Szinguláris érték felbontás és függvényillesztés

Kormányos Andor

Komplex Rendszerek Fizikája Tanszék

2020. február 20.

Szinguláris érték szerinti felbontás
Singular value decomposition (SVD)

Tetszőleges $M \times N$-es ($M \geq N$) $\mathbf{A}$ mátrix felírható a következő formában:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^H$$

vagyis

$$\left(\quad \mathbf{A} \quad \right) = \left(\quad \mathbf{U} \quad \right) \cdot \begin{pmatrix} w_1 & & \\ & \dots & \\ & & w_N \end{pmatrix} \cdot \left(\quad \mathbf{V}^H \quad \right)$$

ahol

- $\mathbf{U}$ $M \times N$-es, $\mathbf{V}$ $N \times N$-es unitér mátrixok: $\mathbf{U}^H \cdot \mathbf{U} = \mathbf{1}$ ($\mathbf{U}^H = (\mathbf{U}^T)^*$), $\mathbf{V}^H \cdot \mathbf{V} = \mathbf{1}$. Mivel $\mathbf{V}^H$ négyzetes, ezért $\mathbf{V} \cdot \mathbf{V}^H = \mathbf{1}$ is igaz.
- $\mathbf{W}$ diagonális mátrix és $w_i \geq 0$ valós számok.

Szinguláris érték szerinti felbontás
Singular value decomposition (SVD)

Tetszőleges $M \times N$-es ($M \geq N$) $\mathbf{A}$ mátrix felírható a következő formában:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^H$$

vagyis

$$\left(\quad \mathbf{A} \quad \right) = \left(\quad \mathbf{U} \quad \right) \cdot \begin{pmatrix} w_1 & & \\ & \dots & \\ & & w_N \end{pmatrix} \cdot \left(\quad \mathbf{V}^H \quad \right)$$

ahol

- $\mathbf{U}$ $M \times N$-es, $\mathbf{V}$ $N \times N$-es unitér mátrixok: $\mathbf{U}^H \cdot \mathbf{U} = \mathbf{1}$ $(\mathbf{U}^H = (\mathbf{U}^T)^*)$, $\mathbf{V}^H \cdot \mathbf{V} = \mathbf{1}$. Mivel $\mathbf{V}^H$ négyzetes, ezért $\mathbf{V} \cdot \mathbf{V}^H = \mathbf{1}$ is igaz.
- $\mathbf{W}$ diagonális mátrix és $w_i \geq 0$ valós számok.

Az SVD számolására kifejlesztett eljárás numerikusan nagyon stabil, függetlenül attól, hogy az $\mathbf{A}$ mátrix esetleg mennyire szinguláris

Szinguláris érték szerinti felbontás

Fontos fogalmak:
i) Egy mátrix *kondicionáltsága* a legnagyobb és a legkisebb w_i hányadosa. Az $\mathbf{A}$ mátrix *szinguláris*, ha ez a hányados végtelen, i.e., egy (vagy több) $w_j = 0$.

Szinguláris érték szerinti felbontás

Fontos fogalmak:
i) Egy mátrix *kondicionáltsága* a legnagyobb és a legkisebb w_i hányadosa. Az $\mathbf{A}$ mátrix *szinguláris*, ha ez a hányados végtelen, i.e., egy (vagy több) $w_j = 0$.

ii) mátrix *mag tere (nullspace)* és *kép tére (range)*. Tekintsük a

következő lineáris egyenletrendszert:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ez egy lineáris leképezés az $\mathbf{x}$ vektortérről a $\mathbf{b}$ vektortérre.

Szinguláris érték szerinti felbontás

Fontos fogalmak:

i) Egy mátrix *kondicionáltsága* a legnagyobb és a legkisebb w_i hányadosa. Az $\mathbf{A}$ mátrix *szinguláris*, ha ez a hányados végtelen, i.e., egy (vagy több) $w_j = 0$.

ii) mátrix *mag tere (nullspace)* és *kép tére (range)*. Tekintsük a

következő lineáris egyenletrendszert:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ez egy lineáris leképezés az $\mathbf{x}$ vektortérről a $\mathbf{b}$ vektortérre.

iii) Ha $\mathbf{A}$ szinguláris, akkor az $\mathbf{x}$ egy alterében levő vektorokra (a *magtér*) teljesül: $\mathbf{A} \cdot \mathbf{x}' = 0$.

