Mátrix kondíciószáma, a LER érzékenysége

1. Határozzuk meg az alábbi mátrixokra a $\mathrm{cond}_1(A)$ és $\mathrm{cond}_2(B)$ mennyiségeket!

$$\text{(a)}\quad A = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \qquad\qquad \text{(b)}\quad B = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

2. Határozzuk meg az A mátrix 1-es és 2-es kondíciószámát!

$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}$$

3. Legyen $\varepsilon > 0$ kicsi pozitív szám, és definiáljuk az A mátrixot a következőképpen:

$$A = \begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix}$$

(a) Mit mondhatunk az A mátrix 1-es kondíciószámáról?

(b) Állítsuk elő az A mátrix LU-felbontását! Vizsgáljuk meg az L és U mátrixok 1-es kondíciószámát!

4. Igazoljuk, hogy a QR felbontás nem változtatja meg a lineáris egyenletrendszer érzékenységét (kondícionáltságát), azaz az $A = QR$ felbontásra

$$\mathrm{cond}_2(A) = \mathrm{cond}_2(R).$$

5. Igazoljuk, hogy szimmetrikus, pozitív definit A mátrix $A = LL^T$ Cholesky-felbontása nem változtatja meg a lineáris egyenletrendszer érzékenységét (kondícionáltságát), azaz

$$\mathrm{cond}_2(A) = \mathrm{cond}_2(L) \cdot \mathrm{cond}_2(L^T) = \left(\mathrm{cond}_2(L)\right)^2.$$

6. Igazoljuk, hogy ha A invertálható mátrix, akkor

$$\mathrm{cond}_2(A^T A) = \left(\mathrm{cond}_2(A)\right)^2.$$

MEGOLDÁS

1. Határozzuk meg az alábbi mátrixokra a $\mathrm{cond}_1(A)$ és $\mathrm{cond}_2(B)$ mennyiségeket!

$$\text{(a)}\quad A = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \qquad\qquad \text{(b)}\quad B = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

Megoldás:

Adott $A \in \mathbb{R}^n$ invertátható mátrix és $\|.\|$ mátrixnorma esetén az A kondíciószáma

$$\boxed{\mathrm{cond}(A) = \|A\| \cdot \|A^{-1}\|.}$$

(a) A kondíciószám definícióját használjuk:

$$\mathrm{cond}_1(A) = \|A\|_1 \cdot \|A^{-1}\|_1$$

Először számítsuk ki az A^{-1}-et:

$$A^{-1} = \frac{1}{(-1)\cdot 2 - 1 \cdot 0} \begin{bmatrix} 2 & 0 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

A normák

$$\|A^{-1}\|_1 = \max_{j=1,2} \left\{ |-1| + \left|\frac{1}{2}\right|, \quad |0| + \left|\frac{1}{2}\right| \right\} = \frac{3}{2}, \qquad \|A\|_1 = 2,$$

Végül

$$\mathrm{cond}_1(A) = \|A\|_1 \cdot \|A^{-1}\|_1 = 2 \cdot \frac{3}{2} = 3.$$

(b) Most $\mathrm{cond}_2(B) = \|B\|_2 \cdot \|B^{-1}\|_2$ értékét szeretnénk meghatározni. Ehhez elvileg ki kellene számolnunk a B^{-1} mátrixot és $\|B^{-1}\|_2$-t. Azonban B **szimmetrikus**, ezért a kettes kondíciószám meghatározására használhatjuk az alábbi összefüggést:

$$\boxed{\mathrm{cond}_2(B) = \varrho(B) \cdot \varrho(B^{-1}) = \frac{\max_{i=1}^2(|\lambda_i(B)|)}{\min_{i=1}^2(|\lambda_i(B)|)}}$$

A B mátrix sajátértékei 2 és 6, ezt felhasználva:

$$\mathrm{cond}_2(B) = \frac{\max\{|6|,\ |2|\}}{\min\{|6|,\ |2|\}} = \frac{6}{2} = 3.$$

2. Határozzuk meg az A mátrix 1-es és 2-es kondíciószámát!

$$\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}$$

Megoldás:

