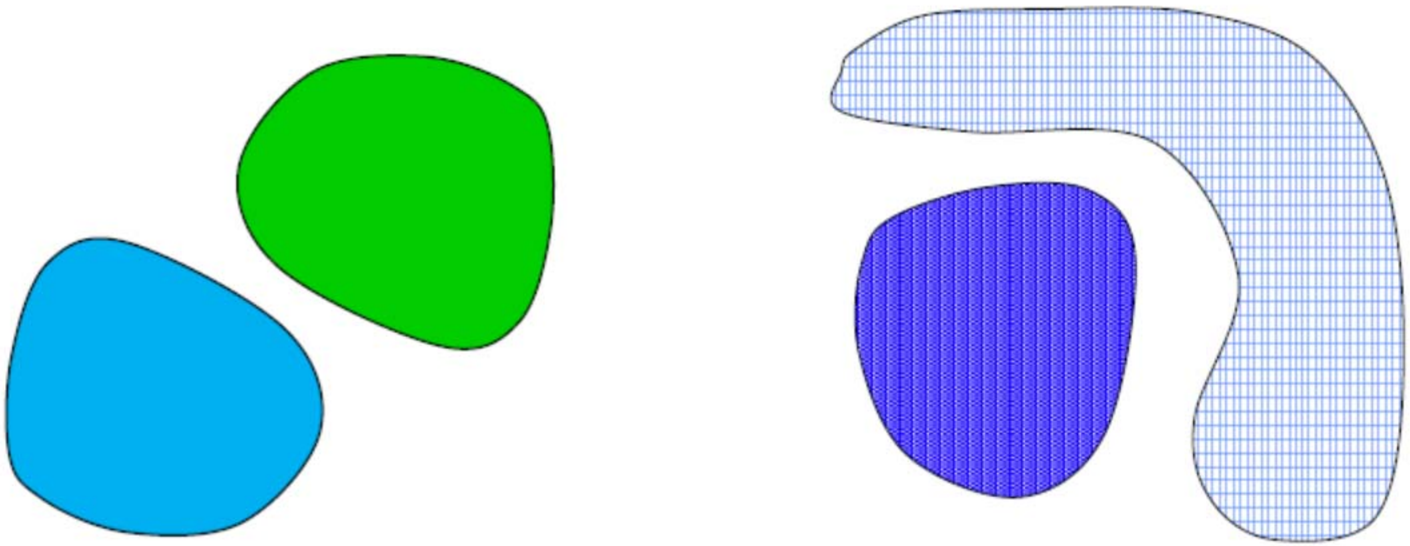
第四章： 线性模型-非线性可分情况下的求解方法

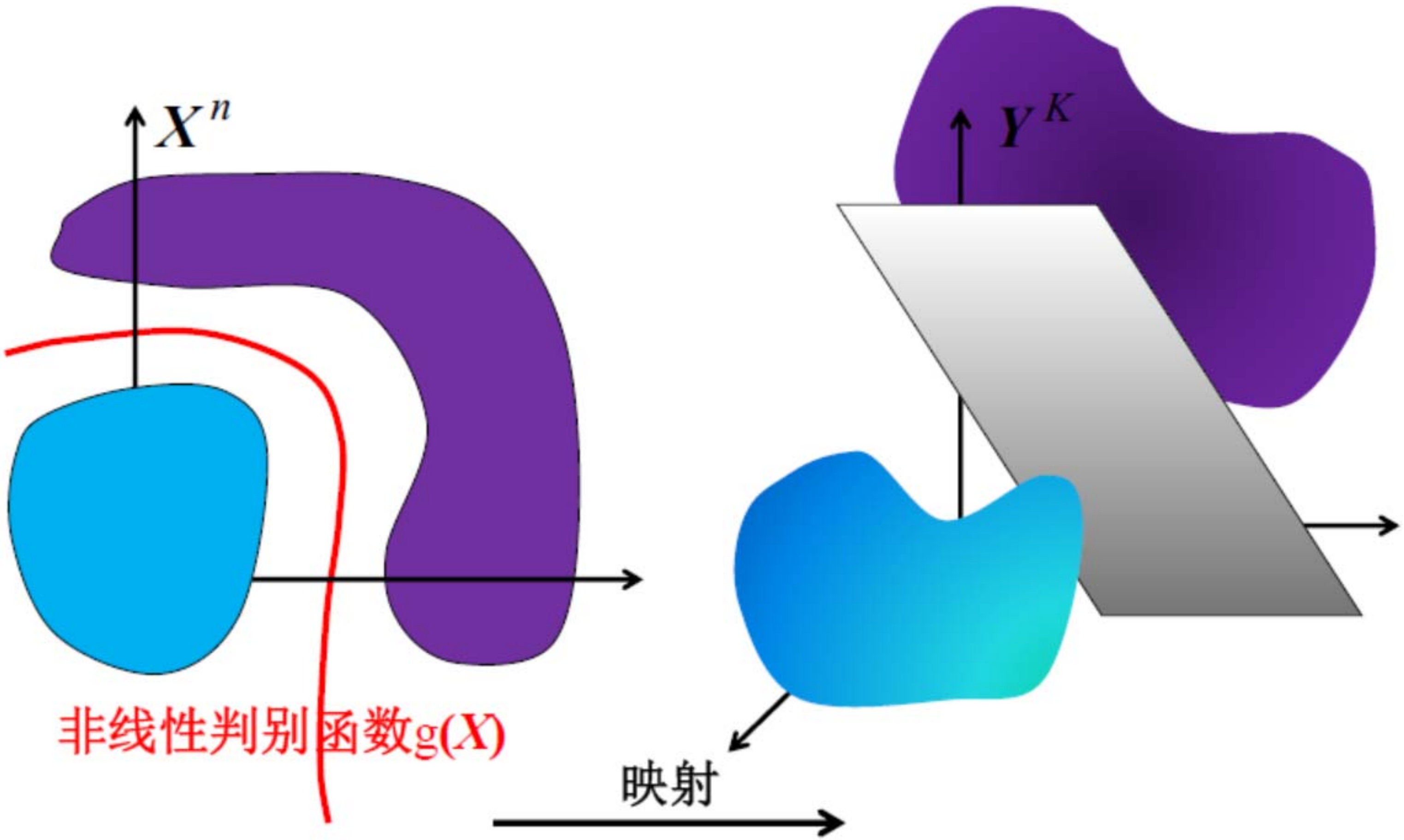
非线性可分情况下的求解方法

- 线性可分
- 非线性可分
  - 广义线性判别函数法
  - 分段线性判别函数近似法
  - 非线性判别函数法



本节PPT主要取自中国科大汪增福教授《模式识别》课件

广义线性判别函数法



广义线性判别函数法

$$X^n \xrightarrow{\text{映射}} Y^K$$

