98101588 HW1 احمدامین رامی

-1

الف) برای هر ماتریس دلخواه P داریم:

$$x^T P x = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} x_i x_j \quad \text{I}$$

$$x^T P x = \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 = \sum_{i=1}^{n-1} x_i^2 + x_{i+1}^2 - 2 x_i x_{i+1} \quad \text{II}$$

از مقایسه I و II به نتیجه زیر می رسیم:

$$P_{ij} = \begin{cases} 1 & i=j=1,n \\ 2 & i=j\neq 1,n \\ -1 & |i-j|=1 \\ 0 & \text{otherwise} \end{cases}$$

برای مثال برای $n=4$، P برابر ماتریس رو به رو است:

$$P = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

ب) شرط کافی برای اینکه نشان دهیم P مثبت نیمه معین است، این است که نشان دهیم $x^T P x \geq 0$ به ازای همه $x \in \mathbb{R}^n$ برقرار است. طبق فرض سوال داریم:

$$x^T P x = \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 \geq 0 \Rightarrow x^T P x \geq 0 \Rightarrow P \succeq 0$$

اما گزاره $P \succ 0$ صحیح نیست. زیرا باید به ازای همه $x \in \mathbb{R}^n$، $x^T P x > 0$ برقرار باشد. اما اگر $x = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ باشد، داریم: $x^T P x = 0$ پس گزاره $P \succ 0$ صحیح نیست.

-2

الف) طبق فرض سوال داریم:

$$\exists x \in R^n : Ax = 0 \implies Ax = 0 \times x$$

پس $\lambda = 0$ یکی از مقادیر ویژه ماتریس A است. همچنین می دانیم که $\det(A)$ برابر با ضرب مقادیر ویژه A است. از آن جایی که یکی از مقادیر ویژه برابر صفر است، داریم:

$$\det(A) = 0$$

ب) ابتدا ثابت می کنیم تمامی مقادیر ویژه ماتریس A بین 1 و -1 قرار دارد. داریم:

$$\|A\|_2 = \max \frac{\|Ax\|}{\|x\|} < 1 \implies \|Ax\| < \|x\| \quad \text{for } \forall x \in R^n$$

حال فرض کنیم v یک بردار ویژه ماتریس A و λ مقدار ویژه آن باشد. داریم:

$$Av = \lambda v \implies \|Av\| = |\lambda| \|v\| < \|v\| \implies |\lambda| < 1 \implies -1 < \lambda < 1$$

پس تمام مقدار ویژه های آن بین 1 و -1 است. مقدار ویژه های $A+I$ برابر $\lambda+1$ است. پس داریم:

$$-1 < \lambda < 1 \implies 0 < \lambda < 2$$

پس تمام مقدار ویژه های $A+I$ غیر صفر هستند. پس دترمینان آن غیر صفر است. پس معکوس پذیر است.

الف) درست است. زیرا:

$$AB=0 \Rightarrow C(A) \subseteq N(A) \Rightarrow Dim(C(A)) \leq Dim(N(A))$$

$$\Rightarrow \min\{m,n\} \leq n-\min\{n,p\} \Rightarrow n \geq \min\{m,n\}+\min\{n,p\}$$

ب) درست است. زیرا اگر λ مقدار ویژه ماتریس A باشد، آنگاه λ^k مقدار ویژه ماتریس A^k است.

حال داریم:

$$A^k=0 \Rightarrow A^kI=0 \Rightarrow \forall i\in\{1,2,\dots,n\}: A^ke_i=0=0\times e_i$$

پس $\lambda=0$ تنها مقدار ویژه ماتریس A^k است. پس تنها مقدار ویژه ماتریس A نیز صفر است.

پس مقادیر ویژه ماتریس $A-I$ برابر -1 است. پس دترمینان آن برابر $(-1)^n$ است. پس معکوس پذیر است.

ج) درست است. فرض کنید v بردار ویژه ماتریس AB و λ مقدار ویژه متناظر آن باشد.

تعریف کنید $u=Bv$.

حالت 1: اگر $\lambda\neq0$ باشد، داریم:

$$BAu=BABv=\lambda Bv=\lambda u \Rightarrow$$ u بردار ویژه ماتریس BA با مقدار ویژه λ است.

