Subject: از این موضوع برای حقیقی بودن مقادیر ویژه A^*A و AA^* ~~استفاده~~ می کنیم .

Date

سوال اول

$A^* = A \Rightarrow$ مقدار ویژه ها حقیقی اند . اثبات :

$$f(x) = \det(xI - A)$$

فرض کنید λ ریشه $f(x) \in \mathbb{C}[x]$ است . پس بردار ناصفر $0 \neq v \in \mathbb{C}^n$

وجود دارد به طوری که $Av = \lambda v$. می دانیم $\langle Av, v\rangle = \langle v, Av\rangle$ چون $A^* = A$

$$\langle Av, v\rangle = v^* A v$$

$$\langle v, Av\rangle = v^* A^* v$$

$$Av = \lambda v \underset{\text{بالا}}{\Rightarrow} \underbrace{\langle \lambda v, v\rangle}_{\lambda v^* v = v^* \lambda v} = \underbrace{\langle v, \lambda v\rangle}_{(\lambda v)^* v = \bar{\lambda} v^* v} \Rightarrow \lambda = \bar{\lambda}$$

$$\Rightarrow \lambda \in \mathbb{R} \quad \square$$

$A \in M_{mn}(\mathbb{F}) \Rightarrow AA^* \in M_m(\mathbb{F})$ و $A^*A \in M_n(\mathbb{F})$

• در واقع داریم: $(AA^*)^* = AA^* \,,\, (A^*A)^* = A^*A \rightarrow$ self-adjoint

★ از طرفی ثابت شد که مقادیر ویژه ماتریس های خودالحاقی، حقیقی اند و نامنفی اند و از این موضوع برای AA^* و A^*A استفاده می کنیم.

$\rightarrow AA^* v = \lambda v$ و $v \neq 0$ و λ مقدار ویژه

$\xRightarrow{\times v^*} v^*AA^*v = v^*\lambda v \xRightarrow{\text{ضرب داخلی است}} \langle A^*v, A^*v \rangle = \lambda \langle v, v \rangle \Rightarrow \|A^*v\|^2 = \lambda \|v\|^2$

$\lambda \in \mathbb{R} \Rightarrow \lambda \geq 0$ □

به طریق مشابه برای A^*A هم می توان نشان داد که مقدار ویژه ها نامنفی اند.

★ اگر در فضای $\mathbb{C}$ نگاه کنیم برای $AA^* \in M_m$ هر m تا مقدار ویژه را خواهیم داشت؛ برای $A^*A \in M_n$ نیز هر n مقدار ویژه را خواهیم داشت. در این صورت AA^*، r تا ناصفر و m مقدار ویژه و m-r تا صفر خواهد داشت. A^*A نیز r تا ناصفر و n تا مقدار ویژه و n-r تا صفر خواهد داشت. (تعداد ناصفرها (r) با احتساب تکرار برابر خواهد بود.)

Subject:
Date

حال برای $U\in M_m(\mathbb{F})$, $V\in M_n(\mathbb{F})$ و ماتریس‌های قطری $D_1\in M_m(\mathbb{R})$ و $D_2\in M_n(\mathbb{R})$
داریم:

$$AA^* = UD_1U^* \quad , \quad A^*A = VD_2V^*$$

$$\Rightarrow D_1 = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{m\times m} \quad , \quad D_2 = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{n\times n}$$

(the first r diagonal entries $\lambda_1,\dots,\lambda_r$; the last $m-r$ (resp. $n-r$) entries 0)

Lemma → $\text{rank}\,A = r = \text{rank}\,AA^* = \text{rank}\,A^*A$

if $\lambda_1 \geq \dots \geq \lambda_r > 0$ are non-zero eigenvalues of AA^* then
$r = \text{rank}\,A = \text{rank}\,AA^* = \text{rank}\,A^*A$

طبق قضیه رتبه ←

$\dim N(A^*A) + \text{rank}(A^*A) = n$

$\dim N(A) + \text{rank}(A) = n$

• کافی است نشان دهیم نقش dim ها
در دو تساوی مقابل باهم برابر است
تا نشان داده باشیم rank ها باهم برابرند.

ادعا: $N(A^*A) = N(A)$ ← اگر $x \in N(A) \Rightarrow Ax = 0 \Rightarrow A^*Ax = 0 \Rightarrow$
$x \in N(A^*A)$ □

اگر $x \in N(A^*A) \Rightarrow A^*Ax = 0 \Rightarrow x^*A^*Ax = 0 \Rightarrow \|Ax\|^2 = 0$
ضرب داخلی
$\Rightarrow Ax = 0 \Rightarrow x \in N(A)$

و حکم اثبات می‌شود □

• برای ماتریس‌های AA^* و A^*A ، مقادیر ویژه ناصفر مشترک با تکرر یکسان دارند.

۱- فرض کنید λ ویژه مقدار ناصفر AA^* باشد، نشان می‌دهیم λ ویژه مقدار A^*A است.
چون λ ویژه مقدار AA^* است بنابراین بردار $v \neq 0$ وجود دارد که $AA^*v = \lambda v$ است و
در نتیجه $A^*A(A^*v) = AA^*v$. اگر $A^*v \neq 0$ آنگاه λ ویژه مقدار A^*A است.
زیرا اگر $A^*v = 0 \Rightarrow AA^*v = 0 \Rightarrow \lambda = 0$ ✗ $\square$

2- به طریق مشابه اگر $\lambda \neq 0$ ویژه مقدار A^*A ، آنگاه ویژه مقدار AA^* است.

3- بایستی برابری تکررها را ثابت کنید.

$$\underbrace{A}_{A}\,\underbrace{A^*}_{B} \Leftrightarrow \underbrace{A^*}_{B}\,\underbrace{A}_{A}$$

در ضرب دو ماتریس چند جمله‌ای مشخصه‌ها چه ویژگی‌ای دارند؟

لم ← اگر $A \in M_{mn}(\mathbb{F})$ و $B \in M_{nm}(\mathbb{F}) \Rightarrow x^n f_{AB}(x) = x^m f_{BA}(x)$

Date

اثبات) <= $AB \in M_m(\mathbb{F})$, $BA \in M_n(\mathbb{F})$ درجه m

فرض کنید : $f_{AB}(x) = \det(xI_m - AB)$ چندجمله‌ای مشخصه AB .

$f_{BA}(x) = \det(xI_n - BA)$ 〃 〃 〃 BA .

درجه n

یک ماتریس باید بسازیم ←

$$\left|\begin{array}{c|c} I_m & -A \\ \hline 0 & xI_n \end{array}\right| \left|\begin{array}{c|c} xI_m & A \\ \hline B & I_n \end{array}\right| = \left|\begin{array}{c|c} xI_m - AB & 0 \\ \hline \underset{= xB}{*} & xI_n \end{array}\right|$$

$$\left|\begin{array}{c|c} I_m & 0 \\ \hline -B & xI_n \end{array}\right| \left|\begin{array}{c|c} xI_m & A \\ \hline B & I_n \end{array}\right| = \left|\begin{array}{c|c} xI_m & A \\ \hline 0 & xI_n - BA \end{array}\right|$$

Q.E.D

* برقراری سری 5 ام
این قسمت اثبات شد.

پس $x^n f_{AA^*}(x) = x^m f_{A^*A}(x)$ □ و این نشان می‌دهد تکرارها یکسان اند .

سوال دوم) الف)

در قضیه‌ای که به اثبات آن می‌پردازیم اگر A ماتریس خودالحاق باشد بدان معنی که $A = A^*$ ، ماتریس وارون پذیری مانند $P \in M_n(\mathbb{F})$ که یکانی است و $P^*P = I$ داریم ؛

$$P^*AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

یعنی ماتریس های خودالحاق قطری شدنی هستند و ماتریسی که آنرا قطری می‌کند یکانی است .

