HOMEWORK - 1

Question 1

(1) Let $x \in R$ be a random variable with mean $E[x] = \mu$. Then, the variance of x,

$$E[(x-\mu)^2] = E[x^2 - 2x\mu + \mu^2]$$
$$= E[x^2] - E[2x\mu] + E[\mu^2]$$
$$= E[x^2] - 2\mu \cdot E[x] + E[\mu^2]$$
$$= E[x^2] - 2\mu \cdot \mu + \mu^2$$
$$= E[x^2] - 2\mu^2 + \mu^2$$
$$\therefore E[(x-\mu)^2] = E[x^2] - \mu^2$$

(2) Let $x \in \mathbb{R}^n$ be an n-dimensional random vector with mean $E[x] = \mu$. Then, the covariance matrix of x,

$$E[(x-\mu)(x-\mu)^T] = E[x \cdot x^T - \mu x^T - x\mu^T + \mu\mu^T]$$
$$= E[x \cdot x^T] - E[\mu x^T] - E[x\mu^T] + E[\mu\mu^T]$$
$$= E[x \cdot x^T] - \mu E[x^T] - \mu^T E[x] + \mu\mu^T$$
$$= E[x \cdot x^T] - \mu \cdot \mu^T - \mu^T \cdot \mu + \mu\mu^T$$
$$= E[x \cdot x^T] - \mu \cdot \mu^T$$

Question 2

① Let the normalization factor be 'N'. Since the pdfs integrate to 1 along the real axis,

$$\therefore P_X(x|L=l) = \int_{-\infty}^{\infty} N \cdot e^{\frac{-|x-a_l|}{b_l}} \cdot dx = 1 \qquad \text{for } l \in 1,2 \text{ and } b_l > 0$$

Due to the modulus involved, the integral is divided as:

$$2\int_{a_l}^{\infty} N \cdot e^{\frac{-|x-a_l|}{b_l}} \cdot dx = 1$$

$$2N \int_{a_l}^{\infty} e^{\frac{-|x-a_l|}{b_l}} \cdot dx = 1$$

$$2N \left[\frac{1}{-1/b_l} \cdot e^{\frac{-|x-a_l|}{b_l}} \right]_{a_l}^{\infty} = 1$$

$$-2Nb_l \left[e^{-\infty} - e^{0} \right] = 1$$

$$-2Nb_l [0-1] = 1$$

$$\therefore 2Nb_l = 1$$

$$\therefore N = \frac{1}{2b_l} \Rightarrow P_X(x|L=l) = \frac{1}{2b_l} e^{\frac{-|x-a_l|}{b_l}}$$

② The log-likelihood ratio ~~function for the case~~ between class labels 1 and 2 evaluated at a given x,

$$l(x) = \ln p(x|L=1) - \ln p(x|L=2)$$

$$= \ln\left(\frac{1}{2b_1} \cdot e^{\frac{-|x-a_1|}{b_1}}\right) - \ln\left(\frac{1}{2b_2} \cdot e^{\frac{-|x-a_2|}{b_2}}\right)$$

$$= \ln\left(\frac{1}{2b_1}\right) + \ln\left(e^{\frac{-|x-a_1|}{b_1}}\right) - \ln\left(\frac{1}{2b_2}\right) - \ln\left(e^{\frac{-|x-a_2|}{b_2}}\right)$$

$$l(x) = -\ln(2b_1) - \frac{|x-a_1|}{b_1} + \ln(2b_2) + \frac{|x-a_2|}{b_2}$$

$$l(x) = \ln\left(\frac{b_2}{b_1}\right) + \frac{|x-a_2|}{b_2} - \frac{|x-a_1|}{b_1}$$

(3) Log-likelihood-ratio function for the case $a_1=0$, $b_1=1$, $a_2=1$, $b_2=2$ is,

$$l(x) = \ln\left(\frac{b_2}{b_1}\right) + \frac{|x-a_2|}{b_2} - \frac{|x-a_1|}{b_1}$$

$$= \ln\left(\frac{2}{1}\right) + \frac{|x-1|}{2} - \frac{|x-0|}{1}$$

$$= \ln(2) + \frac{|x-1|}{2} - |x|$$

NOTE: Plot of data-set along with its code is on GitHub. The link is:

github.com/nandayvk/Intro-to-Machine-Learning-HW-1

Question 3

For a two-class classification,

$$\frac{P(x|L=1)}{P(x|L=2)} \underset{D=2}{\overset{D=1}{\gtrless}} \gamma \quad \text{—— (1)} \qquad \text{where} \quad \gamma = \frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \cdot \frac{P(L=2)}{P(L=1)}$$

If $P(L=1) = P(L=2)$, then $\gamma = 1$.

For the given classes, $P(x|L=1) = \frac{1}{b-a}$

and $P(x|L=2) = \frac{1}{t-r}$

Therefore, substituting in (1),

$$\frac{\frac{1}{b-a}}{\frac{1}{t-r}} \underset{D=2}{\overset{D=1}{\gtrless}} 1$$

$$\therefore \quad t-r \underset{D=2}{\overset{D=1}{\gtrless}} b-a$$

~~Since~~ This is the minimum probability of error classification rule for the given distributions.

Question 4

(1) Let class 1 have pdf of $\mathcal{N}(0,1)$ and let class 2 have a pdf of $\mathcal{N}(\mu,\sigma^2)$.

