Sajátérték-problémák

2019. március 4.

Az alapfeladat

Adott a következő egyenlet:

$$\mathbf{Av} = \lambda \mathbf{v}, \tag{1}$$

ahol

- $\mathbf{A}$ egy ismert mátrix
- $\mathbf{v}$ ismeretlen, nem zérus vektor
- λ ismeretlen szám

Azok a $\mathbf{v}, \lambda$ kombinációk, amikre az egyenlet teljesül az $\mathbf{A}$ mátrix sajátvektorai és sajátértékei.

- minden sajátvektor számszorosára is igaz (1)
- ezeket nem tekintjük külön sajátvektornak
- a sajátvektorok ezért normalizálhatók
- a $\mathbf{v} = 0$ triviális megoldás nem sajátvektor

A karakterisztikus egyenlet

Az eredeti egyenlet akkor és csak akkor teljesül, ha

$$\det |\mathbf{A} - \lambda \mathbf{I}| = 0$$

is teljesül.

- ez egy polinomiális egyenlet λ-ra
- csak $N < 4$ esetben érdemes ezt megoldani
- lehetnek nem valós és degenerált gyökök is
- nagyobb mátrixok esetében más módszerek lesznek

Sajátvektorok

Megkülönböztetünk jobb és baloldali sajátvektorokat

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

$$\mathbf{u}\mathbf{A} = \lambda\mathbf{u}$$

- $\mathbf{v}$: jobboldali sajátvektor, oszlopvektor
- $\mathbf{u}$: baloldali sajátvektor, sorvektor
- a kétoldali sajátvektorokhoz tartozó sajátértékek megegyeznek
- valós szimmetrikus mátrixra a sajátvektorok is

Négyzetes mátrixok állatkertje

Szimmetrikus mátrix	$\mathbf{A} = \mathbf{A}^\mathsf{T}$	$a_{ij} = a_{ji}$
Hermitikus mátrix	$\mathbf{A} = \mathbf{A}^\mathsf{H}$	$a_{ij} = a_{ji}^*$
Ortogonális mátrix	$\mathbf{A}^\mathsf{T} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^\mathsf{T} = \mathbf{I}$	
Normálmátrix	$\mathbf{A}^\mathsf{H} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^\mathsf{H}$	
Unitér mátrix	$\mathbf{A}^\mathsf{H} = \mathbf{A}^{-1}$	

Valós mátrixokra

- hermitikus $\equiv$ szimmetrikus
- unitér $\equiv$ ortogonális

Speciális mátrixok tulajdonságai

Hermitikus mátrix: $\mathbf{A} = \mathbf{A}^{\mathrm{H}}$

- sajátértékei mind valósak

Nem szimmetrikus valós mátrix

- sajátértékei valósak vagy komplex konjugált párok

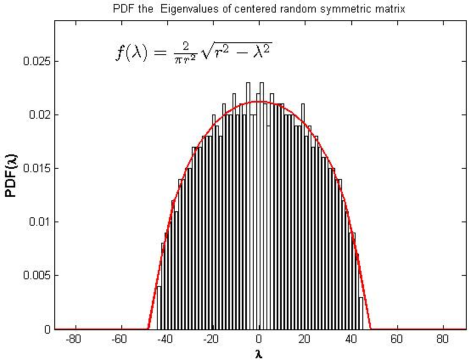
Általános komplex mátrix

- sajátértékei tetszőleges komplex számok

Érdekesség: Wigner-féle félkörtörvény

Valós, szimmetrikus mátrix random elemekkel feltöltve

- A sajátértékek valószínűségi eloszlása egy félkör, $N \to \infty$ esetben
- Viszonylag kevés megkötés az egyes elemek valószínűségi eloszlására:
- Azonos eloszlásból, függetlenül választva



- A sajátértékek eloszlása random mátrixokra tehát jól meghatározott
- Algoritmusok tesztelésére a random mátrixok nem jók!