از طرفی اگر این مورد را نشان دهیم و با توجه به آنکه A^*A و AA^* خودالحاق هستند ، یعنی

$$(A^*A)^* = A^*A \quad و \quad (AA^*)^* = AA^*$$

پس می‌توان نتیجه گرفت ($A \in M_{mn}(\mathbb{F}) \Rightarrow AA^* \in M_{mm}(\mathbb{F})$ و $A^*A \in M_{nn}(\mathbb{F})$)

ماتریس های $U \in M_{mm}(\mathbb{F})$ و $V \in M_{nn}(\mathbb{F})$ وجود دارند که (در واقع حکم P را دارند) یکانی هستند و اگر D ماتریسی قطری حاوی $\lambda_1, \dots, \lambda_r$ باشد $\Rightarrow$

$$AA^* = UD_1U^* \quad و \quad A^*A = VD_2V^*$$

دقت کنید در قضیه بالا ماتریس قطری ما درایه های $\lambda_1, \dots, \lambda_n$ را بر قطر خود دارد ولی می‌تواند عناصر آن از λ_{r+1} تا λ_n همگی صفر باشند و همان ماتریس D فرض سوال ساخته شود .

پس صرفاً با این توضیحات کافی است قضیه بالا اثبات شود .

ماتریس خود الحاقی (ویژگی)

اگر A خود الحاقی باشد ، یک ماتریس وارون پذیر $P \in M_n(\mathbb{F})$ وجود دارد طوری که $P^*P = I$

و $P^*AP = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ قطری شدنی است A

ویژگی مهم خود الحاق ها $\boxed{P^{-1} = P^*}$

(قطری شدنی ها : λ_i ها مقدار ویژه و ستون های P بردار ویژه ها اند ($Av_i = \lambda v_i$)

$$P^*P = \begin{bmatrix} v_1^* \\ \vdots \\ v_n^* \end{bmatrix} [v_1 \; \dots \; v_n] = \begin{bmatrix} v_1^*v_1 & \dots & v_1^*v_n \\ \vdots & \ddots & \vdots \\ v_n^*v_1 & \dots & v_n^*v_n \end{bmatrix} = I \Rightarrow \underbrace{v_i^*v_j}_{\langle v_j, v_i \rangle} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

* پس ستون های P برهم عمود اند و $\langle v_i, v_i \rangle = 1$: $\overbrace{\text{ستون های } P \text{ یکه و متعامد اند}}^{\text{بردار ویژه اند *}}$.

اگر روی $\mathbb{C}$ باشیم که حتماً بردار ویژه داریم .

* فقط Gهای مثبت معین ضرب داخلی را می سازند .

① A خود الحاق $\Leftarrow$ مقدار ویژه های حقیقی

② $T(x) = Ax$, $T : \mathbb{F}^n \to \mathbb{F}^n$

③ استقرا ✓

Subject: از این موضوع برای حقیقی بودن مقادیر ویژه A^*A و AA^* استفاده می‌کنیم.
Date

$A^* = A \Rightarrow$ مقدار ویژه‌ها حقیقی‌اند. اثبات:

$f(x) = \det(xI - A)$

فرض کنید λ ریشه $f(x) \in \mathbb{C}[x]$ است. پس بردار ناصفر $0 \neq v \in \mathbb{C}^n$
وجود دارد به طوری که $Av = \lambda v$. می‌دانیم $\langle Av, v\rangle = \langle v, Av\rangle$ چون $A^* = A$.

$\langle Av, v\rangle = v^* A v$

$\langle v, Av\rangle = v^* A^* v$

$$Av = \lambda v \underset{\text{با ازای}}{\Rightarrow} \underbrace{\langle \lambda v, v\rangle}_{v^*\lambda v = \lambda v^* v} = \underbrace{\langle v, \lambda v\rangle}_{(\lambda v)^* v = \bar{\lambda} v^* v} \Rightarrow \lambda = \bar{\lambda}$$

$$\Rightarrow \lambda \in \mathbb{R} \quad \square$$

وقتی اثبات کردیم بردار ویژه‌ها عموداند (ستون‌های متعامد)، می‌توان P ای یافت که $\underbrace{P^*}_{=P^{-1}} P = I$

و $P^*AP = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \xrightarrow{\times P} AP = P\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

$P = [v_1 \quad \dots \quad v_n]$

$A[v_1 \quad \dots \quad v_n] = [Av_1 \quad \dots \quad Av_n] = [\lambda_1 v_1 \quad \dots \quad \lambda_n v_n]$

$\Rightarrow Av_i = \lambda v_i \quad \forall i \quad 1 \leq i \leq n$

پس ستون‌های P که دنبال آن هستیم بردار ویژه‌های متناظر با λ_1 تا λ_n می‌باشد.

چون P وارون‌پذیر بود vها تشکیل پایه می‌دهند (مستقل خطی اند) $\leftarrow B = \{v_1, \dots, v_n\}$
که ویژه بردار هم هستند و $P^*P = I$

با استقرا نشان می‌دهیم: $T: F^n \to F^n$ و $T(x) = Ax$ با استقرا روی n نشان می‌دهیم
یک پایه متعامد از بردار ویژه‌ها وجود دارد.

PAPCO

فرض کنید λ بردار ویژه A است، بنابراین $0 \neq v \in \mathbb{F}^n$ وجود دارد که $Tv = Av = \lambda v$ است. $W := \text{Span}(\{v\})$ و بنابر قضیه $\mathbb{F}^n = W \oplus W^\perp$

فرض استقرا ← برای $n-1$: $T' : V \to V$ و $\dim V = n-1$ و $[T']^*_B = [T']_B$ آنگاه V یک پایه متعامد از بردار ویژه ها دارد.

باید نشان دهیم تحدید T به $W^\perp$ از $W^\perp$ به $W^\perp$ باشد یعنی

$$T|_{W^\perp} : W^\perp \to \underbrace{\mathbb{F}^n}_{W^\perp} \quad : v \in W^\perp \quad T|_{W^\perp}(v) \in W^\perp$$

ادعا: $T(W^\perp) \subseteq W^\perp$. فرض کنید $u \in W^\perp$ باشد، باید نشان دهیم $Tu \in W^\perp$ یعنی باید نشان دهیم Tu بر هر بردار W عمود است. لذا کافی است نشان دهیم Tu بر v عمود است.

$$\langle Tu, v \rangle = \langle Au, v \rangle = \underbrace{v^*}A u$$

$$v^* A = (A^* v)^* = (Av)^* = (\lambda v)^* = \lambda v^*$$

$\lambda \in \mathbb{R}$ ← حقیقی است. ← خودالحاقی بود.

$$= \underbrace{\lambda \langle u, v \rangle = 0}_{\substack{v \in W \\ u \in W^\perp}} \quad \square$$

پس $T|_{W^\perp} : W^\perp \to W^\perp$ و $\dim W^\perp = n-1$. بنابر فرض استقرا پایه‌ای از بردار ویژه های $T|_{W^\perp}$ مانند B' وجود دارد که اعضای آن متعامد و یکه است. پس قرار دهید:

$$B = B' \cup \left\{\frac{v}{\|v\|}\right\}$$

ادعا: B ویژگی های مطلوب را دارد. زیرا: 1- همه بردارهای B بردار ویژه های T هستند.

2- همه بردارهای B متعامد و یکه اند زیرا:

- $\frac{v}{\|v\|}$ یکه است.
- همه بردارهای B' یکه اند و متعامد اند.
- به ازای هر $u \in B'$ ، $\langle u, \frac{v}{\|v\|} \rangle = 0$ زیرا $u \in W^\perp$ و $v \in W$

پس B پایه مطلوب است.

