حل تمرین سری ۱ - ۹۹۱۰۲۱۹۱

A

سؤال ۱:

$$E[Y] = \int_{-1}^{1} X^2 \, dX = E[X^2]$$

حال $\operatorname{cov}(X,Y)$ را محاسبه می‌کنیم:

$$\operatorname{cov}(X,Y) = \iint (X-E[X])\cdot(Y-E[Y])\cdot p(X,Y)\, dX\, dY$$

$$= \int_{-1}^{1} X\left(X^2 - E[X^2]\right) dX = 0 \Rightarrow \rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} = 0$$

سؤال ۲:

$$E(XZ|Y) = \int xz\, p_{x,z|y}(x,z|y)\, dz\, dx = \int xz\, p_{x|y}(x|y)\, p_{z|y}(z|y)\, dz\, dx$$

$$= \int x\, p_{x|y}(x|y)\, dx \int z\, p_{z|y}(z|y)\, dz = E(X|Y)\,E(Z|Y) \quad *$$

$$\Rightarrow E(XZ) = E_Y\left(E(XZ|Y)\right) \overset{*}{=} E_Y\left(E(X|Y)\,E(Z|Y)\right) \quad ①$$

با توجه به اینکه X و Z به صورت شرطی مستقل از Y هستند، می‌توان نوشت:

$$p_{X|Y}(X|Y) = \mathcal{N}(\mu_{X|Y}, \Sigma_{X|Y}), \quad \Sigma_{XY} = \operatorname{cov}(X,Y) = \rho_{x,y}\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}$$

$$, \quad \Sigma_{YY}^{-1} = \frac{1}{\operatorname{Var}(Y)} \xrightarrow[\text{با توجه به فرمول (۲) داریم}]{} E(X|Y) = \mu_x + \rho_{x,y}\sqrt{\frac{\operatorname{Var}(X)}{\operatorname{Var}(Y)}}(y-\mu_y) \quad ②$$

$$p_{Z|Y}(Z|Y) = \mathcal{N}(\mu_{Z|Y}, \Sigma_{Z|Y}) \xrightarrow[\text{و روابط بالا}]{\text{فرمول (۲)}} E(Z|Y) = \mu_z + \rho_{z,y}\sqrt{\frac{\operatorname{Var}(Z)}{\operatorname{Var}(Y)}}\, (y-\mu_y) \quad ③$$

$$\xrightarrow[②]{①,③} E(XZ) = E_y\Big(\mu_x\mu_z + \left(\mu_x \rho_{z,y}\sqrt{\operatorname{Var}(Z)} + \mu_z \rho_{x,y}\sqrt{\operatorname{Var}(X)}\right)\frac{y-\mu_y}{\sqrt{\operatorname{Var}(Y)}}$$

$$+ \rho_{x,y}\rho_{z,y}\frac{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)}}{\operatorname{Var}(Y)}(y-\mu_y)^2\Big)$$

$$= \mu_x\mu_z + \frac{\mu_x\rho_{z,y}\sqrt{\operatorname{Var}(Z)} + \mu_z\rho_{x,y}\sqrt{\operatorname{Var}(X)}}{\sqrt{\operatorname{Var}(Y)}}\, \cancel{E_y(y-\mu_y)}^{\,0} + \rho_{x,y}\rho_{z,y}\frac{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)}}{\operatorname{Var}(Y)}\, \underbrace{E\left((y-\mu_y)^2\right)}_{\operatorname{Var}(Y)}$$

$$= \mu_x\mu_z + \rho_{x,y}\rho_{z,y}\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)} \rightarrow \rho_{x,z} = \frac{\operatorname{cov}(X,Z)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)}} = \frac{E(XZ) - E(X)E(Z)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)}}$$

$$\rightarrow \rho_{x,z} = \frac{\mu_x\mu_z + \rho_{x,y}\rho_{z,y}\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)} - \mu_x\mu_z}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Z)}} = \rho_{x,y}\rho_{z,y}$$

سوال ۳ قه

$$\bar{y}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} y_i^{(1)} \quad , \quad \bar{y}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i^{(2)}$$

prior $\rightarrow$ $\mu_{\mu_z} = 0 \quad , \quad \Sigma_{\mu_z} = \infty$ ، likelihood : $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ و $\mu_y = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\Sigma_y = \begin{bmatrix} \frac{v_1}{n_1} & 0 \\ 0 & \frac{v_2}{n_2} \end{bmatrix}$

