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

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

$$\therefore E[X_1^2] = \bar{x}_1^2 + \sigma_1^2$$

$$\text{and } E[X_2^2] = \bar{x}_2^2 + \sigma_2^2$$

$$c = E[X_1 X_2] - E[X_1]E[X_2]$$

$$\therefore E[X_1 X_2] = c + \bar{x}_1 \bar{x}_2$$

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 $\pi x_1$ from $\pi x_2$ 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 $\sigma_1^2 = \sigma_2^2$

Sheet(4) md Borhan Uddin

Problem 2(a):

Given the sets of,

$$X_1 = \{a_1 = 1, a_2 = 2, a_3 = 3\}$$
$$X_2 = \{b_1 = 1.5, b_2 = 2, b_3 = -1\}$$

and,

$$p(a_i) = p(b_i) = 1/3, \text{ where } i = 1,2,3.$$
$$T = [t_{ij}]_{i,j=1}^{3} \in R^{3\times3} \quad (t_{ij} \geq 0).$$
$$J(T) = \sum_{i,j=1}^{3} t_{ij} |b_i - a_j|^2.$$

To solve this problem, we assume that we have three water bottle that filled with 1/3 unit of water. From the baseline the bottle heights are $a_1, a_2, a_3$. Now, we target to re distribute the water into another sets of water bottle at height $b_1, b_2, b_3$. The required energy required to move a unit from one bottle to Another is proportion of square of height difference. To find a strategy that minimizes energy Consumption,

if $D_{ij}$ energy consumption per unit from $a_j$ to $b_i$, $D_{ij}$ can be expressed by following matrix

$$[D_{ij}] = \begin{pmatrix} (a_1-b_1)^2 & (a_1-b_2)^2 & (a_1-b_3)^2 \\ (a_2-b_1)^2 & (a_2-b_2)^2 & (a_2-b_3)^2 \\ (a_3-b_1)^2 & (a_3-b_2)^2 & (a_3-b_3)^2 \end{pmatrix}$$

value

$$\begin{bmatrix} a_1 = 1, & b_1 = 1.5 \\ a_2 = 2, & b_2 = 2 \\ a_3 = 3, & b_3 = -1 \end{bmatrix}$$

$$= \begin{pmatrix} 0.25 & 1 & 4 \\ 0.25 & 0 & 9 \\ 2.25 & 1 & 16 \end{pmatrix}$$

In the third column, the cost from anywhere to $b_3$ is higher in every scenarios. So, from here we can start rearrange-trick. We want to fill $b_3$ as much as from $a_1$, and if it is not possible, the we may use water from $a_2$ and so on.

In the Question, we have exactly 1/3 unit in $a_1$ and exactly 1/3 will be transferred into $b_3$. So,

$$[T_{ij}] = \begin{pmatrix} 0 & 0 & 1/3 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 0 \end{pmatrix}$$

Now, we repeat for the rest of bottles. Now cost to fill $b_1$ dominates one required to fill $b_2$, we optimize $b_1$ first by Assign 1/3 from $a_2$ to $b_1$, namly $t_{21} = 1/3$.

So, the resulting optimal strategy is,

$$[T_{ij}^*] = \begin{pmatrix} 0 & 0 & 1/3 \\ 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \end{pmatrix}.$$

If we change the order of indices to visualize the strategy. By re-arranging, we have:

$$[T_{ij}] = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$$

| $[T_{ij}]$ | $b_3$ | $b_1$ | $b_2$ |
|---|---|---|---|
| $a_1$ | 1/3 | 0 | 0 |
| $a_2$ | 0 | 1/3 | 0 |
| $a_3$ | 0 | 0 | 1/3 |

$\Rightarrow$ Re-arrange order $[T_{ij}]$

| $[D_{ij}]$ | $b_3$ | $b_1$ | $b_2$ |
|---|---|---|---|
| $a_1$ | 4 | 0.25 | 1 |
| $a_2$ | 9 | 0.25 | 0 |
| $a_3$ | 16 | 2.25 | 1 |

$\Rightarrow$ Rearrang order $[D_{ij}]$

SANJAY RAJPUROHIT

Problem 2b

The cost function is defined as

$$J(T) = \sum_{i,j=1}^{3} t_{ij} |b_i - a_j|^2$$

We need to calculate $J(T^*)$, where the elements of the optimal strategy matrix

$$[T^*_{ij}] = \begin{bmatrix} 0 & 0 & 1/3 \\ 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$$ from 2a

and $|b_i - a_j|^2 = D_{ij}$

$$[D_{ij}] = \begin{bmatrix} 0.25 & 1 & 4 \\ 0.25 & 0 & 9 \\ 2.25 & 1 & 16 \end{bmatrix}$$ from 2a

So, $J(T^*) = \sum_{i,j=1}^{3} [T^*_{ij}][D_{ij}]$

$$= \sum \begin{bmatrix} 0 & 0 & 0.33 \\ 0.33 & 0 & 0 \\ 0 & 0.33 & 0 \end{bmatrix} \times$$

$$\begin{bmatrix} 0.25 & 1 & 4 \\ 0.25 & 0 & 9 \\ 2.25 & 1 & 16 \end{bmatrix}$$

$$= \sum \begin{bmatrix} 2.25 \times 0.33 & 0.33 & 0.33 \times 16 \\ 0.33 \times 0.25 & 0.33 & 0.33 \times 4 \\ 0.33 \times 0.25 & 0 & 0.33 \times 9 \end{bmatrix}$$

$$= \sum \begin{bmatrix} 0.74 & 0.33 & 5.28 \\ 0.08 & 0.33 & 1.32 \\ 0.08 & 0 & 2.97 \end{bmatrix}$$

$$= (0.74 + 0.33 + 5.28 + 0.08 + 0.33 + 0 + 1.32 + 0.08 + 2.97)$$

$$= 11.13$$