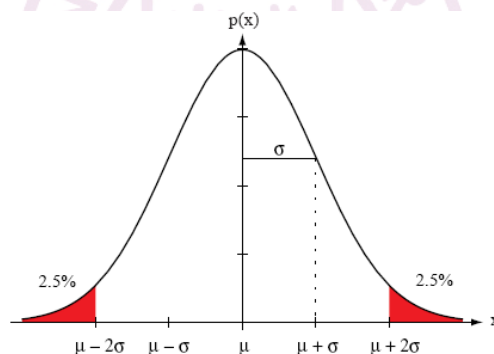


# 正态分布模式的贝叶斯分类器

南开大学网络空间安全学院 计算机学院

## 正态概率密度分布函数

- ❖ 实际工作中经常遇到的概率密度分布函数
- ❖ 经常使用正态概率密度函数的原因
  - ✎ 容易计算
  - ✎ 研究透彻
  - ✎ 建模符合实际



南开大学网络空间安全学院 计算机学院

2

## 正态概率密度分布函数

- 在  $d$  维模式特征空间中,  $\omega_i$  类模式特征  $\mathbf{x}$  的似然函数遵循多元正态概率密度分布

$$p(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right), i=1, \dots, C$$

其中:

- $\omega_i$  类样本的均值  $\boldsymbol{\mu}_i = \mathbb{E}[\mathbf{x}]$
- $d \times d$  维协方差矩阵  $\Sigma_i = \mathbb{E}[(\mathbf{x}-\boldsymbol{\mu}_i)(\mathbf{x}-\boldsymbol{\mu}_i)^T]$
- $|\Sigma_i|$  表示  $\Sigma_i$  的行列式
- $\mathbb{E}[\cdot]$  表示数学期望
- 符号  $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$  表示均值为  $\boldsymbol{\mu}$  协方差矩阵为  $\Sigma$  的正态概率密度分布函数

## 正态分布模式的贝叶斯分类器

- 最大后验概率决策规则可以表示为判别式函数的形式

$$\text{Decide } \mathbf{x} \in \omega_i \text{ if } g_i(\mathbf{x}) > g_j(\mathbf{x}) \quad \forall j \neq i$$

其中  $g_i(\mathbf{x}) = f(\Pr(\omega_i|\mathbf{x}))$ 。

- 当模式特征遵循正态概率密度分布时, 最大后验概率决策规则将具有更加简洁的形式。
- 依据贝叶斯规则, 后验概率表示为

$$\begin{aligned} \Pr(\omega_i|\mathbf{x}) &= \frac{p(\mathbf{x}|\omega_i)\Pr(\omega_i)}{p(\mathbf{x})} \\ &= \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right] \frac{\Pr(\omega_i)}{p(\mathbf{x})} \end{aligned}$$

## 正态分布模式的贝叶斯分类器

除去对所有模式类别都相同的常数项，得到

$$\Pr(\omega_i | \mathbf{x}) \propto |\Sigma_i|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right] \Pr(\omega_i)$$

令单增函数  $f(\cdot) = \ln(\cdot)$ ，判别式函数为

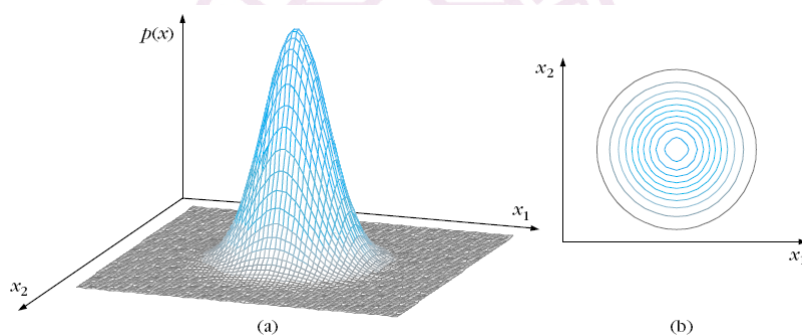
$$g_i(\mathbf{x}) = -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\Sigma_i|) + \ln(\Pr(\omega_i))$$

