

## Homework #1

Name: Rohin Arora

NUID: 001302812

**Problem 1**

(10 %)

(1) Variance(x) -  $Var(x)$  ; Expectation(x) -  $E(x) - \mu$ 

$$\begin{aligned}
Var(x) &= E[(x - \mu)^2] \\
&= E[x^2 + \mu^2 - 2 * x * \mu] \\
&= E[x^2] + E[\mu^2] - 2 * E[x * \mu] && \text{linearity of expectation} \\
&= E[x^2] + \mu^2 - 2 * \mu * E[x] && \text{linearity of expectation} \\
&= E[x^2] + \mu^2 - 2 * \mu^2 \\
&= E[x^2] - \mu^2
\end{aligned}$$

■

(2)  $E[\vec{x}] = \vec{\mu}$  ; Covariance( $\vec{x}$ )= $cov(\vec{x})$ 

$$\begin{aligned}
cov(\vec{x}) &= E[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^T] \\
&= E[(\vec{x} * \vec{x}^T - \vec{x} * \vec{\mu}^T - \vec{\mu} * \vec{x}^T + \vec{\mu} * \vec{\mu}^T)] \\
&= E[\vec{x} * \vec{x}^T] - E[\vec{x} * \vec{\mu}^T] - E[\vec{\mu} * \vec{x}^T] + E[\vec{\mu} * \vec{\mu}^T] && \text{linearity of expectation} \\
&= E[\vec{x} * \vec{x}^T] - E[\vec{x}] * \vec{\mu}^T - \vec{\mu} * E[\vec{x}^T] + E[\vec{\mu} * \vec{\mu}^T] && \text{linearity of expectation} \\
&= E[\vec{x} * \vec{x}^T] - \vec{\mu} * \vec{\mu}^T - \vec{\mu} * \vec{\mu}^T + \vec{\mu} * \vec{\mu}^T && \text{linearity of expectation} \\
&= E[\vec{x} * \vec{x}^T] - \vec{\mu} * \vec{\mu}^T
\end{aligned}$$

■

**Problem 2**

(20 %)

(1)

- Class conditional probabilities shall sum to 1
- Let the constant of integration with class  $l$  be  $k_l$

$$\int_{-\infty}^{\infty} P(X|L=l)dx = k_l * \int_{-\infty}^{\infty} e^{-\frac{|x-a_l|}{b_l}} dx$$

- This integral should sum to 1.
- Consider a modified integral.

$$f(x) = k_l * \int_{a_l}^{\infty} e^{-\frac{x-a_l}{b_l}} dx$$

- The above integral should sum to 1/2.
- Continuing with this  $f(x)$

$$\begin{aligned} f(x) &= k_l * \int_{a_l}^{\infty} e^{-\frac{x-a_l}{b_l}} dx \\ &= k_l * e^{\frac{a_l}{b_l}} * \int_{a_l}^{\infty} e^{-\frac{x}{b_l}} dx \\ &= k_l * e^{\frac{a_l}{b_l}} * (-b_l) * e^{-\frac{x}{b_l}} \Big|_{a_l}^{\infty} \\ &= k_l * e^{\frac{a_l}{b_l}} * (-b_l) * (0 - e^{-\frac{a_l}{b_l}}) \\ &= k_l * e^{\frac{a_l}{b_l}} * (b_l) * (e^{-\frac{a_l}{b_l}}) \\ &= k_l * b_l \\ &= \frac{1}{2} \end{aligned}$$

$$k_l = \frac{1}{2 * b_l}$$

$$P(X|L=1) = \frac{1}{2 * b_1} * e^{-\frac{|x-a_1|}{b_1}}$$

$$P(X|L=2) = \frac{1}{2 * b_2} * e^{-\frac{|x-a_2|}{b_2}}$$

(2)

$$\begin{aligned} l(x) &= \ln P(X|L=1) - \ln P(X|L=2) \\ &= \ln \frac{1}{2 * b_1} - \ln \frac{1}{2 * b_2} - \frac{|x-a_1|}{b_1} + \frac{|x-a_2|}{b_2} \\ &= \frac{|x-a_2|}{b_2} - \frac{|x-a_1|}{b_1} + \ln \frac{b_2}{b_1} \end{aligned}$$

