به نام خدا

علی قاسم‌زاده   ۴۰۱۱۰۶۳۳۹   تمرین اول بهینه‌سازی محدب

۱- الف) در ابتدا فرض می‌کنیم مقادیر ویژه‌ی A و B مرتب شده هستند یعنی: $\lambda_1(A) \geq \dots \geq \lambda_n(A)$, $\lambda_1(B) \geq \dots \geq \lambda_n(B)$

حالا چون ماتریس A متقارن است دارای تجزیه‌ی طیفی به صورت $A = Q_A \Lambda_A Q_A^T$ است.

حالا داریم:

$$Tr(AB) = Tr(Q_A \Lambda_A Q_A^T B) = Tr(\Lambda_A Q_A^T B Q_A)$$

طبق خاصیت شیفت چرخشی Tr

حالا مقادیر ویژه‌ی B و $\tilde{B} = Q_A^T B Q_A$ یکسان هستند زیرا که ماتریس Q_A یک ماتریس یکانی هست و $Q_A^{-1} = Q_A^T$

← ماتریس‌های B و $\tilde{B}$ مشابه‌اند و مقادیر ویژه‌ی یکسانی خواهند داشت، به عبارتی:

$$|B - \lambda I| = 0 \longrightarrow$$

$$|Q_A^T B Q_A - \lambda I| = 0 \longrightarrow |Q_A^T B Q_A - \lambda Q_A^T Q_A| = |Q_A^T (B - \lambda I) Q_A| = 0$$

Q_A و Q_A^T، فول رنک و مربعی هستند، پس تنها حالتی که دترمینان صفر شود، این است که دترمینان $|B-\lambda I|$ برابر صفر شود، زیرا که داریم:

$$|Q_A^T B Q_A - \lambda I| = |Q_A^T (B - \lambda I) Q_A| = \underbrace{|Q_A^T|}_{1} \, |B - \lambda I| \, \underbrace{|Q_A|}_{1} = |B - \lambda I|$$

← هر مقدار ویژه‌ای که برای B داشته باشیم برای $\tilde{B}$ نیز داریم و برعکس.

حالا بدون کم شدن از کلیت مسئله A را قطری در نظر می‌گیریم، چون اگر هم قطری نباشد، $Tr(\Lambda \tilde{B})$ را داریم حساب می‌کنیم و در این حالت $\tilde{B}$ شامل $\lambda_i(B)$ است و Λ شامل $\lambda_i(A)$. حالا با این فرض داریم:

$$A = diag(\lambda_1(A), \dots, \lambda_n(A)) \Rightarrow Tr(AB) = \sum_{i=1}^{n} \lambda_i(A) B_{ii}$$

← حکم سوال معادل نشان دادن عبارت زیر است:

$$\sum_{i=1}^{n} \lambda_i(A) B_{ii} \leq \sum_{i=1}^{n} \lambda_i(A) \lambda_i(B)$$

← ابعاد ماتریس

برای نشان دادن این عبارت روی n استقرا می‌زنیم یعنی فرض می‌کنیم به ازای n-1 این رابطه برقرار است.

برای $n=1$ که نامساوی حالت تساوی اش بدست می‌آید ← حکم برقرار است.

فرض کنیم به ازای هر ماتریس $(n-1) \times (n-1)$ این رابطه برقرار است، نشان می‌دهیم که به ازای $n \times n$ هم برقرار است،

$$Tr(AB) = \sum_{i=1}^{n} \lambda_i(A) B_{ii} = \sum_{i=1}^{n-1} (\lambda_i(A) - \lambda_n(A)) B_{ii} + \sum_{i=1}^{n} \lambda_n(A) B_{ii} = tr(\bar{\Lambda} - \bar{B}) + \lambda_n(A)\, tr(B$$

$\bar{\Lambda} =$ ماتریس $(n-1)*(n-1)$ سمت چپ بالای Λ

$\bar{B} =$ ماتریس $(n-1)\times(n-1)$ سمت چپ بالای B

حالا ماتریس
$$P = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots 1 & 0 \end{bmatrix}_{(n-1)\times n}$$
را در نظر بگیرید، داریم که $\bar{B} = PBP^T$

خود P سطر آخر را از B جدا می‌کند، P^T ستون آخر را از B جدا می‌کند. همچنین داریم که $PP^T = I$

⇐ برای هر بردار u داریم که،

$$\|P^T u\|_2^2 = u^T P P^T u = u^T I u = u^T u = \|u\|_2^2 \Rightarrow \|P^T u\|_2 = \|u\|_2$$

حالا u_i ها را بردار ویژه متناظر $\lambda_i(\bar{B})$ در نظر بگیرید و $S_j = span\{u_1, \ldots, u_j\}$،

$$\Rightarrow \lambda_j(\bar{B}) = \min_{x \in S_j, \|x\|_2 = 1} x^T \bar{B} x = \min_{x \in S_j, \|x\|_2 = 1} x^T P B P^T x$$

حالا $P^T x$ را y می‌نامیم و می‌دانیم که چون $\|x\|_2 = 1$ ⇐ $\|P^T x\|_2 = 1$ ⇐ $\|y\|_2 = 1$ ⇐ داریم،

طبق سوال ۵ که بعدا اثبات می‌کنیم، داریم، ↓

$$\lambda_j(B) = \min_{x \in S_j, \|y\|_2 = 1} y^T B y \le \underbrace{\max_{\{S : dim(S) = j\}} \min_{\{y \in S, \|y\| = 1\}} y^T B y}_{\lambda_j(B)} = \lambda_j(B)$$

$$\Rightarrow \lambda_j(\bar{B}) \le \lambda_j(B)$$

$$\Rightarrow tr(AB) = tr(\bar{A}\bar{B}) + \lambda_n\, tr(B) \leq \sum_{i=1}^{n-1}\left(\lambda_i(A) - \lambda_n(A)\right)\lambda_i(\bar{B}) + \lambda_n(A)\sum_{i=1}^{n}\lambda_i(B)$$

$$\leq \sum_{i=1}^{n-1}\left(\lambda_i(A) - \lambda_n(A)\right)\lambda_i(B) + \lambda_n(A)\sum_{i=1}^{n}\lambda_i(B) = \sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)$$

$$\Rightarrow \underline{tr(AB) \leq \sum_{i=1}^{n}\lambda_i(A)\lambda_i(B)}$$

ب) ابتدا بدون کم شدن از کلیت مسئله فرض می‌کنیم $m \leq n$ اثبات $m > n$ نیز مشابه همین است.

$$A = U_A \Sigma_A V_A^T \Rightarrow tr(AB) = tr(U_A \Sigma_A V_A^T B) = tr(\Sigma_A \underbrace{V_A^T B U_A}_{\bar{B}})$$

مقادیر تکین $\bar{B}$ و B برابر است زیرا مقادیر تکین B عبارت است از جذر مقادیر ویژه $B B^T$ ، مقادیر تکین $\bar{B}$ عبارت است از جذر مقادیر ویژه $\bar{B}\bar{B}^T$. زیرا که داریم:

$$\bar{B}\bar{B}^T = U_A B^T V_A^T V_A B U_A^T = U_A B B U_A^T \rightarrow$$

$$|\bar{B}^T\bar{B} - \lambda I| = 0 \longrightarrow |U_A B^T B U_A^T - \lambda I| = |U_A (B^T B - \lambda I) U_A^T| = 0$$

$$= \underbrace{|U_A|}_{1}\,|B^T B - \lambda I|\,\underbrace{|U_A^T|}_{1} = |B^T B - \lambda I| \rightarrow$$

همان مقادیر ویژه $B^T B$, $\bar{B}^T\bar{B}$ برابر اند ← مقادیر تکین B, $\bar{B}$ برابر اند.

⇐ A را به صورت زیر در نظر می‌گیریم بدون کم شدن از کلیت مسئله:

$$A = \begin{bmatrix} \sigma_1(A) & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \sigma_m(A) \cdots & 0 \end{bmatrix}_{m\times n} \longrightarrow Tr(A^T B) = \sum_{i=1}^{m}\sigma_i(A) B_{ii}$$

حالا باید نشان دهیم:

$$\sum_{i=1}^{m}\sigma_i(A) B_{ii} \leq \sum_{i=1}^{m}\sigma_i(A)\sigma_i(B)$$

روی m استقرا می‌زنیم، پایه، $m=1$ ، نامساوی به وضوح حالت تساوی اش رخ می‌دهد.

حالا فرض کنیم برای $(m-1)\times n$ نامساوی برقرار باشد، در نتیجه باید نتیجه بگیریم برای $m\times n$ $(1\le m\le n)$ نیز برقرار است:

$$Tr(A^TB) = \sum_{i=1}^{m-1}\left(\sigma_i(A)-\sigma_m(B)\right)B_{ii} + \sum_{i=1}^{m}\sigma_m(A)B_{ii} = tr(\bar{\Sigma}^T\bar{B}) + \sigma_m(A)\sum_{i=1}^{m}B_{ii}$$

$\bar{\Sigma}$ ماتریس $(m-1)\times n$ سمت چپ بالای Σ

$$\begin{bmatrix} \sigma_1(A)-\sigma_m(A) & \cdots & & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & \sigma_{m-1}(A)-\sigma_m(A) & 0\cdots 0 \end{bmatrix}$$

$\bar{B}$ هم ماتریس $(m-1)\times n$ سمت چپ بالای B.