حالت 2: اگر $\lambda=0$ باشد:

اگر $Bv\neq0$ آنگاه $Au=0$ پس داریم:

$$BAu=B\times0=0\times u \Rightarrow$$ پس $\lambda=0$ یک مقدار ویژه BA است.

پس در هر صورت مقادیر ویژه AB و BA یکسان هستند. پس حکم درست است.

-4

الف)

فرض کنید ستون iام ماتریس A را با A_i نشان دهیم. داریم:

$$\|A_F\| = \sqrt{Tr\{A^TA\}} = \sqrt{\sum_{j=1}^{n} |A_j|^2} \qquad |A_j|^2 = \sum_{i=1}^{m} A_{ij}^2$$

$$\Longrightarrow \|A_F\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2}$$

ب) مقادیر ویژه AA^T برابر با توان دو مقادیر تکین A است. از طرفی $Tr\{A^TA\}$ برابر مجموع مقادیر ویژه AA^T است.

همچنین مقادیر ویژه AA^T برابر توان دو مقادیر تکین ماتریس A است. پس داریم:

$$\|A\|_F = \sqrt{Tr\{A^TA\}} = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$$

ج)

$$\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2} \geq \sigma_{max}(A)$$

$$\sqrt{r}\,\sigma_{max}(A) = \sqrt{r\,\sigma_{max}^2} \geq \sqrt{\sum_{i=1}^{r} \sigma_i^2} = \|A\|_F$$

$$\Longrightarrow \sqrt{r}\,\sigma_{max}(A) \geq \|A\|_F \Longrightarrow \sigma_{max}(A) \leq \|A\|_F \leq \sqrt{r}\,\sigma_{max}$$

الف) دستگاه معادلات زیر را در نظر بگیرید:

$$\begin{bmatrix} A & -BD^{-1} \\ C & -I \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} Ax - BD^{-1}y = b \\ Cx - y = 0 \Rightarrow y = Cx \Rightarrow Ax - BD^{-1}Cx = b \Rightarrow x = (A - BD^{-1}C)^{-1}b \quad \text{I} \end{cases}$$

$$\Rightarrow \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & -BD^{-1} \\ C & -I \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}\begin{bmatrix} b \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} A & -BD^{-1} \\ 0 & CA^{-1}BD^{-1} - I \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ -CA^{-1}b \end{bmatrix}$$

$$\Rightarrow Ax - BD^{-1}y = b \Rightarrow x = A^{-1}(b + BD^{-1}y)$$

$$(CA^{-1}BD^{-1} - I)y = CA^{-1}b \Rightarrow y = (CA^{-1}BD^{-1} - I)^{-1}CA^{-1}b$$

$$\Rightarrow x = A^{-1}(b + BD^{-1}(CA^{-1}BD^{-1} - I)^{-1}CA^{-1}b) = A^{-1}(I + B(CA^{-1}B - D)^{-1}CA^{-1})b$$

$$\Rightarrow x = (A^{-1} + A^{-1}B(CA^{-1}B - D)^{-1}CA^{-1})b \quad \text{II}$$

$$\Rightarrow (A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(CA^{-1}B - D)^{-1}CA^{-1}$$

$$\text{I}, \text{II} \Rightarrow (A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(CA^{-1}B - D)^{-1}CA^{-1}$$

ب) ابتدا به مقادیر ویژه ماتریس uv^T توجه می کنیم. چون رنک آن 1 است، تنها یک مقدار ویژه غیر صفر دارد. ادعا می کنیم مقدار ویژه غیر صفر آن v^Tu است و بردار ویژه آن u است. زیرا:

$$uv^Tu = \underbrace{(v^Tu)}_{\text{مقدار ویژه}}u$$

پس مقادیر ویژه ماتریس uv^T برابر صفر و u^Tv است. پس مقادیر ویژه ماتریس $I+uv^T$ برابر مقادیر 1 و $1+u^Tv$ است. برای اینکه $I+uv^T$ معکوس پذیر باشد، این مقادیر باید ناصفر باشند، باید داشته باشیم:

$$1+u^Tv \neq 0 \quad \text{پس:} \quad u^Tv \neq -1$$

پس شرط معکوس پذیری برابر است با: $\boxed{u^Tv \neq -1}$

حال ادعا کنیم:

$$(I+uv^T)^{-1} = I - \frac{uv^T}{1+v^Tu}$$

زیرا طبق تعریف ماتریس وارون داریم:

$$(I+uv^T)\left(I - \frac{uv^T}{1+u^Tv}\right) = I - \frac{uv^T}{1+u^Tv} + uv^T - \frac{uv^Tuv^T}{1+u^Tv}$$

$$= I + uv^T + \frac{(-1-v^Tu)uv^T}{1+v^Tu} = I + uv^T - uv^T = I \implies \text{ادعا درست است}$$

ج) طبق تعریف ماتریس وارون داریم:

$$(A+uv^T)\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\right)$$

$$= I - \frac{uv^TA^{-1}}{1+v^TA^{-1}u} + uv^TA^{-1} - \frac{uv^TA^{-1}uv^TA^{-1}}{1+v^TA^{-1}u} = I + uv^TA^{-1} - \frac{(1+v^TA^{-1}u)uv^TA^{-1}}{1+v^TA^{-1}u}$$

$$= I + uv^TA^{-1} - uv^TA^{-1} = I$$

د) با توجه به schur complement، ماتریس M را می‌توان به صورت زیر تجزیه کرد:

$$M = \underbrace{\begin{bmatrix} I & BC^{-1} \\ 0 & I \end{bmatrix}}_{N} \underbrace{\begin{bmatrix} A-BC^{-1}B^T & 0 \\ 0 & C \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} I & 0 \\ C^{-1}B^T & I \end{bmatrix}}_{N^T}$$

حال ثابت می کنیم ماتریس P مثبت معین است. فرض کنید داشته باشیم: (اگر M مثبت معین باشد)

x^TPx .

حال چون روی قطر اصلی ماتریس N، I قرار دارد، معکوس پذیر است. تعریف کنید $y=(N^T)^{-1}x$ حال داریم:

$$y=(N^T)^{-1}x \Rightarrow x=N^Ty \Rightarrow x^TPx = y^TNPN^Ty = y^TMy>0$$

$$\Rightarrow x^TPx>0 \Rightarrow P>0$$

حال ثابت می کنیم $C>0$ و $A-BC^{-1}B^T>0$ می باشند. داریم:

$x\in R^n$, Define $x=\begin{bmatrix}x_1\\x_2\end{bmatrix}$, $x_1\in R^r$, $x_2\in R^{n-r}$

ثابت کردیم که P مثبت معین است. پس داریم:

$$x^TPx>0 \Rightarrow [x_1^T\ \ x_2^T]\begin{bmatrix}A-BC^{-1}B^T&0\\0&C\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}>0$$

$$\Rightarrow x_1^T(A-BC^{-1}B^T)x_1+x_2^TCx_2>0 \quad \forall x_1\in R^r, \forall x_2\in R^{n-r}$$

رابطه فوق برای هر x_1 و x_2 برقرار است. پس قرار دهید: x_1 دلخواه و $x_2=0$. آنگاه داریم:

$$x_1^T(A-BC^{-1}B^T)x_1>0 \Rightarrow A-BC^{-1}B^T>0$$

این بار قرار دهید: $x_1=0$ و x_2 دلخواه. آنگاه داریم:

$$x_2^TCx_2>0\Rightarrow C>0$$

پس ثابت کردیم که اگر $M>0$ باشد آنگاه $C>0$ و $A-BC^{-1}B^T>0$ است.

حال ثابت می کنیم اگر $C>0$ و $A-BC^{-1}B^T>0$ باشند، $M>0$ است. ابتدا توجه کنید که اگر $C>0$ و $A-BC^{-1}B^T>0$ باشند، آنگاه $P>0$ نیز برقرار است. حال داریم:

$x^TMx=x^TNPN^Tx$ Define $y=N^Ty$ then:

$$x^TMx=x^TNPN^Tx = y^TPy>0 \Rightarrow x^TMx>0 \Rightarrow M>0$$

پس ثابت کردیم اگر C و $A-BC^{-1}B^T$ مثبت معین باشند، M نیز مثبت معین است.

