ECE 417 Multimedia Signal Processing
Solutions to Homework 3

UNIVERSITY OF ILLINOIS
Department of Electrical and Computer Engineering

Assigned: Tuesday, 9/28/2021; Due: Tuesday, 10/5/2021
Reading: , Sections 1-3

Problem 3.1

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

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

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

Solution: Your sketch should be an ellipse, centered at the point $\vec{\mu} = [1, 0]^T$, with a radius of $\frac{\sigma_1}{2} = \frac{3}{2}$ in the x_1 direction, and with a radius of $\frac{\sigma_2}{2} = 1$ in the x_2 direction.

Problem 3.2

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

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

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 $\Pr\{-1 < X_1 < 1, -1 < X_2 < 1\}$.

Solution:

$$\begin{aligned}\Pr\{-1 < X_1 < 1, -1 < X_2 < 1\} &= \Pr\{-1 < X_1 < 1\}\Pr\{-1 < X_2 < 1\} \\ &= \left(\Phi\left(\frac{1-1}{3}\right) - \Phi\left(\frac{-1-1}{3}\right)\right)\left(\Phi\left(\frac{1-0}{2}\right) - \Phi\left(\frac{-1-0}{2}\right)\right)\end{aligned}$$

Problem 3.3

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

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

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

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

where η 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 Σ_0 and Σ_1:

$$p_{\vec{X}|Y}(\vec{x}|y) = \mathcal{N}\left(\vec{x}|\vec{\mu} = \left[\begin{array}{c} 0 \\ 0 \end{array}\right], \Sigma_y\right) \tag{3.3-3}$$

Suppose that Σ_0 and Σ_1 are known to be the identity matrix, and the scaled identity matrix, respectively:

$$\Sigma_0 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right], \quad \Sigma_1 = \left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array}\right]$$

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

Solution: The Gaussian classifier chooses $\hat{y} = 1$ under any of the following conditions, all of which are equivalent:

$$\begin{aligned} p_{\vec{X}|Y}(\vec{x}|1) &> \eta p_{\vec{X}|Y}(\vec{x}|0) \\ \ln p_{\vec{X}|Y}(\vec{x}|1) &> \ln p_{\vec{X}|Y}(\vec{x}|0) + \ln \eta \\ -\frac{1}{2}\ln|\Sigma_1| - \frac{1}{2}d_{\Sigma_1}(\vec{x},\vec{0}) &> -\frac{1}{2}\ln|\Sigma_0| - \frac{1}{2}d_{\Sigma_0}(\vec{x},\vec{0}) + \ln\eta \\ -\frac{1}{2}\ln(4) - \frac{1}{2}\left(\frac{x_1^2+x_2^2}{2}\right) &> -\frac{1}{2}\left(x_1^2+x_2^2\right) + \ln\eta \\ \|\vec{x}\|^2\left(\frac{1}{2}-\frac{1}{4}\right) &> \ln\eta + \frac{1}{2}\ln(4) \\ \|\vec{x}\| &> \sqrt{4\ln\eta + 2\ln(4)} \end{aligned}$$

This is exactly the same as Eq. (3.3-2) if $\sqrt{4\ln\eta + 2\ln(4)} = 1.5$, i.e.,

$$\eta = e^{\frac{(1.5)^2 - 2\ln(4)}{4}}$$

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

$$p_X(x) = \sum_{k=0}^{1} c_k \mathcal{N}(x|\mu_k, \sigma_k^2) \tag{3.4-1}$$

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 μ_0, μ_1, Σ_0 and Σ_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, μ and σ^2.

Solution: The gamma probabilities are

$$\gamma_1(1) = \frac{0.5\mathcal{N}(-1|1,1)}{0.5\mathcal{N}(-1|0,1) + 0.5\mathcal{N}(-1|1,1)}$$
$$\gamma_2(1) = \frac{0.5\mathcal{N}(0|1,1)}{0.5\mathcal{N}(0|0,1) + 0.5\mathcal{N}(0|1,1)}$$
$$\gamma_3(1) = \frac{0.5\mathcal{N}(1|1,1)}{0.5\mathcal{N}(1|0,1) + 0.5\mathcal{N}(1|1,1)}$$
$$\gamma_4(1) = \frac{0.5\mathcal{N}(2|1,1)}{0.5\mathcal{N}(2|0,1) + 0.5\mathcal{N}(2|1,1)}$$

and similarly for γ_0. The re-estimated parameters are written in terms of the gamma probabilities, so they will be

$$\mu_0 = -\left(\frac{\mathcal{N}(-1|0,1)}{\mathcal{N}(-1|0,1) + \mathcal{N}(-1|1,1)}\right) + \left(\frac{\mathcal{N}(1|0,1)}{\mathcal{N}(1|0,1) + \mathcal{N}(1|1,1)}\right) + 2\left(\frac{\mathcal{N}(2|0,1)}{\mathcal{N}(2|0,1) + \mathcal{N}(2|1,1)}\right)$$
$$\mu_1 = -\left(\frac{\mathcal{N}(-1|1,1)}{\mathcal{N}(-1|0,1) + \mathcal{N}(-1|1,1)}\right) + \left(\frac{\mathcal{N}(1|1,1)}{\mathcal{N}(1|0,1) + \mathcal{N}(1|1,1)}\right) + 2\left(\frac{\mathcal{N}(2|1,1)}{\mathcal{N}(2|0,1) + \mathcal{N}(2|1,1)}\right)$$
$$\sigma_0^2 = \left(\frac{\mathcal{N}(-1|0,1)}{\mathcal{N}(-1|0,1) + \mathcal{N}(-1|1,1)}\right) + \left(\frac{\mathcal{N}(1|0,1)}{\mathcal{N}(1|0,1) + \mathcal{N}(1|1,1)}\right) + 4\left(\frac{\mathcal{N}(2|0,1)}{\mathcal{N}(2|0,1) + \mathcal{N}(2|1,1)}\right)$$
$$\sigma_1^2 = 4\left(\frac{\mathcal{N}(-1|1,1)}{\mathcal{N}(-1|0,1) + \mathcal{N}(-1|1,1)}\right) + \left(\frac{\mathcal{N}(0|1,1)}{\mathcal{N}(0|0,1) + \mathcal{N}(0|1,1)}\right) + \left(\frac{\mathcal{N}(2|1,1)}{\mathcal{N}(2|0,1) + \mathcal{N}(2|1,1)}\right)$$