

NUMERIKUS MÓDSZEREK

BEVEZETÉS A NUMERIKUS MÓDSZEREKBE

ELŐADÁS JEGYZET – 2020. NOVEMBER 17.

KÉSZÍTETTE: DR. NAGY NOÉMI

Mátrixok sajátértékei, sajátvektorai

Definíció: Legyen $A \in \mathbb{C}^{n \times n}$ tetszőleges mátrix. A $\lambda \in \mathbb{C}$ számot az A mátrix sajátértékének, az $x \in \mathbb{C}^n$ ($x \neq 0$) vektor pedig a λ sajátértékhez tartozó sajátvektornak nevezzük ha $Ax = \lambda x$ teljesül. Azaz a sajátvektort az $x \mapsto Ax$ leképezés saját hatásvonalán hagyja.

Alkalmazások:

- főirányok, melyekben csak húzófeszültségek ébrednek, melyek nagysága éppen a sajátérték
- Page Rank algoritmus internetes oldalak rangsorolására.

Definíció: A $\det(A - \lambda E) = 0$ egyenletet az A mátrix karakterisztikus egyenletének nevezzük. A determináns kifejtésével kapott n -edfokú polinomot pedig karakterisztikus polinomnak.

Példa:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}, \quad \det(A - \lambda E) = \begin{vmatrix} 1 - \lambda & 3 \\ 2 & 2 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda) - 3 \cdot 2 = \lambda^2 - 3\lambda - 4 = 0.$$

$$\lambda_1 = 4, \lambda_2 = -1.$$

A karakterisztikus polinom zérushelyei a mátrix sajátértékei.

Tétel: Egy $A \in \mathbb{C}^{n \times n}$ mátrixnak a multiplicitásokat is figyelembe véve pontosan n sajátértéke van.

Tétel: Ha x a λ sajátértékhez tartozó sajátvektor, akkor minden $t \in \mathbb{C} \setminus \{0\}$ esetén $t \cdot x$ is az.

Definíció: Egy mátrix sajátértékeinek összességét a mátrix spektrumának nevezzük.

Definíció: Legyen $A \in \mathbb{C}^{n \times n}$. A spektrálnormája:

$$\|A\|_2 = \sqrt{A^T \cdot A \text{ vagy } A \cdot A^T \text{ legnagyobb sajátértéke}}$$

Példa:

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \quad A^T \cdot A = \begin{bmatrix} 5 & 7 \\ 7 & 13 \end{bmatrix}$$
$$0 = \begin{vmatrix} 5 - \lambda & 7 \\ 7 & 13 - \lambda \end{vmatrix} = (5 - \lambda)(13 - \lambda) - 7 \cdot 7 = \lambda^2 - 18\lambda + 16.$$

$$\lambda_{1,2} = \frac{18 \pm \sqrt{18^2 - 4 \cdot 16}}{2} = \frac{18 \pm \sqrt{260}}{2} = \frac{18 \pm 16.1245}{2}$$

$$\lambda_1 = 17.06225 \quad \lambda_2 = 0.93775$$

$$\|A\|_2 = \sqrt{17.06225} = 4.1306$$

Tételek a sajátértékekkel kapcsolatban

- (1) A sajátértékek szorzata a mátrix determinánsával egyezik meg.
- (2) A sajátértékek összege a mátrix nyoma (trace), azaz a főátlóbeli elemek összege: $\text{tr}(A) = \lambda_1 + \dots + \lambda_n$.
- (3) Egy mátrix akkor és csak akkor invertálható, ha egyik sajátértéke sem zérus.
- (4) Egy mátrix akkor és csak akkor pozitív definit, ha minden sajátértéke pozitív.
- (5) Ha az A mátrix sajátértékei $\lambda_1, \lambda_2, \dots, \lambda_n$, akkor A^k sajátértékei $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$.
Ugyanis $A^k v = A(A(\dots A(Av))) = \lambda^k v$, mivel minden egyes lépésben egy Av tag átírható λv alakba belülről kifelé haladva.
- (6) Ha A invertálható és λ egy sajátértéke, x pedig a λ -hoz tartozó sajátvektor, akkor A^{-1} -nek egy sajátértéke $\frac{1}{\lambda}$ és x a hozzá tartozó sajátvektor.
Mivel $Av = \lambda v$, ezért $v = A^{-1} \lambda v$, melyből $\frac{1}{\lambda} v = A^{-1} v$.
- (7) Diagonális és háromszög mátrix esetén a sajátértékek megegyeznek a főátlóbeli elemekkel.