پس حکم ثابت شد.

حال ثابت می‌کنیم اگر $C>0$ باشد و $M\geq 0$ باشد، آنگاه $A-BC^{-1}B^T\geq 0$ است. در بخش قبل ثابت کردیم که اگر $M\geq 0$ باشد، ماتریس P نیز که در بخش قبل تعریف کردیم نیز مثبت نیمه معین است.

به عبارتی داریم:

$$x^TPx\geq 0 \ , \ \text{Define } x=\begin{bmatrix}x_1\\x_2\end{bmatrix}, \ x_1\in R^r , \ x_2\in R^{n-r}$$

$$\Rightarrow [x_1^T \ \ x_2^T]\begin{bmatrix}A-BC^{-1}B^T & 0\\ 0 & C\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}\geq 0$$

$$\Rightarrow x_1^T(A-BC^{-1}B^T)x_1+x_2^TCx_2\geq 0$$

حال قرار دهید $x_2=0$ آنگاه:

$$x_1^T(A-BC^{-1}B^T)x_1\geq 0 \ \text{پس} \ A-BC^{-1}B^T\geq 0$$

پس اگر M مثبت نیمه معین باشد، $A-BC^{-1}B^T$ نیز مثبت نیمه معین است.

حال ثابت می‌کنیم اگر $C>0$ باشد و $A-BC^{-1}B^T\geq 0$ باشد، آنگاه $M\geq 0$ است. داریم:

$$\text{Define } x\in R^n, \ x=\begin{bmatrix}x_1\\x_2\end{bmatrix}, \ x_1\in R^r , \ x_2\in R^{n-r}$$

$$\Rightarrow x^TPx \Rightarrow x^TPx=[x_1^T \ \ x_2^T]\begin{bmatrix}A-BC^{-1}B^T & 0\\ 0 & C\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

$$=\underbrace{x_1^T(A-BC^{-1}B^T)x_1}_{\geq 0}+\underbrace{x_2^TCx_2}_{>0} \Rightarrow x^TPx\geq 0$$

حال داریم:

$$y^TMy=y^TNPN^T\underbrace{y}_{x}=x^TPx\geq 0 \Rightarrow y^TMy\geq 0 \Rightarrow M\geq 0$$

پس M مثبت نیمه معین است.

پس حکم ثابت شد.

الف) باید نشان دهیم:

$$\underset{X:\operatorname{rank}(X)\le k}{\arg\min} \|A-X\|_F^2 = A_k$$

ابتدا توجه کنید که اگر $X=A_k$ باشد داریم:

$$\|A-A_k\|_F^2 = \sum_{i=k+1}^{n} \sigma_i^2(A)$$

برای اینکه حکم را ثابت کنیم، عبارت زیر را نشان خواهیم داد:

$$\forall X: \operatorname{rank}(X)\le k \quad \|A-X\|_F^2 \ge \sum_{i=k+1}^{n} \sigma_i^2(A)$$

می دانیم اگر $A=A'+A''$ باشد، داریم: $\sigma_1(A')+\sigma_1(A'')\ge\sigma_1(A)$

حال داریم:

$$\sigma_i(A')+\sigma_j(A'')=\sigma_1(A'-A'_{i-1})+\sigma_1(A''-A''_{j-1})\ge\sigma_1(A'+A''-A'_{i-1}-A''_{j-1})$$

$$=\sigma_1(A-A'_{i-1}-A''_{j-1})\ge\sigma_1(A-A_{i+j-2})=\sigma_{i+j-1}(A)$$

حال فرض کنید X_k یک ماتریس رنک k باشد، قرار دهید $A'=A-X_k$ و $A''=X_k$.

