(KEMNA0302) Alkalmazott lineáris algebra

Dr. Facskó Gábor, PhD
tudományos főmunkatárs
facskog@gamma.ttk.pte.hu

Pécsi Tudományegyetem, Természettudományi Kar, Matematikai és Informatikai Intézet, 7624 Pécs, Ifjúság útja 6.
Wigner Fizikai Kutatóközpont, Űrfizikai és Űrtechnikai Osztály, 1121 Budapest, Konkoly-Thege Miklós út 29-33.
https://facsko.ttk.pte.hu

2025. április 17.

Ütemterv

- **Hátralévő órák: 2025. április 17, 30, május 7.**
- Szinguláris érték, szinguláris vektor, SVD, PCA.
- ~~Mátrixok összehasonlítása, pozitív mátrixok, nemnegatív mátrixok, irreducibilis mátrixok, SMRC, NMF.~~
- Reakcióegyenletek sztöchiometrikus rendezése.
- ~~Lineáris programozási feladatok mátrixaritmetikai megoldhatósága.~~
- ~~Power of matrices. Applications: linear recursions, power of incidence matrixes.~~
- ~~Gram-Schmidt ortogonalization.~~
- ~~Further applications.~~

Szinguláris érték, szinguláris vektor, SVD I

- (Szinguláris érték) A szimmetrikus mátrix ortogonális diagonalizálásának általánosítása tetszőleges mátrixra, egy helyett két ortonormált bázis megkeresésével.
- Információtömörítésre, egyenletrendszerek megoldására használják.
- Az ortogonális diagonalizáció általánosítása, A sajátértékek szerepét a szinguláris értékek veszik át, a sajátfelbontásét a szinguláris érték szerinti felbontás (SVD).
- Egy $\mathbf{A} \in \mathbb{R}^{m \times n}$ mátrixhoz tartozó mátrixleképezés kölcsönösen egyértelmű a sortér és az oszloptér között.
- Ebben a két altérben keresünk ortonormált bázisokat, melyek közt a leképezés mátrixa diagonális alakot ölt.
- Ezeket kiegészítjük az $\mathbb{R}^n$ és az $\mathbb{R}^m$ ortonormált bázisává.
- A komplex terekben mindez hasonlóképpen megvalósítható.

Szinguláris érték, szinguláris vektor, SVD II

- Az $\mathbf{A} \in \mathbb{R}^{m \times n}$ mátrixhoz olyan ortonormált $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subset \mathbb{R}^n$ és $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^m$ bázisokat keresünk, melyekben $\mathbf{A}$ mátrixa diagonálissá válik. Azaz, léteznek olyan σ_i valós értékek, hogy $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$, ahol $1 \leq i \leq min(m, n)$. Azaz, olyan egymásra merőleges vektorokat keresünk, melyek képei is merőlegesek egymásra.
- Állítás: Ha az egymásra merőleges $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektorok legalább egyike az $\mathbf{A}^T\mathbf{A}$ sajátvektora, akkor az $\mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \in \mathbb{R}^m$ vektorok is merőlegesek egymásra.
  Bizonyítás: A Wettl jegyzetben.
- Az $A : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ leképezés kölcsönösen egyértelmű a sortér és az oszloptér között. Így először csak e két altérben keressünk megfelelő bázist. Közös dimenziójuk a ranggal egyenlő, jelölje ezt r.

Szinguláris érték, szinguláris vektor, SVD III

- Mivel $\mathbf{A}^T\mathbf{A}$ szimmetrikus és pozitív szemidefinit, ezért ortogonálisan diagonalizálható és a sajátértékei nem negatívak. A sajátvektorokból kiválasztható $\mathbb{R}^n$ egy ortonormált bázisa, melyek közül a pozitív sajátértékekhez tartozók kifeszítik a nulltér, azaz az $\mathcal{N}\left(\mathbf{A}^T\mathbf{A}\right) = \mathcal{N}(\mathbf{A})$ tér merőleges kiegészítő alterét, tehát a sorteret.
- Jelölje a sortér e bázisát $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. E vektorok egyike sincs a nulltérben, és $\mathbf{A}$ általi képeik páronként merőlegesek egymásra, így az $\mathbf{A}\mathbf{v}_i$ vektorok ortogonális bázist alkotnak az oszloptérben. Ha $\mathbf{v}_i$ az $\mathbf{A}^T\mathbf{A}$ mátrix $\lambda_i > 0$ sajátértékéhez tartozó egységnyi sajátvektora, akkor $|\mathbf{A}\mathbf{v}_i| = \lambda_i$, ugyanis

$$|\mathbf{A}\mathbf{v}_i|^2 = \mathbf{v}_i^T\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \lambda_i\,|\mathbf{v}_i|^2 = \lambda_i.$$