上式称为二次判别式函数。

## 情形一

- 各个模式类别具有相等的对角化协方差矩阵且各个模式特征的方差也相等，即

$$\sigma_{ij} = \begin{cases} \sigma_0^2 & i = j \\ 0 & i \neq j \end{cases}$$



## 情形一

### ❖ 判别式函数

$$\begin{aligned}
 g(\mathbf{x}_i) &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T (\sigma_0^2)^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\sigma_0^2|) + \ln(\Pr(\omega_i)) \\
 &= -\frac{1}{2\sigma_0^2}(\mathbf{x} - \boldsymbol{\mu}_i)^T (\mathbf{x} - \boldsymbol{\mu}_i) - \underbrace{\frac{1}{2} d \ln(\sigma_0^2)}_{\text{independent of class } i} + \ln(\Pr(\omega_i)) \\
 &= -\frac{1}{2\sigma_0^2}(\mathbf{x} - \boldsymbol{\mu}_i)^T (\mathbf{x} - \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i)) \\
 &= -\frac{1}{2\sigma_0^2}(\mathbf{x}^T \mathbf{x} - 2\boldsymbol{\mu}_i^T \mathbf{x} + \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i))
 \end{aligned}$$

南开大学网络空间安全学院 计算机学院

7

## 情形一

### ❖ 除去与模式类别无关的项，有

$$\begin{aligned}
 g_i(\mathbf{x}) &= -\frac{1}{2\sigma_0^2}(-2\boldsymbol{\mu}_i^T \mathbf{x} + \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i)) \\
 &= \mathbf{w}_i^T \mathbf{x} + w_{i0}
 \end{aligned}$$

其中：

$$\mathbf{w}_i = \frac{\boldsymbol{\mu}_i}{\sigma_0^2}$$

$$w_{i0} = -\frac{1}{2\sigma_0^2} \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i + \ln(\Pr(\omega_i))$$

### ❖ 判别式函数是线性的，决策面 $g_i(\mathbf{x}) = g_j(\mathbf{x})$ 是超平面。

南开大学网络空间安全学院 计算机学院

8

## 情形一

- ❖ 如果假设模式类别具有相同的先验概率，则

$$g_i(\mathbf{x}) = -\frac{1}{2\sigma_0^2}(\mathbf{x} - \boldsymbol{\mu}_i)^T(\mathbf{x} - \boldsymbol{\mu}_i)$$

