9810052 صبا نصیری

۱

آ

$$A = U \Sigma V^T = U \begin{bmatrix} \sigma_1 & & 0 \\ 0 & \ddots & \vdots \\ 0 & \cdots & \sigma_r \\ & & 0 \end{bmatrix} V^T \quad , \; r = \text{rank}(A) \; , \; \sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$$

$$\|A\|_2 = \max_{\|v\|_2 = 1} \|Av\| = \max(\sigma) = \sigma_1$$

$$\|A\|_F^2 = \sum_i \sum_j A(i,j)^2 = tr(A^T A) = \sum_{i=1}^{r} \underbrace{\lambda_i}_{\text{eigenvalue}(A^TA)} = \sum_{i=1}^{r} \sigma_i^2$$

$$\Longrightarrow$$

$$\|A\|_2 = \sigma_1 = \sqrt{\sigma_1^2} \leq \|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2} \leq \sqrt{r}\, \|A\|_2 =$$

$$\sqrt{r}\, \sigma_1 = \sqrt{r\, \sigma_1^2} = \sqrt{\underbrace{\sigma_1^2 + \cdots \sigma_1^2}_{r}} \Longrightarrow$$

$$\boxed{\|A\|_2 \leq \|A\|_F \leq \sqrt{\text{rank}(A)}\, \|A\|_2}$$

ب

i

$$E(X) = \sum x\, p(x) \overset{x, p \geq 0}{\Longrightarrow}$$

$$\underbrace{\sum_{x \geq 0} x\, p(x)}_{E(X)} \geq \sum_{x \geq a} x\, p(x) \geq \sum_{x \geq a} a\, p(x) = a \underbrace{\sum_{x \geq a} p(x)}_{p(X \geq a)}$$

$$\Longrightarrow$$

$$E(X) \geq a\, p(X \geq a) \Longrightarrow \boxed{p(X \geq a) \leq \frac{E(X)}{a}}$$

impulsado por CamScanner

(ii)

$$Y = (X-\mu)^2 \xRightarrow{\text{marcov inequality}} P(Y \geq a^2) \leq \frac{E(Y)}{a^2}$$

$$E(Y) = E((X-\mu)^2) = Var(X) = \sigma^2$$

$$P(|X-\mu| \geq a) = P(|X-\mu|^2 \geq a^2) = P(Y \geq a^2) \leq \frac{\overset{\sigma^2}{E(Y)}}{a^2} \Rightarrow$$

$$\boxed{P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}}$$

(iv)

متغیر برنولی با $p = \pi/4$

$$P(|X - \pi/4| \geq 0.01 \times \pi/4) \leq \frac{\sigma^2}{n} \frac{1}{(\pi/4 \times 0.01)^2} =$$

$$\frac{[\pi/4(1-\pi/4)]}{n} \times \frac{1}{(\pi/4 \times 0.01)^2} \leq 0.05$$

$$\frac{2732.39}{n} \leq 0.05 \Rightarrow 54647.9 \leq n \Rightarrow$$

$$\boxed{n = 54648}$$

impulsado por CamScanner

2

الف

$$* \quad a^T x = a_1 x_1 + \cdots + a_n x_n \Rightarrow$$

$$\frac{\partial a^T x}{\partial x} = \left[\frac{\partial a^T x}{\partial x_1}, \frac{\partial a^T x}{\partial x_2}, \ldots, \frac{\partial a^T x}{\partial x_n} \right] =$$

$$[a_1, a_2, \ldots, a_n] = \boxed{a^T}$$

$$*$$

$$\frac{\partial x^T A x}{\partial x} = \frac{\partial}{\partial x}\left(a_{11} x_1^2 + \cdots + a_{nn} x_n^2 + (a_{12}+a_{21}) x_1 x_2 + \cdots + (a_{1n}+a_{n1}) x_1 x_n\right) = \frac{\partial}{\partial x}\left(\sum_{i,j=1}^{n} a_{ij} x_i x_j\right) =$$

$$\left[2a_{11} x_1 + \sum_{j=1, j\neq 1}^{n} a_{1j} a_{j1} x_j, \ldots, 2a_{nn} x_n + \sum_{j=1, j\neq n}^{n} a_{nj} a_{jn} x_j\right] =$$

$$[x_1, \ldots, x_n] \cdot \underbrace{\begin{bmatrix} 2a_{11} & \cdots & a_{1n}+a_{n1} \\ \vdots & & \vdots \\ a_{n1}+a_{1n} & \cdots & 2a_{nn} \end{bmatrix}}_{A^T+A} = \boxed{x^T (A^T + A)}$$

impulsado por CamScanner

ب

$$\frac{\partial A^{-1}}{\partial B} = \lim_{\Delta B \to 0} \frac{(A+\Delta A)^{-1} - A^{-1}}{\Delta B} =$$

$$\lim_{\Delta B \to 0} \frac{(A+\Delta A)^{-1} A A^{-1} - (A+\Delta A)^{-1}(A+\Delta A) A^{-1}}{\Delta B} =$$

$$\lim_{\Delta B \to 0} \frac{(A+\Delta A)^{-1} (\cancel{A A^{-1}} - \cancel{A A^{-1}} - \Delta A A^{-1})}{\Delta B} = \lim_{\Delta B \to 0} \frac{(A+\Delta A)^{-1} \times (-\Delta A) A^{-1}}{\Delta B} =$$

$$- A^{-1} \lim_{\Delta B \to 0} \frac{\Delta A}{\Delta B} A^{-1} = \boxed{- A^{-1} \frac{\partial A}{\partial B} A^{-1}}$$

