

KÖZELÍTŐ ÉS SZIMBOLIKUS SZÁMÍTÁSOK

6. GYAKORLAT

Sajátérték, Gersgorin körök

Készítette:

Gelle Kitti

Csendes Tibor

Somogyi Viktor

Vinkó Tamás

London András

Deák Gábor

jegyzetei alapján

1. Mátrixok sajátértékei és sajátvektorai

Legyen adott egy A négyzetes mátrix. Adjuk meg a λ számot és az $\mathbf{x} \neq \mathbf{0}$ vektort úgy, hogy

$$A\mathbf{x} = \lambda\mathbf{x} \quad (1)$$

legyen. Ezt nevezzük az A mátrix *sajátérték egyenletének*. Itt λ az A mátrix *sajátértéke*, \mathbf{x} pedig a *sajátvektora*. Egy hasonló definícióval értelmezhető a baloldali sajátérték és sajátvektor is:

$$\mathbf{y}^T A = \lambda\mathbf{y}^T.$$

A továbbiakban csak az (1) egyenletet fogjuk vizsgálni, mivel a bal oldali sajátérték visszavezethető a jobb oldalra. Ezen definícióhoz további fontos fogalom kapcsolódik. Az első a $\lambda(A)$ -val jelölt *mátrix spektruma*, ami a sajátértékeinek a halmaza, a második pedig a *spektrálsugara*, ami a legnagyobb abszolútértékű sajátértékének abszolútértéke, azaz: $\rho(A) = \max\{|\lambda| : \lambda \in \lambda(A)\}$.

Megadható egy szemléletes geometriai értelmezése a sajátértéknek és a sajátvektornak. Az (1) egyenletünk bal oldalán az A mátrix egy geometriai transzformációs mátrix. Ez azt jelenti, hogy ez például egy forgatást, eltolást vagy skálázást (esetleg komplexebb, az előbbiekből összerakott egyéni transzformációt, pl. skálázást és forgatást egyszerre, vagy valami egyéb lineáris transzformációt, például nyírást) ír le. Amennyiben ezzel a mátrixszal beszorunk valamilyen vektort, az annyit tesz, mintha alkalmaznánk ezt a transzformációt a vektorra.

Ez az egyenlet tehát azt fejezi ki, hogy az egyes sajátvektorok irányába eső vektorokat a transzformáció egyszerűen megnyújtja a sajátértékeknek megfelelően.

A sajátértékeket, sajátvektorokat meghatározhatjuk az

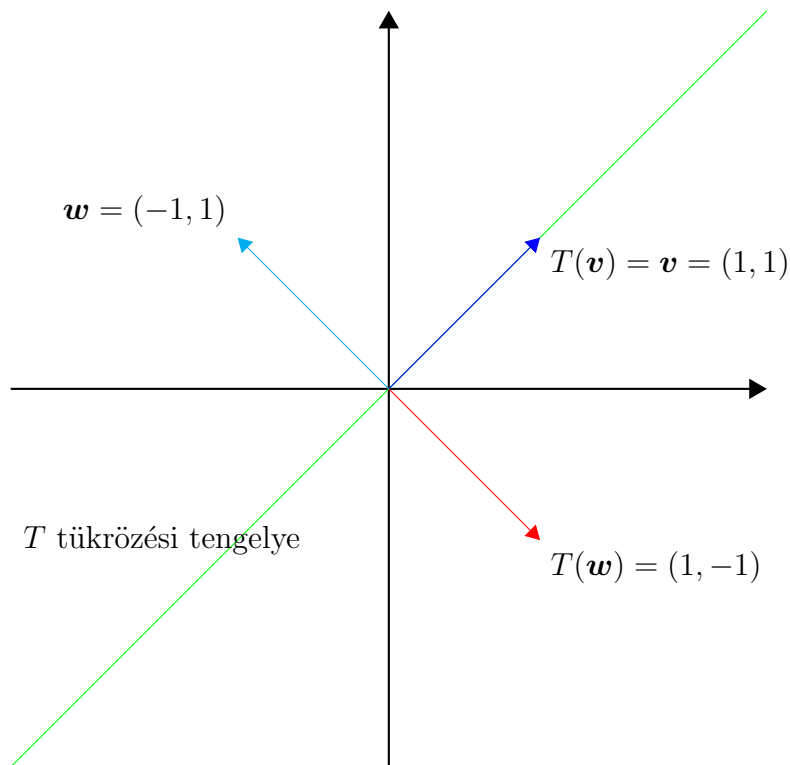
$$(A - \lambda I)\mathbf{x} = \mathbf{0} \quad (2)$$

homogén lineáris egyenletrendszer megoldásával. Ennek pontosan akkor van a nullvektortól különböző megoldása, ha az $A - \lambda I$ mátrix szinguláris. Emiatt a sajátértékeket leírja a

$$\det(A - \lambda I) = 0 \quad (3)$$

egyenlet. Ennek a bal oldalán levő n -edfokú polinomot az A mátrix *karakterisztikus polinomjának* nevezzük (példát látunk mindjárt erre is). A mátrix karakterisztikus polinomjának gyökei lesznek a mátrix sajátértékei.