Q.E.D

حال تعریف می‌کنیم $\sqrt{\lambda_i} = \delta_i$ و داریم : سوال ۹ ب

$$\Sigma = \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_r \end{bmatrix} \Rightarrow D_1 = \Sigma\Sigma^* \text{ و } D_2 = \Sigma^*\Sigma$$

طبق مطالب بالا داریم $\Rightarrow$ $AA^* = UD_1U^* = U\Sigma\Sigma^*U^* = U\Sigma V^*V\Sigma U^*$

$= (U\Sigma V^*)(U\Sigma V^*)^*$

$$A^*A = VD_2V^* = V\Sigma^*\Sigma V^* = V\Sigma^*U^*U\Sigma V^* = (U\Sigma V^*)^*(U\Sigma V^*)$$

و به نظر می‌رسد برای انتخاب A یک کاندیدای مناسب داریم $\Rightarrow A = U\Sigma V^*$

فرم "SVD"

Subject:
Date

چرا این تساوی و این کاندیدا برای A درست است؟ چون $UU^*=I$ و $VV^*=I$

اگر ستون های U و V را در نظر بگیریم یک پایه متعامد یکه برای $\mathbb{F}^n$ تشکیل می‌شود و طبق فصل ۳ می‌دانیم داشتن پایه متعامد یکه بسیار ارزشمند است چون وقتی به صورت ترکیب خطی می‌نوشتیم ضرایب آن مشخص بود.

$V=[v_1 \ \dots \ v_n]$ ← این ستون ها علاوه بر پایه متعامد و یکه، ویژه بردارهای نظیر λ_1 تا λ_n نیز هستند.

$\lambda_j=\delta_j^2$

$$A^*Av_j=\delta_j^2 v_j$$

$$\xrightarrow{\times A} AA^*(Av_j)=\delta_j^2 Av_j$$

اگر بگوییم $0\neq Av_j$ در این صورت Av_j

از طرفی ویژه بردارهای AA^* نیز ستون های U هستند و ویژه بردار نظیر δ_j^2 است.

ارتباطی میان آنها وجود دارد.

در واقع با نرمال کردن Av_j می‌توان یک تناظر با ستون های U برقرار کرد $\Leftarrow$ $u_j=\frac{1}{\|Av_j\|}Av_j$

و $Av_j=0$ چون در این صورت تساوی راست یعنی $\delta_j^2 Av_j$ باید صفر شود ولی چون $v_j\neq 0$ است پس $\delta_j^2=0$ می‌شود و این تناقض است چون فرض کرده بودیم

$0\neq\lambda_j=\delta_j^2$ □

همچنین $\Leftarrow$ $\|Av_j\|^2=v_j^*A^*Av_j=\delta_j^2\|v_j\|^2=\delta_j^2$

$$\Rightarrow \|Av_j\|=\delta_j \ ✓$$

در نهایت $Av_j=\delta_j u_j$ $\Leftarrow$

$$AV=A[v_1 \ \dots \ v_n]$$
$$=[Av_1 \ \dots \ Av_r \ 0 \ \dots \ 0]$$
$$=[\delta_1u_1 \ \dots \ \delta_r u_r \ 0 \ \dots \ 0]$$

$$\Rightarrow [u_1 \ \dots \ u_r \ 0 \ \dots \ 0]\begin{bmatrix}\delta_1 & & & \\ & \ddots & & \\ & & \delta_r & \\ & & & 0 \ddots 0\end{bmatrix}=U\Sigma$$

این تساوی چون در صفر ضرب می‌شود و بی اثر است به جای $u_{r+1} \dots u_m$ صفر قرار می‌دهیم (در واقع می‌تواند دلخواه باشد).

PAPCO

پس تجزیه SVD تمام می‌شود و داریم:

$$\delta_1u_1v_1^*+\dots+\delta_r u_r v_r^*=A=U\Sigma V^*$$

Q.E.D

ادامه
سوال 2 ب) اثبات SVD

v ها را بردار ویژه های A^*A درنظر بگیرید.

$$A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^*\Sigma V$$

در سمت راست تساوی ماتریس بردار ویژه ها یعنی V را داریم که برای ماتریس مثبت (نیمه) معین A^*A تشکیل شده است. در این صورت چون A^*A با ماتریس $\Sigma^*\Sigma$ مشابه است و طبق مطالب فصل 5 می دانیم اعضای $\Sigma^*\Sigma$ را مقادیر ویژه های A^*A تشکیل می دهند.

هر σ^2 برابر $\lambda(A^*A)$ است.

حال طبق سوال سوم که داریم $Av_i = \sigma_i u_i$ به ما بردار یکه های u_1 تا u_r را می دهد.

تمام دلیلی که SVD صحیح می باشد این است که بردارهای یکه u_1 تا u_r به صورت خودکار بر یکدیگر عمود اند (چون v ها متعامد اند).

گام کلیدی $\longrightarrow$ $u_i^* u_j = \left(\frac{Av_i}{\sigma_i}\right)^*\left(\frac{Av_j}{\sigma_j}\right) = \frac{v_i^* A^* A v_j}{\sigma_i \sigma_j} = \frac{\sigma_j^2}{\sigma_i \sigma_j} v_i^* v_j = 0$

v ها بردار ویژه های A^*A (متقارن) بوده اند. آنها متعامد اند و در نتیجه u ها نیز متعامد اند.
(در واقع u هایی که بردار ویژه از AA^* هستند)

در نهایت n تا v و m تا u که پایه متعامد می سازند و یکه نیز هستند (orthonormal) را برای فضای پوچ $N(A)$ و $N(A^*)$ داریم.

V ، U و Σ را داریم پس $A = U\Sigma V^*$ را خواهیم داشت = Q.E.D

اگر ستون های U و V را در نظر بگیریم یک پایه متعامد یکه برای $\mathbb{F}^n$ تشکیل می‌شود و طبق فصل ۳ می‌دانیم یافتن پایه متعامد یکه بسیار ارزشمند است چون وقتی به صورت ترکیب خطی می‌نوشتیم ضرایب آن مشخص بود.

$V = [v_1 \ \dots \ v_n]$

$\rightarrow$ این ستون‌ها علاوه بر پایه متعامد و یکه، ویژه بردارهای نظیر λ_1 تا λ_n نیز هستند.

سوال سوم

$\lambda_j = \delta_j^2$

$$A^*Av_j = \delta_j^2 v_j$$

$$\xrightarrow{\times A} AA^*(Av_j) = \delta_j^2 Av_j$$

اگر بگوییم $Av_j \neq 0$ در این صورت Av_j از طرفی ویژه بردارهای AA^* نیز ستون های U هستند و ویژه بردار زیر $\lambda_j = \delta_j^2$ است.

ارتباطی میان آنها وجود دارد.

در واقع با نرمال کردن Av_j می‌توان یک تناظر با ستون های U برقرار کرد $\Rightarrow$ $u_j = \dfrac{1}{\|Av_j\|} Av_j$

و $Av_j = 0$ چون در این صورت سمت راست یعنی $\delta_j^2 Av_j$ باید صفر شود ولی چون $v_j \neq 0$ است پس $\delta_j^2 = 0$ می‌شود و این تناقض است چون فرض کرده بودیم

$0 \neq \lambda_j = \delta_j^2$ $\square$

همچنین $\Rightarrow$ $\|Av_j\|^2 = v_j^* A^* A v_j = \delta_j^2 \|v_j\|^2 = \delta_j^2$

$$\Rightarrow \|Av_j\| = \delta_j \ \checkmark$$

در نهایت $Av_j = \delta_j u_j$ $\Rightarrow$

$$\begin{aligned} AV &= A[v_1 \ \dots \ v_n] \\ &= [Av_1 \ \dots \ Av_r \ 0 \ \dots \ 0] \\ &= [\delta_1 u_1 \ \dots \ \delta_r u_r \ 0 \ \dots \ 0] \end{aligned}$$

$$\Rightarrow [u_1 \ \dots \ u_r \ \underbrace{0 \ \dots \ 0}] \begin{bmatrix} \delta_1 & & & & \\ & \ddots & & & \\ & & \delta_r & & \\ & & & 0 & \\ & & & & \ddots \ 0 \end{bmatrix} = U\Sigma$$

این قسمت چون در صفر ضرب می‌شود و بی اثر است به جای $u_{r+1} \dots u_m$ صفر قرار می‌دهیم (در واقع می‌تواند دلخواه باشد).