(3) Setting the values  $a_1 = 0, b_1 = 1, a_2 = 1, b_2 = 2$

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

This function is plotted in the following notebook- [Q2.Git.Repo](#)

**Problem 3**

(20 %)

Minimum probability of error classification rule implies 0 - 1 loss. Also, the 2 classes have equal priors. Therefore, the decision rule simplifies to maximum likelihood estimator. Suppose we have 2 labels, 1 and 2. In such a scenario, the decision rule is-

**Decide label 1 if  $P(x|L = 1) > P(x|L = 2)$  Else decide label 2** (contentions resolved arbitrarily)

Rephrasing-

When  $a < x < r$  - **class 1**

When  $r < x < b$  - **class 1** if  $1/(b - a) > 1/(t - r)$  else **class 2**

When  $b < x < t$  - **class 2**

$$P(x|L = 1) = \begin{cases} \frac{1}{b - a} & a \leq x \leq b \\ 0 & elsewhere \end{cases} \quad (0.1)$$

(0.2)

$$P(x|L = 2) = \begin{cases} \frac{1}{t - r} & r \leq x \leq t \\ 0 & elsewhere \end{cases} \quad (0.3)$$

(0.4)

$$\frac{P(x|L = 1)}{P(x|L = 2)} \underset{class2}{\overset{class1}{\geq}} 1$$

$$\frac{1}{b - a} \underset{class2}{\overset{class1}{\geq}} \frac{1}{t - r}$$

The accompanying code for this question with visual example is present here- [Q3\\_Git\\_Repo](#)

**Problem 4**

(30 %)

(1)

- We need to find a decision rule that achieves minimum probability of error. This implies we need to do maximum a posteriori estimation (0-1 loss). It's also given the classes have equal priors. This implies its a special case of MAP- maximum likelihood estimation. The decision rule in such a case is -

$$\frac{P(x|L=1)}{P(x|L=2)} \geq_{class2}^{class1} 1$$

$$P(x|L=1) \geq_{class2}^{class1} P(x|L=2)$$

$$P(x|L=1) \sim \mathcal{N}(0, 1)$$

$$P(x|L=2) \sim \mathcal{N}(\mu, \sigma^2)$$

$$\frac{1}{\sqrt{2\pi}} e^{-(x)^2/2} \geq_{class2}^{class1} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

$$\sigma * e^{-(x)^2/2} \geq_{class2}^{class1} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

Take log of both sides

$$\ln \sigma - \frac{x^2}{2} \geq_{class2}^{class1} -\frac{(x-\mu)^2}{2\sigma^2}$$

$$2\sigma^2 * \ln \sigma - \sigma^2 * x^2 \geq_{class2}^{class1} -(x-\mu)^2$$

$$2\sigma^2 * \ln \sigma - \sigma^2 * x^2 + (x-\mu)^2 \geq_{class2}^{class1} 0$$

$$x^2 + \mu^2 - 2x\mu - x^2\mu^2 + 2\sigma^2 \ln(\sigma) \geq_{class2}^{class1} 0$$

$$x^2(1-\sigma^2) - 2x\mu + \mu^2 + 2\sigma^2 \ln \sigma \geq_{class2}^{class1} 0$$

- The decision boundary is a parabola
- When this parabola is  $> 0$ , we decide class 1. Else we decide class 2

(2)

- If  $\mu = 1$  and  $\sigma^2 = 2$ , the decision boundary becomes-

$$-x^2 - 2x + 1 + 2\ln 2 \geq_{class2}^{class1} 0$$

- This is a parabola facing downwards. The zeros of this equation are  $x = -2.84019$  and  $x = 0.840189$
- **The decision rule is-**

