

正态分布下的 贝叶斯分类器

张俊超

中南大学
航空航天学院





贝叶斯分类器-正态分布及其性质

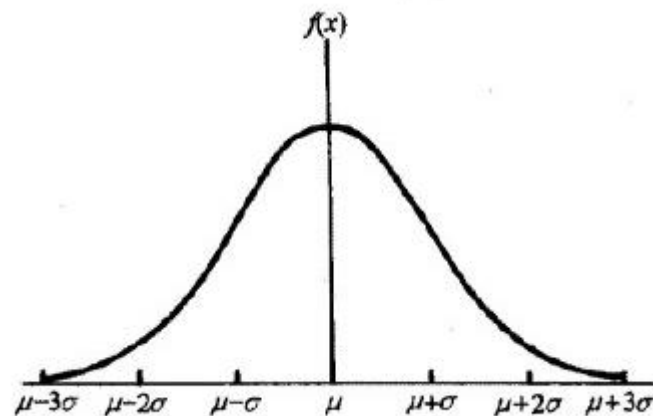
• 单变量正态分布

单变量正态分布概率密度函数：

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right\}$$

$$\mu = E\{x\} = \int_{-\infty}^{\infty} xp(x)$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)$$





贝叶斯分类器-正态分布及其性质

• 多元正态分布

多元正态分布概率密度函数：

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

标
量

$$\mathbf{x} = (x_1, x_2, \dots, x_d)^T$$

$$\boldsymbol{\mu} = E\{\mathbf{x}\} = (\mu_1, \mu_2, \dots, \mu_d)^T$$

$$\mathbf{\Sigma} = E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\} \in \mathbb{R}^{d \times d}$$

协方差矩阵 (对称非负定阵 $|\mathbf{\Sigma}| \geq 0$)

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1d} & \sigma_{2d} & \cdots & \sigma_{dd} \end{pmatrix}$$

协方差矩阵计算的是不同维度之间的协方差，而不是不同样本之间的



贝叶斯分类器-正态分布及其性质

➤ 多元正态分布的性质:

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

① 均值向量和协方差矩阵决定分布

均值向量 $\boldsymbol{\mu} = E\{\mathbf{x}\} = (\mu_1, \mu_2, \dots, \mu_d)^T$

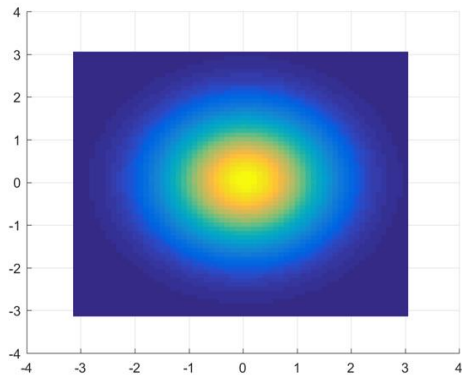
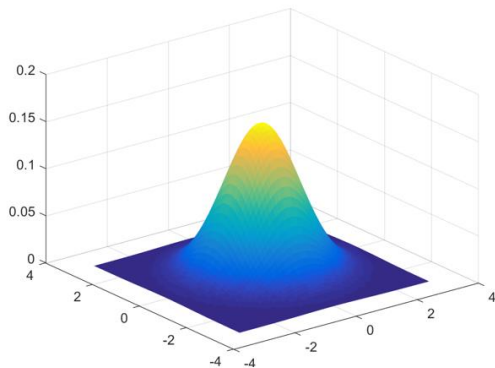
协方差矩阵 $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1d} & \sigma_{2d} & \cdots & \sigma_{dd} \end{pmatrix} \in \mathbb{R}^{d \times d}$

$$\begin{aligned} \sigma_{ij} &= E\{(x_i - \mu_i)(x_j - \mu_j)\} \\ &= E\{x_i x_j\} - E\{x_i\} E\{x_j\} \end{aligned}$$

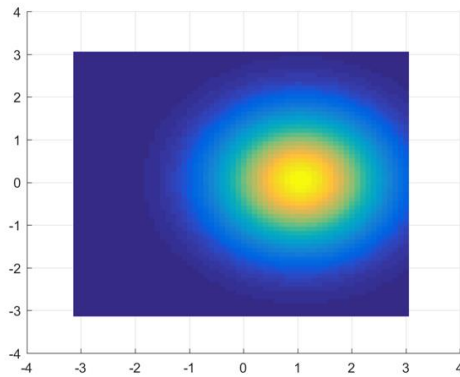
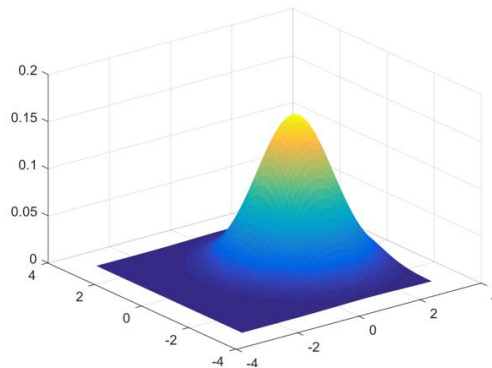
参数总数: $d(d+1)/2 + d$

