Subject:

Year. Month. Date. ( ) ۹۹۱۰۵۲۹۴

محمدجواد باطنی تمرین ۱ یادگیری ماشین

سوال ۱) الف)

$x=\begin{bmatrix}x_1\\ \vdots \\ x_n\end{bmatrix}$ $A=[A_1\ A_2\ \dots\ A_n] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ \vdots & & & \\ a_{n1} & \dots & & a_{nn}\end{bmatrix}$

می‌خواهیم اثبات کنیم

$$\frac{\partial x^TAx}{\partial x} = x^T(A+A^T)$$

روش ۱) از تعریف مشتق استفاده می‌کنیم

$$\frac{\partial x^TAx}{\partial x} = \lim_{h\to 0}\frac{(x+h)^TA(x+h)-x^TAx}{h}$$

$$=\lim_{h\to 0} \frac{x^TAh+h^TAx+h^TAh}{h} = \lim_{h\to 0} \frac{x^TAh}{h}+\lim_{h\to 0}\frac{h^TAx}{h}+\lim_{h\to 0}\frac{h^TAh}{h}$$

$$= x^TA+(Ax)^T+0 = x^T(A+A^T) \ (*) \overset{if\ A=A^T}{=} 2x^TA$$

(*) توجه داشته باشید تنها اگر $A$ متقارن باشد می‌توان گفت که حاصل فوق برابر $2x^TA$ است لذا صورت این سوال احتمالاً ناقص است

روش ۲)

$$x^TAx = [x^TA_1\ \ x^TA_2\ \dots\ x^TA_n]\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} = x^TA_1x_1+\dots+x^TA_nx_n$$

$$\Rightarrow \frac{\partial x^TAx}{\partial x} = \left[\frac{\partial x^TAx}{\partial x_1}\ \ \frac{\partial x^TAx}{\partial x_2}\ \dots\ \frac{\partial x^TAx}{\partial x_n}\right]$$

$$\frac{\partial x^TAx}{\partial x_i} = a_{i1}x_1 + a_{in}x_n + x^TA_i = x_1(a_{i1}+a_{1i})+x_2(a_{i2}+a_{2i})+\dots+x_n(a_{in}+a_{ni})$$

$$= x^T(A+A^T)_{(i)} \Rightarrow$$ عنصر iام $x^T(A+A^T)$

$$\Rightarrow \frac{\partial x^TAx}{\partial x} = x^T(A+A^T) \overset{if\ A=A^T}{=} 2x^TA$$

ب) توجه داریم که $x^TAx$ یک عدد ($\mathbb{R}^1$) خواهد شد چرا که: $\underset{1\times n}{x^T}\ \underset{n\times n}{A}\ \underset{n\times 1}{x} \Rightarrow 1\times 1$

لذا $trace(x^TAx)=x^TAx$ و می‌توان گفت جواب این قسمت نیز همان جواب قسمت قبل یعنی

$$x^T(A+A^T) \underset{if\ A=A^T}{=} 2x^TA$$

است

سوال ۲ (۱.۰) الف) روش ۱)

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

$$\lambda_1 + \cdots + \lambda_n = trace(A)$$

$$Av = \lambda v \Rightarrow (A - \lambda I)v = 0 \Rightarrow \det(A - \lambda I) = 0$$

چند جمله‌ای مشخصه مربوط به مقادیر ویژه :

$$(-1)^n (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n) = 0 \quad \text{I}$$

$$\det(A - \lambda I) \overset{\lambda = x}{=} \det \begin{bmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ \vdots & a_{22} - x & & \\ & & \ddots & \\ a_{n1} & & & a_{nn} - x \end{bmatrix} = (a_{11} - x) \cdots (a_{nn} - x) + g(x^{n-2})$$

تابعی با درجه n-۲ برای $x$

$$= (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) x^{n-1} + g'(x^{n-2}) \quad \text{II}$$

$$\text{I} \cap \text{II} \Rightarrow (-1)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) x^{n-1} = (-1)^{n-1}(a_{11} + \cdots + a_{nn}) x^{n-1}$$

$$\Rightarrow \lambda_1 + \cdots + \lambda_n = trace(A)$$

روش دوم) الف) تجزیه Schur decomposition می‌توان ماتریس A را به صورت $QUQ^{-1}$ نوشت که Q یک ماتریس unitary و U یک ماتریس بالامثلثی است که روی قطر اصلی آن مقادیر ویژه قرار دارند

$$\Rightarrow A = QUQ^{-1}$$

$$\Rightarrow trace(A) = trace(QUQ^{-1}) = trace(QQ^{-1}U) = trace(U) = \sum_{i=1}^{n} \lambda_i$$

$$\Rightarrow \lambda_1 + \cdots + \lambda_n = trace(A)$$

ب) روش ۱) مشابه قسمت قبل با استفاده از Schur decomposition خواهیم داشت

$$\det(A) = \det(QUQ^{-1}) = \det(QQ^{-1}U) = \det(U)$$

$$= \det \begin{bmatrix} \lambda_1 & & \\ 0 & \lambda_2 & \\ \vdots & & \ddots \\ 0 & & \lambda_n \end{bmatrix} = \lambda_1 \lambda_2 \cdots \lambda_n \Rightarrow \det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