$$X=(x_1,x_2,\ldots,x_n)^T \qquad Y=(y_1,y_2,\ldots,y_K)^T$$

若$g(X)$由形式已知的有限个非线性函数所构成，则

$$g(X)=w_1 f_1(X)+w_2 f_2(X)+\Lambda+w_K f_K(X)+w_{K+1}$$

$f_i(X), i=1,2,\ldots K$　　特征空间中的模式向量X的单值实函数

$w_i, i=1,2,\ldots,K+1$　　加权系数

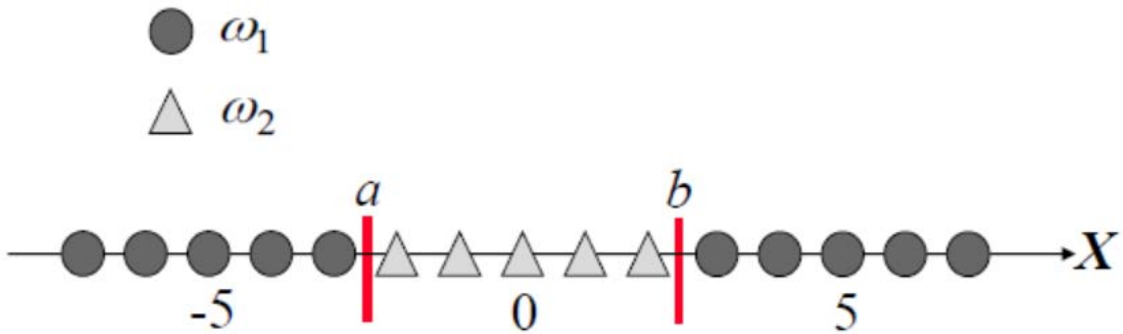
引入变换　　$y_i=f_i(X),\qquad i=1,2,\ldots K$

