Decision boundary for normal distributions

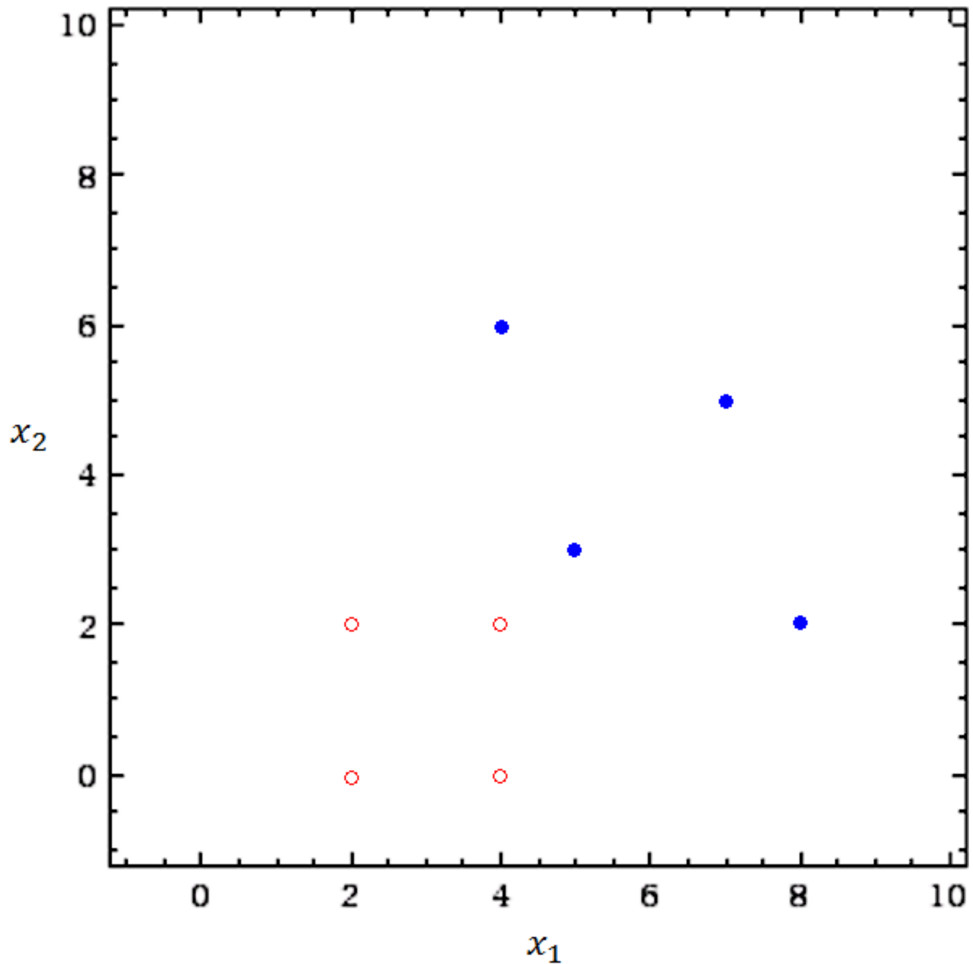
$S^{(1)} = \{(2,0),(4,0),(2,2),(4,2)\}$
$S^{(2)} = \{(4,6),(5,3),(7,5),(8,2)\}$

These two sets are feature vectors that originate from two bivariate normal distributions.

1. Estimate the corresponding covariance matrixes using maximum likelihood estimation
2. Find the analytical form of the optimal decision boundary between the two classes