روش ۲) مطابق قسمت قبل برای معادله مشخصه مقادیر ویژه داریم

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

اگر در معادله فوق $\lambda = 0$ را جایگذاری کنیم داریم:

$$\overset{\lambda = 0}{\Rightarrow} \det A = \lambda_1 \lambda_2 \cdots \lambda_n$$

سوال ۱۰۳) الف)

$$f_{\mu XY}(\mu=t, X=x, Y=y) = f_{\mu XY}(X=x, Y=y \mid \mu=t)\, f_{\mu}(\mu=t)$$

X,Y مستقل‌اند

$$= f_{\mu X}(X=x \mid \mu=t) \times f_{\mu Y}(Y=y \mid \mu=t) \times f_{\mu}(\mu=t)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-t)^2}{2}} \times f_{\mu}(\mu=t)$$

$$\Rightarrow f_{\mu XY}(\mu=t, X=x, Y=y) = \begin{cases} \frac{1}{2\pi} e^{-\left(\frac{(x-t)^2+(y-t)^2}{2}\right)} & t\in[0,1] \\ 0 & \text{else where} \end{cases}$$

ب)

$$\hat{\mu}_{MAP} = \arg\max_{\mu} f(\mu \mid X, Y) = \arg\max_{\mu} f(X, Y \mid \mu) f(\mu)$$

$$X=\{X_1, \dots, X_n\}$$
$$Y=\{Y_1, \dots, Y_n\}$$

$X_i, Y_i$ مستقل‌اند

$$\Rightarrow \hat{\mu}_{MAP} = \arg\max_{\mu} \underbrace{f(X\mid\mu)\, f(Y\mid\mu)\, f(\mu)}_{A}$$

$$A = \left(\frac{1}{\sqrt{2\pi}}\right)^{2n} \times e^{-\frac{(X_1-\mu)^2+\dots+(X_n-\mu)^2+(Y_1-\mu)^2+\dots+(Y_n-\mu)^2}{2}} \times f(\mu)$$

if $0<\mu<1$ (I)

$$A \propto e^{-\frac{(X_1-\mu)^2+\dots+(X_n-\mu)^2+(Y_1-\mu)^2+\dots+(Y_n-\mu)^2}{2}}$$

$$\frac{dA}{d\mu} = \left((X_1-\mu)+\dots+(X_n-\mu)+(Y_1-\mu)+\dots+(Y_n-\mu)\right) e^{-\frac{(X_1-\mu)^2+\dots+(Y_n-\mu)^2}{2}}$$

$$\frac{dA}{d\mu}=0 \Rightarrow \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i - 2n\hat{\mu} = 0 \Rightarrow \hat{\mu} = \frac{\sum X_i + \sum Y_i}{2n}$$

توجه داشته باشید طبق (I) باید $0<\hat{\mu}<1$ که تا حاصل $f(\mu\mid X,Y)$ صفر نشود لذا شرط زیر را برقرار باشد

$$0 < \sum X_i + \sum Y_i < 2n$$

$\hat{\mu} = \frac{\sum X_i + \sum Y_i}{2n}$ تخمین MAP $\mu$ است.

از طرفی $\hat{\mu}$ باید بین صفر و یک باشد چرا که خارج این بازه بودن این متغیر احتمال صفر دارد لذا داریم:

$$\mu=[0,1] \quad \begin{cases} \hat{\mu} = \frac{\sum X_i + \sum Y_i}{2n} & \text{if } 0<\sum X_i+\sum Y_i<2n \\ \hat{\mu}=0 & \text{else if } \sum X_i+\sum Y_i<0 \\ \hat{\mu}=1 & \text{else if } \sum X_i+\sum Y_i>2n \end{cases}$$

Subject:

Year. Month. Date. ( )

(۱.۴ سوال

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}| = \max(۸, ۷, ۵) = ۸$$

$$\|A\|_F = \sqrt{\sum_j \sum_i |a_{ij}|^۲} = \sqrt{۲۵+۱۶+۴+۱+۴+۹+۴+۱+۰} = ۸$$

$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}| = \max(۱۱, ۶, ۳) = ۱۱$$

$$\|A\|_۲ = \sqrt{\max(\text{eigenvalue}(A^T A))}$$

$$A^T A = \begin{bmatrix} ۳۰ & -۲۴ & ۷ \\ -۲۴ & ۲۱ & -۲ \\ ۷ & -۲ & ۱۳ \end{bmatrix} \Rightarrow \det(A^T A - \lambda I) = ۰ \Rightarrow -\lambda^۳ + ۶۴\lambda^۲ - ۶۶۴\lambda + ۲۲۵ = ۰$$

