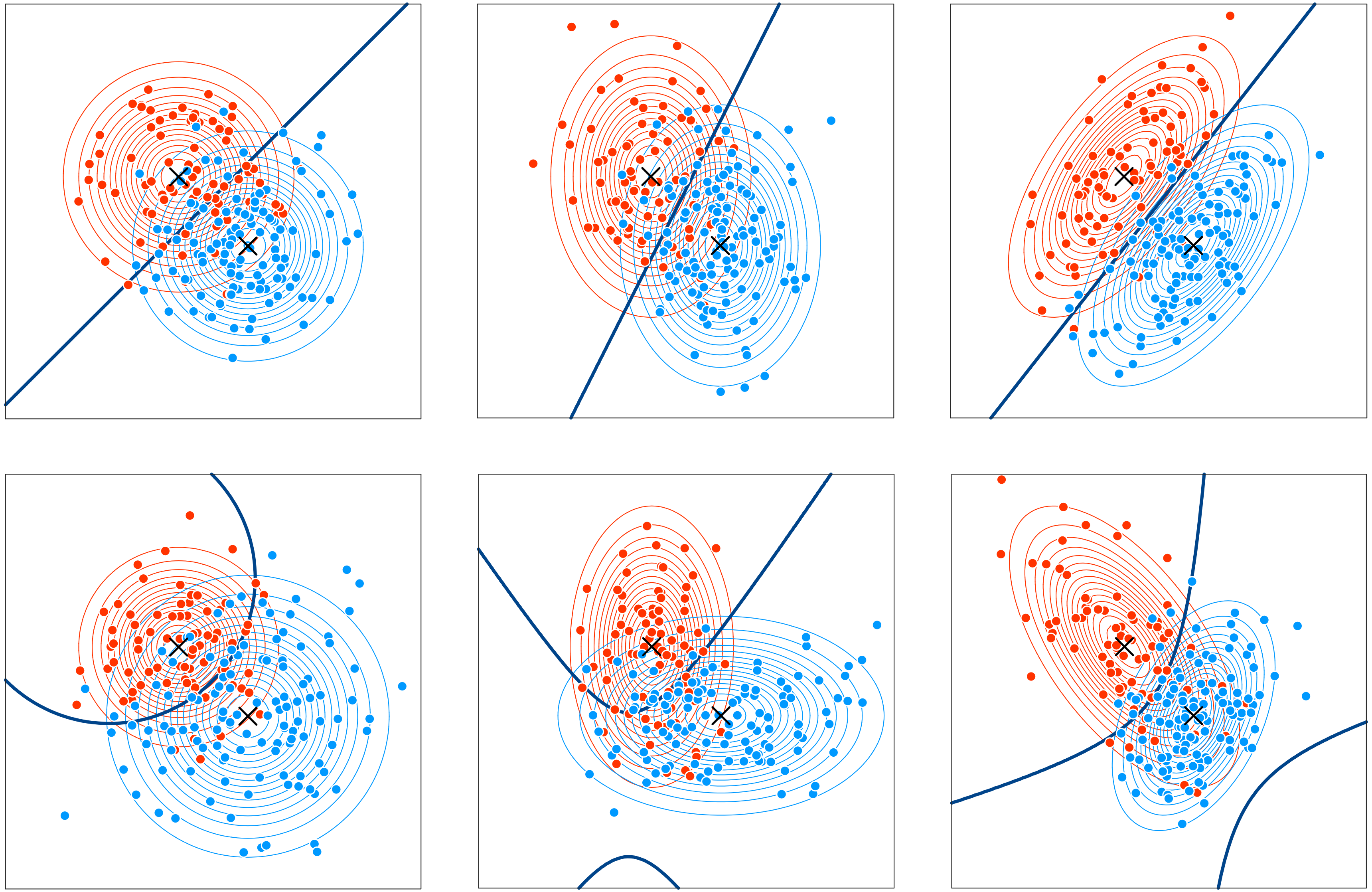
高斯判别分析原理



知乎

高斯判别分析优化目标

$$\hat{y} = \arg\min_{C_m} \sum_{k=1}^{K} p\left(C_k \middle| \mathbf{x}\right) \cdot c\left(C_m \middle| C_k\right)$$

$$c\left(C_m \middle| C_k\right) = \begin{cases} 1 & m \neq k \\ 0 & m = k \end{cases}$$

后验概率

$$p(C_k \mid \mathbf{x}) = \frac{p(\mathbf{x} \cap C_k)}{p(\mathbf{x})} = \frac{p(\mathbf{x} \mid C_k)\mathrm{P}(C_k)}{p(\mathbf{x})}$$

$$p(C_k \mid \mathbf{x}) \propto p(C_k \cap \mathbf{x})$$

bilibili

知乎

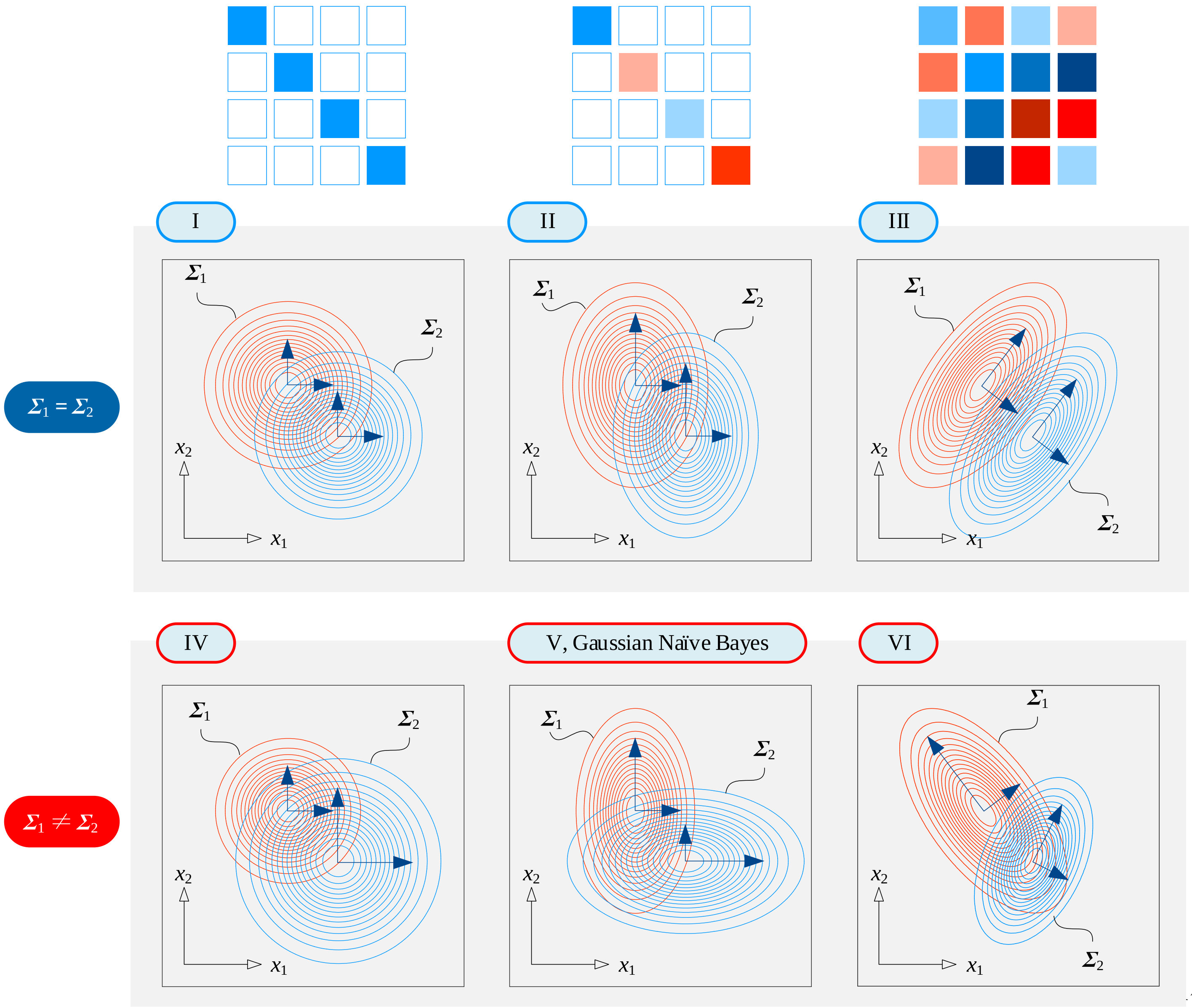
似然概率服从多元高斯分布

$$p\left(\boldsymbol{x} \mid C_k\right) = \frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_k\right)^{\mathrm{T}} \boldsymbol{\Sigma}_k^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_k\right)\right)}{\sqrt{(2\pi)^D \left|\boldsymbol{\Sigma}_k\right|}}$$

六类协方差矩阵

<table>
<tr><th></th><th>Σ_k</th><th>Σ_k 方差 (对角线元素)</th><th>Σ_k 特点</th><th>PDF 等高线</th><th>决策边界</th></tr>
<tr><td>第一类</td><td rowspan="3">相同</td><td>相同</td><td rowspan="2">对角阵
(特征条件独立)</td><td>正圆，形状相同</td><td rowspan="3">直线</td></tr>
<tr><td>第二类</td><td rowspan="2">不限制</td><td>正椭圆，形状相同</td></tr>
<tr><td>第三类</td><td>非对角阵</td><td>任意椭圆，形状相同</td></tr>
<tr><td>第四类</td><td rowspan="3">不相同</td><td>相同</td><td rowspan="2">对角阵
(特征条件独立)</td><td>正圆</td><td>正圆</td></tr>
<tr><td>第五类</td><td rowspan="2">不限制</td><td>正椭圆</td><td>正圆锥曲线</td></tr>
<tr><td>第六类</td><td>非对角阵</td><td>任意椭圆</td><td>圆锥曲线</td></tr>
</table>

六大类判别分析高斯分布椭圆形状



决策边界解析解

$$g_k(\boldsymbol{x}) = \ln\left(p\left(\boldsymbol{x} \cap C_k\right)\right) = \ln\left(p\left(\boldsymbol{x} \mid C_k\right)\mathrm{P}\left(C_k\right)\right)$$

$$= \ln\left(\frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_k\right)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_k\right)\right)}{\sqrt{\left(2\pi\right)^{D}\left|\boldsymbol{\Sigma}_k\right|}}\mathrm{P}\left(C_k\right)\right)$$

$$= -\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}_k\right)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_k\right) - \frac{D}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left|\boldsymbol{\Sigma}_k\right| + \ln\mathrm{P}\left(C_k\right)$$