Maximum likelihood estimation of the distribution parameters

$$S^{(1)} = \{(2,0),(4,0),(2,2),(4,2)\}$$
$$S^{(2)} = \{(4,6),(5,3),(7,5),(8,2)\}$$



$$S^{(1)} = \{(2,0), (4,0), (2,2), (4,2)\}$$
$$S^{(2)} = \{(4,6), (5,3), (7,5), (8,2)\}$$

Maximum likelihood estimation of the mean:

$$\mu^{(1)} = \left(\frac{1}{n}\sum x_1, \frac{1}{n}\sum x_2\right)$$

$$\mu^{(1)} = \left(\frac{1}{4}(2+4+2+4), \frac{1}{4}(0+0+2+2)\right)$$

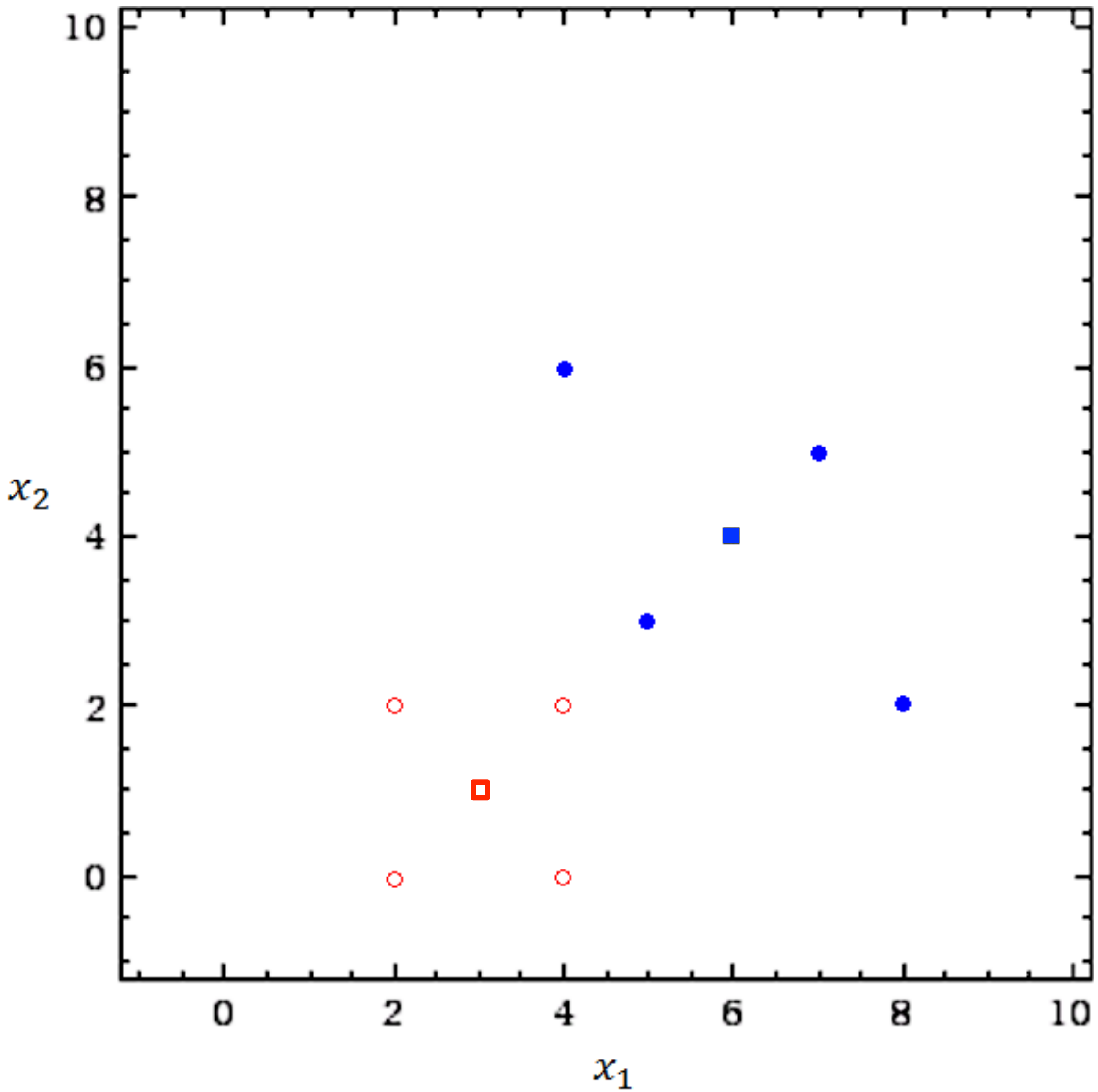
$$\mu^{(1)} = (3,1) \Rightarrow \mu^{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\mu^{(2)} = \left(\frac{1}{n}\sum x_1, \frac{1}{n}\sum x_2\right)$$

