Exercise SHEET 4 | Lamia Islam

Problem 1

a) Let's consider the case where X_1 and X_2 are jointly Gaussian with

$$\mathbb{E}[(X_1, X_2)] = (\bar{x}_1, \bar{x}_2) \quad \text{and}$$

$$\operatorname{Var}[(X_1, X_2)] = \begin{pmatrix} \sigma_1^2 & c \\ c & \sigma_2^2 \end{pmatrix} \quad \text{where } c = \text{covariance}$$

$$\therefore c = \operatorname{Cov}[X_1, X_2] = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$$

Now, considering the square of the L^2-Wasserstein distance for the sake of simplicity:

$$(W_2(\pi X_1, \pi X_2))^2 = \inf_c \mathbb{E}[(X_1 - X_2)^2]$$

$$= \inf_c \mathbb{E}[X_1^2 - 2X_1X_2 + X_2^2] \quad \text{— ①}$$

We know, $\operatorname{Var}(X) = E[X^2] - (E[X])^2$

$$\therefore E[X^2] = (E[X])^2 + \operatorname{Var}(X) = \cancel{x_1}$$

$$\therefore E[X_1^2] = \bar{x}_1^2 + \sigma_1^2 \qquad \Big| \quad c = E[X_1X_2] - E[X_1]E[X_2]$$

$$\text{and } E[X_2^2] = \bar{x}_2^2 + \sigma_2^2 \qquad \Big| \quad \therefore E[X_1X_2] = c + \bar{x}_1\bar{x}_2$$

From ① $\Rightarrow$ $$(W_2(\pi X_1, \pi X_2))^2 = \inf_c (\bar{x}_1^2 + \sigma_1^2) \cancel{2c} + (\bar{x}_2^2 + \sigma_2^2) - 2(c + \bar{x}_1 \cdot \bar{x}_2)$$

$$= \inf_c (\bar{x}_1 - \bar{x}_2)^2 + \sigma_1^2 + \sigma_2^2 - 2c$$

The maximum eligible value of c is given by the Cauchy-Schwarz inequality:

$$(\operatorname{Cov}[X_1, X_2])^2 \leq \operatorname{Var}[X_1]\operatorname{Var}[X_2] \Rightarrow c \leq \sigma_1 \sigma_2$$

Therefore,

$$(W_2(\pi X_1, \pi X_2))^2 = (\bar{x}_1 - \bar{x}_2)^2 + \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2$$

$$= (\bar{x}_1 - \bar{x}_2)^2 + (\sigma_1 - \sigma_2)^2$$

(b) The Kullback-Leibler (KL) divergence of πx_1 from πx_2 is given by:

$$D_{KL}(\pi x_1 \| \pi x_2) := \int_R \log\left(\frac{\pi x_1(x)}{\pi x_2(x)}\right) \pi x_1(x)\, dx$$

$$= \int_R [\log \pi x_1(x) - \log \pi x_2(x)]\, \pi x_1(x)\, dx \quad \text{—①}$$

For Gaussian distribution, PDF

$$\pi X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

~~Substituting the formula for the KL divergence gives:~~

$\cancel{D_{KL}(\pi x_1 \| \pi x_2) = \int_R}$

Taking logarithm on both sides, we get,

$$\log \pi X(x) = \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)\right)$$

$$= \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \log\left(\exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)\right)$$

$$= -\tfrac{1}{2}\log(2\pi\sigma^2) - \frac{(x-\bar{x})^2}{2\sigma^2}$$

So, ① $\Rightarrow$ $\int_R \left[-\tfrac{1}{2}\log(2\pi\sigma_1^2) - \frac{(x-\bar{x}_1)^2}{2\sigma_1^2} + \tfrac{1}{2}\log(2\pi\sigma_2^2) + \frac{(x-\bar{x}_2)^2}{2\sigma_2^2}\right] \pi x_1(x)\, dx$

$$= \mathbb{E}\left[\tfrac{1}{2}\log\frac{\sigma_2^2}{\sigma_1^2} - \frac{1}{2\sigma_1^2}(x_1-\bar{x}_1)^2 + \frac{1}{2\sigma_2^2}(x_1-\bar{x}_2)^2\right]$$

Here, the first term is constant, and the second term is the variance of X_1, that is,

$$\frac{1}{2\sigma_1^2}\mathbb{E}[(X_1-\bar{x}_1)^2] = \frac{1}{2\sigma_1^2}\cdot\sigma_1^2 = \frac{1}{2}$$

For the third term,

$$\mathbb{E}[(X_1-\bar{x}_2)^2] = \mathbb{E}[X_1^2 - 2\bar{x}_2 X_1 + \bar{x}_2^2]$$

$$= (\sigma_1^2 + \bar{x}_1^2) - 2\bar{x}_1\bar{x}_2 + \bar{x}_2^2$$

$$= \sigma_1^2 + (\bar{x}_1 - \bar{x}_2)^2$$

Collecting all three terms, we conclude that,

$$D_{KL}(\pi x_1 \| \pi x_2) = \tfrac{1}{2}\left[\log\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2 + (\bar{x}_1-\bar{x}_2)^2}{\sigma_2^2} - 1\right]$$

$$= \tfrac{1}{2}\left[\log\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} + \frac{(\bar{x}_1-\bar{x}_2)^2}{\sigma_2^2} - 1\right]$$

Note that this value becomes zero, if $\bar{x}_1 = \bar{x}_2$ and $\sigma_1^2 = \sigma_2^2$