Először kiszámítjuk az A mátrix inverzét:

$$\left[\begin{array}{ccc|ccc} 3 & 1 & -1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ -1 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 3 & 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & -8 & -4 & 1 & -3 & 0 \\ 0 & 4 & 4 & 0 & 1 & 1 \end{array}\right] \sim$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 4 & 4 & 0 & 1 & 1 \\ 0 & -8 & -4 & 1 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 0 & 1 & 0 \\ 0 & 4 & 4 & 0 & 1 & 1 \\ 0 & 0 & 4 & 1 & -1 & 2 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & -1/4 & 5/4 & -2/4 \\ 0 & 4 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 1/4 & -1/4 & 2/4 \end{array}\right] \sim$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2/4 & -1/4 & 1/4 \\ 0 & 1 & 0 & -1/4 & 2/4 & -1/4 \\ 0 & 0 & 1 & 1/4 & -1/4 & 2/4 \end{array}\right] \qquad \Longrightarrow \qquad A^{-1} = \frac{1}{4}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Az A mátrix 1-es kondíciószáma

$$\mathrm{cond}_1(A) = \|A\|_1 \cdot \|A^{-1}\|_1 = 5 \cdot (\frac{1}{4} \cdot 4) = 5.$$

A szimmetrikus A mátrix 2-es kondíciószáma

$$\mathrm{cond}_2(A) = \frac{\max_{i=1}^{2}(|\lambda_i(A)|)}{\min_{i=1}^{2}(|\lambda_i(A)|)}$$

Ehhez számítsuk ki az A sajátértékeit! A

$$p(\lambda) = \begin{vmatrix} 3-\lambda & 1 & -1 \\ 1 & 3-\lambda & 1 \\ -1 & 1 & 3-\lambda \end{vmatrix}$$

determinánst az első sora szerint kifejtve

$$p(\lambda) = (3-\lambda)\left[(3-\lambda)^2 - 1^2\right] - (3-\lambda+1) - (1+3-\lambda) = (3-\lambda)(2-\lambda)(4-\lambda) - 2(4-\lambda) =$$

$$= (4-\lambda)[6+\lambda^2-5\lambda-2] = (4-\lambda)(\lambda^2-5\lambda+4) = (4-\lambda)(\lambda-4)(\lambda-1) = 0,$$

tehát a sajátértékek 1 és 4. Végül

$$\mathrm{cond}_2(A) = \frac{\max_{i=1}^{2}(|\lambda_i(A)|)}{\min_{i=1}^{2}(|\lambda_i(A)|)} = \frac{4}{1} = 4.$$

3. Legyen $\varepsilon > 0$ kicsi pozitív szám, és definiáljuk az A mátrixot a következőképpen:

$$A = \begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix}$$

(a) Mit mondhatunk az A mátrix 1-es kondíciószámáról?

Megoldás:

Számítsuk ki az A^{-1} mátrixot!

$$\det(A) = \varepsilon \cdot 1 - 1 \cdot 1 = \varepsilon - 1$$

$$A^{-1} = \frac{1}{\varepsilon - 1} \begin{bmatrix} 1 & -1 \\ -1 & \varepsilon \end{bmatrix}$$

Feltételezéseink szerint $\varepsilon > 0$ kicsi pozitív szám (azaz sokkal kisebb, mint 1), ezért

$$\|A\|_1 = \max\Big\{|\varepsilon| + |1|,\ |1| + |1|\Big\} = 2,$$

$$\|A^{-1}\|_1 = \frac{1}{|\varepsilon - 1|} \max\Big\{|1| + |-1|,\ |-1| + \varepsilon\Big\} = \frac{2}{1-\varepsilon} \approx 2.$$

Így tehát

$$\mathrm{cond}_1(A) = \|A\|_1 \cdot \|A^{-1}\|_1 \approx 2 \cdot 2 = 4.$$

(b) Állítsuk elő az A mátrix LU-felbontását! Vizsgáljuk meg az L és U mátrixok 1-es kondíciószámát!

Megoldás:

Az A mátrix LU felbontását GE segítségével állítjuk elő:

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \quad \xrightarrow[S_2 - \frac{1}{\varepsilon} S_1]{} \quad \begin{bmatrix} \varepsilon & 1 \\ 0 & 1 - \frac{1}{\varepsilon} \end{bmatrix},$$

tehát

$$L = \begin{bmatrix} 1 & 0 \\ \frac{1}{\varepsilon} & 1 \end{bmatrix}, \qquad U = \begin{bmatrix} \varepsilon & 1 \\ 0 & 1 - \frac{1}{\varepsilon} \end{bmatrix}.$$

A kondíciószám meghatározásához számítsuk ki az L és U mátrixok inverzét:

$$L^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{\varepsilon} & 1 \end{bmatrix} \qquad U^{-1} = \frac{1}{\varepsilon - 1} \begin{bmatrix} 1 - \frac{1}{\varepsilon} & -1 \\ 0 & \varepsilon \end{bmatrix}$$