Szinguláris érték szerinti felbontás

Fontos fogalmak:
i) Egy mátrix *kondicionáltsága* a legnagyobb és a legkisebb w_i hányadosa. Az $\mathbf{A}$ mátrix *szinguláris*, ha ez a hányados végtelen, i.e., egy (vagy több) $w_j = 0$.

ii) mátrix *mag tere (nullspace)* és *kép tére (range)*. Tekintsük a

következő lineáris egyenletrendszert:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ez egy lineáris leképezés az $\mathbf{x}$ vektortérről a $\mathbf{b}$ vektortérre.

iii) Ha $\mathbf{A}$ szinguláris, akkor az $\mathbf{x}$ egy alterében levő vektorokra (a *magtér*) teljesül: $\mathbf{A} \cdot \mathbf{x}' = 0$.

iv) A $\mathbf{b}$ vektortér egy része "elérhető" az $\mathbf{A}$ lineáris leképezés segítségével, vagyis vannak olyan $\mathbf{x}$ vektorok, amelyek ide képződnek le: ez az $\mathbf{A}$ *képtere (range)*. Ennek a térnek a dimenziója az $\mathbf{A}$ *rank*-je.

Szinguláris érték szerinti felbontás

Fontos fogalmak:
i) Egy mátrix *kondicionáltsága* a legnagyobb és a legkisebb w_i hányadosa. Az **A** mátrix *szinguláris*, ha ez a hányados végtelen, i.e., egy (vagy több) $w_j = 0$.

ii) mátrix *mag tere (nullspace)* és *kép tére (range)*. Tekintsük a

következő lineáris egyenletrendszert:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ez egy lineáris leképezés az **x** vektortérről a **b** vektortérre.

iii) Ha **A** szinguláris, akkor az **x** egy alterében levő vektorokra (a *magtér*) teljesül: $\mathbf{A} \cdot \mathbf{x}' = 0$.

iv) A **b** vektortér egy része "elérhető" az **A** lineáris leképezés segítségével, vagyis vannak olyan **x** vektorok, amelyek ide képződnek le: ez az **A** *képtere (range)*. Ennek a térnek a dimenziója az **A** *rank*-je.

v) Ha **A** nem szinguláris, akkor $rank = N$, egyébként $rank < N$.

Szinguláris érték szerinti felbontás

Az SVD tulajdonképpen egy ortonormált bázist ad az **A** magterére és képterére:

- az **U** mátrix azon oszlopai, amelyeknek megfelelő w_i értékek nem nullák, bázisvektornak használhatóak a képtérben
- a **V** mátrix azon oszlopai, melyeknek megfelelő w_i értékek nullák, bázisvektornak használhatóak a magtérben

Szinguláris érték szerinti felbontás

Az SVD tulajdonképpen egy ortonormált bázist ad az $\mathbf{A}$ magterére és képterére:

- az $\mathbf{U}$ mátrix azon oszlopai, amelyeknek megfelelő w_i értékek nem nullák, bázisvektornak használhatóak a képtérben
- a $\mathbf{V}$ mátrix azon oszlopai, melyeknek megfelelő w_i értékek nullák, bázisvektornak használhatóak a magtérben

Tekintsük újra az egyenletrendszert

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Szinguláris érték szerinti felbontás

Az SVD tulajdonképpen egy ortonormált bázist ad az $\mathbf{A}$ magterére és képterére:

- az $\mathbf{U}$ mátrix azon oszlopai, amelyeknek megfelelő w_i értékek nem nullák, bázisvektornak használhatóak a képtérben
- a $\mathbf{V}$ mátrix azon oszlopai, melyeknek megfelelő w_i értékek nullák, bázisvektornak használhatóak a magtérben

Tekintsük újra az egyenletrendszert

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

- Ha $\mathbf{A}$ nem szinguláris, akkor a megoldás elvileg $\mathbf{x} = \mathbf{V}[diag(1/w_i)]\mathbf{U}^H\mathbf{b}$
- Nem szokták a Gauss elimináció helyett alkalmazni, mert lassabb

Szinguláris érték szerinti felbontás

Tekintsük újra az egyenletrendszert

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

- Ha $\mathbf{A}$ szinguláris, akkor az a kérdés, hogy $\mathbf{b}$ benne van-e az $\mathbf{A}$ képterében. Ha igen, akkor létezik $\mathbf{x}$ megoldásvektor, csak nem egyértelmű: bármely vektor, amely a magtérben van (melyet a $\mathbf{V}$ oszlopai feszítenek ki) hozzáadható tetszőleges lineárkombinációban az $\mathbf{x}$-hez
- ezen megoldások kozül választhatjuk pl azt, amelyre teljesül, hogy $|\mathbf{x}|^2$ minimális. Hogyan kapható meg ez a partikuláris megoldás?
- ha $w_j = 0$, akkor az $\mathbf{x} = \mathbf{V}[diag(1/w_j)]\mathbf{U}^H\mathbf{b}$-ben $1/w_j$ helyett a megfelelő átlós elemet írd át 0-ra

Szinguláris érték szerinti felbontás

Tekintsük újra az egyenletrendszert

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ha **b** nincs benne az **A** képterében, akkor az egyenletrendszernek nincs megoldása.