Az 1. ábrán láthatunk egy T transzformációt (amely egy tengelyes tükrözés), ennek sajátvektorait, illetve sajátvektorokon elvégzett transzformációkat. Azt is észrevehetjük, hogy az elvárt viselkedést tanúsítja, azaz egy számmal való skálázással azonos a sajátvektorok irányában történő transzformáció.



1. **Ábra.:** A T transzformáció, és sajátvektorai

A T -hez tartozó A mátrix a következő: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. A fenti (3) egyenlettel határozzuk meg a sajátértékeit:

$$\det(A - \lambda I) = \det \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) = \det \left(\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} \right) =$$

$$= \lambda^2 - 1 = 0 \implies \lambda_1 = 1 \text{ és } \lambda_2 = -1.$$

A hozzá tartozó sajátvektorok pedig a (2) egyenlet megoldásaként keletkeznek. Ezeket minden λ értékre el kell végezni:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda_i & 0 \\ 0 & \lambda_i \end{pmatrix} \begin{pmatrix} x_1^{(i)} \\ x_2^{(i)} \end{pmatrix} = 0 \implies \begin{pmatrix} -\lambda_i & 1 \\ 1 & -\lambda_i \end{pmatrix} \begin{pmatrix} x_1^{(i)} \\ x_2^{(i)} \end{pmatrix} = 0$$

Behelyettesítve $\lambda_1 = 1$ -et:

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = 0$$

Egyenletrendszeres formában:

$$\begin{aligned} -x_1^{(1)} + x_2^{(1)} &= 0 \\ x_1^{(1)} - x_2^{(1)} &= 0 \end{aligned}$$

Ezen jól látszik, hogy az első egyenlet nem független a másodiktól, tehát bármely értéket választhatok x_1 -nek (ekkor ugyanannyi lesz x_2 is). Az ábrához hűen most maradjunk az 1-nél, tehát

$$\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

A λ_2 esetében hasonló a helyzet, de ott $x_1^{(2)} = -x_2^{(2)}$ jön ki, ami azt jelenti, hogy a sajátvektorok elemei egymás ellentettjei. Tehát a -1 -hez tartozó sajátvektorok felírhatók a következő formában:

$$\begin{pmatrix} x_1^{(2)} \\ -x_2^{(2)} \end{pmatrix}, \text{ például: } \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

2. Mátrixok sajátértékeinek algoritmikus meghatározása

A sajátértékszámítás egyik kulcsfogalma a hasonlóság, amelyre a numerikus módszereinket is építjük.

Definíció 2.1. Legyen A tetszőleges mátrix, T pedig tetszőleges reguláris mátrix. Az

$$A \mapsto T^{-1}AT$$

hozzárendelést hasonlósági transzformációnak nevezzük. Azt mondjuk, hogy A hasonló a B mátrixhoz (jelben: $A \sim B$), ha létezik olyan reguláris T mátrix, hogy $B = T^{-1}AT$.

Tétel 2.1. Legyenek $A, B \in \mathbb{C}^{n \times n}$ tetszőleges mátrixok.

1. Ha $A \sim B$, akkor sajátértékeik megegyeznek.
2. Ha $A \sim B$ és A egy sajátvektorrendszere $\{u_1, u_2, \dots, u_n\}$ akkor B egy sajátvektorrendszere fölírható $\{v_1, v_2, \dots, v_n\}$ formában, ahol $v_j = T^{-1}u_j$ bármely $1 \leq j \leq n$ esetén.

Abból, hogy az A és a B mátrix sajátértékei megegyeznek, még nem következik, hogy hasonlóak is.

Mivel ötöd- vagy magasabb-fokú polinomra nincs megoldóképlet, így ennek a gyökeit valamilyen közelítő algoritmussal kell meghatározni. Erre két módszert nézünk meg.

2.1 LR transzformáció

Az LR transzformáció egy, az LR felbontáson alapuló módszer, amely egy mátrixsorozatot ad meg. Ezzel a mátrixsorozattal meghatározhatóak A sajátértékei.