贝叶斯分类器-正态分布及其性质

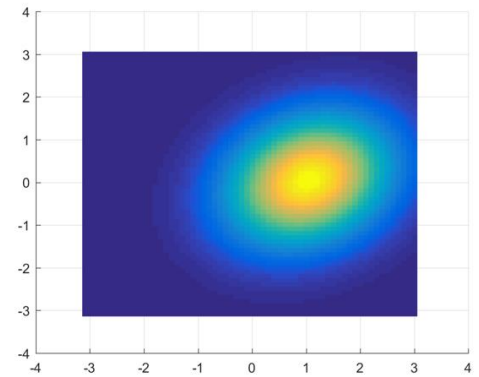
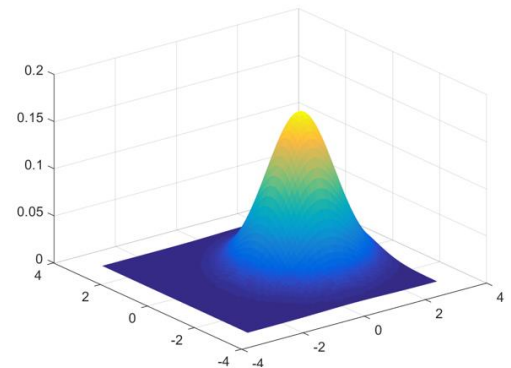
$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



$$\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



$$\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix}$$





贝叶斯分类器-正态分布及其性质

```
%% 多元正态分布
mu=[0;
    1];
sigma=[1 0.2;
        0.2 1];

%%
det_sigma = sqrt(det(sigma));
inv_sigma = inv(sigma);
Item = 1/(2*pi*det_sigma);
x=-pi:0.1:pi;
y=-pi:0.1:pi;
[X,Y]=meshgrid(x,y); % 产生网格数据并处理
p = zeros(size(X));
for i = 1:length(x)
    for j = 1:length(y)
        xx = [x(i);y(j)];
        xx = xx-mu;
        p(i,j) = Item*exp(-0.5*xx'*inv_sigma*xx);
    end
end

figure(1), surf(X,Y,p), shading flat
figure(2), surf(X,Y,p), shading flat, view(2)

%% 等高线
figure(3), contour(X,Y,p)
```




贝叶斯分类器-正态分布及其性质

② 等密度点的轨迹为一超椭球面

常数

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

指数项为常数，概率密度 $p(\mathbf{x})$ 不变，即，等密度点应满足：

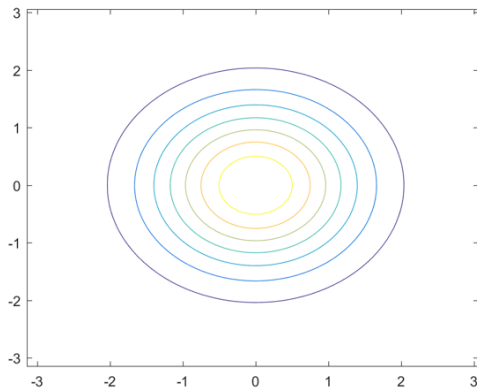
$$(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = \text{常数}$$

轨迹是一超椭球面(二维：椭圆)，中心由均值向量决定，
主轴方向由协方差矩阵的特征向量决定，
主轴的长度与协方差矩阵的特征值决定。

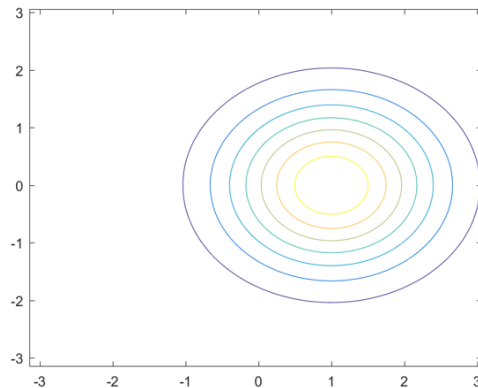


贝叶斯分类器-正态分布及其性质

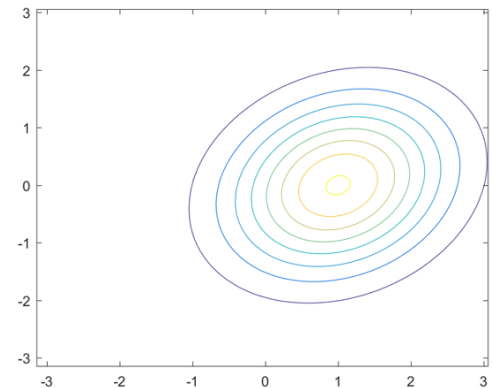
$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



