1. (a) $p(\underline{x}|S_i) = N(\underline{x}, \underline{m}_i, \underline{\underline{\Sigma}}_i) = \frac{1}{(2\pi)^{D/2}|\underline{\underline{\Sigma}}_i|^{1/2}} \exp\{-\frac{1}{2} d_M^2(\underline{x}, \underline{m}_i, \underline{\underline{\Sigma}}_i)\}$

$p(\underline{x}|S_1)\cdot \cancel{P(S_1)} \underset{S_2}{\overset{S_1}{\gtrless}} p(\underline{x}|S_2)\cdot \cancel{P(S_2)} \quad \because P(S_1) = P(S_2) = 0.5$

$g_1(\underline{x}) = \ln\{p(\underline{x}|S_1)\} = -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| - \frac{1}{2}(\underline{x}-\underline{m}_1)^T \underline{\underline{\Sigma}}^{-1}(\underline{x}-\underline{m}_1)$

$= \underbrace{-\frac{D}{2}\ln(2\pi)}_{\text{can be neglect}} - \frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| - \frac{1}{2}(\underline{x}^T\underline{\underline{\Sigma}}_1^{-1}\underline{x} - \underbrace{2\underline{x}^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1}_{2\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{x}} + \underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1)$

$g_1(\underline{x}) \underset{S_2}{\overset{S_1}{\gtrless}} g_2(\underline{x})$

① $\Rightarrow -\frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| - \frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_1^{-1}\underline{x} + \underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{x} - \frac{1}{2}\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1 \underset{S_2}{\overset{S_1}{\gtrless}} -\frac{1}{2}\ln|\underline{\underline{\Sigma}}_2| - \frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_2^{-1}\underline{x} + \underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{x} - \frac{1}{2}\underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{m}_2$

$\Rightarrow (-\frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_1^{-1} + \underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1} + \frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_2^{-1} - \underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1})\underline{x} \underset{S_2}{\overset{S_1}{\gtrless}} \frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| - \frac{1}{2}\ln|\underline{\underline{\Sigma}}_2| + \frac{1}{2}\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1 - \frac{1}{2}\underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{m}_2$

$\Rightarrow (-\frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_1^{-1} + \underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1} + \frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_2^{-1} - \underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1})\underline{x} \underset{S_2}{\overset{S_1}{\gtrless}} -(\frac{1}{2}\ln|\underline{\underline{\Sigma}}_2| - \frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| + \frac{1}{2}\underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{m}_2 - \frac{1}{2}\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1)$

$\Rightarrow -\frac{1}{2}\underline{x}^T(\underline{\underline{\Sigma}}_1^{-1} - \underline{\underline{\Sigma}}_2^{-1})\underline{x} + (\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1} - \underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1})\underline{x} \underset{S_2}{\overset{S_1}{\gtrless}} \frac{1}{2}\ln\frac{|\underline{\underline{\Sigma}}_1|}{|\underline{\underline{\Sigma}}_2|} + \frac{1}{2}(\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1 - \underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{m}_2)$

$\because$ it's quadratic function, it cannot be expressed in a linear form.

(b) Look ① in (a),

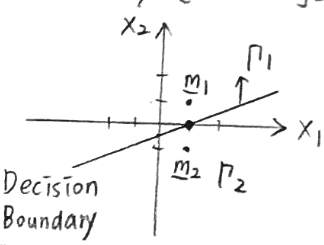
② $-\cancel{\frac{1}{2}\ln|\underline{\underline{\Sigma}}_1|} - \cancel{\frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_1^{-1}\underline{x}} + \underline{m}_1\underline{\underline{\Sigma}}_1^{-1}\underline{x} - \frac{1}{2}\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1 \gtrless -\cancel{\frac{1}{2}\ln|\underline{\underline{\Sigma}}_2|} - \cancel{\frac{1}{2}\underline{x}^T\underline{\underline{\Sigma}}_2^{-1}\underline{x}} + \underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{x} - \frac{1}{2}\underline{m}_2^T\underline{\underline{\Sigma}}_2^{-1}\underline{m}_2$