  Legyen $\sigma_i = \sqrt{\lambda_i}$ és $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$. Ekkor az $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ ortonormált bázist alkotnak az oszloptérben.

Szinguláris érték, szinguláris vektor, SVD IV

- Definíció: (Szinguláris érték). Az r rangú $\mathbf{A} \in \mathbb{R}^{m \times n}$ mátrix szinguláris értékeinek nevezzük az $\mathbf{A}^T\mathbf{A}$ mátrix pozitív sajátértékeinek négyzetgyökeit, melyeket $\sigma_1, \sigma_2, \ldots, \sigma_r$ jelöl. Az $\mathbf{A}^T\mathbf{A}$ mátrix valamely σ^2 sajátértékéhez tartozó egységnyi $\mathbf{v}$ sajátvektorát az $\mathbf{A}$ jobb szinguláris vektorának, az $\mathbf{u} = \frac{1}{\sigma}\mathbf{A}\mathbf{v}$ egységvektort pedig a $\mathbf{v}$-hez tartozó bal szinguláris vektorának nevezzük.
- A szinguláris érték különösen akkor hasznos eszköz például az egyenletrendszerek megoldásában, amikor az együtthatómátrix szinguláris.
- E definíció egyik fontos következménye, hogy mivel $|\mathbf{u}_i| = 1$, ahol $(i = 1, 2, \ldots, r)$, ezért $|\mathbf{A}\mathbf{v}_i| = \sigma_i$.
- Ha σ t-szeres multiplicitású szinguláris értéke $\mathbf{A}$-nak, akkor σ_2 az $\mathbf{A}^T\mathbf{A}$-nak t-szeres algebrai (és t-szeres geometriai) multiplicitású sajátértéke, így a σ-hoz tartozó jobb szinguláris vektorok a sortérben, a hozzá tartozó bal szinguláris vektorok pedig az oszloptérben t-dimenziós alteret feszítenek ki.

Szinguláris érték, szinguláris vektor, SVD V

- Ha $\mathbf{Av} = \sigma\mathbf{u}$, akkor $\mathbf{A}^T\mathbf{u} = \frac{1}{\sigma}\mathbf{A}^T\mathbf{Av} = \frac{1}{\sigma}\sigma^2\mathbf{v} = \sigma\mathbf{v}$, tehát az
$$\mathbf{Av} = \sigma\mathbf{u}, \ \mathbf{A}^T\mathbf{u} = \sigma\mathbf{v}$$
összefüggések párba állítják a jobb és bal szinguláris vektorokat.
- Ha $\mathbf{A} \in \mathbb{C}^{n\times n}$, akkor a sortér helyett a sortér konjugáltjával,azaz az $\mathcal{O}\left(\mathbf{A}^H\right)$ altérrel, transzponálás helyett adjungálással, tehát $\mathbf{A}^T\mathbf{A}$ helyett az $\mathbf{A}^H\mathbf{A}$ mátrixszal igazak maradnak a fentiek.
- Wettle jegyzet 10.3-as példa.
- (Szinguláris felbontás) Egy mátrix szinguláris értékei és vektorai egy mátrixfelbontást adnak, ezt fogjuk szinguláris érték szerinti felbontásnak nevezni.