$$\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



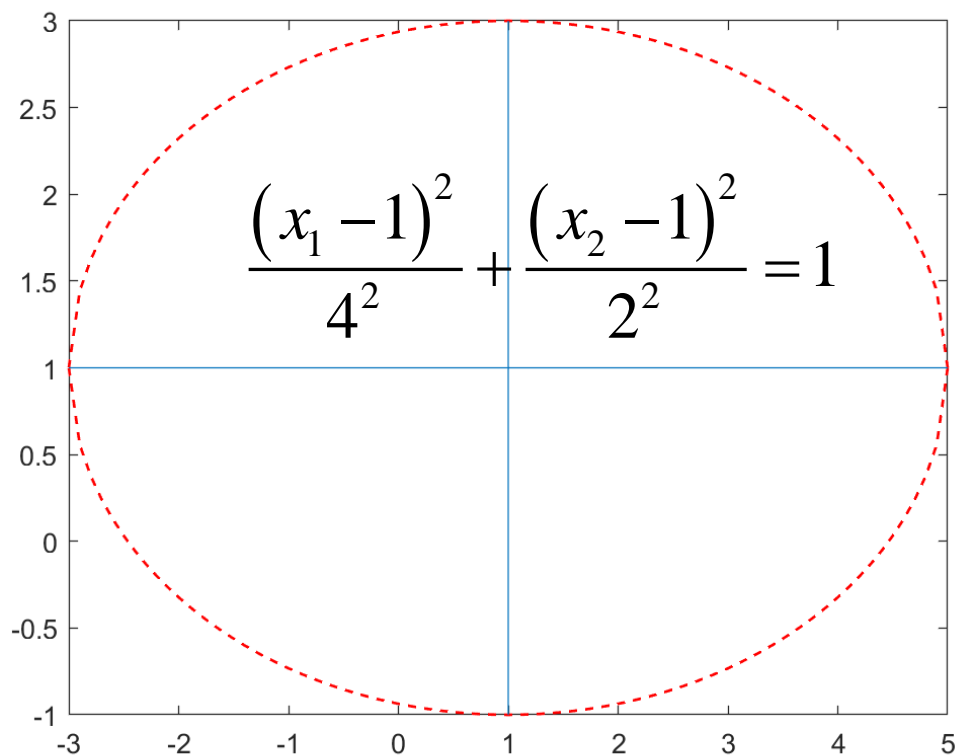
$$\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix}$$





贝叶斯分类器-正态分布及其性质

为什么主轴方向由协方差矩阵的特征向量决定，
主轴的长度与协方差矩阵的特征值决定？

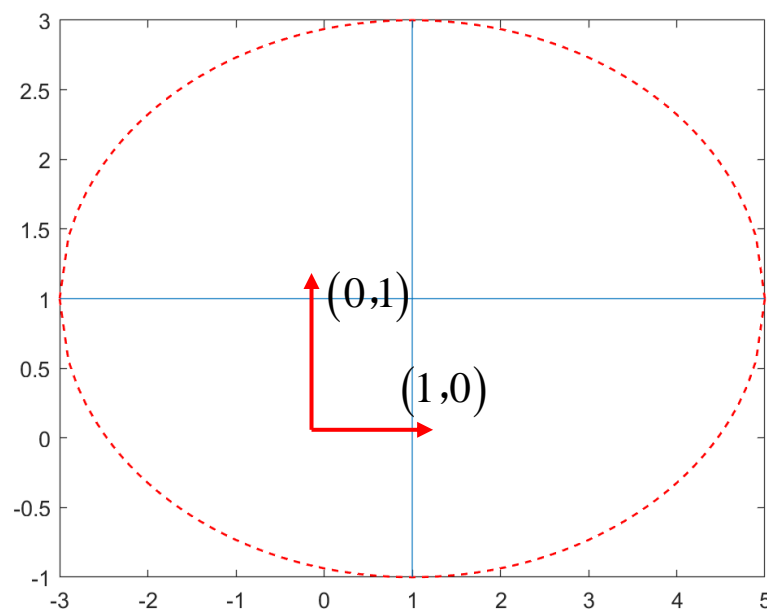




贝叶斯分类器-正态分布及其性质

$$(x_1 - 1 \quad x_2 - 1) \begin{pmatrix} \frac{1}{4^2} & 0 \\ 0 & \frac{1}{2^2} \end{pmatrix} \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} = 1$$

$\begin{pmatrix} 4^2 & 0 \\ 0 & 2^2 \end{pmatrix}$ 的特征值(16,4),
特征向量: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$