Sajátvektorokból alkotott bázis

Ha egy normálmátrix sajátértékei mind különbözőek, akkor

- sajátvektorai ortogonális bázist alkotnak
- normálás után: $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$
- degenerált sajátérték esetén saját altér van
- ortogonalizálható: pl. Gram–Schmidt-eljárás

Általános mátrix esetében

- a sajátvektorok általában nem ortogonálisak
- nem biztos, hogy teljes bázist alkotnak
- de a legtöbbször azért a bázis teljes

Gram–Schmidt-ortogonalizáció

$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{|\mathbf{u}_1|} \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2), & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{|\mathbf{u}_2|} \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3), & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{|\mathbf{u}_3|} \\
\mathbf{u}_4 &= \mathbf{v}_4 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_3}(\mathbf{v}_4), & \mathbf{e}_4 &= \frac{\mathbf{u}_4}{|\mathbf{u}_4|} \\
&\vdots & &\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k), & \mathbf{e}_k &= \frac{\mathbf{u}_k}{|\mathbf{u}_k|}.
\end{aligned}
$$

Vektor vetítése másik vektorra:

$$
\mathrm{proj}_{(\mathbf{b})}\mathbf{a} = |\mathbf{a}| \cos\theta \cdot \frac{\mathbf{b}}{|\mathbf{b}|} = \left(\mathbf{a} \cdot \frac{\mathbf{b}}{|\mathbf{b}|}\right) \cdot \frac{\mathbf{b}}{|\mathbf{b}|} \tag{2}
$$

Jobb és baloldali sajátvektorok mátrixa

Vezessük be a következő mátrixokat:

- $\mathbf{U}$: baloldali sorvektorok alkotta mátrix
- $\mathbf{V}$: jobboldali oszlopvektorok alkotta mátrix
- $\boldsymbol{\lambda}$: a sajátértékekből alkotott diagonális mátrix

Ezzel a jelöléssel a sajátérték-egyenletek:

$$\begin{aligned}\mathbf{U} \cdot \mathbf{A} &= \boldsymbol{\lambda} \cdot \mathbf{U} \\ \mathbf{A} \cdot \mathbf{V} &= \mathbf{V} \cdot \boldsymbol{\lambda}\end{aligned}$$

Bal és jobboldali sajátvektorok ortogonalitása

A sajátértékegyenletek:

$$\begin{aligned} \mathbf{U} \cdot \mathbf{A} &= \boldsymbol{\lambda} \cdot \mathbf{U} \\ \mathbf{A} \cdot \mathbf{V} &= \mathbf{V} \cdot \boldsymbol{\lambda} \end{aligned}$$

Az elsőt balról **V**-vel, a másodikat jobbról **U**-val szorozva, majd a két egyenlet különbségét képezve:

$$(\mathbf{UV})\boldsymbol{\lambda} = \boldsymbol{\lambda}(\mathbf{UV})$$

Különböző elemekből álló diagonális mátrixszal, mint pl. $\boldsymbol{\lambda}$, csak olyan mátrix kommutál, ami maga is diagonális.

- következmény: **UV** diagonális
- következmény: $\mathbf{u}_i \mathbf{v}_j = \delta_{ij}$
- persze **u**-k és **v**-k megfelelően normálandók

Invertálható mátrix faktorizációja

Ha létezik egy **A** mátrix inverze, akkor a következő áll fenn a bal- és jobboldali sajátvektorokból képzett **U** és **V** mátrixokra:

$$\mathbf{U}^{-1} = \mathbf{V}$$

Ha létezik a mátrix inverze, akkor a következőképpen faktorizálható

$$\begin{aligned} \mathbf{A} &= \mathbf{V}\boldsymbol{\lambda}\mathbf{V}^{-1} \\ \mathbf{A}^{-1} &= \mathbf{V}\boldsymbol{\lambda}^{-1}\mathbf{V}^{-1} \end{aligned}$$

illetve igaz, hogy

$$\boldsymbol{\lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$$

Invertálható mátrix diagonalizálása

Előző eredmény:

$$\boldsymbol{\lambda} = \mathbf{UAV}$$

- viszont tudjuk:

$$\mathbf{U} = \mathbf{V}^{-1}$$

- vagyis:

$$\mathbf{V}^{-1}\mathbf{AV} = \boldsymbol{\lambda}$$

Valós szimmetrikus mátrixokra:

- fenn áll, hogy $\mathbf{V}^{-1} = \mathbf{U} = \mathbf{V}^{\mathsf{T}}$
- így a **V**-vel való szendvicselés ortogonális transzformáció is

Általános esetben ezt a „szendvicselést” hasonlósági transzformációnak nevezzük:

$$\mathbf{A} \quad \rightarrow \quad \mathbf{Z}^{-1}\mathbf{A}\mathbf{Z},$$

ahol $\mathbf{Z}$ nyilvánvalóan invertálható

A sajátérték-problémát megoldó programok ezen alapulnak

- megfelelően választott $\mathbf{P}_i$ transzformációkkal
- az $\mathbf{A}$ mátrixot egyre közelítjük a diagonálishoz:

$$\mathbf{A} \quad \rightarrow \quad \mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1 \quad \rightarrow \quad \mathbf{P}_2^{-1}\mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1\mathbf{P}_2 \quad \rightarrow \quad \dots \quad \rightarrow \quad \boldsymbol{\lambda}$$