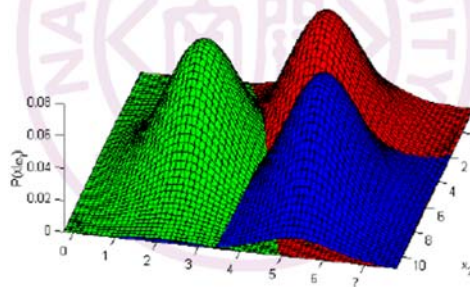
### 说明

- ✧ 该判别式函数称为最小（欧氏）距离分类器或最接近均值分类器
- ✧ 每个模式类别的常数概率轨迹是超球面
- ✧ 当单位方差时，距离变为欧几里得距离

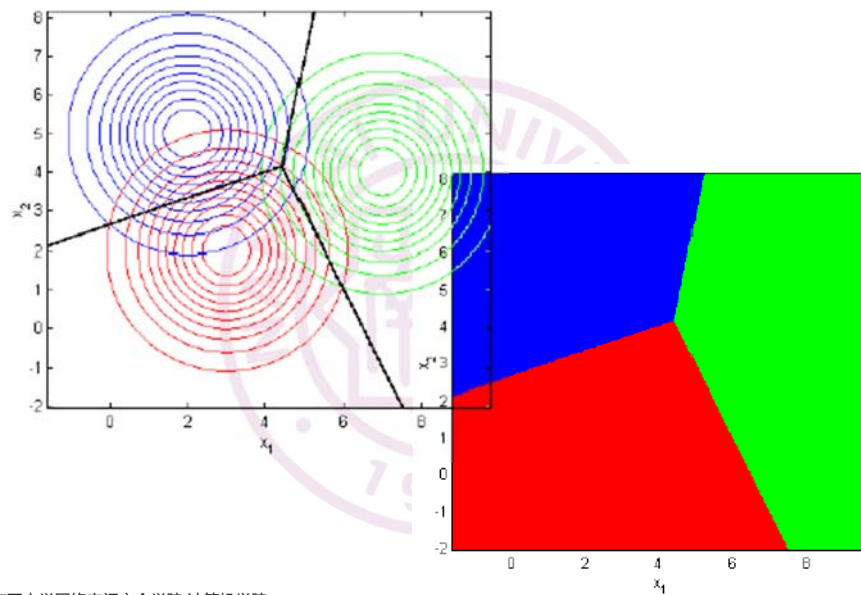
## 情形一

- ❖ 例：三个模式类别分别具有如下均值矢量、协方差矩阵和相等的先验概率，求解在模式特征空间中的决策界面。

$$\begin{aligned} \boldsymbol{\mu}_1 &= (3 \ 2)^T & \boldsymbol{\mu}_2 &= (7 \ 4)^T & \boldsymbol{\mu}_3 &= (2 \ 5)^T \\ \Sigma_1 &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} & \Sigma_2 &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} & \Sigma_3 &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$



## 情形一



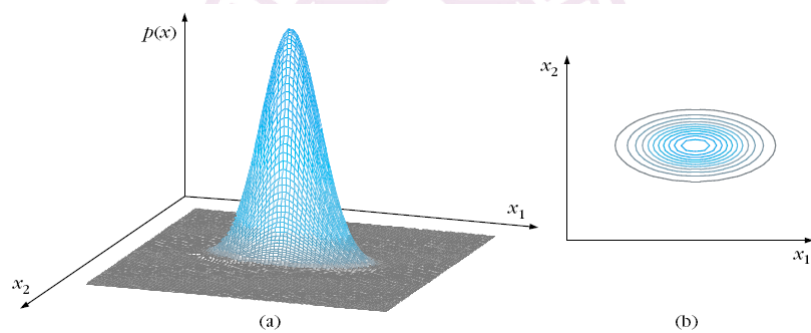
南开大学网络空间安全学院 计算机学院

11

## 情形二

- 各个模式类别具有相等的对角化协方差矩阵且各个模式特征的方差不同，即

$$\sigma_{ij} \begin{cases} \neq 0 & i = j \\ = 0 & i \neq j \end{cases}$$



南开大学网络空间安全学院 计算机学院

12

## 情形二

### ❖ 判别式函数

$$\begin{aligned}
 g_i(\mathbf{x}) &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\boldsymbol{\Sigma}_i|) + \ln(\Pr(\omega_i)) \\
 &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \begin{bmatrix} \sigma_1^{-2} & & \\ & \ddots & \\ & & \sigma_d^{-2} \end{bmatrix} (\mathbf{x} - \boldsymbol{\mu}_i) \\
 &\quad - \frac{1}{2} \ln \begin{vmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_d^2 \end{vmatrix} + \ln(\Pr(\omega_i)) \\
 &= -\frac{1}{2} \sum_{l=1}^d \frac{(x_l - \mu_{i,l})^2}{\sigma_l^2} - \frac{1}{2} \ln \prod_{l=1}^d \sigma_l^2 + \ln(\Pr(\omega_i))
 \end{aligned}$$

南开大学网络空间安全学院 计算机学院

13

## 情形二

$$= -\frac{1}{2} \sum_{l=1}^d \frac{x_l^2 - 2x_l \mu_{i,l} + \mu_{i,l}^2}{\sigma_l^2} - \frac{1}{2} \ln \prod_{l=1}^d \sigma_l^2 + \ln(\Pr(\omega_i))$$

其中： $x_l$  和  $\mu_{i,l}$  分别是  $\mathbf{x}$  和  $\boldsymbol{\mu}_i$  的第  $l$  维分量

### ❖ 除去与模式类别无关的项，有

$$g_i(\mathbf{x}) = -\frac{1}{2} \sum_{l=1}^d \frac{-2x_l \mu_{i,l} + \mu_{i,l}^2}{\sigma_l^2} + \ln(\Pr(\omega_i))$$

### ❖ 说明

- ✎ 因为判别式函数是线性的，故决策界面  $g_i(\mathbf{x}) = g_j(\mathbf{x})$  是超平面。
- ✎ 每个模式类别的常数概率轨迹是沿着特征轴方向的超椭圆。
- ✎ 与情形一中分类器的区别是每根轴上的距离被该轴的方差归一化。