$\because \underline{\underline{\Sigma}}_1 = \underline{\underline{\Sigma}}_2 \Rightarrow \underline{m}_1^T\underline{\underline{\Sigma}}^{-1}\underline{x} - \frac{1}{2}\underline{m}_1^T\underline{\underline{\Sigma}}^{-1}\underline{m}_1 \gtrless \underline{m}_2^T\underline{\underline{\Sigma}}^{-1}\underline{x} - \frac{1}{2}\underline{m}_2^T\underline{\underline{\Sigma}}^{-1}\underline{m}_2 \qquad \underline{\underline{\Sigma}} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2.25 \end{bmatrix} \Rightarrow \underline{\underline{\Sigma}}^{-1} = \frac{1}{2}\begin{bmatrix} 2.25 & -0.5 \\ -0.5 & 1 \end{bmatrix}$

$\Rightarrow [1\ 1]\cancel{\frac{1}{2}}\begin{bmatrix} 2.25 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2}[1\ 1]\cancel{\frac{1}{2}}\begin{bmatrix} 2.25 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \gtrless [1\ -1]\cancel{\frac{1}{2}}\begin{bmatrix} 2.25 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2}[1\ -1]\cancel{\frac{1}{2}}\begin{bmatrix} 2.25 & -0.5 \\ -0.5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\Rightarrow [1.75\ 0.5]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{2.25}{2} \underset{S_2}{\overset{S_1}{\gtrless}} [2.75\ -1.5]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{4.25}{2}$

$\Rightarrow 1.75x_1 + 0.5x_2 - 2.75x_1 + 1.5x_2 + 1 \underset{S_2}{\overset{S_1}{\gtrless}} 0 \quad \Rightarrow -x_1 + 2x_2 + 1 \underset{S_2}{\overset{S_1}{\gtrless}} 0$



Decision Boundary

2. (a) $g_1(\underline{x}) = \ln\{p(\underline{x}|S_1)\} = \ln\{\prod_{j=1}^{D} p(x_j|S_1)\} = \sum_{j=1}^{D} \ln\{p(x_j|S_1)\}$, $\quad p(x_j|S_1) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_j}\exp\{-\frac{(x_j-m_j)^2}{2\sigma_j^2}\}$

$= \sum_{j=1}^{D}\left[-\frac{1}{2}\ln(2\pi) - \ln\sigma_j - \frac{1}{2}\frac{(x_j-m_j)^T(x_j-m_j)}{\sigma_j^2}\right]$

$= -\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| - \frac{1}{2}(\underline{x}-\underline{m}_1)^T\underline{\underline{\Sigma}}_1^{-1}(\underline{x}-\underline{m}_1)$

$= \underbrace{-\frac{D}{2}\ln(2\pi)}_{\text{can be neglected}} - \frac{1}{2}\ln|\underline{\underline{\Sigma}}_1| - \frac{1}{2}(\underline{x}^T\underline{\underline{\Sigma}}_1^{-1}\underline{x} - 2\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{x} + \underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1)$ which is the same with problem 1 (a)

We can use the same results in problem 1 (a)

$\Rightarrow -\frac{1}{2}\underline{x}^T(\underline{\underline{\Sigma}}_1^{-1} - \underline{\underline{\Sigma}}_2^{-1})\underline{x} + (\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1} - \underline{m}_2\underline{\underline{\Sigma}}_2^{-1})\underline{x} \underset{S_2}{\overset{S_1}{\gtrless}} \frac{1}{2}\ln\frac{|\underline{\underline{\Sigma}}_1|}{|\underline{\underline{\Sigma}}_2|} + \frac{1}{2}(\underline{m}_1^T\underline{\underline{\Sigma}}_1^{-1}\underline{m}_1 - \underline{m}_2\underline{\underline{\Sigma}}_2^{-1}\underline{m}_2)$

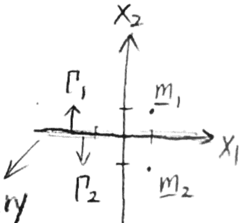
Because we know that the features are assumed independent, and $\sigma_1^{(i)2} = 1$, $\sigma_2^{(i)2} = 2.25$

