Bayesian Classification Part-5

Example:

Two-class problem $(M=2)$

One-dimensional feature vector $(l=1)$

Likelihood functions are Gaussian distributions with:

$\mu_1 = 0, \quad \sigma_1^2 = 1/2 \; ; \; \mu_2 = 1, \quad \sigma_2^2 = 1/2$

Prior probabilities of both classes are equal $\left(P(\omega_1) = P(\omega_2) = 1/2\right)$

Compute the threshold values (x_0) for

(a) minimum error probability (P_e)

(b) minimum risk if the loss matrix is: $L = \begin{bmatrix} 0 & 0.5 \\ 1.0 & 0 \end{bmatrix}$

Solution

$$p(x|\omega_1) = \frac{1(\sqrt{2})}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{\cancel{2}}(\cancel{2})\right) = \frac{1}{\sqrt{\pi}} \exp\left(-x^2\right)$$

$$p(x|\omega_2) = \frac{1(\cancel{\sqrt{2}})}{\cancel{\sqrt{2}}\sqrt{\pi}} \exp\left(-\frac{(x-1)^2}{\cancel{2}}(\cancel{2})\right) = \frac{1}{\sqrt{\pi}} \exp\left(-(x-1)^2\right)$$

(a) $$\cancel{\frac{1}{\sqrt{\pi}}} \exp\left(-x_0^2\right) = \cancel{\frac{1}{\sqrt{\pi}}} \exp\left(-(x_0-1)^2\right)$$

$$\Rightarrow \quad -\cancel{x_0^2} = -(\cancel{x_0^2} + 1 - 2x_0) \Rightarrow \boxed{x_0 = \frac{1}{2}}$$

(b) $$p(x_0|\omega_1) = \left(\frac{\lambda_{21}}{\lambda_{12}}\right) p(x_0|\omega_2)$$

$$\Rightarrow \quad \cancel{\frac{1}{\sqrt{\pi}}} \exp\left(-x_0^2\right) = \left(\frac{1}{0.5}\right)\left(\cancel{\frac{1}{\sqrt{\pi}}}\right) \exp\left(-(x_0-1)^2\right)$$

Taking $\log_e(\,)$ on B.S.

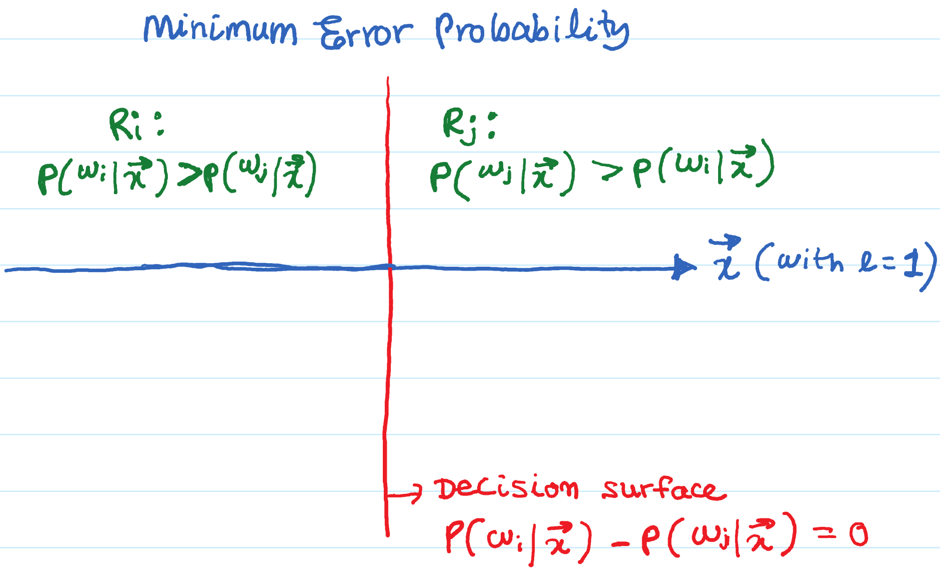
$$\Rightarrow \quad -\cancel{x_0^2} = \ln(2) - (x_0-1)^2 = \ln 2 - \cancel{x_0^2} - 1 + 2x_0$$

$$\Rightarrow \quad x_0 = \frac{1}{2}(1 - \ln 2) \simeq \frac{1}{2}(0.3069) = \underline{0.153}$$

$$\Rightarrow \quad \boxed{x_0 = 0.153}$$

Decision Surfaces:

- Minimizing the error probability or the risk is equivalent to partitioning the feature space into M regions.
- If these regions are contiguous, then they are separated by decision surfaces in the l-dimensional feature space.



Discriminant Function:

Let $f(\cdot)$ be a montonically increasing function,

Let $g_i(\vec{x}) \triangleq f(P(\omega_i|\vec{x}))$ be the discriminant function defined for performing classification.

If $g_i(\vec{x}) - g_j(\vec{x}) > 0$,

then, $\vec{x} \in \omega_i$.

If $g_i(\vec{x}) - g_j(\vec{x}) < 0$,

then, $\vec{x} \in \omega_j$.

In general, discriminant functions are defined independent of the Bayesian rule. They are computationally more tracable, and in practice, they may also lead to better solutions. We will discuss more about them in the upcoming classes!

2.4 Bayesian Classification for Normal Distribution:

2.4.1. Gaussian Probability Density Function:

1D Case (Univariate Gaussian Function):

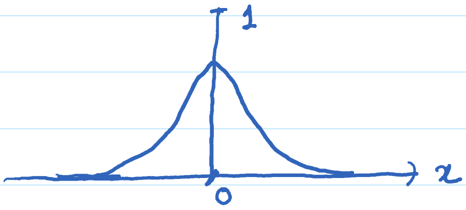
$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

where μ = mean value of $p(x)$

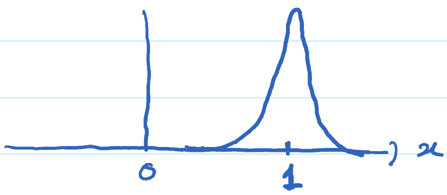
$$\mu = E[x] = \int_{-\infty}^{+\infty} x\, p(x)\, dx$$

σ^2 = variance of $p(x)$

$$= E\left[\left(x - E[x]\right)^2\right] = \int_{-\infty}^{+\infty} (x-\mu)^2\, p(x)\, dx.$$



$\mu = 0, \; \sigma^2 = 1$



$\mu = 1, \; \sigma^2 = 0.2$