Subject:
Date

$P(Y=1) = P(Y=2) = P(Y=3) = \frac{1}{3}$ $\quad 1 \leq i \leq 3 \cdot (X|Y=i) \sim N(\mu_i, \Sigma_i)$. 1

$\mu_1 = [0,0]^T \; , \mu_2 = [1,1]^T \; , \mu_3 = [1,1]^T$

$\Sigma_1 = \begin{bmatrix} 0.7 & 0 \\ 0 & 0.7 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 0.8 & 0.3 \\ 0.3 & 0.2 \end{bmatrix}, \Sigma_3 = \begin{bmatrix} 0.7 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$

$$\Rightarrow (Y|X=x) = \underset{i}{\operatorname{argmax}} \; N(x|\mu_i, \Sigma_i)$$

$$= \underset{i}{\operatorname{argmax}} \; (2\pi)^{-k/2} |\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu_i) \Sigma_i^{-1} (x-\mu_i)\right)$$

$$= \underset{i}{\operatorname{argmin}} \; \lg|\Sigma_i| + (x-\mu_i)\Sigma_i^{-1}(x-\mu_i)$$

99105901

$$= \underset{i}{\operatorname{argmin}} \begin{cases} -0.71 + x^T \begin{bmatrix} 1.42 & 0 \\ 0 & 1.42 \end{bmatrix} x, \\ -2.65 + (x-1)^T \begin{bmatrix} 2.85 & -4.28 \\ -4.28 & 11.42 \end{bmatrix} (x-1), \\ -0.51 + (x-1)^T \begin{bmatrix} 1.53 & -0.38 \\ -0.38 & 1.34 \end{bmatrix} (x-1) \end{cases}$$

$x = [50, 0.5] \Rightarrow (Y|X=x) = 1$

$x = [0.5, 0.5] \Rightarrow (Y|X=x) = 2$

Subject:
Date

$$E_D(\omega) = \frac{1}{2}\sum_{n=1}^{N}\left(y(x_n,\omega) - y_n\right)^2 \quad -2$$

$$\tilde{E}_D(\omega) = \frac{1}{2}\sum_{n=1}^{N} E\left\{ y_n^2 + y(x_n+\epsilon_n, \omega)^2 - 2y_n\, y(x_n+\epsilon, \omega)\right\}$$

$$= \frac{1}{2}\sum_{n=1}^{N}\left(y_n^2 - 2y_n E\{y(x_n+\epsilon_n,\omega)\} + E\{y(x_n+\epsilon_n,\omega)^2\}\right)$$

$$E\{y(x_n+\epsilon_n,\omega)\} = \omega_0 + \sum_{i=1}^{D}\omega_i E\{x_{ni}+\epsilon_{ni}\}$$

$$= \omega_0 + \sum_{i=1}^{D}\omega_{i} x_{ni} = y(x_n,\omega)$$

$$E\{y(x_n+\epsilon_n,\omega)^2\} = \omega_0^2 + E\left\{\left(\sum_{i=1}^{D}\omega_i(x_{ni}+\epsilon_{ni})\right)^2\right\} + 2\omega_0\sum_{i=1}^{D}\omega_i E\{x_{ni}+\epsilon_{ni}\}$$

$$= \omega_0^2 + 2\omega_0\sum_{i=1}^{D}\omega_{i} x_{ni} + E\left\{\left(\sum_{i}^{D}\omega_i x_{ni} + \sum_{i}^{D}\omega_i\epsilon_{ni}\right)^2\right\}$$

$$E\left\{\Big(\underbrace{\sum_{i=1}^{D}\omega_i x_{ni}}_{a} + \underbrace{\sum_{i=1}^{D}\omega_i \epsilon_{ni}}_{b}\Big)^2\right\} = \underbrace{E\{a^2\}}_{\text{Const}} + E\{b^2\} + E\{2ab\}$$

$$= a^2 + E\left\{\left(\sum_{i=1}^{D}\omega_i\epsilon_{ni}\right)^2\right\} + E\left\{2\sum_{i=1}^{D}\sum_{j=1}^{D}\omega_i\omega_j x_{ni}\epsilon_{nj}\right\}$$

↳ each one is zero so entire would be zero

$$\Rightarrow E\left\{\left(\sum\omega_i x_{ni} + \sum\omega_i\epsilon_{ni}\right)^2\right\}$$

$$\xrightarrow{\epsilon_{ni}\text{ i.i.d}} = \left(\sum_{i=1}^{D}\omega_i x_{ni}\right)^2 + \sum_{i=1}^{D}\omega_i^2 E\{\epsilon_{ni}^2\} = \sigma_i^2\sum\omega_i^2$$

by assembling all the parts, we only get one more residual part: $\sigma_i^2\sum_{i=1}^{D}\omega_i^2$