PAPCO

پس تجزیه تمام SVD می‌شود و داریم:

$$\delta_1 u_1 v_1^* + \dots + \delta_r u_r v_r^* = A = U\Sigma V^*$$

Q.E.D

سوال چهارم ۱

طبق مطالبی که تا به اینجا چه از سوالات قبلی و چه در کلاس یاد گرفته ایم

می دانیم برای A با رتبه r که $\delta_1 \dots \delta_r$ مقدار تکین‌های ما هستند داریم ؛

$$A = U \Sigma V^*$$

و در سوال سوم نشان داده‌ایم که : $AV = U\Sigma$

$$AV = A[v_1 \; \dots \; v_n]$$

$$= [Av_1 \; \dots \; Av_r \; \overbrace{0 \; 0 \; \dots \; 0}^{\text{بی اهمیت}}]$$

$$= [\delta_1 u_1 \; \dots \; \delta_r u_r \; 0 \; \dots \; 0]$$

و نکته حائز اهمیت این است که

با توجه به آنکه r تا مقدار تکین درگیر مسئله هستند و باعث می‌شوند ماتریس ما رتبه r داشته باشد

و با توجه به فرم خاص ماتریس Σ که یک بلوک $S = [\delta_1 \; \dots \; \delta_r]$ تولید می‌کند و عملاً

اهمیتی ندارد باقی اعضای U و V چگونه انتخاب شوند و فقط r تای اول آنها مهم اند

در این صورت داریم :

$$A = [u_1 \; \dots \; u_r \; \underbrace{0 \; \dots \; 0}_{\text{بی اهمیت}}] \left[\begin{array}{c|c} S_{r\times r} & 0 \\ \hline 0 & 0 \end{array}\right] \begin{bmatrix} v_1^* \\ \vdots \\ v_r^* \\ \left.\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}\right\} \text{بی اهمیت} \end{bmatrix}$$

$$\Rightarrow A = [u_1 \; \dots \; u_r][S_{r\times r}]\begin{bmatrix} v_1^* \\ \vdots \\ v_r^* \end{bmatrix} = U_r \Sigma_r V_r^*$$

Q.E.D

سوال پنجم)

با توجه به سوال چهارم نشان دادیم که $A = U_r \Sigma_r V_r^*$

در این صورت با توجه به قاعده ضرب ماتریس ها داریم :

ماتریس ۱x۱ ← تجزیه SVD

$$A = [u_1 \; \dots \; u_r]\begin{bmatrix}\sigma_1 & & 0\\ & \ddots & \\ 0 & & \sigma_r\end{bmatrix}\begin{bmatrix}v_1^*\\ \vdots \\ v_r^*\end{bmatrix} \rightarrow U_{1\times r}\Sigma_{r\times r}V^*_{r\times 1} = A_{1\times 1}$$

$$= \begin{bmatrix} u_1\sigma_1 & & \\ & \ddots & \\ & & u_r\sigma_r\end{bmatrix}\begin{bmatrix}v_1^*\\ \vdots\\ v_r^*\end{bmatrix} = [u_1\sigma_1 v_1^* + \dots + u_r\sigma_r v_r^*]$$

Q.E.D

سوال ششم ا برای A داریم <=

$$A = U\Sigma V^* = \sum_{i=1}^{r} \delta_i u_i v_i^*$$

سوال هشتم در واقع از این موضوع استفاده می کند

ماتریس های $Z_i = u_i v_i^*$ و

$$= \sum_{i=1}^{r} \delta_i Z_i$$

مقادیر تکین $\delta_1 \geq \dots \geq \delta_r$ را داریم .

از طرفی ماتریس های $\{Z_1, \dots, Z_r\}$ ماتریس های متعامد و یکه اند چون <=

$$\langle Z_i, Z_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

از طرفی چون U و V یکانی اند

$$Z_i^* Z_j = (v_i u_i^*)(u_j v_j^*)$$

$UU^* = I$ و $VV^* = I$

و داریم :

$$\langle u_j, u_i \rangle = u_i^* u_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$\langle v_j, v_i \rangle = v_i^* v_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

در واقع وقتی $j = i$ باشد $\langle Z_i, Z_i \rangle$ را خواهیم داشت که اندازه 1 می دهد (یکه بودن) و برای $i \neq j$ ، $\langle Z_i, Z_j \rangle$ صفر می شود (متعامد بودن)

از طرفی ثابت کنیم داریم :

$\delta_i = \langle Z_i, A \rangle$ و این یعنی

در نتیجه هر مولفه δ_i از A هم در جهت Z_i خواهد بود .

$$\langle A, Z_i \rangle = Z_i^* A = (v_i u_i^*)(\delta_1 u_1 v_1^* + \dots + \delta_i u_i v_i^* + \dots + \delta_r u_r v_r^*)$$

از طرفی طبق مطالب بالا برای $j \neq i$ مقدار صفر بر می گرداند حاصل ضرب بالا و برای $i = j$ مقدار 1 .

در نهایت از A فقط $\delta_i u_i v_i^*$ مهم است است

که داریم $(v_i u_i^*)(\delta_i u_i v_i^*) = \delta_i (v_i \underbrace{u_i^* u_i}_{1} v_i^*)$

$$= \delta_i \quad \square$$

Q.E.D

سوال هشتم)

گام اول) به طور کلی برای ماتریس های $A_{m\times n}$ و $B_{n\times p}$ یک قضیه وجود دارد که به صورت زیر است :

$$rank(AB) \le \min(rank(A), rank(B))$$

اثبات) ابتدا توجه کنید که برای رتبه ضرب داریم : $rank(AB) = rank(B) - \dim N(A)\cap R(B)$

اثبات) پایه $S=\{v_1,\dots,v_s\}$ را برای $N(A)\cap R(B)$ در نظر بگیرید و توجه کنید که $N(A)\cap R(B)\subseteq R(B)$ است. اگر $\dim R(B)=s+t$ در این صورت با گسترش پایه می توان پایه جدید $B=\{v_1,\dots,v_s,u_1,\dots,u_t\}$ را یافت که پایه ای برای $R(B)$ است.

باید نشان دهیم $\dim R(AB)=t$ و این کار را با تعریف ماتریس T به شکل $T=\{Au_1,\dots,Au_t\}$ که پایه ای برای $R(AB)$ است نشان دهیم. T ، $R(AB)$ را span می کند چون اگر $b\in R(AB)$ باشد در این صورت $b=Ay$ برای y ولی $By\in R(B)$ یعنی

$$By=\sum_{i=1}^{s}\alpha_i v_i+\sum_{i=1}^{t}\beta_i u_i$$

$$\Rightarrow b=A\left(\sum_{i=1}^{s}\alpha_i v_i+\sum_{i=1}^{t}\beta_i u_i\right)=\sum_{i=1}^{s}\alpha_i Av_i+\sum_{i=1}^{t}\beta_i Au_i$$

$$=\sum_{i=1}^{t}\beta_i Au_i$$

T مستقل خطی است چون

$$0=\sum_{i=1}^{t}c_i Au_i$$

$$=A\sum_{i=1}^{t}c_i u_i$$

در نهایت $\sum_{i=1}^{t}c_i u_i\in N(A)\cap R(B)$ و در نتیجه اسکالرهای g_i وجود دارند که

$$\sum_{i=1}^{t}c_i u_i=\sum_{j=1}^{s}g_j v_j$$

یا معادلاً :

$$\sum_{i=1}^{t}c_i u_i-\sum_{j=1}^{s}g_j v_j=0$$

چون B مجموعه مستقل بود تنها جواب برای عبارت بالا برای c_i ها و g_i ها جواب صفر است.