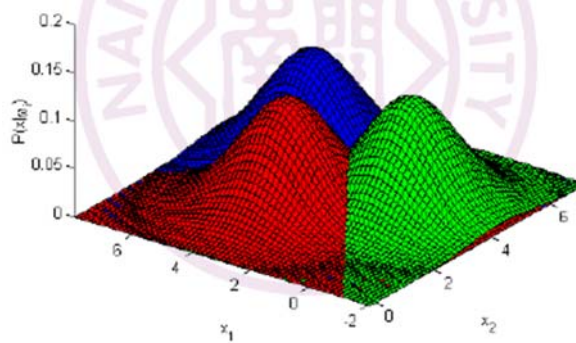
南开大学网络空间安全学院 计算机学院

14

## 情形二

❖ 例：三个模式类别分别具有如下均值矢量、协方差矩阵和相等的先验概率，求解在模式特征空间中的决策界面。

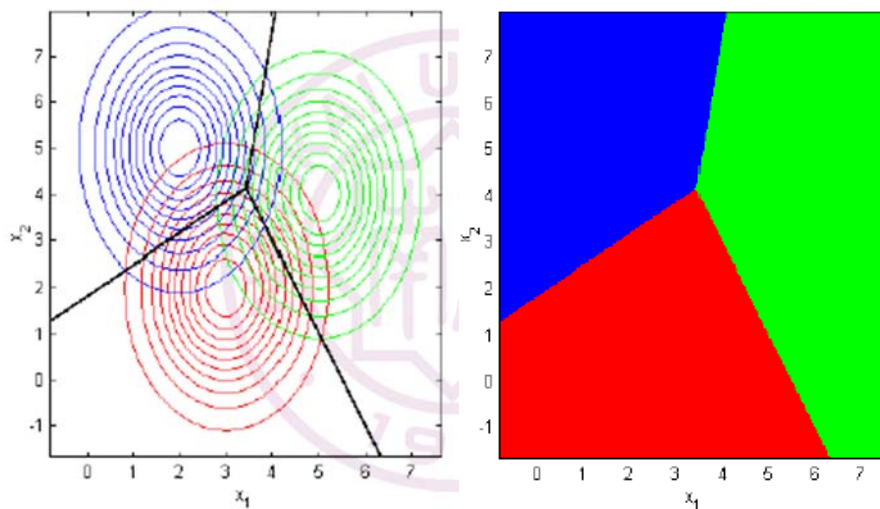
$$\begin{aligned}\mu_1 &= (3 \ 2)^T & \mu_2 &= (5 \ 4)^T & \mu_3 &= (2 \ 5)^T \\ \Sigma_1 &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} & \Sigma_2 &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} & \Sigma_3 &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\end{aligned}$$



南开大学网络空间安全学院 计算机学院

15

## 情形二



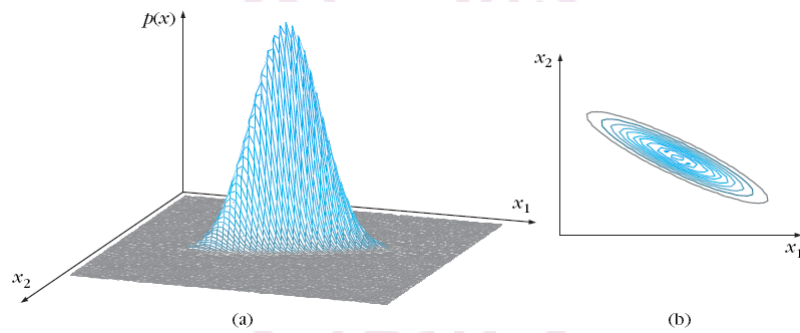
南开大学网络空间安全学院 计算机学院

16



### 情形三

- 各个模式类别具有相等的非对角化协方差矩阵且各个模式特征的方差不同。



南开大学网络空间安全学院 计算机学院

17

### 情形三

- 判别式函数

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\boldsymbol{\Sigma}_i|) + \ln(\Pr(\omega_i)) \\ &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\boldsymbol{\Sigma}|) + \ln(\Pr(\omega_i)) \end{aligned}$$

- 除去与模式类别无关的项，有

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i))$$

- 展开二次项，得到

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i)) \\ &= -\frac{1}{2}(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - 2\boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i)) \end{aligned}$$

