Assignment Classification

Question 3

This problem relates to the QDA model, where observations within each class are drawn from a normal distribution with class-specific mean and variance. We consider the case $p = 1$ (one feature). Suppose that for class k,

$$X \mid Y = k \sim \mathcal{N}(\mu_k, \sigma_k^2).$$

We show that the Bayes classifier is not linear and is instead quadratic.

Step 1: Bayes Classifier

The Bayes classifier assigns x to the class maximizing

$$P(Y = k \mid X = x).$$

Using Bayes' theorem, this is equivalent to maximizing

$$\pi_k f_k(x)$$

where $\pi_k = P(Y = k)$ and $f_k(x)$ is the class-conditional density.

Step 2: Gaussian Density

The density of the normal distribution is

$$f_k(x) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right).$$

Thus we maximize

$$\pi_k \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right).$$

Step 3: Log Discriminant Function

Taking logarithms:

$$\delta_k(x) = \log \pi_k - \frac{1}{2}\log(2\pi\sigma_k^2) - \frac{(x-\mu_k)^2}{2\sigma_k^2}.$$

Ignoring constants shared across classes:

$$\delta_k(x) = \log \pi_k - \frac{1}{2}\log(\sigma_k^2) - \frac{(x-\mu_k)^2}{2\sigma_k^2}.$$

Step 4: Expand the Square

$$(x-\mu_k)^2 = x^2 - 2\mu_k x + \mu_k^2$$

Substitute:

$$\delta_k(x) = \log \pi_k - \frac{1}{2}\log(\sigma_k^2) - \frac{x^2}{2\sigma_k^2} + \frac{\mu_k x}{\sigma_k^2} - \frac{\mu_k^2}{2\sigma_k^2}.$$

Step 5: Quadratic Form

This can be written as

$$\delta_k(x) = a_k x^2 + b_k x + c_k$$

where

$$a_k = -\frac{1}{2\sigma_k^2}, \qquad b_k = \frac{\mu_k}{\sigma_k^2}$$

and c_k is constant in x.
Therefore $\delta_k(x)$ is quadratic in x.

Step 6: Decision Boundary

The boundary between classes k and j satisfies

$$\delta_k(x) = \delta_j(x)$$

which results in a quadratic equation in x because the x^2 terms do not cancel when $\sigma_k^2 \neq \sigma_j^2$.

Conclusion

Since the discriminant function contains x^2 terms that remain when comparing classes with different variances, the Bayes classifier is not linear. Instead, the decision boundary is quadratic, which explains the name Quadratic Discriminant Analysis (QDA).