Az eljárás az $A_1 = A$ mátrixból indul, és a következő két lépésből áll:

$$A_k = L_k R_k,$$

$$A_{k+1} = R_k L_k.$$

Tehát az aktuális mátrixot felbontjuk egy alsó-és egy felsőtrianguláris mátrixra, majd a következő iteráció mátrixát ezek fordított sorrendben való összeszorzásával kapjuk. Az L_k alsó háromszögmátrix főátlójában egyesek vannak.

Legyen $T_k = L_1 L_2 \dots L_k$ és $U_k = R_k R_{k-1} \dots R_1$. Ekkor érvényes a következő

Segédteétel 2.1. *Ha minden $k = 1, 2, \dots$ indexre létezik a fent definiált A_k mátrix, akkor*

1. $A_k = (L_1 L_2 \dots L_{k-1})^{-1} A_1 (L_1 L_2 \dots L_{k-1}) = T_{k-1}^{-1} A_1 T_{k-1}$ (így tehát az $A = A_1$ mátrix hasonló mindegyik A_k -hoz) és
2. $A_1^k = T_k U_k$.

Fontos tehát, hogy minden lépésben az A -hoz hasonló mátrixot állítunk elő, ami azt jelenti, hogy az A_k mátrix sajátértékei megegyeznek az eredeti A mátrix sajátértékeivel.

Tétel 2.2. *(Az LR transzformáció konvergencia-tétele) Ha az A valós négyzetes mátrix teljesíti az alábbi feltételeket:*

1. az $A_k = L_k R_k$ trianguláris felbontás létezik minden $k \geq 1$ -re,
2. az A sajátértékeit alkalmasan indexelve érvényes $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$ (amiből adódik, hogy A reguláris és sajátértékei mind valósak és egyszerűek), és
3. megadható az A mátrixot diagonalizáló olyan X mátrix, amellyel egyrészt $X^{-1} A X = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, másrészt létezik mind az $X = \tilde{L} \tilde{R}$, mind az $X^{-1} = \bar{L} \bar{R}$ trianguláris felbontás, akkor az $\{A_k\}$, $\{R_k\}$ és az $\{L_k\}$ mátrix sorozatok mind konvergensek:

$$\lim_{k \rightarrow \infty} A_k = \lim_{k \rightarrow \infty} R_k = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad \text{és} \quad \lim_{k \rightarrow \infty} L_k = I$$

Az első két sorozat tehát egy olyan mátrixhoz tart, amelynek főátlójában az A mátrix sajátértékei vannak nagyság szerint rendezve.

Az LR algoritmus megvalósításához bármilyen felbontást használhatunk, ami egy alsó és egy felsőtrianguláris mátrixra bontja fel az eredeti A mátrixot. Ennek ellenére az algoritmus legfőbb hátránya, hogy még erős feltételek teljesülése esetén is előfordulhat, hogy valamely k -ra az A_k mátrix nem bontható fel $L_k R_k$ alakban.

2.2 QR-algoritmus

A leggyakrabban alkalmazott sajátértékszámító módszer az ortogonális-trianguláris fölbontást felhasználó QR-algoritmus. Ehhez az $A_1 \in \mathbb{R}^{n \times n}$ mátrixból kiindulva képezzük az

$$\begin{aligned} A_k &= Q_k R_k \\ A_{k+1} &= R_k Q_k \end{aligned}$$

rekurzív definícióval adott $\{A_k\}$ mátrixsorozatot. Mivel tetszőleges A_k mátrix ortogonális-trianguláris fölbontása létezik, a sorozat tagjai bármely A mátrix esetén képezhetők tetszőleges QR faktorizációs eljárás felhasználásával. A kapott mátrixsorozat legfontosabb tulajdonságait a következő segéd-tételben foglaljuk össze.

Segéd-tétel 2.2. *Legyenek $V_k = Q_1 Q_2 \dots Q_k$ ortogonális és $U_k = R_k R_{k-1} \dots R_1$ felső trianguláris mátrixok. Ekkor bármely $k \geq 1$ -re igaz*

1. $A_{k+1} = (Q_1 Q_2 \dots Q_k)^T A_1 (Q_1 Q_2 \dots Q_k) = V_k^T A_1 V_k$, vagyis az A_k mátrix hasonló A_1 -hez (hiszen $V_k^T = V_k^{-1}$, mert V_k is ortogonális), továbbá
2. $A_1^k = V_k U_k$.

Ahogy az előzőben, itt is egy fontos tulajdonság, hogy a QR-algoritmus minden lépésében a kiindulási mátrixhoz hasonló mátrixot kapunk, tehát sajátértékeik megegyeznek.

