ISLR 4.8.3

is maximized.

3. This problem relates to the QDA model, in which the observations within each class are drawn from a normal distribution with a class-specific mean vector and a class specific covariance matrix. We consider the simple case where $p = 1$; i.e. there is only one feature.

   Suppose that we have K classes, and that if an observation belongs to the kth class then X comes from a one-dimensional normal distribution, $X \sim N(\mu_k, \sigma_k^2)$. Recall that the density function for the one-dimensional normal distribution is given in (4.16). Prove that in this case, the Bayes classifier is *not* linear. Argue that it is in fact quadratic.

   Hint: For this problem, you should follow the arguments laid out in Section 4.4.1, but without making the assumption that $\sigma_1^2 = \ldots = \sigma_K^2$.

Answer

Normal Density fn. in a one dimensional setting

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

It is given that $p = 1$. Hence, the normal density fn. is as shown above.

μ_k and σ_k^2 are the mean and variance parameters for the k^{th} class.

We also know the Bayes Theorem states that $\Pr(Y = k|X = x) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{K} \pi_l f_l(x)}$.

Using the $f_k(x)$ formula, we can say:

$$p_k(x) = \frac{\pi_k \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(\frac{-1}{2\sigma_k^2}(x-\mu_k)^2\right)}{\sum_{l=1}^{K} \pi_l \frac{1}{\sqrt{2\pi}\sigma_l} \exp\left(\frac{-1}{2\sigma_l^2}(x-\mu_l)^2\right)}$$

The Bayes Classifier involves assigning

an observation $X = x$ to the class for which $p_k(x)$ is the largest.

Since the log function is monotonically increasing, we can take the log on both sides of the equation of $p(k)$ and find the largest k for which $\log(p_k(x))$ is the largest.

$$p_k(x) = \frac{\frac{\pi_k}{\sigma_k} \exp\left(-\frac{1}{2\sigma_k^2}(x-\mu_k)^2\right)}{\sum_{l=1}^{K} \frac{\pi_l}{\sigma_l} \exp\left(-\frac{1}{2\sigma_l^2}(x-\mu_l)^2\right)}$$

$$\log(p_k(x)) = \log\left(\frac{\pi_k}{\sigma_k}\right) - \frac{1}{2\sigma_k^2}(x-\mu_k)^2$$

$$- \log \sum_{l=1}^{K} \frac{\pi_l}{\sigma_l} \exp\left(\frac{-1}{2\sigma_l^2}(x-\mu_l)^2\right)$$

Since the last term is independent of k, we just need to find k for which

$$\log \frac{\pi_k}{\sigma_k} - \frac{1}{2\sigma_k^2}(x-\mu_k)^2$$ is max

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

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

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

The above expression is clearly not linear in x (as x^2 is also present, and the x^2 coefficient is dependent on k).

Hence, the Bayes Classifier is in fact <u>Quadratic</u>.