二分类

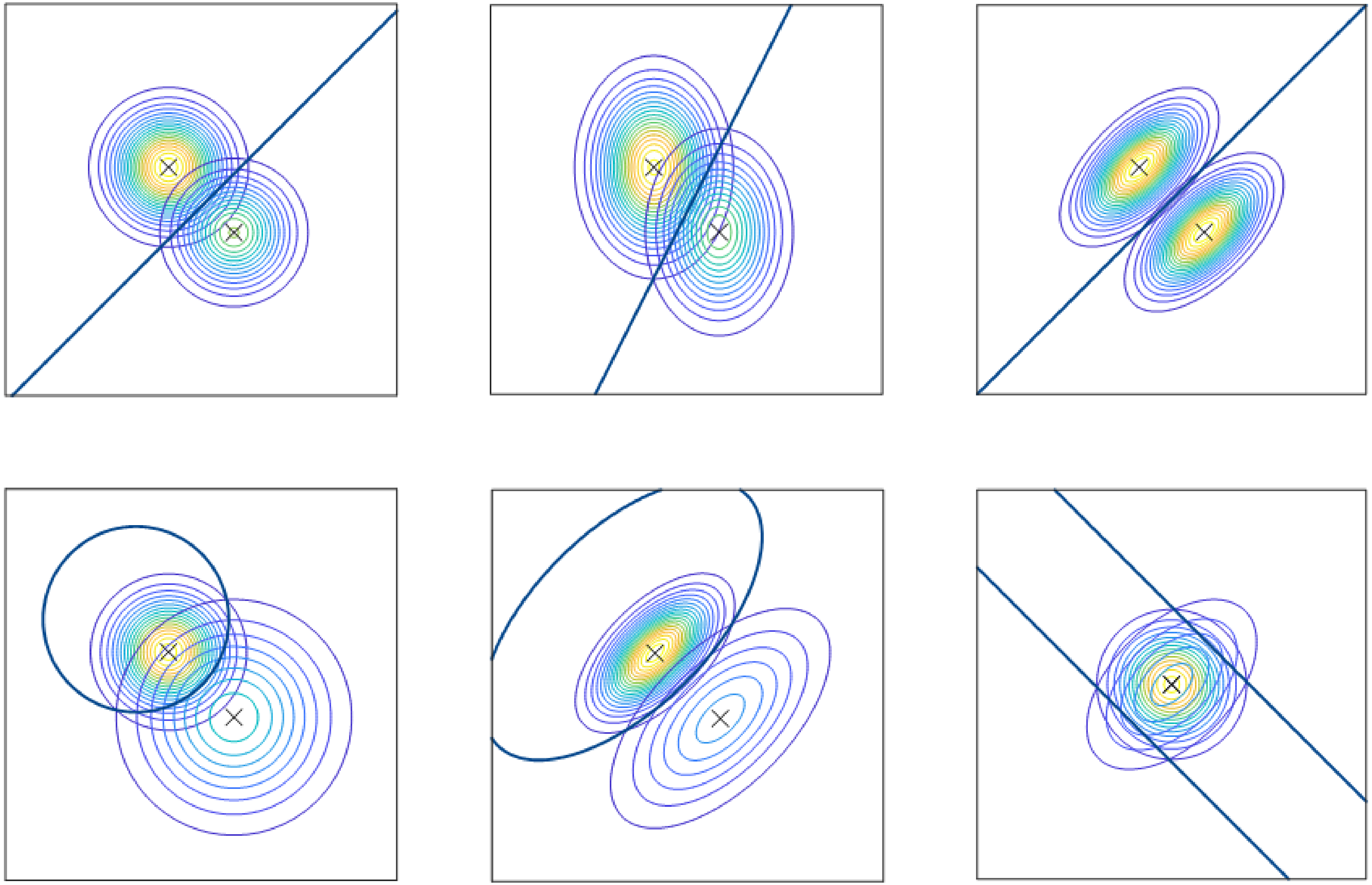
$$\begin{cases} g_1(\boldsymbol{x}) = -\dfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_1)^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_1) - \dfrac{D}{2}\ln(2\pi) - \dfrac{1}{2}\ln|\boldsymbol{\Sigma}_1| + \ln \mathrm{P}(C_1) \\ g_2(\boldsymbol{x}) = -\dfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_2)^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_2) - \dfrac{D}{2}\ln(2\pi) - \dfrac{1}{2}\ln|\boldsymbol{\Sigma}_2| + \ln \mathrm{P}(C_2) \end{cases}$$

二分类

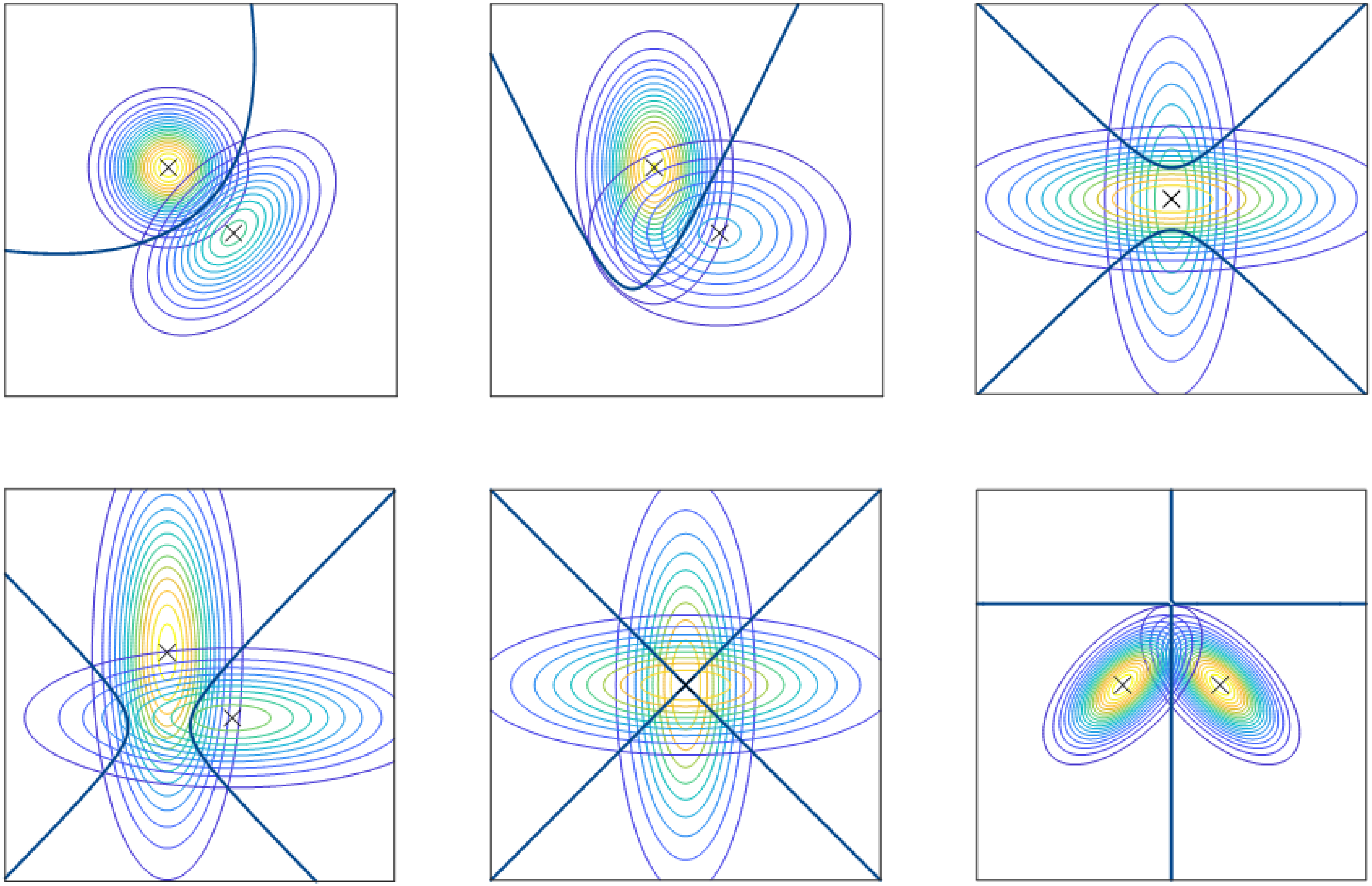
$$g_1(\boldsymbol{x}) = g_2(\boldsymbol{x})$$

$$\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_1)^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_1)-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_2)^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_2)=\ln \mathrm{P}(C_1)-\ln \mathrm{P}(C_2)+\left(\frac{1}{2}\ln|\boldsymbol{\Sigma}_2|-\frac{1}{2}\ln|\boldsymbol{\Sigma}_1|\right)$$

