

**EE E6887 (Statistical Pattern Recognition)**  
**Solutions for homework 2**

P.1 Let  $p(\mathbf{x}|\omega_i) \sim N(\mu_i, \Sigma)$  for  $i = 1, 2$  in a two-category  $d$ -dimensional problem with the same covariance but arbitrary means and prior probabilities. Consider the square Mahalanobis distance:

$$r_i^2 = (\mathbf{x} - \mu_i)^t \Sigma^{-1} (\mathbf{x} - \mu_i)$$

(a) Show that the gradient of  $r_i^2$  is given by

$$\nabla r_i^2 = 2\Sigma^{-1}(\mathbf{x} - \mu_i)$$

(b) Show that any position on a given line through  $\mu_i$  the gradient  $\nabla r_i^2$  points in the same direction. Must this direction be parallel to that line?

(e) True or False: For a two-category problem involving normal densities with arbitrary means and covariances, and  $P(\omega_1) = P(\omega_2) = 1/2$ , the Bayes decision boundary consists of the set of points of equal Mahalanobis distance from the respective sample means. Explain.

**Answer:**

(a)  $\nabla r_i^2$  can be written as:

$$\begin{aligned} \nabla r_i^2 &= \left[ \Sigma^{-1}(\mathbf{x} - \mu_i) + ((\mathbf{x} - \mu_i)^t \Sigma^{-1})^t \right] \\ &= \left[ \Sigma^{-1}(\mathbf{x} - \mu_i) + (\Sigma^{-1})^t (\mathbf{x} - \mu_i) \right] \end{aligned}$$

Since the covariance matrix  $\Sigma$  is symmetry,  $(\Sigma^{-1}) = \Sigma^{-1}$ .

$$\nabla r_i^2 = 2\Sigma^{-1}(\mathbf{x} - \mu_i)$$

(b) In a given line through  $\mu_i$ , for any point  $\mathbf{x}$  on this line,  $\nabla r_i^2 = 2\Sigma^{-1}(\mathbf{x} - \mu_i) = 2\|\mathbf{x} - \mu_i\| \Sigma^{-1} \left( \frac{\mathbf{x} - \mu_i}{\|\mathbf{x} - \mu_i\|} \right)$ .  $\Sigma^{-1} \left( \frac{\mathbf{x} - \mu_i}{\|\mathbf{x} - \mu_i\|} \right)$  is a constant for any point on the line. Thus for all points on this line,  $\nabla r_i^2$  points to the same direction.

When  $\Sigma = \sigma^2 \mathbf{I}$ , for any  $\sigma > 0$ , the direction parallel to the line. Otherwise, the direction doesn't parallel to the line.

(e) This statement is true. When  $P(\omega_1) = P(\omega_2) = 1/2$  and the covariance matrix are equal for the two classes, the Bayes decision function turns to:

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \mu_i)^t \Sigma^{-1}(\mathbf{x} - \mu_i) = -\frac{1}{2}r_i^2$$

The decision boundary corresponds to the points which have

$$g_1(\mathbf{x}) = g_2(\mathbf{x}) \Rightarrow r_1^2 = r_2^2$$

Thus the Bayes decision boundary consists of the set of points of equal Mahalanobis distance from the respective sample means. Otherwise, this will be false if the covariance matrix are not equal.

P.2 Suppose we have three categories in two dimensions with the following distributions:

- $p(\mathbf{x}|\omega_1) \sim N(\mathbf{0}, \mathbf{I})$
- $p(\mathbf{x}|\omega_2) \sim N([1, 1]^t, \mathbf{I})$
- $p(\mathbf{x}|\omega_3) \sim \frac{1}{2}N([0.5, 0.5]^t, \mathbf{I}) + \frac{1}{2}N([-0.5, 0.5]^t, \mathbf{I})$

with  $P(\omega_i) = 1/3$ ,  $i = 1, 2, 3$ .

- (a) By explicit calculation of posterior probabilities, classify the point  $\mathbf{x} = [0.3, 0.3]^t$  for minimum probability error.
- (b) Suppose that for a particular test point the first feature is missing. That is, classify  $\mathbf{x} = [*, 0.3]^t$ .
- (c) Suppose that for a particular test point the second feature is missing. That is, classify  $\mathbf{x} = [0.3, *]^t$ .

