Problem 3.1

Suppose $\mathbf{X} = [X_1, X_2]^T$ is a Gaussian random vector with mean and covariance given by

$$\mathbf{\mu} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

Sketch the set of points such that $f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{12\pi} e^{-\frac{1}{8} x^2}$, where $f_{\mathbf{X}}(\mathbf{x})$ is the pdf of $\mathbf{X}$.

Problem 3.2

Suppose $\mathbf{X} = [X_1, X_2]^T$ is a Gaussian random vector with mean and covariance given by

$$\mathbf{\mu} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

Define $\Phi(z)$ as follows:

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} du$$

In terms of $\Phi(z)$, find the probability $\Pr\{ -1 < X_1 < 1, -1 < X_2 < 1 \}$.

Problem 3.3

Suppose that, for a particular classification problem, the observations are $\mathbf{x} \in \mathbb{R}^2$, and the labels are $y \in \{0, 1\}$. It just so happens that the correct label of every data point is as follows:

$$y^*(\mathbf{x}) = \begin{cases} 1 & \|\mathbf{x}\|_2 > 1.5 \\ 0 & \|\mathbf{x}\|_2 < 1.5 \end{cases} \quad (3.3-1)$$

Unfortunately, you aren’t allowed to use the correct labeling function. Instead, are required to learn a Gaussian classifier with the following form:

$$\hat{y}(\mathbf{x}) = \begin{cases} 1 & \frac{p_{\mathbf{X}|Y}(\mathbf{x}|1)}{p_{\mathbf{X}|Y}(\mathbf{x}|0)} > \eta \\ 0 & \frac{p_{\mathbf{X}|Y}(\mathbf{x}|1)}{p_{\mathbf{X}|Y}(\mathbf{x}|0)} < \eta \end{cases} \quad (3.3-2)$$

where $\eta$ is a parameter called the likelihood ratio threshold, and where the probability models for for both classes are zero-mean Gaussians, with different covariance matrices $\Sigma_0$ and $\Sigma_1$:

$$p_{\mathbf{X}|Y}(\mathbf{x}|y) = \mathcal{N}(\mathbf{x} | \mathbf{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma_y) \quad (3.3-3)$$
Suppose that \( \Sigma_0 \) and \( \Sigma_1 \) are known to be the identity matrix, and the scaled identity matrix, respectively:

\[
\Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}
\]

Find a value of \( \eta \) so that the Gaussian classifier (Eq. (3.3-2)) gives exactly the same decision boundary as the correct decision rule (Eq. (3.3-1)).

**Problem 3.4**

Suppose you have a scalar random variable \( X \), with training examples \([x_1, x_2, x_3, x_4] = [-1, 0, 1, 2]\). You want to try to model these data using a Gaussian mixture model, with the form

\[
p_X(x) = \sum_{k=0}^{1} c_k \mathcal{N}(x|\mu_k, \sigma_k^2)
\]

You have initial parameter estimates \( \mu_0 = 0, \mu_1 = 1, \sigma_0^2 = \sigma_1^2 = 1 \), and \( c_0 = c_1 = 0.5 \). Perform one iteration of EM training. What are the new values of \( \mu_0, \mu_1, \Sigma_0 \) and \( \Sigma_1 \), after one round of EM training? Write your answer in terms of the normal probability density functions \( \mathcal{N}(x|\mu, \sigma^2) \), for numerical values of \( x, \mu \) and \( \sigma^2 \).