بنابراین T پایه ای برای $R(AB)$ است و $t=\dim R(AB)=rank(AB)$

$$\Rightarrow rank(B)=\dim R(B)=s+t=\underbrace{\dim N(A)\cap R(B)+rank(AB)}_{(I)}\qquad\square$$

ادامه سوال هشتم)

حال با توجه به اثبات (I) داریم :

$$rank(AB) = rank(B) - \dim N(A) \cap R(B) \leq rank(B)$$

و این عبارت به ما می گوید رتبه حاصلضرب ماتریس ها نمی تواند از $rank(B)$ (عبارت سمت راست) بیشتر شود . برای آنکه نشان دهیم $rank(AB) \leq rank(A)$

می دانیم ترانهاده کردن رتبه را تغییر نمی دهد پس : $rank(AB) = rank(AB)^T$

$$= rank(B^T A^T) \leq rank(A)^T = rank(A)$$

Q.E.D

پس برای B از رتبه k و ماتریس V_{k+1} می توان نتیجه گرفت

$$rank(BV_{k+1}) \leq rank(B) = k \qquad \text{Q.E.D}$$

در این صورت می توان نتیجه گرفت که $1 \leq \dim N(BV_{k+1}) = k+1 - rank(BV_{k+1})$

در این صورت وجود دارد $x \in N(BV_{k+1})$ که $\|x\|_2 = 1 \Rightarrow BV_{k+1}x = 0$

و $\rightarrow$

$$(*) \; \overbrace{AV_{k+1}x = U\Sigma V^T F x}^{} = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T F x$$

که $\leftarrow$ $D = \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_r \end{bmatrix}_{r \times r}$

$$= U \begin{bmatrix} S & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} = U \begin{bmatrix} Sx \\ 0 \\ 0 \end{bmatrix}$$

$$S = \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_{k+1} \end{bmatrix}$$

حال طبق تعریف اگر به جای z ، y را قرار دهیم $\Rightarrow$

$$\|A - B\|_F = \max_{\|y\|_2 = 1} \|(A-B)y\|_2$$

و داریم $\|V_{k+1}x\|_2 = \|x\|_2 = 1$

در این صورت چون $y = V_{k+1}x \Rightarrow$ $(*)$ $Ay = AV_{k+1}x$ $\rightarrow$ که درستی هم با توجه به (*) بالایی اثبات می شود .

در نهایت $\Rightarrow$

$$\|A - B\|_F^2 \geq \|(A-B)\underbrace{V_{k+1}x}_{=y}\|_2^2 = \|Sx\|_2^2$$

$$= \sum_{i=1}^{k+1} \delta_i^2 x_i^2 \geq \delta_{k+1}^2 \sum_{i=1}^{k+1} x_i^2 = \delta_{k+1}^2$$

در واقع در نهایت اثبات کردیم $\delta_{k+1} = \min_{rank(B) = k} \|A - B\|_2$

Q.E.D

و این همان نزدیکترین ماتریس از رتبه k به A است $\leftarrow$ ماتریس B منظور است .

سوال هشتم)

طبق فرضیات مسئله ما ماتریس A را به صورت تجزیه SVD نوشته ایم

یعنی ؛

$$A = U\Sigma V^* = \sum_{i=1}^{r} \delta_i \underbrace{u_i v_i^*}_{Z_i} = \underbrace{\delta_1 Z_1 + \dots + \delta_r Z_r}_{} \quad (I)$$

$Z_i := u_i v_i^*$

اگر در نظر بگیریم ،

$$\delta_1 \geq \dots \geq \delta_r$$

ماتریس های Z_i متعامد خواهند بود در واقع داریم ؛

$$\langle Z_i , Z_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

(مشابه مطالبی که قبلاً هم داشتیم)

در نهایت داریم ؛

$$SNR(\delta_1 Z_1) \geq \dots \geq SNR(\delta_r Z_r)$$

اگر مقدار تکین δ_{K+1} نسبتاً کوچک باشد در این صورت عبارت $\delta_{K+1} Z_{K+1}$ نیز مقدار SNR کوچکی خواهد داشت . حال اگر این مقدار را از تساوی (I) حذف کنیم ، از کل سیگنال ها بخش کوچکی را از دست می دهیم اما یک بخش بزرگ از کل سیگنال های ماتریس A حذف می شود در این صورت متوجه می شویم تجزیه ناقص که جمله های کمتری نسبت به A دارند $\left(A^{(K)} = \delta_1 Z_1 + \dots + \delta_K Z_K\right)$ باعث می شود بدون از دست دادن اطلاعات زیادی درمورد سیگنال های A داده های حاوی نویز را پالایش کنیم .

Q.E.D

مولوتر جست و جو

سوال نهم)

در واقع با توجه به آنکه ماتریس A شامل ستون های D_i و سطرهای T_j است پس یک ماتریس $m\times n$ است که به صورت زیر تعریف می‌شود :

$$A = \begin{array}{c|ccc} & D_1 & \cdots & D_n \\ \hline T_1 & f_{11} & \cdots & f_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ T_m & f_{m1} & \cdots & f_{mn} \end{array}_{m\times n}$$

در واقع f_{ij} برابر است با تعداد دفعاتی که عبارت T_i در متن D_j ظاهر می‌شود .

از طرفی $q = e_{i1} + \cdots + e_{il}$ و این به معنای آن است که ماند یک Indicator Variable هرگاه T_i ظاهر می‌شود q مقدار ۱ را برمی‌گرداند و در غیر این صورت صفر را برمی‌گرداند .

عبارت پیشنهادی سوال $\cos\theta_j = \dfrac{q^T D_j}{\|q\|_2\|D_j\|_2}$ است که با توجه به تعاریف خود آن نیز برابر است با :

$$\frac{q^T A e_j}{\|q\|_2\|Ae_j\|_2}$$

اگر ستون های ماند q در ابتدا متعامد باشند و یکه باشند در این صورت می‌خواهیم بگوییم چرا $\|q^T A\| = [\|\cos\theta_1\|, \ldots, \|\cos\theta_n\|]$ اطلاعاتی درباره رتبه بندی متن‌ها می‌دهند !

در واقع عبارت ↑ نویزهایی درون خود دارد و برای پالایش این نویزها مانند سوال هشتم کافی است A را به صورت یک تجزیه SVD بنویسیم <=

$$A = \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^* = U\Sigma V^*$$

و مانند سوال هشتم برای تجزیه ناقص $A^{(k)}$ داریم ؛

$$A^{(k)} = \sigma_1 u_1 v_1^* + \cdots + \sigma_k u_k v_k^*$$

و از آن به جای A استفاده می‌کنیم

می‌توان گفت به ازای ϵ هایی که $\|\cos\theta_j\| \geq \epsilon$ است ، متن D_j قابل قبول است .

Q.E.D

ادامه سوال نهم) سوال دهم) چرا مقایسه D_j با $\cos\varphi_j = \frac{q^T A^{(k)} e_j}{\|q\|_2 \|A^{(k)} e_j\|_2}$ راهکاری مناسب برای پالایش نویزها است؟

الف)

در آخر سوال نهم گفتیم با تجزیه SVD و تجزیه ناقص آن می‌توان پالایش نویزها را بهبود بخشید. این کار یعنی جایگزینی $A^{(k)}$ با A به عبارتی دیگر جایگزینی $\cos\theta_j$ با $\cos\varphi_j$ است و در ادامه مقایسه D_j با آن.

