Model based approach

1 Our goal is to build parametric model for conditional distribution $P(Y=k|X=x)$

$$Pr(y=k|X=x) = \frac{Pr(X=x|Y=k)\,Pr(Y=k)}{\sum_{l=1}^{K} Pr(X=x|Y=l)\,Pr(Y=l)}$$

$$Pr(Y=k|X=x) = \frac{f_k(x)\,\pi_k}{\sum_{l=1}^{K} f_l(x)\,\pi_l}$$

π_k prior probability

A common choice $f_k(x) = \frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{(x-\mu_k)^2}{\sigma_k^2}}$

Estimate the parameters (μ_k, σ_k^2) from the data

Classify to the class with the highest value of $P(Y=k|X=x)$

$$\log \frac{Pr(Y=k \mid X=x)}{Pr(Y=j \mid X=x)} =$$

$$= \log \frac{f_k(x)}{f_j(x)} + \log \frac{\pi_k}{\pi_j} =$$

$$= \log \frac{\pi_k}{\pi_j} - \frac{1}{2}(\mu_k + \mu_j)^T \Sigma^{-1} (\mu_k + \mu_j) +$$

$$+ x^T \Sigma^{-1} (\mu_k - \mu_j)$$

Discriminant function

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k \Sigma^{-1} \mu_k + \log(\mu_k)$$

- Decide on class based on $\hat{Y}(x) = \arg\max_k \delta(x)$
- We usually do this with EM.