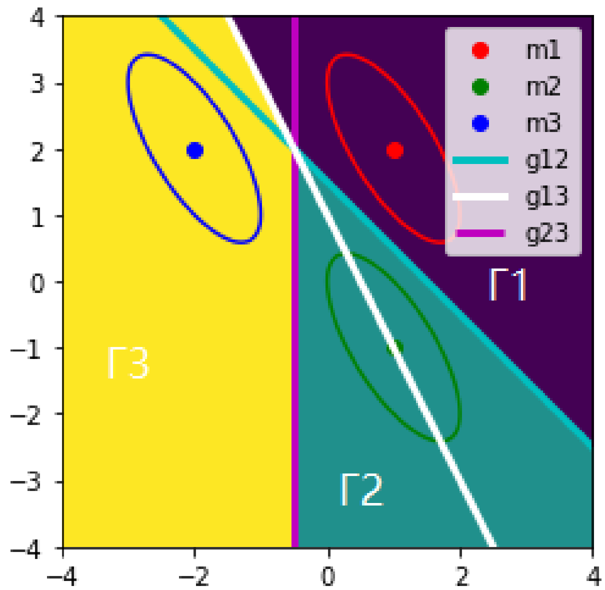

Fig 1. Plot of 1-(c)-(iii)
The horizontal axis denotes x1, the vertical axis denotes x2,
The red, blue, green circle denote $d_M^2(\underline{x}, \underline{m_1}) = 1$, $d_M^2(\underline{x}, \underline{m_2}) = 1$, $d_M^2(\underline{x}, \underline{m_3}) = 1$ respectively,
The purple, cyan, yellow region denote decision regions of Γ1, Γ2, Γ3 respectively,
The cyan, white, purple line denote decision boundaries for class 1 and 2, class 1 and 3, class 2 and 3 respectively.

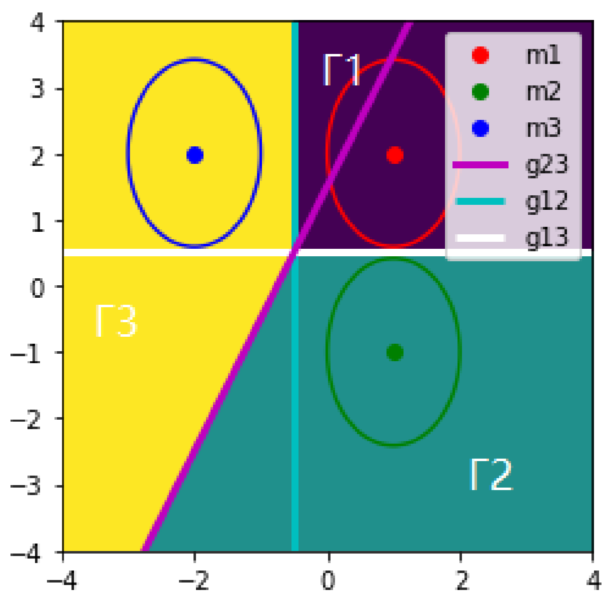

Fig 2. Plot of 2-(c)-(iii)
The horizontal axis denotes x1, the vertical axis denotes x2,
The red, blue, green circle denote $d_M^2(\underline{x}, \underline{m_1}) = 1$, $d_M^2(\underline{x}, \underline{m_2}) = 1$, $d_M^2(\underline{x}, \underline{m_3}) = 1$ respectively,
The purple, cyan, yellow region denote decision regions of Γ1, Γ2, Γ3 respectively,
The cyan, white, purple line denote decision boundaries for class 1 and 2, class 1 and 3, class 2 and 3 respectively.