Szinguláris érték, szinguláris vektor, SVD VI

- Képezzük a szinguláris értékekből a diagonális

$$\Sigma_1 = diag\left(\sigma_1, \ldots, \sigma_r\right) = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \end{pmatrix}$$

mátrixot, valamint a szinguláris vektorokból az $\mathbf{U}_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ és a $\mathbf{V}_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ mátrixokat. Ekkor az $\mathbf{A}\mathbf{v}_i = \sigma\mathbf{u}_i$ egyenlőségek az

$$\mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\mathbf{\Sigma}_1,$$

azaz az

Szinguláris érték, szinguláris vektor, SVD VII

$$\mathbf{A}(\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r) = (\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_r)\begin{pmatrix}\sigma_1 & 0 & \ldots & 0\\ 0 & \sigma_2 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \sigma_r\end{pmatrix}$$

alakot öltik. Mivel $\mathbf{V}_1$ szemiortogonális, és oszlopvektorai a sortér ONB-át alkotják, ezért $\mathbf{V}_1\mathbf{V}_1^T$ a sortérre való merőleges vetítés mátrixa, így $\mathbf{A}\mathbf{V}_1\mathbf{V}_1^T = \mathbf{A}$, ezért jobbról szorozva a fenti egyenletet $\mathbf{V}_1^T$ -tal kapjuk, hogy $\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T$.

- Definíció: (Redukált szinguláris felbontás és diadikus alakja). A valós (komplex) $\mathbf{A}$ mátrix

$$\mathbf{A} = \mathbf{U}_1\boldsymbol{\Sigma}_1\mathbf{V}_1^T \quad \left(\mathbf{A} = \mathbf{U}_1\boldsymbol{\Sigma}_1\mathbf{V}_1^H\right)$$

felbontását redukált szinguláris felbontásnak nevezzük, ha Σ_1 négyzetes,

Szinguláris érték, szinguláris vektor, SVD VIII

diagonális mátrix, főátlójában monoton csökkenően rendezett pozitív valós számokkal, $\mathbf{U}_1$ és $\mathbf{V}_1$ szemiortogonálisak (komplex esetben $\mathbf{U}_1^H\mathbf{U}_1 = \mathbf{V}_1^H\mathbf{V}_1 = \mathbf{E}_r$). Ha ezt a szorzatot az $\mathbf{U}_1$ oszlopvektorokra és a $\mathbf{V}_1^T$ sorvektorokra blokkosított alakjára írjuk fel, akkor az $\mathbf{A}$ mátrix egy diadikus felbontását kapjuk, melyet szinguláris érték szerinti diadikus felbontásnak nevezünk:

$$\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T .$$

(komplex esetben transzponált helyett adjungálttal).

⋮

Szinguláris érték, szinguláris vektor, SVD IX

- Definíció: (Szinguláris felbontás). A valós (komplex) $\mathbf{A}$ mátrix

$$\mathbf{A} = \mathbf{U\Sigma V}^T \quad \left(\mathbf{A} = \mathbf{U\Sigma V}^H\right)$$

alakú felbontását A szinguláris érték szerinti felbontásának, vagy röviden szinguláris felbontásának nevezzük, ha $\mathbf{U}$ és $\mathbf{V}$ ortogonális (unitér) és Σ diagonális, főátlójában monoton csökkenően rendezett nem negatív valós számokkal.
- Wettl jegyzet: 10.6-os példa.
- Tétel: (Az SVD létezése és Σ egyértelműsége). Minden valós vagy komplex mátrixnak létezik szinguláris érték szerinti felbontása. A szinguláris értékek monoton csökkenő sorozata egyértelmű, de a felbontás nem.
Bizonyítás: Wettl jegyzet.

Szinguláris érték, szinguláris vektor, SVD X

- ~~Szinguláris felbontás geometriai interpretációja~~
- ~~Egységgömb képe~~
- ~~Polárfelbontás~~
- ~~Pszeudoinverz~~

Principal Component Analysis I

- Principal Component Analysis (PCA), avagy főkomponens-analízis vagy főkomponens-elemzés egy többváltozós statisztikai eljárás, mely az adatredukciós módszerek közé sorolható.
- Lényege, hogy egy nagy adathalmaz - melynek változói kölcsönös kapcsolatban állnak egymással - dimenzióit lecsökkentse, miközben a jelen lévő varianciát a lehető legjobban megtartsa.
- Ezt úgy hajtja végre, hogy egy ortogonális (merőleges) transzformáció segítségével az adathalmaz lehetségesen korreláltatható változóit lineárisan korrelálatlan változók értékkészletévé alakítja át, melyeket főkomponenseknek nevezünk.
- A főkomponensek száma kisebb vagy egyenlő az eredeti változók számával.

