Answer Key for Homework 2

CS 591Q/791V - Pattern Recognition
Instructor: Dr. Arun Ross
Posted on: March 20, 2009

---

1. The Bayes classification rule for an arbitrary pattern x:

   Assign x to C_1 if $P(C_1|x) > P(C_2|x)$; else assign x to C_2.

   If $P(C_1|x) > P(C_2|x)$, then

$$\frac{p(x|C_1).p(C_1)}{p(x)} > \frac{p(x|C_2).p(C_2)}{p(x)} \qquad \text{(By Bayes formula)}$$
$$\Rightarrow p(x|C_1) > p(x|C_2) \qquad \text{(Since } p(C_1) = p(C_2)\text{)}$$
$$\Rightarrow e^{-\lambda_1}\frac{\lambda_1^x}{x!} > e^{-\lambda_2}\frac{\lambda_2^x}{x!}$$
$$\Rightarrow [-\lambda_1 + x\ln(\lambda_1) - \ln(x!)] > [-\lambda_2 + x\ln(\lambda_2) - \ln(x!)] \qquad \text{(Applying ln on both sides)}$$
$$\Rightarrow x[\ln(\lambda_1) - \ln(\lambda_2)] > \lambda_1 - \lambda_2$$
$$\Rightarrow x > \frac{(\lambda_1 - \lambda_2)}{\ln(\lambda_1/\lambda_2)}$$

   Thus, the required classification rule is:

   Assign x to C_1 if $x > \frac{(\lambda_1-\lambda_2)}{\ln(\lambda_1/\lambda_2)}$; else assign x to C_2. $\square$

2. The posterior probability for each of the 3 classes, $P(C_i|\boldsymbol{x}), i = 1, 2, 3$, has to be explicitly computed. The Bayes decision rule is:

   Assign $\boldsymbol{x}$ to C_k if $P(C_k|\boldsymbol{x}) \geq P(C_j|\boldsymbol{x}), \quad \forall j, k = 1, 2, 3.$

   We first compute the likelihoods:

$$p\left(\boldsymbol{x} = (0.3, 0.3)^t|C_1\right) = \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5\begin{bmatrix}0.3 & 0.3\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.3\\ 0.3\end{bmatrix}\right)$$
$$\Rightarrow \quad p(\boldsymbol{x} = (0.3, 0.3)^t|C_1) = 0.14. \tag{1}$$

$$p\left(\boldsymbol{x} = (0.3, 0.3)^t|C_2\right) = \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5\begin{bmatrix}0.3-1 & 0.3-1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.3-1\\ 0.3-1\end{bmatrix}\right)$$
$$\Rightarrow \quad p(\boldsymbol{x} = (0.3, 0.3)^t|C_2) = 0.09. \tag{2}$$

$$p(\boldsymbol{x} = (0.3, 0.3)^t|C_3) = 0.5 \times \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5\begin{bmatrix}0.3-0.5 & 0.3-0.5\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.3-0.5\\ 0.3-0.5\end{bmatrix}\right)$$
$$+0.5 \times \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5\begin{bmatrix}0.3+0.5\\ 0.3-0.5\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.3+0.5\\ 0.3-0.5\end{bmatrix}\right)$$
$$\Rightarrow \quad p(\boldsymbol{x} = (0.3, 0.3)^t|C_3) = 0.13. \tag{3}$$

By Bayes formula,

$$P(C_i|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|C_i) \times P(C_i)}{p(x)} \tag{4}$$

$$\text{where} \quad p(x) = \sum_{k=1}^{3} P(\boldsymbol{x}|C_k).P(C_k)$$

Hence,

$$p(x) = (0.14 + 0.09 + 0.13) \times \frac{1}{3} = 0.12 \tag{5}$$

Thus, based on equations 1, 2, 3 and 5 we have:

$$\begin{aligned}
P(C_1|(0.3, 0.3)^t) &= \frac{0.14 \times \frac{1}{3}}{0.12} = 0.38. \\
P(C_2|(0.3, 0.3)^t) &= \frac{0.09 \times \frac{1}{3}}{0.12} = 0.25. \\
P(C_3|(0.3, 0.3)^t) &= \frac{0.13 \times \frac{1}{3}}{0.12} = 0.36. \\
\Rightarrow \boldsymbol{x} &= (0.3, 0.3)^t \in C_1.\square
\end{aligned}$$

3. Substitute the given values for $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{x} = (x_1, x_2, x_3)^t$ in the equation for the multivariate normal distribution and you will obtain $p([0.5, 0, 1]^t) = 0.0082$. $\square$

