CS171 PS1

- Name: Tsung-Ying Chen
- SID: 861310198
- Date: 10/14/2017

We will assume a binary classification problem with classes A and B with prior probabilities of p_A and p_B, and that the width or variance of the class-conditional distributions, σ, is the same for both classes. Thus, if μ_A is the center of $p(x \mid y = A)$ and μ_B is the same for $y = B$,

$$p(x \mid y = A) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu_A)^T(x-\mu_A)} \qquad p(y = A) = p_A$$

$$p(x \mid y = B) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu_B)^T(x-\mu_B)} \qquad p(y = B) = p_B$$

- **If these are the true distributions, what is the Bayes-optimal decision rule?**

if $p(x, y = A) > p(x, y = B)$
    return A
else
    return B

$$\begin{cases} p(x, y = A) = p(y = A)\, p(x \mid y = A) = p_A \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu_A)^T(x-\mu_A)} \\ p(x, y = B) = p(y = B)\, p(x \mid y = B) = p_B \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu_B)^T(x-\mu_B)} \end{cases}$$

- **Demonstrate that this implies a linear decision boundary between the two classes.**

The expression is $if\ p(x, y = A) > p(x, y = B)$, so the boundary is $p(x, y = A) = p(x, y = B)$

$$\Rightarrow p_A \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu_A)^T(x-\mu_A)} = p_B \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu_B)^T(x-\mu_B)}$$

$$\Rightarrow e^{-\frac{1}{2\sigma^2}(x-\mu_A)^T(x-\mu_A)} = \frac{p_B}{p_A} e^{-\frac{1}{2\sigma^2}(x-\mu_B)^T(x-\mu_B)}$$

$$\Rightarrow -\frac{1}{2\sigma^2}(x-\mu_A)^T(x-\mu_A) = \ln\frac{p_B}{p_A} - \frac{1}{2\sigma^2}(x-\mu_B)^T(x-\mu_B)$$

$$\Rightarrow (x-\mu_A)^T(x-\mu_A) - (x-\mu_B)^T(x-\mu_B) = -2\sigma^2 \ln\frac{p_B}{p_A}$$

Which is a linear function since it can always reduce to a 2-variable (x_1, x_2) function as next question showed.

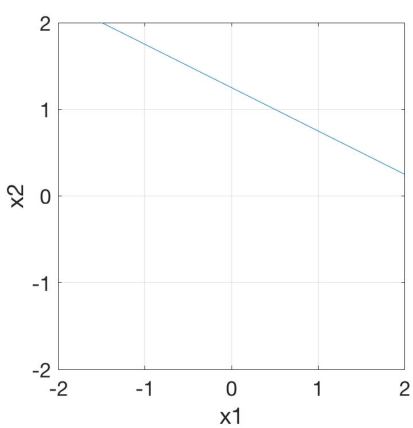
- **Draw the decision boundary for the four 2d cases below.**

$(a)\ p_A = p_B = 0.5, \mu_A = [0\ 0]^T, \mu_B = [1\ 2]^T, \sigma = 1$

$$\left(\begin{bmatrix}x_1\\x_2\end{bmatrix} - \begin{bmatrix}0\\0\end{bmatrix}\right)^T \left(\begin{bmatrix}x_1\\x_2\end{bmatrix} - \begin{bmatrix}0\\0\end{bmatrix}\right) - \left(\begin{bmatrix}x_1\\x_2\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix}\right)^T \left(\begin{bmatrix}x_1\\x_2\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix}\right) = 0$$

$$\Rightarrow ({x_1}^2 + {x_2}^2) - ((x_1 - 1)^2 + (x_2 - 2)^2) = 0$$

$$\Rightarrow 2x_1 + 4x_2 = 5$$

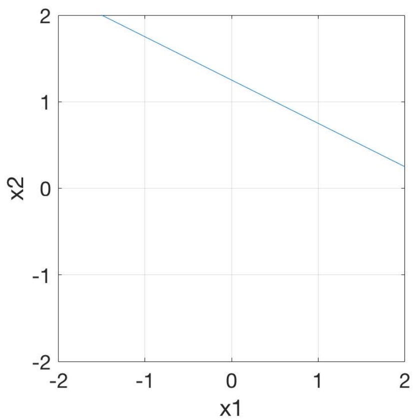


$(b)\ p_A = p_B = 0.5, \mu_A = [0\ 0]^T, \mu_B = [1\ 2]^T, \sigma = 3$

$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = 0$$

$\Rightarrow ({x_1}^2 + {x_2}^2) - ((x_1 - 1)^2 + (x_2 - 2)^2) = 0$

$\Rightarrow 2x_1 + 4x_2 = 5$

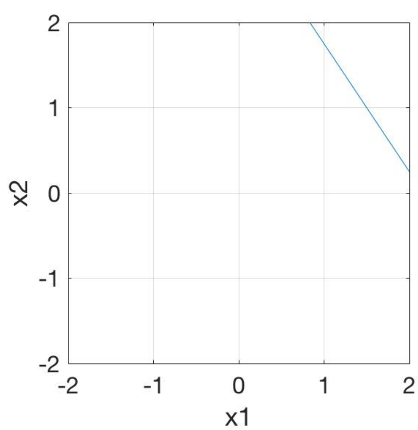


$(c)\ p_A = p_B = 0.5, \mu_A = [0\ 0]^T, \mu_B = [3\ 2]^T, \sigma = 1$

$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix}\right) = 0$$

$\Rightarrow ({x_1}^2 + {x_2}^2) - ((x_1 - 3)^2 + (x_2 - 2)^2) = 0$

$\Rightarrow 6x_1 + 4x_2 = 13$



$(d)\ p_A = 0.25, p_B = 0.75, \mu_A = [0\ 0]^T, \mu_B = [1\ 2]^T, \sigma = 1$

$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) - \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = -2 \ln 3$$

$$\Rightarrow ({x_1}^2 + {x_2}^2) - ((x_1 - 1)^2 + (x_2 - 2)^2) = -2 \ln 3$$

$$\Rightarrow 2x_1 + 4x_2 = 5 - 2 \ln 3$$