ج

* laplace expansion

$$|X| = \sum_{j=1}^{n} x_{1j} A_{1j} = \sum_{j=1}^{n} x_{2j} A_{2j} = \dots = \sum_{j=1}^{n} x_{nj} A_{nj} =$$

$$\sum_{j=1}^{n} x_{1j}\, \mathrm{adj}(x)_{j1} = \sum_{j=1}^{n} x_{2j}\, \mathrm{adj}(x)_{j2} = \dots = \sum_{j=1}^{n} x_{nj}\, \mathrm{adj}(x)_{jn}$$

where: $\mathrm{adj}(x) = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}^T$

impulsado por CamScanner

$$\frac{\partial |x|}{\partial x} = \nabla |x| = \begin{vmatrix} \frac{\partial |x|}{\partial x_{11}} & \cdots & \frac{\partial |x|}{\partial x_{1n}} \\ \vdots & & \\ \frac{\partial |x|}{\partial x_{n1}} & \cdots & \frac{\partial |x|}{\partial x_{nn}} \end{vmatrix} =$$

$$\begin{vmatrix} \frac{\partial \sum x_{1j}\, adj(x)_{j1}}{\partial x_{11}} & \frac{\partial \sum x_{1j}\, adj(x)_{j1}}{\partial x_{12}} & \cdots & \frac{\partial \sum x_{1j}\, adj(x)_{j1}}{\partial x_{1n}} \\ \frac{\partial \sum x_{2j}\, adj(x)_{j2}}{\partial x_{21}} & & & \vdots \\ \vdots & & & \\ \frac{\partial \sum x_{nj}\, adj(x)_{jn}}{\partial x_{n1}} & \cdots & \cdots & \frac{\partial \sum x_{nj}\, adj(x)_{jn}}{\partial x_{nn}} \end{vmatrix} =$$

$$\begin{vmatrix} adj(x)_{11} & adj(x)_{21} & \cdots & adj(x)_{n1} \\ adj(x)_{12} & & & \vdots \\ \vdots & & & \\ adj(x)_{1n} & \cdots & \cdots & adj(x)_{nn} \end{vmatrix} =$$

$$(adj(x))^T = \Big(|x| \cdot \underbrace{\frac{adj(x)}{|x|}}_{x^{-1}}\Big)^T = \boxed{|x|\ x^{-T}}$$

impulsado por CamScanner

$$* \quad \frac{\partial \log|A|}{\partial A} = \frac{\partial \log|A|}{\partial |A|} \times \left(\frac{\partial |A|}{\partial A}\right)^T = \frac{1}{|A|} \times \left(|A| \cdot A^{-T}\right)^T =$$

$$\frac{1}{|A|} \times |A| \, A^{-1} = \boxed{A^{-1}}$$

3

الف) A را به صورت فرم قطری آن می‌نویسیم:

$$P^{-1} A P = \lambda \Rightarrow A = P \lambda P^{-1} \Rightarrow$$

$$tr(A) = tr(P \lambda P^{-1}) = tr(P^{-1} P \lambda) = tr(\lambda) = \boxed{\sum_{i=1}^{n} \lambda_i}$$

ب) چندجمله‌ای مشخصه

$$f(x) = \det(A - xI) = (-1)^n (x-\lambda_1)^{r_1} \dots (x-\lambda_n)^{r_n}$$

اگر $A_{n\times n}$ بیشتر n تا مقدار ویژه داشته باشیم:

$$\det(A - \lambda I) = (-1)^n (\lambda - \lambda_1) \dots (\lambda - \lambda_n) =$$

$$(-1)(\lambda - \lambda_1)(-1)(\lambda - \lambda_2) \dots (-1)(\lambda - \lambda_n) =$$

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda)$$

اگر $\lambda = 0$:

$$\boxed{\det(A) = \lambda_1 \lambda_2 \dots \lambda_n}$$

impulsado por CamScanner

4

الف

ابتدا ثابت می‌کنیم A^TA وارون‌پذیر و fullrank است:

اثبات:

$$Ax = 0 \Rightarrow A^TAx = 0 \Rightarrow$$

$$① \quad x \in \text{Nullspace}(A^TA)$$

$$A^TAx = 0 \Rightarrow x^TA^TAx = 0 \Rightarrow$$

$$(Ax)^TAx = 0 \Rightarrow Ax = 0 \Rightarrow x \in \text{nullspace}(A) \quad ②$$

$$①, ② \overset{*}{\Rightarrow} \text{nullspace}(A^TA) = \text{nullspace}(A)$$

$$\cancel{\dim(A)}^{\,n} = \cancel{\dim(\underbrace{A^TA}_{n\times n})}^{\,n} \overset{*}{\Rightarrow} \text{rank}(A^TA) = \text{rank}(A) = n$$

پس A^TA یک ماتریس $n\times n$ با رنک n است که یعنی وارون‌پذیر است

$$(A^TA)A^+ = \left(A^T(AA^+)\right)^{TT} = \left((AA^+)^TA\right)^T = (\underbrace{A^+A}_{I}A)^T = A^T \Rightarrow$$

$$\boxed{A^+ = (A^TA)^{-1}A^T}$$

impulsado por CamScanner

حال ثابت می‌کنیم ۴ خواص زیر برای $A^\dagger$ به دست آمده برقرار است:

① $AA^\dagger A = A$

② $A^\dagger A A^\dagger = A^\dagger$

③ $(AA^\dagger)^T = AA^\dagger$

④ $(A^\dagger A)^T = A^\dagger A$

① $AA^\dagger A = A\overbrace{(A^TA)^{-1}(A^TA)}^{I} = A$ ✓

② $A^\dagger A A^\dagger = \underbrace{(A^TA)^{-1}(A^TA)}_{I}(A^TA)^{-1}A^T = (A^TA)^{-1}A^T = A^\dagger$

③ $(AA^\dagger)^T = \left(A(A^TA)^{-1}A^T\right)^T = \left((A^TA)^{-1}A^T\right)^T A^T =$

$\left(A\,(\underbrace{A^TA}_{\text{متقارن}})^{-T}\right)A^T = A(A^TA)^{-1}A^T = AA^\dagger$ ✓

④ $(A^\dagger A)^T = \left((A^TA)^{-1}A^TA\right)^T = A^TA\,(\underbrace{A^TA}_{\text{متقارن}})^{-T} = (A^TA)(A^TA)^{-1} =$

$I = (A^TA^{-1})(A^TA) = A^\dagger A$ ✓

پس $\boxed{A^\dagger = (A^TA)^{-1}A^T}$.