$$\mu^{(2)} = \left(\frac{1}{4}(4+5+7+8), \frac{1}{4}(6+3+5+2)\right)$$

$$\mu^{(2)} = (6,4) \Rightarrow \mu^{(2)} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$\mu^{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad \mu^{(2)} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$



Maximum likelihood estimation of the covariance matrices:

$$\Sigma = \frac{1}{n}\begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{2,2} \end{bmatrix}$$

$$\Sigma = \frac{1}{n}\begin{bmatrix} \sum_{i=0}^{n}(x_{1_i} - \mu_1)(x_{1_i} - \mu_1) & \sum_{i=0}^{n}(x_{1_i} - \mu_1)(x_{2_i} - \mu_2) \\ \sum_{i=0}^{n}(x_{2_i} - \mu_2)(x_{1_i} - \mu_1) & \sum_{i=0}^{n}(x_{2_i} - \mu_2)(x_{2_i} - \mu_2) \end{bmatrix}$$

This is a biased estimator (division by n). For an unbiased estimator, we need to divide by $\mathrm{n} - 1$.

Estimation of the covariance matrix Σ_1 of the $S^{(1)}$

$$\mu^{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

In the first case: ${\mu_1}^{(1)} = 3, {\mu_2}^{(1)} = 1$

$$\Sigma_1 = \frac{1}{4}\begin{bmatrix} \sum_{i=0}^{4}\left(x_{1_i}-3\right)^2 & \sum_{i=0}^{4}\left(x_{1_i}-3\right)\left(x_{2_i}-1\right) \\ \sum_{i=0}^{4}\left(x_{2_i}-1\right)\left(x_{1_i}-3\right) & \sum_{i=0}^{4}\left(x_{2_i}-1\right)^2 \end{bmatrix}$$

$$\begin{aligned} \sigma_{1,1} &= (2-3)^2 + (4-3)^2 + (2-3)^2 + (4-3)^2 \\ &= (-1)^2 + 1^2 + (-1)^2 + 1^2 \\ &= 4 \end{aligned}$$

$\sigma_{1,2}$ of set
$S_1 = \{(2,0), (4,0), (2,2), (4,2)\}$

$$\Sigma_1 = \frac{1}{4}\begin{bmatrix} (2-3)^2 + (4-3)^2 + (2-3)^2 + (4-3)^2 & \boxed{(2-3)(0-1)} + (4-3)(0-1) + (2-3)(2-1) + (4-3)(2-1) \\ (0-1)(2-3) + (0-1)(4-3) + (2-1)(2-3) + (2-1)(4-3) & (0-1)^2 + (0-1)^2 + (2-1)^2 + (2-1)^2 \end{bmatrix}$$

$$\Sigma_1 = \frac{1}{4}\begin{bmatrix} (-1)^2 + 1^2 + (-1)^2 + 1^2 & (-1\cdot-1) + (1\cdot-1) + (-1\cdot 1) + (1\cdot 1) \\ (-1\cdot-1) + (1\cdot-1) + (-1\cdot 1) + (1\cdot 1) & (-1)^2 + (-1)^2 + 1^2 + 1^2 \end{bmatrix}$$

$$\Sigma_1 = \frac{1}{4}\begin{bmatrix} 1+1+1+1 & 1-1-1+1 \\ 1-1-1+1 & 1+1+1+1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Maximum likelihood estimation of Σ_2

