THE BAYES CLASSIFIER

Consider (X, Y) where

$$X \in \mathbb{R}^d$$

$$Y \in \{1, \ldots, K\}$$

Let $f : \mathbb{R}^d \longrightarrow \{1, \ldots, K\}$ be a <u>classifier</u>

The <u>probability of error</u> of f is

$$R(f) := \Pr\{f(X) \neq Y\}$$

Denote the <u>a posteriori</u> class probabilities by

$$\eta_k(x) :=$$

$k = 1, \ldots, K$.

Theorem The classifier

$$f^*(x) := \arg\max_k \eta_k(x)$$

satisfies

$$R(f^*) = \min R(f)$$

where the min is over all classifiers.

Terminology

- f^* is called the
- $R(f^*)$ is called the

Proof For convenience, assume $X \mid Y=k$ is a continuous random variable with density $g_k(x)$.

Let $\pi_k := \Pr\{Y=k\}$, the <u>a priori</u> class probabilities.

Consider an arbitrary classifier f. Denote the decision regions

$$\Gamma_k(f) := \{x : f(x) = k\}.$$

Then

$$1 - R(f) = \Pr\{f(X) = Y\}$$

A $$=$$

$$=$$

To maximize this expression, we should select f such that

B $$x \in \Gamma_k(f) \iff$$

Therefore, the optimal f has

$$f^*(x) =$$

By Bayes rule,

(c) $$\eta_k(x) =$$

Variations

Different ways of expressing the Bayes classier:

- $$f^*(x) = \arg\max_k \eta_k(x)$$

- $$f^*(x) = \arg\max_k \pi_k g_k(x)$$

- When $K = 2$

$$\frac{g_2(x)}{g_1(x)} \underset{1}{\overset{2}{\gtrless}} \frac{\pi_1}{\pi_2}$$ (likelihood ratio test)

- When $\pi_1 = \pi_2 = \cdots = \pi_K$

$$f^*(x) = \arg\max g_k(x)$$

(maximum likelihood classifier/detector)

$\boxed{\text{Key}}$ A. $= \sum_{k=1}^{K} \pi_k \Pr\{f(X)=k \mid Y=k\}$

$$= \sum_{k=1}^{K} \pi_k \int_{\Pi_k(f)} g_k(x)\,dx$$

B. $x \in \Gamma_k(x) \iff \pi_k g_k(x)$ is maximal

$$\Rightarrow f^*(x) = \arg\max_k \pi_k g_k(x)$$

C. $\eta_k(x) = P\{Y=k \mid X=x\}$

$$= \frac{\pi_k g_k(x)}{\underbrace{\sum_{l=1}^{K} \pi_l g_l(x)}_{\text{independent of } k}}$$