Tétel 2.3. *(A QR-algoritmus konvergencia-tétele) Ha az $A_1 \in \mathbb{R}^{n \times n}$ mátrix teljesíti az alábbi feltételeket:*

1. az A_1 sajátértékeit alkalmasan indexezve $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$
2. megadható az A_1 -et diagonalizáló olyan X mátrix, amellyel

$$X^{-1} A_1 X = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n);$$

továbbá létezik az $X^{-1} = \bar{L}\bar{R}$ trianguláris fölbontás, akkor a QR-algoritmus által generált $\{A_k\}$ mátrixsorozat konvergens:

$$\lim_{k \rightarrow \infty} A_k = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

A QR algoritmus előnye, hogy bármely mátrix esetén előállítható annak ortogonális-trianguláris felbontása, viszont a fenti tétel alapján látható, hogy a konvergenciát itt sem minden mátrix esetén tudjuk biztosítani.

3. Gersgorin körök

Tétel 3.1. (Gersgorin). Az A mátrix összes sajátértéke benne van a

$$K_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{k=1, k \neq i}^n |a_{ik}| \right\}$$

körök $K = \cup_{i=1}^n K_i$ egyesítésében.

Példa. Vegyük a következő mátrixot:

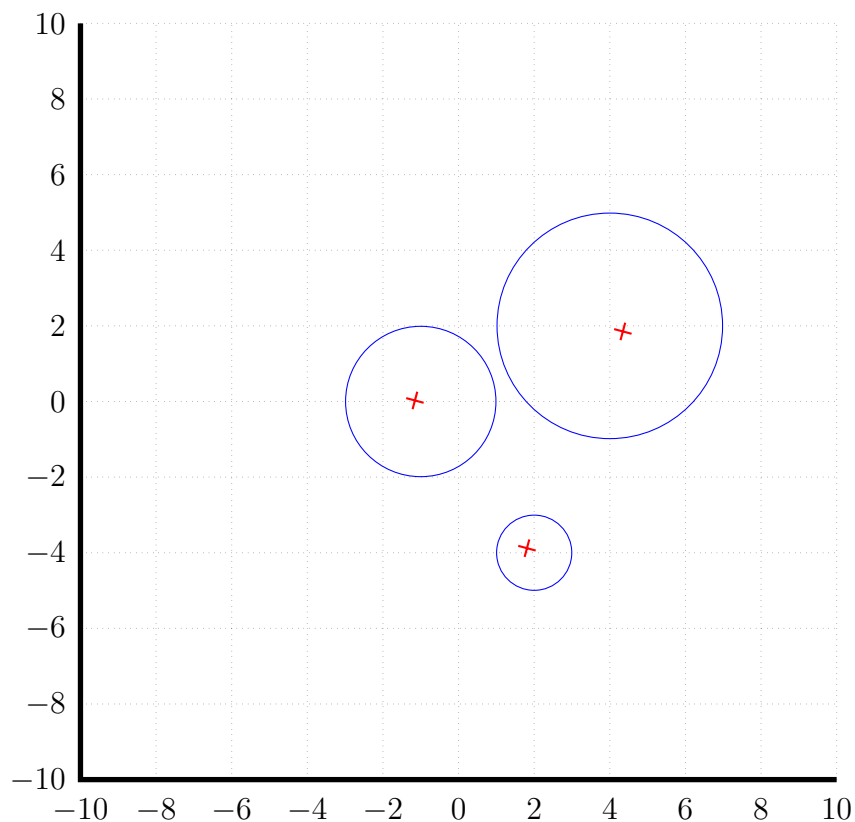
$$A = \begin{pmatrix} -1 & 1 & i \\ 2 & 4 + 2i & 1 \\ 1 & 0 & 2 - 4i \end{pmatrix}$$

Itt a következő köröket tudjuk felírni:

$$K_1 = |z + 1| \leq |1| + |i| \implies K_1 = |z + 1| \leq 2$$

$$K_2 = |z - (4 + 2i)| \leq 2 + 1 \implies K_2 = |z - (4 + 2i)| \leq 3$$

$$K_3 = |z - (2 - 4i)| \leq 1$$



2. Ábra.: A Gersgorin körök és a bennük levő sajátértékek

Tehát a körök középpontjait a főátlóbeli elemekből kapjuk, míg a sugarakat az adott sorban lévő elemek abszolút értékeinek összegeiből. Az A sajátértékei:

$$\lambda_1 = -1.1542 + 0.0272i$$

$$\lambda_2 = 4.3452 + 1.8526i$$

$$\lambda_3 = 1.8089 - 3.8798i$$

Ha ábrázoljuk a köröket a komplex számsíkon, akkor tényleg látszik, hogy a sajátértékek benne vannak a körök egyesítésében (2. Ábra).