حال قرار دهید $j=k+1$ و $i\ge 1$ داریم:

$$\sigma_i(A-X_k)+\underbrace{\sigma_{k+1}(X_k)}_{0}\ge\sigma_{i+k}(A)\Rightarrow\sigma_i(A-X_k)\ge\sigma_{i+k}(A)$$

$$\Rightarrow\sigma_i^2(A-X_k)\ge\sigma_{i+k}^2(A)\Rightarrow\sum_{i=1}^{n}\sigma_i^2(A-X_k)\ge\sum_{i=k+1}^{n}\sigma_i(A)$$

$$\Rightarrow\|A-X_k\|_F^2\ge\sum_{i=k+1}^{n}\sigma_i^2(A)=\|A-A_k\|_F^2\Rightarrow\|A-X_k\|_F^2\ge\|A-A_k\|_F^2$$

$$\Rightarrow\underset{X:\operatorname{rank}(X)\le k}{\arg\min}\|A-X\|_F^2=A_k \Longrightarrow$$ حکم ثابت شد

ب)

می‌خواهیم مسئله بهینه‌سازی رو‌به‌رو را حل کنیم:

$$\min_{X:\ \text{rank}(X)\leq K} \|A-X\|_2^2$$

ابتدا توجه کنید که اگر $X = A_K$ باشد، داریم: $\|A-A_K\|_2^2 = \sigma_{K+1}^2(A)$

حال اثبات می‌کنیم که A_K جواب مسئله بهینه‌سازی فوق است. فرض کنید ماتریس B رتبه حداکثر K دارد و مقدار بهتری نسبت به A_K دهد یعنی: $\|A-B\|_2^2 < \sigma_{K+1}^2(A)$

حال $K+1$ ستون ماتریس V را در نظر بگیرید و آن را مجموعه $\{v_1, v_2, \ldots, v_{K+1}\}$ بنامید.
ادعا می‌کنیم برداری مانند z در اشتراک $N(B)$ و $\text{span}(\{v_1, v_2, \ldots, v_{K+1}\})$ وجود دارد
یعنی: $z \in \text{span}(\{v_1, v_2, \ldots, v_{K+1}\}) \cap N(B)$ دلیل اینکه این بردار وجود دارد این است که:

$$\text{Dim}(N(B)) \geq n-K \implies \text{Dim}(N(B)) + \text{Dim}(\{v_1, v_2, \ldots, v_{K+1}\}) \geq n+1$$

$$\text{Dim}(\{v_1, v_2, \ldots, v_{K+1}\}) = K+1$$

چون جمع بعدهای آن‌ها از $n+1$ بزرگتر است، پس حتما این دو زیرفضا اشتراک دارند.
بردار z در اشتراک آن‌ها را به‌گونه‌ای انتخاب می‌کنیم که $|z|=1$ همچنین z را می‌توان به صورت رو‌به‌رو نوشت: $z = \sum_{j=1}^{K+1} a_j v_j$ حال مقدار $\|(A-B)z\|^2$ را محاسبه می‌کنیم:

$$\|(A-B)z\|^2 = \|Az - \underbrace{Bz}_{0}\|^2 = \|Az\|^2 = \|(\sum_{i=1}^{n} \sigma_i u_i v_i^T)(\sum_{j=1}^{K+1} a_j v_j)\|^2$$

$$= \|\sum_{i=1}^{n}\sum_{j=1}^{K+1} \sigma_i a_j u_i v_i^T v_j\|^2 = \|\sum_{i=1}^{K+1} \sigma_i u_i a_i\|^2 = \sum_{i=1}^{K+1} a_i^2 \sigma_i^2$$

$$\sum_{i=1}^{K+1} a_i^2 \sigma_i^2 \geq \sigma_{K+1}^2 \sum_{i=1}^{K+1} a_i^2 = \sigma_{K+1}^2 \|z\|^2 = \sigma_{K+1}^2$$

$$\implies \|(A-B)z\|^2 \geq \sigma_{K+1}^2$$

گزاره $\|(A-B)z\|_2^2 \geq \sigma_{K+1}^2$ با فرض ما که $\|(A-B)z\|_2^2 < \sigma_{K+1}^2$ بود در تناقض است. پس به تناقض رسیدیم و کمترین مقدار $\|A-X\|_2^2$ برابر σ_{K+1}^2 است که هنگامی که $X = A_K$ است حاصل می‌شود.