حالا مشابه سوال قبلی ماتریس $(m-1)\times m$ ، P را در نظر می‌گیریم:

$$P = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \Rightarrow \bar{B} = PB$$

حالا داریم:

$$\bar{B}\bar{B}^T = PBB^TP^T$$

حالا از همان چیزی که در الف نشان دادیم استفاده می‌کنیم، خواهیم داشت:

$$\lambda_j(\bar{B}\bar{B}^T) \le \lambda_j(BB^T) \longrightarrow \sqrt{\lambda_j(\bar{B}\bar{B}^T)} \le \sqrt{\lambda_j(BB^T)} \Rightarrow \sigma_j(\bar{B}) \le \sigma_j(B)$$

طبق فرض استقرا و نابرابری $\sum_{i=1}^{m}B_{ii} \le \sum_{i=1}^{m}\sigma_i(B)$

$$\Rightarrow Tr(A^TB) = Tr(\bar{\Sigma}^T\bar{B}) + \sigma_m(A)\sum_{i=1}^{m}B_{ii} \le \sum_{i=1}^{m-1}\left(\sigma_i(A)-\sigma_m(A)\right)\sigma_i(\bar{B}) + \sigma_m(A)\sum_{i=1}^{m}\sigma_i(B)$$

$$\le \sum_{i=1}^{m-1}\left(\sigma_i(A)-\sigma_m(A)\right)\sigma_i(B) + \sigma_m(A)\sum_{i=1}^{m}\sigma_i(B) = \sum_{i=1}^{m}\sigma_i(A)\sigma_i(B)$$

یک شرط برقراری حالت تساوی این است که A و B باید بردارهای ویژه یکسانی داشته باشند.

یعنی بتوان B نوشت: $A = U\Sigma_AV^T$, $B = U\Sigma_BV^T \rightarrow Tr(A^TB) = Tr(\Sigma_A^T\Sigma_B) \rightarrow \sum\sigma_i(A)\sigma_i(B)$

⇐ در این حالت تساوی اتفاق می افتد

ج) طبق ب داریم که $Tr(A^TB) \le \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)\sigma_i(B)$

$$Tr(A^TB) \le \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)\sigma_i(B) \le \sum_{i=1}^{\min} \sigma_1(A)\,\sigma_i(B) = \underbrace{\sigma_1(A)}_{\|A\|_2}\sum_{i=1}^{\min\{m,n\}} \sigma_i(B)$$

$$\Rightarrow B_* = \sup\{tr(A^TB) \mid \|A\|_2 \le 1\} \le \sup\left\{\|A\|_2 \sum_{i=1}^{\min\{m,n\}} \sigma_i(B) \;\middle|\; \|A\|_2 \le 1\right\} \le \sum_{i=1}^{\min\{m,n\}} \sigma_i(B)$$

از طرفی: $B = U_B \Sigma_B V_B^T \longrightarrow U_B, V_B^T \longrightarrow$ orthonormal matrices

حالا اگر قرار دهیم $A = U_B V_B^T$ ، خواهیم داشت:

$$Tr(A^TB) = Tr(V_B \underbrace{\cancel{U_B^T U_B}}_{I} \Sigma_B V_B^T) = tr(\Sigma_B \underbrace{\cancel{V_B^T V_B}}_{I}) = tr(\Sigma_B) = \sum_{i=1}^{\min\{m,n\}} \sigma_i(B)$$

در این حالت

$$\|A\|_2 = \|U_B V_B^T\|_2 = \|U_B I V_B^T\|_2 = \sigma_1(I) = 1$$

در نتیجه یک مقدار A برای ترم بالا $Tr(A^TB)$ پیدا کردیم که اگر A را برابر آن مقدار قرار دهیم $Tr(A^TB)$ مقدار ترم بالایش را می دهد ⇐

$$\|B\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i(B)$$

د) کاربردهای نرم اتمی در بهینه سازی:

- بازیابی سیگنال های نادر: یکی از کاربردهای نرم اتمی در مسائل بازیابی سیگنال های نادر است. مثل مدل های فشرده سازی و بازیابی سیگنال. سیگنال به عنوان ترکیبی از چند اتم انتخابی از یک مجموعه اتمی مدل می‌شود. با استفاده از نرم اتمی مسئله بازیابی را به یک مسئله بهینه سازی محدب تبدیل می کند و حلش می کند.

- مدل سازی سیگنال های پیوسته: در بهینه سازی های مرتبط با بازیابی سیگنال پیوسته استفاده می‌شود و نرم اتمی این اجازه را می دهد که سیگنال های با ساختار ساده را بهتر و دقیق تر مدل کنیم.

- فشرده سازی اطلاعات : نرم اتمی به عنوان یک ابزار برای بازسازی دقیق سیگنال های نادر از داده های ناقص استفاده می شود. این روش در فشرده سازی تصویر و پردازش صورت و سیگنال اهمیت دارد.

۲- الف) برای اینکه نشان دهیم $x-x^* \in N(A)$ کافی است نشان دهیم $A(x-x^*)=0$

داریم که $Ax^*=b$ ، x هر یک جواب برای $Ax=b$ هست پس داریم،

$$A(x-x^*) = b-b=0 \longrightarrow x-x^* \in N(A)$$

ب) $x^* = A^T(AA^T)^{-1}b$ ، پس x^* در $\text{Range}(A^T)$ قرار دارد $\leftarrow$ یعنی x^* عضو $\text{Rowspace}(A)$ است، همچنین $x-x^* \in N(A)$ است طبق الف $\leftarrow$ خواهیم داشت که

$x^* \perp x-x^*$ طبق قضیه اساسی جبر خطی.

ج)

$$\|x\|_2^2 = \|x-x^*+x^*\|_2^2 = \|x-x^*\|_2^2 + \|x^*\|_2^2 + \underbrace{2(x-x^*)^T(x^*)}_{\text{طبق ب صفر است}}$$

$$\Rightarrow \|x\|_2^2 = \|x-x^*\|_2^2 + \|x^*\|_2^2 \Rightarrow \|x\|_2^2 \geq \|x^*\|_2^2$$

د) می‌خواهیم تصویر روی A^T را بدست آوریم، یعنی داریم $A^Tx=b$ و برای حل این باید نزدیک‌ترین نقطه را انتخاب کنیم که نزدیک‌ترین نقطه تصویر b روی A^T است،

حالا داریم $\min_x \|b-A^Tx\|_2 \leftarrow \min_x \|b-A^Tx\|_2^2$

$$\longrightarrow \min_x (b-A^Tx)^T(b-A^Tx) = \min_x \underbrace{+x^TAA^Tx - x^TAb - b^TA^Tx + b^Tb}_{F}$$

$$\longrightarrow \frac{\partial}{\partial x}F = +2AA^Tx - 2Ab = 0 \rightarrow AA^Tx = Ab \rightarrow x = (AA^T)^{-1}Ab$$

حالا تصویر x روی A^T برابر است با A^Tx یعنی $A^T(AA^T)^{-1}Ab$ که برابر است با Px

$\leftarrow P = A^T(AA^T)^{-1}A$ ، حالا نگاشت متناظر پوچی عبارت است از:

هر بردار X را ما می توان به صورت جمع تصویرش روی $Null(A)$, $Row(A)$ ~~Row(A)~~ نوشت :

$$X = P_A X + P_{null} X \longrightarrow (I - P_A) X = P_{null} X \longrightarrow$$

$$(I - A^T(AA^T)^{-1}A)X = P_{null} X \longrightarrow P_{null} = I - A^T(AA^T)^{-1}A$$

ماتریس تصویر روی nullspace A ،

Projection

3- الف)

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \longrightarrow \left[\begin{array}{cc|cc} A & B & I & 0 \\ C & D & 0 & I \end{array}\right] \xrightarrow[\text{اول}]{\text{بلوک سطری دوم منهای بلوک سطری } CA^{-1}}$$

$$\left[\begin{array}{cc|cc} A & B & I & 0 \\ 0 & D - CA^{-1}B & -CA^{-1} & I \end{array}\right] \xrightarrow{\text{بلوک سطری اول منهای } B(D-CA^{-1}B)^{-1} \text{ برابر بلوک سطری دوم}}$$

از سمت چپ

$$\left[\begin{array}{cc|cc} A & 0 & I + \cancel{CA^{-1}}B(D-CA^{-1}B)^{-1}CA^{-1} & -B(D-CA^{-1}B)^{-1} \\ 0 & (D-CA^{-1}B) & -CA^{-1} & I \end{array}\right] \xrightarrow[\text{بلوک سطر دوم } (D-CA^{-1}B)^{-1}]{\text{بلوک سطری اول } \times A^{-1}}$$

(1)

$$\left[\begin{array}{cc|cc} I & 0 & A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1} \\ 0 & I & -(D-CA^{-1}B)^{-1}CA^{-1} & (D-CA^{-1}B)^{-1} \end{array}\right]$$

حالا یکبار دیگر به روشی دیگر rref می کنیم :

$$\left[\begin{array}{cc|cc} A & C & I & 0 \\ B & D & 0 & I \end{array}\right] \xrightarrow[\text{دوم}]{\text{بلوک سطری اول منهای} \ D^{-1}C \ \text{برابر بلوک سطری}} \left[\begin{array}{cc|cc} A-BD^{-1}C & 0 & I & D^{-1}C \\ B & D & 0 & I \end{array}\right]$$

$$\xrightarrow[\text{(A-BD^{-1}C)^{-1}B \ \text{برابر بلوک سطری اول}}]{\text{بلوک سطری دوم منهای}} \left[\begin{array}{cc|cc} A-BD^{-1}C & 0 & I & D^{-1}C \\ 0 & D & -(A-BD^{-1}C)^{-1}B & I-D^{-1}C(A-BD^{-1}C)^{-1}B \end{array}\right] \xrightarrow[\text{بلوک سطری دوم ضربدر } D^{-1}]{\text{بلوک سطری اول ضربدر } (A-BD^{-1}C)^{-1}}$$

$$\rightarrow \left[\begin{array}{cc|cc} I & 0 & (A-BD^{-1}C)^{-1} & * \\ 0 & I & * & * \end{array}\right] \longrightarrow$$

②

هرجوری rref کنیم باید به یک خروجی برسیم $\Leftarrow$

$$①\text{و}② \Longrightarrow (A-BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1}$$

