Bayesian Classification Part-7

$$p(\vec{x}) = \frac{1}{(\sqrt{2\pi})^{l} |\Sigma|^{1/2}} \cdot \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})\right)$$

Bayesian Classifier for Normally Distributed classes:

$p(\omega_i|\vec{x})$, $i = 1:m$

$$p(\omega_i|\vec{x}) = \frac{p(\vec{x}|\omega_i)\, p(\omega_i)}{p(\vec{x})}$$

$$\boxed{\max_i \; p(\vec{x}|\omega_i)\, p(\omega_i)}$$

$p(\vec{x}|\omega_i)$: Likelihood prob.

$p(\omega_i)$: prior probability.

Gaussian distribution

$$f_i(\vec{x}) = p(\vec{x}|\omega_i)\, p(\omega_i)$$

$$f_j(\vec{x}) = p(\vec{x}|\omega_j)\, p(\omega_j)$$

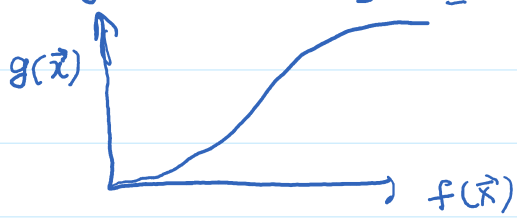
$$f_{ij}(\vec{x}) = f_i(\vec{x}) - f_j(\vec{x}) = 0$$

$$f_i(\vec{x}) = \frac{1}{(\sqrt{2\pi})^{l} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu}_i)^T \Sigma_i (\vec{x}-\vec{\mu}_i)\right) p(\omega_i)$$

$$g_i(\vec{x}) = \ln(f_i(\vec{x}))$$

$$g_j(\vec{x}) = \ln(f_j(\vec{x}))$$

$g \rightarrow$ montonically $\uparrow$ fn



If $f_i(\vec{x}) > f_j(\vec{x})$

$g_i(\vec{x}) > g_j(\vec{x})$

$$g_{ij}(\vec{x}) = g_i(\vec{x}) - g_j(\vec{x}) = 0$$

$$f_i(\vec{x}) = \frac{1}{(\sqrt{2\pi})^{l} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x}-\vec{\mu}_i)\right) p(\omega_i)$$

$$\Rightarrow g_i(\vec{x}) = -\frac{1}{2}(\vec{x}-\vec{\mu}_i)^T \Sigma_i (\vec{x}-\vec{\mu}_i) + \ln(p(\omega_i)) \cancel{- \frac{l}{2}\ln(2\pi)} - \frac{1}{2}\ln(|\Sigma_i|)$$

$$g_j(\vec{x}) = -\frac{1}{2}(\vec{x}-\vec{\mu}_j)^T \Sigma_j (\vec{x}-\vec{\mu}_j) + \ln(p(\omega_j)) \cancel{- \frac{l}{2}\ln(2\pi)} - \frac{1}{2}(\ln(|\Sigma_j|))$$

If $g_{ij}(\vec{x}) \triangleq g_i(\vec{x}) - g_j(\vec{x}) = 0 \Rightarrow$ Decision Hyper-plane

Special-Case

$$L = 2 \Rightarrow \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \vec{\mu}_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \end{bmatrix}$$

$$\Sigma_i = \begin{bmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_i^2 \end{bmatrix} = \sigma_i^2 \begin{bmatrix} 1 & 0 \\ & 1 \end{bmatrix}$$

$\Rightarrow$ All components of the feature vector are stat. independent

$$\Rightarrow g_i(\vec{x}) = -\frac{1}{2}(\vec{x} - \vec{\mu}_i)^T \underbrace{\Sigma_i^{-1}}(\vec{x} - \vec{\mu}_i) + \ln(p(\omega_i)) - \cancel{\frac{\ell}{2}\ln(2\pi)} - \frac{1}{2}\underbrace{\ln|\Sigma_i|}$$

$$g_i(\vec{x}) = -\left(\frac{1}{2\sigma^2}\right)\left[\underbrace{(\vec{x} - \vec{\mu}_i)^T(\vec{x} - \vec{\mu}_i)}\right] + \ln(p(\omega_i)) - \frac{1}{2}\ln(|\Sigma_i|)$$

$$\left((\vec{x})^T - (\vec{\mu}_i)^T\right)(\vec{x} - \vec{\mu}_i)$$

$$= (\vec{x})^T\vec{x} - \underbrace{(\vec{x})^T\vec{\mu}_i}_{1\times 1} - \underbrace{(\vec{\mu}_i)^T\vec{x}}_{1\times 1} + (\vec{\mu}_i)^T\vec{\mu}_i.$$

$$= (\vec{x})^T\vec{x} - 2(\vec{\mu}_i)^T\vec{x} + (\vec{\mu}_i)^T\vec{\mu}_i$$

$$\Rightarrow g_i(\vec{x}) = -\frac{1}{2\sigma^2}(\vec{x})^T\vec{x} + \frac{1}{\sigma^2}(\vec{\mu}_i)^T\vec{x} - \frac{1}{2\sigma^2}\vec{\mu}_i^T\vec{\mu}_i + \ln(p(\omega_i)) - \frac{1}{2}\ln(|\Sigma_i|)$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \vec{\mu}_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \end{bmatrix}$$

$$g_i(\vec{x}) = -\frac{1}{2\sigma_i^2}\left(x_1^2 + x_2^2 - 2(\mu_{i1}x_1 + \mu_{i2}x_2) + \underbrace{(\mu_{i1}^2 + \mu_{i2}^2)}\right) + \ln(p(\omega_i)) - \frac{1}{2}\ln(|\Sigma_i|)$$

$g_i(\vec{x}) - g_j(\vec{x})$

Discriminant $\Rightarrow$

Quadrics (ellipsoids, paraboloas, hyperboloas etc.)