Remember
$S_2 = \{(4,6), (5,3), (7,5), (8,2)\}$

$$\Sigma_2 = \frac{1}{4}\begin{bmatrix} (4-6)^2+(5-6)^2+(7-6)^2+(8-6)^2 & (4-6)(6-4)+(5-6)(3-4)+(7-6)(5-4)+(8-6)(2-4) \\ (6-4)(4-6)+(3-4)(5-6)+(5-4)(7-6)+(2-4)(8-6) & (6-4)^2+(3-4)^2+(5-4)^2+(2-4)^2 \end{bmatrix}$$

$$\Sigma_2 = \frac{1}{4}\begin{bmatrix} (-2)^2+(-1)^2+1^2+2^2 & (-2\cdot 2)+(-1\cdot -1)+(1\cdot 1)+(2\cdot -2) \\ (2\cdot -2)+(-1\cdot -1)+(1\cdot 1)+(-2\cdot 2) & 2^2+(-1)^2+1^2+(-2)^2 \end{bmatrix}$$

$$\Sigma_2 = \frac{1}{4}\begin{bmatrix} 4+1+1+4 & -4+1+1-4 \\ -4+1+1-4 & 4+1+1+4 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 10 & -6 \\ -6 & 10 \end{bmatrix} = \begin{bmatrix} 2\frac{1}{2} & -1\frac{1}{2} \\ -1\frac{1}{2} & 2\frac{1}{2} \end{bmatrix}$$

Result of the maximum likelihood estimation

$$\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 2\frac{1}{2} & -1\frac{1}{2} \\ -1\frac{1}{2} & 2\frac{1}{2} \end{bmatrix}$$

Multivariate normal density

$$p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^t\Sigma^{-1}(x-\mu)}$$

- $x \in R^d$ is a d-dimensional vector ($d = 2$ in our example)
- μ is the mean (d-dimensional vector)
- Σ is the covariance matrix ($|\Sigma|$ is the determinant and Σ^{-1} is the inverse)

Common notation:

$$p(x) \sim N(\mu, \Sigma)$$

Reminder matrix determinant and inverse

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\det(A) = |A| = ad - bc$$

The inverse exists only iff $\det(A) \neq 0$

$$inv(A) = A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)}$$

- $x \in R^d, d = 2$
- $\mu^{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mu^{(2)} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$
- $\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow |\Sigma_1| = 1, \quad {\Sigma_1}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- $\Sigma_2 = \begin{bmatrix} 2\frac{1}{2} & -1\frac{1}{2} \\ -1\frac{1}{2} & 2\frac{1}{2} \end{bmatrix} \rightarrow |\Sigma_2| = 4, \quad {\Sigma_2}^{-1} = \begin{bmatrix} {}^5/_8 & {}^3/_8 \\ {}^3/_8 & {}^5/_8 \end{bmatrix}$

Probability density for the distribution modeling set $S^{(1)}$

$$p\left(x|N\left(\mu^{(1)},\Sigma_1\right)\right)=\frac{1}{(2\pi)^{2/2}1^{1/2}}e^{-\frac{1}{2}\left(x-\mu^{(1)}\right)^t\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\left(x-\mu^{(1)}\right)}$$

And for the second matrix:

$$p\left(x|N\left(\mu^{(2)},\Sigma_2\right)\right)=\frac{1}{(2\pi)^{2/2}4^{1/2}}e^{-\frac{1}{2}\left(x-\mu^{(2)}\right)^t\begin{bmatrix}5/8 & 3/8\\ 3/8 & 5/8\end{bmatrix}\left(x-\mu^{(2)}\right)}$$

Optimal decision boundary

Assuming equal prior probabilities:

$$P(\omega_1) = P(\omega_2) = 0.5$$

Solve the equation xxxxxxxxxxxxxxx $\cancel{p(\omega_1) = p(\omega_2)}$ P(x | ω_1) = P(x | ω_2)

$$\frac{1}{(2\pi)^{2/2}1^{1/2}}\,e^{-\frac{1}{2}\left(x-\mu^{(1)}\right)^t\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\left(x-\mu^{(1)}\right)} = \frac{1}{(2\pi)^{2/2}4^{1/2}}\,e^{-\frac{1}{2}\left(x-\mu^{(2)}\right)^t\begin{bmatrix}5/8 & 3/8\\ 3/8 & 5/8\end{bmatrix}\left(x-\mu^{(2)}\right)}$$

$$0{,}1592e^{-\frac{1}{2}\left(x-\mu^{(1)}\right)^t\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\left(x-\mu^{(1)}\right)} = 0.0796e^{-\frac{1}{2}\left(x-\mu^{(2)}\right)^t\begin{bmatrix}5/8 & 3/8\\ 3/8 & 5/8\end{bmatrix}\left(x-\mu^{(2)}\right)}$$