$$Y=(y_1,y_2,\ldots,y_K)^T \Rightarrow Y^K$$

$$g(Y)=w_1 y_1+w_2 y_2+\Lambda+w_K y_K+w_{K+1}$$　　线性判别函数

施行非线性变换的过程往往伴随着特征空间维数的增加！

广义线性判别函数法



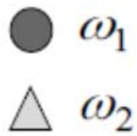
决策面函数

$$g(x) = (x-a)(x-b) = \boxed{x^2} - (a+b)\boxed{x} + ab$$

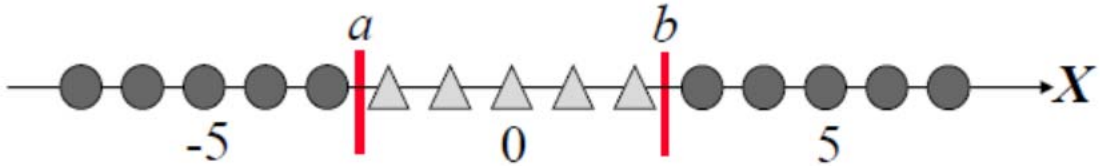
判决规则
$$\begin{cases} 若 x < a 或 x > b, & 则 x \in \omega_1 \\ 若 a < x < b, & 则 x \in \omega_2 \end{cases}$$

$$\begin{cases} g(x) > 0 \\ g(x) < 0 \end{cases} \Rightarrow \begin{cases} x \in \omega_1 \\ x \in \omega_2 \end{cases}$$

广义线性判别函数法



$$g(x)=(x-a)(x-b)=\boxed{x^2}-(a+b)\boxed{x}+ab$$

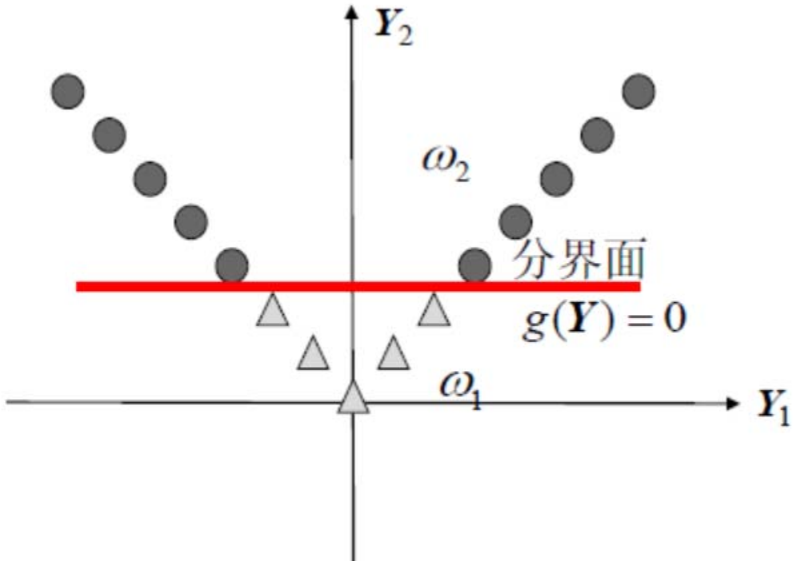


定义变换

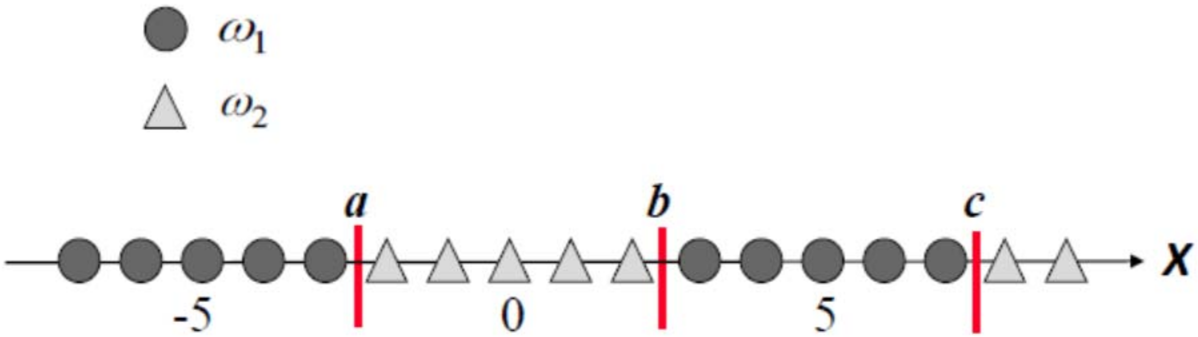
$$\begin{cases} y_1 = x^2 \\ y_2 = x \end{cases}$$

