Haladó lineáris algebra

Hatodik házi feladat

1. A feladat: egy véletlen mátrixhoz (pl. rand függvénnyel az Octave/Matlab programokban generált legalább $\mathbf{n} = 6$ darab legalább $\mathbf{m} = 5$-dimenziós vektort adó mátrixhoz) keressük meg a legnagyobb szórást adó egyenes egységnyi irányvektorát, és vetítsük a vektorokat erre az egyenesre, majd írjuk ki a vetületül kapott számokat. A maximális szórást számoljuk ki többféleképp is! Mi köze mindennek a szinguláris értékekhez és vektorokhoz?

Legyenek ezek a véletlen oszlopvektorok mérések és ekkor rakjuk őket össze egy mátrixba

$$\mathbf{A} = \begin{bmatrix} | & \dots & | \\ \mathbf{a_1} & \dots & \mathbf{a_6} \\ | & \dots & | \end{bmatrix} = \begin{bmatrix} 0{,}5606 & 0{,}8790 & 0{,}9900 & 0{,}4981 & 0{,}5860 & 0{,}6609 \\ 0{,}9296 & 0{,}9889 & 0{,}5277 & 0{,}9009 & 0{,}2467 & 0{,}7298 \\ 0{,}6967 & 0{,}0005 & 0{,}4795 & 0{,}5747 & 0{,}6664 & 0{,}8908 \\ 0{,}5828 & 0{,}8654 & 0{,}8013 & 0{,}8452 & 0{,}0835 & 0{,}9823 \\ 0{,}8154 & 0{,}6126 & 0{,}2278 & 0{,}7386 & 0{,}6260 & 0{,}7690 \end{bmatrix},$$

majd középpontosítsuk őket az origó körül úgy, hogy a 6 db oszlopvektor átlagvektorát vonjuk ki az oszlopvektorokból. Ezek a következő MATLAB kódokat használtam.

```
%%- A kiindulasi adatmatrix merete
M = 5; N = 6;

%%- Legyenek veletlen valos szamok 0 es 1 kozott
A = rand([M,N]);

%%-- Szamoljunk atlagot
tmp = zeros(M,1);
for i= 1:N
    tmp = tmp + A(:,i);
end
%%- Az atlagvektor
avg = 1/N* tmp;

%%- Toljuk el az adatok sulypontjat az origoba
%%- ugy, hogy az A oszlopainak atlaga a zerusvektor legyen
B = zeros(M,N);
for j = 1:N
    B(:,j) = A(:,j) - avg;
end
```

Eredményül a következő **B** mátrixot kaptam

$$\mathbf{B} = \begin{bmatrix} -0{,}1564 & 0{,}2097 & 0{,}3847 & -0{,}2134 & 0{,}1443 & -0{,}1456 \\ 0{,}2126 & 0{,}3196 & -0{,}0776 & 0{,}1894 & -0{,}1950 & -0{,}0768 \\ -0{,}0203 & -0{,}6688 & -0{,}1257 & -0{,}1368 & 0{,}2247 & 0{,}0842 \\ -0{,}1342 & 0{,}1961 & 0{,}1961 & 0{,}1337 & -0{,}3582 & 0{,}1757 \\ 0{,}0984 & -0{,}0567 & -0{,}3774 & 0{,}0272 & 0{,}1842 & -0{,}0375 \end{bmatrix}.$$

A feladat szerint találni kell ehhez az adatmátrixhoz egy olyan egyenest, amelyre vetítve a vetítések között szórás(négyzet) a lehető legnagyobb. Legyen ennek az egyenesnek az egység-irányvektora $\mathbf{x}$, így formálisan a feladat

$$\mathbf{x} = max_{\mathbf{x}}\Big(Var\big\{proj_{\mathbf{x}}(\mathbf{b_i})\big\}\Big), \tag{1}$$

mindegyik $\mathbf{b_i}$ vektorra. Tehát az a feladatunk, hogy a lehető legnagyobb „hézagok" legyenek az egyenesre vetített vektorok között. A variancia – arra az esetre, amikor már az átlag köré csoportosítottuk az adatokat – a következő

$$Var(\mathbf{b_i}) = \frac{1}{M}\sum_{i=1}^{M} b_i^2 = \frac{1}{M}(b_1^2 + \ldots + b_6^2) = \frac{1}{M}\mathbf{b_i^T}\mathbf{b_i}.$$

Az $\mathbf{x}$-re való merőleges vetítés pedig

$$proj_{\mathbf{x}}(\mathbf{b_i}) = \frac{\mathbf{b_i^T x}}{\mathbf{x^T x}} = \mathbf{b_i^T x},$$

mivel az $\mathbf{x}$ egységvektor. A $\mathbf{b_i}$ helyére pedig írjuk be a $\mathbf{B}$ mátrixot, ekkor

$$proj_{\mathbf{x}}(\mathbf{B}) = \mathbf{B^T x}.$$

Az 1 egyenletbe írjuk vissza, amit kaptunk és használjuk fel a variancia előbb meghatározott képletét

$$\mathbf{x} = max_{\mathbf{x}}\Big(Var\{proj_{\mathbf{x}}(\mathbf{B})\}\Big) = max_{\mathbf{x}}\Big(\frac{1}{N}(\mathbf{B^T x})^{\mathbf{T}}\mathbf{B^T x}\Big) = max_{\mathbf{x}}\Big(\frac{1}{N}\mathbf{x^T B B^T x}\Big) = \frac{1}{N}max_{\mathbf{x}}\Big(\mathbf{x^T B B^T x}\Big), \qquad (2)$$