$\underline{\underline{\Sigma}}_1 = \underline{\underline{\Sigma}}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2.25 \end{bmatrix} = \underline{\underline{\Sigma}} \qquad \underline{\underline{\Sigma}}^{-1} = \begin{bmatrix} 2.25 & 0 \\ 0 & 1 \end{bmatrix} \times \frac{1}{2.25}$

from ② in problem 1. (b) that $\underline{m}_1^T\underline{\underline{\Sigma}}^{-1}\underline{x} - \frac{1}{2}\underline{m}_1^T\underline{\underline{\Sigma}}^{-1}\underline{m}_1 \gtrless \underline{m}_2^T\underline{\underline{\Sigma}}^{-1}\underline{x} - \frac{1}{2}\underline{m}_2^T\underline{\underline{\Sigma}}^{-1}\underline{m}_2$

$\Rightarrow [1\ 1]\frac{1}{2.25}\begin{bmatrix} 2.25 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2}[1\ 1]\frac{1}{2.25}\begin{bmatrix} 2.25 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \gtrless [1\ -1]\begin{bmatrix} 2.25 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2}[1\ -1]\frac{1}{2.25}\begin{bmatrix} 2.25 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\Rightarrow [2.25\ 1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \cancel{\frac{3.25}{2}} \gtrless [2.25\ -1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \cancel{\frac{3.25}{2}} \Rightarrow x_2 \underset{S_2}{\overset{S_1}{\gtrless}} 0$, x-axis is the decision boundary



3. $\nearrow$ conditional risk when $\underline{x}$ is defined as class 1

$R(\alpha_1|\underline{x}) = \lambda_{11}P(S_1|\underline{x}) + \lambda_{12}P(S_2|\underline{x})$

$R(\alpha_2|\underline{x}) = \lambda_{21}P(S_1|\underline{x}) + \lambda_{22}P(S_2|\underline{x})$

decide S_1 if $R(\alpha_1|\underline{x}) < R(\alpha_2|\underline{x})$

$\Sigma_1 = \Sigma_2 = \Sigma = \begin{bmatrix} 1 & 0,5 \\ 0,5 & 2,25 \end{bmatrix}$

$\Sigma^{-1} = \frac{1}{2}\begin{bmatrix} 2,25 & -0,5 \\ -0,5 & 1 \end{bmatrix}$

$$\Rightarrow \underbrace{\lambda_{11}}_{0}P(S_1|\underline{x}) + \underbrace{\lambda_{12}}_{=2\lambda_{21}}P(S_2|\underline{x}) \underset{S_2}{\overset{S_1}{\lessgtr}} \lambda_{21}P(S_1|\underline{x}) + \underbrace{\lambda_{22}}_{0}P(S_2|\underline{x})$$

$$\Rightarrow 2\cancel{\lambda_{21}}P(S_2|\underline{x}) \underset{S_2}{\overset{S_1}{\lessgtr}} \cancel{\lambda_{21}}P(S_1|\underline{x})$$

$$\Rightarrow 2\frac{P(S_2\cap\underline{x})}{\cancel{P(\underline{x})}} \underset{S_2}{\overset{S_1}{\lessgtr}} \frac{P(S_1\cap\underline{x})}{\cancel{P(\underline{x})}} \Rightarrow 2\,p(\underline{x}|S_2)\cdot P(S_2) \underset{S_2}{\overset{S_1}{\lessgtr}} p(\underline{x}|S_1)\,P(S_1)$$

$$P(S_1) = P(S_2) = 0,5 \Longrightarrow 2\,p(\underline{x}|S_2) \underset{S_2}{\overset{S_1}{\lessgtr}} p(\underline{x}|S_1)$$

$$\Rightarrow \ln 2 + \ln\{p(\underline{x}|S_2)\} \underset{S_2}{\overset{S_1}{\lessgtr}} \ln\{p(\underline{x}|S_1)\}$$