判别分析常见决策边界



判别分析常见决策边界



第一类高斯判别分析

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \sigma^2 \boldsymbol{I}$$

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

$P(C_1) = P(C_2)$

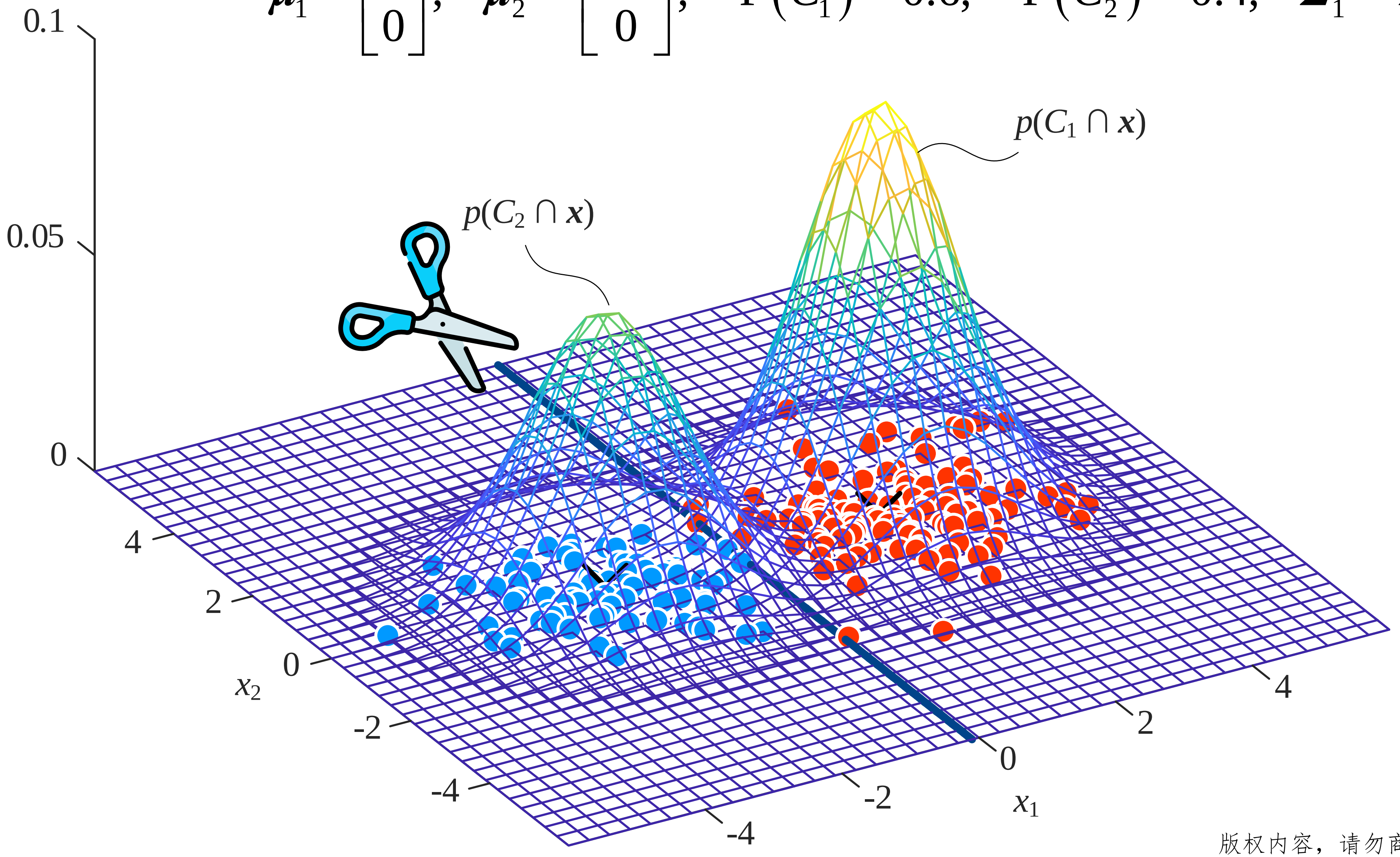
$$\left(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1\right)^{\mathrm{T}} \boldsymbol{x}-\frac{1}{2}\left(\boldsymbol{\mu}_2^{\mathrm{T}} \boldsymbol{\mu}_2-\boldsymbol{\mu}_1^{\mathrm{T}} \boldsymbol{\mu}_1\right)=0$$

$$\Rightarrow\left(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1\right)^{\mathrm{T}} \boldsymbol{x}-\frac{1}{2}\left(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1\right)^{\mathrm{T}}\left(\boldsymbol{\mu}_2+\boldsymbol{\mu}_1\right)=0$$

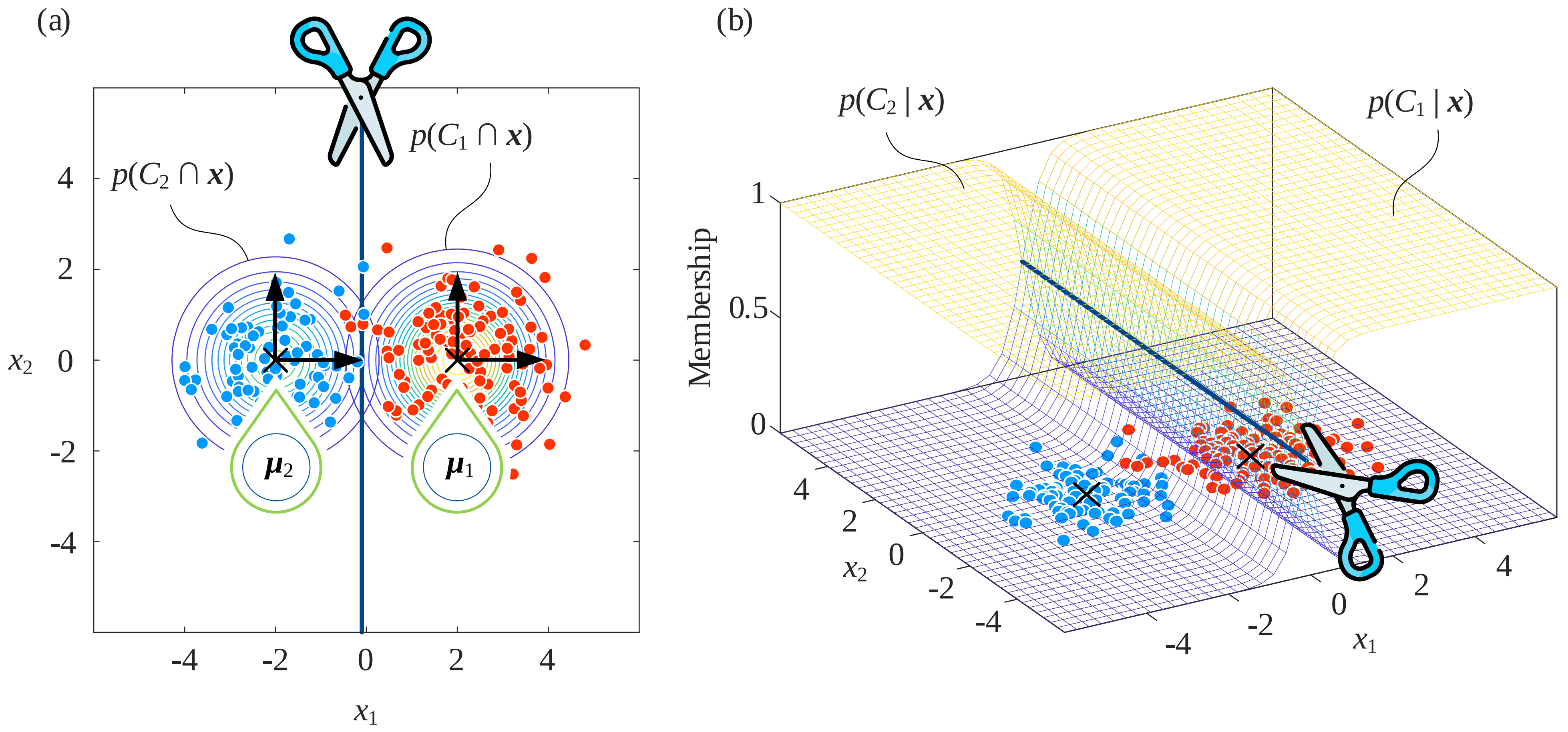
$$\Rightarrow\left(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1\right)^{\mathrm{T}}\left[\boldsymbol{x}-\frac{1}{2}\left(\boldsymbol{\mu}_2+\boldsymbol{\mu}_1\right)\right]=0$$

第一类高斯判别分析

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \quad \mathrm{P}(C_1) = 0.6, \quad \mathrm{P}(C_2) = 0.4, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



第一类高斯判别分析


(a)
$p(C_1 \cap \boldsymbol{x})$
$p(C_2 \cap \boldsymbol{x})$
$\boldsymbol{\mu}_2$
$\boldsymbol{\mu}_1$
x_1
x_2
(b)
$p(C_2 \mid \boldsymbol{x})$
$p(C_1 \mid \boldsymbol{x})$
Membership
x_1
x_2

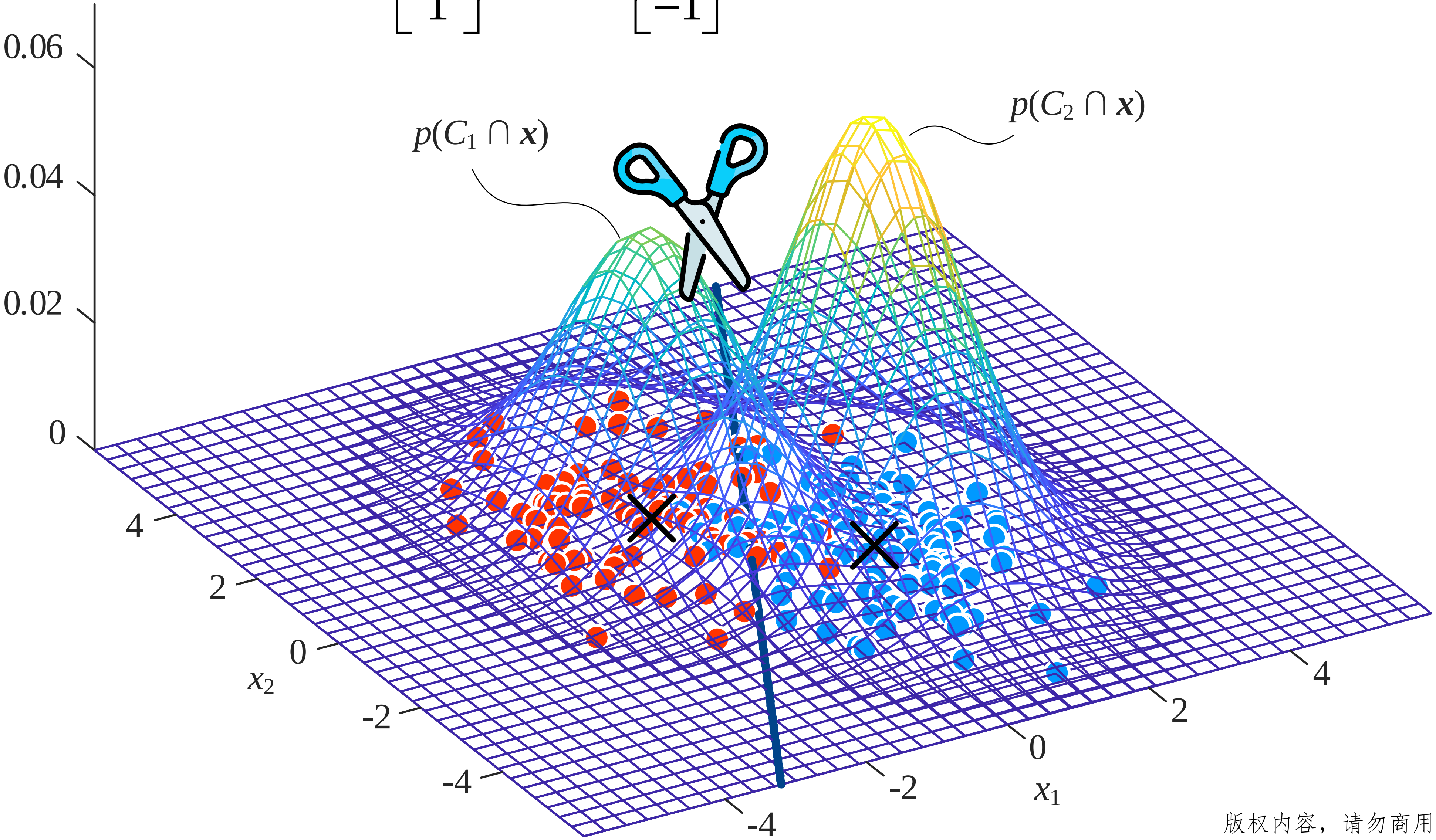
第二类高斯判别分析

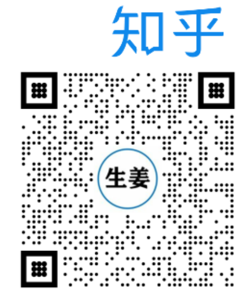
$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}(C_1) = 0.4, \quad \mathrm{P}(C_2) = 0.6, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