Ekkor a sajátvektorokat a következő módon kapjuk

$$\mathbf{V} = \mathbf{P}_1 \cdot \mathbf{P}_2 \cdot \dots$$

Jacobi-transzformáció

Keressük a **P** transzformációkat síkbeli forgatások alakjában:

$$\boldsymbol{P}_{pq} = \begin{pmatrix} 1 & & & & & & \\ & \cdots & & & & & \\ & & c & \cdots & s & & \\ & & \vdots & 1 & \vdots & & \\ & & -s & \cdots & c & & \\ & & & & & \cdots & \\ & & & & & & 1 \end{pmatrix}$$

ahol a c és s számokra fennáll, hogy $c^2 + s^2 = 1$

- ez a mátrix csak a pq, qp, pp és qq helyeken nem 0 vagy 1
- szendvicseljük vele **A**-t, és számoljuk végig

Egyenlet a Jacobi-transzformáció paramétereire

Elvégezzük a beszorzást[1]

- kiszámoljuk az $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ mátrix elemeit
- csak a p. és q. sorok, illetve oszlopok elemei változnak
- tegyük a'_{pq}-t nullává!
- a következő egyenletre jutunk

$$\operatorname{ctg} 2\phi = \frac{c^2 - s^2}{2sc} = \frac{a_{qq} - a_{pp}}{2a_{pq}}$$

[1] házi feladat

$\mathbf{A}' = \mathbf{P}^{-1}\mathbf{AP}$ kiszámítása

$$
\begin{aligned}
a'_{rp} &= ca_{rp} - sa_{rq} \qquad \text{ha} \qquad r \neq p, r \neq q \\
a'_{rq} &= ca_{rq} + sa_{rp} \qquad \text{ha} \qquad r \neq p, r \neq q \\
a'_{pp} &= c^2 a_{pp} + s^2 a_{qq} - 2sca_{pq} \\
a'_{qq} &= s^2 a_{pp} + c^2 a_{qq} + 2sca_{pq} \\
a'_{pq} &= \left(c^2 - s^2\right) a_{pq} + sc\left(a_{pp} - a_{qq}\right) = a'_{qp}
\end{aligned}
$$

A Jacobi-transzformáció elemi lépése

Az előbb felvázolt lépéssel a mátrix a_{pq} elemét nulláztuk

- ezt a lépés tetszőlegesen sokszor ismételhetjük
- az újabb lépés a korábban lenullázott elemeket ugyan elrontja,
- de belátható:
  a nem diagonális elemek négyzetösszege monoton csökken
- vagyis a mátrix tart a diagonálishoz

A Jacobi-transzformáció tulajdonságai

- érdemes mindig a legnagyobb nem diagonális elemet nullázni
- az algoritmus N^2 elem nullázását igényli
- de többször végig kell menni az egész mátrixon

A hatványiteráció

Keressük egy mátrix *legnagyobb* sajátértékéhez tartozó sajátvektorát

- a sajátvektorai teljes ortogonális bázist alkotnak
- a teljes sajátérték-probléma megoldása nem feltétlen éri meg
- gyorsabb, iteratív megoldást keresünk

Ha $\mathbf{x}_0$ tetszőleges vektor

- megszorozzuk az $\mathbf{A}$ mátrixszal, majd
- leosztjuk a kapott vektort a legnagyobb elemével

$$\mathbf{y}_0 = \frac{1}{c_0}\mathbf{A}\mathbf{x}_0, \quad \text{ahol} \quad c_0 = \max_i \left(\mathbf{A}\mathbf{x}_0\right)^{(i)}$$

- ezután képezzük a következő iterációt

$$\mathbf{x}_k = \mathbf{y}_{k-1}$$

A hatványiteráció konvergenciája

Fejtsük ki az $\mathbf{x}_0$ vektort a sajátvektorok bázisán:

$$\mathbf{x}_0 = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n$$

Mivel a $\mathbf{v}$ vektorok sajátvektorok, ezért az $\mathbf{A}$ mátrix hatása:

$$\mathbf{A}\mathbf{x}_0 = \lambda_1 \alpha_1 \mathbf{v}_1 + \lambda_2 \alpha_2 \mathbf{v}_2 + \ldots + \lambda_n \alpha_n \mathbf{v}_n$$