$$\Rightarrow \lambda_1 \simeq ۰,۳۵ \quad \lambda_۲ \simeq ۱۲,۵۶ \quad \lambda_۳ \simeq ۵۱,۰۹$$

$$\Rightarrow \|A\|_۲ = \sqrt{۵۱,۰۹} \simeq \underline{۷,۱۴}$$

سوال ۲۰۱)

$$E(X^3+2X-7) \overset{LOTUS}{=} \int_{-\infty}^{\infty} (x^3+2x-7)\, f_X(x)\, dx$$

$$= \int_1^2 (x^3+2x-7)(2x-2)\, dx = \int_1^2 2x^4-2x^3+4x^2-4x-14x+14\, dx$$

$$\left(\frac{2}{5}x^5 - \frac{x^4}{2} + \frac{4x^3}{3} - 9x^2 + 14x\right)\Big|_1^2 = \left(\frac{64}{5} - 8 + \frac{32}{3} - 36 + 28\right) - \left(\frac{2}{5} - \frac{1}{2} + \frac{4}{3} - 9 + 14\right)$$

$$= \frac{62}{5} - 21 + \frac{28}{3} + \frac{1}{2} = \underline{\frac{37}{30}}$$

سوال ۲۰۲)

$$Y_1 = \max(X_1, \dots, X_n)$$

$$F_{Y_1}(y) = f_{Y_1}(Y_1 \leq y) = f_{X_1X_2\dots X_n}(\max(X_1, \dots, X_n) \leq y) = f_{X_1\dots X_n}(X_1 \leq y, \dots, X_n \leq y)$$

$X_i$s are independent

$$= f_{X_1}(X_1 \leq y)\, f_{X_2}(X_2 \leq y) \dots f_{X_n}(X_n \leq y) = F_{X_1}(y) \dots F_{X_n}(y)$$

$$\Rightarrow f_{Y_1}(Y_1 = y) = \frac{\partial F_{Y_1}(y)}{\partial y} = \frac{\partial\left(F_{X_1}(y) \times F_{X_2}(y) \times \dots \times F_{X_n}(y)\right)}{\partial y}$$

$$= \underline{n\, f_X(X=y)\, F_X^{n-1}(y)}$$

$$Y_2 = \min(X_1, \dots, X_n) \Rightarrow F_{Y_2}(y) = f_{Y_2}(Y_2 \leq y) = 1 - f_{Y_2}(Y_2 > y)$$

$$= 1 - f_{X_1\dots X_n}(X_1 > y, X_2 > y, \dots, X_n > y) \overset{independent}{=} 1 - \left(f_{X_1}(X_1 > y) \times \dots \times f_{X_n}(X_n > y)\right)$$

$$= 1 - \left((1 - F_{X_1}(y)) \times (1 - F_{X_2}(y)) \times \dots \times (1 - F_{X_n}(y))\right)$$

$$\Rightarrow f_{Y_2}(Y_2=y) = \frac{\partial F_{Y_2}(y)}{\partial y} = \frac{\partial\left(1-\left((1-F_{X_1}(y))\times\cdots\times(1-F_{X_n}(y))\right)\right)}{\partial y}$$

$$= n\, f_X(x=y)\,(1-F_X(y))^{n-1}$$

سوال ۲.۳) فرض می‌کنیم $\lambda$ و $v$ به ترتیب مقدار و بردار ویژه‌های $P^{-1}MP$ باشند

$$\Rightarrow P^{-1}MPv = \lambda v$$

توجه داریم که اگر $P$ فول رنک باشد می‌توان گفت $P$ معکوس پذیر است، لذا $P^{-1}$ وجود دارد.

$$\times P \Rightarrow MPv = \lambda Pv$$

$Pv$ را به صورت یک بردار $v'$ در نظر بگیریم خواهیم داشت

$$Mv' = \lambda v'$$

لذا می‌توان گفت هر مقدار ویژه $P^{-1}MP$ مقدار ویژه $M$ نیز است.

اکنون اگر اثبات کنیم هر مقدار ویژه $M$ مقدار ویژه $P^{-1}MP$ است اثبات سوال تمام می‌شود

فرض می‌کنیم $\lambda$ و $v$ مقدار و بردار ویژه $M$ هستند

$$Mv = \lambda v$$

می‌توان گفت از آن جایی که $P$ معکوس پذیر است داریم:

$$MPP^{-1}v = \lambda v$$

$$\times P^{-1} \Rightarrow P^{-1}MP\underbrace{P^{-1}v}_{v'} = \lambda \underbrace{P^{-1}v}_{v'} \Rightarrow P^{-1}MPv' = \lambda v'$$

$\Leftarrow$ هر مقدار ویژه $M$ مقدار ویژه $P^{-1}MP$ نیز است.

لذا در مجموع می‌توان گفت $P^{-1}MP$ و $M$ یک مجموعه مقدار ویژه یکسان دارند.