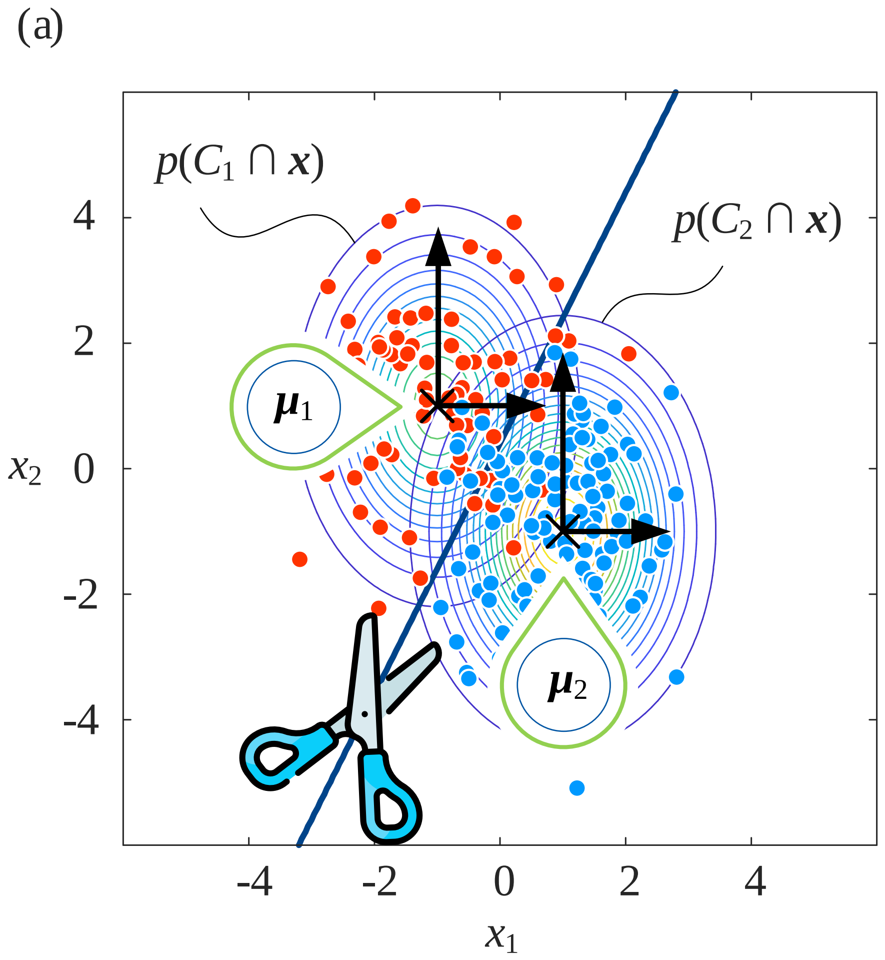
第二类高斯判别分析

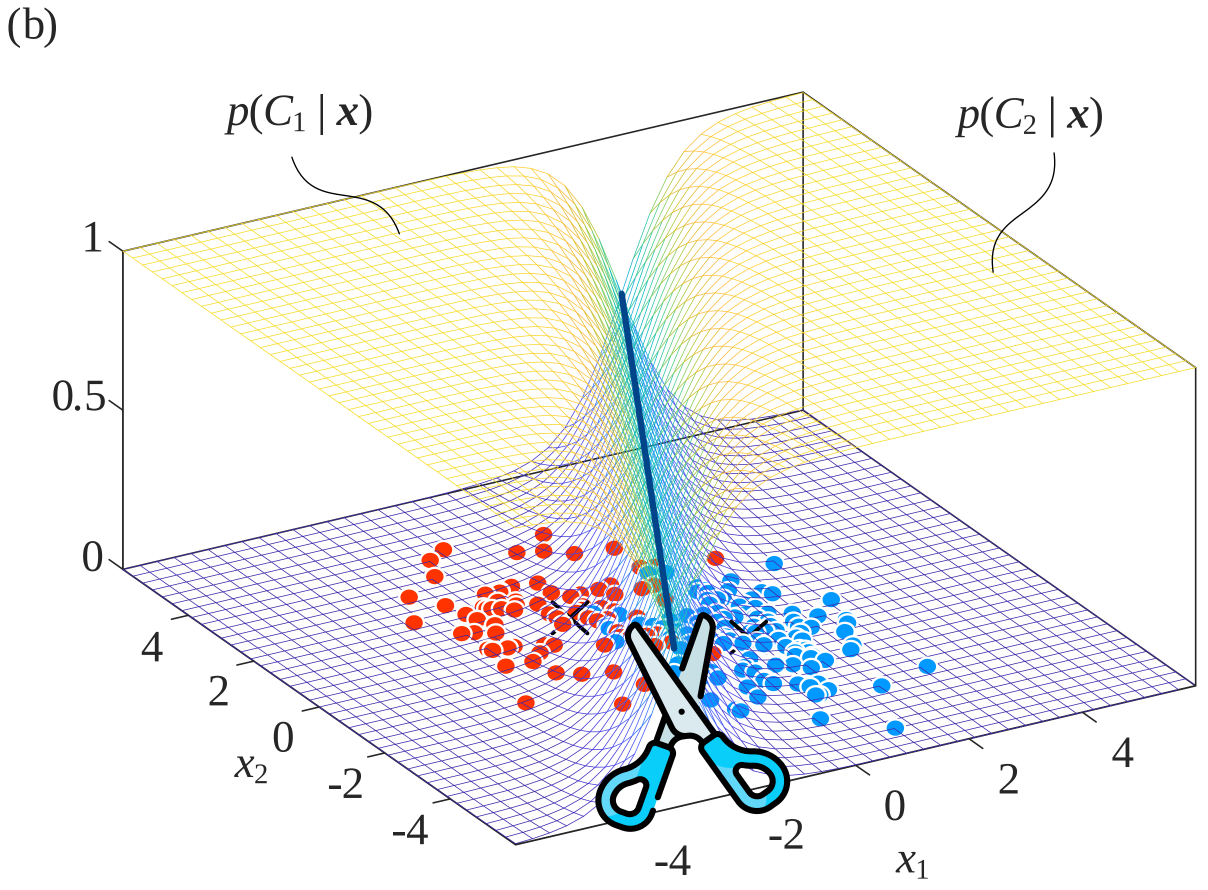
$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}(C_1) = 0.4, \quad \mathrm{P}(C_2) = 0.6, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$





第二类高斯判别分析

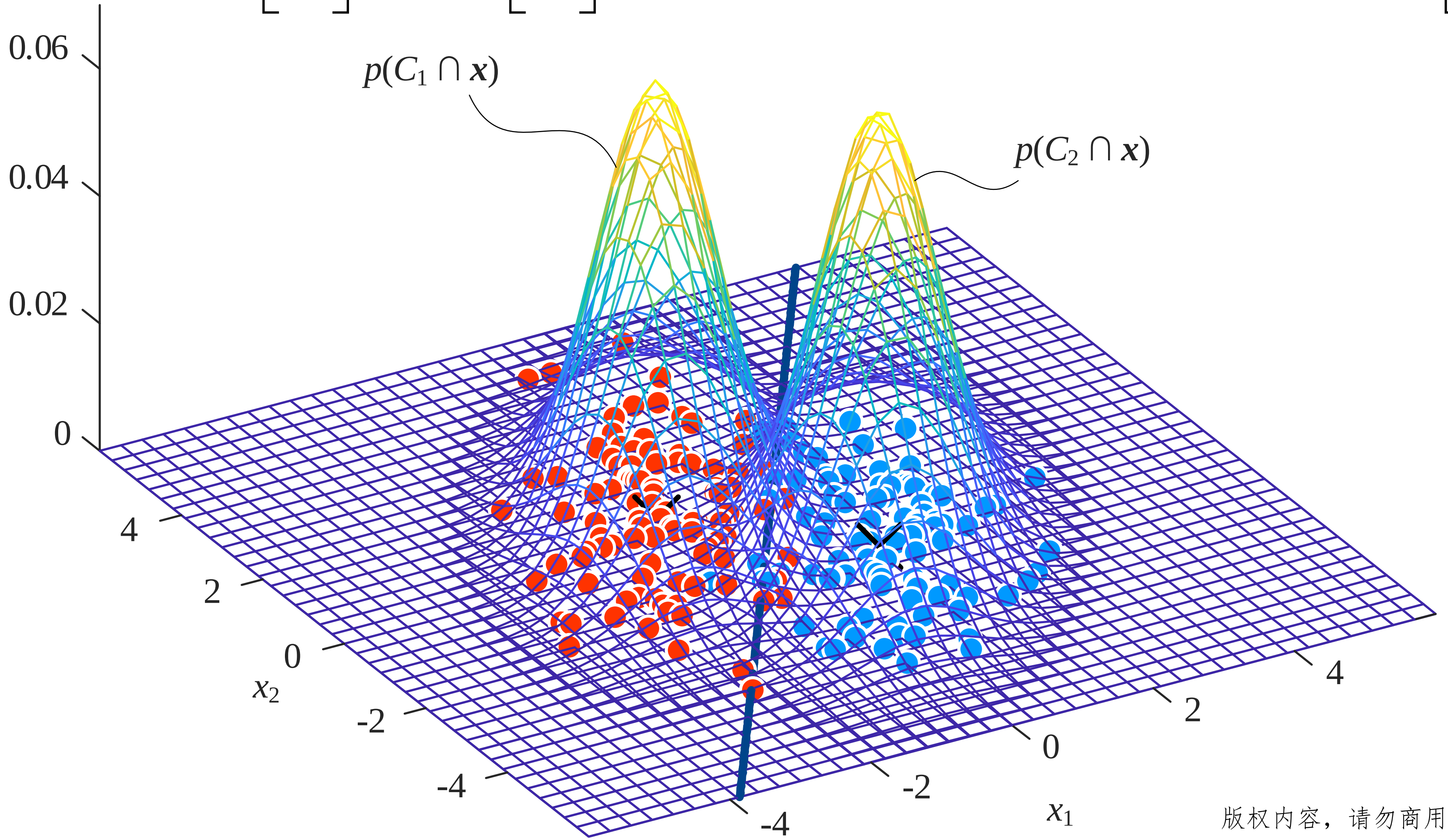

(a)
$p(C_1 \cap \boldsymbol{x})$
$p(C_2 \cap \boldsymbol{x})$
$\boldsymbol{\mu}_1$
$\boldsymbol{\mu}_2$
x_1
x_2
4
2
0
-2
-4