---

* راه دوم:

$(A^TA)^{-1}A^T = \left((U\Sigma V^T)^T U\Sigma V^T\right)^{-1}(U\Sigma V^T)^T =$

impulsado por CamScanner

$$\left(V \Sigma \underbrace{U^T U}_{I} \Sigma V^T\right)^{-1} \left(U \Sigma V^T\right)^T =$$

$$\left(V \Sigma^2 V^T\right)^{-1} V \Sigma U^T = V \Sigma^{-2} \underbrace{V^T V}_{I} \Sigma U^T =$$

$$\boxed{V \Sigma^{-1} U^T = A^+}$$

impulsado por CamScanner

ب

ابتدا ثابت می‌کنیم (AA^T) وارون‌پذیر و full rank است:

اثبات:

$$A^T x = 0 \Rightarrow A A^T x = 0 \Rightarrow$$

$$① \quad x \in \text{nullspace}(A A^T)$$

$$A A^T x = 0 \Rightarrow x^T A A^T x = 0 \Rightarrow$$

$$(A^T x)^T (A^T x) = 0 \Rightarrow A^T x = 0 \Rightarrow$$

$$② \quad x \in \text{nullspace}(A^T)$$

$$①, ② \overset{*}{\Rightarrow} \text{nullspace}(A^T) = \text{nullspace}(AA^T)$$

$$\dim(\overset{m}{A^T}) = \dim(\overset{m}{AA^T}) \overset{*}{\Rightarrow}$$

$$\text{rank}(A^T) = \text{rank}(AA^T) = m$$

پس AA^T یک ماتریس $m \times m$ با رنک m است یعنی وارون‌پذیر است:

$$A^\dagger (AA^T) = ((A^\dagger A) A^T)^{TT} = (A^\dagger (A^\dagger A)^T)^T =$$

$$(\underbrace{A A^\dagger A}_{I})^T = A^T \Rightarrow \boxed{A^\dagger = A^T (AA^T)^{-1}}$$

impulsado por CamScanner

حال ثابت می‌کنیم ۴ خاصیت زیر برای $A^\dagger$ بدست آمده برقرار است:

① $AA^\dagger A = A$

② $A^\dagger A A^\dagger = A^\dagger$

③ $(AA^\dagger)^T = AA^\dagger$

④ $(A^\dagger A)^T = A^\dagger A$

① $AA^\dagger A = \underbrace{AA^T(AA^T)^{-1}}_{I}A = A$ ✓

② $A^\dagger A A^\dagger = A^T\underbrace{(AA^T)^{-1}A\,A^T}_{I}(AA^T)^{-1} = A^T(AA^T)^{-1} = A^\dagger$ ✓

③ $(AA^\dagger)^T = \left(AA^T(AA^T)^{-1}\right)^T = \left(A^T(AA^T)^{-1}\right)^T A^T =$

$\left((AA^T)^{-T}A\right)A^T = (AA^T)^{-1}(AA^T) = I = AA^T(AA^T)^{-1} = AA^\dagger$ ✓

④ $(A^\dagger A)^T = A^T(A^\dagger)^T = A^T\left(A^T(AA^T)^{-1}\right)^T =$

$A^T\left((AA^T)^{-T}A\right) = A^T\left((AA^T)^{-1}A\right) = \left(A^T(AA^T)^{-1}\right)A =$

$A^\dagger A$ ✓

پس $\boxed{A^\dagger = A^T(AA^T)^{-1}}$

---

* راه دوم:

$A^T(AA^T)^{-1} = (U\Sigma V^T)^T\left((U\Sigma V^T)(U\Sigma V^T)^T\right)^{-1} =$

impulsado por CamScanner

$$(V \Sigma U^T)(U \Sigma \underbrace{V^T V}_{I} \Sigma U^T)^{-1} = (V \Sigma U^T)(U \Sigma^2 U^T)^{-1} =$$

$$V \Sigma \underbrace{U^T U}_{I} \Sigma^{-2} U^T = \boxed{V \Sigma^{-1} U^T = A^{\dagger}}$$

impulsado por CamScanner

(5)

الف)

ابتدا ثابت می‌کنیم اگر $M' = \begin{bmatrix} A_{n\times n} & B_{n\times k} \\ 0_{k\times n} & D_{k\times k} \end{bmatrix}$ باشد $Det(M') = Det(A)Det(D)$.

از استقرا استفاده می‌کنیم.

پایه استقرا:

$n=1$:

$$M' = \begin{bmatrix} a_{11} & B_{1\times k} \\ 0_{k\times 1} & D_{k\times k} \end{bmatrix} \Rightarrow$$

$$Det(M') = a_{11}\, Det(D) = det(A)\, det(D) \checkmark$$

حال فرض می‌کنیم حکم برای $A_{(n-1)\times(n-1)}$ برقرار است. حکم را برای $A_{n\times n}$ اثبات می‌کنیم.

$$M' = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & b_{n1} & \cdots & b_{nk} \\ 0 & \cdots & 0 & d_{11} & \cdots & d_{1k} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & d_{k1} & \cdots & d_{kk} \end{bmatrix}$$

A_{ij} cofactor، a_{ij} و B^i ماتریس حاصل از حذف سطر i از B است.

$\Rightarrow$

$$det(M') = a_{11}\begin{vmatrix} A_{11} & B^1 \\ 0 & D \end{vmatrix} - a_{21}\begin{vmatrix} A_{21} & B^2 \\ 0 & D \end{vmatrix} + \cdots +$$

$$(-1)^{n+1} a_{n1} \begin{vmatrix} A_{n1} & B^n \\ 0 & D \end{vmatrix}$$

ماتریس

$det\left(\begin{bmatrix} A_{ij} & B^i \\ 0 & D \end{bmatrix}\right)$ برابر $det(A_{ij})\,det(D)$ است زیرا $(n-1)\times(n-1)$ است. پس:

impulsado por CamScanner

$$\det(M') = a_{11}\det(A_{11})\det(D) - a_{21}\det(A_{21})\det(D) + \cdots + a_{n1}(-1)^{n+1}\det(A_{n1})\det(D) =$$

$$\det(D)\underbrace{\left[a_{11}\det(A_{11}) - a_{21}\det(A_{21}) + \cdots + a_{n1}(-1)^{n+1}\det(A_{n1})\right]}_{\det(A)} = \boxed{\det(A)\det(D)}$$

