Solutions

Theory Sheet 4

Solution T-3.1: Decision boundaries

The decision rule for a MAP classifier is given by $\omega_{opt} = \arg\max_{\omega_k} P(\omega_k|\mathbf{x})$. A corresponding discriminant function is (see lecture):

$$\begin{aligned} g_i(\mathbf{x}) &= P(\omega_i|\mathbf{x}) \\ &= p(\mathbf{x}|\omega_i)P(\omega_i) \end{aligned}$$

Since the log-function is monotonically increasing the following discriminant function has the same classification result:

$$g_i(\mathbf{x}) = \log p(\mathbf{x}|\omega_i) + \log P(\omega_i)$$

The decision boundary for a two-class classification task is given by $g_i(\mathbf{x}) = g_j(\mathbf{x})$ and thus

$$\log p(\mathbf{x}|\omega_i) + \log P(\omega_i) = \log p(\mathbf{x}|\omega_j) + \log P(\omega_j)$$

If the priors have equal values (special condition) the log-likelihood ratio is simply given by:

$$\begin{aligned} \log p(\mathbf{x}|\omega_i) - \log p(\mathbf{x}|\omega_j) &= 0 \\ \log \frac{p(\mathbf{x}|\omega_i)}{p(\mathbf{x}|\omega_j)} &= 0 \qquad (q.e.d.) \end{aligned}$$

Solution T-3.2: Decision boundary for two-dimensional Gaussian data

The parameters of the two classes are found to be:

$$\mu_1 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}; \qquad \Sigma_1 = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}; \qquad \mu_2 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}; \qquad \Sigma_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

The inverse matrices are then

$$\Sigma_1^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \qquad \text{and} \qquad \Sigma_2^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

and the determinants are given by

$$\det \Sigma_1 = 1; \qquad \det \Sigma_2 = 4$$

We start with the discriminant functions (cf. Chapter 2.3 of the lecture):

$$\begin{aligned} g_1(\mathbf{x}) &= \ln p(\mathbf{x}|\omega_1) + \ln P(\omega_1) \\ g_2(\mathbf{x}) &= \ln p(\mathbf{x}|\omega_2) + \ln P(\omega_2) \end{aligned}$$

and set

$$g_1(\mathbf{x}) = g_2(\mathbf{x})$$

With equal prior probabilities we obtain

$$\ln p(\mathbf{x}|\omega_1) = \ln p(\mathbf{x}|\omega_2) \tag{1}$$

Since the given 2D data is normally distributed we use the general multivariate normal density to model the two likelihoods, i.e.:

$$p(\mathbf{x}|\omega) = \frac{1}{\sqrt{((2\pi)^2 \cdot \det \Sigma)}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^t \Sigma^{-1}(\mathbf{x}-\mu)}$$

Note, that the data is statistically independent and may therefore also be modeled by a Gaussian density with independent components.
Taking the natural logarithm

$$\ln p(\mathbf{x}|\omega) = -\frac{1}{2} \ln[(2\pi)^2 \cdot \det \Sigma] - \frac{1}{2}(\mathbf{x}-\mu)^t \Sigma^{-1}(\mathbf{x}-\mu) \cdot \ln e$$

and entering the known parameters μ and Σ of the two classes leads to

$$\begin{aligned} \ln p(\mathbf{x}|\omega_1) &= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(\det \Sigma_1) + (\mathbf{x}-\mu_1)^t \Sigma_1^{-1}(\mathbf{x}-\mu_1)\right\} \\ &= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(1) + (\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 3 \\ 6 \end{pmatrix})^t \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} (\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 3 \\ 6 \end{pmatrix})\right\} \\ &= -\frac{1}{2}\left\{\ln(2\pi)^2 + (x_1 - 3 \quad x_2 - 6) \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} x_1 - 3 \\ x_2 - 6 \end{pmatrix}\right\} \\ &= -\frac{1}{2}\left\{\ln(2\pi)^2 + (2(x_1-3) \quad \frac{1}{2}(x_2-6)) \begin{pmatrix} x_1 - 3 \\ x_2 - 6 \end{pmatrix}\right\} \\ &= -\frac{1}{2}\left\{\ln(2\pi)^2 + 2(x_1-3)^2 + \frac{1}{2}(x_2-6)^2\right\} \end{aligned} \tag{2}$$

and

$$\begin{aligned}
\ln p(\mathbf{x}|\omega_2) &= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(\det\Sigma_2) + (\mathbf{x}-\mu_2)^t\Sigma_2^{-1}(\mathbf{x}-\mu_2)\right\} \\
&= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(4) + (\begin{pmatrix}x_1\\x_2\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix})^t \begin{pmatrix}1/2 & 0\\ 0 & 1/2\end{pmatrix}(\begin{pmatrix}x_1\\x_2\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix})\right\} \\
&= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(4) + (x_1-3 \quad x_2+2)\begin{pmatrix}1/2 & 0\\ 0 & 1/2\end{pmatrix}\begin{pmatrix}x_1-3\\x_2+2\end{pmatrix}\right\} \\
&= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(4) + (\frac{1}{2}(x_1-3) \quad \frac{1}{2}(x_2+2))\begin{pmatrix}x_1-3\\x_2+2\end{pmatrix}\right\} \\
&= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(4) + \frac{1}{2}(x_1-3)^2 + \frac{1}{2}(x_2+2)^2\right\} \qquad (3)
\end{aligned}$$

