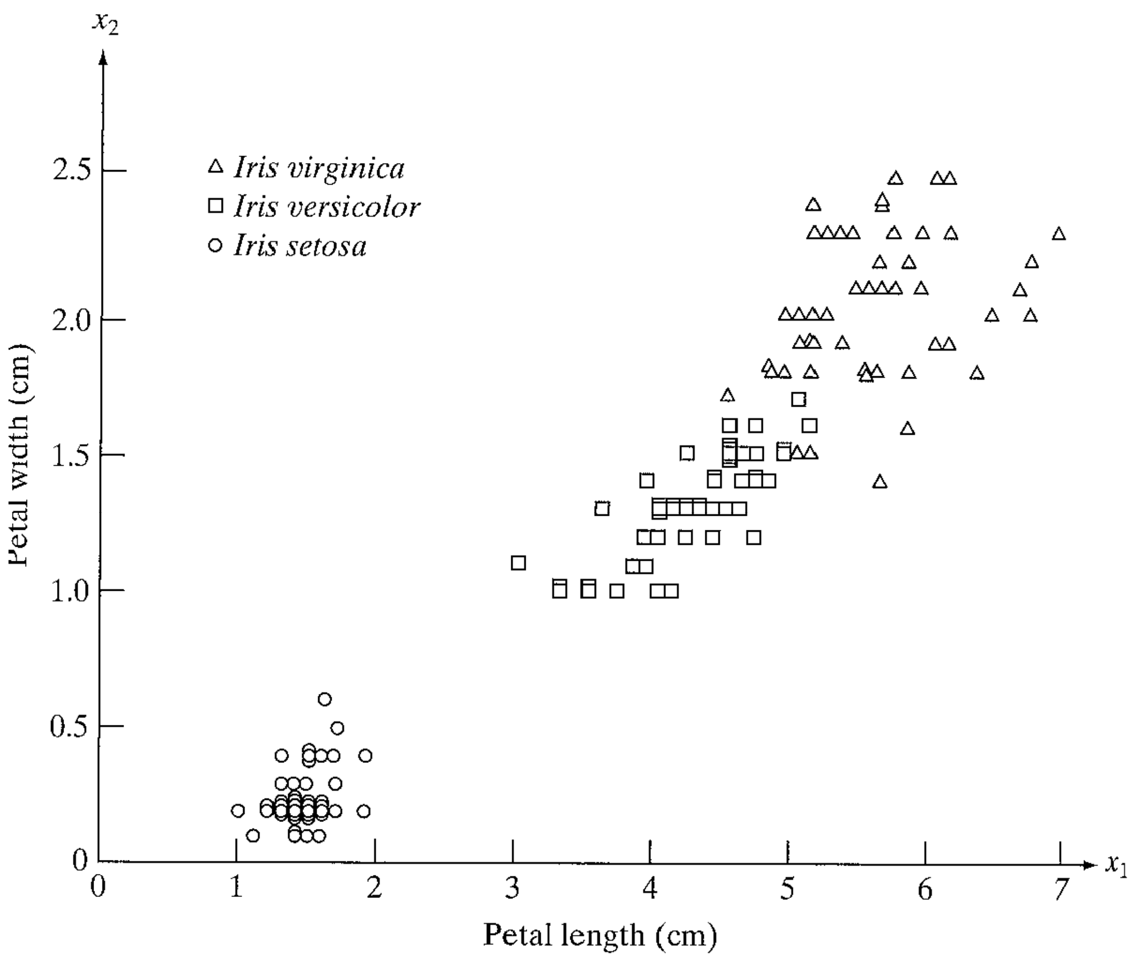
Solution of selected exercises

Exercises INF 4300 related to the lecture 10.10.16

2. Finding the decision functions for a minimum distance classifier.

A classifier that uses diagonal covariance matrices is often called a minimum distance classifier, because a pattern is classified to class that is closest when distance is computed using Euclidean distance.



a. In the above figure, find the class means just by looking at the plot.

b. If this data is classified using a minimum distance classifier, sketch the decision boundaries on the plot.