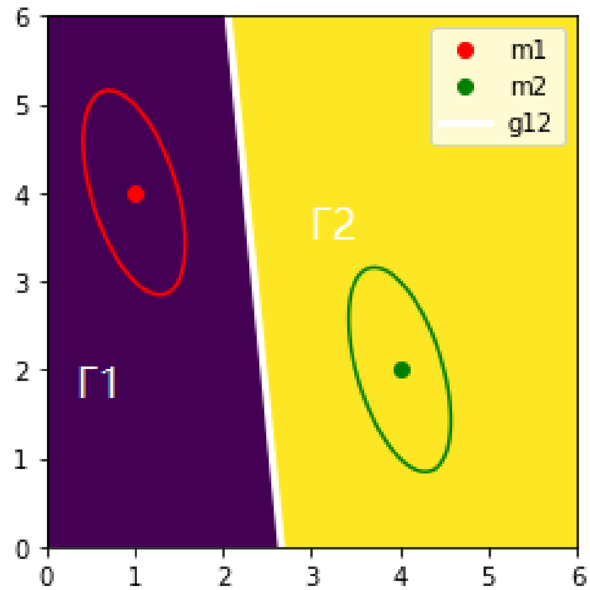

Fig 3. Plot of 3-(d)
The horizontal axis denotes x1, the vertical axis denotes x2,
The red, green circle denote $d_M^2(\underline{x}, \underline{m_1}) = 1$, $d_M^2(\underline{x}, \underline{m_2}) = 1$ respectively,
The purple, yellow region denote decision regions of Γ1, Γ2 respectively,
The white line denotes decision boundaries for class 1 and 2.

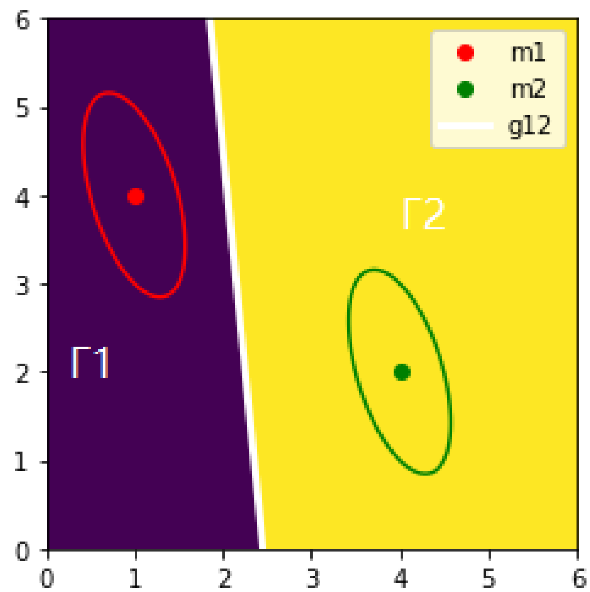

Fig 4. Plot of 3-(e)
The horizontal axis denotes x1, the vertical axis denotes x2,
The red, green circle denote $d_M^2(\underline{x}, \underline{m_1}) = 1$, $d_M^2(\underline{x}, \underline{m_2}) = 1$ respectively,
The purple, yellow region denote decision regions of Γ1, Γ2 respectively,
The white line denotes decision boundaries for class 1 and 2.

1.(a)(i). $g_i(\underline{x}) = -\frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(\underline{x}-\underline{m}_i)^T\Sigma_i^{-1}(\underline{x}-\underline{m}_i) + \ln P(S_i)$.

$= -\frac{1}{2}d_M^2(\underline{x}, \underline{m}_i) - \frac{1}{2}\ln|\Sigma_i| + \ln\pi_i$.

(ii) $g_i(\underline{x}) = -\frac{1}{2}\ln|\Sigma_i| + \ln\pi_i - \frac{1}{2}(\underline{x}^T\Sigma_i^{-1} - m_i\Sigma_i^{-1})(\underline{x}-\underline{m}_i)$

$= -\frac{1}{2}\ln|\Sigma_i| + \ln\pi_i - \frac{1}{2}(\underline{x}^T\Sigma_i^{-1}\underline{x} - 2\underline{m}_i^T\Sigma_i^{-1}\underline{x} + \underline{m}_i^T\Sigma_i^{-1}\underline{m}_i)$.

$= \underline{x}^T(-\frac{1}{2}\Sigma_i^{-1})\underline{x} + (\underline{m}_i^T\Sigma_i^{-1})\underline{x} - \frac{1}{2}\underline{m}_i^T\Sigma_i^{-1}\underline{m}_i - \frac{1}{2}\ln|\Sigma_i| + \ln\pi_i$.

$W_i = -\frac{1}{2}\Sigma_i^{-1}$, $\quad w_i = \underline{m}_i^T\Sigma_i^{-1}$ $\quad w_0^{(i)} = -\frac{1}{2}\underline{m}_i^T\Sigma_i^{-1}\underline{m}_i - \frac{1}{2}\ln|\Sigma_i| + \ln\pi_i$.