از LGS می‌دانیم:

$$\Sigma^{-1}_{\mu_z|y} = \Sigma^{-1}_{\mu_z} + A^T \Sigma_y^{-1} A = 0 + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{v_1}{n_1} & 0 \\ 0 & \frac{v_2}{n_2} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{n_1}{v_1} + \frac{n_2}{v_2}$$

$$\mu_{\mu_z|y} = \Sigma^{-1}_{\mu_z|y}\left[A^T \Sigma_y^{-1} \bar{y} + \Sigma^{-1}_{\mu_z} \mu_{\mu_z}\right] = \Sigma^{-1}_{\mu_z|y}\left[\frac{n_1 \bar{y}_1}{v_1} + \frac{n_2 \bar{y}_2}{v_2}\right]$$

سوال ۲ - الف)

$$NLL(\theta) \Rightarrow NLL(\theta) = -\log \prod_{n=1}^{N} p(y_n|\theta) = -\log \prod_{n=1}^{N} \theta^{\mathbb{I}(y_n=1)} (1-\theta)^{\mathbb{I}(y_n=0)}$$

$$= -\sum_{n=1}^{N} [\mathbb{I}(y_n=1)\log\theta + \mathbb{I}(y_n=0)\log(1-\theta)] = -[N_1 \log(\theta) + N_0 \log(1-\theta)]$$

برای بدست آوردن $\hat{\theta}_{mle}$ داریم:

$$\frac{d}{d\theta} NLL(\theta) = \frac{-N_1}{\theta} + \frac{N_0}{1-\theta} = 0 \Rightarrow \hat{\theta}_{mle} = \frac{N_1}{N_0+N_1}$$

ب)

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \; 0 < x < \infty \quad 0<\theta<\infty$$

$$\Rightarrow L(\theta) = L(\theta; x_1, \dots, x_n) = \left(\frac{1}{\theta} e^{-\frac{x_1}{\theta}}\right)\left(\frac{1}{\theta} e^{-\frac{x_2}{\theta}}\right)\dots\left(\frac{1}{\theta} e^{-\frac{x_n}{\theta}}\right) = \frac{1}{\theta^n} e^{\frac{-\sum_{i=1}^{n} x_i}{\theta}}$$

$$\Rightarrow LL(\theta) = -n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i \Rightarrow \frac{d(LL(\theta))}{d\theta} = \frac{-n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \hat{\theta}_{mle} = \frac{\sum_{i=1}^{n} x_i}{n}$$

ج)

$$NLL(\mu, \sigma^2) = -\sum_{n=1}^{N_D} \log\left[\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} \exp\left(-\frac{(y_n-\mu)^2}{2\sigma^2}\right)\right] = \frac{1}{2\sigma^2}\sum_{n=1}^{N}(y_n-\mu)^2 + \frac{N_D}{2}\log(2\pi\sigma^2)$$

$$\Rightarrow \frac{\partial}{\partial\mu} NLL(\mu, \sigma^2) = 0 \Rightarrow \hat{\mu}_{mle} = \frac{1}{N}\sum_{n=1}^{N} y_n = \bar{y}$$

$$, \; \frac{\partial}{\partial\sigma^2} NLL(\mu, \sigma^2) = 0 \Rightarrow \hat{\sigma}^2_{mle} = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{\mu}_{mle})^2 = \frac{1}{N}\sum_{n=1}^{N}[y_n^2 + \hat{\mu}^2_{mle} - 2\hat{\mu}_{mle} y_n]$$

سوال ۸: ابتدا μ و Σ را برای $p(x_1, x_2)$ محاسبه می‌کنیم:

$$p(x_1, x_2) \propto \exp\left[-\frac{1}{2}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}^T \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}\right]$$

می‌توان نوشت:

$$\Longrightarrow p(x_1, x_2) \propto \exp\left[-\frac{1}{2}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}^T \begin{pmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{pmatrix}\begin{pmatrix} (\Sigma/\Sigma_{22})^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix}\right.$$

$$\left. \times \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}\right]$$

$$= \exp\left[-\frac{1}{2}(x_1-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2))^T(\Sigma/\Sigma_{22})^{-1}(x_1-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2))\right]$$

$$\times \exp\left[-\frac{1}{2}(x_2-\mu_2)^T\Sigma_{22}^{-1}(x_2-\mu_2)\right] \quad (1)$$

$$\Rightarrow p(x_1, x_2) = p(x_1|x_2)\,p(x_2) \Rightarrow p(x_1|x_2) \propto N(x_1|\mu_{1|2}, \Sigma_{1|2})\,N(x_2|\mu_2, \Sigma_{22}) \quad (2)$$

$$(1), (2) \Rightarrow \mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2), \quad \Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$