Decision Boundary: Multivariate Unimodal Gaussian(Normal) Distribution

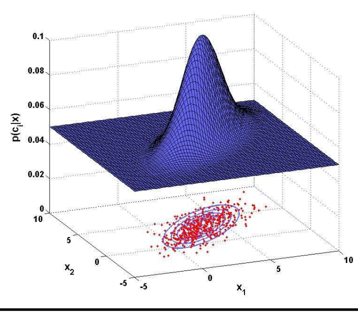
Multivariate Unimodal Gaussian Distribution

- Data in d-dimensional space

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$= \frac{1}{(2\pi)^{d/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} \underbrace{(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}_{\textbf{Mahalanobis distance}} \right)$$

  - $\boldsymbol{\mu}$ is the mean vector
  - $\boldsymbol{\Sigma}$ is the covariance matrix
- Bivariate Gaussian distribution: d=2



$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} E[x_1] \\ E[x_2] \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} E\left[(x_1-\mu_1)^2\right] & E\left[(x_1-\mu_1)(x_2-\mu_2)\right] \\ E\left[(x_2-\mu_2)(x_1-\mu_1)\right] & E\left[(x_2-\mu_2)^2\right] \end{bmatrix}$$

Multivariate Unimodal Gaussian Distribution

- Training data for class i

  $\mathcal{D}_i = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{Ni}\}$
- Likelihood or Class conditional density:

$$p(\mathbf{x}|C_i) \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$$

$$p(\mathbf{x}|C_i) = \frac{1}{(2\pi)^{\frac{d}{2}}|\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right)$$

  - $\boldsymbol{\mu}_i$ is the mean vector of class i
  - $\boldsymbol{\Sigma}_i$ is the covariance matrix of class i

Discriminant Function in Multivariate Unimodal Gaussian Distribution

- Discriminant function:

$$g_i(\mathbf{x}) = ln\big(p(\mathbf{x}|C_i)\big) + ln\big(p(C_i)\big)$$

$$p(\mathbf{x}|C_i) = \frac{1}{(2\pi)^{\frac{d}{2}}|\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right)$$

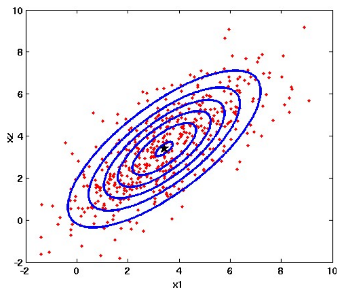
$$ln\big(p(\mathbf{x}|C_i)\big) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}ln(2\pi) + \frac{d}{2}ln(|\boldsymbol{\Sigma}_i|)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}ln(2\pi) + \frac{d}{2}ln(|\boldsymbol{\Sigma}_i|) + ln\big(p(C_i)\big)$$

- The discriminant function and its shape depends on the covariance matrix

Bivariate Gaussian Distribution: Illustaration - 01

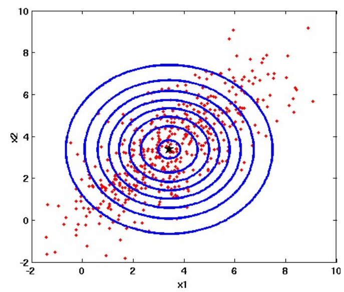
- Full covariance matrix:



$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 3.92 & 2.87 \\ 2.87 & 3.85 \end{bmatrix}$$

- Diagonal covariance matrix:

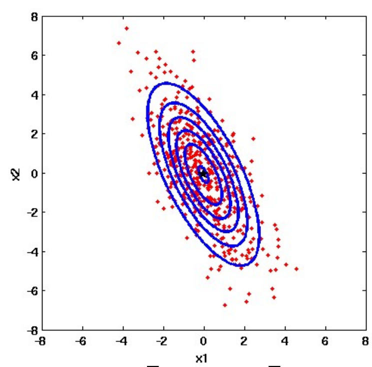


$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 3.92 & 0 \\ 0 & 3.85 \end{bmatrix}$$

Bivariate Gaussian Distribution: Illustaration - 02

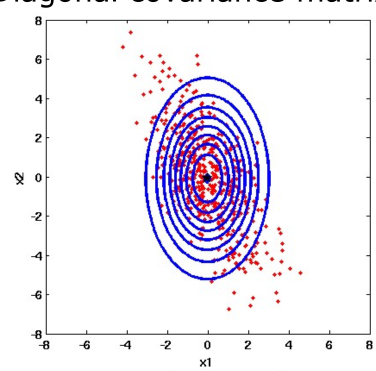
- Full covariance matrix:



$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 2.09 & -2.37 \\ -2.37 & 5.96 \end{bmatrix}$$

- Diagonal covariance matrix:

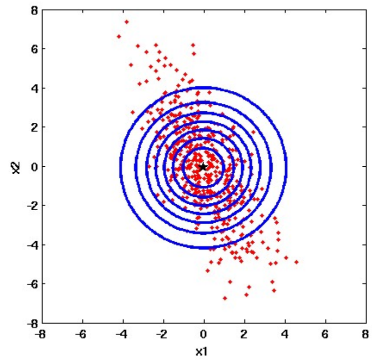


$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 2.09 & 0 \\ 0 & 5.96 \end{bmatrix}$$

Bivariate Gaussian Distribution: Illustration - 03

- Diagonal covariance matrix with equal variance:

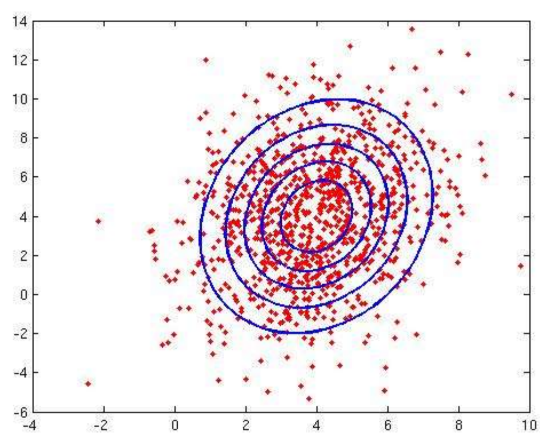


$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

- **Note**: variance is the average of two diagonal variances

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4.03 & 0 \\ 0 & 4.03 \end{bmatrix}$$

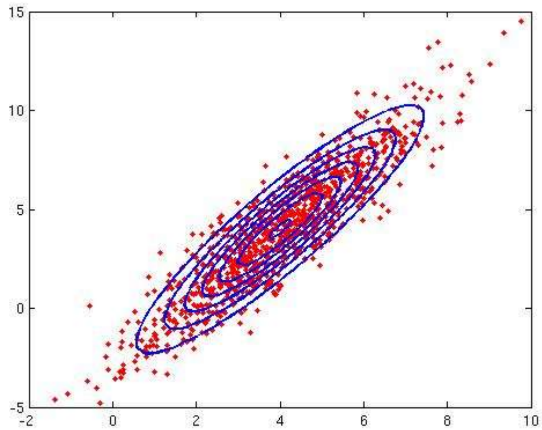
Bivariate Gaussian Distribution: Illustration - 04



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 3 & 1 \\ 1 & 10 \end{bmatrix}$$

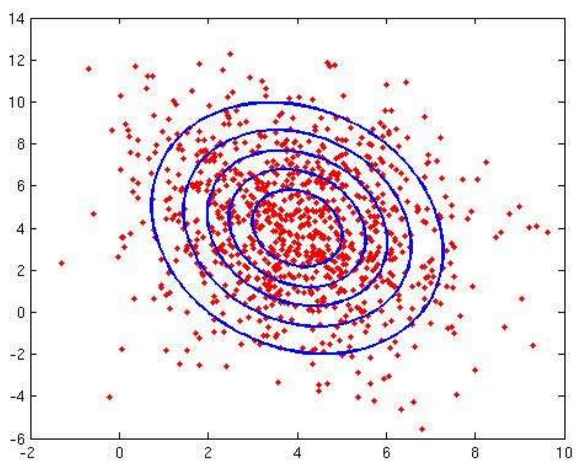
Bivariate Gaussian Distribution: Illustration - 05



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 3 & 5 \\ 5 & 10 \end{bmatrix}$$