南开大学网络空间安全学院 计算机学院

18

### 情形三

❖ 除去常数项  $\mathbf{x}^T \Sigma^{-1} \mathbf{x}$ , 有

$$g_i(\mathbf{x}) = -\frac{1}{2}(-2\boldsymbol{\mu}_i^T \Sigma^{-1} \mathbf{x} + \boldsymbol{\mu}_i^T \Sigma^{-1} \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i))$$

重新排列得到

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

其中:

$$\mathbf{w}_i = \Sigma^{-1} \boldsymbol{\mu}_i$$

$$w_{i0} = -\frac{1}{2} \boldsymbol{\mu}_i^T \Sigma^{-1} \boldsymbol{\mu}_i + \ln(\Pr(\omega_i))$$

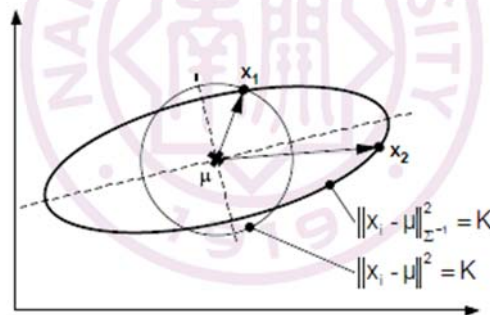
### 马氏距离

❖ 马氏距离是统计模式识别中非常重要的距离度量

❖ 计算公式:  $\|\mathbf{x} - \mathbf{y}\|_{\Sigma^{-1}}^2 = (\mathbf{x} - \mathbf{y})^T \Sigma^{-1} (\mathbf{x} - \mathbf{y})$

❖ 马氏距离是一个使用  $\Sigma^{-1}$  为模的矢量距离

- ✧  $\Sigma^{-1}$  可以认为是空间的伸缩因子
- ✧ 对于单位协方差矩阵, 马氏距离转化为熟悉的欧氏距离。



### 情形三

#### ❖ 说明

- ✎ 因为判别式函数是线性的，所以决策界面是超平面。
- ✎ 每个模式类别的常数概率轨迹是沿着协方差矩阵特征矢量方向上的超椭圆。
- ✎ 当所有模式类别的先验概率相等时，判别式函数为

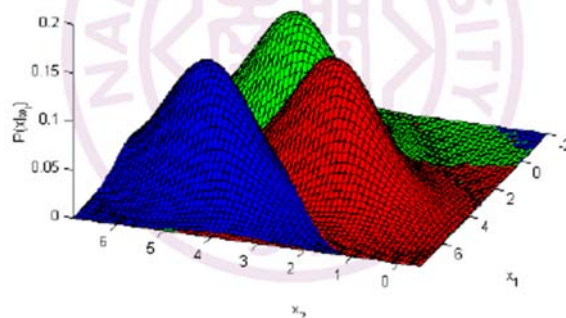
$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)$$