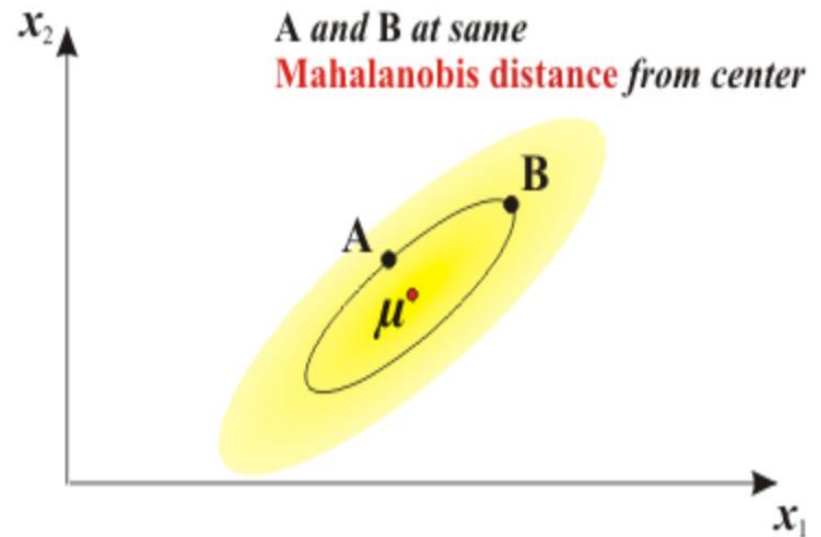
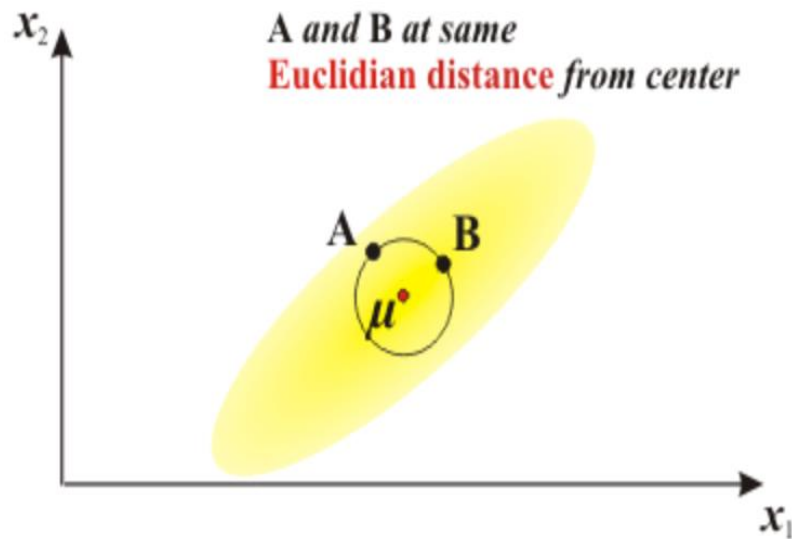
贝叶斯分类器-正态分布及其性质

- 马氏距离(Mahalanobis)
- 欧式距离

$$\gamma^2 = (\mathbf{x}_i - \mathbf{x}_j)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{x}_j)$$

$$d^2 = (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)$$

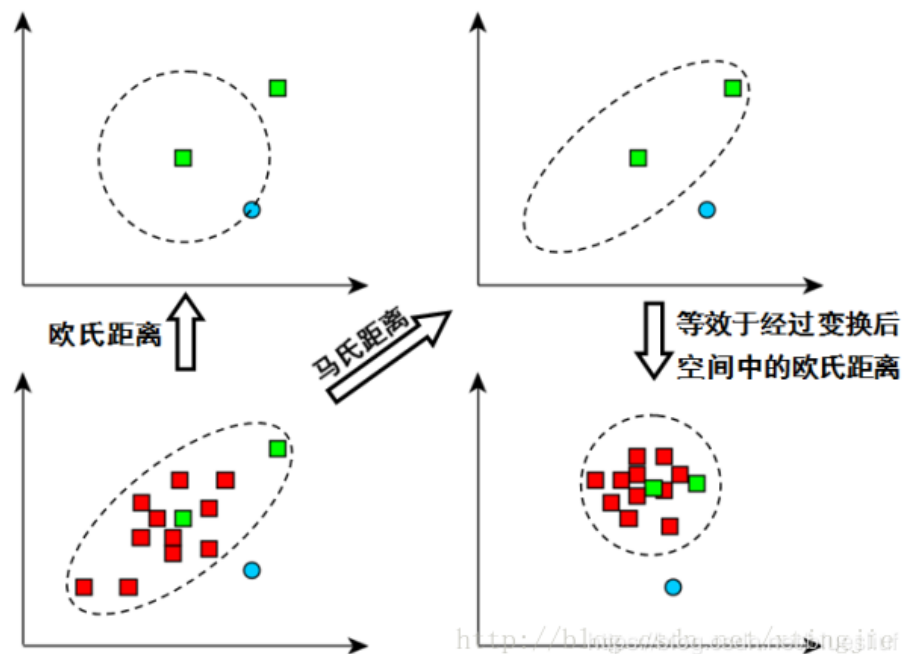
不同程度
(单位)
之间做了
归一化





贝叶斯分类器-正态分布及其性质

- 马氏距离(Mahalanobis) $\gamma^2 = (\mathbf{x}_i - \mathbf{x}_j)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{x}_j)$
- 欧式距离 $d^2 = (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)$





贝叶斯分类器-补充知识

- 自学

- 如何理解矩阵的特征值分解?

