Gaussian Random Vectors, Gaussian Discriminant Analysis

Submit a PDF of your answers to Canvas.

1. If a random vector $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a Gaussian random vector, then the marginals are Gaussian, i.e, $p(x_j) \sim \mathcal{N}(\mu_j, \sigma_j)$. In this problem, we consider the converse. Let $X_1 \sim \mathcal{N}(0, 1)$ and define the random variable

$$X_2 = \begin{cases} X_1 & \text{with probability } 1/2 \\ -X_1 & \text{with probability } 1/2. \end{cases}$$

   a) What is $p(x_2)$?

   b) Sketch $p(x_1, x_2)$ in the x_1, x_2 plane (or describe the joint distribution of X_1, X_2).

   c) Is $[X_1 \; X_2]^T$ a Gaussian random vector?

2. For binary classification the MAP classification rule can be expressed using the log-likelihood ratio: if

$$\log \left(\frac{p(\boldsymbol{x}|y = 0)}{p(\boldsymbol{x}|y = 1)} \right) > 0 \tag{1}$$

   then $\widehat{y} = 0$, else $\widehat{y} = 1$.

   In *Gaussian discriminant analysis*, we assume that the class conditional distributions are Gaussian:

$$p(\boldsymbol{x}|y = 0) = \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}_0|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu}_0)^T \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_0) \right)$$

   and

$$p(\boldsymbol{x}|y = 1) = \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}_1|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_1) \right).$$

   a) In linear discriminant analysis, $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_1$, and (1) is equivalent to the following linear classification rule: if

$$\boldsymbol{w}^T \boldsymbol{x} > c$$

   then $\widehat{y} = 0$, else $\widehat{y} = 1$. Find $\boldsymbol{w}$ and c in terms of $\boldsymbol{\Sigma}_0$, $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$.

   b) In quadratic discriminant analysis, $\boldsymbol{\Sigma}_0 \neq \boldsymbol{\Sigma}_1$, and (1) is equivalent to the following quadratic rule: if

$$\boldsymbol{x}^T \boldsymbol{B} \boldsymbol{x} + \boldsymbol{w}^T \boldsymbol{x} > c$$

   then $\widehat{y} = 0$, else $\widehat{y} = 1$. Find an expression for $\boldsymbol{B}$, $\boldsymbol{w}$ and c in terms of $\boldsymbol{\Sigma}_0$, $\boldsymbol{\Sigma}_1$, $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$.