Class 1	if $-2.84019 < x < 0.840189$
Class 2	otherwise

- Some math demonstrating the posterior calculation -

$$P(x|L = 1) \sim \mathcal{N}(0, 1)$$

$$P(x|L = 2) \sim \mathcal{N}(1, 2)$$

$$P(L = 1|x) = \frac{P(x|L = 1)P(L = 1)}{P(x)}$$

$$P(L = 2|x) = \frac{P(x|L = 2)P(L = 2)}{P(x)}$$

$$\text{where } P(x) = P(x|L = 1)P(L = 1) + P(x|L = 2)P(L = 2)$$

$$\begin{aligned} P(L = 1|x) &= \frac{P(x|L = 1)P(L = 1)}{P(x|L = 1)P(L = 1) + P(x|L = 2)P(L = 2)} \\ &= \frac{\mathcal{N}(0, 1)}{\mathcal{N}(0, 1) + \mathcal{N}(1, 2)} \end{aligned}$$

$$\begin{aligned} P(L = 2|x) &= \frac{P(x|L = 2)P(L = 2)}{P(x|L = 1)P(L = 1) + P(x|L = 2)P(L = 2)} \\ &= \frac{\mathcal{N}(1, 2)}{\mathcal{N}(0, 1) + \mathcal{N}(1, 2)} \end{aligned}$$

**Problem 5**

(20 %)

(1)

$$\vec{x} = \mathbf{A}\vec{z} + \vec{b} \quad \text{where } z \in \mathcal{N}(0, \mathbf{I}) \quad \text{and } \vec{x} \in \mathbb{R}^n$$

**By linearity of expectation**

$$\begin{aligned} E[\vec{x}] &= E[\mathbf{A}\vec{z} + \vec{b}] \\ &= E[\mathbf{A}\vec{z}] + E[\vec{b}] \quad \mathbf{A} \text{ and } \vec{b} \text{ are constants} \\ &= \mathbf{A}E[\vec{z}] + \vec{b} \end{aligned}$$

$$E[\vec{z}] = 0$$

$$\text{Therefore, } E[\vec{x}] = \vec{b}$$

$$\begin{aligned} \text{CoVar}(\vec{x}) &= \text{Covar}[\mathbf{A}\vec{z} + \vec{b}] \\ &= E[(\mathbf{A}\vec{z} + \vec{b} - \mu_{\mathbf{A}\vec{z} + \vec{b}})(\mathbf{A}\vec{z} + \vec{b} - \mu_{\mathbf{A}\vec{z} + \vec{b}})^T] \\ &= E[(\mathbf{A}\vec{z} + \vec{b} - \vec{b})(\mathbf{A}\vec{z} + \vec{b} - \vec{b})^T] \\ &= E[(\mathbf{A}\vec{z})(\mathbf{A}\vec{z})^T] \\ &= E[(\mathbf{A}\vec{z})(\vec{z}^T \mathbf{A}^T)] \\ &= \mathbf{A}E[(\vec{z})(\vec{z}^T)]\mathbf{A}^T \end{aligned}$$

$$E[(\vec{z})(\vec{z}^T)] = \text{CoVar}(\vec{z}) \quad \text{as } \vec{z} \text{ has zero mean}$$

$$\text{CoVar}(\vec{z}) = \mathbf{I} \quad \text{given in the question}$$

$$\begin{aligned} \text{Therefore, } \text{CoVar}(\vec{x}) &= \mathbf{A} * \mathbf{I} * \mathbf{A}^T \\ &= \mathbf{A}\mathbf{A}^T \end{aligned}$$

$$\vec{x} \sim \mathcal{N}(\vec{b}, \mathbf{A}\mathbf{A}^T)$$

■

(2)

Let the random vector be  $\vec{y}$ 

$$\begin{aligned} \vec{y} &= \mathbf{A}\vec{z} + \vec{b} \\ \vec{y} &\sim \mathcal{N}(\vec{\mu}, \Sigma) \end{aligned}$$

From previous sub problem,

$$\begin{aligned} \vec{b} &= \vec{\mu} \\ \mathbf{A}\mathbf{A}^T &= \Sigma \end{aligned}$$

$\mathbf{A}$  can be found by [Cholesky decomposition](#) of  $\Sigma$ . Indeed  $\Sigma$  is the covariance matrix of a Gaussian distribution, and hence positive semidefinite. Therefore cholesky decomposition can be applied here.

(3)

The code for this question can be found here- [Q5\\_Git\\_Repo](#)

## References and Acknowledgements

1. [Python](#)
2. [Scientific python stack](#)
3. [Variance](#)
4. [Latex](#)
5. [Lecture Notes](#)
6. [Iridescent](#)
7. [DON'T PANIC](#)