- <https://www.matongxue.com/madocs/228.html>

- 马氏距离VS欧式距离

- <https://zhuanlan.zhihu.com/p/46626607>

- 协方差矩阵

- <https://zhuanlan.zhihu.com/p/37609917>



贝叶斯分类器-正态分布及其性质

③ 不相关等价于独立

$$\text{独立: } p(x_i x_j) = p(x_i) p(x_j)$$

$$\text{不相关: } E\{x_i x_j\} = E\{x_i\} \cdot E\{x_j\}$$

※多元随机向量的各分量相互独立，则协方差矩阵是**对角阵**。

$$\Sigma = \begin{pmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{dd} \end{pmatrix}$$

④ 多元正态分布的边缘分布和条件分布仍是正态分布

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

二元正态向量 $(x_1, x_2)^T$ ，有

$$p(x_1) \sim N(\mu_1, \sigma_{11}) \quad p(x_2) \sim N(\mu_2, \sigma_{22})$$



贝叶斯分类器-正态分布及其性质

⑤ 多元正态随机向量的线性变换仍满足正态分布

$$p(\mathbf{x}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Rightarrow p(\mathbf{y} = \mathbf{A}\mathbf{x}) \sim N(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T), \mathbf{A} \text{是非奇异的}$$

存在某个 \mathbf{A} 使得 $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T$ 是对角阵 \Rightarrow \mathbf{y} 的各个分量不相关

⑥ 多元正态随机向量的线性组合，是**一维**的正态随机变量

$$p(\mathbf{x}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Rightarrow p(\mathbf{a}^T \mathbf{x}) \sim N(\mathbf{a}^T \boldsymbol{\mu}, \mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a})$$



贝叶斯分类器-正态分布及其性质

- 小结:

- 6条性质

- 正态分布下，不相关等价于独立

- 马氏距离

- 线性变换的正态性($\mathbf{A}\mathbf{\Sigma}\mathbf{A}^T$ 是对角阵)

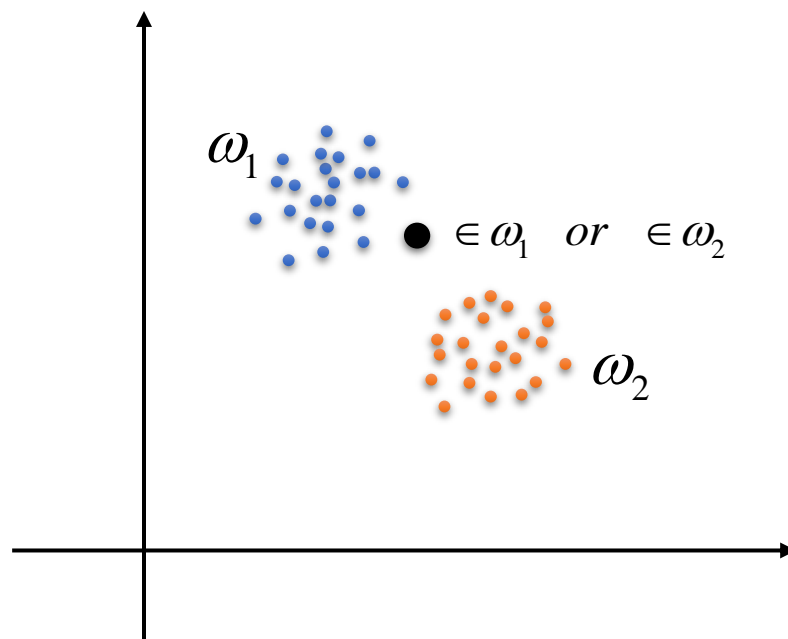
$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\mathbf{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$



贝叶斯分类器——正态分布下的最小错误率贝叶斯决策

最小错误率贝叶斯决策：

$$\text{若 } P(\omega_i | \mathbf{x}) = \max_{j=1,2,\dots,c} P(\mathbf{x} | \omega_j) P(\omega_j), \text{ 则 } \mathbf{x} \in \omega_i$$



首先，要估计类条件概率密度和先验概率。



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- 讨论类条件概率密度是正态分布

$$g_i(\mathbf{x}) = P(\mathbf{x}|\omega_i)P(\omega_i)$$

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



取对数，仍用 g 表示

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



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

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



去掉常数项

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



决策面方程: $g_i(\mathbf{x}) = g_j(\mathbf{x})$

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



$$g_j(\mathbf{x}) = \ln(P(\omega_j)) - \frac{1}{2} \ln(|\Sigma_j|) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j)$$



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

➤ 多元正态分布下的最小错误率**判别函数**为：

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

➤ 多元正态分布下的最小错误率贝叶斯**决策面方程**为：

$$g_i(\mathbf{x}) = g_j(\mathbf{x})$$

⇓

$$-\frac{1}{2} \left[(\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - (\mathbf{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j) \right] - \frac{1}{2} \ln \left(\frac{|\Sigma_i|}{|\Sigma_j|} \right) + \ln \left(\frac{P(\omega_i)}{P(\omega_j)} \right) = 0$$