According to the classification rule for a two-class setting,

$$\frac{P(X|L=1)}{P(X|L=2)} \underset{D=2}{\overset{D=1}{\gtrless}} \gamma$$

When the classes have equal class priors, $\gamma = 1$.

Therefore,

$$\frac{P(X|L=1)}{P(X|L=2)} \underset{D=2}{\overset{D=1}{\gtrless}} 1$$

$$P(x|L=1) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \cdot e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

$$= (2\pi)^{-n/2} \cdot 1 \cdot e^{-\frac{1}{2} \cdot x^T \cdot x}$$

$$P(x|L=2) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \cdot e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

$$= (2\pi)^{-n/2} \cdot \sigma^{-1} \cdot e^{-\frac{1}{2}(x-\mu)^T \cdot \sigma^{-2} \cdot (x-\mu)}$$

Therefore, substituting in the classification rule,

$$\frac{\cancel{(2\pi)^{-n/2}} \cdot e^{-\frac{1}{2} x^T \cdot x}}{\cancel{(2\pi)^{-n/2}} \cdot \sigma^{-1} \cdot e^{-\frac{1}{2}(x-\mu)^T \cdot \sigma^{-2} \cdot (x-\mu)}} \underset{D=2}{\overset{D=1}{\gtrless}} 1$$

$\therefore$ $\cancel{e^{-\frac{1}{2}x^T x}}$ Therefore, taking ln (log) on both sides,

$$-\frac{1}{2} \cdot x^T \cdot x + \ln(\sigma) - \left(-\frac{1}{2}(x-\mu)^T \cdot \sigma^{-2} \cdot (x-\mu)\right) \underset{D=2}{\overset{D=1}{\gtrless}} 0$$

$$\therefore \ln(\sigma) - \frac{1}{2}\left(x^T \cdot x + \frac{(x-\mu)^T(x-\mu)}{\sigma^2}\right) \underset{D=2}{\overset{D=1}{\gtrless}} 0$$

$$\therefore \ln(\sigma) = \sum_{D=2}^{D=1} \frac{\sigma^2 \cdot x^T \cdot x + (x-\mu)^T (x-\mu)}{2\sigma^2}$$

$$\therefore 2\sigma^2 \ln(\sigma) \sum_{D=2}^{D=1} \sigma^2 \cdot x^T \cdot x + (x-\mu)^T (x-\mu)$$

Question 5

(1) Given: $x = Az + b$ $\quad A \in \mathbb{R}^{n \times m}, \; b \in \mathbb{R}^n, \; z \sim \mathcal{N}(0, I)$

$$E[x] = E[Az + b]$$
$$= E[Az] + E(b)$$
$$= A\,E[z] + b$$
$$= A \cdot (0) + b$$
$$E(x) = b = \mu_x$$

$$Var[x] = Var[Az + b] \qquad Var[x] = E[(x-\mu)(x-\mu)^T]$$
$$= E[(Az + b - \mu_x)(Az + b - \mu_x)^T]$$
$$= E[(Az + b - b)(Az + b - b)^T]$$
$$= E[(Az)(Az)^T]$$
$$= A \cdot E[z \cdot z^T] \cdot A^T$$
$$= A \Sigma_z A^T$$
$$= A I A^T$$
$$Var[x] = AA^T = \sigma_x^2$$

$$\therefore x \sim \mathcal{N}(\mu_x, \sigma_x^2)$$

$$\therefore x \sim \mathcal{N}(b, AA^T)$$

(2) Let $x = Az + b \Rightarrow z = A^{-1}(x-b)$

$$\therefore \frac{\partial x}{\partial z} = A$$

We know that, while transforming from one PDF to another,

$$P_x(x) = \frac{P_z(z)}{\left|\frac{\partial x}{\partial z}\right|} = \frac{1}{|A|} \cdot (2\pi)^{-n/2} \cdot |\Sigma|^{-1/2} e^{-\frac{1}{2}(z-\mu)^T \Sigma^{-1} (z-\mu)}$$

Substituting value of z in eqn.,

$$P_x(x) = \frac{1}{|A|} \cdot (2\pi)^{-n/2} \cdot |\Sigma_z|^{-1/2} \cdot e^{-\frac{1}{2}(A^{-1}(x-b) - \mu_z)^T \Sigma_z^{-1} (A^{-1}(x-b) - \mu_z)}$$

$$= \frac{1}{|A|} \cdot (2\pi)^{-n/2} \cdot |\Sigma_z|^{-1/2} \cdot e^{-\frac{1}{2}\left(x-b-\frac{\mu_z}{A^{-1}}\right)^T (A^{-1})^T \cdot \Sigma_z^{-1} \cdot (A^{-1}) \left(x-b-\frac{\mu_z}{A^{-1}}\right)}$$

$$= \frac{1}{|A|} \cdot (2\pi)^{-n/2} \cdot |\Sigma_z|^{-1/2} e^{-\frac{1}{2}(x-b-A\mu_z)^T \cdot (A \Sigma_z A^T)^{-1} \cdot (x-b-A\mu_z)}$$

Comparing with PDF of ~~x~~ X, ~~we ca~~

$$P_x(x) = (2\pi)^{-n/2} \cdot |\Sigma|^{-1/2} \cdot e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

we can see that, $b + A\mu_z = \mu$ and $A\Sigma_z A^T = \Sigma$

$$\Rightarrow A = \frac{\mu - b}{\mu_z}$$ ~~and~~

$$\Rightarrow b = \mu - A\mu_z$$