علیرضا دهقان پور فراشاه ۹۸۱۰۱۵۵۵

سؤال یک:

الف)

$$\det(\lambda I - A) = (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)\cdots(\lambda-\lambda_n)$$

$$= \lambda^n + c_{n-1}\lambda^{n-1} + \cdots c_1\lambda + c_0$$

$$c_{n-1} = -\sum_{i=1}^{n} \lambda_i$$

$$|A-\lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & & a_{nn}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) + Q(\lambda)$$

که در آن $Q(\lambda)$ حداکثر $n-2$ تا از عبارت های $(a_{ii}-\lambda)$ را دارد پس حداکثر دارای عبارات با درجه λ^{n-2} می باشد بنابراین ضریب $(-\lambda)^{n-1}$ در دترمینان $A-\lambda I$ می شود $\sum_{i=1}^{n} a_{ii}$ که همان $trace(A)$ است.

$$|\lambda I - A| = -|A-\lambda I| \Rightarrow -c_{n-1} = trace(A) \Rightarrow$$

$$\sum_{i=1}^{n} \lambda_i = trace(A)$$

ب) برای بخش ب اگر $\lambda=0$ را در چندجمله ای جایگذاری کنیم داریم

$$|A-\lambda I| = (\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda) \xrightarrow{\lambda=0} |A| = \lambda_1\lambda_2\cdots\lambda_n$$

ج) اثبات:

$$\lambda \in \sigma(AB) \Rightarrow \exists x \neq 0 \quad ABx = \lambda x \Rightarrow BA(Bx) = \lambda(Bx)$$

$$x \neq 0 \Rightarrow \lambda x \neq 0 \Rightarrow ABx = \lambda x \neq 0 \Rightarrow Bx \neq 0 \Rightarrow \lambda \in \sigma(BA) \Rightarrow \sigma(AB) \subset \sigma(BA)$$

به طریق مشابه برعکس هم ثابت می شود یعنی $\sigma(BA) \subset \sigma(AB)$ بنابراین $\sigma(AB) = \sigma(BA)$

د)

$$|A-\lambda I| = |(A-\lambda I)^T| = |A^T - \lambda I^T| = |A^T - \lambda I|$$

پس چندجمله ای مشخصه برای A^T و A یکسان است بنابراین مقادیر ویژه یکسانی دارند.

Scanned by CamScanner

سوال دو :

الف)

$$Var(E[X|Y]) = E[(E(X|Y))^2] - (E[E[X|Y]])^2$$

$$E[var(X|Y)] = E[E[X^2|Y] - (E[X|Y])^2] =$$

$$= E[E[X^2|Y]] - E[(E[X|Y])^2]$$

$$Var(E[X|Y]) + E[var(X|Y)] =$$

$$= \underline{E[(E[X|Y])^2]} - \underbrace{E[E[X|Y]]^2}_{E^2[X]} + \underbrace{E[E[X^2|Y]]}_{E[X^2]} - \underline{E[(E[X|Y])^2]} =$$

$$= E[X^2] - E^2[X] = Var(X)$$

ب)

$$Cov(X,Y|Z) = E[(X-E[X|Z])(Y-E[Y|Z])|Z] =$$

$$= E[XY - Y\underbrace{E[X|Z]}_{\text{تابع برحسب } Z} - X\underbrace{E[Y|Z]}_{\text{تابع برحسب } Z} + \underbrace{E[X|Z]\,E[Y|Z]}_{\text{تابع برحسب } Z} | Z] =$$

$$= E[XY|Z] - E[Y\underbrace{E[X|Z]}_{\text{ثابت}}|Z] - E[X\underbrace{E[Y|Z]}_{\text{ثابت}}|Z] + E[\underbrace{E[X|Z]\,E[Y|Z]}_{\text{ثابت}}|Z] \overset{\text{لم}}{\overset{*}{=}}$$

$$= E[XY|Z] - E[X|Z]\,E[Y|Z] - E[Y|Z]\,E[X|Z] + E[X|Z]\,E[Y|Z] =$$

$$= E[XY|Z] - E[X|Z]\,E[Y|Z]$$

ج)