حال برای اثبات حکم خواسته شده، ضرب زیر را در نظر می‌گیریم:

$$\underbrace{\begin{bmatrix} A^{-1} & 0 \\ -CA^{-1} & I \end{bmatrix}}_{\text{به طور مشابه طبق حکم ثابت شده در بالای صفحه}} \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{M} = \underbrace{\begin{bmatrix} I & A^{-1}B \\ 0 & -CA^{-1}B+D \end{bmatrix}}_{\text{طبق حکم ثابت شده در بالای صفحه}} \xRightarrow{\text{از طرفین}} \text{می‌گیریم}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\underbrace{\det(A^{-1})}_{\frac{1}{\det(A)}}\cancel{\det(I)}^{\,1}\det(M) = \cancel{\det(I)}^{\,1}\det(-CA^{-1}B+D) \Rightarrow$$

$$\boxed{\det(M) = \det(A)\det(-CA^{-1}B+D)} \checkmark\checkmark$$

impulsado por CamScanner

ب

برای اثبات این بخش، از نتایج بخش قبل و ضرب زیر، استفاده می‌کنیم:

$$\underbrace{\begin{bmatrix} 0 & D^{-1} \\ I & -BD^{-1} \end{bmatrix}}_{\text{به طور ثابت به طبق حکم ثابت شده در بخش قبل}} \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{M} = \underbrace{\begin{bmatrix} D^{-1}C & I \\ A-BD^{-1}C & 0 \end{bmatrix}}_{\text{به طور ثابت به طبق حکم ثابت شده در بخش قبل}} \xRightarrow[\text{دترمینان}]{} \text{گرفتن}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\underbrace{\det(D^{-1})}_{\frac{1}{\det(D)}} \cancel{\det(I)}^{\,1} \det(M) = \cancel{\det(I)}^{\,1} \det(A-BD^{-1}C) \Rightarrow$$

$$\boxed{\det(M) = \det(D)\det(A-BD^{-1}C)} \checkmark\checkmark$$

ج

$$\begin{bmatrix} A & B \\ C & -D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix} \Rightarrow$$

$$\begin{cases} Ax + By = I \Rightarrow \text{I} \quad x = A^{-1} - A^{-1}By \\ Cx - Dy = 0 \Rightarrow \text{II} \quad y = D^{-1}Cx \end{cases} \Rightarrow$$

$$y = D^{-1}C(A^{-1} - A^{-1}By) = D^{-1}CA^{-1} - D^{-1}CA^{-1}By \Rightarrow$$

$$(I + D^{-1}CA^{-1}B)y = D^{-1}CA^{-1} \xRightarrow{\times D}$$

impulsado por CamScanner

$$(D + CA^{-1}B)y = CA^{-1} \implies y = (D + CA^{-1}B)^{-1}CA^{-1}$$

$$Ax + By = Ax + B(D + CA^{-1}B)^{-1}CA^{-1} = I \implies$$

$$* \quad x = A^{-1} - A^{-1}B(D + CA^{-1}B)^{-1}CA^{-1}$$

$$\left.\begin{array}{ll} \text{I} \implies & x = A^{-1} - A^{-1}By \\ \text{II} \implies & y = D^{-1}Cx \end{array}\right\} \implies x = A^{-1} - A^{-1}BD^{-1}Cx \implies$$

$$(I + A^{-1}BD^{-1}C)x = A^{-1} \xrightarrow{\times A}$$

$$(A + BD^{-1}C)x = I \implies x = (A + BD^{-1}C)^{-1} \quad **$$

$$*, ** \implies \boxed{(A + BD^{-1}C)^{-1} = A^{-1} - A^{-1}B(D + CA^{-1}B)^{-1}CA^{-1}}$$

impulsado por CamScanner

ماتریس های زیر را تعریف می‌کنیم:

$$\underbrace{\overbrace{\begin{bmatrix} A-uv^T & u \\ 0 & I \end{bmatrix}}^{A_1}}_{} = \underbrace{\begin{bmatrix} A_{n\times n} & u_{n\times 1} \\ v^T_{1\times n} & 1 \end{bmatrix}}_{A_2} \underbrace{\begin{bmatrix} I & 0 \\ -v^T & I \end{bmatrix}}_{A_3}$$

طبق قضیه اثبات شده در بخش اول:

$$\det(A_1) = \det(A-uv^T)\underbrace{\det(I)}_{1} = \det(A-uv^T)$$

طبق بخش الف:

$$\left.\begin{aligned} \det(A_2) &= \det(A)\det(1-v^TA^{-1}u) \\ \det(A_3) &= \underbrace{\det(I)}_{1}\underbrace{\det(I)}_{1} = 1 \end{aligned}\right\} \quad 2)$$

$$\det(A_1) = \det(A_2)\det(A_3) \Rightarrow$$

$$\boxed{\det(A-uv^T) = \det(A)\det(1-v^TA^{-1}u)}$$

impulsado por CamScanner

الف

$$A = \begin{bmatrix} B & x \\ y^* & a \end{bmatrix}$$

$\det(tI-A)$ را محاسبه خواهیم:

$$tI-A = \begin{bmatrix} tI-B & x \\ y^* & t-a \end{bmatrix}$$

$\det(tI-A)$ را حول سطر آخر آن بسط می‌دهیم:

$$\det(tI-A) = \sum_{i=1}^{n} y_i^* (-1)^{n+1+i} \det(C_{n+1,j}) + (t-a)\det(tI-B)$$

$C_{n+1,j}$ از حذف سطر $n+1$ و ستون j ماتریس A بدست می‌آید.

حال $\det(C_{n+1,j})$ را حول ستون آخر آن که x است، بسط می‌دهیم:

$$\det(C_{n+1,j}) = \sum_{j=1}^{n} x_j (-1)^{n+j} \det(d_{j,n})$$

$d_{j,n}$ از حذف سطر j و ستون n از ماتریس $C_{n+1,j}$ بدست می‌آید.