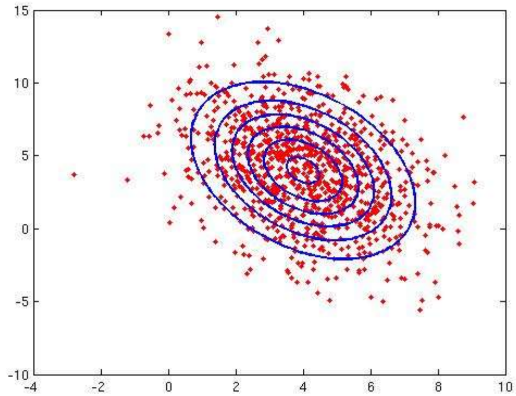
Bivariate Gaussian Distribution: Illustration - 06



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 3 & -1 \\ -1 & 10 \end{bmatrix}$$

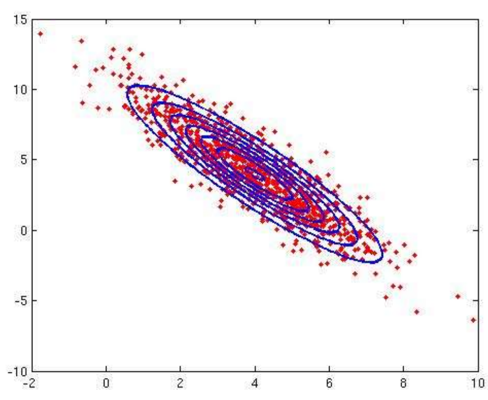
Bivariate Gaussian Distribution: Illustration - 07



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 3 & -2 \\ -2 & 10 \end{bmatrix}$$

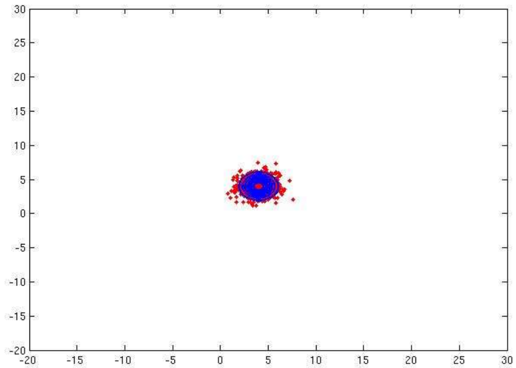
Bivariate Gaussian Distribution: Illustration - 08



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 3 & -5 \\ -5 & 10 \end{bmatrix}$$

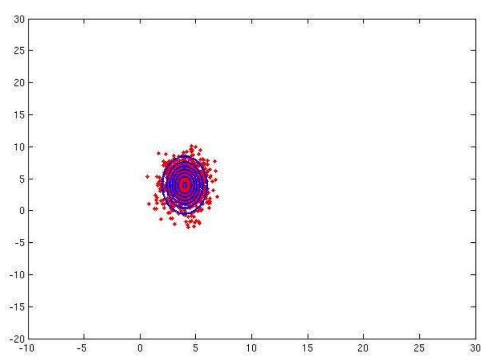
Bivariate Gaussian Distribution: Illustration - 09



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

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

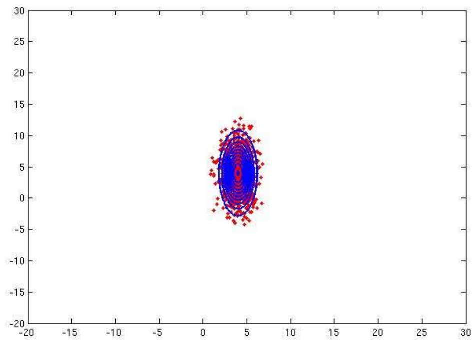
Bivariate Gaussian Distribution: Illustration - 10



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

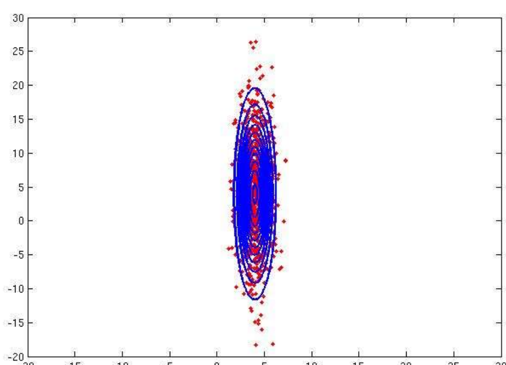
Bivariate Gaussian Distribution: Illustration - 11