← در این صورت : $A^{(k)} = U_k \Sigma_k V_k^*$ ، $\|A^{(k)} e_j\|_2 = \|U_k \Sigma_k V_k^* e_j\|_2$

اگر قرار دهیم $S_k = \Sigma_k V_k^* = [S_1 \dots S_k]$

در این صورت داریم:

$$\|A^{(k)} e_j\|_2 = \|U_k S_j\|_2 = \|S_j\|_2$$

$$\Rightarrow \cos\varphi_j = \frac{q^T U_k S_j}{\|q\|_2 \|S_j\|_2}$$

در این صورت بردار های در U_k و S_k بدون محاسبه کل SVD به دست می‌آیند و با محاسبات نه چندان زیاد $\cos\varphi_j$ محاسبه می‌شود. به علاوه در این صورت می‌توانیم جملات بیشتری در تجزیه ناقص را حذف کنیم و با توجه به آنکه عبارات زیادی که ممکن است در متن D_j تولید شده باشند و در این صورت مقدار خیلی زیادی نویز در A تولید می‌کنند و با این کار و حذف جملات A و رسیدن به تجزیه ناقص با جملات به مراتب کمتر از A همانند آنچه در سوال هشتم گفته شد یک رویکرد مناسبی برای پالایش نویزها خواهد بود.

به طور خلاصه به دنبال جایگزینی $A^{(k)}$ با جملات کمتر از A هستیم تا بتوانیم پالایش بهتری از نویزهای در A داشته باشیم Q.E.D

ب) یک رویکرد جایگزین می‌تواند جایگزینی q در $\cos\theta_j = \frac{q^T D_j}{\|q\|_2 \|D_j\|}$ انگلیش شده آن است

اگر تعریف کنیم $P_{C(A)} = U_r U_r^*$ به طوری که $q_p = P_{C(A)} q$ که تصویر عمودی بر فضای ستونی A یا $C(A)$ است که بر راستای $C(A)^\perp$ خواهد بود.

در فصل سوم به طور کامل در مورد ساخت q_p بحث شد. می‌دانیم $q_p \in C(A)$ یعنی برداری در فضای ستونی A یا به عبارتی در فضای متن‌های D_j قرار دارد و نزدیکترین نقطه به q است و در واقع می‌توان با استفاده از q_p به جای q که بهترین تقریب از آن است می‌تواند رویکردی جایگزین و مناسب برای ما باشد.

ادامه صفحه بعد

ادامه سوال اهم ب) دقت کنید که چون q_p در فضای ستونی A یعنی همان D_j ها قرار دارد پس با ترکیب خطی ای از آنها تشکیل می‌شود.

حال توجه کنیم که: $q_p^T A = q^T A \Rightarrow \|q_p\|_2 \leq \|q\|_2$

و این بدان معنا است که $\cos\theta_{p_j} \geq \cos\theta_j$ و در این صورت وقتی از تصویر q (با توجه به $C(A)$) استفاده می‌کنیم به نظر می‌توان مقادیر بیشتری از متن‌ها را قبول کرد

$$\Rightarrow \cos\theta_{p_j} = \frac{q_p^T A e_j}{\|q_p\|_2 \|A e_j\|_2}$$

به نحوی مانند مطالب قبلی که سعی می‌کردیم تجزیه ناقص یعنی $A^{(k)}$ را جایگزین A کنیم

در این صورت $\cos\varphi_{p_j} = \dfrac{q^T U_k s_j}{\|U_k^T q\|_2 \|s_j\|}$ دقیقاً باشد همان برای $\cos\varphi_j$.

و ...

Q.E.D

سوال یازدهم)

الف) طبق تجزیه SVD داریم $A = U\Sigma V^*$

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix} \quad , \quad \Sigma^\dagger = \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_r \\ & & \end{bmatrix}$$

از طرفی نشان دادیم که

$$(I) \quad A = U\Sigma V^* = \sigma_1 u_1 v_1^* + \dots + \sigma_r u_r v_r^*$$

حال برای $A^\dagger$ داریم: $A^\dagger = (U\Sigma V^*)^\dagger = V\Sigma^\dagger U^*$

و کاملا مشابه به تفریقی که در I داشتیم می توان نوشت:

$$(II) \quad A^\dagger = \frac{1}{\sigma_1} v_1 u_1^* + \dots + \frac{1}{\sigma_r} v_r u_r^*$$

در واقع صرفاً جای V با U عوض شد و ماتریس ها ترانهاده و مزدوج شدند و عناصر قطری در $\Sigma^\dagger$ از σ_i در Σ به $1/\sigma_i$ ها تغییر یافتند (از ضرب استفاده کردیم).

Q.E.D

ب) 1- $AA^\dagger A = A$ ← طبق SVD داریم: $A = U\Sigma V^*$ که V و U یکانی اند.

$$\Rightarrow AA^\dagger A = (U\Sigma \underbrace{V^*)(V}_{I}\Sigma^\dagger \underbrace{U^*)(U}_{I}\Sigma V^*)$$

$$\Rightarrow U\Sigma\Sigma^\dagger\Sigma V^* = U\Sigma V^* = A \quad \square$$

دقت کنید Σ $\Sigma^\dagger$ یک بلوک I می سازند و همان برای ما اهمیت دارد.

2- $A^\dagger A A^\dagger = A^\dagger \Leftarrow$

$$(V\Sigma^\dagger \underbrace{U^*)(U}_{I}\Sigma \underbrace{V^*)(V}_{I}\Sigma^\dagger U^*)$$

$$= V\Sigma^\dagger\Sigma\Sigma^\dagger U^* = V\Sigma^\dagger U^* = A^\dagger \quad \square$$

3- $(AA^\dagger)^* = AA^\dagger \Leftarrow$

$$(U\Sigma \underbrace{V^* V}_{I}\Sigma^\dagger U^*)^* = (U\Sigma\Sigma^\dagger U^*)^*$$

$$= U(\Sigma\Sigma^\dagger)^* U^* = U(\Sigma\Sigma^\dagger)U^* \xrightarrow{V^*V = I} U\Sigma I \Sigma^\dagger U^*$$

$$= \underbrace{U\Sigma V^*}_{A}\underbrace{V\Sigma^\dagger U^*}_{A^\dagger} = AA^\dagger \quad \square$$

4- $(A^\dagger A)^* = A^\dagger A \Leftarrow$

$$(V\Sigma^\dagger \underbrace{U^* U}_{I}\Sigma V^*)^* = (V\Sigma^\dagger\Sigma V^*)^*$$

$$= V(\Sigma^\dagger\Sigma)^* V^* = V(\Sigma^\dagger\Sigma)V^* \xrightarrow{U^*U = I} V\Sigma^\dagger I \Sigma V^*$$

$$= V\Sigma^\dagger U^* U\Sigma V^* = A^\dagger A \quad \square$$

ادامه سوال یازدهم)

ج) $C(A^{\dagger}) = C(A^{T}) \leftarrow$ یعنی فضای سطری A با فضای ستونی شبه وارون یا $A^{\dagger}$ برابر است.

ولی در $\mathbb{F}$ ها هم تساوی داریم فقط A^* داریم.

الف) $A \in M_{mn}(\mathbb{R})$ (چون در صورت سوال A^T مطرح شده فرض کردیم در $\mathbb{R}$ است).