表示一个最小马氏距离分类器。

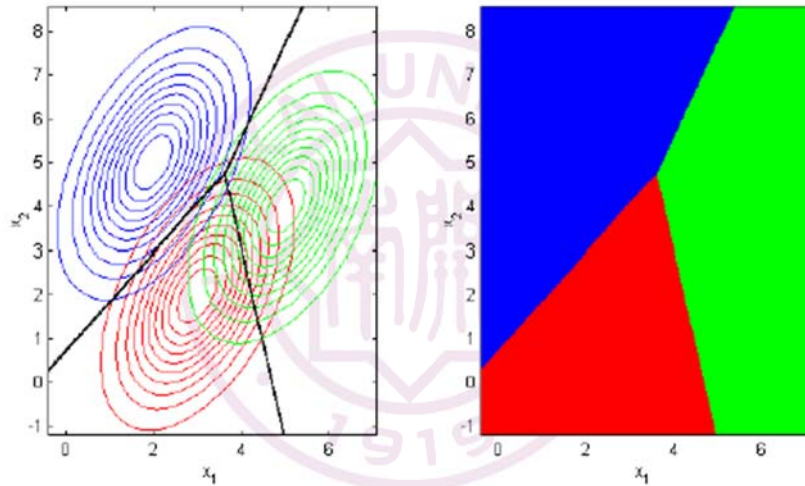
### 情形三

- ❖ 例：三个模式类别分别具有如下均值矢量、协方差矩阵和相等的先验概率，求解在模式特征空间中的决策界面。

$$\begin{aligned} \boldsymbol{\mu}_1 &= (3 \ 2)^T & \boldsymbol{\mu}_2 &= (5 \ 4)^T & \boldsymbol{\mu}_3 &= (2 \ 5)^T \\ \boldsymbol{\Sigma}_1 &= \begin{pmatrix} 1 & 0.7 \\ 0.7 & 2 \end{pmatrix} & \boldsymbol{\Sigma}_2 &= \begin{pmatrix} 1 & 0.7 \\ 0.7 & 2 \end{pmatrix} & \boldsymbol{\Sigma}_3 &= \begin{pmatrix} 1 & 0.7 \\ 0.7 & 2 \end{pmatrix} \end{aligned}$$



### 情形三



南开大学网络空间安全学院 计算机学院

23

### 情形四

- ❖ 各个模式类别具有不同的协方差矩阵，但它们与单位矩阵成比例，即  $\Sigma_i = \sigma_i^2 \mathbf{I}$ 。

- ❖ 判别式函数

$$\begin{aligned} g_i(\mathbf{x}) &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\Sigma_i|) + \ln(\Pr(\omega_i)) \\ &= -\frac{1}{2\sigma_i^2}(\mathbf{x} - \boldsymbol{\mu}_i)^T(\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} d \ln(\sigma_i^2) + \ln(\Pr(\omega_i)) \end{aligned}$$

- ❖ 说明

- ✧ 决策边界是二次的：超椭圆
- ✧ 每个模式类别的常数概率轨迹是沿着协方差矩阵特征轴方向的超球面

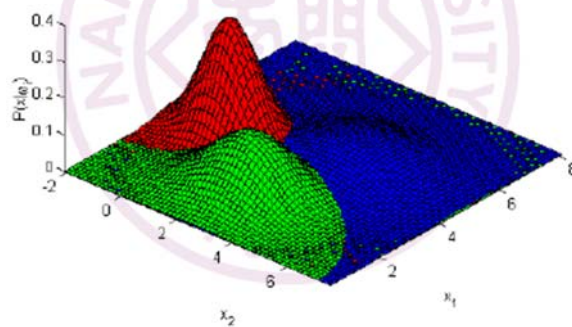
南开大学网络空间安全学院 计算机学院

24

## 情形四

❖ 例：三个模式类别分别具有如下均值矢量、协方差矩阵和相等的先验概率，求解在模式特征空间中的决策界面。

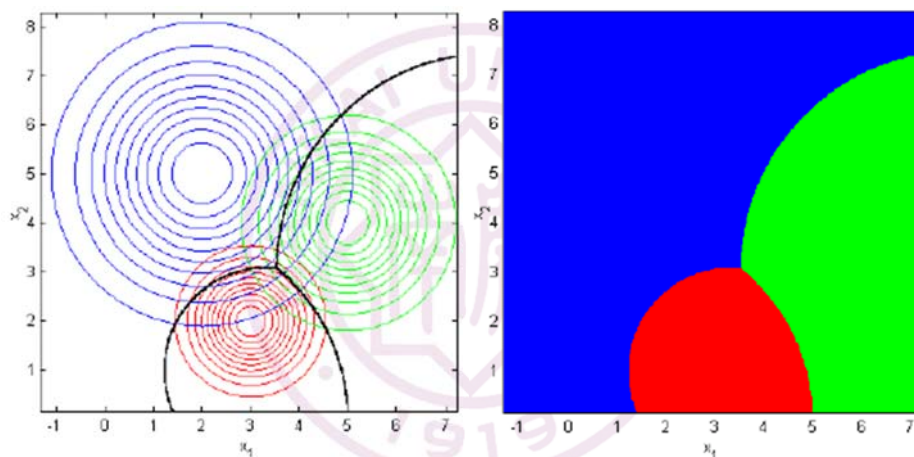
$$\begin{aligned}\mu_1 &= (3 \ 2)^T & \mu_2 &= (5 \ 4)^T & \mu_3 &= (2 \ 5)^T \\ \Sigma_1 &= \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} & \Sigma_2 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \Sigma_3 &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\end{aligned}$$