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}$$

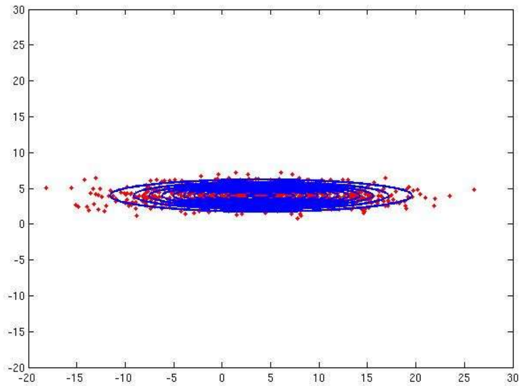
Bivariate Gaussian Distribution: Illustration - 12



$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 50 \end{bmatrix}$$

Bivariate Gaussian Distribution: Illustration - 13

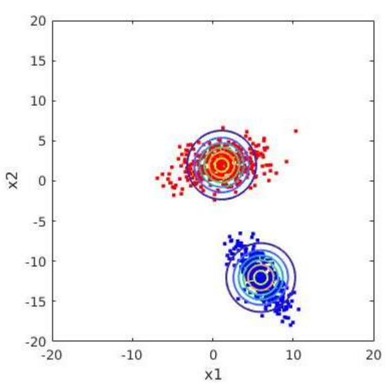


$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 50 & 0 \\ 0 & 1 \end{bmatrix}$$

Case 1: $\Sigma_i = \sigma^2 \mathbf{I}$

- Occurs when the features are statistically independent and each feature has same variance
- Covariance matrix is diagonal and all variance is same
- Isotropic covariance matrix
- Covariance matrix for each class is same



$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4.03 & 0 \\ 0 & 4.03 \end{bmatrix}$$

Case 1: $\Sigma_i = \sigma^2 \mathbf{I}$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln(|\boldsymbol{\Sigma}|) + \ln(P(C_i))$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\frac{(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}(\mathbf{x}-\boldsymbol{\mu}_i)}{\sigma^2} - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln(|\boldsymbol{\Sigma}|) + \ln(P(C_i))$$

$$g_i(\mathbf{x}) = -\frac{1}{2\sigma^2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln(P(C_i))$$

Case 1: $\Sigma_i = \sigma^2 \mathbf{I}$

$$g_i(\mathbf{x}) = \frac{1}{\sigma^2}\boldsymbol{\mu}^{\mathrm{T}}_i\mathbf{x} - \frac{1}{2\sigma^2}\boldsymbol{\mu}^{\mathrm{T}}_i\boldsymbol{\mu}_i + \ln(P(C_i))$$

$$g_i(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}_i\mathbf{x} + w_{i0}$$

Where,

$$\mathbf{w}_i = \frac{1}{\sigma^2}\boldsymbol{\mu}_i \qquad w_{i0} = -\frac{1}{2\sigma^2}\boldsymbol{\mu}^{\mathrm{T}}_i\boldsymbol{\mu}_i + \ln(P(C_i))$$

Case 1: $\Sigma_i = \sigma^2 \mathbf{I}$

Decision boundary:

$$g(\mathbf{x}) = g_1(\mathbf{x}) - g_2(\mathbf{x}) = 0$$

$$\left(\mathbf{w}_1^{\mathrm{T}} - \mathbf{w}_2^{\mathrm{T}}\right)\mathbf{x} + \left(w_{10} - w_{20}\right) = 0$$

Where,

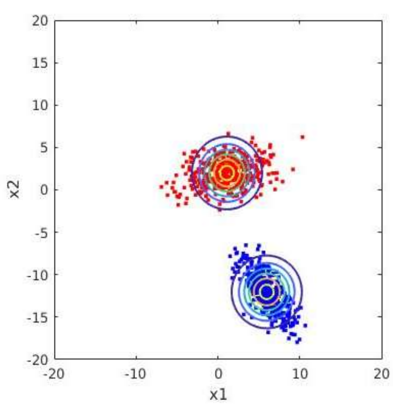
$$\mathbf{w}^{\mathrm{T}} \quad \mathbf{x} + \quad w_0 \quad = 0$$