ب)

$$(I+uv^T)x = y \rightarrow x + uv^Tx = y \rightarrow y - u\underbrace{v^Tx}_{s} = x \rightarrow x = y - su$$

$$\xrightarrow{v^T} v^Tx = v^Ty - s\,v^Tu \rightarrow s = v^Ty - s\,v^Tu \rightarrow s(1+v^Tu) = v^Ty \rightarrow s = \frac{v^Ty}{1+v^Tu}$$

$$\rightarrow x = y - \frac{v^Ty}{1+v^Tu}u = y - u\frac{v^Ty}{1+v^Tu} = \left(I - \frac{uv^T}{1+v^Tu}\right)y \Rightarrow$$

$$x = \underbrace{\left(I - \frac{uv^T}{1+v^Tu}\right)}_{\text{وارون ماتریس}} y \longrightarrow \boxed{(I+uv^T)^{-1} = I - \frac{uv^T}{1+v^Tu}}$$

* شرایط معکوس پذیری هم $v^Tu \neq -1$ است که مخرج مخالف صفر شود

ع

$$(A+uv^T)x = y \longrightarrow Ax + uv^Tx = y \longrightarrow Ax = y - u\overset{s}{(v^Tx)} \longrightarrow Ax = y - su$$

$$\longrightarrow x = A^{-1}(y-su) = A^{-1}y - sA^{-1}u \xrightarrow{v^T} \underbrace{v^Tx}_{s} = v^TA^{-1}y - sv^TA^{-1}u$$

$$\longrightarrow s(1+v^TA^{-1}u) = v^TA^{-1}y \longrightarrow s = \frac{v^TA^{-1}y}{1+v^TA^{-1}u} \longrightarrow x = A^{-1}y - \frac{(A^{-1}u)\,v^TA^{-1}y}{1+v^TA^{-1}u}$$

$$\Rightarrow x = \underbrace{\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}y}{1+v^TA^{-1}u}\right)}_{\text{وارون ماتریس .}} y \longrightarrow (A+uv^T)^{-1} = \left(A^{-1} - \frac{A^{-1}uv^TA^{-1}y}{1+v^TA^{-1}u}\right)$$

۴- الف) ابتدا فرض می‌کنیم $A>0$ و $S>0$، نشان می‌دهیم که $M>0$:

$$M=\begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

می‌دانیم که با عملیات سطری PD بودن یک ماتریس عوض نمی‌شود و در فرم rref باید تمام عناصر روی قطر اصلی مثبت باشند.

← چون ماتریس فول رنک است، هیچ سطر صفری نداریم و از طرفی ماتریس مربعی است ← pivot ها همان عناصر روی قطر اصلی‌اند.

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \xrightarrow[\text{برابر بلوک اول از سمت چپ ضرب شود}]{\text{بلوک دوم سطری منهای } B^TA^{-1} \text{ برابر بلوک سطری اول}} \begin{bmatrix} A & B \\ 0 & C-B^TA^{-1}B \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & S \end{bmatrix} \xrightarrow[S^{-1}B \text{ برابر بلوک سطری دوم}]{S>0 \text{ حالا بلوک اول سطری منهای}}$$

$$\begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix} \xrightarrow[\text{سطری اول و دوم جدا از هم می‌زنیم}]{\text{حالا rref روی بلوک}}$$

چون A, S وارون‌پذیر هستند، در بلوک سطری اول A^{-1} در بلوک سطری دوم S^{-1} ضرب می‌کنیم،

$$\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

حاصل می‌شود ← تمام عناصر روی قطر اصلی 1 هستند و مثبت‌اند ← M مثبت معین است.

← if $A>0, S>0 \Rightarrow M>0$

حالا طرف دیگر این را نشان می‌دهیم یعنی، if $M>0 \Rightarrow S>0, A>0$

طبق تست سیلوستر می‌دانیم که هر بلوک روی قطر M را بگیریم دترمینانش مثبت است

$$\begin{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} & B \\ B^T & C \end{bmatrix}$$

تمام بلوک هایی که در A هستند، باید det شان مثبت باشد ← چون هر ماتریس با این ویژگی PD هست، $A>0$.

$$\longrightarrow \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \xrightarrow{\text{عملیات مانند بالا}} \begin{bmatrix} A & B \\ 0 & S \end{bmatrix} \longrightarrow \begin{bmatrix} I & A^{-1}B \\ 0 & S \end{bmatrix}$$

حالا برای نوشتن rref این ماتریس که معادل $rref(M)$ است، باید به ماتریسی برسیم که تمام pivot هایش مثبت باشد، در این عملیات ها ما روی بلوک پایینی حرکت می‌زنیم زیرا که اگر بلوک بالا را هم دخیل کنیم، بلوک صفرمان ناصفر می‌شود و دیگر فرم rref نداریم.

با انجام عملیات های سطری مقدماتی روی بلوک پایین، سمت چپ که تماماً صفر است تغییر نمی‌کند ولی S تغییر می‌کند و یک ماتریس بدست می‌آید با pivot های پیش

حالا اگر S فول رنک نباشد، حداقل یکی از سطرهایش در این عملیات های سطری مقدماتی کاملاً صفر می‌شود سمت چپ این سطر هم تا صفر اول صفر است ← M یک سطر تماماً صفر خواهد داشت ← M، PD نمی‌شود بلکه PSD می‌شود ← S باید فول رنک باشد تا $M>0$، حالا در فرم $rref(S)$، تمام عناصر روی قطر اصلی همان pivot ها هستند، این عناصر pivot های M نیز هستند ← چون تمام pivot ها M، مثبت اند ← تمام pivot های S هم مثبت اند ← S هم PD است ← $S>0$

$$\Longrightarrow \quad M>0 \quad \text{if and only if} \quad A>0,\ S>0$$

[redacted]

ب) اثباتی دیگر برای الف.

$$M=\begin{bmatrix} A & B \\ B^T & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ B^TA^{-1} & I \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & (A^T)^{-1}B \\ 0 & I \end{bmatrix}$$

$$= Q^T L Q$$

$$\text{if } M>0 : \forall x \quad x^TMx>0 \Rightarrow \forall x : \underbrace{x^TQ^T}_{y^T} L \underbrace{Qx}_{y} > 0 \Rightarrow$$

$$\Rightarrow \forall y : y^TLy>0 \Rightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}^T \begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = y_1^TAy_1 + y_2^TAy_2 > 0$$

حالا چون y بردار دلخواه غیر صفر است، برای $y_1\neq0 , y_2=0$ هم برقرار است ←

به ازای هر $y_1\neq0$ و $y_2=0$ هم باید برقرار باشد:

y^TMy

$$\forall y_1\neq0 : \; y_1^TAy_1>0 \longrightarrow A>0$$

حالا چون y بردار دلخواه غیر صفر است، برای $y_1=0 , y_2\neq0$ هم برقرار است ←

به ازای هر $y_2\neq0$, $y_1=0$ هم باید برقرار باشد،

y^TMy

$$\forall y_2\neq0 : \; y_2^TSy_2>0 \longrightarrow S>0$$

حالا برعکسش را اثبات می‌کنیم.

$$\left.\begin{aligned} A>0 &: \forall y_1 : y_1^TAy_1>0 \\ S>0 &: \forall y_2 : y_2^TSy_2>0 \end{aligned}\right\} \Rightarrow \forall y_1,y_2 : y_1^TAy_1 + y_2^TSy_2 > 0$$

$$\Rightarrow [y_1^T \; y_2^T]\begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} > 0 \Rightarrow \forall y, \; y^T\begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix} y > 0 \Rightarrow$$

$$\forall y : y^TLy>0 \longrightarrow \forall x : x^TQ^TLQx>0$$

$$y=Qx$$

$$\Rightarrow \forall x, \; x^TMx>0 \Rightarrow M>0$$

ب)

$$A>0 \longrightarrow \forall y_1 \neq 0: \; y_1^T A y_1 > 0$$
$$S \geq 0 \longrightarrow \forall y_2 \neq 0: \; y_2^T S y_2 \geq 0$$

$$\longrightarrow \begin{cases} \forall y_1, y_2, \; y_1 \neq 0, y_2 \neq 0: & \overbrace{y_1^T A y_1 + y_2^T S y_2}^{y^T M y} > 0 \\ \forall y_1, \; y_1 \neq 0, y_2 = 0: & y_1^T A y_1 > 0 \\ \forall y_2, \; y_2 \neq 0, y_1 = 0: & y_2^T S y_2 \geq 0 \end{cases}$$

$$\Rightarrow \forall y \neq 0: \; y^T M y \geq 0$$

y همان Qx است.

$$\Rightarrow \forall x \neq 0: \; x^T \overbrace{Q^T L Q}^{M} x \geq 0 \Leftrightarrow$$

اگر $A>0$ ، آنگاه اگر $S \geq 0 \Leftarrow M \geq 0$

حالا طرف دیگر را نشان می دهیم

$$A>0 \longrightarrow \forall y_1 \neq 0: \; y_1^T A y_1 > 0$$
$$M \geq 0 \longrightarrow \forall x \neq 0: \; x^T M x \geq 0$$
$$\longrightarrow \underbrace{x^T Q^T}_{y^T} L \underbrace{Qx}_{y} \geq 0 \rightarrow y^T L y \geq 0$$

$$\Rightarrow y^T \begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix} y \geq 0 \longrightarrow \begin{bmatrix} y_1^T & y_2^T \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq 0$$

حالا به ازای $y_1=0$، روی تمام مقادیر $y_2 \neq 0$ ها باید حاصل عبارت بزرگتر مساوی صفر شود.

$$\forall y_2 \neq 0: \; y_2^T S y_2 \geq 0 \longrightarrow S \geq 0$$

ج) چون ممکن است A وارون نداشته باشد (زیرا PSD است نه PD) پس $S = C - B^T A^+ B$ می‌شود.

حال با استفاده از لم این مورد را نشان می‌دهیم و سپس لم را اثبات می‌کنیم.

لم: برای هر ماتریس متقارن $P \in S^n$ و b، $f: \mathbb{R}^n \to \mathbb{R}$ ، $f(x) = x^T P x + 2 x^T b$

دارای کمینه است اگر و تنها اگر $P \geq 0$ و $(I - PP^+) b = 0$ داشته باشیم ← f دارای کمینه

با مقدار $-b^T P^+ b$ است.

