HST.582J / 6.555J / 16.456J Biomedical Signal and Image Processing
Spring 2007

HST-582J/6.555J/16.456J - Biological Signal and Image Processing - Spring 2007

Problem Set 5
Due April 26, 2007

Problem 1

Decision Boundaries: Two-dimensional Gaussian Case

The optimal **Bayesian** decision rule can be written:

$$\phi(x) = \begin{cases} 1 & ; \ \frac{p_1(x)}{p_0(x)} > \frac{P_0}{P_1} \\ 0 & ; \ \text{otherwise} \end{cases}$$

It is sometimes useful to express the decision in the log domain, or equivalently

$$\phi(x) = \begin{cases} 1 & ; \ \ln(p_1(x)) - \ln(p_0(x)) > \ln\left(\frac{P_0}{P_1}\right) \\ 0 & ; \ \text{otherwise} \end{cases}$$

The **decision boundary** is defined as the locus of points, x, where the ratios are equal, that is

$$\ln(p_1(x)) - \ln(p_0(x)) = \ln\left(\frac{P_0}{P_1}\right)$$

If $x = [x_1, x_2]$ is a two-dimensional Gaussian variable, its PDF is written:

$$p_i(x) = \frac{1}{2\pi |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(x - m_i)^T \Sigma_i^{-1} (x - m_i)\right)$$

where m_i, Σ_i are the class-conditional means and covariances, respectively. Plugging this into the log form of the decision boundary above yields:

$$-\frac{1}{2}(x - m_1)^T \Sigma_1^{-1}(x - m_1) + \frac{1}{2}(x - m_0)^T \Sigma_0^{-1}(x - m_0) + \frac{1}{2}\ln\left(\frac{|\Sigma_0|}{|\Sigma_1|}\right) = \ln\left(\frac{P_0}{P_1}\right)$$

Suggestion: You may want to do part (d) of this problem first as a way of checking your answers to the first three parts although it is not necessary to do so.

a) Suppose

$$P_1 = P_0 = \frac{1}{2} \quad x = \begin{bmatrix} x_1 \\ x_0 \end{bmatrix} \quad m_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad m_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\Sigma_1 = \Sigma_0 = \begin{bmatrix} 1 & \frac{9}{10} \\ \frac{9}{10} & 1 \end{bmatrix} \quad \Sigma_1^{-1} = \Sigma_0^{-1} = \begin{bmatrix} \frac{100}{19} & \frac{-90}{19} \\ \frac{-90}{19} & \frac{100}{19} \end{bmatrix} \quad |\Sigma_1| = |\Sigma_0| = \frac{19}{100}$$

express the decision boundary in the form $x_2 = f(x_1)$.

b) If we keep all values from part (a), but set

$$\frac{P_0}{P_1} = \exp\left(-\frac{1}{2}\right)$$

how does the decision boundary change in terms of its relationship to m_1 and m_0? Express the decision boundary in the form $x_2 = f(x_1)$ using the new value of the ratio of P_0 to P_1 and the means and covariances from part (a).

c) Suppose now that

$$\Sigma_1 = \Sigma_0 = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$$

where $|r| < 1$ (which is simply a constraint to ensure Σ_i is a valid covariance matrix) keeping all other relevant terms from part (a). How does this change the decision boundary as compared to the result of part (a)?

d) Now let

$$\Sigma_0 = \begin{bmatrix} 1 & \frac{-9}{10} \\ \frac{-9}{10} & 1 \end{bmatrix} \quad \Sigma_0^{-1} = \begin{bmatrix} \frac{100}{19} & \frac{90}{19} \\ \frac{90}{19} & \frac{100}{19} \end{bmatrix} \quad |\Sigma_0| = \frac{19}{100}$$

setting all other parameters, **except** P_1 and P_0, the same as in part (a). Use matlab **contour** function to plot the decision boundary as a function of the ratio of prior probabilities of each class for the values $P_0/P_1 = [1/4, 1/2, 1, 2, 4]$. Here is *some* of the code you will need (where "function" is the left side of the decision boundary equation, $\ln(p_1(x)) - \ln(p_0(x))$):

```
[x1,x2] = meshgrid(-4:0.1:4,-4:0.1:4);
d = function(x1,x2);
[c,h] = contour(x1,x2,d,log([1/4,1/2,1,2,4]));
clabel(c,h);
```

Problem 2

Suggestion: read the entire question, the answer can be stated in one sentence with no calculations.

Suppose you have a 3-dimensional measurement vector $x = [x_1, x_2, x_3]$ for a binary classification problem where $0 < P_1 < 1$ (i.e. it is strictly greater then 0 and less then 1). Recall that the class-conditional marginal distribution of x_1, x_2 is

$$\begin{aligned} p_i(x_1, x_2) &= \int p_i(x_1, x_2, x_3)\, dx_3 \\ &= \int p_i(x_1, x_2 | x_3)\, p_i(x_3)\, dx_3 \end{aligned}$$

and that the unconditioned marginal density of any single measurement is

$$p(x_k) = \sum_{i=0}^{1} P_i p_i(x_k)$$

where $k = 1$, 2, or 3.

Now consider 2 *different* decision functions. The first $\phi(x_1, x_2, x_3)$ is the optimal classifier using the full measurement vector $[x_1, x_2, x_3]$, while the second $\varphi(x_1, x_2)$ is the optimal classifier using only $[x_1, x_2]$. In general the probability of error using $\phi(x_1, x_2, x_3)$ will be lower then when using $\varphi(x_1, x_2)$ (i.e. when we ignore the third measurement). State a condition under which both classifiers will achieve the same probability of error.