(b). decision rule: $\begin{cases} \text{if } g_1(\underline{x}) > g_2(\underline{x}),\ g_1(\underline{x}) > g_3(\underline{x}), & \underline{x}\in S_1 \\ \text{if } g_2(\underline{x}) > g_1(\underline{x}),\ g_2(\underline{x}) > g_3(\underline{x}), & \underline{x}\in S_2 \\ \text{if } g_3(\underline{x}) > g_1(\underline{x}),\ g_3(\underline{x}) > g_2(\underline{x}), & \underline{x}\in S_3 \end{cases}$

if $\Sigma_1 = \Sigma_2 = \Sigma_3$, it is linear classifier, else, it is quadratic classifier

(c) (i) Because $\Sigma_1 = \Sigma_2 = \Sigma_3$, $g_i(\underline{x})$ could be simplified to $g_i(\underline{x}) = (\underline{m}_i^T\Sigma_i^{-1})\underline{x} - \frac{1}{2}\underline{m}_i^T\Sigma_i^{-1}\underline{m}_i + \ln\pi_i$.

~~So g~~ $\Sigma_i^{-1} = \begin{bmatrix}1 & -1\\ -1 & 2\end{bmatrix}^{-1} = \begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}$

So $g_1(\underline{x}) = [1, 2]\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\underline{x} - \frac{1}{2}[1, 2]\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1\\ 2\end{bmatrix} + \ln\pi_1$.

$= 4x_1 + 3x_2 - 5 + \ln\pi_1$.

$g_2(\underline{x}) = [1, -1]\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\underline{x} - \frac{1}{2}[1, -1]\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix} + \ln\pi_2$

$= x_1 - \frac{1}{2} + \ln\pi_2$

$g_3(\underline{x}) = [-2, 2]\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\underline{x} - \frac{1}{2}[-2, 2]\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}-2\\ 2\end{bmatrix} + \ln\pi_3$

$= -2x_1 - 2 + \ln\pi_3$ This is a linear classifier

(ii) decision boundary: 1,2: $g_1(\underline{x}) = g_2(\underline{x}) \Rightarrow x_1 + x_2 - \frac{3}{2} = 0$

1,3: $g_1(\underline{x}) = g_3(\underline{x}) \Rightarrow 2x_1 + x_2 - 1 = 0$

2,3: $g_2(\underline{x}) = g_3(\underline{x}) \Rightarrow x_1 + \frac{1}{2} = 0$

(iii) $d_M^2(\underline{x}, \underline{m}_i) = 1 \Rightarrow 2(x_1 - m_1)^2 + 2(x_1 - m_1)(x_2 - m_2) + (x_2 - m_2)^2 = 1$

$d_M^2(\underline{x}, \underline{m}_1) = 1 \Rightarrow 2(x_1 - 1)^2 + 2(x_1 - 1)(x_2 - 2) + (x_2 - 2)^2 = 1$

$d_M^2(\underline{x}, \underline{m}_2) = 1 \Rightarrow 2(x_1 - 1)^2 + 2(x_1 - 1)(x_2 + 1) + (x_2 + 1)^2 = 1$

$d_M^2(\underline{x}, \underline{m}_3) = 1 \Rightarrow 2(x_1 + 2)^2 + 2(x_1 + 2)(x_2 - 2) + (x_2 - 2)^2 = 1$

Plot is shown in Fig. 1.

2. (a). $g_i(x) = -\frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x-m_i)^T \Sigma_i^{-1}(x-m_i) + \ln P(S_i)$.

$|\Sigma_i| = (\sigma_1^{(i)})^2(\sigma_2^{(i)})^2$. $\Sigma_i^{-1} = \mathrm{diag}\left(\frac{1}{(\sigma_1^{(i)})^2}, \frac{1}{(\sigma_2^{(i)})^2}\right)$. $-\frac{1}{2}\ln|\Sigma_i| = -\ln \sigma_1^{(i)}\sigma_2^{(i)}$

$$g_i(x) = -\ln(\sigma_1^{(i)}\sigma_2^{(i)}) + \ln\pi_i - \frac{1}{2(\sigma_1^{(i)})^2}(x_1 - m_1^{(i)})^2 - \frac{1}{2(\sigma_2^{(i)})^2}(x_2 - m_2^{(i)})^2$$