决策面函数

$$g(Y)=w_1y_1+w_2y_2+w_3$$



广义线性判别函数法



$$g(x)=(x-a)(x-b)(x-c)=x^3-(a+b+c)x^2+(ab+ac+bc)x-abc$$

$$y_1=x^3,\quad y_2=x^2,\quad y_3=x$$

$$g(\boldsymbol{Y})=w_1y_1+w_2y_2+w_3y_3+w_4$$

维数灾

如何确定非线性变换？

- 先验知识
- 函数逼近

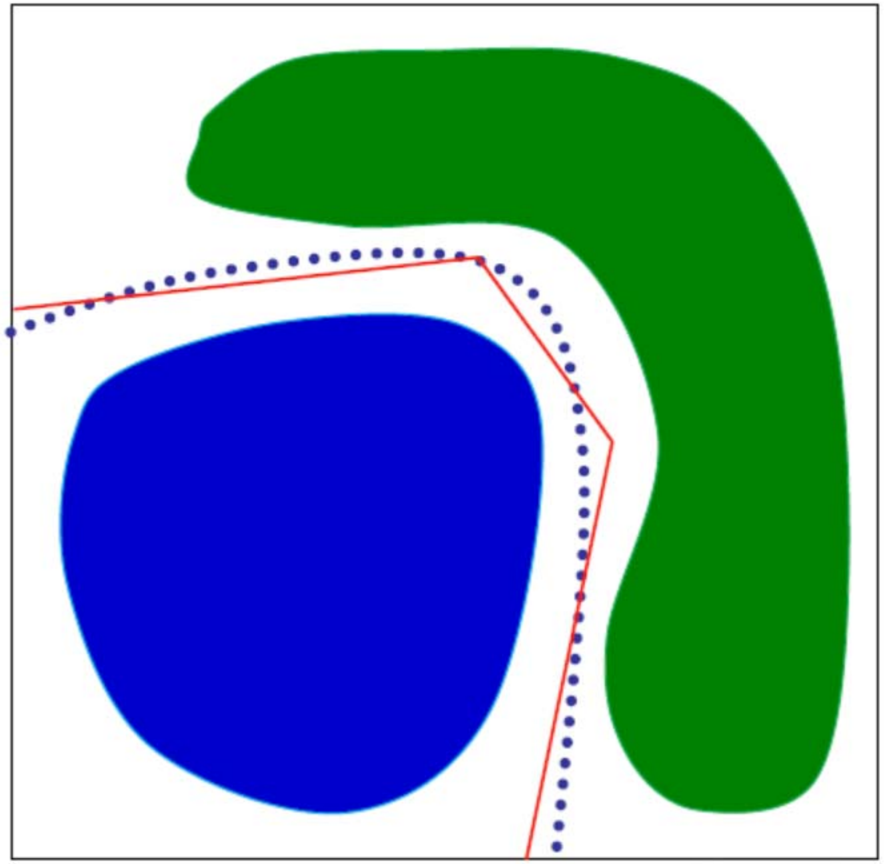
如果待求非线性决策面函数可用一个连续函数表示，则根据函数逼近理论，只要多项式次数取得足够高，总可以在预定容限范围内任意地逼近该函数。