接下来，讨论特殊情况下的判别函数的表达形式



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- **Case 1:** 每类的协方差矩阵都相等，且类内各特征相互独立，具有相同的方差。各类的先验概率也相等。

$$\Sigma_i = \sigma^2 \mathbf{I}, P(\omega_i) = \text{常数}, i = 1, 2, \dots, c$$

\Downarrow

$$|\Sigma_i| = \sigma^{2d}, \Sigma_i^{-1} = \frac{1}{\sigma^2} \mathbf{I}$$

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

\Downarrow

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



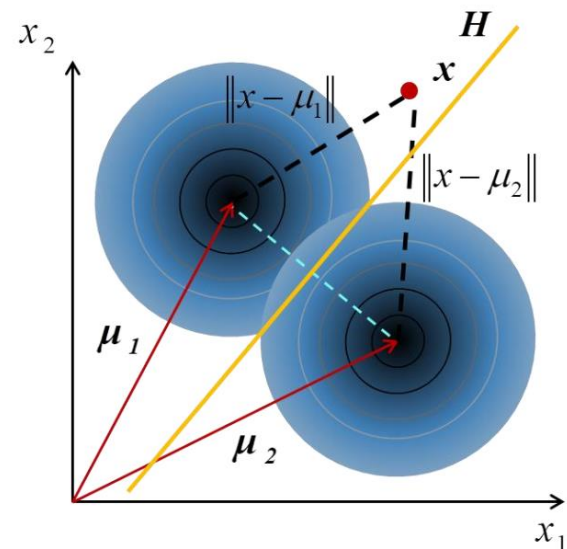
贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

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

$$g_i(\mathbf{x}) = \max_{i=1,2,\dots,c} \left\{ -\frac{1}{2\sigma^2}(\mathbf{x} - \boldsymbol{\mu}_i)^T(\mathbf{x} - \boldsymbol{\mu}_i) \right\} \Rightarrow \mathbf{x} \in \omega_i$$
$$= \min_{i=1,2,\dots,c} \left\{ (\mathbf{x} - \boldsymbol{\mu}_i)^T(\mathbf{x} - \boldsymbol{\mu}_i) \right\}$$

欧式距离下的
最小距离分类器

决策面方程呢？



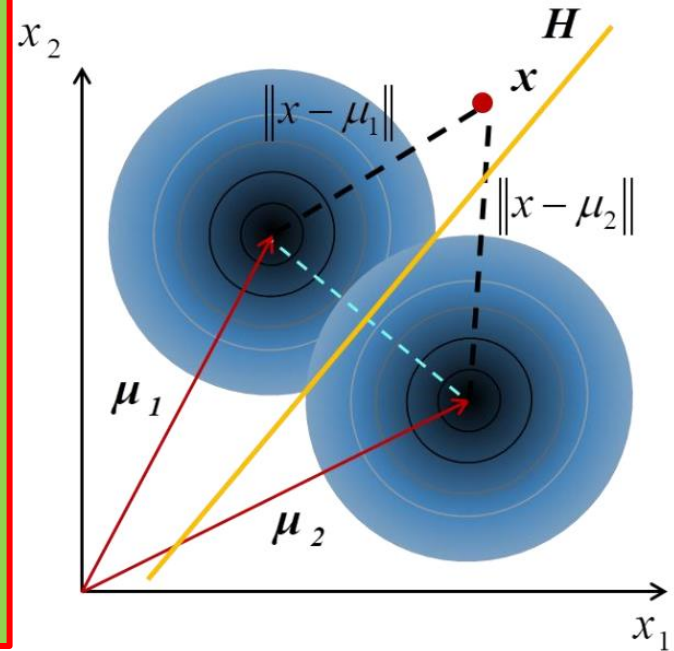
贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

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

↓ 决策面方程

$$\begin{aligned}(\mathbf{x} - \boldsymbol{\mu}_i)^T(\mathbf{x} - \boldsymbol{\mu}_i) &= (\mathbf{x} - \boldsymbol{\mu}_j)^T(\mathbf{x} - \boldsymbol{\mu}_j) \\ \mathbf{x}^T \mathbf{x} - 2\boldsymbol{\mu}_i^T \mathbf{x} + \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i &= \mathbf{x}^T \mathbf{x} - 2\boldsymbol{\mu}_j^T \mathbf{x} + \boldsymbol{\mu}_j^T \boldsymbol{\mu}_j \\ 2(\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \mathbf{x} &= \boldsymbol{\mu}_i^T \boldsymbol{\mu}_i - \boldsymbol{\mu}_j^T \boldsymbol{\mu}_j \\ 2(\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \mathbf{x} &= (\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T)(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) \\ (\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \left(\mathbf{x} - \frac{\boldsymbol{\mu}_i + \boldsymbol{\mu}_j}{2} \right) &= 0\end{aligned}$$

分界面是正交平分
均值向量的连线





贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- **Case 2:** 每类的协方差矩阵都相等，且类内各特征相互独立，具有相同的方差。

$$\Sigma_i = \sigma^2 \mathbf{I}, i = 1, 2, \dots, c$$

\Downarrow

$$|\Sigma_i| = \sigma^{2d}, \Sigma_i^{-1} = \frac{1}{\sigma^2} \mathbf{I}$$

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

\Downarrow

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



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

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

\Downarrow

线性分类器

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

\Downarrow

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

\Downarrow

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

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

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

决策面方程呢？



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

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

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



决策面方程

$$(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_i^T \boldsymbol{\mu}_i - \boldsymbol{\mu}_j^T \boldsymbol{\mu}_j) + \sigma^2 \cdot \ln \frac{P(\omega_i)}{P(\omega_j)} = 0$$