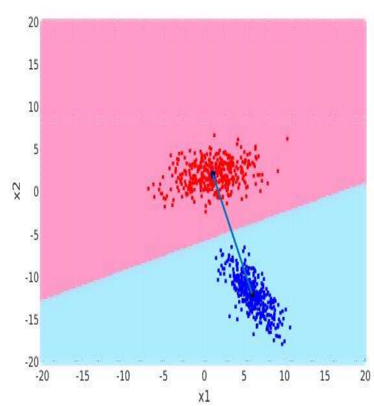
$$\mathbf{w} = \frac{1}{\sigma^2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \qquad w_0 = -\frac{1}{2\sigma^2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathrm{T}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \ln\left(\frac{P(C_1)}{P(C_2)}\right)$$

$$g(\mathbf{x}) = \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathrm{T}}}{\sigma^2}\left[\mathbf{x} - \underbrace{\frac{(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)}{2} + \frac{\sigma^2(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathrm{T}}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}\ln\left(\frac{P(C_1)}{P(C_2)}\right)}_{\mathbf{x}_0}\right] = 0$$

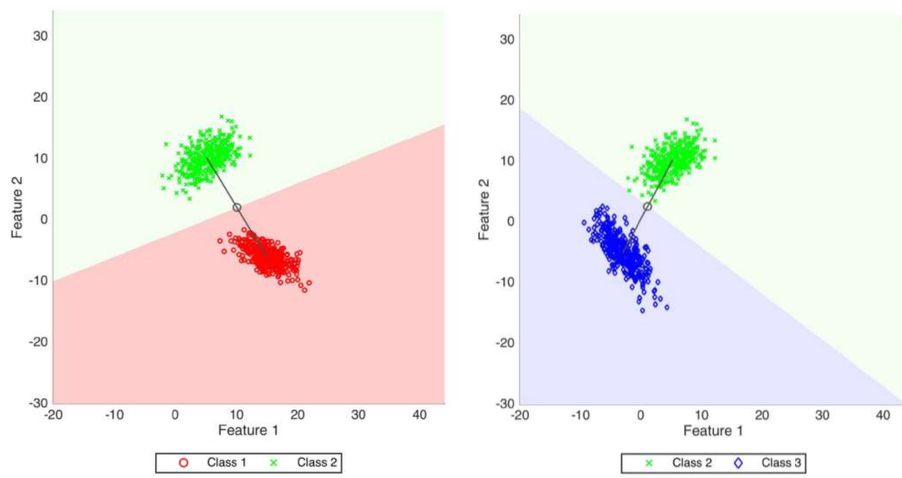
$$g(\mathbf{x}) = \mathbf{w}^{\mathrm{T}} \left[\mathbf{x} - \mathbf{x}_0\right] = 0$$

Case 1: $\Sigma_i = \sigma^2 \mathbf{I}$

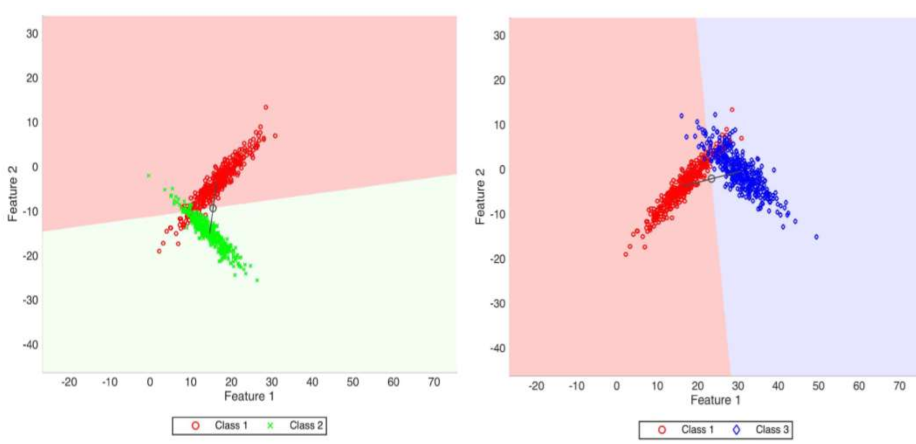




Case 1: $\Sigma_i = \sigma^2 \mathrm{I}$

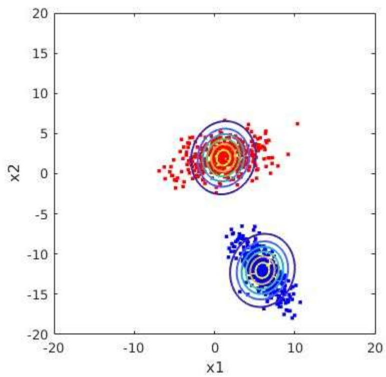


Case 1: $\Sigma_i = \sigma^2 \mathrm{I}$



Case 2: $\Sigma_i = \Sigma$

- Covariance matrix is same for each class
- Covariance matrix is arbitrary (i.e. full covariance matrix) or diagonal
- All the variance is not same