$$\det(tI-A) = \sum_{i=1}^{n} y_i^* (-1)^{n+i+1} \sum_{j=1}^{n} x_j (-1)^{n+j} \det(d_{j,n}) +$$

$$+ (t-a)\det(tI-B) =$$

impulsado por CamScanner

$$= (t-a)\det(tI-B) + \sum_{i=1}^{n} y_i^* \sum_{j=1}^{n} x_j (-1)^{i+j+1} \det(d_{j,n}) =$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} -y_i^* x_j \underbrace{(-1)^{i+j}\det(d_{j,n})}_{\operatorname{adj}(tI-B)_{i,j}} + (t-a)\det(tI-B) =$$

$$(t-a)\det(tI-B) + \sum_{i=1}^{n}\sum_{j=1}^{n} -y_i^* x_j \operatorname{adj}(tI-B)_{i,j} =$$

$$(t-a)\underbrace{\det(tI-B)}_{P_B} - y^* \operatorname{adj}(tI-B)\, x \Rightarrow$$

$$\boxed{\det(tI-A) = P_A = (t-a)P_B - y^* \operatorname{adj}(tI-B)\, x}$$

ب) طبق قضیه‌ی courant-fischer ، زیرفضاهایی وجود دارند به نام S_A و S_B که $\dim(S_A) = i+j$ و $\dim(S_B) = n-j$ و در این صورت:

$$\lambda_{i+j}(A) = \max_{\substack{x \in S_A \\ \|x\|=1}} (Ax, x)$$

$$\lambda_{n-j}(B) = \max_{\substack{x \in S_B \\ \|x\|=1}} (Bx, x)$$

impulsado por CamScanner

همچنین می‌دانیم:

$$\dim(S_A) + \dim(S_B) - \dim(S_A \cap S_B) \leq n \Rightarrow$$

$$\dim(S_A \cap S_B) \geq i + \cancel{j} + n - \cancel{j} - \cancel{n} \Rightarrow$$

$$\dim(S_A \cap S_B) \geq i$$

حال تفسیر Courant-fischer را برای $A+B$ اعمال می‌کنیم:

$$\lambda_i(A+B) = \min_{\dim(S)=i} \max_{\substack{x \in S \\ \|x\|=1}} \langle (A+B)x, x \rangle \leq$$

$$\min_{\substack{\dim(S)=i \\ S \subseteq S_A \cap S_B}} \max_{\substack{x \in S \\ \|x\|=1}} \left(\langle Ax, x \rangle + \langle Bx, x \rangle \right) \leq$$

$$\min_{\substack{\dim(S)=i \\ S \subseteq S_A \cap S_B}} \left(\max_{\substack{x \in S \\ \|x\|=1}} \langle Ax, x \rangle + \max_{\substack{x \in S \\ \|x\|=1}} \langle Bx, x \rangle \right) =$$

$$\lambda_{i+j}(A) + \lambda_{n-j}(B) \Rightarrow$$

$$\boxed{\lambda_i(A+B) \leq \lambda_{i+j}(A) + \lambda_{n-j}(B), \quad j = 0, \dots, n-i, \quad i = 1, \dots, n}$$

impulsado por CamScanner

ب

طبق بخش الف:

$$\left.\begin{aligned} \det(tI-A) &= (t-a)\,P_B - y^* \operatorname{adj}(tI-B)\,y \\ \underbrace{\det(tI-B)}_{P_B} &= \frac{\operatorname{adj}(tI-B)}{(tI-B)^{-1}} \end{aligned}\right\} \Rightarrow$$

$$\underbrace{\det(tI-A)}_{P_A} = (t-a)\,P(B) - y^*\,P_B\,(tI-B)^{-1}\,y$$

$$P_A = P_B\left(t-a-y^*(tI-B)^{-1}y\right)$$

ریشه های چندجمله‌ای مشخصه ی هر ماتریس، λ های آن است. طبق رابطه ی بالا، اگر $P_B=0$ آنگاه $P_A=0$ پس تمام λ های B، λ های A نیز هستند ولی A دارای یک مقدار ویژه‌ی $\lambda = a+y^*(tI-B)^{-1}y$ نیز است. فرض می‌کنیم این مقدار ویژه برابر λ_i باشد پس:

$$\lambda_1(A)=\lambda_1(B)\ /\ \lambda_2(A)=\lambda_2(B)\ ,\ \dots\ /\ \lambda_i(A)\ /$$

$$\lambda_{i+1}(A)=\lambda_i(B)\ /\ \dots\ /\ \lambda_{n+1}(A)=\lambda_n(B)$$

پس:

$$\lambda_1(A)=\lambda_1(B) \leq \lambda_2(A)=\lambda_2(B) \leq \dots \leq \lambda_i(A) \leq$$

$$\lambda_{i+1}(A)=\lambda_i(B) \leq \dots \leq \lambda_{n+1}(A)=\lambda_n(B) \Rightarrow$$

impulsado por CamScanner

$$\lambda_1(A) \leq \lambda_1(B) \leq \lambda_2(A) \leq \lambda_2(B) \leq \cdots \leq$$

$$\lambda_i(B) \leq \lambda_{i+1}(A) \leq \cdots \leq \lambda_n(B) \leq \lambda_{n+1}(A)$$

یک حکم ثابت می‌شود.

impulsado por CamScanner

⑦ اگر نمونه‌ها iid باشند:

$$f = f(x_1, x_2, \dots, x_n) = \left(\frac{1}{\theta^2}\right)^n x_1 \cdots x_n \, e^{-\frac{(x_1 + \dots + x_n)}{\theta}}$$

$$\frac{\partial f}{\partial \theta} = 0 \Rightarrow (x_1 \cdots x_n) \left[n \left(\frac{1}{\theta^2}\right)^{n-1} \times -2 \frac{1}{\theta^3} e^{-\frac{(x_1 + \dots + x_n)}{\theta}} + \right.$$

$$\left. \left(\frac{1}{\theta^2}\right)^n \times e^{-\frac{(x_1 + \dots + x_n)}{\theta}} \times -(x_1 + \dots + x_n) \times \frac{-1}{\theta^2} \right] = 0$$

$$\Rightarrow$$

$$\frac{-2n}{\theta^3} + \frac{1}{\theta^2} \times (-x_1 - \dots - x_n) \times \frac{-1}{\theta^2} = 0 \Rightarrow$$

$$\frac{\sum_{i=1}^{n} x_i}{\cancel{\theta^4}\,\theta} = \frac{2n}{\cancel{\theta^3}} \Rightarrow \boxed{\theta_{MLE} = \frac{\sum_{i=1}^{n} x_i}{2n}}$$