Principal Component Analysis II

- A transzformáció oly módon meghatározott, hogy az első főkomponens rendelkezik a lehető legnagyobb varianciával, majd minden utána következő komponens a fennmaradó legnagyobb varianciával fog rendelkezni, ha megfelel annak a feltételnek, hogy merőleges (azaz korrelálatlan) az azt megelőző komponensekre.
- A főkomponensek merőlegesek, mivel a kovarianciamátrix - ami szimmetrikus - sajátvektorai.
- A főkomponens-analízis érzékeny az eredeti változók relatív skálázására.
- A főkomponens-analízis felfogható úgy is, mint ha egy n dimenziós ellipszoidot próbálnánk az adatokra illeszteni, ahol az ellipszoid mindegyik tengelye egy főkomponens lenne. Ha az ellipszoid valamelyik tengelye kicsi, akkor a tengely menti variancia is kicsi lesz, és ha elhagyjuk ezt a tengelyt és a hozzá tartozó főkomponenst az adathalmaz reprezentációjából, akkor csak egy ezzel arányosan kis adatmennyiséget veszítünk el.

Principal Component Analysis III

- Ahhoz, hogy megtaláljuk az ellipszoid tengelyeit, elsőként ki kell vonnunk minden változó átlagát az adathalmazból, hogy az adatokat az origó köré igazítsuk. Ezután kiszámítjuk az adatok kovarianciamátrixát és a sajátértékeket, illetve a kovarianciamátrixhoz tartozó sajátvektorokat. Utána merőlegesítjük (ortogonalizáljuk) és normalizáljuk a sajátvektor halmazt (ONB) hogy egységvektorokat kapjunk. Ezt követően a kapott, kölcsönösen merőleges egységvektorokra úgy tekinthetünk, mint az ellipszoid adatra illesztett tengelyeire. A varianciák azon hányada, melyet az egyes sajátvektorok képviselnek kiszámítható úgy, hogy a sajátvektorhoz tartozó sajátértéket elosztjuk az összes sajátérték összegével.
- Ez az eljárás érzékeny az adatok skálázására, s nem létezik konszenzus arra vonatkozólag, hogy hogyan kell az adatokat skálázni az optimális eredmények eléréséhez.

Principal Component Analysis IV

- A főkomponens-analízis menete:
  1. Bevezetjük az adatokat és létrehozzuk az adathalmazt. Legyenek x és y.
  2. Kivonjuk az átlagokat minden adat dimenzióból, vagyis minden x értékből kivonjuk a hozzá tartozó $\overline{X}$ átlagot, illetve minden y értékből a hozzá tartozó $\overline{Y}$ átlagot.
  3. Kiszámoljuk a kovariáns mátrixot a fenti képlet alapján. Ez esetben a kovariáns mátrix 2×2-es lesz.
  4. Kiszámoljuk a kovarianciamátrixhoz tartozó sajátvektorokat és sajátértékeket. A főkomponens-analízis esetében fontos, hogy a sajátvektorok egységvektorok legyenek (hosszuk 1). A sajátvektoroknak egymásra merőlegeseknek kell lenniük.
  5. Kiválasztjuk a komponenseket és kialakítjuk a vonás vektort. A sajátvektorokat sajátérték szerint sorrendbe állítjuk, legnagyobbtól a legkisebbig, megkapva így a komponenseket szignifikanciájuk szerint. Ha a sajátértékek kellően kicsit, akkor az adatmennyiség amit elveszítünk ezzel arányos mértékben szintén kicsi lesz. Ezt követően létrehozzuk a vonásvektort a megtartott sajátvektorokból.

Principal Component Analysis V

6. Leszármaztatjuk az új adathalmazt. Ez úgy történik, hogy az előző lépésben kapott vonás vektort, ami oszlopvektor, transzponáljuk, vagyis átalakítjuk sorvektorrá úgy, hogy a legszignifikánsabb sajátvektor lesz az első érték. Az így kapott vektort megszorozzuk jobbról az eredeti adathalmaz transzponáltjával.

- Kovariancia: két változó közötti kapcsolatot adja meg.
- Kovariancia mátrix: 3D-s vektornál 3×3-as mátrix.

Vége

Köszönöm a figyelmüket!