分段线性判别函数法

避免“维数灾”问题的方策1：分段线性逼近

问题:
在二维情况下，如何
用折线近似曲线？

推而广之:
在多维情况下，如何
用超平面近似超曲面？

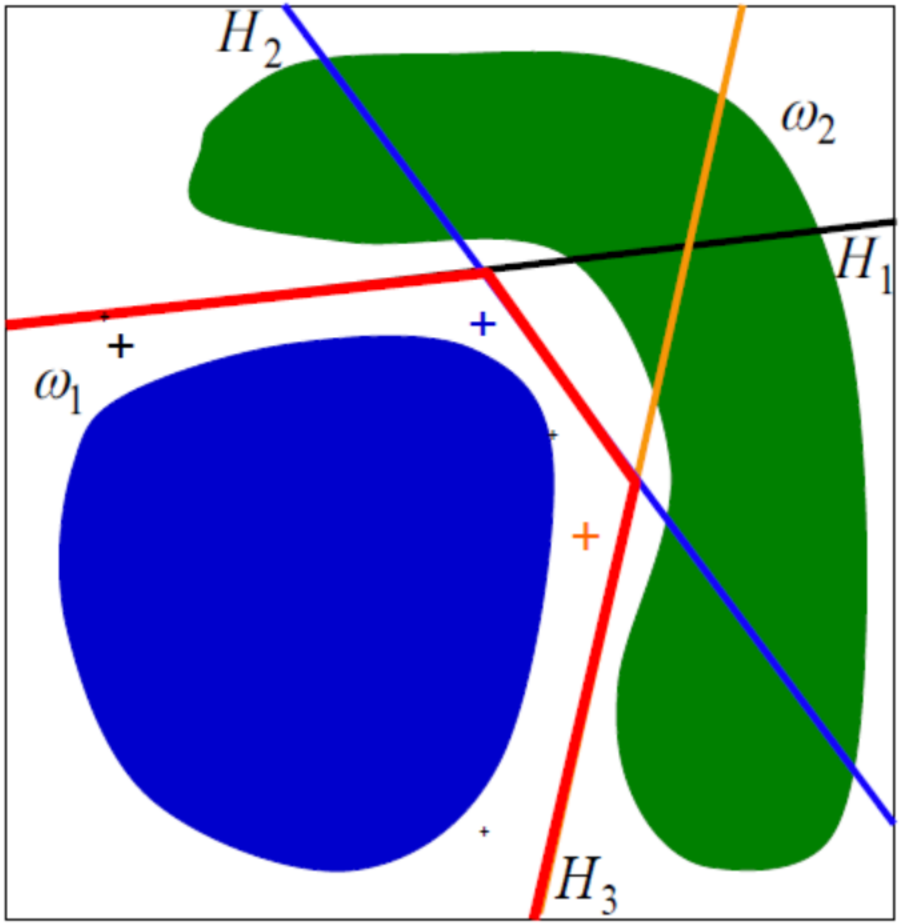


分段线性判别函数法：单超平面法

起始时整个特征空间为一个区域。

对每一个包含两个不同类别样本的区域递归地使用超平面进行再分类，直到每一个区域只包含同一类别的样本为止。

对得到的所有分类区域进行综合形成最终的分段超平面。

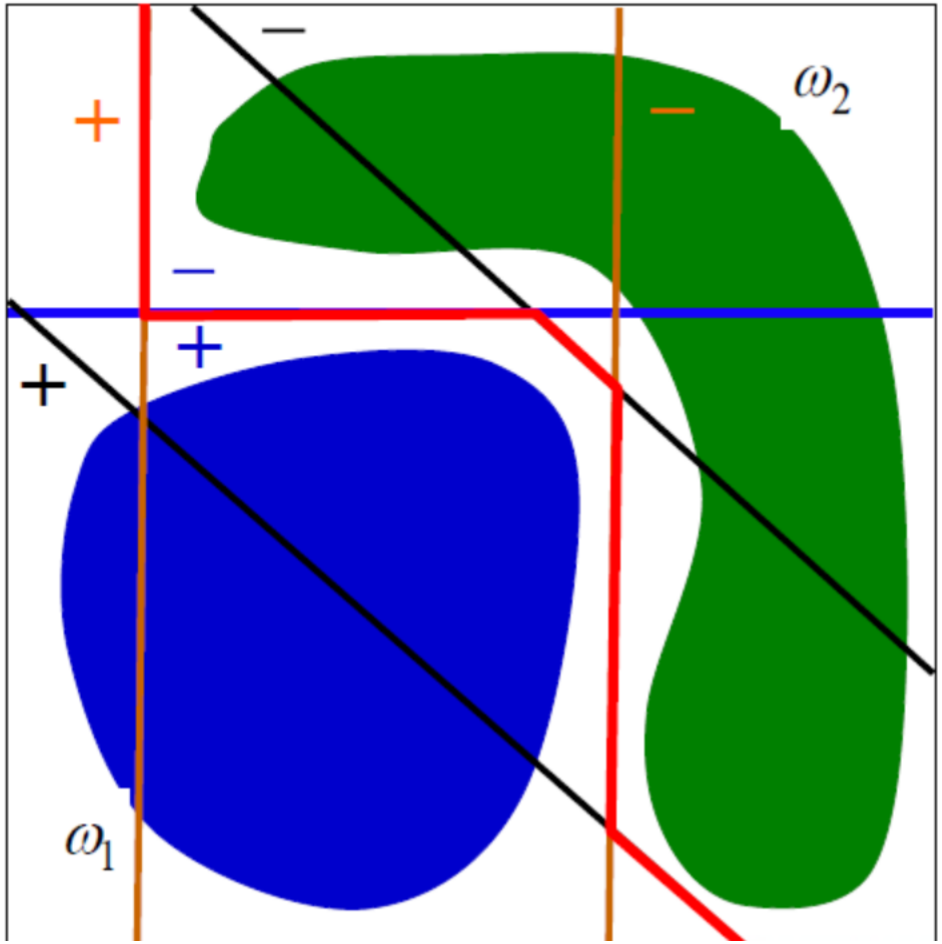


分段线性判别函数法：双超平面法

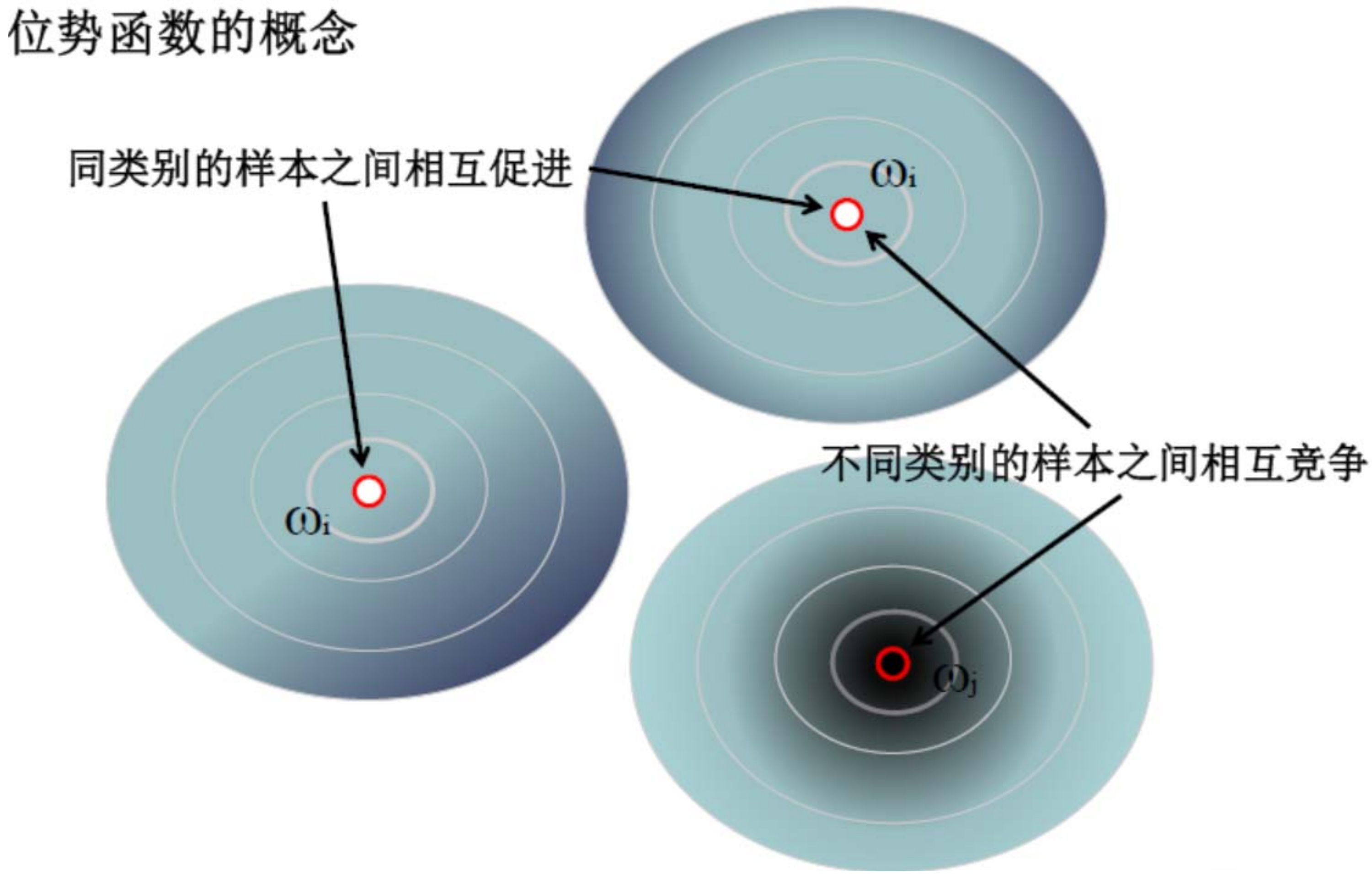
