4.7.3

Bays theorem can be stated as :
(Conditional probability)

~~BA~~

$$P(C_K \mid x_1, x_2 \ldots x_n)$$

$$= P(C_K \mid x) = \frac{P(C_K) \cdot P(x \mid C_K)}{P(x)}$$

considering the gaussian naive Bayes for continuos data :

$$P(x = v \mid C_K) = \frac{1}{\sqrt{2\pi\sigma_K^2}} \cdot e^{\frac{(v - u_K)^2}{2\sigma_K^2}}$$

where u_K : mean values in x associated with K.

σ_K^2 : variance of values in x associated with K

Now for Bayes classifier, ~~wt~~ we have

$$p_K(x) = \frac{\pi_K \cdot \frac{1}{\sqrt{2\pi}\sigma_K} \cdot \exp\left(-\frac{1}{2\sigma_K^2}(x-\mu_K)^2\right)}{\sum_{l=1}^{K} \pi_l \cdot \frac{1}{\sqrt{2\pi}\sigma_l} \cdot \exp\left(\frac{1}{2\sigma_l^2}(x-\mu_l)^2\right)}$$

Our estimate should be k with largest $\delta(x)$, where taking logarithm

$$\cancel{\delta(x) = \log \pi_K - \frac{\mu_K^2}{2\sigma^2} + x \cdot \frac{\mu_K}{\sigma^2}}$$

$$\cancel{\sigma} \quad \delta(x) = \log \pi_K + \log \frac{1}{\sqrt{2\pi}\sigma_K} - \frac{1}{2\sigma_K^2} \cdot (x-\mu_K)^2$$

$$\Rightarrow \quad \sigma_K(X) = \log \frac{1}{\sqrt{2\sigma_K \pi}} - \frac{1}{2\sigma_K^2} \cdot x^2 + \frac{\mu_K x}{\sigma_K^2} - \frac{\mu_K^2}{2\sigma_K^2} + \log \pi_K$$

A quadratic term $\frac{1}{2\sigma_K^2} \cdot x^2$ is present and cannot be ignored since σ^2 is not being shared.
Hence it proves that the Bayes Classifier is not linear and is quadratic in nature.