Lecture 3

Bayes Classifiers

Minimum Distance Classifier

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}_j \qquad j = 1, 2, \ldots, W$$

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \qquad j = 1, 2, \ldots, W$$

$$\|\mathbf{a}\| = (\mathbf{a}^T\mathbf{a})^{1/2}$$ is the Euclidean Norm

e.g., $a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$a^T = [1 \quad 2]$, $\quad a^T a = \underset{1\times 2}{[1 \quad 2]} \underset{2\times 1}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}} = 1\times 1 + 2\times 2 = 5$

$\|a\| = \sqrt{a^T a} = \sqrt{5}$

```
>> a

a =

   1
   2
```

```
>> norm(a)

ans =

   2.2361
```

△ Decision function

$$(\underset{1\times n}{x^T} \underset{n\times 1}{m_1})^T = m_1^T x$$

$$\|x - m_1\| \overset{?}{\lesseqgtr} \|x - m_2\|$$

$$(x - m_1)^T (x - m_1) = (x^T - m_1^T)(x - m_1) = x^T x - \underbrace{x^T m_1 - m_1^T x}_{} + m_1^T m_1 \quad \cdots (1)$$

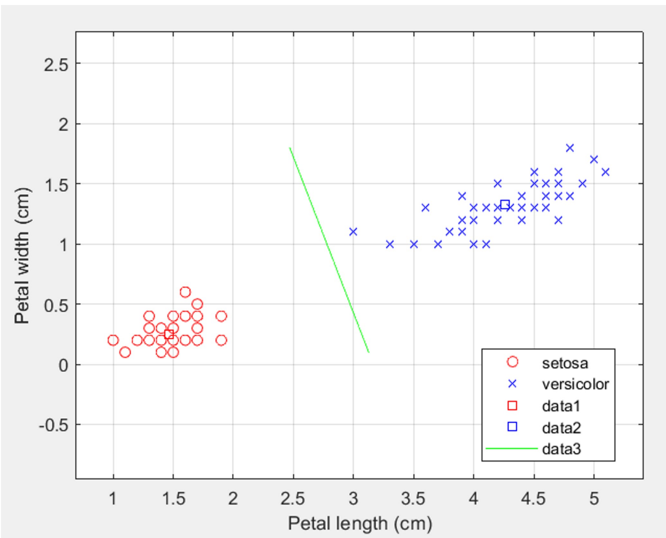
$$(x - m_2)^T (x - m_2) = (x^T - m_2^T)(x - m_2) = x^T x - x^T m_2 - m_2^T x + m_2^T m_2 \quad \cdots (2)$$

$$-2x^T m_1 + m_1^T m_1 < -2x^T m_2 + m_2^T m_2$$

$$x^T m_1 - \frac{1}{2} m_1^T m_1 > x^T m_2 - \frac{1}{2} m_2^T m_2$$

- It can be shown that it is equivalent to selecting a class that can maximize the following decision function:

$$d_j(\mathbf{x}) = \mathbf{x}^T\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{m}_j \qquad j = 1, 2, \ldots, W$$



"md_classifer.m"

```
% Calculate the coefficients for the decision boundary line
% Straight line equation: ax + by + c = 0
a = mean_versi_x - mean_seto_x;
b = mean_versi_y - mean_seto_y;
c = -0.5*(mean_versi_x^2 - mean_seto_x^2 + ...
   mean_versi_y^2 - mean_seto_y^2);
```

Derivation of the Bayes Classifier

$$r_j(\boldsymbol{x}) = \sum_{k=1}^{N} L_{kj}\, p(\boldsymbol{x}|c_k) P(c_k) \quad \text{and} \quad L_{kj} = 1 - \delta_{kj}$$

$$r_j(\boldsymbol{x}) = \sum_{k=1}^{N} (1 - \delta_{ij}) p(\boldsymbol{x}|c_k) P(c_k)$$

$$= \sum_{k=1}^{N} p(\boldsymbol{x}|c_k) P(c_k) - \sum_{k=1}^{N} \delta_{ij}\, p(\boldsymbol{x}|c_k) P(c_k)$$

$$= p(\boldsymbol{x}) - p(\boldsymbol{x}|c_j) P(c_j)$$

$$\sum_{k=1}^{N} p(x, c_k) = p(x)$$

Similarly,

$$r_i(\boldsymbol{x}) = p(\boldsymbol{x}) - p(\boldsymbol{x}|c_i) P(c_i)$$

Decision Rule

- classifier assigns an unknown pattern $\boldsymbol{x}$ to class c_i if
  $r_i(\boldsymbol{x}) < r_j(\boldsymbol{x})$ for $j = 1, 2, \ldots, N; j \neq i$.

$$p(\boldsymbol{x}) - p(\boldsymbol{x}|c_i) P(c_i) < p(\boldsymbol{x}) - p(\boldsymbol{x}|c_j) P(c_j),$$

or equivalently

$$\boxed{p(\boldsymbol{x}|c_i) P(c_i) > p(\boldsymbol{x}|c_j) P(c_j)}$$