起始时整个特征空间为一个区域。

对每一个包含两个不同类别样本的区域递归地使用相互平行的两个超平面进行再分类，直到两个超平面所夹内侧区域区域只包含同一类别的样本为止。

对得到的所有分类区域进行综合形成最终的分段超平面。



非线性判别函数法：位势函数法



非线性判别函数法：位势函数法

位势函数的概念

$K(X, X_n)$ → 样本点

样本点X_n在X处引起的电位值

位势函数的条件

1. $K(X, X_n) = K(X_n, X)$
2. 当且仅当$X = X_n$时，$K(X,X_n)$达到极值。
3. 当X 和X_n之间的距离$\to \infty$时，$K(X,X_n) \to 0$。
4. $K(X,X_n)$是X 和X_n之间距离的单调和光滑函数。

非线性判别函数法：位势函数法

位势函数举例

1. $K(X, X_n) = \sum_{i=1}^{m} \phi_i(X)\phi_i(X_n)$ 完备正交多项式

2. $K(X, X_n) = \exp\{-\boxed{\alpha}\|X - X_n\|^2\}$

3. $K(X, X_n) = \dfrac{1}{1+\boxed{\alpha}\|X - X_n\|^2}$

4. $K(X, X_n) = \left|\dfrac{\sin \boxed{\alpha}\|X - X_n\|^2}{\alpha\|X - X_n\|^2}\right|$

