



第三章 判别域代数界面方程法

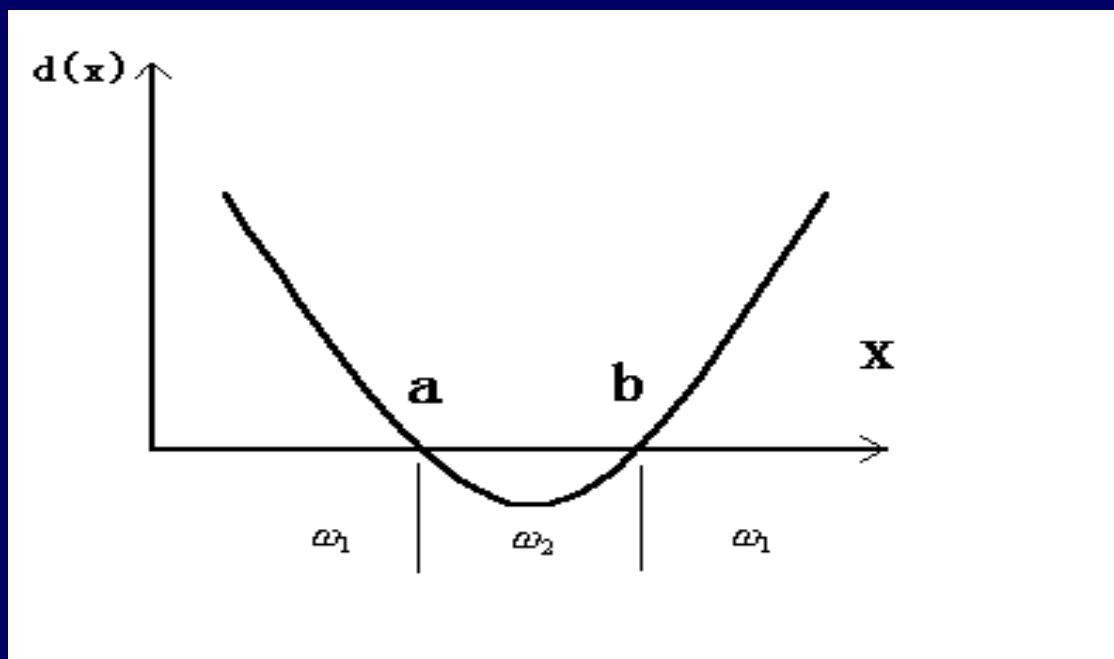
- 3.1 判别域界面方程分类的概念
- 3.2 线性判别函数
- 3.3 判别函数值的鉴别意义、权空间及解空间
- 3.4 fisher线性判别
- 3.5 线性可分条件下判别函数权矢量算法
- 3.6 一般情况下的判别函数权矢量算法
- 3.7 非线性判别函数
- 3.8 最近邻方法



3.7 非线性判别函数

3.7.1 广义线性判别函数

设一维两类模式 x 在一维空间——坐标轴上分布下图所示，两类的类域为 $\Omega_1 : (-\infty, a) \cup (b, \infty)$ 和 $\Omega_2 : (a, b)$



一维特征空间中非线性可分的图示

模式识别



3.7 非线性判别函数

3.7.1 广义线性判别函数

显然，上述的两类不是线性可分的，它不能用一个分界点将两类分开。但是，对于一条穿过 a 、 b 两点的二次曲线，

$$d(x) = (x-a)(x-b) = x^2 - (a+b)x + ab$$

当 $x \in \omega_1$ 即 $x < a$ 或 $x > b$ 时, $d(x) > 0$

当 $x \in \omega_2$ 即 $a < x < b$ 时, $d(x) < 0$



3.7 非线性判别函数

3.7.1 广义线性判别函数

如果作非线性变换：
$$\begin{cases} y_1 = x^2 \\ y_2 = x \end{cases}$$

则原来的一维特征空间映射为二维特征空间，二次判别函数 $d(x)$ 变换后在二维特征空间中映射成一次判别函数：
$$d(\vec{y}) = y_1 - (a + b)y_2 + ab$$

原来的一维非线性可分的模式在所映射的二维特征空间中是线性可分的，即：

$$\begin{cases} \vec{y} \in \omega_1 & d(\vec{y}) > 0 \\ \vec{y} \in \omega_2 & d(\vec{y}) < 0 \end{cases}$$