① $E[Cov(X,Y|Z)] \underset{\text{رابطه بالا}}{=} E[E[XY|Z] - E[X|Z]\,E[Y|Z]] =$

$$= E[XY] - E[E[X|Z]\,E[Y|Z]]$$

② $Cov(E[X|Z], E[Y|Z]) = E[E[X|Z]\,E[Y|Z]] - E[E[X|Z]]\,E[E[Y|Z]]$

$$= E[E[X|Z]\,E[Y|Z]] - E[X]\,E[Y]$$

$\Rightarrow$ ①+② $= E[XY] - E[E[X|Z]\,E[Y|Z]] + E[E[X|Z]\,E[Y|Z]] - E[X]\,E[Y] =$

$$= E[XY] - E[X]\,E[Y] = Cov(X,Y)$$

توجه کنید که در بخش ب تساوی * به وسیله لم صفحه بعد اثبات شده است.

Scanned by CamScanner

لم :

$$E[E[X|Z]Y|Z] = E[X|Z]\,E[Y|Z]$$

اثبات :

$$\int E[X|Z]\,y f_{Y|Z}(y)\,dy = \int y f_{Y|Z}(y) \int x f_{X|Z}(x)\,dx\,dy$$

$$= \int y f_{Y|Z}(y)\,dy \times \int x f_{X|Z}(x)\,dx = E[Y|Z]\,E[X|Z]$$

لم :

$$E[E[X|Z]\,E[Y|Z]\,|\,Z] = E[X|Z]\,E[Y|Z]$$

اثبات :

طبق لم بالا

$$E[E[X|Z]\,\underbrace{E[Y|Z]}_{T}\,|\,Z] = E[X|Z]\,E[T|Z] =$$

$$= E[X|Z]\,E[\underbrace{E[Y|Z]}_{\text{تابعی برحسب } Z}\,|\,Z] = E[X|Z]\,E[Y|Z]$$

Scanned by CamScanner

سؤال سه :

الف) توجه کنید طبق صورت سؤال جواب را بردار سطری اعلام کرده است.

$$x^TAx = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} x_j$$

$$x^TAx = \sum_{\substack{i=1 \\ i\neq k}}^{n} \sum_{\substack{j=1 \\ j\neq k}}^{n} x_i a_{ij} x_j + x_k^2 a_{kk} + \sum_{\substack{j=1 \\ j\neq k}}^{n} x_k a_{kj} x_j + \sum_{\substack{i=1 \\ i\neq k}}^{n} x_i a_{ik} x_k$$

$$\frac{\partial x^TAx}{\partial x_k} = 2x_k a_{kk} + \sum_{\substack{j=1 \\ j\neq k}}^{n} a_{kj} x_j + \sum_{\substack{i=1 \\ i\neq k}}^{n} x_i a_{ik} = \sum_{i=1}^{n} x_i a_{ik} + \sum_{j=1}^{n} x_j a_{kj}$$

$$= x^T \underbrace{a_{:k}}_{\text{ستون kام}} + x^T \underbrace{a_{k:}}_{\text{سطر kام}} \Rightarrow \frac{\partial x^TAx}{\partial x} = x^TA + x^TA^T \overset{\text{متقارن } A}{=} 2x^TA$$

ب) می‌دانیم در X^TAX درایه ستون iام و سطر iام از $x_i^TAx_i$ بدست می‌آید که x_i ستون iام از ماتریس X است.

$$\mathrm{trace}(X^TAX) = \sum_{i=1}^{n} x_i^T A x_i$$

$$\frac{\partial\, \mathrm{trace}(X^TAX)}{\partial x_{ij}} = \frac{\partial\, x_j^T A x_j}{\partial x_{ij}} \overset{\text{بخش قبل}}{=} \sum_{m=1}^{n} x_{mj} a_{mi} + \sum_{l=1}^{n} x_{lj} a_{il}$$

$$= x_j^T \underbrace{a_{i:}}_{\text{سطر iام}} + x_j^T \underbrace{a_{:i}}_{\text{ستون iام}} = \text{درایه ijام ماتریس مشتق} \Rightarrow$$

$$\frac{\partial\, \mathrm{trace}(X^TAX)}{\partial \underbrace{x_i}_{\text{ستون iام}}} = x_i^TA + x_i^TA^T \Longrightarrow$$

$$\frac{\partial\, \mathrm{trace}(X^TAX)}{\partial X} = \begin{bmatrix} \frac{\partial x_1^TAx_1}{\partial x_1} \\ \vdots \\ \frac{\partial x_n^TAx_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} x_1^T(A+A^T) \\ \vdots \\ x_n^T(A+A^T) \end{bmatrix} = X^T(A+A^T)$$

توجه کنید که اگر X یک بردار باشد x^TAx یک عدد خواهد بود و trace آن خودش است و دقیقاً همان الف می‌شود با این تفاوت که A متقارن نیست.