南开大学网络空间安全学院 计算机学院

25

## 情形四



南开大学网络空间安全学院 计算机学院

26

## 情形五

### ❖ 一般情形下的判别式函数

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - \frac{1}{2} \ln(|\boldsymbol{\Sigma}_i|) + \ln(\Pr(\omega_i))$$

重新排列各项, 得到

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

其中

$$\mathbf{W}_i = -\frac{1}{2} \boldsymbol{\Sigma}_i^{-1}$$

$$\mathbf{w}_i = \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\mu}_i$$

$$w_{i0} = -\frac{1}{2} \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\mu}_i - \frac{1}{2} \ln(|\boldsymbol{\Sigma}_i|) + \ln(\Pr(\omega_i))$$

## 情形五

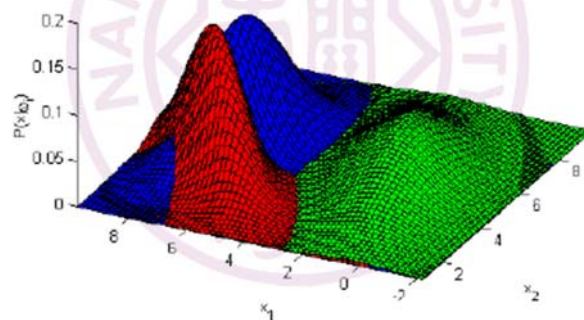
### ❖ 说明

- ✎ 决策界面是二次的: 超椭球或超抛物面
- ✎ 每个模式类别的常数概率轨迹是取向于该类协方差矩阵  $\boldsymbol{\Sigma}_i$  的特征矢量方向上的超椭球。
- ✎ 判别式函数中的二次项与马氏距离成比例

## 情形五

❖ 例：三个模式类别分别具有如下均值矢量、协方差矩阵和相等的先验概率，求解在模式特征空间中的决策界面。

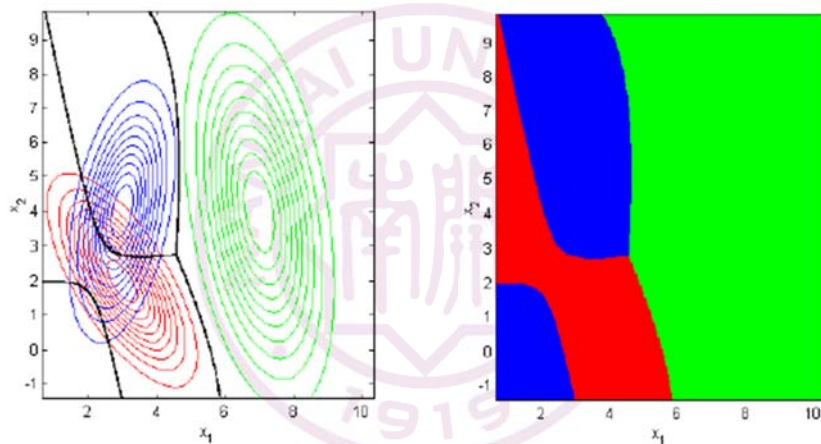
$$\begin{aligned}\mu_1 &= (3 \ 2)^T & \mu_2 &= (5 \ 4)^T & \mu_3 &= (2 \ 5)^T \\ \Sigma_1 &= \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} & \Sigma_2 &= \begin{pmatrix} 1 & -1 \\ -1 & 7 \end{pmatrix} & \Sigma_3 &= \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 3 \end{pmatrix}\end{aligned}$$



南开大学网络空间安全学院 计算机学院

29

## 情形五



南开大学网络空间安全学院 计算机学院

30



## 分类器设计实例

- ❖ **问题：**由下面正态分布定义了三维空间中的两类问题，试写出分类器的决策规则。

$$\boldsymbol{\mu}_1 = (0 \ 0 \ 0)^T, \boldsymbol{\mu}_2 = (1 \ 1 \ 1)^T$$

$$\Sigma_1 = \Sigma_2 = \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}$$