With (2) and (3) Equation (1) becomes

$$\begin{aligned}
\ln p(\mathbf{x}|\omega_1) &= \ln p(\mathbf{x}|\omega_2) \\
-\frac{1}{2}\left\{\ln(2\pi)^2 + 2(x_1-3)^2 + \frac{1}{2}(x_2-6)^2\right\} &= -\frac{1}{2}\left\{\ln(2\pi)^2 + \ln(4) + \frac{1}{2}(x_1-3)^2 + \frac{1}{2}(x_2+2)^2\right\} \\
-(x_1-3)^2 - \frac{1}{4}(x_2-6)^2 &= -\frac{1}{2}\ln(4) - \frac{1}{4}(x_1-3)^2 - \frac{1}{4}(x_2+2)^2 \\
-x_1^2 + 6x_1 - 9 - \frac{1}{4}x_2^2 + 3x_2 - 9 &= -\frac{1}{2}\ln(4) - \frac{1}{4}x_1^2 + \frac{3}{2}x_1 - \frac{9}{4} - \frac{1}{4}x_2^2 - x_2 - 1 \\
4x_2 &= \frac{3}{4}x_1^2 - \frac{9}{2}x_1 + \frac{59}{4} - \frac{1}{2}\ln(4) \\
x_2 &= \frac{3}{16}x_1^2 - \frac{9}{8}x_1 + \frac{59}{16} - \frac{1}{8}\ln(4) \\
x_2 &= 0.1875x_1^2 - 1.125x_1 + 3.514
\end{aligned} \qquad (4)$$

This equation describes a parabola shown in Figure 1. Note that the decision boundary lies slightly lower than the point midway between the two means. This is because for the ω_1 distribution, the probability distribution is squeezed in the x_1 direction more so than for the ω_2 distribution.

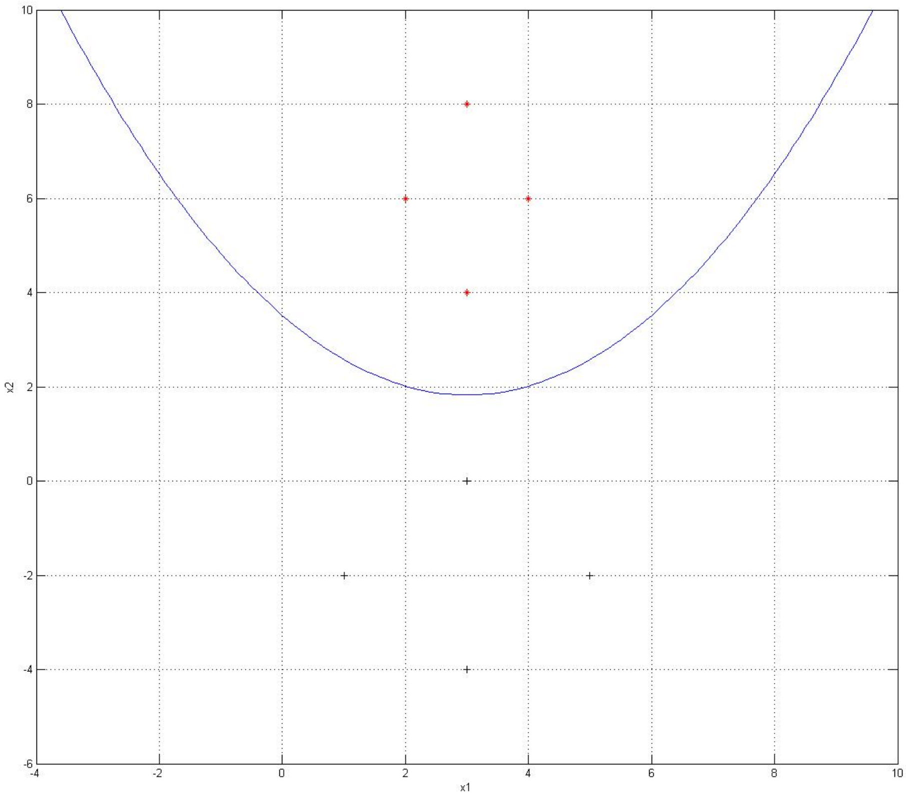

Figure 1. Computed Bayes decision boundary with the two distributions.