Scanned by CamScanner

سوال چهار:

$Y = X^2$

برای حل این سوال از قضیه بنیادی آمار و احتمال استفاده می‌کنیم:

$y = g(x) = x^2 , \quad g'(x) = 2x$

$x = \pm\sqrt{y} \implies f_Y(y) = \dfrac{f_X(\sqrt{y})}{|2\sqrt{y}|} + \dfrac{f_X(-\sqrt{y})}{|-2\sqrt{y}|} = \dfrac{1}{2\sqrt{y}} \quad y \in (0,1]$

$$\implies f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}} & 0 < y \leq 1 \\ 0 & \text{o.w} \end{cases}$$

$Y = \sqrt{X} , \quad g(x) = \sqrt{x} = y \implies \overset{*}{y > 0}, \; x = y^2$

$g'(x) = \dfrac{1}{2\sqrt{x}}$

$\implies f_Y(y) = \dfrac{f_X(y^2)}{\dfrac{1}{2\sqrt{y^2}}} = \dfrac{1}{\dfrac{1}{2|y|}} = 2|y| \overset{*}{=} 2y$

$$\implies f_Y(y) = \begin{cases} 2y & 0 \leq y \leq 1 \\ 0 & \text{o.w} \end{cases}$$

سوال پنج:

$\lambda \in \sigma(M) \implies Mx = \lambda x \implies P^{-1}MP(P^{-1}x) = P^{-1}(Mx) = P^{-1}(\lambda x) = \lambda(P^{-1}x)$

P فول رنک پس P^{-1} نیز فول رنک است پس $P^{-1}x \neq 0$ است.

$\implies \lambda \in \sigma(P^{-1}MP)$

$\lambda' \in \sigma(P^{-1}MP) \implies P^{-1}MPx = \lambda' x \implies M(Px) = \lambda'(Px) \implies \lambda' \in \sigma(M)$

چون P فول رنک است

$Px \neq 0$ است.

بنابراین این مجموعه مقادیر ویژه یکسان است.

Scanned by CamScanner

الف)

سؤال شش: چون ماتریس متقارن است یعنی $A = A^T$ و تمامی مقادیر ویژه مثبت است آن‌گاه قطری پذیر است که P بردار ویژه‌ها هستند و D مقادیر ویژه که چون A معین مثبت و متقارن است $\Rightarrow A = P^{-1}DP$

بردارهای ویژه متعامد نرمال هستند یعنی $P^T = P^{-1}$ پس داریم:

$$A = P^{-1}DP = P^TDP = P^TD^{1/2} \times D^{1/2}P = (D^{1/2}P)^T(D^{1/2}P)$$

چون P و $D^{\frac{1}{2}}$ هر دو وارون پذیر هستند $D^{\frac{1}{2}}P$ نیز وارون پذیر است بنابراین دارای تجزیه یکتای QR است که R ماتریس بالامثلثی با قطر مثبت است. (اگر صفر باشد یعنی A وارون ناپذیر است که تناقض است)

$$A = (QR)^T(QR) = R^T\underbrace{Q^TQ}_{I}R = R^TR$$

که R^T ماتریس پایین مثلثی با قطر مثبت است.

پس $L = R^T$ داریم $A = LL^T$

اثبات یکتایی:

$$A = L_1L_1^T = L_2L_2^T \Rightarrow (L_2^{-1}L_1)(L_1^TL_2^{-T}) = I$$

$$\Rightarrow \underline{L_2^{-1}L_1} = L_2^TL_1^{-T} = (\underline{L_2^{-1}L_1})^T$$

L_2^{-1} پایین مثلثی و L_1 پایین مثلثی است پس سمت چپ پایین مثلثی و سمت راست بالامثلثی است پس $L_2^{-1}L_1$ ماتریس قطری است که و $(L_2^{-1}L_1)(L_2^{-1}L_1)^T = (L_2^{-1}L_1)^2 = I$ یعنی $L_2^{-1}L_1$ ماتریس قطری با قطرهای ± 1 است و چون L_1 و L_2 هر دو قطر مثبت دارند و قطر L_2^{-1} معکوس قطر L_2 است یعنی قطر L_2^{-1} نیز مثبت است و L_2^{-1} و L_1 هر دو پایین مثلثی و قطر $L_2^{-1}L_1$ حاصل ضرب دو قطر است پس قطر $L_2^{-1}L_1$ نیز مثبت است یعنی $L_2^{-1}L_1 = I$.

