SMAI HW-6 Problem 3

a) $$P(\omega_1) = \frac{\text{Number of samples in } \omega_1}{\text{Total no. of samples}}$$

$$= \frac{7}{14} = \frac{1}{2}$$

$$P(\omega_2) = \frac{\#\text{ of samples in } \omega_2}{\text{Total no. of samples}}$$

$$= \frac{7}{14} = \frac{1}{2}.$$

b) $$\text{Mean of } \omega_1 = \begin{bmatrix} \frac{0+0+2+3+3+2+2}{7} \\ \frac{0+1+0+2+3+2+0}{7} \end{bmatrix}$$

$$= \begin{bmatrix} 1.714 \\ 1.143 \end{bmatrix}$$

$$Cov(x, Y) = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{7}$$

$$= 0.881.$$

$$Var(x) = \frac{\Sigma (x_i - \bar{x})^2}{7} = 1.571$$

$$Var(y) = \frac{\Sigma (y_i - \bar{y})^2}{7} = 1.476.$$

$$\text{Covariance matrix of } \omega_1 = \begin{bmatrix} 1.571 & 0.881 \\ 0.881 & 1.476 \end{bmatrix}$$

Mean of ω_2 = $\begin{bmatrix} \dfrac{7+8+9+8+7+8+7}{7} \\ \\ \dfrac{7+6+7+10+10+9+11}{7} \end{bmatrix}$

$$= \begin{bmatrix} 7.714 \\ 8.571 \end{bmatrix}$$

$$\text{Cov}(x,y) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{7}$$

$$= -0.643$$

$$\text{Var}(x) = \frac{\sum_i (x_i - \bar{x})^2}{7} = 0.571$$

$$\text{Var}(y) = \frac{\sum_i (y_i - \bar{y})^2}{7} = 3.619$$

$\therefore$ Covariance matrix of ω_2 = $\begin{bmatrix} 0.571 & -0.643 \\ -0.643 & 3.619 \end{bmatrix}$

(c) The decision boundary is the locus of all points where ~~$P(\omega_1) =$~~ posterior probability of sample belonging to class ω_1 and ω_2 are equal.

Let $\vec{v}$ be our sample

$$P(\vec{v} \mid \omega_1)\, P(\omega_1) = P(\vec{v} \mid \omega_2)\, P(\omega_2)$$

$$\frac{1}{\sqrt{2\pi}\, |\Sigma_1|^{1/2}} \exp\left(-\frac{1}{2} (\vec{v} - \mu_1)^T \Sigma_1^{-1} (\vec{v} - \mu_1) \right) P(\omega_1)$$

$$= \frac{1}{\sqrt{2\pi}\, |\Sigma_2|^{1/2}} \exp\left(-\frac{1}{2} (\vec{v} - \mu_2)^T \Sigma_2^{-1} (\vec{v} - \mu_2) \right) P(\omega_2)$$

$$\Rightarrow -\ln |\Sigma_1|^{1/2} - \frac{1}{2} [\vec{v} - \mu_1]^T \Sigma_1^{-1} [\vec{v} - \mu_1]$$

$$= -\ln |\Sigma_2|^{1/2} - \frac{1}{2} [\vec{v} - \mu_2]^T \Sigma_2^{-1} [\vec{v} - \mu_2]$$

$$\Rightarrow -\frac{1}{2} \ln |\Sigma_1| - \frac{1}{2} \left[(\vec{v} - \vec{\mu}_1)^T \Sigma_1^{-1} (\vec{v} - \mu_1) \right]$$

$$= -\frac{1}{2} \ln |\Sigma_2| - \frac{1}{2} \left[(\vec{v} - \vec{\mu}_2)^T \Sigma_2^{-1} (\vec{v} - \vec{\mu}_2) \right]$$

$$\Rightarrow \ln |\Sigma_1| + (\vec{v} - \vec{\mu}_1)^T \Sigma_1^{-1} (\vec{v} - \vec{\mu}_1)$$

$$= \ln |\Sigma_2| + (\vec{v} - \vec{\mu}_2) \Sigma_2^{-1} (\vec{v} - \vec{\mu}_2).$$

$\ln |\Sigma_1| - \ln |\Sigma_2| = 0.0695$ (written code)

$$\begin{bmatrix} x - 1.714 & y - 1.142 \end{bmatrix} \begin{bmatrix} 0.956 & -0.571 \\ -0.571 & 1.018 \end{bmatrix} \begin{bmatrix} x - 1.714 \\ y - 1.142 \end{bmatrix}$$

$$= \begin{bmatrix} x - 7.714 & y - 8.571 \end{bmatrix} \begin{bmatrix} 2.187 & 0.3884 \\ 0.3884 & 0.3453 \end{bmatrix} \begin{bmatrix} x - 7.714 \\ y - 8.571 \end{bmatrix}$$

$$+ 0.0695.$$

Note: I wrote code to compute Σ_1^{-1} & Σ_2^{-1}. Used inbuilt pinv().

$$\Rightarrow \begin{bmatrix} x-1.714 & y-1.142 \end{bmatrix} \begin{bmatrix} 0.956x - 1.639 - 0.571y + .652 \\ -0.571x + 0.979 + 1.018y - 1.163 \end{bmatrix}$$

$$- 0.0695 .$$

$$= \begin{bmatrix} x-7.714 & y-8.5714 \end{bmatrix} \begin{bmatrix} 2.187x - 16.871 + 0.3884y - 3.308 \\ 0.3884x - 2.996 + 0.3453y - 2.941 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x-1.714 & y-1.142 \end{bmatrix} \begin{bmatrix} 0.956x - 0.571y - 0.987 \\ -0.571x + 1.018y - 0.184 \end{bmatrix}$$

$$- 0.0695$$

$$= \begin{bmatrix} x-7.714 & y-8.571 \end{bmatrix} \begin{bmatrix} 2.187x - 20.179 + .3884y \\ 0.3884x - 5.937 + .3453y \end{bmatrix}$$