Take the ln of both sides:

$$\ln(0{,}1592) - \frac{1}{2}\left(x-\mu^{(1)}\right)^t\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\left(x-\mu^{(1)}\right) = \ln(0.0796) - \frac{1}{2}\left(x-\mu^{(2)}\right)^t\begin{bmatrix}5/8 & 3/8\\ 3/8 & 5/8\end{bmatrix}\left(x-\mu^{(2)}\right)$$

First we simplify the left-hand side of the equation:

$$\ln(0{,}1592) - \frac{1}{2}(x-\mu)^t \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} (x-\mu)$$

t = transpose

$$\mu^{(1)} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$\ln(0{,}1592) - \frac{1}{2}(x-\mu)^t \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} (x-\mu)$$

$$= \ln(0{,}1592) - \frac{1}{2}\begin{bmatrix} x_1 - 3 & x_2 - 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 - 3 \\ x_2 - 1 \end{bmatrix}$$

$$= \ln(0{,}1592) - \frac{1}{2}\begin{bmatrix} x_1 - 3 & x_2 - 1 \end{bmatrix} \begin{bmatrix} x_1 - 3 \\ x_2 - 1 \end{bmatrix}$$

$$= \ln(0{,}1592) - \frac{1}{2}((x_1 - 3)^2 + (x_2 - 1)^2)$$

$$= -\frac{1}{2}x_1^2 + 3x_1 - \frac{1}{2}x_2^2 + x_2 - 6{,}8376$$

Now the right-hand side of the equation:

$$\ln(0.0796) - \frac{1}{2}(x-\mu)^t \begin{bmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{bmatrix} (x-\mu)$$

$\mu^{(2)} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$

$$= \ln(0.0796) - \frac{1}{2}\begin{bmatrix} x_1 - 6 & x_2 - 4 \end{bmatrix} \begin{bmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{bmatrix} \begin{bmatrix} x_1 - 6 \\ x_2 - 4 \end{bmatrix}$$

$$= \ln(0.0796) - \frac{1}{2}\begin{bmatrix} \frac{5}{8}(x_1 - 6) + \frac{3}{8}(x_2 - 4) & \frac{3}{8}(x_1 - 6) + \frac{5}{8}(x_2 - 4) \end{bmatrix} \begin{bmatrix} x_1 - 6 \\ x_2 - 4 \end{bmatrix}$$

$$= \ln(0.0796) - \frac{1}{2}\left(\left(\frac{5}{8}(x_1 - 6) + \frac{3}{8}(x_2 - 4)\right)(x_1 - 6) + \left(\frac{3}{8}(x_1 - 6) + \frac{5}{8}(x_2 - 4)\right)(x_2 - 4)\right)$$

$$= -\frac{5}{16}x_1^2 - \frac{3}{8}x_1 x_2 + 5\frac{1}{4}x_1 - \frac{5}{16}x_2^2 + 4\frac{3}{4}x_2 - 27.7807$$

And we solve the final equation:

$$-\frac{1}{2}x_1^2+3x_1-\frac{1}{2}x_2^2+x_2-6,8376=-\frac{5}{16}x_1^2-\frac{3}{8}x_1x_2+5\frac{1}{4}x_1-\frac{5}{16}x_2^2+4\frac{3}{4}x_2-27.7807$$

Result:

$$-\frac{3}{16}x_1^2+\frac{3}{8}x_1x_2-2\frac{1}{4}x_1-\frac{3}{16}x_2^2-3\frac{3}{4}x_2+20.9431=0$$