(b)
$p(C_1 \mid \boldsymbol{x})$
$p(C_2 \mid \boldsymbol{x})$
1
0.5
0
x_1
x_2
4
2
0
-2
-4

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}\left(C_1\right) = 0.5, \quad \mathrm{P}\left(C_2\right) = 0.5, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 2 \end{bmatrix}$$

第三类高斯判别分析

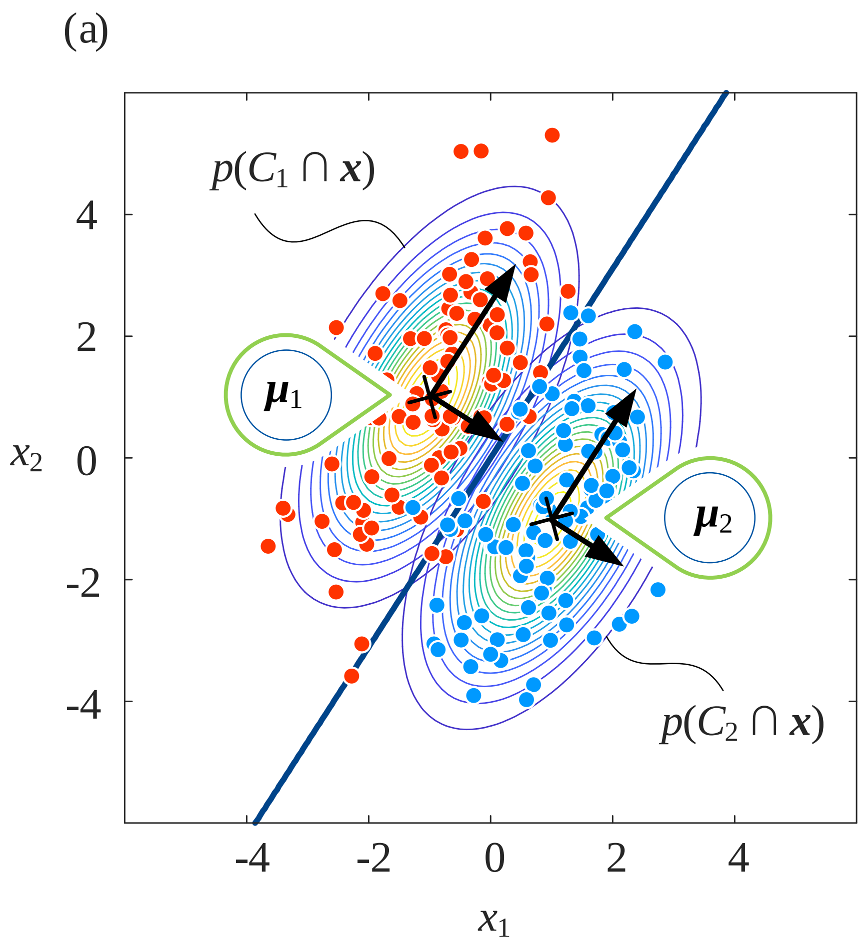
$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}(C_1) = 0.5, \quad \mathrm{P}(C_2) = 0.5, \quad \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 2 \end{bmatrix}$$



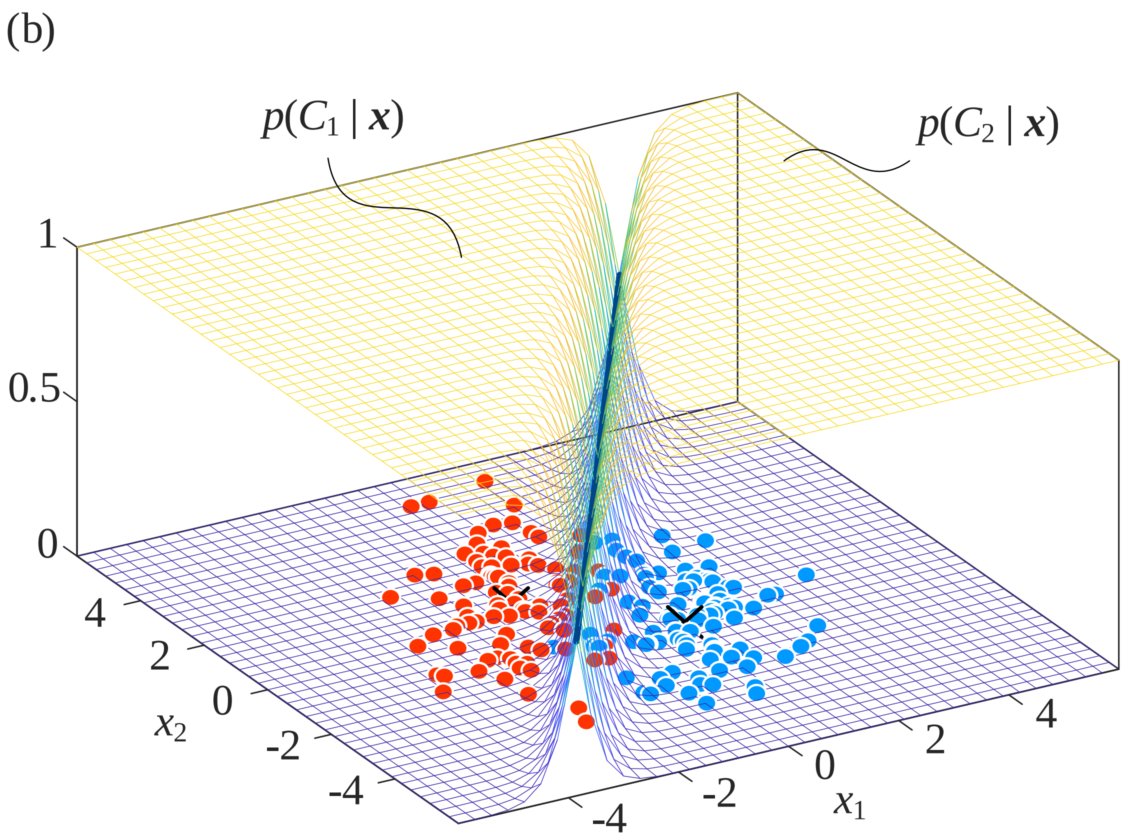
bilibili

知乎

第三类高斯判别分析



(a)
$p(C_1 \cap \boldsymbol{x})$
$\boldsymbol{\mu}_1$
$\boldsymbol{\mu}_2$
$p(C_2 \cap \boldsymbol{x})$
x_1
x_2



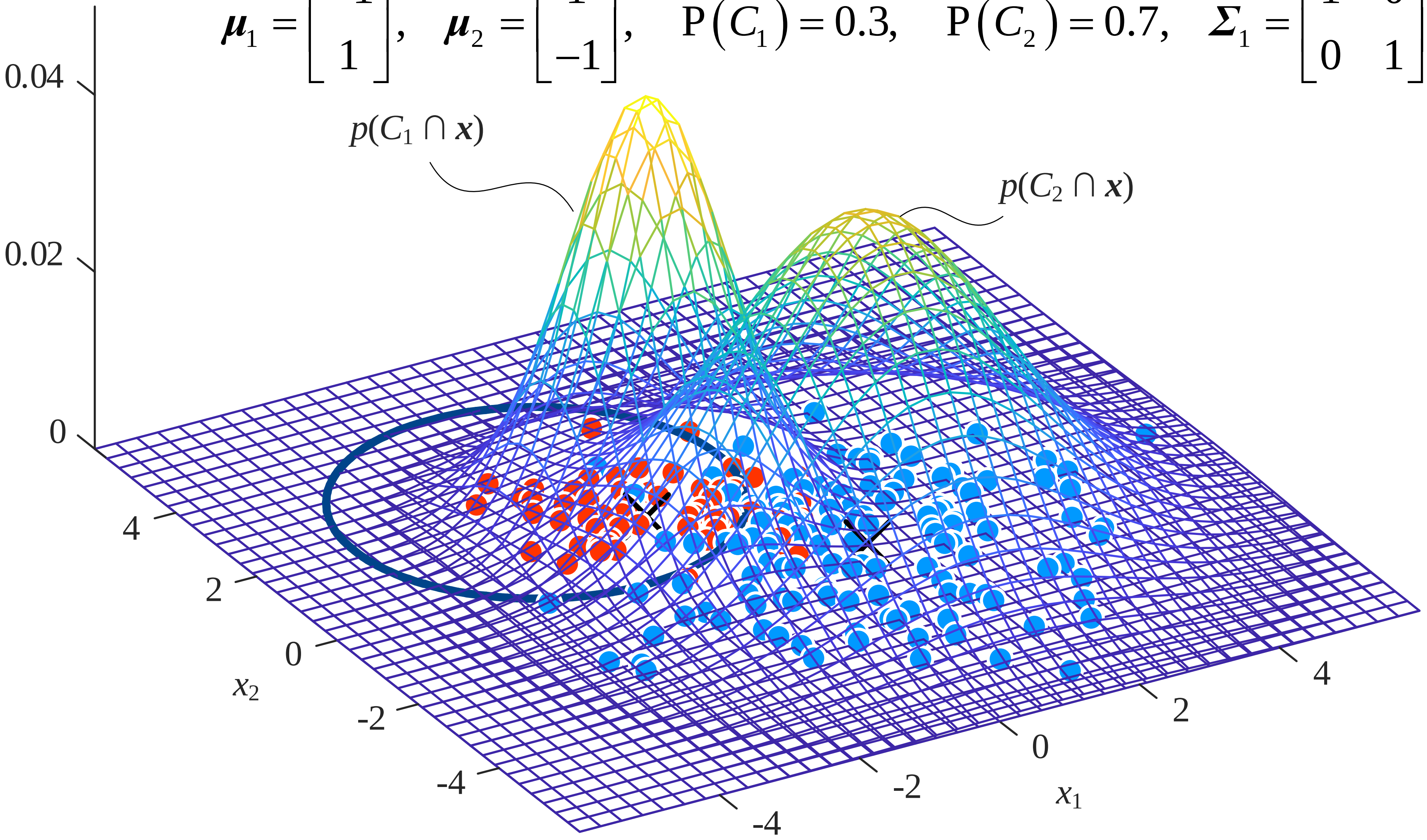
(b)
$p(C_1 \mid \boldsymbol{x})$
$p(C_2 \mid \boldsymbol{x})$
x_1
x_2