A sajátértékek elhelyezkedése

Tétel: Legyen λ az A mátrix tetszőleges sajátértéke. Indukált mátrixnorma esetén fennáll a $|\lambda| \leq \|A\|$ egyenlőtlenség.

Bizonyítás: $|\lambda| \|x\| = \|\lambda x\| = \|Ax\| \leq \|A\| \|x\|$. Mivel $\|x\| \neq 0$, ezért oszthatunk vele.

Gersgorin tétele: Legyen

$$A \in \mathbb{C}^{n \times n}, \quad r_i = \sum_{j=1, j \neq i}^n |a_{ij}|$$

és $D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$ $i = 1, 2, \dots, n$ (Gersgorin-körök). Ekkor az A mátrix minden λ sajátértékére fennáll, hogy $\lambda \in \bigcup_{i=1}^n D_i$.

Bizonyítás: Legyen $v \neq 0$ egy tetszőleges λ sajátértékhez tartozó sajátvektor. Az $Av = \lambda v$ egyenletrendszernek vegyük azt az egyenletét, ahol $|v_i|$ maximális:

$$\lambda v_i = \sum_{j=1}^n a_{ij} v_j = a_{ii} v_i + \sum_{j=1, j \neq i}^n a_{ij} v_j.$$

$$|\lambda v_i - a_{ii}v_i| = |(\lambda - a_{ii})v_i| = |\lambda - a_{ii}| |v_i| = \left| \sum_{j=1, j \neq i}^n a_{ij}v_j \right| \leq \left| \sum_{j=1, j \neq i}^n a_{ij}v_i \right| \leq \sum_{j=1, j \neq i}^n |a_{ij}| |v_i| = |v_i| \sum_{j=1, j \neq i}^n |a_{ij}|$$

Mindkét oldalt osztva $|v_i|$ -vel:

$$|\lambda - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}| = r_i,$$

tehát λ benne van a D_i körlemezben, azaz a D_i -k egyesítése (uniója) az összes sajátértéket tartalmazza.

A hatványmódszer

A sajátértékeket nehéz kiszámítani, illetve gyakran nincs is szükség az összes sajátértékre, hanem csak egyre. A Von Miesestől származó eljárás a legnagyobb abszolút értékű (domináns) sajátértékek és a hozzá tartozó sajátvektorok előállítására szolgál.

Tegyük fel, hogy $A \in \mathbb{R}^{n \times n}$ sajátértékei valós számok, és egyetlen legnagyobb sajátérték van: $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$.

Jelölje x_i a λ_i -hez tartozó sajátvektort, és legyen $v^{(0)}$ olyan vektor, mely felírható a sajátvektorok lineáris kombinációjaként úgy, hogy x_1 is szerepel a felírásban:

$$v^{(0)} = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

ahol $\alpha_1 \neq 0$.

Képezzük a $v^{(k)} = Av^{(k-1)} = A^k v^{(0)}$ vektorsorozatot:

$$v^{(1)} = Av^{(0)} = A(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) = \alpha_1 Ax_1 + \alpha_2 Ax_2 + \dots + \alpha_n Ax_n = \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \dots + \alpha_n \lambda_n x_n$$

Illetve

$$v^{(k)} = Av^{(k-1)} = \alpha_1 \lambda_1^k x_1 + \alpha_2 \lambda_2^k x_2 + \dots + \alpha_n \lambda_n^k x_n = \lambda_1^k \left(\alpha_1 x_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right)$$

Igazolható, hogy ha $v^{(k)}$ i-edik koordinátája nem 0, akkor

$$\frac{v_i^{(k+1)}}{v_i^{(k)}} \rightarrow \lambda_1,$$

a $v^{(k)}$ sorozat pedig a λ_1 -hez tartozó sajátvektorhoz tart.