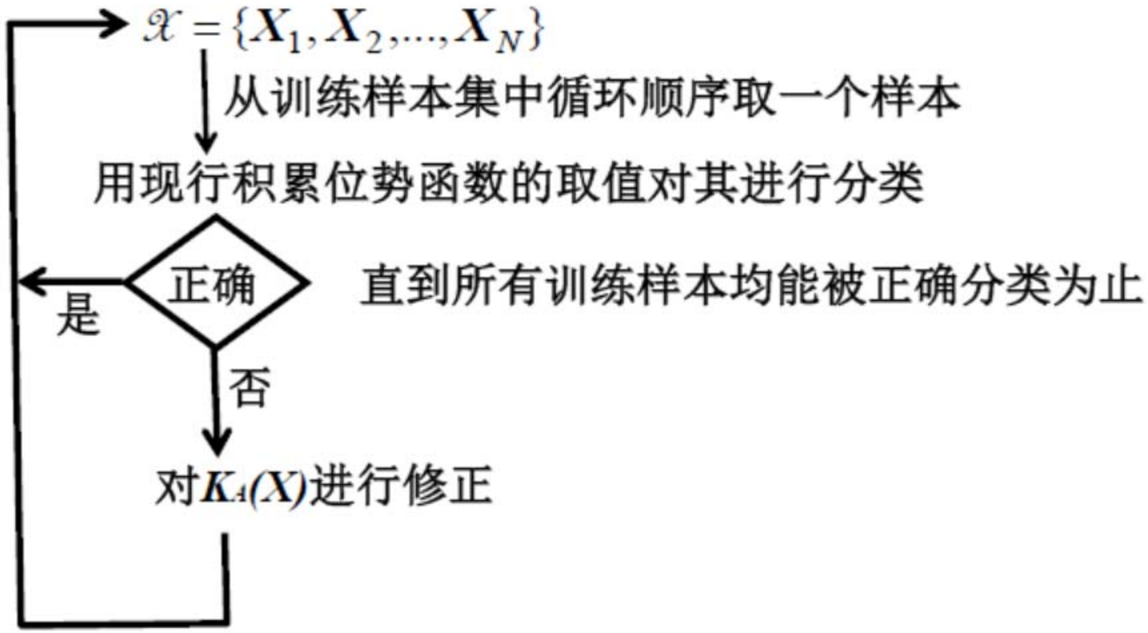
α 用于控制位势函数衰减速度的常数

非线性判别函数法：位势函数法

积累位势函数$K_A(X)$ $\longrightarrow$ 判别函数

由相关样本点在任一点X处产生的位势总和。

注意：不是各样本点位势的简单叠加，而是一种加权和。



非线性判别函数法：位势函数法

积累位势函数$K_A(X)$的修正规则

$$\mathscr{X} = \{X_1, X_2, \ldots, X_N\} \begin{cases} \omega_i \\ \omega_j \end{cases}$$

$\downarrow$

$K_{A,k}(X)$，使对所有训练样本正确分类

第 k步积累位势函数

$$K_{A,k}(X) \begin{cases} > 0 \\ = 0 \\ < 0 \end{cases}$$

正确分类

$$\begin{cases} X \in \omega_i \text{时}, K_{A,k}(X) > 0 \\ X \in \omega_j \text{时}, K_{A,k}(X) < 0 \end{cases}$$

错误分类

$$\begin{cases} X \in \omega_i \text{时}, K_{A,k}(X) \le 0 \\ X \in \omega_j \text{时}, K_{A,k}(X) \ge 0 \end{cases}$$

非线性判别函数法：位势函数法

积累位势函数$K_A(X)$的修正规则

$$K_{A,0}(\boldsymbol{X}) = 0$$

$$K_{A,k}(\boldsymbol{X}) = K_{A,k-1}(\boldsymbol{X}) + \gamma_k K(\boldsymbol{X}, \boldsymbol{X}_k)$$

$$\gamma_k = \begin{cases} 0 & \text{若 } \boldsymbol{X}_k \in \omega_i\text{，且} K_{A,k-1}(\boldsymbol{X}_k) > 0 \\ 0 & \text{若 } \boldsymbol{X}_k \in \omega_j\text{，且} K_{A,k-1}(\boldsymbol{X}_k) < 0 \\ 1 & \text{若 } \boldsymbol{X}_k \in \omega_i\text{，且} K_{A,k-1}(\boldsymbol{X}_k) \le 0 \\ -1 & \text{若 } \boldsymbol{X}_k \in \omega_j\text{，且} K_{A,k-1}(\boldsymbol{X}_k) \ge 0 \end{cases}$$

非线性判别函数法：位势函数法

Step1. 赋初值：$k=0$，$K_{A,0}(X)=0$，$N_c=0$；

Step2. 读入训练样本集合 $\mathscr{X}=\{X_0, X_1, \ldots, X_{N-1}\}$；

Step3. 取样本$X'=X_{[k]_N}$ $[k]_N=k \mod(N)$；

Step4. 根据修正规则，更新积累位势函数：

若 $X'\in\omega_i$，且$K_{A,k}(X')>0$，则$K_{A,k+1}(X)=K_{A,k}(X)$，$N_c+=1$;

若 $X'\in\omega_i$，且$K_{A,k}(X')\le 0$，则$K_{A,k+1}(X)=K_{A,k}(X)+K(X,X')$，$N_c=0$;

若 $X'\in\omega_j$，且$K_{A,k}(X')<0$，则$K_{A,k+1}(X)=K_{A,k}(X)$，$N_c+=1$;

若 $X'\in\omega_j$，且$K_{A,k}(X')\ge 0$，则$K_{A,k+1}(X)=K_{A,k}(X)-K(X,X')$，$N_c=0$;

Step5. 若 $N_c\ge N$，输出 $K_{A,k+1}(X)$，算法结束；

否则 $k=k+1$，返回step3。

非线性判别函数法：位势函数法

小结

- 分类能力很强
- 存储量和计算量大
- 分界面可能相当复杂