$$\Pr(\omega_2) = 2\Pr(\omega_1)$$

- ❖ **解：**

- ✎ 根据题意，可得判别式函数计算公式如下

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) + \ln(\Pr(\omega_i))$$

## 分类器设计实例

- ✎ 每个模式类别的判别式函数为

$$g_1(\mathbf{x}) = -\frac{1}{2} \begin{bmatrix} x-0 \\ y-0 \\ z-0 \end{bmatrix}^T \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x-0 \\ y-0 \\ z-0 \end{bmatrix} + \ln \frac{1}{3}$$

$$= -2(x^2 + y^2 + z^2) + \ln \frac{1}{3}$$

$$g_2(\mathbf{x}) = -\frac{1}{2} \begin{bmatrix} x-1 \\ y-1 \\ z-1 \end{bmatrix}^T \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x-1 \\ y-1 \\ z-1 \end{bmatrix} + \ln \frac{2}{3}$$

$$= -2((x-1)^2 + (y-1)^2 + (z-1)^2) + \ln \frac{2}{3}$$



## 分类器设计实例

决策规则

$$g_1(\mathbf{x}) \underset{\omega_2}{\overset{\omega_1}{>}} g_2(\mathbf{x}) \Rightarrow$$

$$-2(x^2 + y^2 + z^2) + \ln \frac{1}{3} \underset{\omega_2}{\overset{\omega_1}{>}} -2((x-1)^2 + (y-1)^2 + (z-1)^2) + \ln \frac{2}{3}$$

$$x + y + z \underset{\omega_1}{\overset{\omega_2}{>}} \frac{6 - \ln 2}{4} \approx 1.32$$

判断测试模式  $\mathbf{x} = [0.1 \ 0.7 \ 0.8]^T$  的模式类别

$$\because 0.1 + 0.7 + 0.8 = 1.6 > 1.32$$

$$\therefore \mathbf{x} \in \omega_2$$

南开大学网络空间安全学院 计算机学院

33

## 结论

- ❖ 对一般形式的正态分布模式类别分类问题，贝叶斯分类器的判别式函数是二次的。
- ❖ 对具有相等协方差矩阵的正态分布模式类别分类问题，贝叶斯分类器的判别函数是一次的。
- ❖ 对最小马氏距离分类器，贝叶斯最优的条件是：
  - ✧ 模式类别样本满足正态分布
  - ✧ 相等的协方差矩阵
  - ✧ 相等的先验概率
- ❖ 对最小欧氏距离分类器，贝叶斯最优的条件是：
  - ✧ 模式类别样本满足正态分布
  - ✧ 相等的协方差矩阵且与单位矩阵成比例
  - ✧ 相等的先验概率
- ❖ 欧氏距离和马氏距离分类器均是线性分类器。

南开大学网络空间安全学院 计算机学院

34

## 结论

- ❖ 最受欢迎的分类器来自于**决策理论**和**简化假设**。
- ❖ 重要的是要认识到使用特定的（如欧氏距离或马氏距离）分类器则暗示着某些相应的**统计假设**。
- ❖ 在实际应用中，这些统计假设是否成立并不重要。
- ❖ 在绝大多数情况下，重要的是“**这个分类器是否真正地解决了实际问题。**”

## 人物介绍：高斯



Carl Friedrich Gauss  
1777–1855

It is said that when Gauss went to elementary school at age 7, his teacher Büttner, trying to keep the class occupied, asked the pupils to sum the integers from 1 to 100. To the teacher's amazement, Gauss arrived at the answer in a matter of moments by noting that the sum can be represented as 50 pairs (1 + 100, 2 + 99, etc.) each of which added to 101, giving the answer 5,050. It is now believed that the problem which was actually set was of the same form but somewhat harder in that the sequence had a larger starting value and a larger increment. Gauss was a German math-

## 人物介绍：高斯

ematician and scientist with a reputation for being a hard-working perfectionist. One of his many contributions was to show that least squares can be derived under the assumption of normally distributed errors. He also created an early formulation of non-Euclidean geometry (a self-consistent geometrical theory that violates the axioms of Euclid) but was reluctant to discuss it openly for fear that his reputation might suffer if it were seen that he believed in such a geometry. At one point, Gauss was asked to conduct a geodetic survey of the state of Hanover, which led to his formulation of the normal distribution, now also known as the Gaussian. After his death, a study of his diaries revealed that he had discovered several important mathematical results years or even decades before they were published by others.