Ezután számítsuk ki (és becsüljük meg) az L, L^{-1}, U, U^{-1} mátrixok 1-es normáját:

$$\|L\|_1 = \max\Big\{|1| + |\tfrac{1}{\varepsilon}|,\ |0| + |1|\Big\} = 1 + \frac{1}{\varepsilon} \approx \frac{1}{\varepsilon},$$

$$\|L^{-1}\|_1 = \max\Big\{|1| + |-\tfrac{1}{\varepsilon}|,\ |0| + |1|\Big\} = 1 + \frac{1}{\varepsilon} \approx \frac{1}{\varepsilon},$$

$$\|U\|_1 = \max\Big\{|\varepsilon| + |0|,\ |1| + |1 - \tfrac{1}{\varepsilon}|\Big\} = 1 + \frac{1}{\varepsilon} - 1 = \frac{1}{\varepsilon},$$

$$\begin{aligned} \|U^{-1}\|_1 &= \frac{1}{|\varepsilon - 1|} \max\Big\{|0| + |1 - \tfrac{1}{\varepsilon}|,\ |-1| + |\varepsilon|\Big\} = \\ &= \frac{1}{1-\varepsilon}\left(\frac{1}{\varepsilon} - 1\right) = 1 \cdot \frac{1}{\varepsilon} = \frac{1}{\varepsilon}. \end{aligned}$$

Végül a kondíciószámok:

$$\operatorname{cond}_1(L) = \|L\|_1 \cdot \|L^{-1}\|_1 \approx \frac{1}{\varepsilon^2},$$

$$\operatorname{cond}_1(U) = \|U\|_1 \cdot \|U^{-1}\|_1 = \frac{1}{\varepsilon^2}.$$

Tudjuk, hogy a lineáris egyenletrendszerek perturbációs tételei szerint a LER numerikus megoldásának hibája a LER mátrixának kondíciószámával közelítőleg egyenesen arányos. Az $Ax = b$ LER „jól kondícionált”, hiszen $\operatorname{cond}_1(A) = 4$ meglehetősen kicsi. Viszont, ha az LU felbontáson keresztül számítjuk a megoldást, akkor az $Ly = b$ és $Ux = y$ egyenletrendszerek mindegyike rosszul kondícionált, hiszen ha ε kicsi pozitív szám, akkor $\frac{1}{\varepsilon^2}$ nagyon nagy pozitív szám, és az imént láttuk, hogy $\operatorname{cond}_1(L) \approx \operatorname{cond}_1(U) = \frac{1}{\varepsilon^2}$.

4. Igazoljuk, hogy a QR felbontás nem változtatja meg a lineáris egyenletrendszer érzékenységét, azaz az $A = QR$ felbontásra

$$\operatorname{cond}_2(A) = \operatorname{cond}_2(R).$$

Megoldás:

Mivel $A = QR$, az inverze $A^{-1} = R^{-1}Q^{-1}$.

A Q mátrix ortogonalitása miatt Q^{-1} is ortogonális mátrix, amiből

$$\|A\|_2 = \|QR\|_2 = \|R\|_2, \quad \text{és} \quad \|A^{-1}\|_2 = \|R^{-1}Q^{-1}\|_2 = \|R^{-1}\|_2,$$

végül

$$\operatorname{cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \|R\|_2 \cdot \|R^{-1}\|_2 = \operatorname{cond}_2(R).$$

5. Igazoljuk, hogy szimmetrikus, pozitív definit A mátrix $A = LL^T$ Cholesky-felbontása nem változtatja meg a lineáris egyenletrendszer érzékenységét (kondícionáltságát), azaz

$$\operatorname{cond}_2(A) = \operatorname{cond}_2(L) \cdot \operatorname{cond}_2(L^T) = (\operatorname{cond}_2(L))^2.$$

Megoldás:

A szimmetrikus, pozitív definit A mátrixnak egyértelműen létezik az $A = LL^T$ Cholesky-felbontása.

$$A = LL^T \quad \Longrightarrow \quad L^{-1}A = L^T \quad \Longrightarrow \quad L^{-1}AL = L^TL, \quad \text{és} \quad L^{-1}(LL^T)L = L^TL$$

miatt az A, az L^TL és az LL^T mátrixok hasonló mátrixok. Ez azt jelenti, hogy a sajátértékeik megegyeznek, és ezáltal a spektrálsugaraik is azonosak:

$$\varrho(A) = \varrho(L^TL) = \varrho(LL^T). \tag{$*$}$$

Ekkor

$$\|L\|_2^2 \underset{\text{def}}{=} \varrho(L^TL) \underset{\text{hasonlóság}}{=} \varrho(LL^T) = \varrho\left((L^T)^TL^T\right) \underset{\text{def}}{=} \|L^T\|_2^2 \quad \Longrightarrow \quad \|L\|_2 = \|L^T\|_2.$$