ahol észrevehető, hogy a maximum függvény „hasában” egy kvadratikus forma szerepel és a $\mathbf{C} = \mathbf{BB^T}$ egy szimmetrikus mátrix (az $1/N$ most nem érdekes,mivel az csak a variancia értékét változtatja). Valós szimmetrikus mátrixok esetében használható a *valós spektráltétel*, azaz hasonlóvá tehetők egy diagonális mátrixhoz. Ez a diagonizálás nagy segítséget fog nyújtani a 2 optimalizálási feladat meghatározása során. A spektráltételnek megfelelően tehát

$$\mathbf{C} = \mathbf{Q}\Lambda\mathbf{Q^T},$$

ahol $\mathbf{Q}$ ortogonális mátrix, az oszlopai sajátvektorok és Λ a sajátértékek diagonális vektora. Persze, az ortogonálisan diagonizálható mátrixok SVD-je éppen a fenti diagonális alakkal egyenlő

$$\mathbf{C} = \mathbf{Q}\Lambda\mathbf{Q^T} = \mathbf{U}\Sigma\mathbf{V^T}.$$

Újra felírva a problémát, az optimalizálási feladat a következőképpen néz ki

$$\mathbf{x} = max_{\mathbf{x}}\Big(\mathbf{x^T C x}\Big).$$

Ebből most nem látjuk, hogy is néz ki valójában a $\mathbf{C}$ mátrix. Használjuk fel a diagonizálást

$$\mathbf{x} = max_{\mathbf{x}}\Big(\mathbf{x^T Q}\Lambda\mathbf{Q^T x}\Big).$$

Nézzük mikor lesz ez a kvadratikus forma maximális

$$\mathbf{x^T Q}\Lambda\mathbf{Q^T x} = \lambda_{\mathbf{1}}(\mathbf{q_1} \cdot \mathbf{x})^{\mathbf{2}} + \ldots + \lambda_{\mathbf{5}}(\mathbf{q_5} \cdot \mathbf{x})^{\mathbf{2}},$$

ahol $\mathbf{C}$ $5{\times}5$-ös volt. Azt látjuk, hogy a $\mathbf{q_i} \cdot \mathbf{x}$ skaláris szorzatok legfeljebb 1-ek lehetnek, mivel a $\mathbf{q_i}$ ortonormált és az $\mathbf{x}$-eket csak az egységvektorok közül választhatjuk ki. Továbbá a négyzetek miatt egy 1-nél kisebb valós szám négyzetre emeléssel kisebb lesz. Ha $\mathbf{x}$-nek a legnagyobb sajátértékhez tartozó egységnyi hosszú sajátvektort választjuk, akkor az előbb említett kvadratikus forma maximális lesz.

Ezt az $\mathbf{x}$-et pld. a $\mathbf{C}$ mátrix SVD-jéből kaphatjuk meg úgy, hogy a legnagyobb szinguláris értékhez – ami itt most a sajátértéke is – tartozó szinguláris vektort tekintjük az $\mathbf{x}$-nek. Tehát

$$\mathbf{x} = \begin{bmatrix} -0{,}1915 \\ -0{,}5745 \\ 0{,}4185 \\ -0{,}6767 \\ 0{,}0137 \end{bmatrix}.$$

Mindezek a MATLAB-ban a következőképp néztek ki.

```
%%- A kvadratikus formaban szereplo matrix
C = B * transpose(B);

%%- SVD-je
[U,S,V] = svd(C);

%%- Az egyenes irany egysegvektora, amire vetitek
x = U(:,1);
```

A vetítés eredménye a következő volt

$$\mathbf{p} = \begin{bmatrix} 0{,}0440 \\ -0{,}5364 \\ -0{,}0542 \\ -0{,}1572 \\ 0{,}7541 \\ -0{,}0502 \end{bmatrix}$$

és a varianciájuk

$$Var\{proj_{\mathbf{x}}(\mathbf{B})\} = \frac{1}{N} max\{\lambda_{\mathbf{i}}(\mathbf{C})\} = \frac{1}{N}\sigma_{\mathbf{1}}(\mathbf{C}) = \frac{1}{6}0{,}8885 = 0{,}1481$$

és szórásuk

$$\sqrt{Var\{proj_{\mathbf{x}}(\mathbf{B})\}} = \sqrt{\frac{1}{N}\left(\sigma_{\mathbf{1}}(\mathbf{B}^{\mathbf{T}})\right)^{\mathbf{2}}} = 0{,}3848.$$

Persze a feladat szövegében szereplő 2-normával is ugyanerre jutottam.

```
%%- A vetites eredmenye
p = B' * x;

%%- A szorasnegyzet
Vari1 = S(1,1) * 1/N;

%%- Es a szoras
Devi1 = sqrt(Vari1)

%%- Maskepp a szoras, B transz.-janak SVD-jebol
%%- F a szigma matrix
[E,F,G] = svd( transpose(B) );
Vari2 = 1/N * F(1,1)^2;
Devi2 = sqrt(1/N * F(1,1)^2);

%%- 2-normaval
Vari3 = 1/N * norm(transpose(B),2);
Devi3 = sqrt(Vari3);
```