$$= -\frac{1}{2(\sigma_1^{(i)})^2}x_1^2 + \frac{1}{(\sigma_1^{(i)})^2}m_1^{(i)}x_1 - \frac{1}{2(\sigma_2^{(i)})^2}x_2^2 + \frac{1}{(\sigma_2^{(i)})^2}m_2^{(i)}x_2 + \ln\pi_i - \ln(\sigma_1^{(i)}\sigma_2^{(i)}) - \frac{m_1^{(i)2}}{2(\sigma_1^{(i)})^2} - \frac{m_2^{(i)2}}{2(\sigma_2^{(i)})^2}$$

So $w_{11}^{(i)} = -\frac{1}{2(\sigma_1^{(i)})^2}$ $\quad w_{12}^{(i)} = 0$ $\quad w_{22}^{(i)} = -\frac{1}{2(\sigma_2^{(i)})^2}$

$w_1^{(i)} = \frac{m_1^{(i)}}{(\sigma_1^{(i)})^2}$ $\quad w_2^{(i)} = \frac{m_2^{(i)}}{(\sigma_2^{(i)})^2}$ $\quad w_0^{(i)} = \ln\pi_i - \ln(\sigma_1^{(i)}\sigma_2^{(i)}) - \frac{m_1^{(i)2}}{2(\sigma_1^{(i)})^2} - \frac{m_2^{(i)2}}{2(\sigma_2^{(i)})^2}$

(b). if ~~all~~ $\Sigma_1 = \Sigma_2 = \Sigma_3$ it is linear. else, it is quadratic.

(c)(i) $g_i(x) = \frac{m_1^{(i)}}{(\sigma_1^{(i)})^2}x_1 + \frac{m_2^{(i)}}{(\sigma_2^{(i)})^2}x_2 + \ln\pi_i - \frac{m_1^{(i)2}}{2(\sigma_1^{(i)})^2} - \frac{m_2^{(i)2}}{2(\sigma_2^{(i)})^2}$

$g_1(x) = x_1 + x_2 - \frac{3}{2} + \ln\pi_1$.

$g_2(x) = x_1 - \frac{1}{2}x_2 - \frac{3}{4} + \ln\pi_2$

$g_3(x) = -2x_1 + x_2 - 3 + \ln\pi_3$ This is a linear classifier

(ii). decision boundary: 1,2: $g_1(x) = g_2(x) \Rightarrow x_2 - \frac{1}{2} = 0$

1,3: $g_1(x) = g_3(x) \Rightarrow x_1 + \frac{1}{2} = 0$

2,3: $g_2(x) = g_3(x) \Rightarrow x_1 - \frac{1}{2}x_2 + \frac{3}{4} = 0$

(iii). $d_M^2(x, m_i) = 1 \Rightarrow (x_1 - m_1)^2 + \frac{1}{2}(x_2 - m_2)^2 = 1$

$d_M^2(x, m_1) = 1 \Rightarrow (x_1 - 1)^2 + \frac{1}{2}(x_2 - 2)^2 = 1$

$d_M^2(x, m_2) = 1 \Rightarrow (x_1 - 1)^2 + \frac{1}{2}(x_2 + 1)^2 = 1$

$d_M^2(x, m_3) = 1 \Rightarrow (x_1 + 2)^2 + \frac{1}{2}(x_2 - 2)^2 = 1$

Plot is shown in Fig. 2.

(d). ~~This~~ Both decision boundaries in 2(c)(iii) and 1(c)(iii) are linear, From 2(c)(iii) we would see that the ~~de~~ Naive Bayes makes the axis of ~~A~~ the oval denoting $d_M^2(x, m_i) = 1$ parallel to the coordinates ~~Because the form of decision boundaries is not changed~~ so it make a substantial difference in decision boundaries and regions.