$\Sigma_1 = \Sigma_2 = \Sigma$

$$\Rightarrow \ln 2 - \cancel{\tfrac{D}{2}\ln(2\pi)} - \cancel{\tfrac{1}{2}\ln|\Sigma_2|} - \tfrac{1}{2}(\underline{x}-\underline{m}_2)^T\Sigma_2^{-1}(\underline{x}-\underline{m}_2) \underset{S_2}{\overset{S_1}{\lessgtr}} -\cancel{\tfrac{D}{2}\ln(2\pi)} - \cancel{\tfrac{1}{2}\ln|\Sigma_1|} - \tfrac{1}{2}(\underline{x}-\underline{m}_1)^T\Sigma_1^{-1}(\underline{x}-\underline{m}_1)$$

$$\Rightarrow \ln 2 - \tfrac{1}{2}(\cancel{\underline{x}^T\Sigma_2^{-1}\underline{x}} - 2\underline{x}^T\Sigma_2^{-1}\underline{m}_2 + \underline{m}_2^T\Sigma_2^{-1}\underline{m}_2) \underset{S_2}{\overset{S_1}{\lessgtr}} -\tfrac{1}{2}(\cancel{\underline{x}^T\Sigma_1^{-1}\underline{x}} - 2\underline{x}^T\Sigma_1^{-1}\underline{m}_1 + \underline{m}_1^T\Sigma_1^{-1}\underline{m}_1)$$

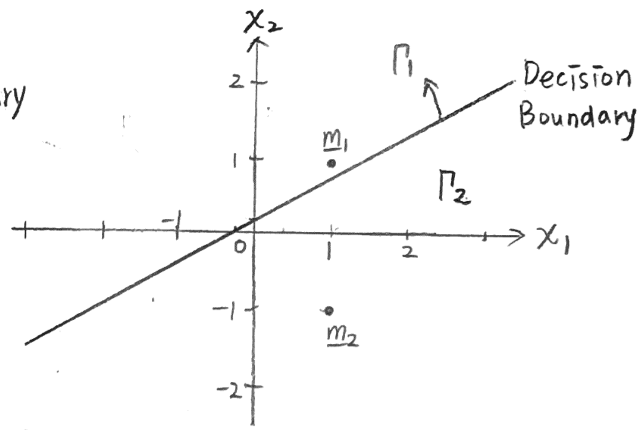
$$\Rightarrow \ln 2 + \underline{x}^T\Sigma^{-1}\underline{m}_2 - \tfrac{1}{2}\underline{m}_2^T\Sigma^{-1}\underline{m}_2 \underset{S_2}{\overset{S_1}{\lessgtr}} \underline{x}^T\Sigma^{-1}\underline{m}_1 - \tfrac{1}{2}\underline{m}_1^T\Sigma^{-1}\underline{m}_1 \longrightarrow \text{Decision boundary}$$

$$\Rightarrow \ln 2 + [x_1\ x_2]\tfrac{1}{2}\begin{bmatrix} 2,25 & -0,5 \\ -0,5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} - \tfrac{1}{2}[1\ -1]\cdot\tfrac{1}{2}\begin{bmatrix} 2,25 & -0,5 \\ -0,5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \underset{S_2}{\overset{S_1}{\lessgtr}} [x_1\ x_2]\tfrac{1}{2}\begin{bmatrix} 2,25 & -0,5 \\ -0,5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \tfrac{1}{2}[1\ 1]\tfrac{1}{2}\begin{bmatrix} 2,25 & -0,5 \\ -0,5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \ln 2 + \tfrac{1}{2}(2,75x_1 - 1,5x_2) - \tfrac{1}{4}\cdot 4,25 \underset{S_2}{\overset{S_1}{\lessgtr}} \tfrac{1}{2}(1,75x_1 + 0,5x_2) - \tfrac{1}{4}\cdot 2,25$$

$$\times 2 \Rightarrow 2\ln 2 + 2,75x_1 - 1,5x_2 - 1 \underset{S_2}{\overset{S_1}{\lessgtr}} 1,75x_1 + 0,5x_2$$

$$\Rightarrow -x_1 + 2x_2 + 1 - \underbrace{2\ln 2}_{\fallingdotseq 1,386} \underset{S_2}{\overset{S_1}{\gtrless}} 0 \rightarrow \text{decision boundary}$$



compare the left plot to the plot of Problem 1, Γ_2 region has grown because of the given losses.

$\because \lambda_{12} = 2\lambda_{21}$

$\Rightarrow \lambda_{12} > \lambda_{21}$

$\Rightarrow S_1$ incurs more loss.