طبق قضیه اساسی جبرخطی داریم:

$$\mathbb{R}^n = R(A^T) \oplus N(A)$$
$$\mathbb{R}^m = R(A) \oplus N(A^T)$$

حال طبق تجزیه SVD $\Leftarrow$

$$A = U\Sigma V^T$$
$$= [U_R \quad U_N] \left[\begin{array}{ccc|c} \delta_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \delta_r & \\ \hline & 0 & & 0 \end{array}\right] \begin{bmatrix} V_R^T \\ V_N^T \end{bmatrix}$$

در واقع با توجه به فرم خاص ماتریس Σ، در حقیقت U_N و V_N چه صفر باشد چه هر مقدار دیگری برای ما بی اهمیت است.

$$\Rightarrow [u_1 \ \ldots \ u_r \ u_{r+1} \ \ldots \ u_m] \begin{bmatrix} S_{r\times r} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \\ v_{r+1}^T \\ \vdots \\ v_n^T \end{bmatrix}$$

Column Vectors	Span
$u_1 \ldots u_r$	$R(A)$
$v_1 \ldots v_r$	$R(A^T)$
$u_{r+1} \ldots u_m$	$N(A)$
$v_{r+1} \ldots v_n$	$N(A^T)$

در نتیجه طبق قضیه بالا داریم $\Leftarrow$

حال طبق SVD: $\leftarrow A^{\dagger} = V\Sigma^{\dagger}U^*$

$$= [V_R \quad V_N] \begin{bmatrix} S^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_R^T \\ U_N^T \end{bmatrix}$$

$$\Rightarrow A^{\dagger} \in M_{nm}(\mathbb{R})$$

و حال برای ترانهاده هم داریم:

$$A^T = V\Sigma^T U^T$$

$$= [V_R \quad V_N] \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_R^T \\ U_N^T \end{bmatrix} \Rightarrow A^T \in M_{nm}(\mathbb{R})$$

و این نتیجه می دهد که $C(A^{\dagger}) = C(A^T)$

Q.E.D

Subject:
Date

سوال یازدهم

$x^{\dagger} := V^* \Sigma^{\dagger} U b = A^{\dagger} b$ $\longrightarrow$ This $x^{\dagger}$ is a solution for $A^*Ax = A^*b$

Proof $\Rightarrow A^*Ax^{\dagger} = A^*AA^{\dagger}b$

$= V\Sigma^* U^* U \Sigma V^* V \Sigma^{\dagger} U^* b$

$= V\Sigma^* \Sigma \Sigma^{\dagger} U^* b$

$= V\Sigma^* U^* b = A^* b$ $\Rightarrow x^{\dagger} = A^{\dagger} b \in \underset{x}{\operatorname{argmin}} \|b - Ax\|$

در واقع همان است مجموعه ای از x ها عبارت بالا را $\min$ کند و $x^{\dagger}$ یکی از آنها است؛ حال گوییم هر $z \in \underset{x}{\operatorname{argmin}} \|b - Ax\|$ ، نرم $\|z\| \geq \|x^{\dagger}\|$ خواهد بود و $x^{\dagger}$ بهترین جواب می باشد.

فرض کنید چنین z ای وجود داشته باشد: $z \in \underset{x}{\operatorname{argmin}} \|b - Ax\|$

$\Rightarrow A^*Az = A^*b$ و $A^*A(\underbrace{A^{\dagger}b}_{x^{\dagger}}) = A^*b$

$\ominus$
$\Rightarrow$

$A^*A(z - A^{\dagger}b) = 0 \Rightarrow \underbrace{z - A^{\dagger}b}_{:= v} \in N(A^*A)$

$\Rightarrow z = v + \underbrace{A^{\dagger}b}_{\in C(A^{\dagger})}$, $v \in N(A^*A)$

عمود برهم اند

$\underbrace{\|z\|_2^2 = \|v\|^2 + \|A^{\dagger}b\|^2}$ $\geq \underbrace{\|A^{\dagger}b\|^2}_{x^{\dagger}}$ $\square$

دقت کنید اگر دستگاه سازگار باشد و b در فضای ستونی A قرار بگیرد $\|z\|$ برابر صفر خواهد شد و ناچار $\|v\|$ نیز صفر می شود و به رابطه $0 \geq 0$ می رسیم و حکم باز هم برقرار است $\square$

طبق مطالب گفته شده ! $\underline{C(A^{\dagger})} \perp \underline{N(A^*A)}$

با r ستون اول V ساخته می شود. (پایه)

با $n - r$ ستون آخر V ساخته می شود (پایه)

Q.E.D

ادامه سوال دوازدهم)

به طور دقیق تر ← حال اگر $Ax=b$ سازگار باشد باز هم ادعا می کنیم $x^\dagger = A^\dagger b$

کمترین نرم اقلیدسی را دارد ←

فرض کنید $Ax_0 = b$ باشد ، حال با توجه به سوال یازدهم می دانیم

$A = AA^\dagger A$ است و خواهیم داشت

$$b = Ax_0 = AA^\dagger A x_0 = AA^\dagger b$$

پس وقتی دستگاه سازگار است ، جواب $x = A^\dagger b$ باعث حل دستگاه می شود ✓

برای آنکه ببینیم $x^\dagger = A^\dagger b$ جوابی با کمترین نرم اقلیدسی است ، جواب عمومی $A^\dagger b + N(A)$ (که یک جواب خاص به علاوه جواب عمومی برای همگن همان است (مانند آنچه در معادلات دیفرانسیل داشتیم)) را در نظر می گیریم و تعریف می کنیم $z = A^\dagger b + v$ که $v \in N(A)$. به راحتی می توان دید که $A^\dagger b \in \underbrace{C(A^\dagger) = C(A^T)}_{\text{طبق سوال ۱۱}}$ است و در نتیجه

$A^\dagger b \perp v$ و طبق مطالب فصل 3 $\Rightarrow$

$$\|z\|^2 = \|A^\dagger b + v\|^2 = \|A^\dagger b\|^2 + \|v\|^2 \geq \|A^\dagger b\|^2$$

حالت تساوی وقتی به وجود می آید که $v=0$ و $A^\dagger b$ در این صورت جواب یکتایی است و چون یکتاست پس عملاً "کمترین نرم اقلیدسی را دارد (در واقع قرار نیست با چیزی مقایسه شود).

Q.E.D

سوال اول (حل مقدار ویژه ها و بردار ویژه ها)

$A^{[k]} =$ ماتریسی که هر درایه از آن به توان k برسد.

فرض کنید $f(x)$ چندجمله‌ای ویژه A باشد $f(x) = x^n + a_1 x^{n-1} + \dots + a_n$

$$f(A) = 0 \Rightarrow A^n + a_1 A^{n-1} + \dots + a_n I = 0$$

$$\xrightarrow{\times A} A^{n+1} + a_1 A^n + \dots + a_n A = 0$$

← حال با $A^{[k]}$ ها جابه‌جای کنیم طبق فرض:

$$A^{[n+1]} + a_1 A^{[n]} + \dots + a_n A^{[1]} = 0$$

یعنی ← هر درایه‌ای از A را در نظر بگیریم در این چند جمله‌ای صفر می‌شوند

پس هر درایه از A ریشه‌ای از چند جمله‌ای $x f(x)$ است. (فرض سوال برای 1 تا n را پوشش می‌دهد)

حال می‌گوییم به ازای $n+1 \leq m$، $A^m = A^{[m]}$

کافی است از این تساوی استفاده کنیم.

* وقتی ریشه‌ای از $x f(x)$ باشد، ریشه‌ای از $x^m f(x)$ نیز است.

$$\Rightarrow \underbrace{A^{[m+n]} + a_1 A^{[m+n-1]} + \dots + a_n A^{[m]} = 0}_{(\mathrm{I})}$$

چون هر درایه از A ← ریشه‌ای از چند جمله‌ای $x^m f(x)$ است.

از طرفی $x^m f(x)$ را اگر A را در آن قرار دهیم صفر می‌شود (طبق قضیه کیلی همیلتون)

$$\Rightarrow A^m f(A) = 0$$

از طرفی $f(A) = A^n + \dots + a_n I = 0$

$$\underbrace{A^{m+n} + \dots + a_n A^m = 0}_{(\mathrm{II})} \xleftarrow{\times A^m} \text{(روی } m)$$

حال کافی است استقرا بزنیم (به ازای هر عدد طبیعی m عبارات (I) و (II) برقرارند و جمله‌های برابرند.