第四类高斯判别分析

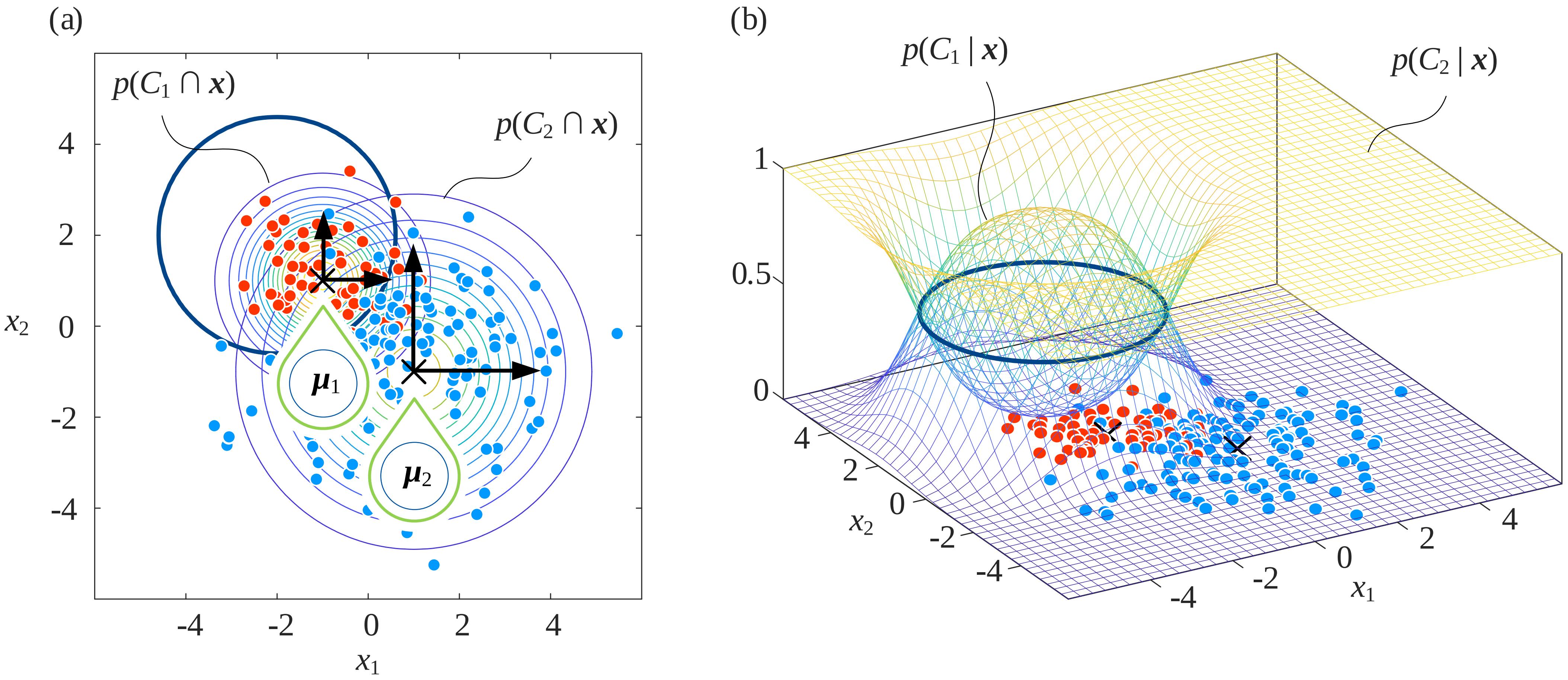
$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}(C_1) = 0.3, \quad \mathrm{P}(C_2) = 0.7, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

第四类高斯判别分析

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}(C_1) = 0.3, \quad \mathrm{P}(C_2) = 0.7, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$



第四类高斯判别分析



bilibili

知乎

第五类高斯判别分析

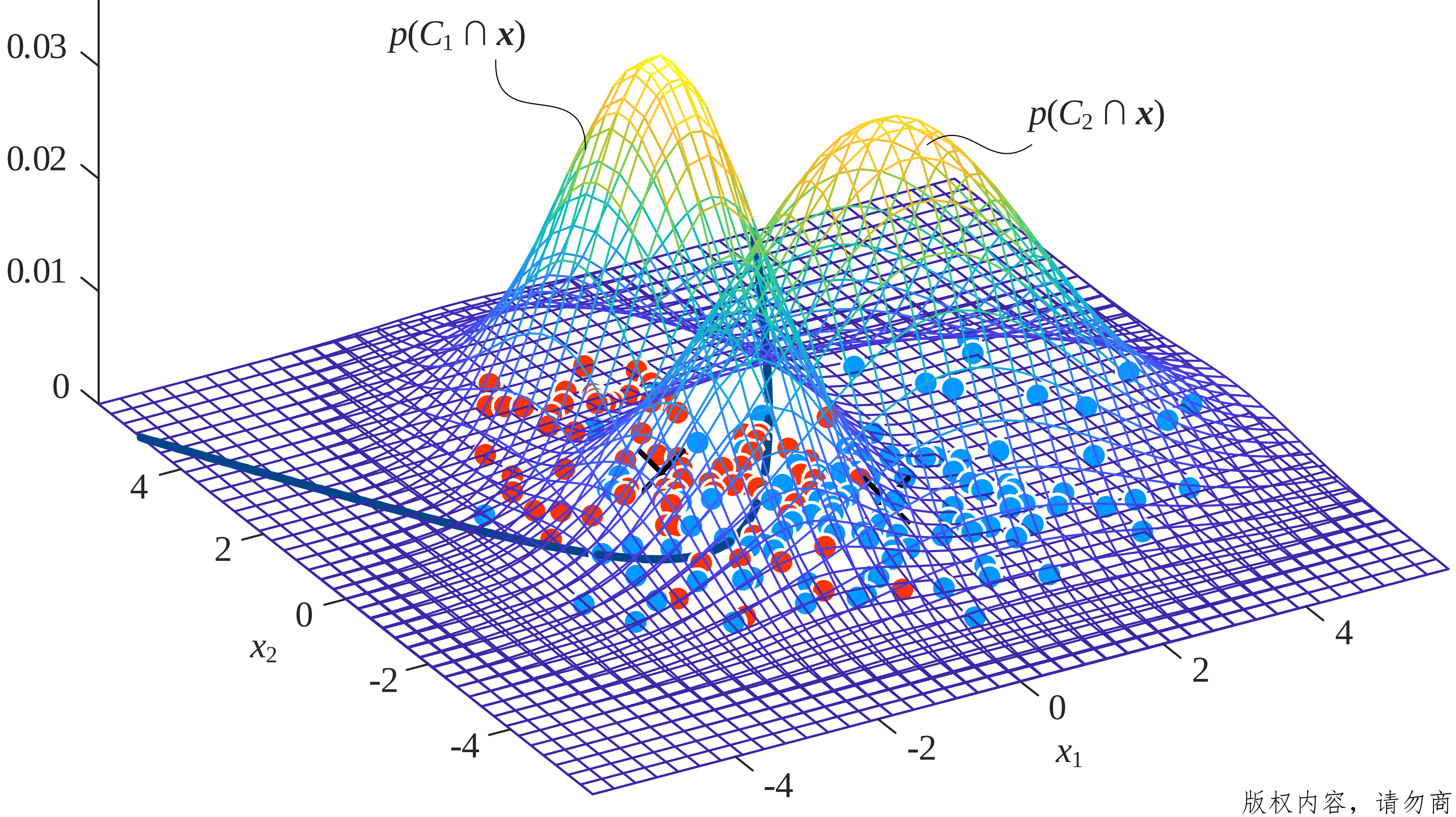
$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathrm{P}\left(C_1\right) = 0.4, \quad \mathrm{P}\left(C_2\right) = 0.6, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

bilibili

知乎

第五类高斯判别分析

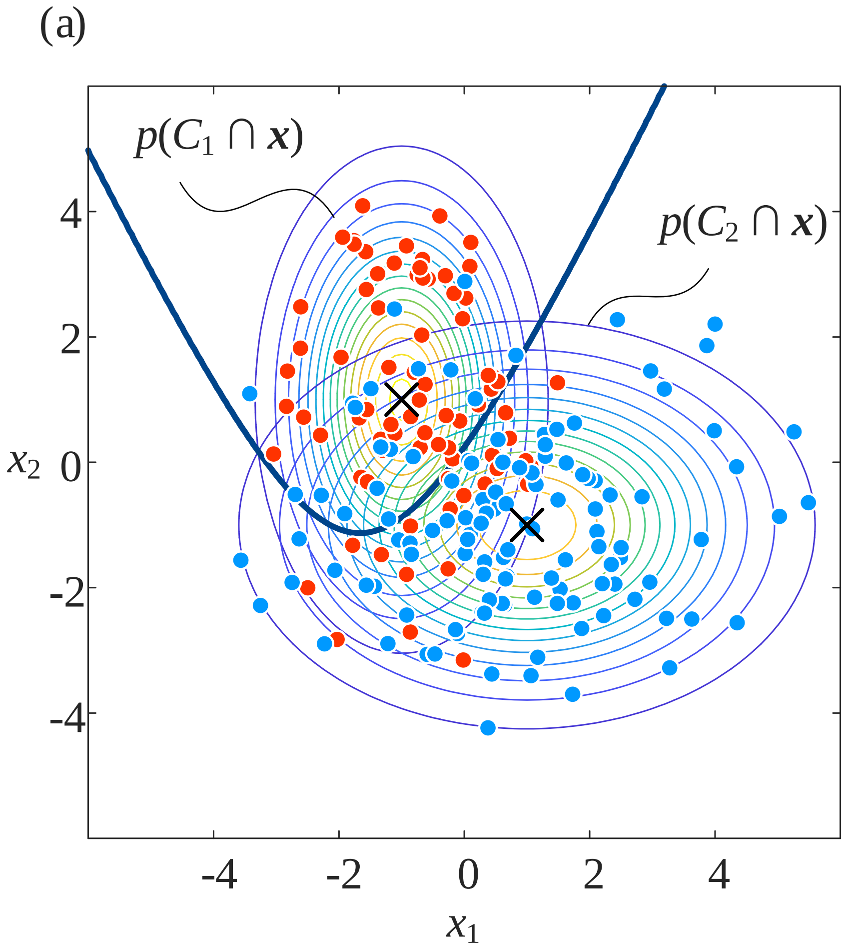
$$\boldsymbol{\mu}_1=\begin{bmatrix}-1\\1\end{bmatrix},\quad \boldsymbol{\mu}_2=\begin{bmatrix}1\\-1\end{bmatrix},\quad \mathrm{P}(C_1)=0.4,\quad \mathrm{P}(C_2)=0.6,\quad \boldsymbol{\Sigma}_1=\begin{bmatrix}1&0\\0&3\end{bmatrix},\quad \boldsymbol{\Sigma}_2=\begin{bmatrix}4&0\\0&2\end{bmatrix}$$