Szinguláris érték szerinti felbontás

Tekintsük újra az egyenletrendszert

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ha **b** nincs benne az **A** képterében, akkor az egyenletrendszernek nincs megoldása.

Ekkor a $\mathbf{x} = \mathbf{V}[diag(1/w_j)]\mathbf{U}^H\mathbf{b}$ "megoldásra" ($1/w_j$ helyett a megfelelő átlós elem 0-ra cserélve) igaz az, hogy ez minimalizálja a $|\mathbf{Ax} - \mathbf{b}|$ mennyiséget a legkisebb négyzetek (vagyis χ^2) értelemben.

Szinguláris érték szerinti felbontás

Tekintsük újra az egyenletrendszert

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Ha **b** nincs benne az **A** képterében, akkor az egyenletrendszernek nincs megoldása.

Ekkor a $\mathbf{x} = \mathbf{V}[diag(1/w_j)]\mathbf{U}^H\mathbf{b}$ "megoldásra" ($1/w_j$ helyett a megfelelő átlós elem 0-ra cserélve) igaz az, hogy ez minimalizálja a $|\mathbf{Ax} - \mathbf{b}|$ mennyiséget a legkisebb négyzetek (vagyis χ^2) értelemben.

Bizonyítás: Numerical Recipies, Eq(2.6.10)

SVD és lineáris függvényillesztés

- Az SVD-t többféle problémában lehet alkalmazni
- Most: lineáris függvényillesztés
- Később: főkomponens analízis

Probléma: a normálegyenletek együtthatómátrixa rosszul kondicionált, ezért az egyenletek megoldásakor (közel) 0 pivot elemmel találkoznánk.

SVD és lineáris függvényillesztés

- Az SVD-t többféle problémában lehet alkalmazni
- Most: lineáris függvényillesztés
- Később: főkomponens analízis

Probléma: a normálegyenletek együtthatómátrixa rosszul kondicionált, ezért az egyenletek megoldásakor (közel) 0 pivot elemmel találkoznánk.

Az **X** tervmátrix és a mért értékeket tartalmazó **b** vektor segítségével a legkisebb négyzetek módszere átfogalmazható:

keressük azt az **a** vektort, amelyre $\chi^2 = |\mathbf{X} \cdot \mathbf{a} - \mathbf{b}|^2$ minimális

SVD és lineáris függvényillesztés

- Az SVD-t többféle problémában lehet alkalmazni
- Most: lineáris függvényillesztés
- Később: főkomponens analízis

Probléma: a normálegyenletek együtthatómátrixa rosszul kondicionált, ezért az egyenletek megoldásakor (közel) 0 pivot elemmel találkoznánk.

Az $\mathbf{X}$ tervmátrix és a mért értékeket tartalmazó $\mathbf{b}$ vektor segítségével a legkisebb négyzetek módszere átfogalmazható:

keressük azt az $\mathbf{a}$ vektort, amelyre $\chi^2 = |\mathbf{X} \cdot \mathbf{a} - \mathbf{b}|^2$ minimális

Egzakt megoldás: $\mathbf{a} = \mathbf{V}[diag(1/w_j)]\mathbf{U}^T\mathbf{b}$ ($\mathbf{U}$ most valós mátrix)

SVD és lineáris függvényillesztés

Egzakt megoldás: $\mathbf{a} = \mathbf{V}[diag(1/w_j)]\mathbf{U}^T\mathbf{b}$

Ha valamely $w_j \approx 0$ akkor helyettesítés: $(1/w_j) \to 0$

SVD és lineáris függvényillesztés

Egzakt megoldás: $\mathbf{a} = \mathbf{V}[diag(1/w_j)]\mathbf{U}^T\mathbf{b}$

Ha valamely $w_j \approx 0$ akkor helyettesítés: $(1/w_j) \to 0$

Bizonyítás nélkül:

$$Cov(a_j, a_k) = \mathbf{V}\mathbf{W}^{-1}[\mathbf{V}\mathbf{W}^{-1}]^T = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^T$$

A diagonális elemek adják meg a $\sigma^2(a_j)$ varianciákat