حالا:

$$f(x, y) = x^T A x + 2 x^T B y + y^T C y$$

حالا $M \geq 0$ است اگر و تنها اگر به ازای هر (x, y) داشته باشیم $0 \leq f(x,y)$

حالا y را فیکس می‌کنیم، با توجه به لم $f(x,y)$ دارای کمینه است اگر و تنها اگر $A \geq 0$, $(I - AA^+) B = 0$

$$\Rightarrow \quad f(x^*, y) = -y^T B^T A^+ B y + y^T C y = y^T (C - B^T A^+ B) y$$

از طرف دیگر f باید به ازای هر y هم دارای کمینه باشد ← $C - B^T A^+ B \geq 0$ باید برقرار باشد و این مقدار کمینه برابر 0 است.

⇐ $M \geq 0$ است اگر و تنها اگر $A \geq 0$, $S \geq 0$, $(I - AA^+) B = 0$.

$$f(z) = (x+P^{\dagger}b)^T P(x+P^{\dagger}b) - 2x^T PP^{\dagger}b - b^T P^{\dagger}b + 2x^T b$$ اثبات لم

$$= (x+P^{\dagger}b)P(x+P^{\dagger}b) + 2x^T(I-PP^{\dagger})b - b^T P^{\dagger}b$$

if f has minimum the $P \succeq 0$, $(I-PP^{\dagger})b = 0$.

حالا فرض کنیم این رابطه برقرار نباشد یعنی $P \nsucceq 0$, $(I-PP^{\dagger})b \neq 0$ ← مقدار ویژه λ منفی داریم و $v \neq 0$ بردار ویژه متناظر آن ← $Pv = \lambda v$ ، حالا $z = \alpha v - P^{\dagger}b$ و α متغیر حقیقی و ناصفر باشد، حال داریم

$$f(z) = \alpha^2 v^T P v + 2(\alpha v - P^{\dagger}b)^T(I-PP^{\dagger})b - b^T P^{\dagger}b = \alpha^2 \lambda \|v\|_2^2 + 2\alpha v^T(I-PP^{\dagger})b$$

$$\underbrace{-2b^T P^{\dagger}(I-PP^{\dagger})b}_{(2b^T P^{\dagger} - 2b^T \cancel{P^{\dagger}PP^{\dagger}})b = 0} - b^T P^{\dagger}b = \alpha^2 \lambda \|v\|_2^2 + 2\alpha v^T(I-PP^{\dagger})b - b^T P^{\dagger}b$$

$P^{\dagger}$

حالا چون $\lambda < 0$ و $\|v\|_2 > 0$ ، این یک این عبارت یک سهمی کوژ با ضریب منفی ← کران پایین ندارد.

← حالا فرض کنیم که حداقل یکی از شرایط باید برقرار باشند.

$(I-PP^{\dagger})b = 0$ برقرار باشد ولی $P \succeq 0$ برقرار نباشد، پس P باید مقدار ویژه منفی داشته باشد.

عبارت بالا همچنان باقی می ماند پس همواره $P \succeq 0$ باید برقرار باشد.

حالا فرض کنیم $(I-PP^{\dagger})b \neq 0$ باشد، آنگاه داریم $b \neq 0$ و $\operatorname{Rank} P < n$ چون اگر $\operatorname{Rank} P = n$ باشد

$PP^{\dagger} = PP^{-1} = I$ ← $I - PP^{\dagger} = 0$ ← حالا طبق قضیه Real spectral داریم که یک basis orthonormal $v_1 — v_n$ که هستند و از مقادیر ویژه P تشکیل شده اند وجود دارد حالا $v_1 — v_n$ مربوط به بردارهای ویژه متناظر مقدار ویژه مثبت باشند و $v_{r+1} — v_n$ بردارهای ویژه متناظر مقادیر ویژه صفر باشند.

$$\operatorname{Range} P = \operatorname{span}\{v_1 —— v_r\}$$

$$\operatorname{Range} P = \operatorname{span}\{v_{r+1} —— v_n\}$$

حالا چون $(I-PP^{\dagger})b \neq 0$ است ← b در Null space P قرار ندارد چون این ماتریس بردارها را روی $N(P)$ تصویر می کند.

→ for all $v \in N(P)$: $v^T(I-PP^{\dagger})b = v^T b$

حالا داریم:

$x = \alpha b - P^{+} b$

$$f(x) = 2\alpha \|b\|_2^2 - b^T P^{+} b$$

حالا اگه $\alpha \to -\infty$ مقدار f هم به سمت $-\infty$

می‌رود یعنی کران پایین نداریم ← پس برای داشتن کران پایین این دو شرط لازم است.

حالا اگه هر دو شرط برقرار باشند، داریم:

$$f(z) = (x + P^{+} b)^T P (x + P^{+} b) - b^T P^{+} b$$

چون که $-b^T P^{+} b$ یک عبارت ثابت است و $P \succeq 0$ ← مقدار کمینه‌اش $f(z)$

برابر است با $-b^T P^{+} b$ زیرا $(x + P^{+} b)^T P (x + P^{+} b) \geq 0$.

~~[illegible]~~ f کران پایین دارد اگر و تنها اگر $(I - PP^{+}) b = 0$, $P \succeq 0$

به طور تعمیم یافته در عبارت خودمان داشتیم که به ازای هر y، $(I - AA^{+}) B y$ صفر باشد

$$\longrightarrow (I - AA^{+}) B = 0 \longrightarrow B^T (I - AA^{+}) = 0$$

د) چون M متقارن است ← A و C هم متقارن اند. حالا داریم که:

$$M>0 \longrightarrow A>0,\ C>0,\ S>0$$

$$\forall y_2\neq 0,\ y_1=0:\quad \begin{bmatrix}y_1\\y_2\end{bmatrix}^T [M] \begin{bmatrix}y_1\\y_2\end{bmatrix} = y_2^T C y_2 > 0 \to C>0$$

حالا داریم که: $S=C-B^TA^{-1}B$

$$\Rightarrow S = C^{1/2}\left(I - C^{-1/2}B^TA^{-1}BC^{-1/2}\right)C^{1/2} > 0$$

$$\Rightarrow \forall y:\ \underbrace{y^TC^{1/2}}_{x^T}\left(I-C^{-1/2}B^TA^{-1}BC^{-1/2}\right)\underbrace{C^{1/2}y}_{x} > 0 \Longrightarrow$$

چونکه $C^{-1/2}$ یک ماتریس فول رنک است، ~~تمام~~ $C^{1/2}y$ تمام بردارها را R^{n-k} را تولید می‌کند.

$$\forall x:\ x^T\left(I-(A^{-1/2}BC^{-1/2})^T(A^{-1/2}BC^{-1/2})\right)x>0$$

$$\Rightarrow I - \underbrace{(A^{-1/2}BC^{-1/2})^T}_{V\Sigma^TU^T}\underbrace{(A^{-1/2}BC^{-1/2})}_{U\Sigma V^T} > 0$$

$$\Rightarrow V(I-\Sigma^T\Sigma)V^T>0 \longrightarrow V \begin{bmatrix}1-\sigma_1^2 & & \\ & \ddots & \\ & & 1-\sigma_k^2\end{bmatrix} V^T > 0$$

حالا چون V هم فول رنک است اگر یک کدام از $1-\sigma_i^2$ ها منفی باشد، آنگاه بردار ویژه متناظر آن را برداریم

$1-\sigma_i^2\le 0$ فرض کنیم آنگاه با برداشتن بردار v_i و ضرب آن حاصل عبارت منفی می‌شود.

ستون i ام V

$$v_i^TV = [0 \dots 0\ \ v_i^Tv_i\ \ 0 \dots 0]$$

از دو طرف داریم:

$$\Rightarrow v_i^TV(I-\Sigma^T\Sigma)V v_i = (v_i^Tv_i)^2(1-\sigma_i^2) \le 0 \quad ※$$

پس به ازای هر σ_i، $1-\sigma_i^2>0$ ← $\sigma_i^2<1$ ← تمامی مقادیر تکین کوچکتر از ۱ هستند.

حالا طرف دیگر را نشان می‌دهیم: اگر تمامی مقادیر ~~سینگولار~~ $A^{-1/2}BC^{-1/2}$ کوچکتر از 1 باشند و A و C مثبت معین باشند، آنگاه $M>0$ است:

$$A>0,\ C>0\ ,\ \sigma_{max}(\underbrace{A^{-1/2}BC^{-1/2}}_{U\Sigma V^T})<1$$

$$\underbrace{(A^{-1/2}BC^{-1/2})^T(A^{-1/2}BC^{-1/2})}_{V\Sigma^T\Sigma V^T} = C^{-1/2}B^TA^{-1}BC^{-1/2}$$

$$\longrightarrow\ I-C^{-1/2}B^TA^{-1}BC^{-1/2} = V(\underbrace{I-\Sigma^T\Sigma}) V^T$$

$$V\begin{bmatrix}1-\sigma_1^2 & & \\ & \ddots & \\ & & 1-\sigma_k^2\end{bmatrix}V^T$$

تمامی مقادیر روی قطر ماتریس قطری مان مثبت اند ← یک ماتریس $\overset{\text{ماتریس } H}{\widetilde{V\Lambda V^T}}$ روبه رو هستیم که مقادیر ویژه H/

تمامی مقادیر ویژه H/ مثبت اند ←

H ماتریسی مثبت و معین است ←

$$\forall y:\ y^T(I-C^{-1/2}B^TA^{-1}BC^{-1/2})y>0$$

حالا چون $C^{1/2}$ فول رنک است، می‌توانیم قرار دهیم: $Y=C^{1/2}X$ و تمامی Xها را اگر در نظر بگیریم، $C^{1/2}X$ تمام Yها را تولید می‌کند ⇐

$$\forall x,\ x^TC^{1/2}(I-C^{-1/2}B^TA^{-1}BC^{-1/2})C^{1/2}x \Rightarrow$$

$$\forall x,\ x^T(C-B^TA^{-1}B)x>0 \longrightarrow \underbrace{C-B^TA^{-1}B}_{S}>0 \Longrightarrow S>0$$