So: $$\tilde{E}_D(\omega) = E_D(\omega) + \sigma_i^2\sum_{i=1}^{D}\omega_i^2$$

Subject:
Date

$$BLR:\quad P(i)=\frac{e^{w_i x+b}}{1+e^{w_i x+b}}$$

-3

الف)

One v one using k-1 BLRs.

$$\hookrightarrow \log\left(\frac{P(k)}{P(K)}\right)=w_i x$$

$$\hookrightarrow P(k)=e^{w_i x}\,p(K)$$

we know $P(K)=1-\sum_{i=1}^{K-1}p(i)=1-\sum_{i=1}^{K-1}P(K)e^{w_i x}$

$w_K=0$
$\hookrightarrow e^{w_K x}=1$

$$\Rightarrow P(K)=\frac{1}{1+\sum_{j=1}^{K-1}e^{w_j x}}\rightarrow P(k)=\frac{e^{w_i x}}{1+\sum_{j=1}^{K-1}e^{w_j x}}$$

$$\hookrightarrow\ \Rightarrow P(y=i\,|\,x)=\frac{e^{w_i x}}{\sum_{j=1}^{K}e^{w_j x}}$$

ب)

$$\ln P(Y=y_i\,|\,x=x_i)=w_{y_i}x_i-\ln\sum_{j=1}^{K}e^{w_j x_i}$$

$$L(w)=\sum_{i=1}^{n}w_{y_i}x_i-\sum_{i=1}^{n}\ln\left(\sum_{j=1}^{K}e^{w_j x_i}\right)$$

پ)

$$\frac{\partial}{\partial w_k}L(w)=\frac{\partial}{\partial w_k}\sum_{i=1}^{n}w_{y_i}x_i-\frac{\partial}{\partial w_k}\sum_{i=1}^{n}\ln\left(\sum_{j=1}^{K}e^{w_j x_i}\right)$$

$$\frac{\partial}{\partial w_k}\sum_{i=1}^{n}w_{y_i}x_i=\sum_{i=1}^{n}x_i\cdot I(y_i=k)$$

$$\frac{\partial}{\partial w_k}\sum_{i=1}^{n}\ln\left(\sum_{j=1}^{K}e^{w_j x_i}\right)=\sum_{i=1}^{n}\frac{\partial}{\partial w_k}\ln\left(\sum_{j=1}^{k}e^{w_j x_i}\right)$$

$$=\sum_{i=1}^{n}\frac{1}{\sum_{j=1}^{K}e^{w_j x_i}}\cdot x_i\cdot e^{w_k x_i}=\sum_{i=1}^{n}\frac{x_i e^{w_k x_i}}{\sum_{j=1}^{K}e^{w_j x_i}}$$

Subject:
Date

$$\Rightarrow \frac{\partial}{\partial \omega_k} L(\omega) = \sum_{i=1}^{n} x_i I(y_i = k) - \sum_{i=1}^{n} \frac{x_i e^{\omega_k x_i}}{\sum_{j=1}^{K} e^{\omega_j x_i}}$$

$$\nabla L(\omega) = \begin{bmatrix} \frac{\partial}{\partial \omega_1} L(\omega) \\ \frac{\partial}{\partial \omega_2} L(\omega) \\ \vdots \\ \frac{\partial}{\partial \omega_{K-1}} L(\omega) \end{bmatrix}$$

$$\nabla f(\omega) = \nabla L(\omega) - \nabla \left(\frac{\lambda}{2} \sum_{j=1}^{K-1} |\omega_j|_2^2 \right)$$

$$\frac{\partial}{\partial \omega_k} \frac{\lambda}{2} \sum_{j=1}^{K-1} |\omega_j|_2^2 = \lambda \omega_k$$

$$\Rightarrow \nabla f(\omega) = \nabla L(\omega) - \begin{bmatrix} \lambda \omega_1 \\ \lambda \omega_2 \\ \vdots \\ \lambda \omega_{K-1} \end{bmatrix}$$

Subject:
Date

$$w_j = \underset{w_j}{\arg\min}(|w_j x_j - Y|)$$ -4

Simply answer of least squares Prove in (ب)

$$\Rightarrow w_j = \underbrace{(x_j^T x_j)^{-1}}_{\text{number}} x_j^T Y = \frac{x_j^T Y}{(x_j^T x_j)}$$ الف)

