Open your mind. LUT.
Lappeenranta University of Technology

LUT Machine Vision and Pattern Recognition 2015-10-05

BM40A0700 Pattern Recognition

Lasse Lensu

Exercise 4 solutions: Bayesian classification and normal distribution

3. Classification in two dimensions (2 points): The following three-class classification problem was considered. Each class is normally distributed with equal covariance matrix and a priori probabilities

$$\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad P(\omega_1) = P(\omega_2) = P(\omega_3), \qquad \mu_1 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad \mu_3 = \begin{pmatrix} 11 \\ -2 \end{pmatrix} \tag{1}$$

A multivariate l-dimensional Gaussian probability density function (PDF)

$$N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \sim p(\mathbf{x}) = \frac{1}{(2\pi)^{l/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right] \tag{2}$$

where $\boldsymbol{\mu} = \mathrm{E}[\mathbf{x}]$ is the mean (vector), and $\boldsymbol{\Sigma}$ is the $l \times l$ covariance matrix defined as $\boldsymbol{\Sigma} = \mathrm{E}\left[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^\intercal\right]$.

By defining a logarithmic discriminant function g_i for each class ω_i based on the multivariate normal distribution, we get

$$\begin{aligned}
g_i(\mathbf{x}) &= \ln(p(\mathbf{x}|\omega_i)P(\omega_i)) & (3)\\
&= \ln p(\mathbf{x}|\omega_i) + \ln P(\omega_i) & (4)\\
&= -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^\intercal \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{l}{2}\ln 2\pi - \frac{1}{2}\ln|\boldsymbol{\Sigma}_i| + \ln P(\omega_i) & (5)\\
&= \mathbf{x}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{x} - 2(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i)^\intercal\mathbf{x} + \boldsymbol{\mu}_i^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \frac{l}{2}\ln 2\pi - \frac{1}{2}\ln|\boldsymbol{\Sigma}_i| + \ln P(\omega_i). & (6)
\end{aligned}$$

Some of the terms are equal for all classes, and for the case where $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, the decision surfaces are

$$\begin{aligned}
g_i(\mathbf{x}) &= (\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i)^\intercal\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_i^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \ln P(\omega_i)\\
&= \mathbf{w}_i^\intercal\mathbf{x} + w_{i0}\\
\mathbf{w}_i &= \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\\
w_{i0} &= -\frac{1}{2}\boldsymbol{\mu}_i^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \ln P(\omega_i). & (7)
\end{aligned}$$

To define a decision surface $g_{ij}(\boldsymbol{x})$ between classes ω_i and ω_j, we require that $g_i(\boldsymbol{x}) - g_j(\boldsymbol{x}) = 0$ and can write

$$\begin{aligned}
g_{ij}(\mathbf{x}) &= \mathbf{w}^\intercal(\mathbf{x}-\mathbf{x}_0) & (8)\\
\mathbf{w} &= \boldsymbol{\mu}_i - \boldsymbol{\mu}_j\\
\mathbf{x}_0 &= \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) - \sigma^2\ln\left(\frac{P(\omega_i)}{P(\omega_j)}\right)\frac{\boldsymbol{\mu}_i - \boldsymbol{\mu}_j}{\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\|^2}
\end{aligned}$$

PO BOX 20, FI-53851 LAPPEENRANTA, FINLAND
E-mail: lasse.lensu@lut.fi

Since the a priori probabilities are equal, $\frac{P(\omega_i)}{P(\omega_j)} = 1$ and $\ln\left(\frac{P(\omega_i)}{P(\omega_j)}\right) = 0$. Therefore, $\mathbf{x}_0 = \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j)$.

As an example calculation, we determine the decision boundary between classes ω_1 and ω_2:

$$\begin{aligned}
\mathbf{w_{12}} &= \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \begin{pmatrix}3\\6\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}0\\8\end{pmatrix} \\
\mathbf{x}_0^{12} &= \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) = \frac{1}{2}\left[\begin{pmatrix}3\\6\end{pmatrix} + \begin{pmatrix}3\\-2\end{pmatrix}\right] = \begin{pmatrix}3\\2\end{pmatrix}.
\end{aligned}$$

From Eq. 8, we get

$$\begin{aligned}
g_{ij}(\mathbf{x}) = \mathbf{w}^\mathsf{T}(\mathbf{x} - \mathbf{x}_0) &= 0 \\
\Rightarrow \begin{pmatrix}0\\8\end{pmatrix}^\mathsf{T}\left(\mathbf{x} - \begin{pmatrix}3\\2\end{pmatrix}\right) &= 0 \\
\Rightarrow \begin{pmatrix}0 & 8\end{pmatrix}\begin{pmatrix}\mathbf{x}_1 - 3\\ \mathbf{x}_2 - 2\end{pmatrix} &= 0 \\
\Rightarrow 8\mathbf{x}_2 - 16 &= 0 \\
\Rightarrow \mathbf{x}_2 &= 2.
\end{aligned}$$

The other two decision boundaries can be solved in a similar way.

To visualise the classes and the decision boundaries, functions `plotclass.m` and `line` can be used in Matlab. An example is given in Fig. 1.

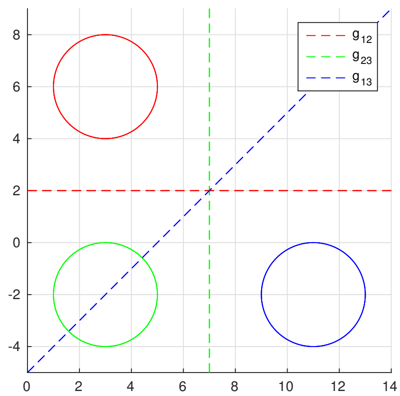

Figure 1: The classes shown as equiprobable circles and the decision boundaries.