حالا چون هم $A>0$ و هم $S>0$ طبق الف نتیجه می‌دهد که $M>0$

۵- الف) داریم که $AX = XB$ ، حالا بردار ویژه های B را v_i و بردار ویژه های A^T را u_i در نظر بگیرید.
مقادیر ویژه ی B را λ_i و مقادیر ویژه ی A^T را μ_i

$$AXv_i = XBv_i = X\lambda_i v_i = \lambda_i X v_i \Rightarrow A(Xv_i) = \lambda_i (Xv_i) \Rightarrow$$

Xv_i بردار ویژه ی A است با مقدار ویژه λ_i $\therefore$ طبق صورت سوال A و B مقدار ویژه ی برابر ندارند $\Rightarrow$
$Xv_i = 0$ است به ازای هر v_i ای.

$$AX = XB \longrightarrow X^T A^T = B^T X^T \longrightarrow X^T \underbrace{A^T u_i}_{\mu_i u_i} = B^T X^T u_i \longrightarrow$$

$$B^T (X^T u_i) = \mu_i (X^T u_i) \longrightarrow$$

$X^T u_i$ بردار ویژه ی B^T است با مقدار ویژه ی μ_i ، می دانیم که ~~...~~ مقادیر ویژه یک ماتریس با معکوس برابر است
$\leftarrow$ μ_i ها مقادیر ویژه ی A و B هستند $\therefore$ طبق صورت سوال A و B مقادیر ویژه برابر ندارند $\Rightarrow$ $X^T u_i = 0$ است به ازای هر u_i ای.

حالا داریم که $\text{for all } u_i : X^T u_i = 0$ ، $\text{for all } v_i : Xv_i = 0$.

از طرفی ~~...~~ اگر ماتریسی فول رنک نباشد، مقدار ویژه ی صفر دارد و اگر ~~...~~ هردو A و B فول رنک نباشند،
یعنی هردو مقدار ویژه ی صفر دارند $\leftarrow$ از بین A و B $\leftarrow$ حداقل یکی فول رنک است.

$\leftarrow$ ماتریسی که فول رنک است، بردار های ویژه اش کل فضا را پوشش می دهند. (به اندازه ی بعد فضا بردار ویژه مستقل خطی داریم)
$\Leftarrow$ ~~...~~ بدون کم شدن از کلیت مسئله فرض کنیم B فول رنک است $\Leftarrow$
v_i ها را در ستون های یک ماتریس می چینیم به اسم V و سپس داریم:

هر سطر از X در تمام v_i ضرب داخلی شود، صفر خروجی
$XV = 0 \longrightarrow$
می دهد، درصورتی که v_i ها مستقل خطی اند و کل فضا را پوشش می دهند $\longrightarrow$ حال که $XV = 0$ است
پس تک تک سطرهای X صفر هستند $\longleftarrow$ X ماتریسی با درایه های تماماً صفر است.

Scanned with CamScanner

به طور مشابه اگر A را غیر فول رنک هم بگیریم و B فول رنک باشد، همین اتفاق می‌افتد.
در کل بردار های ویژه مستقل خطی اند و کل فضایشان را می‌سازند؛ مثلاً کل فضای n بعدی را پوشش می‌دهند
← در هر سطر x ضرب داخلی اش در تمام بردار های ویژه صفر است ← $x=0$

ب) A متقارن است ← n بردار ویژه‌ی متعامد دارد که مستقل خطی هم هستند و فضای n بعدی مان را هم span می‌کنند، این بردار ها را $\{q_1 — q_n\}$ در نظر می‌گیریم، خواهیم داشت:

$$\langle q_i, q_j\rangle = 0$$

هر q_i هم مرتبط با λ_i است. (λ_i مقدار ویژه‌ی i ام A است.)

حالا می‌خواهیم نشان دهیم:

$$\lambda_k = \max_{\substack{S\subseteq \mathbb{R}^n \\ \dim(S)=k}} \ \min_{\substack{x\in S \\ x\neq 0}} \frac{x^TAx}{x^Tx}$$

حالا عملیات max به ازای بعد k باید k تا از پایه ها را بردارد. پایه‌های ما هم $q_1 — q_n$ هستند، هر پایه‌ی دیگر هم در نظر بگیریم x عنصر از آن می‌تواند k تا از q_i ها را تولید کند. حالا برای سادگی و بدون کم شدن از کلیت مسئله، عملیات max باید k تا از q_i ها را بردارد، در این صورت اگر $\lambda_1 — \lambda_k$ را بردارد، عملیات min که کوچکترین آنها را برمی‌دارد. با برداشتن بردار ویژه متناظر λ_k مقدار λ_k را تولید می‌کند. اگر ضریبی از سایر بردار های ویژه وجود داشته باشد، مقدارمان بیشتر از λ_k می‌شود ← min بین حالت ها λ_k است.)

$$\frac{q_k^TAq_k+\eta^2 q_j^TAq_j}{\underbrace{(q_k+\eta q_j)^T(q_k+\eta q_j)}_{q_k^Tq_k+\eta^2 q_j^Tq_j}}$$

$$\frac{q_k^TAq_k}{q_k^Tq_k}=\lambda_k \qquad \frac{q_j^TAq_j}{q_j^Tq_j}=\lambda_j \longrightarrow \lambda_k \le \frac{q_k^TAq_k+\eta^2 q_j^TAq_j}{q_k^Tq_k+\eta^2 q_j^Tq_j} \le \lambda_j$$

بزرگتر از λ_k

در هر صورت λ_k، min است.

اگر هم max بردار ویژه‌های متناظر با مقادیر ویژه کوچکتر هم داشته باشد نسبت به $\lambda_1 — \lambda_k$، آنگاه min با انتخاب آن به مقداری کمتر از λ_k می‌رسد. پس بهترین انتخاب برای max برداشتن بردارهای $q_1 — q_k$ است.

حالا طرف دیگر را نشان می‌دهیم:

$$\lambda_k = \min_{\substack{C\subseteq \mathbb{R}^n \\ \dim(C)=n-k+1}} \ \max_{\substack{x\in C \\ x\neq 0}} \frac{x^TAx}{x^Tx}$$

مشابه قبلی، عملیات min باید $n-k+1$ از پایه‌های q_i را بردارد، حالا اگر q_k — q_n را بردارد، عملیات max، q_k را برمی‌دارد که بزرگترین مقدار ویژه بین λ_n — λ_k متناظر است حالا اگر min بردارهای غیراز q_k — q_n را بیاورد یکی از آنها بردارد، max آن بردار جدید را چون مقدار ویژه متناظرش بزرگتر است، انتخاب می‌کند و به مقدار بزرگتری می‌رسیم پس بهترین انتخاب برای min مجموعه q_n — q_{n-k+1} است و

⇐ حاصل عبارت λ_k است.

ادامه‌ی ۵ ب):

$$\min_{\substack{C\subseteq R^n\\ \dim(C)=n-k+1}} \max_{\substack{x\in C\\ x\neq 0}} \frac{\|Ax\|_2}{\|x\|_2} = \sqrt{\min_{\substack{C\subseteq R^n\\ \dim(C)=n-k+1}} \max_{\substack{x\in C\\ x\neq 0}} \frac{\|Ax\|_2^2}{\|x\|_2^2}} \rightarrow \frac{x^TA^TAx}{x^Tx}$$

مشابه حالت قبل این عبارت عبارت است از $\lambda_k(A^TA)$

$$\text{حاصل عبارت} = \sqrt{\lambda_k(A^TA)} = \sigma_k(A)$$

ج) کمینه کردن نرم معادل کمینه کردن توان ۲ نرم است ⇐

⇐ چون نامنفی است.

$$\min_{\alpha, Q} \|A-\alpha Q\| \equiv \min \|A-\alpha Q\|^2$$

$$= \min_{\alpha, Q} Tr\left((A-\alpha Q)^T(A-\alpha Q)\right) = \min_{\alpha, Q} \overbrace{Tr(A^TA)}^{\text{ثابت}} - 2\alpha\, Tr(A^TQ) + \alpha^2 Tr(\cancel{Q^TQ})$$

$$\equiv \min_{\alpha, Q} -2\alpha\, Tr(A^TQ) + n\alpha^2 \qquad A = U\Sigma V^T$$

حال باید $Tr(A^TQ)$ را بیشینه کنیم و سپس براساس α مقدار کلی را کمینه کنیم.

با بیشینه کردن $Tr(A^TQ)$ به ازای یک α خاص مقدار تابع کمینه می‌شود

$$Tr(A^TQ) = Tr(V\Sigma U^TQ) = Tr(\Sigma U^TQV)$$

حاصلضرب دوتا ماتریس یکانی یک ماتریس یکانی است زیرا:

$$(QV)^TQV = V^T\underbrace{Q^TQ}_{I}V = V^TV = I \quad \checkmark \rightarrow U^TQV$$

$$(QV)(QV)^T = Q\underbrace{VV^T}_{I}Q^T = QQ^T = I$$

خود یک ماتریس یکانی است.