Megjegyzés: A hatványmódszer algoritmus a $|\lambda_2/\lambda_1|$ nagyságrendjétől függő konvergencia sebességgel rendelkezik. A módszer erősen érzékeny a $v^{(0)}$ kezdővektor megválasztására is.

Az inverz hatványmódszer

Ha A^{-1} -re alkalmazzuk a hatványmódszert, akkor A^{-1} domináns sajátértékének meghatározása után reciprokvétellel kapjuk A legkisebb abszolút értékű sajátértékét, hiszen korábban szerepelt, hogy ha A sajátértékei $|\lambda_1| < |\lambda_2| < \dots < |\lambda_n|$, akkor A^{-1} sajátértékei:

$$\left| \frac{1}{\lambda_1} \right| > \left| \frac{1}{\lambda_2} \right| > \dots > \left| \frac{1}{\lambda_n} \right|$$

A QR-módszer

Definíció: Legyenek $A, B \in \mathbb{C}^{n \times n}$. Azt mondjuk, hogy az A és B mátrixok hasonlóak, ha van olyan $T \in \mathbb{C}^{n \times n}$ invertálható mátrix, hogy $B = T^{-1}AT$.

Tétel: Hasonló mátrixok sajátértékei megegyeznek. Ha x az A mátrix sajátvektora, akkor $y = T^{-1}x$ a B sajátvektora.

Bizonyítás: $Ax = \lambda x$ -ből következik, hogy $T^{-1}Ax = \lambda T^{-1}x$, melyet kiegészítve mindkét oldalon $T T^{-1}$ -gyel az x tag előtt: $T^{-1}AT T^{-1}x = \lambda T^{-1}T T^{-1}x$ adódik.

Csoportosítva ezeket: $(T^{-1}AT)(T^{-1}x) = \lambda(T^{-1}T)(T^{-1}x)$.

Felhasználva, hogy $B = T^{-1}AT$: $B(T^{-1}x) = \lambda(T^{-1}x)$.

Ha bevezetjük az $y = T^{-1}x$ jelölést, azt kapjuk, hogy $By = \lambda y$, vagyis $y = T^{-1}x$ valóban a B sajátvektora.

Tétel (négyzetes mátrix QR-felbontása): Minden $A n \times n$ négyzetes mátrix előáll $A = QR$ alakban, ahol Q ortogonális mátrix (azaz $Q^T Q = E$) és R felső háromszögmátrix.

Egy mátrix QR-felbontását leggyakrabban Gram-Schmidt ortogonalizációs eljárással határozzák meg.

A QR-módszer: Általános eljárás egy mátrix összes sajátértékeinek meghatározására.

$A_1 = A = Q_1 R_1$ (nem elfelejtendő, hogy a Q -k ortogonálisak, a R mátrixok pedig felső háromszögmátrixok).

Az $A_2 = R_1 Q_1$ mátrix hasonló A_1 -hez, azaz A -hoz, mert $R_1 = Q_1^{-1} A_1$ és $A_2 = Q_1^{-1} A_1 Q_1$.

Ezért az $A_k = R_{k-1} Q_{k-1} = Q_k R_k$ ($k = 2, 3, \dots$) sorozat minden eleme hasonló A -hoz, így a sorozat elemeinek ugyanazok a sajátértékei.

Ezt a sorozatot használhatjuk a sajátértékek meghatározására: bizonyos feltételek mellett az A_k sorozat főátlóbeli elemei a sajátértékekből álló vektorokhoz közelítenek:

$$A_1 = A = Q_1 R_1$$

$$A_2 = R_1 Q_1 = Q_2 R_2$$

$$A_3 = R_2 Q_2 = Q_3 R_3.$$

Tehát minden lépésben felbontjuk az aktuális ortogonális és felső háromszögmátrix szorzatára, melyet a következő lépésben fordított sorrendben szorzunk össze. Az így kapott szorzatot újra felbontjuk. Az így kapott sorozat főátlóbeli elemei elemenként tartanak az eredeti A mátrix sajátértékeihez.