$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 3.92 & 2.87 \\ 2.87 & 3.85 \end{bmatrix}$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln(|\mathbf{\Sigma}|) + \ln(P(C_i))$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x}\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x}\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[- \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boxed{\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x}} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[- \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[- 2\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[-2\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\boxed{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i} - \frac{1}{2}\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \ln\left(P(C_i)\right)$$

Case 2: $\Sigma_i = \Sigma$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[-2\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right] + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = \left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right)^{\mathrm{T}}\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = \mathbf{w}_i^{\mathrm{T}}\mathbf{x} + w_{i0}$$ Linear discriminant function

Case 1: $\Sigma_i = \Sigma$

Decision boundary:

$$g(\mathbf{x}) = g_1(\mathbf{x}) - g_2(\mathbf{x}) = 0$$

$$\left(\mathbf{w}_1^{\mathsf{T}} - \mathbf{w}_2^{\mathsf{T}}\right)\mathbf{x} + \left(w_{10} - w_{20}\right) = 0$$

Where,

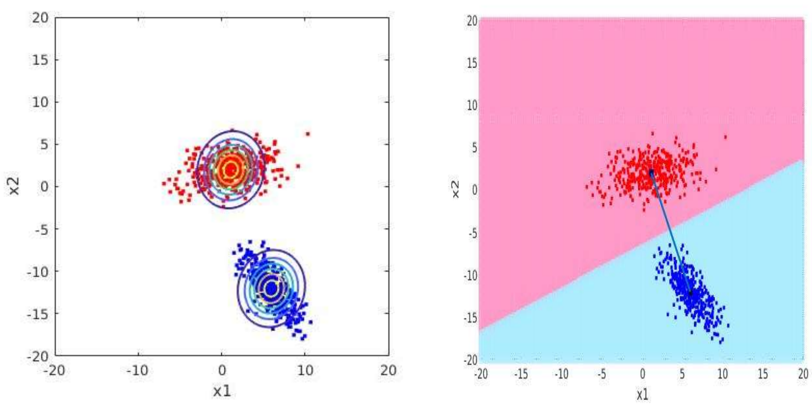
$$\mathbf{w}^{\mathsf{T}} \quad \mathbf{x} \; + \quad w_0 \quad = 0$$

$$\mathbf{w} = \mathbf{\Sigma}^{-1}\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right) \qquad w_0 = -\frac{1}{2}\left(\boldsymbol{\mu}_1^{\mathsf{T}}\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^{\mathsf{T}}\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2\right) + \ln\left(\frac{P(C_1)}{P(C_2)}\right)$$

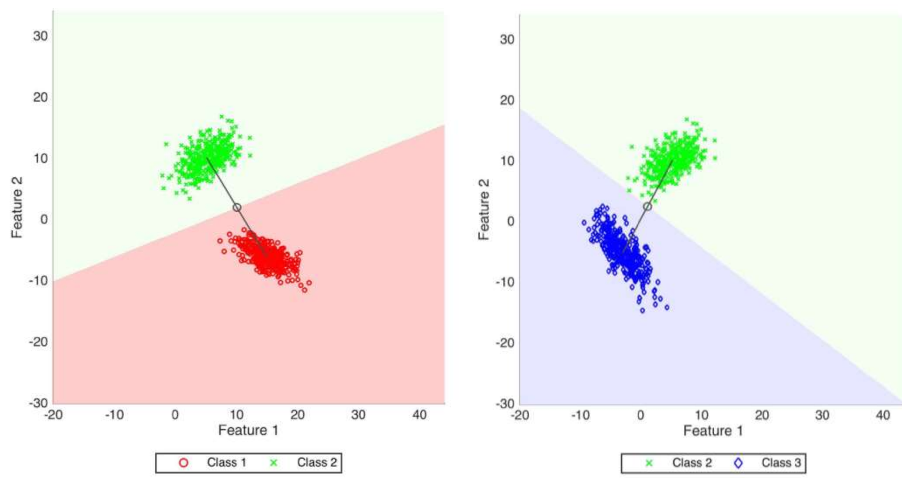
$$g(\mathbf{x}) = \mathbf{\Sigma}^{-1}\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right)^{\mathsf{T}}\left[\mathbf{x} - \underbrace{\frac{\left(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2\right)}{2} + \frac{\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right)}{\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right)^{\mathsf{T}}\mathbf{\Sigma}^{-1}\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right)}\ln\left(\frac{P(C_1)}{P(C_2)}\right)}\right] = 0$$

$$g(\mathbf{x}) = \quad \mathbf{w}^{\mathsf{T}} \quad \left[\mathbf{x} - \quad \mathbf{x}_0 \quad\right] = 0$$