پس هر درایه اش کوچکتر مساوی یک است زیرا مجموع مجذورات درایه‌های هر سطر 1 است

$$\Rightarrow Tr(\Sigma \underbrace{U^TQV}_{M}) = \sum_{i=1}^{n} \sigma_i M_{ii} \le \sum_{i=1}^{n} \sigma_i$$

$$Q = UV^T$$

حال اگر $M=I$ باشد این نامساوی به تساوی تبدیل می‌شود و $Tr(A^TQ)$ بیشینه می‌شود ⇐

با این انتخاب برای Q داریم که $Tr(A^TQ) = \sum_{i=1}^{n} \sigma_i$ ، حالا باید یک مقدار بهینه برای α پیدا کنیم،

$$\min_{\alpha} \; n\alpha^2 - 2\alpha\left(\sum_{i=1}^{n}\sigma_i\right)$$

$$\frac{\partial}{\partial\alpha}\left(n\alpha^2 - 2\alpha\left(\sum_{i=1}^{n}\sigma_i\right)\right) = 0 \rightarrow \alpha = \frac{1}{n}\sum_{i=1}^{n}\sigma_i$$

$\Longrightarrow$

مقدار بهینه $\|A-\alpha Q\|$ عبارت است از:

$$\min_{\alpha,Q} \sqrt{Tr(A^TA) - 2\alpha\, Tr(A^TQ) + n\alpha^2} = \sqrt{\|A\|^2 - \frac{2}{n}\left(\sum\sigma_i\right)^2 + \frac{1}{n}\left(\sum\sigma_i\right)^2}$$

$$= \boxed{\sqrt{\|A\|^2 - \frac{1}{n}\left(\sum\sigma_i\right)^2}}$$

۶

$$\lambda_{max}\left(\frac{A+A^H}{2}\right) = \max_{\|x\|_2=1} x^H\left(\frac{A+A^H}{2}\right)x$$

$$x^H\left(\frac{A+A^H}{2}\right)x = \frac{x^HAx}{2} + \frac{x^HA^Hx}{2} = \mathrm{Real}(x^HAx)$$

$$\lambda_{max}\left(\frac{A+A^H}{2}\right) = \max_{\|x\|_2=1} \mathrm{Real}(x^HAx)$$

$$|x^HAx| \leq \|Ax\| \overbrace{\|x\|}^{1} \leq \sigma_{max}(A)$$

$$\sigma_{max}(A) = \max_{\|x\|_2=1} \|Ax\|$$

$$\mathrm{Real}(x^HAx) \leq |x^HAx| \leq \sigma_{max}(A)$$

$$\mathrm{Real}(x^HAx) \leq \sqrt{\mathrm{Real}(x^HAx)^2 + \mathrm{c}(x^HAx)^2} \rightarrow \lambda_{max}\left(\frac{A+A^H}{2}\right) \leq \sigma_{max}(A)$$

۶- ابتدا قضیه Gershgorin circle را اثبات می‌کنیم و سپس از آن استفاده می‌کنیم برای حل سوال:

ماتریس A را در نظر بگیرید با بردار ویژه X ← دلخواه $Ax=\lambda x$. حالا x_i را المنتی در X که بزرگترین قدرمطلق را دارد در نظر بگیرید $x_i = \underset{x_j\in x}{\arg\max}|x_j|$. می‌دانیم که داریم:

$$\sum_j a_{ij}x_j = \lambda x_i \longrightarrow (\lambda-1)x_i = \sum_{j\neq i} a_{ij}x_j \rightarrow$$

$$\xrightarrow[\text{مثلثی داریم}]{\text{طبق نامساوی}} \lambda-1 = \sum_{j\neq i} a_{ij}\frac{x_j}{x_i} \longrightarrow |\lambda-1| = \left|\sum_{j\neq i} a_{ij}\frac{x_j}{x_i}\right| \xrightarrow{\text{نامساوی مثلثی}}$$

چون ماکزیمم قدرمطلق را برای x_i برداشته بودیم → $\frac{|x_j|}{|x_i|}\le 1$

$$|\lambda-1| \le \sum_{j\neq i}|a_{ij}|\frac{|x_j|}{|x_i|} \le \sum_{j\neq i}|a_{ij}| \le 1$$

چون A هرمیتی است تمام مقادیر ویژه حقیقی اند.

$$\Rightarrow |\lambda-1|\le 1 \longrightarrow -1\le\lambda-1\le 1 \rightarrow 0\le\lambda\le 2 \longrightarrow$$

تمامی مقادیر ویژه بزرگتر مساوی صفر هستند ← A مثبت نیمه معین است. ← الف)

همچنین بدست آمد که $\lambda\le 2$ ← ب)

حال به بخش ج) می‌پردازیم یعنی باید نشان دهیم $0\le\det(A)\le 1$ ، می‌دانیم که:

$$\boxed{\det(A)=\prod_{i=1}^{n}\lambda_i(A)} \quad , \quad \boxed{tr(A)=\sum_{i=1}^{n}\lambda_i(A) = \sum_i a_{ii} = n}$$

$$\Rightarrow \sum_{i=1}^{n}\lambda_i(A) = n \quad , \quad \max\prod_{i=1}^{n}\lambda_i(A) = ?$$

طبق نامساوی هندسی حسابی می‌دانیم که برای $a_1 \dots a_n \in [0,+\infty)$ داریم که

$\frac{\sum a_i}{n} \ge \prod a_i$

Scanned with CamScanner

همچنین چون A هرمیتی است تمام مقادیر ویژه‌ی آن حقیقی است و از طرفی تمامی آنها بزرگتر مساوی صفراند ←

$$\det(A) = \prod_{i=1}^{n} \lambda_i(A) \leq \underbrace{\sum_{i=1}^{n} \lambda_i(A)/n}_{tr(A)/n = 1} \longrightarrow \det(A) \leq 1$$

از طرفی چون تمام λ_iها بزرگتر مساوی صفراند ← حاصلضرب آنها یعنی $\det(A)$ هم بزرگتر مساوی صفر است.

$$\Rightarrow \quad 0 \leq \det(A) \leq 1$$

اثبات نامساوی حسابی هندسی،

با استقرای توان،

پایه: $P=2$: $\frac{a+b}{2} \geq \sqrt{ab}$: $(\sqrt{a}-\sqrt{b})^2 \geq 0 \to a+b-2\sqrt{ab} \geq 0 \longrightarrow \frac{a+b}{2} \geq \sqrt{ab}$ ✓

حالا به ازای هر P_n ، P_{2n} را نتیجه می‌گیریم، جملات را دوتا دوتا در نظر می‌گیریم و نامساوی مثلث می‌زنیم،

$$\underline{a_1+a_2}+ \text{———} + \underline{a_{2n-1}+a_{2n}} \geq 2\sqrt{a_1a_2} + \cdots + 2\sqrt{a_{2n-1}\,a_{2n}} \xrightarrow[n\text{تایی برقرار است }(P_n)]{\text{حالا چون نامساوی مثلث}}$$

$$a_1+a_2+ \text{———} + a_{2n-1}+a_{2n} \geq 2\left(\sqrt{a_1a_2} + \text{———} + \sqrt{a_{2n-1}a_{2n}}\right) \geq 2n\sqrt[n]{\sqrt{a_1a_2}\cdots\sqrt{a_{2n-1}a_{2n}}}$$

$$\Rightarrow \frac{a_1+\cdots+a_{2n}}{2n} \geq \sqrt[n]{\sqrt{a_1a_2}\cdots\sqrt{a_{2n-1}a_{2n}}} = \sqrt[2n]{a_1\cdots a_{2n}} \quad ✓$$

حالا به ازای هر P_n ، P_{n-1} را نتیجه می‌گیریم:

$$a_1, a_2, \text{———}, a_{n-1}, g : \quad \frac{a_1 + \text{———} + a_{n-1} + g}{n} \geq \sqrt[n]{a_1a_2 \cdots a_{n-1}g} \quad \text{for all } g > 0$$

$$\longrightarrow \text{choose } g = \sqrt[n-1]{a_1 \cdots a_n} \Rightarrow a_1 + \cdots + a_{n-1} + g \geq n\underbrace{\sqrt[n]{a_1 \cdots a_{n-1}\sqrt[n-1]{a_1 \cdots a_{n-1}}}}_{\sqrt[n]{(a_1\cdots a_{n-1})^{\frac{n}{n-1}}} = \sqrt[n-1]{a_1 \cdots a_{n-1}}} = ng$$

⇐ داریم که:

$$a_1+\dots+a_{n-1}+g \ge ng \longrightarrow a_1+\dots+a_{n-1} \ge (n-1)g$$

$$\Rightarrow g \le \frac{a_1+\dots+a_{n-1}}{n-1} \Rightarrow \sqrt[n-1]{a_1\cdots a_{n-1}} \le \frac{a_1+\dots+a_{n-1}}{n-1}$$

بدین ترتیب به ازای $P=2$ که برقرار است ← به ازای $(P_n, n=2^k)$ ها هم برقرار می‌شود به ازای هر $k \ge 1$، همچنین

به ازای هر P_n که $n=2^k$ برقرار است $n-1$، سپس $n-2$، ... اثبات می‌شود ⇐ برای کل

n ها برقرار است یعنی همواره این نامساوی برقرار است ■

(در این اثبات از استقرا قهقرایی استفاده کردیم.)

۷- ابتدا بخش ب را حل می‌کنیم:

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T \longrightarrow A - A_k = \sum_{i=k+1}^{r} \sigma_i u_i v_i^T$$

نرم 2 بزرگترین مقدار تکین را می‌داد. $\longrightarrow \|A - A_k\|_2 = \sigma_{k+1}$

همان طور که در سوال ۸ اثبات می‌کنیم. حالا بین بردارهای $\{v_1 — v_{k+1}\}$ که بردارهای ویژه A هستند و به ترتیب متناظر با بزرگترین مقدار تکین، دومین بزرگترین مقدار تکین و ... $k+1$ امین مقدار تکین بزرگ هستند.