⑧

$\boxed{\text{MLE:}}$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$f = f(x_1, \dots, x_n) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}} \Rightarrow$$

$$\frac{\partial f}{\partial \mu} = 0 \Rightarrow \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}} \times \frac{+1}{2\sigma^2} \times \sum_{i=1}^{n} 2(x_i-\mu) = 0$$

$$\Rightarrow$$

impulsado por CamScanner

$$\sum_{i=1}^{n}(x_i-\mu)=0 \Rightarrow \sum_{i=1}^{n} x_i = n\mu \Rightarrow$$

$$\boxed{\mu_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n}}$$

$$E[(\overset{\mu_{MLE}}{\mu'}-\mu)^2] = E[(\mu'-E(\mu')+E(\mu')-\mu)^2] =$$

$$E[(\mu'-E(\mu'))^2] + 2E[\underbrace{\mu'-E(\mu')}_{0}](\mu-E(\mu')) +$$

$$(E(\mu')-\mu)^2 = var(\mu') + (E(\mu')-\mu)^2 =$$

$$Var(\bar{x}) + (E(\bar{x})-\mu)^2 = \frac{\sigma^2}{n} + \frac{\sum_{i=1}^{n} E(x_i)}{n} - \mu =$$

$$\frac{\sigma^2}{n} + \frac{n\mu}{n} - \mu = \frac{\sigma^2}{n}$$

$$\lim_{n\to\infty} E((\mu'-\mu)^2) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

پس میانگین مربعات اختلاف μ و μ_{MLE} اگر $n\to\infty$، به صفر میل می‌کند.

پس هرچه نمونه بیشتری بگیریم، تخمین دقیق‌تری از μ خواهیم داشت و اگر $n\to\infty$

μ_{MLE} دقیقاً برابر μ خواهد شد.

impulsado por CamScanner

$\boxed{\text{MAP:}}$

$$p(\mu|x)\,p(x) = p(x|\mu)\,p(\mu) \Rightarrow$$

$$p(\mu|x) = \frac{p(x|\mu)\,p(\mu)}{p(x)} =$$

$$\frac{p(x_1|\mu)\cdots p(x_n|\mu)\,p(\mu)}{p(x)}$$

$$\mu_{map} = \underset{\mu}{\operatorname{argmax}}\; p(\mu|x) = \underset{\mu}{\operatorname{argmax}}\; \frac{p(x|\mu)\,p(\mu)}{p(x)} =$$

$$\underset{\mu}{\operatorname{argmax}}\; p(x|\mu)\,p(\mu) = \underset{\mu}{\operatorname{argmax}}\; p(x_1|\mu)\cdots p(x_n|\mu)\,p(\mu) =$$

$$\underset{\mu}{\operatorname{argmax}} \underbrace{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\beta}\, e^{-\frac{(\mu-\gamma)^2}{2\beta^2}}}_{A}$$

$$\frac{\partial A}{\partial \mu} = 0 \Rightarrow \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \times \frac{1}{\sqrt{2\pi}\beta} \times \left[e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}} \times \frac{+1}{\cancel{2}\sigma^2} \times \sum \cancel{2}(x_i-\mu) \times \right.$$

$$e^{-\frac{(\mu-\gamma)^2}{2\beta^2}} + e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}} \times e^{-\frac{(\mu-\gamma)^2}{2\beta^2}} \times \frac{-1}{2\beta^2} \times 2(\mu-\gamma) = 0$$

$$\Rightarrow$$

impulsado por CamScanner

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) - \frac{1}{\beta^2}(\mu-\gamma) = 0 \Rightarrow$$

$$\sum_{i=1}^{n} x_i - n\mu = \frac{\sigma^2}{\beta^2}\mu - \frac{\sigma^2}{\beta^2}\gamma \Rightarrow$$

$$\sum_{i=1}^{n} x_i + \frac{\sigma^2}{\beta^2}\gamma = \left(\frac{\sigma^2}{\beta^2}+n\right)\mu \Rightarrow$$

$$\boxed{\mu_{MAP} = \frac{\sum_{i=1}^{n} x_i + \frac{\sigma^2}{\beta^2}\gamma}{\left(\frac{\sigma^2}{\beta^2}+n\right)}}$$

$$\lim_{n\to\infty}\mu_{MAP} = \lim_{n\to\infty}\frac{\sum_{i=1}^{n} x_i}{n+\frac{\sigma^2}{\beta^2}} + \lim_{n\to\infty}\cancel{\frac{\frac{\sigma^2}{\beta^2}\gamma}{\frac{\sigma^2}{\beta^2}+n}}^{\,0} =$$

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow$$ این مقدار برابر همان تخمین MLE است که در بخش قبل ثابت کردیم که اگر $n\to\infty$، تخمین دقیق‌تری از μ به دست آورده و به آن میل می‌کنیم. پس اگر $n\to\infty$، μ_{MAP} به μ نزدیک‌تر می‌شود و اگر $n\to\infty$، μ_{MAP} دقیقاً برابر μ خواهد شد.

impulsado por CamScanner

الف

متغیر جدیدی به نام z به صورت زیر تعریف می‌کنیم:

$$z = x_1 + A x_2 \ , \ A = -\Sigma_{12}\Sigma_{22}^{-1}$$

ابتدا ثابت می‌کنیم که z و x_2 uncorrelated هستند سپس به ادامه اثبات، از این خاصیت استفاده می‌کنیم:

$$\mathrm{cov}(z, x_2) = \mathrm{cov}(x_1 + A x_2, x_2) = \mathrm{cov}(x_1, x_2) +$$

$$A \underbrace{\mathrm{cov}(x_2, x_2)}_{\mathrm{Var}(x_2)} = \Sigma_{12} - \Sigma_{12} \underbrace{\Sigma_{22}^{-1}\Sigma_{22}}_{I} = 0$$

(مشترکاً نرمال هستند)

x_1 و x_2 دارای تابع توزیع چندمتغیره نرمال هستند، پس مجموع آن‌ها از جمله $z = x_1 + A x_2$ نرمال است. z و x_2 نرمال و uncorrelated اند پس مستقل هم هستند.