Most fejtsük ki az $\mathbf{x}_k$ vektort a sajátvektorok bázisán:

$$\mathbf{x}_k = \frac{1}{c_1 c_2 \ldots c_k} \left[\lambda_1^k \alpha_1 \mathbf{v}_1 + \lambda_2^k \alpha_2 \mathbf{v}_2 + \ldots + \lambda_n^k \alpha_n \mathbf{v}_n \right]$$

Emeljük ki a legnagyobb sajátérték k. hatványát:

$$\mathbf{x}_k = \frac{\lambda_1^k}{c_1 c_2 \ldots c_k} \left[\alpha_1 \mathbf{v}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \ldots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{v}_n \right]$$

A hatványiteráció konvergenciája

Előző diáról:

$$\mathbf{x}_k = \frac{\lambda_1^k}{c_1 c_2 \ldots c_k} \left[\alpha_1 \mathbf{v}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \ldots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{v}_n \right]$$

- ahol λ_1 a legnagyobb sajátérték
- emiatt az iteráció során az összes együttható kihal
- kivéve α_1-et
- a sorozat a legnagyobb sajátértékhez tartozó sajátvektorhoz konvergál

$$\begin{aligned} \lim_{k \to \infty} \mathbf{x}_k &= \mathbf{v}_1 \\ \lim_{k \to \infty} c_k &= \lambda_1 \end{aligned}$$

Fontos tétel

Az első egyenlet mindkét oldalához $\tau\mathbf{v}$-et adva:

$$\mathbf{Av} + \tau\mathbf{v} = \lambda\mathbf{v} + \tau\mathbf{v}$$

- a sajátértékek eltolódnak, de
- a sajátvektorok változatlanok, ugyanis

$$(\mathbf{A} + \tau\mathbf{I}) \cdot \mathbf{v} = (\lambda + \tau) \cdot \mathbf{v}$$

Ezt a tulajdonságot sok numerikus módszer kihasználja

Inverz hatványiteráció

Gyakran előfordul, hogy megsejtjük valamelyik sajátértéket

- hogyan találhatjuk meg a hozzá tartozó sajátvektort?

Tekintsük a következő egyenletet

$$(\mathbf{A} - \tau \mathbf{I}) \cdot \mathbf{y}_0 = \mathbf{b}_0$$

- legyen $\mathbf{b}_0$ tetszőleges vektor
- τ pedig közel van valamelyik λ sajátértékhez
- $\mathbf{y}_0$ a fenti egyenlet megoldása

Belátjuk, hogy $\mathbf{b}_n = \mathbf{y}_{n-1}$ helyettesítéssel

- az iteráció konvergens
- $\mathbf{b}_n$ a λ-hoz tartozó sajátvektorhoz tart

Az inverz hatványiteráció konvergenciája

Fejtsük ki $\mathbf{y}$-t és $\mathbf{b}$-t az $\mathbf{v}_i$ sajátvektorok alkotta bázison:

$$\mathbf{y}_0 = \sum_j \alpha_j \mathbf{v}_j \qquad\qquad \mathbf{b}_0 = \sum_j \beta_j \mathbf{v}_j$$

Ezzel az egyenlet:

$$\sum_j \alpha_j (\lambda_j - \tau) \mathbf{v}_j = \sum_j \beta_j \mathbf{v}_j,$$

vagyis

$$\alpha_j = \frac{\beta_j}{\lambda_j - \tau} \qquad \text{és} \qquad \mathbf{y}_0 = \sum_j \frac{\beta_j \mathbf{v}_j}{\lambda_j - \tau}$$

Az így kapott $\mathbf{y}_0$-t helyettesítjük be $\mathbf{b}_0$ helyére, ebből:

$$\mathbf{y}_1 = \sum_j \frac{\beta_j \mathbf{v}_j}{(\lambda_j - \tau)^2}, \quad \text{stb.}$$

Inverz hatványiteráció

Előző diáról:

$$\mathbf{y}_1 = \sum_j \frac{\beta_j \mathbf{v}_j}{(\lambda_j - \tau)^2}, \quad \text{stb.}$$

A további iterációk során a nevezőben $\lambda_j - \tau$ hatványai jelennek meg:

$$\mathbf{y}_k = \sum_j \frac{\beta_j \mathbf{v}_j}{(\lambda_j - \tau)^{(k-1)}}$$

- Ha τ közel van a j. sajátértékhez, azaz $|\lambda_j - \tau| \ll |\lambda_i - \tau|, \quad \forall i \neq j$, akkor
- a szummában a j. tag fog dominálni
- vagyis $\mathbf{y}_k \to \mathbf{v}_j$