از طرفی $Rank(X) = k \longrightarrow Nullity(X) = n-k$

$$\longleftarrow Rank(\{v_1 — v_{k+1}\}) + Nullity(X) = n+1 \longrightarrow \exists w \in Null(X) \cap (\{v_1 — v_{k+1}\})$$

w را طوری انتخاب می‌کنیم که $\|w\|_2 = 1$ (اسکیل می‌کنیم)

حالا داریم که

$$\|A - X\|_2 = \sup_{v \neq 0} \|(A-X)v\|_2 \geq \|(A-X)w\|_2 = \|Aw\|_2 = w^T A^T A w$$

$$\Longrightarrow \|A-X\|_2 \geq w^T V \Sigma^2 V^T w \xrightarrow{\sum_{i=1}^{k+1} w_i v_i} = \sum_{i=1}^{k+1} w_i^2 \sigma_i^2 \geq \sigma_{k+1}^2 \underbrace{\sum_{i=1}^{k} w_i^2}_{\|A-A_k\|_2}$$

$$\Longrightarrow \overbrace{\|A - X\|_2 \geq \|A - A_k\|_2}^{\text{* برقراری نامساوی}} \Longrightarrow \boxed{\sigma_{k+1}(A) \leq \sigma_1(A - X)}$$

حالا به بخش اول می‌پردازیم، فرض کنیم $A = B + C$، داریم که:

$$\sigma_i(B) + \sigma_j(C) = \sigma_1(B - B_{i-1}) + \sigma_1(C - C_{j-1}) \geq \sigma_1(B + C - B_{i-1} - C_{j-1})$$

$$* \; B_{i-1} + C_{j-1} = X \longrightarrow Rank(X) \leq i+j-2$$

طبق بخش ب

$$\longrightarrow \sigma_1(A - B) \geq \sigma_{i+j-2}(A) \rightsquigarrow \|A - \cancel{B}\|_F^2 = \sum_{i=1}^{n} \sigma_i(A-X)^2 \geq \sum \sigma_{i+k}(A)^2 = \sum_{i=k+1}^{n} \sigma_i(A)^2$$

$$\longrightarrow \|A - A_k\|_F^2$$

۸- الف)

$$\|A\|_2 = \sup_u \frac{\|Au\|_2}{\|u\|_2} = \sup \|Au\|_2 \quad s.t. \quad \|u\|_2 = 1$$

می‌دانیم که ماکزیمم کردن $\|Au\|_2$ با ماکزیمم کردن $\|Au\|_2^2$. حالا داریم:

ماتریس

$$\sup_u u^T A^T A u = \sup_u u^T V \Sigma \underbrace{U^T U}_{I} \Sigma V^T u = \sup_u u^T V \Sigma^2 V^T u = \sup_u \|\Sigma V^T u\|_2$$

$$\|V^T u\|_2^2 = u^T \underbrace{V V^T}_{I} u = u^T u = \|u\|_2^2 = 1 \longrightarrow$$ $V^T u$ یک بردار با نرم یک است.

این بردار را s می‌نامیم، داریم:

$$\Rightarrow \sup_u \|\Sigma V^T u\|_2 = \sup_s \|\Sigma s\|_2 \rightsquigarrow$$

معادل است با $\sup_s \|\Sigma s\|_2^2 \longleftarrow$

$$\sup_s s^T \Sigma^2 s = \sum_{i=1}^{n} \sigma_i^2 s_i^2 \le \sigma_{max}^2 \underbrace{\sum_{i=1}^{n} s_i^2}_{1} = \sigma_{max}^2$$

$$\Rightarrow \|\Sigma s\|_2^2 \le \sigma_{max}^2 \longrightarrow \|\Sigma s\|_2 \le \sigma_{max}$$

$$\Rightarrow \sup_u \frac{\|Au\|_2}{\|u\|_2} \le \sigma_{max}$$

حالا فرض کنیم در تجزیه مان σ_i ها به ترتیب از بزرگ به کوچک در تجزیه مان قرار داده ایم $\longleftarrow \sigma_1 = \sigma_{max}$

حالا بردار ویژه متناظر σ_1، V_1 است، حالا اگر s برابر باشد با $\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Leftarrow s^T \Sigma s = \sigma_{max}$

حالا می‌دانیم که $V^T u = s \longleftarrow \underbrace{V V^T}_{I} u = V s \longleftarrow u = V s \longleftarrow$

$$u = [V_1 | \cdots | V_n] \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = V_1 \longrightarrow$$ اگر $u = V_1$ باشد، آنگاه $\underline{\|Au\|_2 = \sigma_{max}}$

← برداری داریم که کران بالا را تولید کند ← $\|A\|_2 = \sigma_{max}(A)$

ب)

$A = U\Sigma V^T \rightarrow A^{-1} = V\Sigma^{-1}U^T, \quad \Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n} \end{bmatrix}$

$\|A\|_2 = \sigma_{max}(A)$

$\|A^{-1}\|_2 = \sigma_{max}(A^{-1}) \quad \Rightarrow \quad \sigma_{max}(A^{-1}) = \frac{1}{\sigma_{min}(A)}$

بزرگترین مقدار تکین A^{-1}، برابر است با 1 تقسیم بر کوچکترین مقدار تکین A.
چرا که کوچکترین مقدار تکین A، بزرگترین معکوس را دارد.

$$\Rightarrow K_A = \|A\|_2 \|A^{-1}\|_2 = \sigma_{max}(A) \frac{1}{\sigma_{min}(A)} = \frac{\sigma_{max}(A)}{\sigma_{min}(A)} \geq 1$$

ج)

$$Ax = b \rightarrow \|Ax\|_2 = \|b\|_2 \longrightarrow \|x\|_2 \underbrace{\Big\| A \overset{\text{بردار یکه}}{\frac{x}{\|x\|_2}} \Big\|_2}_{} = \|b\|_2$$

طبق تعریف بالاتر این عبارت کوچکتر مساوی نرم 2 ماتریس A است.

$$\Rightarrow \|x\|_2 \|A\|_2 \geq \|b\|_2 \rightarrow \frac{1}{\|x\|_2} \leq \frac{\|A\|_2}{\|b\|_2}$$

$$Ax - A\tilde{x} = b - \tilde{b} = e \rightarrow A(x - \tilde{x}) = e \rightarrow x - \tilde{x} = A^{-1}e \longrightarrow \|x - \tilde{x}\|_2 = \|A^{-1}e\|_2$$

← طبق بالا مشابه کاری که کردیم داریم:

$$\|A^{-1}e\|_2 = \|e\|_2 \Big\| A^{-1} \underset{\text{بردار یکه}}{\frac{e}{\|e\|_2}} \Big\|_2 \leq \|A^{-1}\|_2 \|e\|_2$$

$$\Rightarrow \|x - \tilde{x}\|_2 \leq \|A^{-1}\|_2 \|e\|_2$$

$$\frac{1}{\|x\|_2} \leq \|A\|_2 / \|b\|_2 \longrightarrow \frac{\|x - \tilde{x}\|_2}{\|x\|_2} \leq \frac{\overbrace{\|A\|_2 \|A^{-1}\|_2}^{K_A} \|e\|_2}{\|b\|_2} = K_A \frac{\|e\|_2}{\|b\|_2}$$

طرف راست نامساوی اثبات شد. حالا طرف دیگر را نشان می دهیم:

$$Ax - A\tilde{x} = b - \tilde{b} = e \longrightarrow \|A(x-\tilde{x})\|_2 = \|e\|_2 \longrightarrow \|x-\tilde{x}\|_2 \left\|A \underbrace{\frac{x-\tilde{x}}{\|x-\tilde{x}\|_2}}_{\text{بردار یکه}}\right\|_2 = \|e\|_2$$

$$\Rightarrow \|x-\tilde{x}\|_2 \, \|A\|_2 \geq \|e\|_2 \longrightarrow \|x-\tilde{x}\|_2 \geq \frac{\|e\|_2}{\|A\|_2}$$

$$Ax = b \longrightarrow x = A^{-1}b \longrightarrow \|x\|_2 = \|A^{-1}b\|_2 = \|b\|_2 \left\|A^{-1} \underbrace{\frac{b}{\|b\|}}_{\text{بردار یکه}}\right\|_2 \leq \|b\|_2 \|A^{-1}\|_2$$

$$\longrightarrow \|x\|_2 \leq \|b\|_2 \|A^{-1}\|_2 \longrightarrow \frac{1}{\|x\|_2} \geq \frac{1}{\|b\| \, \|A^{-1}\|_2}$$

$$\Rightarrow \frac{\|x-\tilde{x}\|_2}{\|x\|_2} \geq \frac{\|e\|_2}{\|b\|_2 \underbrace{\|A\|_2 \|A^{-1}\|_2}_{K_A}} = K_A^{-1} \frac{\|e\|_2}{\|b\|_2}$$

$$\Rightarrow \boxed{\frac{\|x-\tilde{x}\|}{\|x\|_2} \geq K_A^{-1} \frac{\|e\|_2}{\|b\|_2}}$$

$$\Rightarrow K_A^{-1} \frac{\|e\|_2}{\|b\|_2} \leq \frac{\|x-\tilde{x}\|_2}{\|x\|_2} \leq K_A \frac{\|e\|_2}{\|b\|_2}$$

$$\left.\begin{array}{l} A\tilde{x} = (\tilde{A}+E)\tilde{x} = \tilde{A}\tilde{x}+E\tilde{x} = \tilde{b}+E\tilde{x} \\ Ax = b \end{array}\right\} \Rightarrow A(\tilde{x}-x) = (\tilde{b}-b) + E\tilde{x}$$

$$\Rightarrow \tilde{x}-x = A^{-1}\left((\tilde{b}-b)+E\tilde{x}\right) \Longrightarrow \|\tilde{x}-x\|_2 \le \|A^{-1}\|_2\left(\underbrace{\|\tilde{b}-b\|_2 + \|E\|_2\,\|\tilde{x}\|_2}\right)$$

