M339G: October 21st, 2024.

Linear Discriminant Analysis w/ p=1.

Fisher in 1936.

Goal: Classify observations into one of the K classes ($K \geq 2$), i.e., figure out:

$$\boxed{p_k(x) := \mathbb{P}[Y = k \mid X = x]}$$ posterior probability ✓

Environment:
- π_k ... prior probability that a randomly chosen observation falls into category $k = 1..K$
- $f_k(x)$ .. density function of X for observations from class k

choice

model

$f_k(x)\,dx$ ... the probability that X falls in $(x, x+dx)$ for points from class k

Then,

$$\mathbb{P}[Y=k \mid X=x] = \frac{\mathbb{P}[Y=k \text{ and } X=x]}{\mathbb{P}[X=x]} = \frac{\overbrace{\mathbb{P}[X=x \mid Y=k]}^{f_k(x)} \cdot \overbrace{\mathbb{P}[Y=k]}^{\pi_k}}{\mathbb{P}[X=x]}$$

$X \in (x, x+dx)$

Bayes Thm

The Law of Total Probability

And,

$$\boxed{p_k(x) = \frac{\pi_k \cdot f_k(x)}{\sum_{j=1}^{K} \pi_j \cdot f_j(x)}}$$

$\Rightarrow$ classify into

$$\boxed{k = \underset{j=1..K}{\operatorname{argmax}} \left(p_j(x)\right)}$$

Linear Discriminant Analysis (LDA).

The choice: f_k are normal densities for each $k=1..K$, i.e.,

$$\boxed{f_k(x) = \frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}}}$$ for $k = 1..K$

w/ μ_k and σ_k being the mean and the standard deviation for the k^{th} class

Additional Assumption: Homogeneity: $\boxed{\sigma_1 = \dots = \sigma_K = \sigma}$

We now return to the posterior probability, i.e.,

$$p_k(x) = \frac{\pi_k \cdot f_k(x)}{\boxed{\sum_{j=1}^{K} \pi_j f_j(x)}}$$

<u>Remember</u>: We're looking for the k for which the above is MAXIMAL.

Since all $p_k(x)$ have the same denominator, it's sufficient to find the k for which

$$\pi_k \cdot f_k(x) \longrightarrow \max$$

Because ln is increasing, this is equivalent to:

$$\ln(\pi_k) + \ln(f_k(x)) \longrightarrow \max$$

$$\Longleftrightarrow$$

$$\ln(\pi_k) + \ln\left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu_k)^2}{2\sigma^2}}\right) \longrightarrow \max$$

$$\Longleftrightarrow$$

$$\ln(\pi_k) - \boxed{\ln(\sigma\sqrt{2\pi})} - \frac{(x-\mu_k)^2}{2\sigma^2} \longrightarrow \max$$

const. in terms of k

$$\Longleftrightarrow$$

$$\ln(\pi_k) \boxed{- \frac{x^2}{2\sigma^2}} + \frac{\cancel{2} x\mu_k}{\cancel{2}\sigma^2} - \frac{\mu_k^2}{2\sigma^2} \longrightarrow \max$$

doesn't depend on k

$$\delta_k(x) := \ln(\pi_k) + \frac{\mu_k}{\sigma^2} x - \frac{\mu_k^2}{2\sigma^2} \longrightarrow \max$$

These are called DISCRIMINANT (SCORES) and they are LINEAR in x

<u>Special Case</u>: $K=2$, $\pi_1 = \pi_2 = \frac{1}{2}$

$$\frac{\mu_k}{\sigma^2} \cdot x - \frac{\mu_k^2}{2\sigma^2} \longrightarrow \max$$

$$\mu_k x - \frac{\mu_k^2}{2} \longrightarrow \max$$

IF $\underline{\mu_1 x - \frac{\mu_1^2}{2} > \mu_2 x - \frac{\mu_2^2}{2}}$, $\underline{\text{then}}$ classify as 1

$$(\mu_1 - \mu_2)x > \frac{\mu_1^2 - \mu_2^2}{2} = \frac{(\mu_1 - \mu_2)(\mu_1 + \mu_2)}{2}$$

Boundary always $\frac{\mu_1+\mu_2}{2}$