4. (a) To compute the Bayes decision boundary:

$$\begin{aligned}
& P(C_1|\boldsymbol{x}) = P(C_2|\boldsymbol{x}) \\
& \Rightarrow \frac{p(\boldsymbol{x}|C_1).p(C_1)}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{x}|C_2).p(C_2)}{p(\boldsymbol{x})} \qquad \text{(By Bayes formula)} \\
& \Rightarrow p(\boldsymbol{x}|C_1) = p(\boldsymbol{x}|C_2) \qquad \text{(Since } p(C_1) = p(C_2)\text{)} \\
& \Rightarrow \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5 \begin{bmatrix} x_1 - 1 & x_2 - 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix}\right) \\
& \Rightarrow \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 1 & x_2 - 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} \qquad \text{(After applying ln on both sides)} \\
& \Rightarrow x_1^2 + x_2^2 = (x_1 - 1)^2 + (x_2 - 1)^2 \\
& \Rightarrow x_1^2 + x_2^2 = (x_1^2 - 2x_1 + 1) + (x_2^2 - 2x_2 + 1) \\
& \Rightarrow x_1 + x_2 - 1 = 0.\square
\end{aligned}$$

The Bayes decision rule:

$$\begin{aligned}
&\text{If} \quad x_1 + x_2 - 1 < 0, \quad \text{then} \quad \boldsymbol{x} = (x_1, x_2)^t \in C_1; \\
&\text{If} \quad x_1 + x_2 - 1 \geq 0, \quad \text{then} \quad \boldsymbol{x} = (x_1, x_2)^t \in C_2.\square
\end{aligned} \tag{6}$$

See Figure 1 for the decision boundary.

(b) Generate 100 points for each class using the *mvnrnd* function. Plug-in each point in equation (6) to determine its class as predicted by the Bayes classifier. The empirical error rate denotes the fraction of points that are incorrectly classified by this classifier.

(c) Same procedure as (b) above.

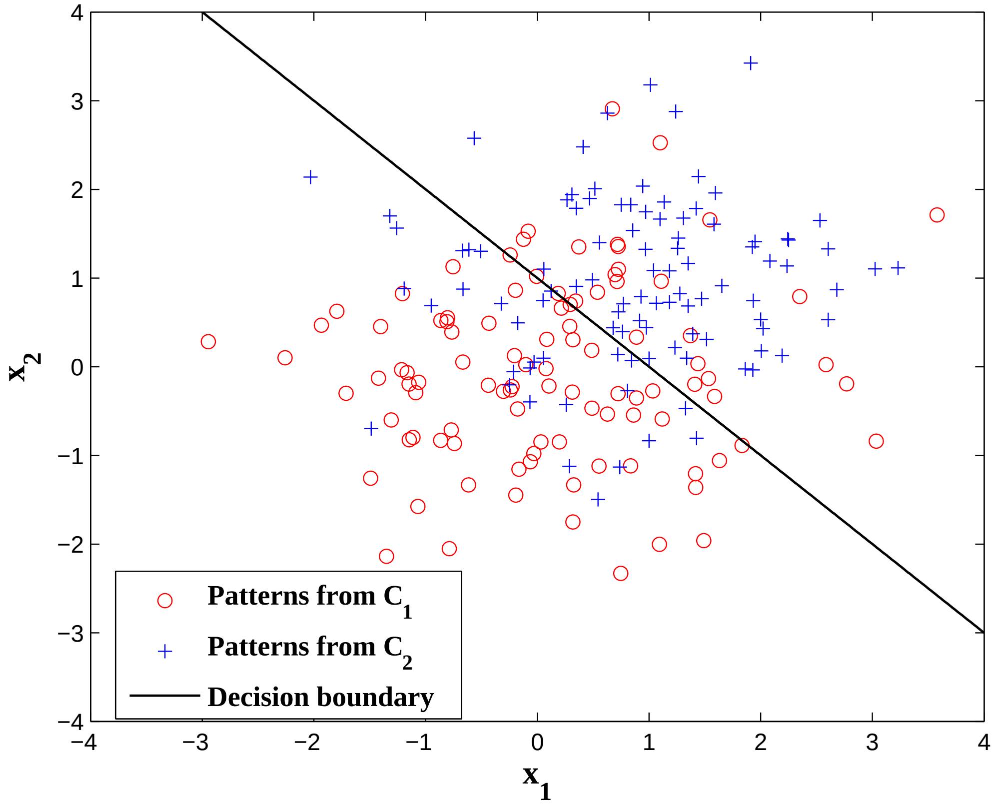

Figure 1: The linear decision boundary of problem 4. The patterns from the two classes, C_1 and C_2, are sampled from $N\left((0,0)^t, I\right)$ and $N\left((1,1)^t, I\right)$, respectively.