(مشترکاً نرمال)

$$* \ E[x_1 | x_2] = E[z - A x_2 | x_2] = E[\overbrace{z | x_2}^{\text{مستقل اند}}] +$$

$$- E[A x_2 | x_2] = E[z] - A x_2 = E(x_1 + A x_2) - A x_2 =$$

$$\mu_1 + A\mu_2 - A x_2 = \mu_1 + A(\mu_2 - x_2) =$$

$$\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) = \boxed{\mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b)}$$

impulsado por CamScanner

$$* \quad Var(x_1|x_2) = Var(z - Ax_2|x_2) = Var(z|x_2) +$$

$$Var(\underbrace{-Ax_2|x_2}_{0}) + A\underbrace{cov(z,-x_2)}_{0} + \underbrace{cov(-x_2,z)}_{0}A^T =$$

$$Var(z|x_2) = var(z) = Var(x_1 + Ax_2) =$$

$$var(x_1) + A\,Var(x_2)A^T + A\underbrace{cov(x_1,x_2)}_{\Sigma_{12}} + \underbrace{cov(x_2,x_1)}_{\Sigma_{21}}A^T =$$

$$\Sigma_{11} + \Sigma_{12}\underbrace{\Sigma_{22}^{-1}\Sigma_{22}}_{I}\Sigma_{22}^{-1}\Sigma_{21} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12} +$$

$$-\Sigma_{21}\Sigma_{22}^{-1}\Sigma_{21} = \Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - 2\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12} =$$

$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \boxed{\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}}$$

* همچنین بدیهی است که این توزیع نرمال است زیرا:

$$P(x_a|x_b) = \frac{\overbrace{P(x_a,x_b)}^{\text{نرمال}}}{\underbrace{P(x_b)}_{\text{نرمال است زیرا}}} \Rightarrow$$ حاصل این عبارت نرمال است

x_a و x_b، مشترکاً نرمال هستند در نتیجه x_b هم نرمال خواهد بود.

impulsado por CamScanner

ب

X_a و X_b مشترکاً نرمال هستند پس در عبارت aX_a+bX_b اگر $a=1$ و $b=0$ ، X_a باید نرمال باشد. حال پارامترهای آن را به دست می‌آوریم:

تبدیل خطی S را به صورت زیر تعریف می‌کنیم

$$S = \begin{array}{c} \\ \\ \\ \\ \\ \end{array}\begin{array}{c} \overbrace{\hphantom{1\ 0\ \cdots\ 0}}^{(n+m-1)} \\ \left[\begin{array}{cccccccc} 1 & 0 & \cdots & & & & & 0 \\ 0 & 1 & 0 & \cdots & & & & 0 \\ 0 & 0 & 1 & 0 & \cdots & & & 0 \\ \vdots & & & & & & & \vdots \\ & & & \underbrace{1}_{n} & & & & \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{array}\right] \end{array} \Big\} n$$

پس:

$$S\,x = \begin{bmatrix} 1 & 0 & \cdots & & & 0 \\ 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & \cdots & 0 & 1 & 0 \cdots 0 \end{bmatrix}_{n\times(n+m)} \begin{bmatrix} x_a{}_{n\times1} \\ x_b{}_{m\times1} \end{bmatrix}_{(n+m)\times1} = x_a{}_{n\times1}$$

طبق خاصیت تأیید تبدیل‌های خطی میانگین μ و Σ ، میانگین و واریانس جدید تحت این تبدیل به صورت زیر خواهد بود:

$$\mu_{new} = S\,\mu = \begin{bmatrix} 1 & 0 & \cdots & & & 0 \\ 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & \cdots & 0 & 1 & 0 \cdots 0 \end{bmatrix}_{n\times(n+m)} \begin{bmatrix} \mu_a{}_{n\times1} \\ \mu_b{}_{m\times1} \end{bmatrix}_{(n+m)\times1} = \frac{\mu_a}{?}$$

$$\Sigma_{new} = S\,\Sigma\,S^T = \begin{bmatrix} 1 & 0 & \cdots & & & 0 \\ 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & \cdots & & 0 \\ \vdots & & & & & \\ 0 & 0 & \cdots & 0 & 1 & 0 \cdots 0 \end{bmatrix} \times \begin{bmatrix} \Sigma_{aa}{}_{n\times n} & \Sigma_{ab}{}_{n\times m} \\ \Sigma_{ba}{}_{m\times n} & \Sigma_{bb}{}_{m\times m} \end{bmatrix} \times$$

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ \vdots & \vdots & 1 & & \vdots \\ \vdots & \vdots & 0 & & \vdots \\ & & & & 0 \\ & & & & 1 \\ & & & & 0 \\ & & & & \vdots \\ 0 & 0 & 0 & & 0 \end{bmatrix}_{(n+m)\times n} = \begin{bmatrix} \Sigma_{aa}{}_{n\times n} & \Sigma_{ab}{}_{n\times m} \end{bmatrix} S^T = \frac{\Sigma_{aa}}{?}{}_{n\times n}$$

$$\Sigma_{aa}{}_{n\times n} \Longrightarrow$$

$$S\,x = x_a \Longrightarrow \boxed{E(x_a) = \mu_a \ , \quad \Sigma_{x_a} = \Sigma_{aa}}$$

$$E(x_a) \to S\mu \qquad \Sigma_{x_a} \to S\Sigma S^T$$

impulsado por CamScanner

۱۰

الف

حالت اول:

اگر ماتریس وجود داشته باشد که $Ax=b$ آنگاه $\|Ax-b\|=0$ و $x=\underbrace{A^{-1}}_{A^{+}}b$ پس حکم ثابت می‌شود.

حالت دوم:

اگر چنین جواب داشته باشیم، دنبال x ای هستیم که $\|Ax-b\|$ را کمینه کند.

$$\frac{\partial \|Ax-b\|}{\partial x} = \frac{\partial (Ax-b)^T(Ax-b)}{\partial x} =$$

$$\nabla f(g(x))^T \cdot \frac{\partial g}{\partial x} = (\nabla_z \underbrace{z^T z}_{z=Ax-b})^T \frac{\partial (Ax-b)}{\partial x} =$$

$$(2z^T)A = 2z^T A = 2(Ax-b)^T A = 0 \Rightarrow$$

$$A^T(Ax-b)=0 \Rightarrow A^TAx - A^Tb = 0 \Rightarrow$$

$$A^TAx = A^Tb \Rightarrow$$

$$x = \underbrace{(A^TA)^{-1}A^T}_{A^{+}}b \Rightarrow$$

$$\boxed{x = A^{+}b}$$

impulsado por CamScanner

حال ثابت می‌کنیم بین تمامی جواب‌ها $A^\dagger$ کمترین نرم b را دارد.