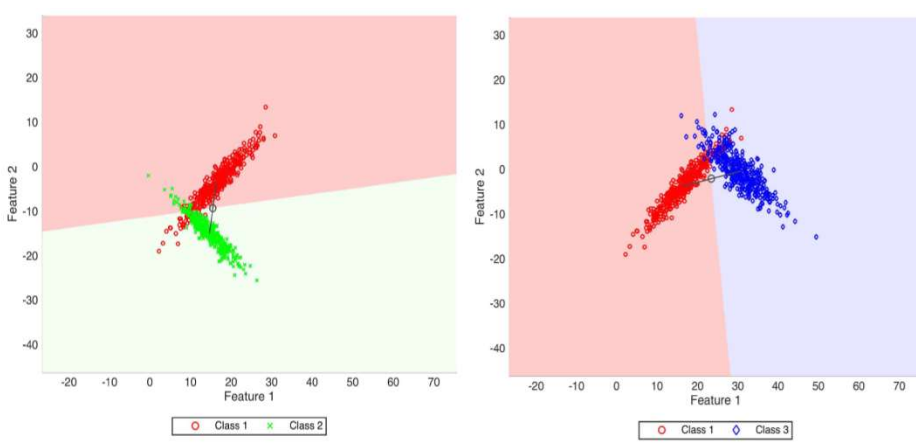
Case 2: $\Sigma_i = \Sigma$



Case 2: $\Sigma_i = \Sigma$

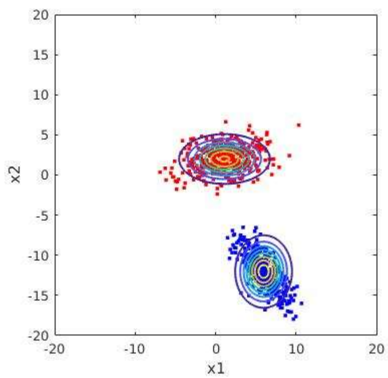


Case 2: $\Sigma_i = \Sigma$

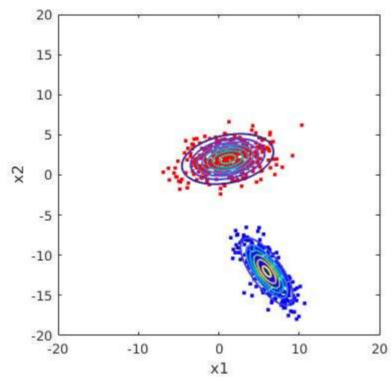


Case 3: Σ_i is Arbitrary

- Covariance matrix is different for each class
- Covariance matrix is arbitrary (i.e. full covariance matrix) or diagonal
- All the variance is not same



Diagonal Covariance Matrix



Full Covariance Matrix

Case 3: Σ_i is Arbitrary

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln(|\boldsymbol{\Sigma}_i|) + \ln(P(C_i))$$

Case 3: Σ_i is Arbitrary

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{1}{2}\ln\left(\left|\boldsymbol{\Sigma}_i\right|\right) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x}\right] - \frac{1}{2}\ln\left(\left|\boldsymbol{\Sigma}_i\right|\right) + \ln\left(P(C_i)\right)$$

Case 3: Σ_i is Arbitrary

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{1}{2}\ln\left(\left|\boldsymbol{\Sigma}_i\right|\right) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\right] - \frac{1}{2}\ln\left(\left|\boldsymbol{\Sigma}_i\right|\right) + \ln\left(P(C_i)\right)$$

Case 3: Σ_i is Arbitrary

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_i|\right) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x}\right] - \frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_i|\right) + \ln\left(P(C_i)\right)$$