ب)

$$A = QR = U\Sigma V^T \Rightarrow R = Q^TU\Sigma V^T$$

Q ماتریس متعامد نرمال و

U نیز نیز ماتریس یکه است و

Q^TU نیز یکه است زیرا:

$$(Q^TU)(Q^TU)^T = Q^TUU^TQ = Q^TQ = I$$

پس تجزیه SVD ماتریس R شبیه A است فقط به جای ماتریس U ماتریس Q^TU قرار دارد.

$$R = U'\Sigma V^T \quad , \quad U' = Q^TU$$

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

Scanned by CamScanner

سؤال هفت:

فرض $A^K = 0$

$$(I-A)\underbrace{(I+A+A^2+\dots A^{K-1})}_{B} = I - A + A - A^2 + A^2, \dots - A^{K-1} + A^{K-1} - A^K =$$

$$= I - \underbrace{A^K}_{0} \overset{\text{فرض}}{=} I \Rightarrow (I-A)^{-1} = (I+A+\dots A^{K-1})$$

حال اگر B را از سمت چپ نیز ضرب کنیم خواهیم:

$$(I+A+\dots A^{K-1})(I-A) = I - A + A - A^2 + \dots - A^{K-1} + A^{K-1} - A^K = I - \underbrace{A^K}_{0} = I$$

پس داریم $B(I-A) = (I-A)B = I$ بنابراین طبق تعریف داریم $BA = AB = I$ پس

$B = A^{-1}$

Scanned by CamScanner

سؤال هفت : الف)

$$F_{M,X,Y}(x,y,t) = P[X\le x, Y\le y, M\le t] \overset{\text{بیز}}{=} P[X\le x, Y\le y \mid M\le t] \times \overbrace{P[M\le t]}^{U(0,1)}$$

$0\le t\le 1$

$$= P[X\le x, Y\le y \mid M\le t]\times t \overset{X,Y\ \text{استقلال}}{=} P[X\le x\mid M\le t]\times P[Y\le y\mid M\le t]\times t$$

$$= F_X(x)\times F_Y(y)\times t$$

$$f_{X,Y,M}(x,y,t) = \frac{\partial\, F_X(x)\, F_Y(y)\times t}{\partial x\,\partial y\,\partial t} = f_X(x)\times f_Y(y) \Longrightarrow$$

$$f_{X,Y,M}(x,y,t) = \begin{cases} \frac{1}{2\pi} e^{-\left(\frac{(x-M)^2}{2}+\frac{(y-M)^2}{2}\right)} & 0\le t\le 1 \\ 0 & \text{o.t} \end{cases}$$

ب)

$$\max_M P(M\mid D) = \frac{P(D\mid M)\times P(M)}{\underbrace{P(D)}_{\text{ثابت}}} \sim \max_M P(D\mid M)\times P(M)$$

$$0\le M\le 1 :\ P(M)=1 \Rightarrow \max_M P(D\mid M) = \max_M \prod_{i=1}^{n} P(D_i\mid M)$$

$$\log P(D\mid M) = \sum_{i=1}^{n} \log P(D_i\mid M) = \sum_{i=1}^{n} -\frac{(x_i-M)^2}{2} - \frac{(y_i-M)^2}{2} + \log\frac{1}{2\pi}$$

$$\frac{d}{dM}(\log P(D\mid M)) = \sum_{i=1}^{n} +(x_i-M)+(y_i-M) = 0 \Rightarrow M = \frac{\sum x_i+y_i}{2n}$$

پس اگر $\frac{\sum x_i+y_i}{2n}$ بین $[0,1]$ باشد جواب مسئله است در غیر این صورت جواب 0 یا 1 است که باید بررسی شود.

Scanned by CamScanner