فرض می‌کنیم 2 جواب x' و $\underbrace{A^\dagger b}_{x^*}$ را داشته باشیم:

$$Ax' = b$$
$$A\underbrace{A^\dagger b}_{x^*} = b \quad \Rightarrow \quad A(x' - x^*) = 0 \Rightarrow$$

$$x' - x^* \in \text{nullspace}(A)$$

$$x^* = A^\dagger b = V\Sigma^{-1}U^T b$$

$$\|x'\|^2 = \|x^* - x^* + x'\|^2 = \|x^* - (x^* - x')\|^2 =$$

$$\|x^*\|^2 + \|x^* - x'\|^2 - 2\underbrace{x^{*T}(x^* - x')}$$

$$x^{*T}(x^* - x') = b^T U \Sigma^{-1} V^T (x^* - x')$$

V در فضای سطری A است و $(x^* - x')$ در فضای پوچی A است. می‌دانیم $\text{Col}(A)$ و $\text{nullspace}(A)$ بر هم عمودند پس $V^T(x^* - x') = 0$ و $x^{*T}(x^* - x') = 0$ پس:

$$\|x'\|^2 = \|x^*\|^2 + \underbrace{\|x^* - x'\|^2}_{\geq 0} \Rightarrow$$

پس $x^* = A^\dagger b$ کمترین نرم در بین تمام $\|x'\|^2 \geq \|x^*\|^2 \Rightarrow$ جواب‌ها را دارد.

impulsado por CamScanner

$$x^{(0)} = 0 \Rightarrow$$

$$x^{(1)} = 0 - \nu A^T(-b) = \nu A^T b$$

$$x^{(2)} = \nu A^T b - \nu A^T(\nu A A^T b - b) = \nu A^T b + \nu(I - \nu A^T A)A^T b$$

* $x^{(n)}$ در حالت کلی به صورت زیر خواهد بود:

$$x^{(n)} = \nu \sum_{i=0}^{n-1} (I - \nu A^T A)^i A^T b$$

متغیر جدید z را به صورت زیر تعریف می‌کنیم:

$$z^{(n)} = V^T x^{(n)} \quad , \quad A = U \Sigma V^T \Rightarrow$$

$$z^{(n)} = V^T \nu \sum_{i=0}^{n-1} (I - \nu A^T A)^i V \Sigma U^T b =$$

$$\nu \sum_{i=0}^{n-1} \underbrace{V^T (I - \nu A^T A)^i V}_{} \Sigma U^T b$$

ثابت می‌کنیم $V^T(I - \nu A^T A)^i V$ دارای $(I - \nu \Sigma^2)^i$ است.

استقرا:

پایه استقرا: $i = 0 \Rightarrow V^T V = I = (I - \nu \Sigma^2)^0 = I$ ✓

impulsado por CamScanner

حال فرض می‌کنیم برای n داشته باشیم:

$$V^T (I - \nu A^T A)^n V = (I - V \Sigma^2)^n$$

حکم را برای $n+1$ ثابت می‌کنیم:

$$V^T (I - \nu A^T A)^{n+1} V = \underbrace{\underbrace{V^T (I - \nu A^T A)^n V}_{(I - \nu \Sigma^2)^n} \underbrace{V^T (I - \nu A^T A) V}_{(I - \nu \Sigma^2)}}_{(I - \nu \Sigma^2)^{n+1}} =$$

پس $z^{(n)}$ به صورت زیر ساده می‌شود:

$$z^{(n)} = \nu \sum_{i=0}^{n-1} (I - \nu \Sigma^2)^i \Sigma u^T b$$

سری بالا همگرا می‌شود اگر مقادیر ویژه ماتریس $(I - \nu \Sigma^2)$ کمتر از یک باشند.

$A^\dagger$

$I - \nu \Sigma^2$ را قطری می‌کنیم:

$$I - \nu \Sigma^2 = P \lambda P^{-1} \rightarrow (I - \nu \Sigma^2)^n = P \lambda^n P^{-1}$$

پس برای همگرایی $(I - \nu \Sigma^2)^n$ باید $\lambda < 1$، یعنی تمامی مقادیر ویژه‌های $I - \nu \Sigma^2$ کمتر از یک باشند پس:

$$|\lambda_i (I - \nu \Sigma^2)| < 1 \Rightarrow |1 - \nu \sigma_i^2| < 1 \Rightarrow$$

$$-1 < 1 - \nu \sigma_i^2 < 1 \Rightarrow -2 < -\nu \sigma_i^2 < 0 \Rightarrow$$

impulsado por CamScanner

$$\Rightarrow \quad 0 < v\sigma_i^2 < 2 \quad \Rightarrow \quad 0 < v < 2/\sigma_i^2$$

برای پیدا کردن کمترین کران بالا باید $\sigma_i = \sigma_{max}(A)$ باشد پس :

$$\boxed{0 < v < 2/\sigma_{max}(A)^2}$$

$$\|A x^{(t+1)} - b\| = \| A x^{(t)} - vAA^T (Ax^{(t)} - b) - b \| =$$

$$\| (Ax^{(t)} - b) - vAA^T (Ax^{(t)} - b) \| =$$

$$\| (I - vAA^T)(Ax^{(t)} - b) \|$$

$$* \quad \underbrace{(I - vAA^T)}_{A'} \underbrace{(Ax^{(t)} - b)}_{x'} \Rightarrow$$

$$\|A'\| = \sigma_{max}(A') = \max \frac{\|A'x'\|}{\|x'\|} \Rightarrow$$

$$\|A x^{(t+1)} - b\| = \|(I - vAA^T)(Ax^{(t)} - b)\| \leq \sigma_{max}$$

$$\sigma_{max}(I - vAA^T) \|Ax^{(t)} - b\| =$$

$$[1 - v\sigma_{max}^2(A)] \|Ax^{(t)} - b\|$$

طبق شرط بالا صفحه که $0 < v < 2/\sigma_{max}^2(A)$ پس $|1 - v\sigma_{max}^2(A)| < 1$

درنتیجه :

$$\boxed{\|Ax^{(t+1)} - b\| \leq \|Ax^{(t)} - b\|}$$

impulsado por CamScanner