$$(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T (\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) + \frac{(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)}{\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\|^2} \cdot \sigma^2 \cdot \ln \frac{P(\omega_i)}{P(\omega_j)} = 0$$

$$(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T \left(\mathbf{x} - \frac{(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j)}{2} + \frac{\sigma^2}{\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_j\|^2} \cdot \ln \frac{P(\omega_i)}{P(\omega_j)} (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \right) = 0$$

$$\mathbf{w}^T (\mathbf{x} - \mathbf{x}_0) = 0$$

贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

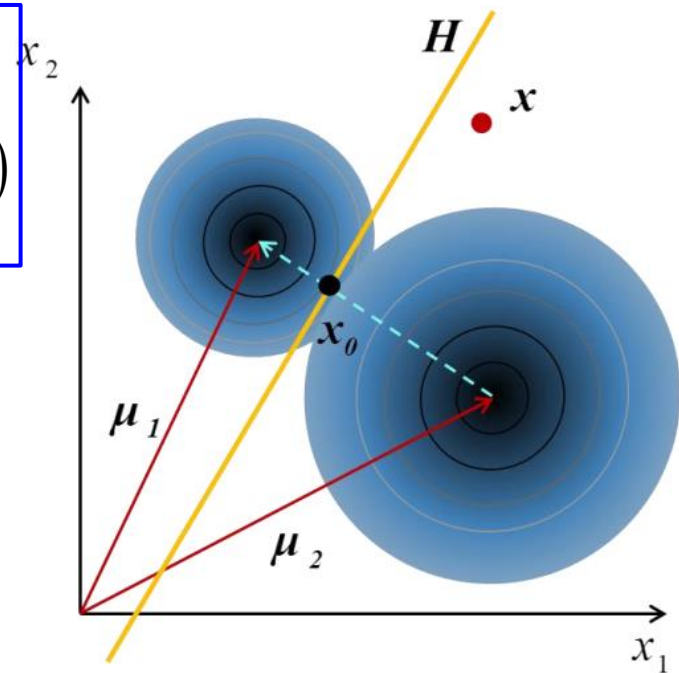
$$(\mu_i - \mu_j)^T \left(\mathbf{x} - \frac{(\mu_i + \mu_j)}{2} + \frac{\sigma^2}{\|\mu_i - \mu_j\|^2} \cdot \ln \frac{P(\omega_i)}{P(\omega_j)} (\mu_i - \mu_j) \right) = 0$$

$$\mathbf{w}^T (\mathbf{x} - \mathbf{x}_0) = 0$$

$$\mathbf{w} = \mu_i - \mu_j$$

$$\mathbf{x}_0 = \frac{(\mu_i + \mu_j)}{2} - \frac{\sigma^2}{\|\mu_i - \mu_j\|^2} \cdot \ln \frac{P(\omega_i)}{P(\omega_j)} (\mu_i - \mu_j)$$

分界面是正交均值向量的连线，向先验概率小的方向偏移





贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- **Case 3:** 每类的协方差矩阵都相等，各类的先验概率也相等。

$$\Sigma_i = \Sigma, P(\omega_i) = \text{常数}, i = 1, 2, \dots, c$$

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

\Downarrow

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

\Downarrow

马式距离下的
最小距离分类器

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



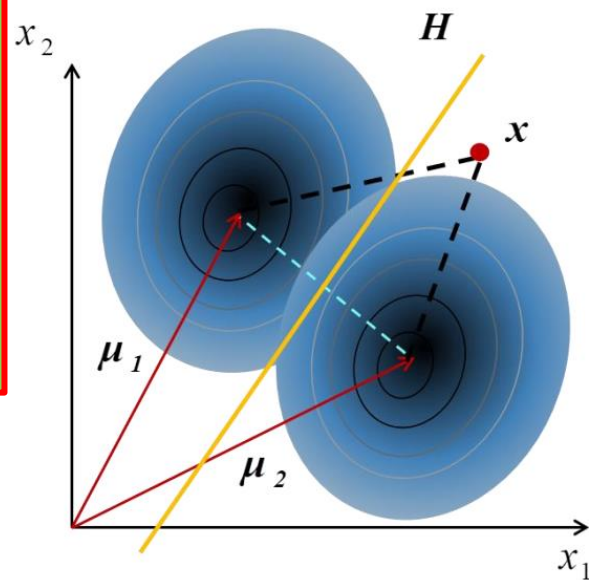
贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

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

↓ 决策面方程

分界面通过均值
向量连线的中点

$$\begin{aligned}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) &= (\mathbf{x} - \boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_j) \\ \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - 2\boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i &= \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - 2\boldsymbol{\mu}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_j^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_j \\ 2(\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \boldsymbol{\Sigma}^{-1} \mathbf{x} &= \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_i - \boldsymbol{\mu}_j^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_j \\ 2(\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \boldsymbol{\Sigma}^{-1} \mathbf{x} &= (\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) \\ (\boldsymbol{\mu}_i^T - \boldsymbol{\mu}_j^T) \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \frac{\boldsymbol{\mu}_i + \boldsymbol{\mu}_j}{2} \right) &= 0\end{aligned}$$





贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- **Case 4:** 每类的协方差矩阵都相等。

$$\Sigma_i = \Sigma, i = 1, 2, \dots, c$$

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

线性分类器

\Downarrow

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

\Downarrow

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

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

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

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

\Downarrow

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



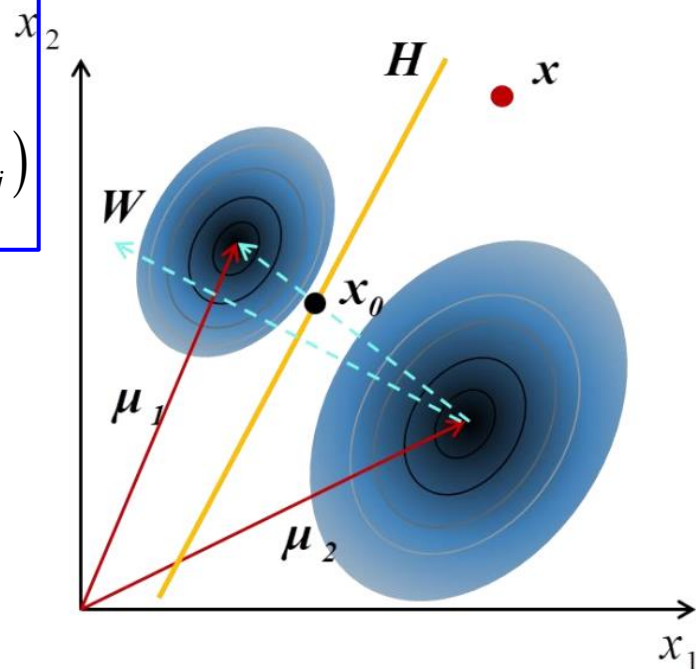
贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

$$\left(\mu_i - \mu_j \right)^T \Sigma^{-1} \left(\mathbf{x} - \frac{(\mu_i + \mu_j)}{2} + \frac{1}{(\mathbf{x} - \mu_i)^T \Sigma^{-1} (\mathbf{x} - \mu_i)} \cdot \ln \frac{P(\omega_i)}{P(\omega_j)} (\mu_i - \mu_j) \right) = 0$$
$$\mathbf{w}^T (\mathbf{x} - \mathbf{x}_0) = 0$$

$$\mathbf{w} = \Sigma^{-1} (\mu_i - \mu_j)$$

$$\mathbf{x}_0 = \frac{(\mu_i + \mu_j)}{2} - \frac{\ln \frac{P(\omega_i)}{P(\omega_j)}}{(\mathbf{x} - \mu_i)^T \Sigma^{-1} (\mathbf{x} - \mu_i)} \cdot (\mu_i - \mu_j)$$

分界面不通过均值向量连线
的中点，向先验概率小的方向
偏移





贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- **Case 5:** 每类的协方差矩阵不相等。

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

\Downarrow

非线性分类器

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

$$\mathbf{W}_i = -\frac{1}{2} \Sigma_i^{-1} \quad \mathbf{w}_i = \Sigma_i^{-1} \boldsymbol{\mu}_i$$

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

决策面可能为超球面、超椭球面、超抛物面、超双曲面或超平面



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- 小结:

➤ 多元正态分布下的最小错误率**判别函数**为:

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

➤ 多元正态分布下的最小错误率贝叶斯**决策面方程**为:

$$g_i(\mathbf{x}) = g_j(\mathbf{x})$$

⇓

$$-\frac{1}{2} \left[(\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - (\mathbf{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j) \right] - \frac{1}{2} \ln \left(\frac{|\Sigma_i|}{|\Sigma_j|} \right) + \ln \left(\frac{P(\omega_i)}{P(\omega_j)} \right) = 0$$



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- 小结:

Case1: $\Sigma_i = \sigma^2 \mathbf{I}, P(\omega_i) = \text{常数}, i = 1, 2, \dots, c$

分界面是正交平分
均值向量的连线

Case2: $\Sigma_i = \sigma^2 \mathbf{I}, i = 1, 2, \dots, c$

分界面是正交均值向
量的连线，向先验概
率小的方向偏移

Case3: $\Sigma_i = \Sigma, P(\omega_i) = \text{常数}, i = 1, 2, \dots, c$

Case4: $\Sigma_i = \Sigma, i = 1, 2, \dots, c$

Case5: $\Sigma_i \neq \Sigma_j, i, j = 1, 2, \dots, c, j \neq i$

分界面通过均值
向量连线的中点

分界面不通过均值向
量连线的中点，向先
验概率小的方向偏移

非线性分类问题



贝叶斯分类器 - 正态分布下的最小错误率贝叶斯决策

- 正态分布下的贝叶斯分类器(步骤)

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

根据训练样本(有标签)
估计先验概率

根据训练样本(有标签)
估计均值向量和协方差矩阵

待识别样本代入判别
函数, 根据最大后验
概率决策进行分类



模式识别-贝叶斯分类器

若干素材取自网络，特此致谢！





模式识别-贝叶斯分类器

谢谢聆听！