نامساوی مثلثی

$$\|\tilde{x}\|_2 \le \|\tilde{x}-x\|_2 + \|x\|_2$$

$$\Rightarrow \|x-\tilde{x}\|_2 \le \|A^{-1}\|_2\left(\|\tilde{b}-b\|_2 + \|E\|_2\,\|x-\tilde{x}\|_2 + \|E\|_2\,\|x\|_2\right)$$

$$\Rightarrow \left(1-\|A^{-1}\|_2\,\|E\|_2\right)\|x-\tilde{x}\|_2 \le \|A^{-1}\|_2\left(\|e\|_2 + \|E\|_2\,\|x\|_2\right)$$

$$\Rightarrow \frac{\|x-\tilde{x}\|_2}{\|x\|_2} \le \frac{\|A^{-1}\|_2}{1-\|A^{-1}\|_2\,\|E\|_2}\left(\frac{\|e\|_2}{\|x\|_2}+\|E\|_2\right) \xrightarrow[\frac{1}{\|x\|_2} \le \frac{\|A\|_2}{\|b\|_2}]{\|b\|_2 \le \|A\|_2\,\|x\|_2}$$

$$\frac{\|x-\tilde{x}\|_2}{\|x\|_2} \le \frac{\|A^{-1}\|_2}{1-k_A\frac{\|E\|_2}{\|A\|_2}}\left(\frac{\|e\|_2\,\|A\|_2}{\|b\|_2}+\|E\|_2\right) \Rightarrow$$

$$\frac{\|x-\tilde{x}\|_2}{\|x\|_2} \le \frac{\frac{\|e\|_2}{\|b\|_2}k_A + \overbrace{\|A^{-1}\|_2\,\|A\|_2}^{k_A}\|E\|_2/\|A\|_2}{1-k_A\frac{\|E\|_2}{\|A\|_2}}$$

$$\Rightarrow \frac{\|x-\tilde{x}\|_2}{\|x\|_2} \le \frac{k_A\left(\frac{\|e\|_2}{\|b\|_2}+\frac{\|E\|_2}{\|A\|_2}\right)}{1-k_A\left(\frac{\|E\|_2}{\|A\|_2}\right)}$$

5)

$$r = A\hat{x} - \hat{\lambda}\hat{x}, \quad A = S\Lambda S^{-1}$$

$$\Rightarrow r = S\Lambda S^{-1}\hat{x} - \hat{\lambda}\hat{x} = S\Lambda S^{-1}\hat{x} - \hat{\lambda} S S^{-1}\hat{x} = S(\Lambda - \hat{\lambda}) S^{-1}\hat{x} = r$$

$$\Rightarrow S \begin{bmatrix} \lambda_1 - \hat{\lambda} & & O \\ & \ddots & \\ O & & \lambda_n - \hat{\lambda} \end{bmatrix} S^{-1}\hat{x} = r$$

صفر

صفر

حالا اگر وجود داشته باشد λ_i ای که $\lambda_i - \hat{\lambda} = 0$، آنگاه همان λ_i را انتخاب می‌کنیم، خواهیم داشت که

$$|\lambda_i - \hat{\lambda}| = 0 \qquad \frac{\|r\|_2}{\|x\|_2} k_s \geq 0$$

$$\Rightarrow |\lambda_i - \hat{\lambda}| \leq \frac{\|r\|_2}{\|x\|_2} k_s$$

یک λ ای هست که این نابرابری برقرار باشد.

حالا اگر هیچ یک از $\lambda_i - \hat{\lambda}$ ها صفر نباشد، ماتریس $\Lambda - \hat{\lambda} I$ وارون پذیر است $\Leftarrow$

$$S(\Lambda - \hat{\lambda}I) S^{-1}\hat{x} = r \xrightarrow{S^{-1}} (\Lambda - \hat{\lambda}I) S^{-1}\hat{x} = S^{-1} r \xrightarrow{(\Lambda - \hat{\lambda}I)^{-1}} S^{-1}\hat{x} = (\Lambda - \hat{\lambda}I)^{-1} S^{-1} r$$

$$\xrightarrow{S} \hat{x} = S(\Lambda - \hat{\lambda}I)^{-1} S^{-1} r \longrightarrow \|x\|_2 = \|S(\Lambda - \hat{\lambda}I)^{-1} S^{-1} r\|_2$$

$$\Rightarrow \|x\|_2 \leq \|S\|_2 \, \|(\Lambda - \hat{\lambda}I)^{-1}\|_2 \, \|S^{-1}\|_2 \, \|r\|_2 \longrightarrow \|x\|_2 \leq k_s \, \|r\|_2 \, \|(\Lambda - \hat{\lambda}I)^{-1}\|_2$$

$$\Rightarrow \frac{1}{\|(\Lambda - \hat{\lambda}I)^{-1}\|_2} \leq k_s \frac{\|r\|_2}{\|x\|_2}$$

ماتریس $(\Lambda - \hat{\lambda}I)^{-1}$ عبارت است از

$$\begin{bmatrix} \frac{1}{\lambda_1 - \hat{\lambda}} & & O \\ & \ddots & \\ O & & \frac{1}{\lambda_n - \hat{\lambda}} \end{bmatrix} \longrightarrow norm2((\Lambda - \hat{\lambda}I)^{-1}) = \sigma_{max}((\Lambda - \hat{\lambda}I)^{-1})$$

$$= 1/|\lambda_i - \hat{\lambda}| \text{ such that } 1/|\lambda_i - \hat{\lambda}| \text{ is maximum}$$

$$\longrightarrow \frac{1}{\frac{1}{|\lambda_i - \hat{\lambda}|}} \leq k_s \frac{\|r\|_2}{\|x\|_2}$$

$\frac{1}{|\lambda_i - \hat{\lambda}|}$ ها بین تمام ماکزیمم است.

$$\longrightarrow |\lambda_i - \hat{\lambda}| \leq k_s \frac{\|r\|_2}{\|x\|_2}$$

پس همواره یک λ ای داریم که $|\lambda-\hat{\lambda}| \le k_3 \frac{\|r\|_2}{\|x\|_2}$ برقرار باشد.

۹- الف) A چون n تا مقدار ویژه‌ی متمایز دارد پس n تا بردار ویژه‌ی متمایز نیز دارد: بردارهای ویژه را v_i درنظر بگیریم، (v_iها مستقل خطی اند)

$$Av_i = \lambda_i v_i \longrightarrow A\underbrace{[v_1|\cdots|v_n]}_{V} = [\lambda_1 v_1|\text{—}|\lambda_n v_n] = V\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix} = V\Lambda$$

$$\Rightarrow AV = V\Lambda \longrightarrow A = V\Lambda V^{-1} \longrightarrow$$ A قطری شدنی است.

مقادیر ویژه‌ی A ← ماتریس حاصل از بردارهای ویژه A

حالا داریم که $AB=BA$ ⇐

$$ABv_i = BAv_i = B(\lambda_i v_i) \rightarrow A(Bv_i) = \lambda_i(Bv_i) \rightarrow$$

Bv_i یه بردار ویژه متناظر با λ_i است ⇔ و v_i هم یه بردار ویژه متناظر λ_i در A است ←

Bv_i یک ضریب از v_i باید باشد ⇐ $Bv_i = c_i v_i$ ⇐ v_iها بردار ویژه‌های B هم هستند

بنابراین B، n بردار ویژه متمایز دارد ← B هم قطری شدنی است.

⇐ $AB\, v_i = \lambda_i c_i v_i$ ⇐ AB هم n تا بردار ویژه متمایز دارد ← قطری شدنی است.

ب) عکس این رابطه برقرار است: در نظر بگیرید A و B همزمان قطری شدنی باشند، آنگاه داریم:

جابه‌جایی پذیرند $\begin{bmatrix}d_1 d_1' & & \\ & \ddots & \\ & & d_n d_n'\end{bmatrix}$

$$D_1 = P^{-1}AP \longrightarrow A = PD_1P^{-1}$$
$$D_2 = P^{-1}BP \longrightarrow B = PD_2P^{-1}$$
$$\longrightarrow AB = PD_1\underbrace{P^{-1}P}_{I}D_2P^{-1} = PD_1D_2P^{-1}$$
$$= PD_2D_1P^{-1} = \underbrace{PD_2P^{-1}}_{B}\underbrace{PD_1P^{-1}}_{A} = BA$$

$$\Rightarrow \boxed{AB=BA}$$

Scanned with CamScanner

ج) اول حالت خاص را نشان می دهیم،

if $A>0$:

ماتریس های A و $A^{-1/2}BA^{-1/2}$ را در نظر بگیرید، این ماتریس متقارن است ← وجود دارد ماتریس orthonormal O که $O^TA^{-1/2}BA^{-1/2}O=D$ ← ماتریس $S=A^{-1/2}O$ در نظر می گیریم،

$$\Rightarrow S^TBS=D$$

حالا S^TAS را در نظر می گیریم،

$$O^T\underbrace{A^{-1/2}AA^{-1/2}}_{I}O=O^TO=I$$

$$\Rightarrow \exists S: S^TBS=D,\ S^TAS=I$$

حالا در نظر بگیرید $\underline{\alpha A+\beta B>0}$ این ماتریس را C می نامیم، C هم متقارن است

$$C^T=\alpha A^T+\beta B^T=\alpha A+\beta B=C$$

ماتریس های A, C را در نظر بگیرید،

همچنین ماتریس $C^{-1/2}AC^{1/2}$ را در نظر بگیرید، این ماتریس متقارن است ⇐ وجود دارد ماتریس orthonormal O به طوریکه $O^TC^{-1/2}AC^{-1/2}O=D$ ← حالا $S=A^{-1/2}O$ را در نظر می گیریم،

$$\Rightarrow S^TAS=D \qquad S^TCS=O^T\underbrace{C^{-1/2}CC^{-1/2}}_{I}O=O^TO=I$$

$$\boxed{* \quad C=Q\Lambda Q^T \to C^{1/2}=Q\Lambda^{1/2}Q^T \rightsquigarrow C^{-1/2}=Q\Lambda^{-1/2}Q^T}$$

$$\Rightarrow S^TCS=I,\ S^TAS=D:\quad S^T(\alpha A+\beta B)S=\alpha S^TAS+\beta S^TBS=I$$

اگر $\beta=0$ باشد که همان اثبات بالا برقرار می شود،

$$\to \beta S^TBS=I-\alpha S^TAS \longrightarrow$$

اگر $\beta\neq0$ باشد،

$$S^TBS=\frac{1}{\beta}\underbrace{(I-\alpha S^TAS)}_{\text{diagonal}}$$

← طرف راست ماتریس قطری است

S^TBS نیز یک ماتریس قطری است.