3(a). decision rule: $R(\alpha_1|\underline{x}) \underset{S_1}{\overset{S_2}{\gtrless}} R(\alpha_2|\underline{x})$

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

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

In both side of the inequality, multiple by $p(\underline{x})$,

$\Rightarrow \lambda_{11} P(\underline{x}|S_1) P(S_1) + \lambda_{12} P(\underline{x}|S_2) P(S_2) \underset{S_1}{\overset{S_2}{\gtrless}} \lambda_{21} P(\underline{x}|S_1) P(S_1) + \lambda_{22} P(\underline{x}|S_2) P(S_2)$

$\Rightarrow (\lambda_{21} - \lambda_{11}) P(\underline{x}|S_1) P(S_1) \underset{S_2}{\overset{S_1}{\gtrless}} (\lambda_{12} - \lambda_{22}) P(\underline{x}|S_2) P(S_2)$

$\Rightarrow \ln(\lambda_{21} - \lambda_{11}) + \ln P(S_1) + \ln P(\underline{x}|S_1) \underset{S_2}{\overset{S_1}{\gtrless}} \ln(\lambda_{12} - \lambda_{22}) + \ln P(S_2) + \ln P(\underline{x}|S_2)$

$\Rightarrow \ln(\lambda_{21} - \lambda_{11}) + \ln P(S_1) - \frac{1}{2}\ln|\Sigma_1| - \frac{1}{2} d_M^2(\underline{x}, \underline{m}_1) \underset{S_2}{\overset{S_1}{\gtrless}} \ln(\lambda_{12} - \lambda_{22}) + \ln P(S_2) - \frac{1}{2}\ln|\Sigma_2| - \frac{1}{2} d_M^2(\underline{x}, \underline{m}_2)$

(b) $P(S_1) = 0.8 \quad P(S_2) = 0.2$

(c). $\Sigma_1^{-1} = \frac{2}{3}\begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}$.

$d_M^2(\underline{x}, \underline{m}_1) = (x_1 - 1, x_2 - 4) \cdot \frac{2}{3}\begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 - 1 \\ x_2 - 4 \end{pmatrix} = \frac{2}{3}[4(x_1-1)^2 + 2(x_1-1)(x_2-4) + (x_2-4)^2]$

$= \frac{2}{3}(4x_1^2 - 8x_1 + 4 + 2x_1x_2 - 8x_1 - 2x_2 + 8 + x_2^2 - 8x_2 + 16)$

$d_M^2(\underline{x}, \underline{m}_2) = (x_1 - 4, x_2 - 2) \cdot \frac{2}{3}\begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 - 4 \\ x_2 - 2 \end{pmatrix} = \frac{2}{3}[4(x_1-4)^2 + 2(x_1-4)(x_2-2) + (x_2-2)^2]$

$= \frac{2}{3}(4x_1^2 - 32x_1 + 64 + 2x_1x_2 - 4x_1 - 8x_2 + 16 + x_2^2 - 4x_2 + 4)$

$\ln(1-0) + \ln 0.8 - \frac{1}{3}(-16x_1 - 10x_2 + 28) \underset{S_2}{\overset{S_1}{\gtrless}} \ln(10-0) + \ln 0.2 - \frac{1}{3}(-36x_1 - 12x_2 + 84)$

$\Rightarrow g_1(\underline{x}) = \ln 1 + \ln 0.8 - \frac{1}{3}(-16x_1 - 10x_2 + 28)$

$g_2(\underline{x}) = \ln 10 + \ln 0.2 - \frac{1}{3}(-36x_1 - 12x_2 + 84)$

~~decision boundary~~ $g_{12}(\underline{x}) = \ln 1 + \ln 0.8 - \ln 10 - \ln 0.2 - \frac{1}{3}(-16x_1 - 10x_2 + 28) + \frac{1}{3}(-36x_1 - 12x_2 + 84)$

$= -\frac{20}{3}x_1 - \frac{2}{3}x_2 + \frac{56}{3} + \ln 0.4$

$\approx -\frac{20}{3}x_1 - \frac{2}{3}x_2 + 17.75$

So decision boundary is $-\frac{20}{3}x_1 - \frac{2}{3}x_2 + \frac{56}{3} + \ln 0.4 = 0$

decision rule: if $-\frac{20}{3}x_1 - \frac{2}{3}x_2 + \frac{56}{3} + \ln 0.4 > 0$, $\underline{x} \in S_1$

if $-\frac{20}{3}x_1 - \frac{2}{3}x_2 + \frac{56}{3} + \ln 0.4 < 0$, $\underline{x} \in S_2$

(d). See Fig. 3

(e). See Fig. 4. the decision boundary moves slightly towards m_1, and in the same range, as Fig. 3, the area of decision region of S_1 become smaller