Most megismételjük a fenti gondolatmenetet az A mátrix inverzére.

$$A^{-1} = (LL^T)^{-1} = (L^T)^{-1}L^{-1} \quad \Longrightarrow \quad A^{-1}L = (L^{-1})^T \quad \Longrightarrow \quad L^{-1}A^{-1}L = L^{-1}(L^{-1})^T,$$

és

$$L\left(L^{-1}(L^{-1})^T\right)L^{-1} = (L^{-1})^T L^{-1}$$

miatt az A^{-1}, az $(L^{-1})^T L^{-1}$ és az $L^{-1}(L^{-1})^T$ mátrixok hasonló mátrixok. Ez azt jelenti, hogy a sajátértékeik megegyeznek, és ezáltal a spektrálsugaraik is azonosak:

$$\varrho(A^{-1}) = \varrho\left((L^{-1})^T L^{-1}\right) = \varrho\left(L^{-1}(L^{-1})^T\right). \tag{$**$}$$

Ekkor

$$\|L^{-1}\|_2^2 \underset{\text{def}}{=} \varrho((L^{-1})^T L^{-1}) \underset{\text{hasonlóság}}{=} \varrho(L^{-1}(L^{-1})^T) = \varrho\left(\left((L^{-1})^T\right)^T (L^{-1})^T\right) \underset{\text{def}}{=} \|(L^{-1})^T\|_2^2,$$

azaz

$$\|L^{-1}\|_2 = \|(L^{-1})^T\|_2.$$

A fentiek alapján

$$\text{cond}_2(L) = \|L\|_2 \cdot \|L^{-1}\|_2 = \|L^T\|_2 \cdot \|(L^T)^{-1}\|_2 = \text{cond}_2(L^T).$$

Mivel A mátrix és inverze is szimmetrikus, a spektrálnormájuk a spektrálsugarukkal egyenlő:

$$\|A\|_2 = \varrho(A) \underset{(*)}{=} \|L\|_2^2, \qquad \text{és} \qquad \|A^{-1}\|_2 = \varrho(A^{-1}) \underset{(**)}{=} \|L^{-1}\|_2^2,$$

és végül

$$\begin{aligned}\text{cond}_2(A) &= \|A\|_2 \cdot \|A^{-1}\|_2 = \|L\|_2^2 \cdot \|L^{-1}\|_2^2 = \|L\|_2 \cdot \|L^T\|_2 \cdot \|L^{-1}\|_2 \cdot \|(L^{-1})^T\|_2 = \\ &= \left(\|L\|_2 \cdot \|L^{-1}\|_2\right)\left(\|L^T\|_2 \cdot \|(L^T)^{-1}\|_2\right) = \text{cond}_2(L) \cdot \text{cond}_2(L^T),\end{aligned}$$

illetve

$$\text{cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \|L\|_2^2 \cdot \|L^{-1}\|_2^2 = \left(\text{cond}_2(L)\right)^2.$$

6. Igazoljuk, hogy ha A invertálható mátrix, akkor

$$\text{cond}_2(A^T A) = \left(\text{cond}_2(A)\right)^2.$$

Megoldás:

$$\text{cond}_2(A^T A) = \|A^T A\|_2 \cdot \|(A^T A)^{-1}\|_2$$

Mivel $A^T A$ szimmetrikus,

$$\|A^T A\|_2 \underset{A^T A \text{ szim.}}{=} \varrho(A^T A) \underset{\text{def}}{=} \|A\|_2^2.$$

Most kiszámítjuk az $(A^T A)^{-1}$ mátrix 2-es (spektrál) normáját.

$$A\left(A^{-1}(A^{-1})^T\right)A^{-1} = (A^{-1})^T A^{-1} \quad \Longrightarrow \quad \varrho(A^{-1}(A^{-1})^T) = \varrho((A^{-1})^T A^{-1})$$

miatt

$$\|(A^T A)^{-1}\|_2^2 = \varrho\left((A^T A)^{-1}\right) = \varrho(A^{-1}(A^T)^{-1}) = \varrho(A^{-1}(A^{-1})^T) \underset{\text{hasonlóság}}{=} \varrho((A^{-1})^T A^{-1}) \underset{\text{def}}{=} \|A^{-1}\|_2^2.$$

Végül

$$\text{cond}_2(A^T A) = \|A^T A\|_2 \cdot \|(A^T A)^{-1}\|_2 = \|A\|_2^2 \cdot \|A^{-1}\|_2^2 = \left(\|A\|_2 \cdot \|A^{-1}\|_2\right)^2 = \left(\text{cond}_2(A)\right)^2.$$