$$

Példa

Példa: Az

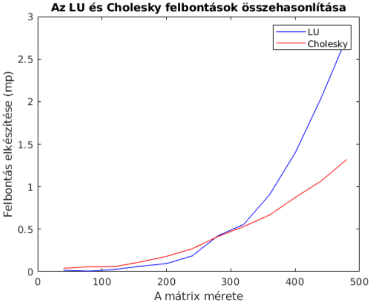
$$\mathbf{A} = \begin{bmatrix} 16 & -4 & 4 \\ -4 & 10 & -1 \\ 4 & 1 & 2 \end{bmatrix}$$

mátrix Cholesky-felbontása a CholeskyFelbontas(A) függvénnyel elkészítve:

$$\mathbf{U} = \begin{bmatrix} 4 & -1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

és így $\mathbf{U}^T\mathbf{U} = \mathbf{A}$.

A futásidő tesztelése Matlabban



A futásidőkön látszik, hogy nagy méretű mátrixokra a Cholesky–felbontás elkészítése gyorsabb, mint az LU-felbontásé. Viszont az sqrt() függvény kiszámítása több időt igényel, mint az alapműveletek, ennek hatását látjuk kis méretű mátrixokra.