# Example Two classes salmon & seabass

Observed data:

$\omega_1 = sl$: Salmon: $\begin{bmatrix}10\\5\end{bmatrix}, \begin{bmatrix}11\\7\end{bmatrix}, \begin{bmatrix}12\\6\end{bmatrix}$, # of observations in $\omega_1$, $m=3$

$\omega_2 = sb$: Seabass: $\begin{bmatrix}5\\10\end{bmatrix}, \begin{bmatrix}7\\11\end{bmatrix}, \begin{bmatrix}6\\12\end{bmatrix}$ # of observations in $\omega_2 \Rightarrow m=3$

$$\hat{\pi}_{sl} = \frac{3}{6} = \hat{\pi}_{sb} = \hat{\pi}_1 = \hat{\pi}_2 \quad \text{(Prior probabilities)}$$

Mean vectors:

$$\hat{\mu}_1 = \frac{1}{3}\left(\begin{bmatrix}10\\5\end{bmatrix} + \begin{bmatrix}11\\7\end{bmatrix} + \begin{bmatrix}12\\6\end{bmatrix}\right) = \begin{bmatrix}11\\6\end{bmatrix}, \quad \hat{\mu}_2 = \frac{1}{3}\left(\begin{bmatrix}5\\10\end{bmatrix} + \begin{bmatrix}7\\11\end{bmatrix} + \begin{bmatrix}6\\12\end{bmatrix}\right) = \begin{bmatrix}6\\11\end{bmatrix}$$

Covariance matrices: Class 1:

$$\hat{\Sigma}_1 = \frac{1}{3-1}\sum_{i=1}^{3}(\vec{x}_1^{(i)} - \hat{\mu}_1)(\vec{x}_1^{(i)} - \hat{\mu}_1)^T = \frac{1}{2}\left(\begin{bmatrix}-1\\-1\end{bmatrix}[-1\ -1] + \begin{bmatrix}0\\1\end{bmatrix}[0\ 1] + \begin{bmatrix}1\\0\end{bmatrix}[1\ 0]\right)$$

$$\hat{\Sigma}_1 = \frac{1}{2}\left(\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\right) = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}\frac{1}{2} = \begin{bmatrix}1 & 0.5\\0.5 & 1\end{bmatrix}$$

$\sigma_{11}^2 = 1$ is the variance of the first entry of the feature vector for class 1.

class 1

$\sigma_{12}^2 = 1$ " " " " " 2nd " " for class 1.

" " " 2nd entry of the feature vector.

1st entry of the feature vector

$E[(x_1 - \mu_{11})(x_2 - \mu_{12})]$ is estimated as $0.5$.

Class 2:

$$\hat{\Sigma}_2 = \frac{1}{3-1}\sum_{i=1}^{3}(\vec{x}_2^{(i)} - \hat{\mu}_2)(\vec{x}_2^{(i)} - \mu_2)^T = \frac{1}{2}\left(\begin{bmatrix}-1\\-1\end{bmatrix}[-1\ -1] + \begin{bmatrix}0\\1\end{bmatrix}[0\ 1] + \begin{bmatrix}1\\0\end{bmatrix}[1\ 0]\right)$$

$$\hat{\Sigma}_2 = \begin{bmatrix}1 & 0.5\\0.5 & 1\end{bmatrix}, \quad \sigma_{21}^2 = 1, \quad \sigma_{22}^2 = 1$$

Variance of the 1st entry of the feature vector for class 2.

It turns out that

$$\hat{\Sigma}_1 = \hat{\Sigma}_2$$

Generative
Models based on Gaussians.

$\omega_1$ has a multivariate " distribution

$$P_1 : \quad N\left(\begin{bmatrix} 11 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right)$$

$\omega_2$ " " " " "

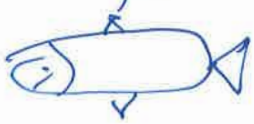
$$P_2 : \quad N\left(\begin{bmatrix} 6 \\ 11 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right)$$

This is the end of training!

---

Testing phase:

Catch a fish

 $\rightarrow \vec{x} = \begin{bmatrix} 7 \\ 12 \end{bmatrix}$

We calculate the likelihood of X for class 1 and 2.

$$P_1(\bar{x}) = P(x|\omega_1) = \frac{1}{\sqrt{(2\pi)^2 |\hat{\Sigma}_1|}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} 7 \\ 12 \end{bmatrix} - \begin{bmatrix} 11 \\ 6 \end{bmatrix}\right)^T \hat{\Sigma}_1^{-1} [-4 \ \ 6]^T\right)$$

$$P_2(\bar{x}) = P(X|\omega_2) = \frac{1}{\sqrt{(2\pi)^2 |\hat{\Sigma}_2|}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} 7 \\ 12 \end{bmatrix} - \begin{bmatrix} 6 \\ 11 \end{bmatrix}\right)^T \hat{\Sigma}_2^{-1} [1 \ \ 1]^T\right)$$

$\leftarrow$ Determinant of $\Sigma_2$ matrix.

where $\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 12 \end{bmatrix} - \begin{bmatrix} 6 \\ 11 \end{bmatrix}$, $\begin{bmatrix} -4 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 12 \end{bmatrix} - \begin{bmatrix} 11 \\ 6 \end{bmatrix}$

$$\pi_2 P_2(\vec{x}) > \pi_1 P_1(\bar{x}) \Rightarrow \text{class } 2$$

$$0.5 \times 0.094 > 0.5 \quad 1.8 \times 10^{-23}$$

---

In Practice ~~[illegible]~~ we also add $cI$ to $\Sigma$ matrices.

$$\Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \rightarrow \Sigma_1 + cI = \begin{bmatrix} 1+c & 0.5 \\ 0.5 & 1+c \end{bmatrix}$$

small number: $c = 0.0001$

It should be smaller than $\sigma_1, \sigma_2$