5. (c) To compute the Bayes decision boundary:

$$
\begin{aligned}
&P(C_1|\boldsymbol{x}) = P(C_2|\boldsymbol{x}) \\
&\Rightarrow \frac{p(\boldsymbol{x}|C_1).p(C_1)}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{x}|C_2).p(C_2)}{p(\boldsymbol{x})} \qquad \text{(By Bayes formula)} \\
&\Rightarrow p(\boldsymbol{x}|C_1) = p(\boldsymbol{x}|C_2) \qquad \text{(Since } p(C_1) = p(C_2)\text{)} \\
&\Rightarrow \frac{1}{2\pi} \times \frac{1}{\sqrt{4}} \times \exp\left(-0.5 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \frac{1}{2\pi} \times \frac{1}{1} \times \exp\left(-0.5 \begin{bmatrix} x_1 - 1 & x_2 - 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \end{bmatrix}\right) \\
&\Rightarrow -\ln(2) - 0.5 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -0.5 \begin{bmatrix} x_1 - 1 & x_2 - 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \end{bmatrix} \qquad \text{(Applying ln on both sides)} \\
&\Rightarrow \ln(2) + 0.5 \begin{bmatrix} \frac{1}{2}x_1 & \frac{1}{2}x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.5 \begin{bmatrix} x_1 - 1 & x_2 - 1 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \end{bmatrix} \\
&\Rightarrow 0.69 + 0.25x_1^2 + 0.25x_2^2 = 0.5x_1^2 - x_1 + 0.5 + 0.5x_2^2 - 2x_2 + 2 \\
&\Rightarrow 0.25x_1^2 + 0.25x_2^2 - x_1 - 2x_2 + 1.81 = 0 \\
&\Rightarrow x_1^2 + x_2^2 - 4x_1 - 8x_2 + 7.24 = 0. \square
\end{aligned}
$$

This is the equation of an ellipse. Thus, the decision boundary is an ellipse. Points within the ellipse are classified as C_2 while those outside the ellipse are classified as C_1. The decision rule will be (see Figure 2 in next page):

$$
\begin{aligned}
&\text{If} \quad x_1^2 + x_2^2 - 4x_1 - 8x_2 + 7.24 > 0, \quad \text{then} \quad \boldsymbol{x} = (x_1, x_2)^t \in C_1; \\
&\text{If} \quad x_1^2 + x_2^2 - 4x_1 - 8x_2 + 7.24 \leq 0, \quad \text{then} \quad \boldsymbol{x} = (x_1, x_2)^t \in C_2. \square
\end{aligned} \tag{7}
$$

For part (b), estimate the parameters from the training data and follow the same procedure indicated above.

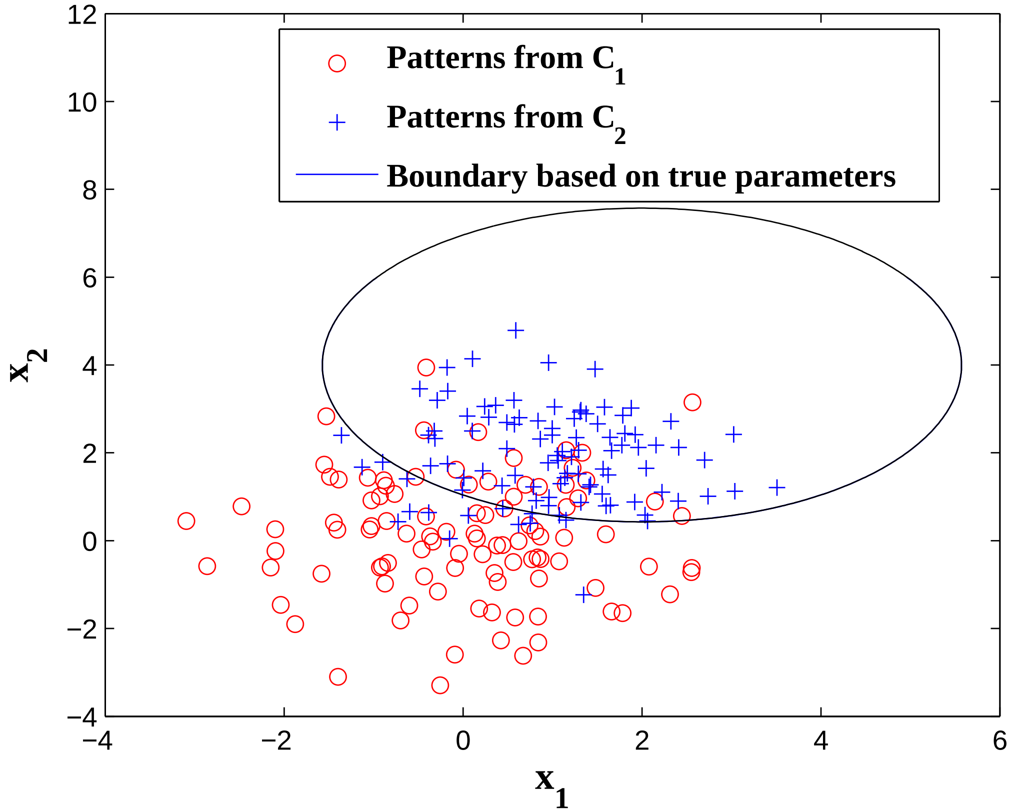

Figure 2: The quadratic decision boundary of problem 5. The sample patterns from the two classes, C_1 and C_2, are obtained from $N([0,0]^t, 2I)$ and $N([1,2]^t, I)$, respectively.