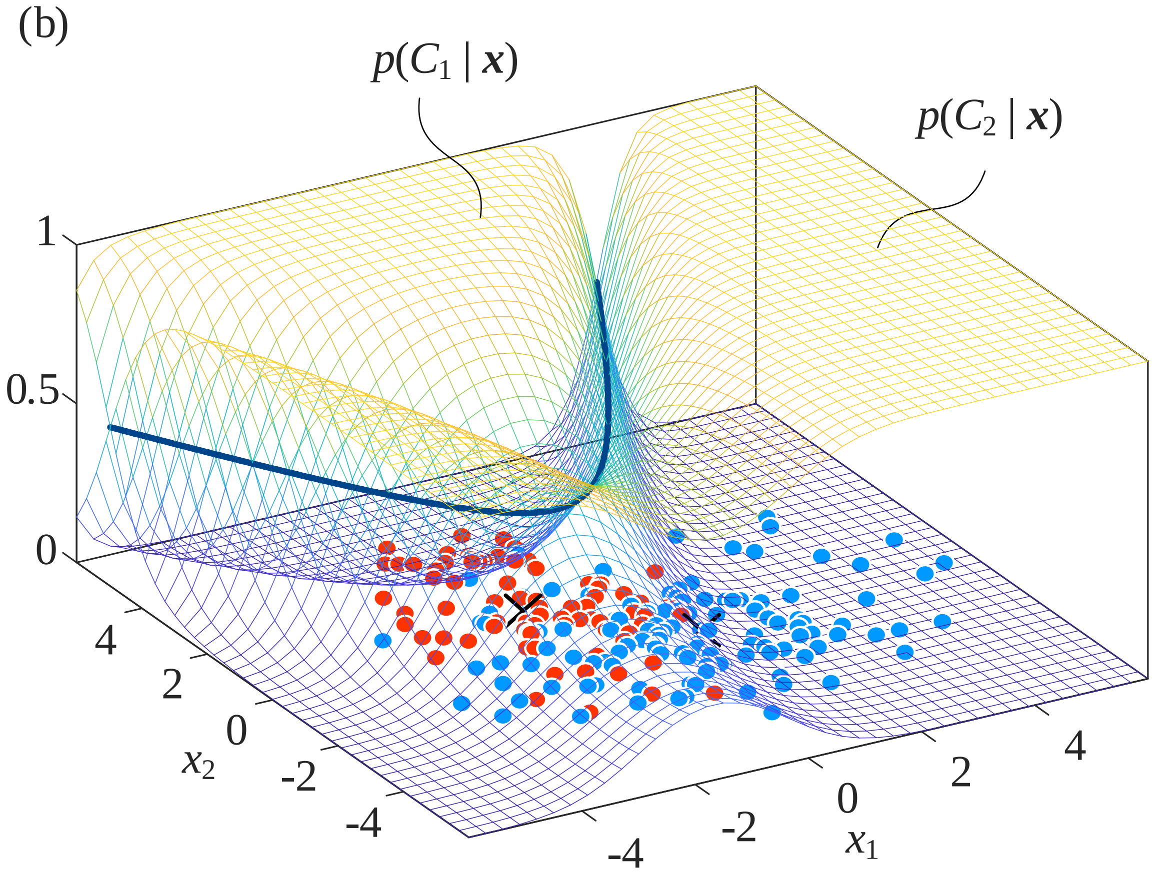
bilibili

知乎

第五类高斯判别分析



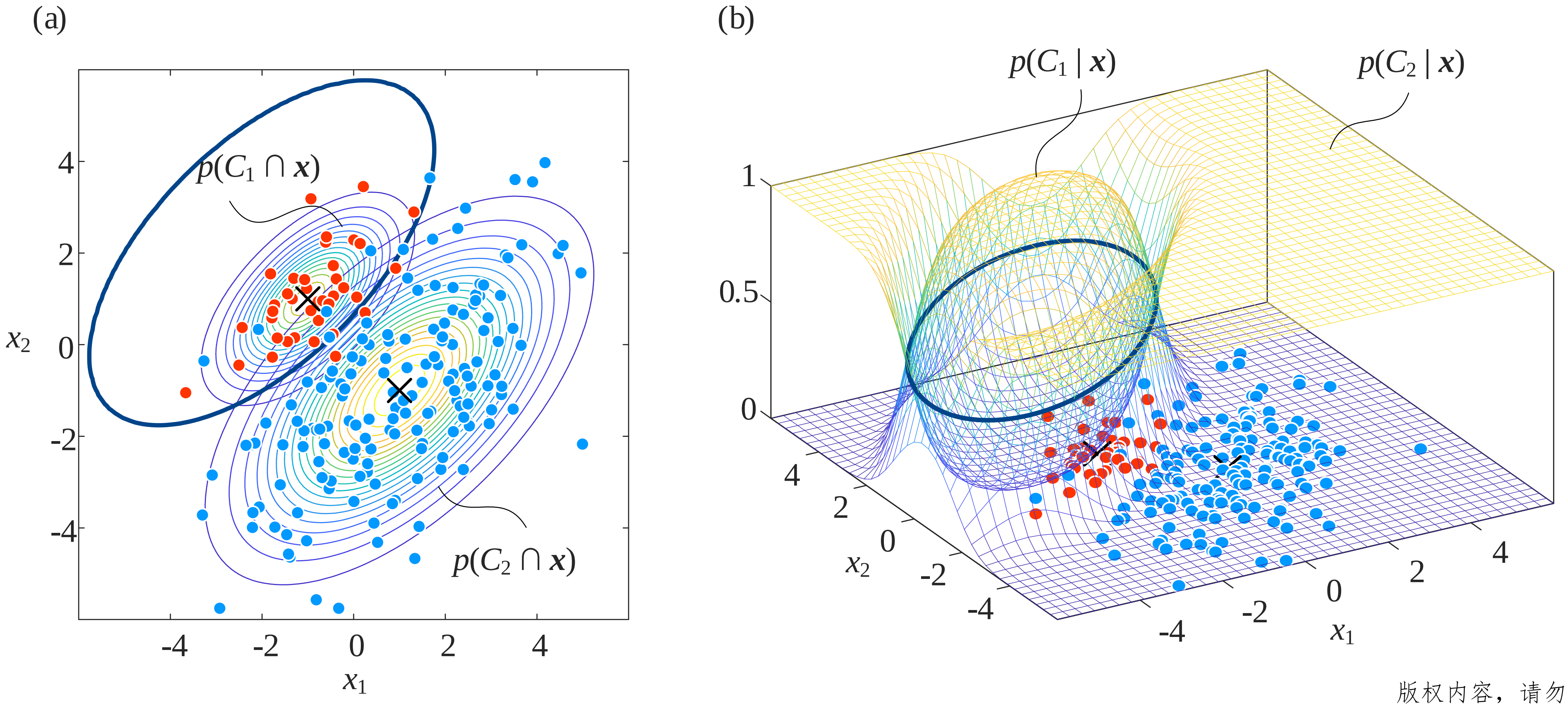
(a)
$p(C_1 \cap \boldsymbol{x})$
$p(C_2 \cap \boldsymbol{x})$
4
2
0
-2
-4
x_2
-4
-2
0
2
4
x_1



(b)
$p(C_1 \mid \boldsymbol{x})$
$p(C_2 \mid \boldsymbol{x})$
1
0.5
0
4
2
0
-2
-4
x_2
-4
-2
0
2
4
x_1