در حالت پایه $m=2$ چون نظیر به نظیر می‌توان (I) و (II) را برابر قرار داد ⇐ یعنی:

$$a_n A^{[2]} = a_n A^2 \quad \dots \quad a_1 A^{[n+1]} = a_1 A^{n+1}$$

نظیر به نظیر برابرند و جمع تساوی‌های (I) و (II) صفر اند بنابراین باید عبارات آخر یعنی

$$A^{m+n} = A^{[m+n]} \Rightarrow m=2 \Rightarrow A^{n+2} = A^{[n+2]}$$

حال برای $m=3$ هم استدلال مشابهی داریم و در نهایت باید $A^{n+3} = A^{[n+3]}$ شود.

در نهایت اگر برای $m-1$ حکم درست باشد جمله‌های نظیر به نظیر را برابر می‌گذاریم و جمله آخر نیز مجبور است باهم برابر باشند و حکم اثبات می‌شود

Q.E.D

سوال دوم قسمت بردار ویژه ها و مقدار ویژه ها) احتمالاً فرض سوال این بوده که $A\in M_n(\mathbb{R})$
است ولی اگر هم نباشد به نظر ایرادی به وجود نمی آید.

برهان خلف) اگر فرض کنیم حداقل ۱ مقدار ویژه غیر حقیقی داریم. از طرفی اگر A دارای عناصر حقیقی باشد چند جمله ای مشخصه ضرایب حقیقی خواهد داشت و ممکن نیست که شما یک مقدار ویژه مختلط داشته باشیم پس تمام ریشه ها حقیقی و یعنی همه مقدار ویژه ها حقیقی اند.

حال در نظر بگیرید که: $\det(xI-A)=D$ (I) چند جمله ای مشخصه

(II) A ماتریس پوچ توان نیست پس همه مقادیر ویژه با هم صفر نیستند.

برای $j=i$: $0\ge a_{ij}a_{ji} \Rightarrow 0\ge a_{ii}^2$ و چون فرض کرده ایم ماتریس حقیقی است پس $a_{ii}=0$ یعنی عناصر قطر صفر هستند. در این صورت برای چند جمله ای مشخصه داریم:

$$\det(xI-A)=x^n+\underbrace{(-\text{trace})x^{n-1}}_{=0}-\Big(\sum_{i<j}a_{ij}a_{ji}\Big)x^{n-2}+\dots$$

monic بودن

این موضوع در کلاس نشان داده شد.

پس D را که تعریف کردیم در (I)

دقت کنید طبق مطالب کلاس برای چند جمله ای مشخصه داریم:

$$f(x)=\det(xI-A)=x^n+\Big(-\sum_{i=1}^n\lambda_i\Big)x^{n-1}+\overbrace{\Big(\sum_{i<j}\lambda_i\lambda_j\Big)}^{=-\sum_{i<j}a_{ij}a_{ji}}x^{n-2}+\dots$$

در واقع برای ضریب c_k داریم $\Rightarrow c_k=(-1)^k\sum_{i_1=1}^n\sum_{i_2=1}^n\dots\sum_{i_k=1}^n\lambda_{i_1}\dots\lambda_{i_k}$

$i_1<i_2<\dots<i_k$ for $k=1,2,\dots,n$

حال برای هر نوع $a_{ij}a_{ji}\le 0$ پس $-a_{ij}a_{ji}\ge 0$

و طبق بسط چند جمله ای داریم:

$$0=\Big(\sum_{i=1}^n\lambda_i\Big)^2=\sum_{i=1}^n\lambda_i^2+2\sum_{i<j}\lambda_i\lambda_j$$

$$\Rightarrow 2\sum_{i<j}\lambda_i\lambda_j=2\sum_{i<j}-a_{ij}a_{ji}\ge 0$$

حال طبق (II) که A پوچ توان نیست پس $\sum_{i=1}^n\lambda_i^2>0$ خواهد شد

ادامه سوال دوم بخش بردار ویژه ها

پس در نهایت می توان نوشت <=

$$\sum_{i=1}^{n} \lambda_i^2 + 2\sum_{i<j} \lambda_i \lambda_j > 0$$

درحالی که ما فرض کرده بودیم

این مقدار صفر است و این تناقض فرض خلف را رد کرده

و حکم اثبات می شود.

Q.E.D

سوالات بخش SVD - $\quad A = U\Sigma V^T$

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$$

۱- ابتدا به یافتن σ_i ها با یافتن مقادیر ویژه AA^* می‌پردازیم (دقت کنید A حقیقی مقدار است پس $A^* = A^T$).

$$AA^T = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

$$\lambda_1 = 25 \quad \lambda_2 = 9$$

$$\Rightarrow \det(AA^T - \lambda I) = \lambda^2 - 34\lambda + 225 = (\lambda - 25)(\lambda - 9)$$

و در این صورت $\sigma_1 = \underbrace{5}_{\sqrt{25}}$, $\sigma_2 = \sqrt{9} = 3$

حال به محاسبه بردار تکین های سمت راست یعنی ستون های V می پردازیم. برای این کار بردارهای ویژه متعامد A^TA را می یابیم. یا به طور متقارن می توانیم این کار را برای ستون های U انجام دهیم

برای A^TA داریم

$$A^TA = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

مقادیر ویژه A^TA برابر با 25 , 9 , 0 هستند و از آنجایی که ماتریس متقارن است بردار ویژه ها متعامدند.

$$\begin{vmatrix} 13-x & 12 & 2 \\ 12 & 13-x & -2 \\ 2 & -2 & 8-x \end{vmatrix} = 2(-24 - 2(13-x)) + 2(-2(13-x) - 24) + (8-x)((13-x)^2 - 144) = 0 \Rightarrow$$

مقادیر ویژه $\rightarrow \quad x(x-25)(x-9) = 0$

$$\lambda = 25 \Rightarrow A^TA - 25I = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix}$$

حال اگر به صورت سطری پلکانی بنویسیم به شکل مقابل در خواهد آمد: $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

حال در Null space این ماتریس بردار یکه به صورت $v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$ خواهد بود.

حال برای $\lambda = 9$ نیز داریم:

$$A^TA - 9I = \begin{bmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{bmatrix}$$

ماتریس تحویل شده سطری پلکانی (با اعمال سطری مقدماتی) $\Rightarrow \begin{bmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{bmatrix}$

و در فضای پوچ آن بردار یکه $v_2 = \begin{bmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix}$ خواهد بود.

ادامه سوال ۱ بخش SVD ←

در نهایت می‌توان برای بردار سوم v_3 یک بردار یافت که بر دو بردار قبلی عمود باشد.

برای این کار داریم: اگر $v_1 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ عمود باشد باید $-a = b$ باشد.

از طرفی $v_2^T v_3 = 0 \Rightarrow$

$$\frac{2a}{\sqrt{18}} + \frac{4c}{\sqrt{18}} = 0 \Rightarrow -a = 2c$$

و از طرفی باید یکه‌تر باشد پس کافی است $a = 2/3$ انتخاب شود $\Rightarrow v_3 = \begin{bmatrix} a \\ -a \\ -a/2 \end{bmatrix}$

$$\Rightarrow v_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix}$$

در این صورت داریم:

$$A = U\Sigma V^T = U \begin{bmatrix} \overset{\sigma_{1A}}{5} & 0 & 0 \\ 0 & \underset{\sigma_{2A}}{3} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

حال طبق رابطه‌ای که در کلاس به آن پرداخته شد: $\sigma u_i = A v_i \Rightarrow u_i = \frac{A v_i}{\sigma}$

برای محاسبه U از آن استفاده می‌کنیم.

$$\Rightarrow U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

=> Singular Value DeComposition for A :

$$A = U\Sigma V^* = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

Q.E.D