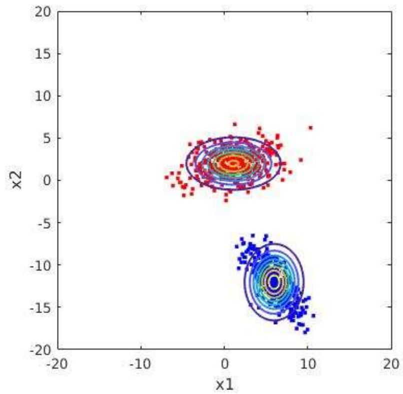
Case 3: Σ_i is Arbitrary

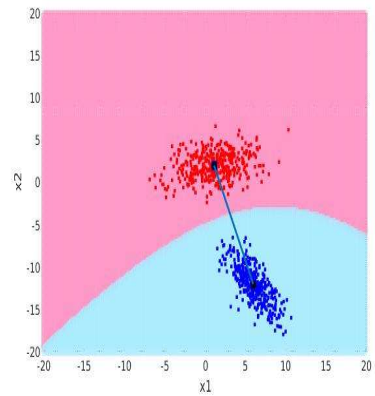
$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_i|\right) + \ln\left(P(C_i)\right)$$

$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\right] - \frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_i|\right) + \ln\left(P(C_i)\right)$$

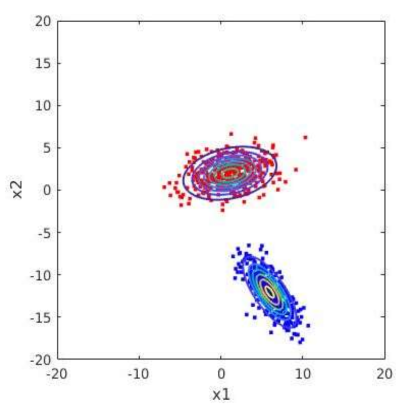
$$g_i(\mathbf{x}) = -\frac{1}{2}\left[\mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} - 2\boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{x} + \boldsymbol{\mu}_i^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\right] - \frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_i|\right) + \ln\left(P(C_i)\right)$$

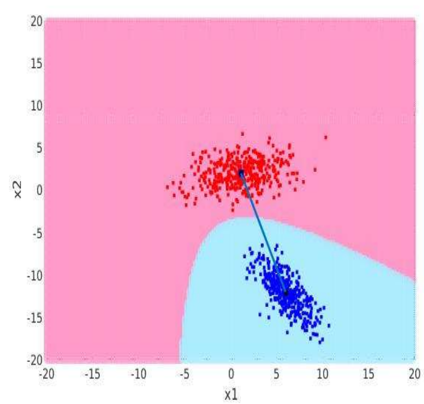
Case 3: Σ_i is Arbitrary and Diagonal



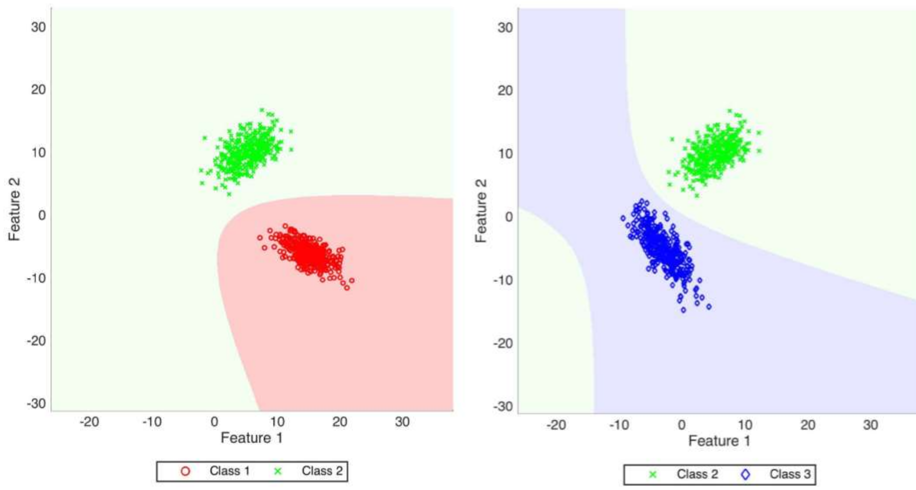


Case 3: Σ_i is Arbitrary and Full





Case 3: Σ_i is Arbitrary



Case 3: Σ_i is Arbitrary