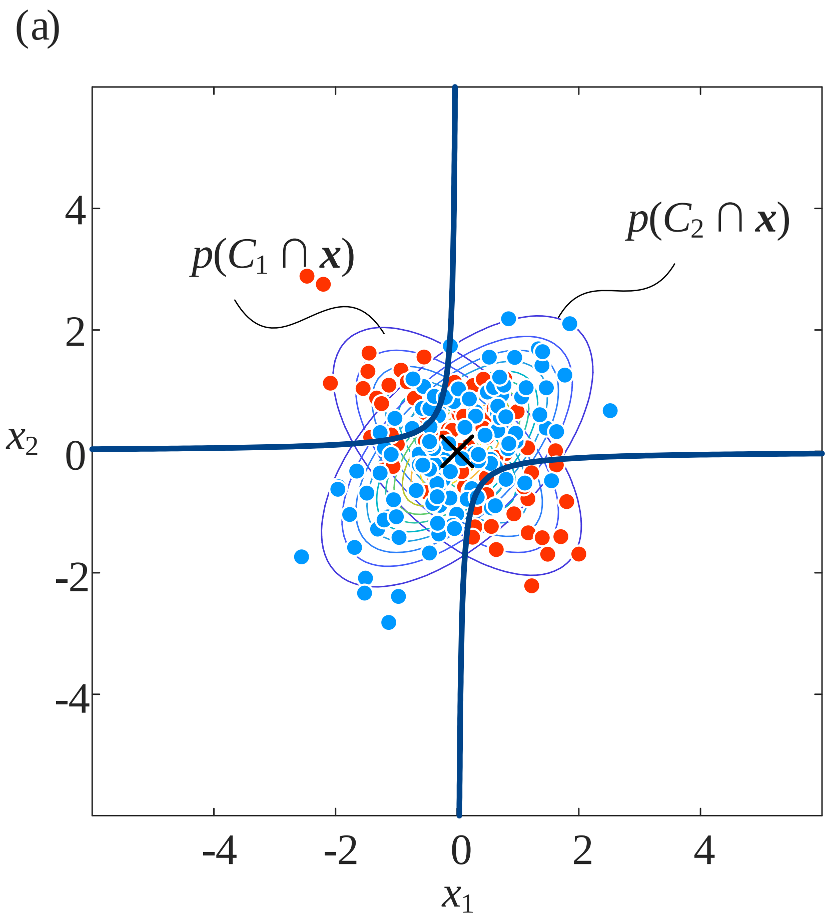
第六类高斯判别分析

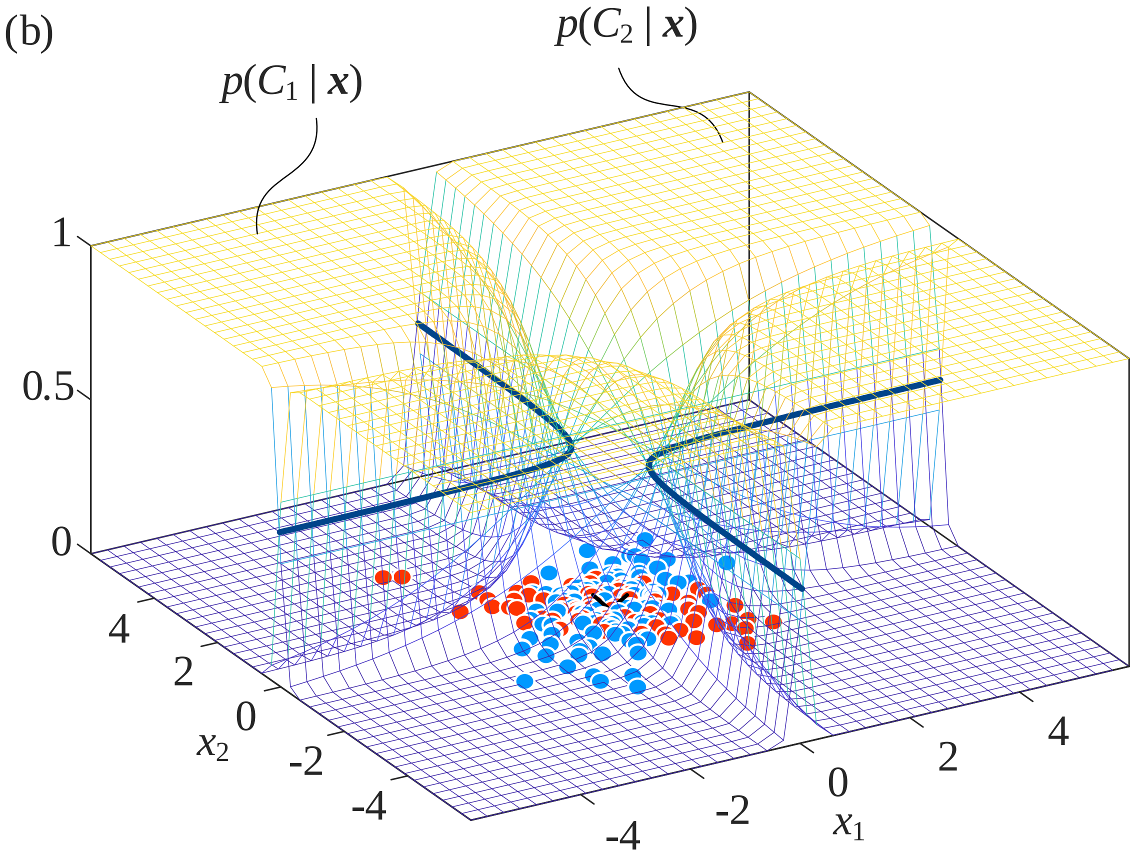
$$\boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \begin{cases} \mathrm{P}(C_1) = 0.2 \\ \mathrm{P}(C_2) = 0.8 \end{cases}, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 3 & 1.8 \\ 1.8 & 3 \end{bmatrix}$$



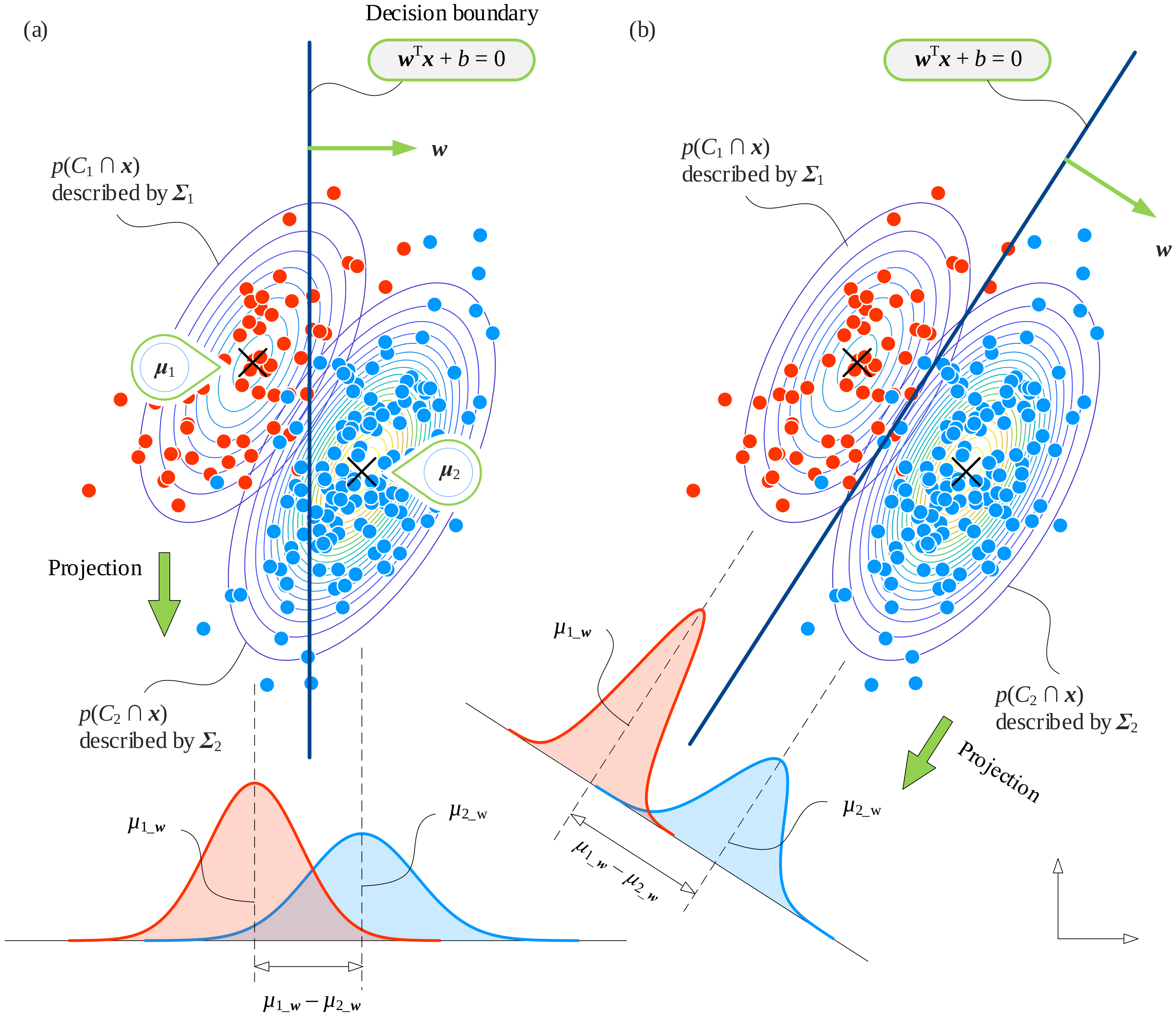
第六类高斯判别分析

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{cases} \mathrm{P}(C_1) = 0.4 \\ \mathrm{P}(C_2) = 0.6 \end{cases}, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & -0.6 \\ -0.6 & 1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix}$$

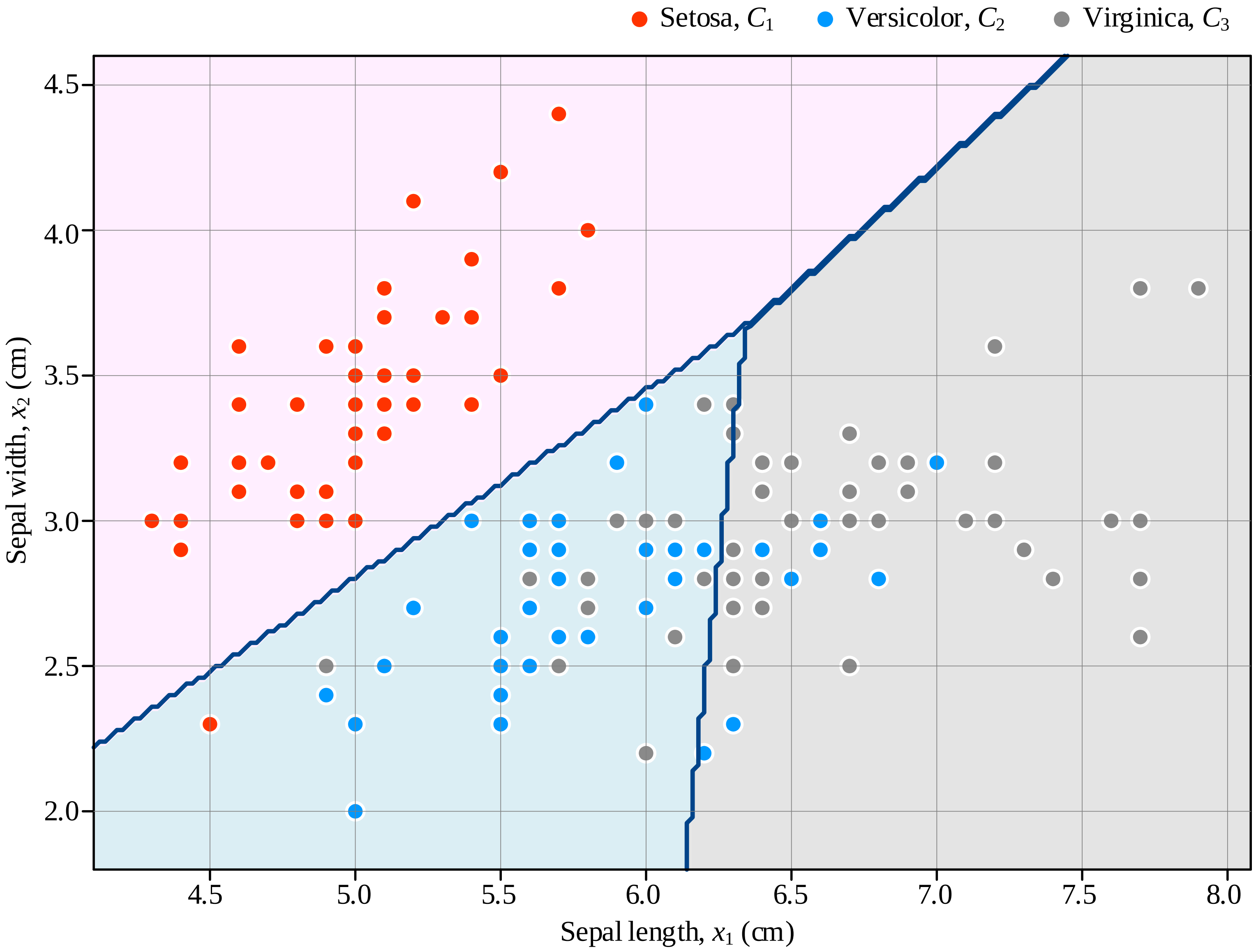




从投影角度解释线性判别分析



线性判别分析分类鸢尾花



bilibili

知乎

二次判别分析分类鸢尾花