Assuming x_j's are uncorrelated ب)

$$\Rightarrow X^T X \rightarrow \text{diagonal} = \text{diag}[\lambda_i]_{i=1}^{m}$$

we know:

$$\hat{w} = \underbrace{(X^T X)^{-1}}_{\text{diag}\left[\frac{1}{\lambda_i}\right]_{i=1}^{m}} X^T Y = \underbrace{\text{diag}\left[\frac{1}{\lambda_i}\right]_{i=1}^{m}}_{D} \underset{nm}{X^T} \underset{n\times 1}{Y}$$

row j of D

$$\Rightarrow \hat{w}_j = [d_j] \cdot X^T \cdot Y$$

Since $[d_j]$ is zero in all columns except j th column

$$\Rightarrow \hat{w}_j = \frac{1}{\lambda_j} x_j^T Y = \frac{x_j^T Y}{\lambda_j}$$

and we know $\lambda_j = x_j^T x_j$ So $\hat{w}_j = \frac{x_j^T Y}{(x_j^T x_j)}$

we know

$$Cov(x_j, y) = \left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right)\frac{1}{n-1} =$$ پ)

$$Var(x_j) = \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)\frac{1}{n-1}$$

Subject:
Date: least squares

taking partial derivatives. $S(\omega, \omega_0) = \sum_{i=1}^{n} (y_i - \omega_0 - \omega x_i)^2$

$$\Rightarrow \frac{\partial S}{\partial \omega_0} = -2 \sum_{i=1}^{n} (y_i - \hat{\omega}_0 - \hat{\omega} x_i) = 0$$

$$\Rightarrow \frac{\partial S}{\partial \omega} = -2 \sum_{i=1}^{n} (y_i - \hat{\omega}_0 - \hat{\omega} x_i) x_i = 0$$

$$\hookrightarrow \sum y_i = n\hat{\omega}_0 + \hat{\omega} \sum x_i \Rightarrow \bar{y} = \hat{\omega}_0 + \hat{\omega}\bar{x}$$

$$\hookrightarrow \sum x_i y_i = \hat{\omega}_0 \sum x_i + \hat{\omega} \sum x_i^2 \qquad \Rightarrow \hat{\omega}_0 = \bar{y} - \hat{\omega}\bar{x}$$

$$\Rightarrow \sum x_i y_i = (\bar{y} - \hat{\omega}\bar{x}) n\bar{x} + \hat{\omega} \sum x_i^2$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} = \hat{\omega} \left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right)$$

$$\Rightarrow \hat{\omega} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{(n-1)\, Cov(x, y)}{(n-1)\, Var(x)} = \frac{Cov(x, y)}{Var(x)}$$

Refrence: Dr. Guangliang Chen | Mathematics & Statistics
San José State Uni.

PAPCO

Subject :
Date

-5

$$E(X) = \int_0^\infty x f(x)\,dx = \int_0^a x f(x)\,dx + \int_a^\infty x f(x)\,dx$$ (۱

$$P(X \geq a) = \int_a^\infty f(x)\,dx$$

$$\hookrightarrow E(x) \geq \int_a^\infty x f(x)\,dx \geq \int_a^\infty a f(x)\,dx = a \int_a^\infty f(x)\,dx$$

$$\Rightarrow E(X) \geq a\,P(X \geq a)$$

$$\Rightarrow P(X \leq a) \leq \frac{E(X)}{a}$$

$$P(|X-\mu| \geq \alpha) = P((X-\mu)^2 \geq \alpha^2) \leq \frac{E((X-\mu)^2)}{\alpha^2}$$ (۲

$$E((X-\mu)^2) = \sigma_x^2$$

$$\Rightarrow P(|X-\mu| \geq \alpha) \leq \frac{\sigma_x^2}{\alpha^2}$$

if we throw a random point inside square (uniformly) (2

it will get inside circle with $p(\text{inside}) = \frac{S_c}{S_s} = \frac{\pi r^2}{(2r)^2} = \frac{\pi}{4}$

So if we throw N times we have a binomial random var.

$$f\left(\frac{N\pi}{4}, N, \frac{\pi}{4}\right) = \binom{N}{\lceil \frac{N\pi}{4} \rceil} \left(\frac{\pi}{4}\right)^{\lceil \frac{N\pi}{4} \rceil} \left(1-\frac{\pi}{4}\right)^{\lceil N(1-\frac{\pi}{4}) \rceil}$$

From chebychev:

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

Subject:
Date

Question: $P\left(\left|\frac{k}{N} - \frac{\pi}{4}\right| \geqslant 0{,}01\,\frac{\pi}{4}\right) \leqslant 0{,}05$

$$\Rightarrow P\left(\left|k - \underbrace{\frac{N\pi}{4}}_{\mu}\right| \geqslant \underbrace{0{,}01\,\frac{N\pi}{4}}_{\alpha}\right) \leqslant 0{,}05 = \frac{\sigma^2}{\alpha^2}$$

binomial $\longrightarrow \mu$

So from chebychev:

$$\frac{\sigma^2}{\alpha^2} \leqslant 0{,}05 \rightarrow \sigma^2 \leqslant 0{,}05 \times \left(0{,}01\,\frac{N\pi}{4}\right)^2$$

$\xrightarrow{\text{binomial}}$ $\sigma^2 = npq$

$$\Rightarrow N\frac{\pi}{4}\left(1-\frac{\pi}{4}\right) \leqslant 5 \times 10^{-4} \cdot N^2 \frac{\pi^2}{16}$$

$$\Rightarrow N \geqslant \left\lceil \frac{\pi}{4}\left(1-\frac{\pi}{4}\right) \cdot \frac{16}{\pi^2} \cdot 2000 \right\rceil$$

$$\Rightarrow N \geqslant 547$$

Subject:
Date

From definition of SVD: -6

$$A = V\Sigma U^T \Rightarrow A^{-1} = V\Sigma^{-1}U^T \quad \text{where } \Sigma = \text{diag}[\sigma_i]^T_{i=1,\text{rank}(A)}$$

Since Σ is diagonal $\rightarrow \Sigma^{-1} = \text{diag}\left[\frac{1}{\sigma_i}\right]_{i=1}^{\text{rank}A}$ الف

$$\rightarrow \sigma_{max}(A) = \Sigma_{1,1} \quad \sigma_{max}(A^{-1}) = \frac{1}{\sigma_{min}(A)}$$

we know: $\frac{\sigma_{max}(A)}{\sigma_{min}(A)} \geqslant 1 \Rightarrow \sigma_{max}(A) \cdot \frac{1}{\sigma_{min}(A)} \geqslant 1$

$$\Rightarrow \sigma_{max}(A) \cdot \sigma_{max}(A^{-1}) \geqslant 1$$

We know $\|A\|_2 = \sigma_{max}(A)$ and $\|A\|_F = \sqrt{\sum_{i=1}^{\text{rank}(A)} \sigma_i^2}$

$$\|A\|_2^2 \leqslant \sum_{i=1}^{\text{rank}(A)} \sigma_i^2 \Rightarrow \|A\|_2 \leqslant \sqrt{\sum_{i=1}^{\text{rank}(A)} \sigma_i^2} = \|A\|_F$$

$$\left(\sqrt{\text{rank}(A)} \cdot \|A\|_2\right)^2 = \text{rank}(A) \cdot \sigma_{max}(A)^2$$

$$= \sum_{i=1}^{\text{rank}(A)} \sigma_{max}(A)^2 \geqslant \sum_{i=1}^{\text{rank}(A)} \sigma_i^2 = \|A\|_F^2$$

$$\Rightarrow \|A\|_F \leqslant \left(\sqrt{\text{rank}(A)} \cdot \|A\|_2\right)$$

$$\Rightarrow \|A\|_2 \leqslant \|A\|_F \leqslant \sqrt{\text{rank}(A)}\,\|A\|_2$$

Subject:
Date

$$\sigma(x) = \frac{1}{1+e^{-x}} = \frac{e^{x}}{e^{x}+1} \qquad \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1}$$ -7

$$\Rightarrow \tanh(x) = 2\sigma(2x) - 1 \Rightarrow \sigma(2x) = \frac{\tanh(x)+1}{2}$$

$$y(x,w) = w_0 + \sum_{j=1}^{n}\left(w_j\, \sigma\left(2\,\frac{x-\mu_j}{s}\right)\right)$$

$$= w_0 + \sum_{j=1}^{n}\left(w_j\, \frac{\tanh\left(\frac{x-\mu_j}{s}\right)+1}{2}\right)$$

$$= w_0 + \frac{1}{2}\sum_{1}^{n} w_j + \sum_{j=1}^{n} \frac{1}{2} w_j \tanh\left(\frac{x-\mu_j}{2}\right)$$

$$= \underbrace{\qquad}_{u_0} + \sum_{j=1}^{n} u_j \tanh\left(\frac{x-\mu_j}{2}\right) = y(x,u)$$

where $u_0 = w_0 + \frac{1}{2}\sum_{j=1}^{n} w_j$

and $u_j = \frac{1}{2} w_j$