**Answer:**

(a)  $p(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{p(\mathbf{x})}$ . Since  $P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$ , we just need to compare the likelihood  $p(\mathbf{x}|\omega_i)$ .

$$p(\mathbf{x}|\omega_1) = \frac{1}{2\pi} \exp \left[ -\frac{1}{2}[0.3, 0.3][0.3, 0.3]^t \right] = 0.1455$$

$$p(\mathbf{x}|\omega_2) = \frac{1}{2\pi} \exp \left[ -\frac{1}{2}[-0.7, -0.7][-0.7, -0.7]^t \right] = 0.0975$$

$$\begin{aligned}
p(\mathbf{x}|\omega_3) &= \frac{1}{4\pi} \exp\left[-\frac{1}{2}[-0.2, -0.2][-0.2, -0.2]^t\right] \\
&\quad + \frac{1}{4\pi} \exp\left[-\frac{1}{2}[0.8, -0.2][0.8, -0.2]^t\right] \\
&= 0.1331
\end{aligned}$$

Since  $p(\mathbf{x}|\omega_1)$  is the largest likelihood,  $\mathbf{x}$  is classified to be from class 1.

(b) Assume that  $\mathbf{x} = [x_1, x_2]^t$ ,

$$p(\omega_i|x_2) = \frac{\int p(x_1, x_2|\omega_i)P(\omega_i)dx_1}{p(x_2)}$$

Still since  $P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$ , we just need to compare  $\int p(x_1, x_2|\omega_i)dx_1$ .

$$\begin{aligned}
\int p(x_1, x_2|\omega_1)dx_1 &= \frac{1}{2\pi} \int \exp\left[-\frac{1}{2}(x_1^2 + 0.09)\right] dx_1 \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{0.09}{2}\right) = 0.3813 \\
\int p(x_1, x_2|\omega_2)dx_1 &= \frac{1}{2\pi} \int \exp\left[-\frac{1}{2}\left((x_1 - 1)^2 + 0.49\right)\right] dx_1 \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{0.49}{2}\right) = 0.3122 \\
\int p(x_1, x_3|\omega_2)dx_1 &= \frac{1}{4\pi} \int \exp\left[-\frac{1}{2}\left((x_1 - 0.5)^2 + 0.04\right)\right] dx_1 \\
&\quad + \frac{1}{4\pi} \int \exp\left[-\frac{1}{2}\left((x_1 + 0.5)^2 + 0.04\right)\right] dx_1 \\
&= \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{0.04}{2}\right) + \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{0.04}{2}\right) \\
&= 0.3910
\end{aligned}$$

Thus  $\mathbf{x}$  is classified to be from class 3.

(c) Still  $\mathbf{x} = [x_1, x_2]^t$ ,

$$p(\omega_i|x_1) = \frac{\int p(x_1, x_2|\omega_i)P(\omega_i)dx_2}{p(x_1)}$$

We need to compare  $\int p(x_1, x_2|\omega_i)dx_2$ .

$$\begin{aligned}\int p(x_1, x_2|\omega_1)dx_2 &= \frac{1}{2\pi} \int \exp\left[-\frac{1}{2}(x_2^2 + 0.09)\right] dx_2 \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{0.09}{2}\right) = 0.3813 \\ \int p(x_1, x_2|\omega_2)dx_2 &= \frac{1}{2\pi} \int \exp\left[-\frac{1}{2}\left((x_2 - 1)^2 + 0.49\right)\right] dx_2 \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{0.49}{2}\right) = 0.3122 \\ \int p(x_1, x_3|\omega_2)dx_2 &= \frac{1}{4\pi} \int \exp\left[-\frac{1}{2}\left((x_2 - 0.5)^2 + 0.04\right)\right] dx_2 \\ &\quad + \frac{1}{4\pi} \int \exp\left[-\frac{1}{2}\left((x_2 - 0.5)^2 + 0.64\right)\right] dx_2 \\ &= \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{0.04}{2}\right) + \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{0.64}{2}\right) \\ &= 0.3403\end{aligned}$$

Thus in this case  $\mathbf{x}$  is classified to be in class 1.

The solution of this problem can be simplified by noticing the special covariance structure of the Gaussian pdf's. Identity covariance matrix implies independence of the random variables for each dimension. Therefore, a 2D Gaussian pdf can be factorized into the product of two 1D Gaussian pdf's. The marginalization step becomes much simplified after the factorization because the integration of a pdf simply gives us one.