Solution:

Problem 12.1

(a) By inspection, the mean vectors of the three classes are, approximately, $\mathbf{m}_1 = (1.5, 0.3)^T$, $\mathbf{m}_2 = (4.3, 1.3)^T$, and $\mathbf{m}_3 = (5.5, 2.1)^T$ for the classes Iris setosa, versicolor, and virginica, respectively. The decision functions are of the form given in Eq. (12.2-5). Substituting the above values of mean vectors gives:

$$
\begin{aligned}
d_1(\mathbf{x}) &= \mathbf{x}^T\mathbf{m}_1 - \frac{1}{2}\mathbf{m}_1^T\mathbf{m}_1 = 1.5x_1 + 0.3x_2 - 1.2 \\
d_2(\mathbf{x}) &= \mathbf{x}^T\mathbf{m}_2 - \frac{1}{2}\mathbf{m}_2^T\mathbf{m}_2 = 4.3x_1 + 1.3x_2 - 10.1 \\
d_3(\mathbf{x}) &= \mathbf{x}^T\mathbf{m}_3 - \frac{1}{2}\mathbf{m}_3^T\mathbf{m}_3 = 5.5x_1 + 2.1x_2 - 17.3
\end{aligned}
$$

(b) The decision boundaries are given by the equations

$$
\begin{aligned}
d_{12}(\mathbf{x}) &= d_1(\mathbf{x}) - d_2(\mathbf{x}) = -2.8x_1 - 1.0x_2 + 8.9 = 0 \\
d_{13}(\mathbf{x}) &= d_1(\mathbf{x}) - d_3(\mathbf{x}) = -4.0x_1 - 1.8x_2 + 16.1 = 0 \\
d_{23}(\mathbf{x}) &= d_2(\mathbf{x}) - d_3(\mathbf{x}) = -1.2x_1 - 0.8x_2 + 7.2 = 0
\end{aligned}
$$

A plot of these boundaries is shown in Fig. P12.1.

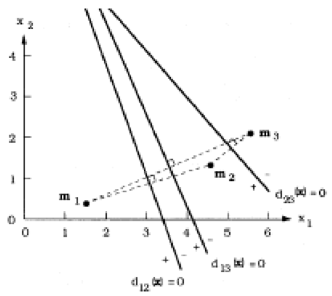


Figure P12.1

3. Discriminant functions

A classifier that uses Euclidean distance computes distance from pattern ***x*** to class j as:

$$D_j(x) = \left\| x - \mu_j \right\|$$

Show that classification with this rule is equivalent to using the discriminant function

$$d_j(x) = x^T\mu_j - \frac{1}{2}\mu_j^T\mu_j$$

Solution:

Problem 12.2

From the definition of the Euclidean distance,

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| = \left[(\mathbf{x} - \mathbf{m}_j)^T(\mathbf{x} - \mathbf{m}_j)\right]^{1/2}$$

Since $D_j(\mathbf{x})$ is non-negative, choosing the smallest $D_j(\mathbf{x})$ is the same as choosing the smallest $D_j^2(\mathbf{x})$, where

$$
\begin{aligned}
D_j^2(\mathbf{x}) &= \|\mathbf{x} - \mathbf{m}_j\|^2 = (\mathbf{x} - \mathbf{m}_j)^T(\mathbf{x} - \mathbf{m}_j) \\
&= \mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{m}_j + \mathbf{m}_j^T\mathbf{m}_j \\
&= \mathbf{x}^T\mathbf{x} - \mathbf{2}\left(\mathbf{x}^T\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{m}_j\right)
\end{aligned}
$$

We note that the term $\mathbf{x}^T\mathbf{x}$ is independent of j (that is, it is a constant with respect to j in $D_j^2(\mathbf{x})$, $j = 1, 2, ...$). Thus, choosing the minimum of $D_j^2(\mathbf{x})$ is equivalent to choosing the maximum of $\left(\mathbf{x}^T\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{m}_j\right)$.

❖ Example:

Given ω_1, ω_2 : $P(\omega_1) = P(\omega_2)$ and $p(\underline{x}|\omega_1) = N(\underline{\mu}_1, \Sigma)$,

$p(\underline{x}|\omega_2) = N(\underline{\mu}_2, \Sigma)$, $\underline{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\underline{\mu}_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $\Sigma = \begin{bmatrix} 1.1 & 0.3 \\ 0.3 & 1.9 \end{bmatrix}$

classify the vector $\underline{x} = \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix}$ using Bayesian classification :

- $\Sigma^{-1} = \begin{bmatrix} 0.95 & -0.15 \\ -0.15 & 0.55 \end{bmatrix}$
  - Compute Mahalanobis d_m from μ_1, μ_2 : $d^2_{m,1} = [1.0, \quad 2.2]$ $\Sigma^{-1} \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix} = 2.952$, $d^2_{m,2} = [-2.0, \quad -0.8]\, \Sigma^{-1} \begin{bmatrix} -2.0 \\ -0.8 \end{bmatrix} = 3.672$

- Classify $\underline{x} \rightarrow \omega_1$. Observe that $d_{E,2} < d_{E,1}$