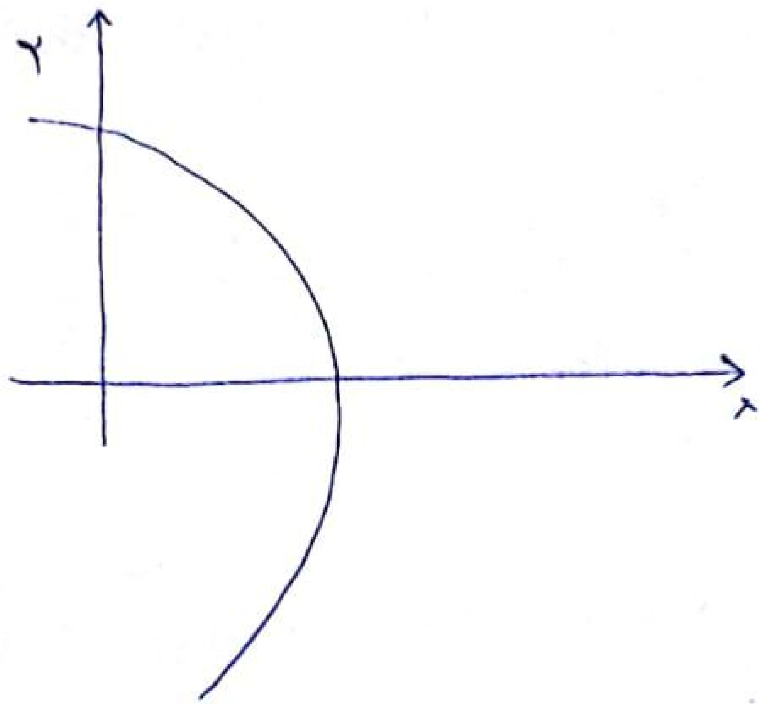
$$\Rightarrow 0.956x^2 + 0.987x - 0.571xy - 1.638x + 1.692$$
$$+ 0.979y - 0.571xy + 1.018y^2 - 0.184y +$$
$$0.652x - 1.163y + 0.2101 - 0.0695$$

$$= 2.187x^2 - 20.179x + 0.3884xy - 16.871x$$
$$+ 155.665 - 2.996y + 0.3884xy - 5.937y$$
$$+ 0.3453y^2 - 3.3291x + 50.888 - 2.9597y$$

$$\Rightarrow 1.231x^2 - 0.6727y^2 + 1.9188xy - 40.3801x - 11.5247y + 204.7203 = 0$$

is the equation of the decision boundary.

Rough sketch.



e) Since misclassifying ω_1 as ω_2 is twice as costly than misclassifying ω_2 as ω_1, this is a clear indication to increase the prior probability of ω_1. We can have many ω_1 and this won't affect us much. In other words classifying ω_2 as ω_1 is fine but classifying ω_1 as ω_2 will recieve a very heavy penalty.

$\therefore P(\omega_1) = 2P(\omega_2)$

$P(\omega_1) = \frac{2}{3} \qquad P(\omega_2) = \frac{1}{3}$.

$\therefore$ Equation of decision boundary gets modified only by the term $\ln|P(\omega_1)| - \ln P(\omega_2)$

$$= \ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right) = \ln 2.$$

$$1.231x^2 - 0.6727y^2 + 1.9188xy - 40.3801x - 11.5487y$$
$$+ 204.72 + 2\ln(2) = 0$$

The net effect is the slight shift of boundary towards ω_2.

Mathematically defining a function to capture our possible errors.

$$f(\vec{v}) = \begin{cases} \omega_2 & , \vec{v} \in \omega_1. \\ \omega_1 & , \vec{v} \in \omega_2 \end{cases}$$

ω_1, ω_2 are 2 classes in pattern space Ω with continuous PDFs $p_1(x)$ & $p_2(x)$ respectively. We can see this as dividing space Ω into Ω_1 & Ω_2.

[$\Omega_1 \cup \Omega_2 = \Omega$ & $\Omega_1 \cap \Omega_2 = \phi$
i.e mutually exclusive & exhaustive]

$$\text{Prob. of error 1} = \int_{\Omega_2} p_1(x)\,dx.$$

$$\text{Prob. of error}_2 = \int_{\Omega_1} p_2(x)\,dx.$$

Let c_1 be cost of misclassifying ω_1 into $\omega_2 = \frac{2}{3}$ &

c_2 be cost of misclassifying ω_2 into ω_1

Total expected error (E)

$$= \frac{2}{3} P(\omega_1) \int_{\Omega_2} p_1(x)\, dx + \frac{1}{3} P(\omega_2) \int_{\Omega_1} p_2(x)\, dx$$

$$= \frac{1}{3} \left[2P(\omega_1) \int_{\Omega_2} p_1(x)\, dx + \{1 - P(\omega_1)\} \int_{\Omega_1} p_2(x)\, dx \right]$$

which is what we have to minimise.

The magnitude of shift in decision boundary or the above error will depend on other parameters.