将上述方法一般化:

设 n 维模式特征集 $\{\vec{x}_i\}$ 在特征空间 X^n 中是非线性可分的, 对各模式作非线性变换

$$T: X^n \Rightarrow Y^d \quad d > n$$

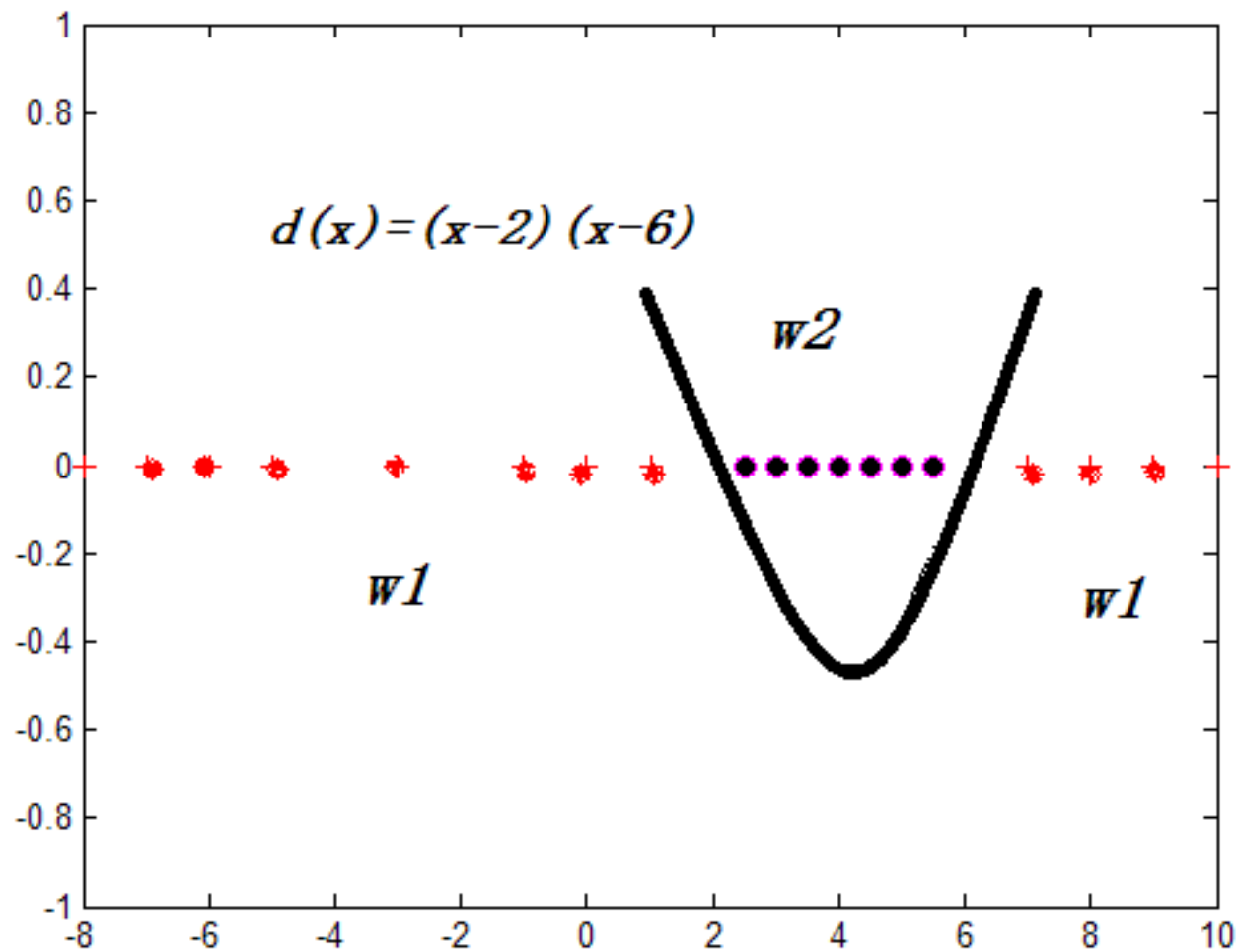
$$\vec{x} = (x_1, x_2, \dots, x_n)' \quad \vec{y} = (y_1, y_2, \dots, y_d)'$$

其中 $y_i = f_i(\vec{x})$ 是 \vec{x} 的单值函数, 选取适当的函数

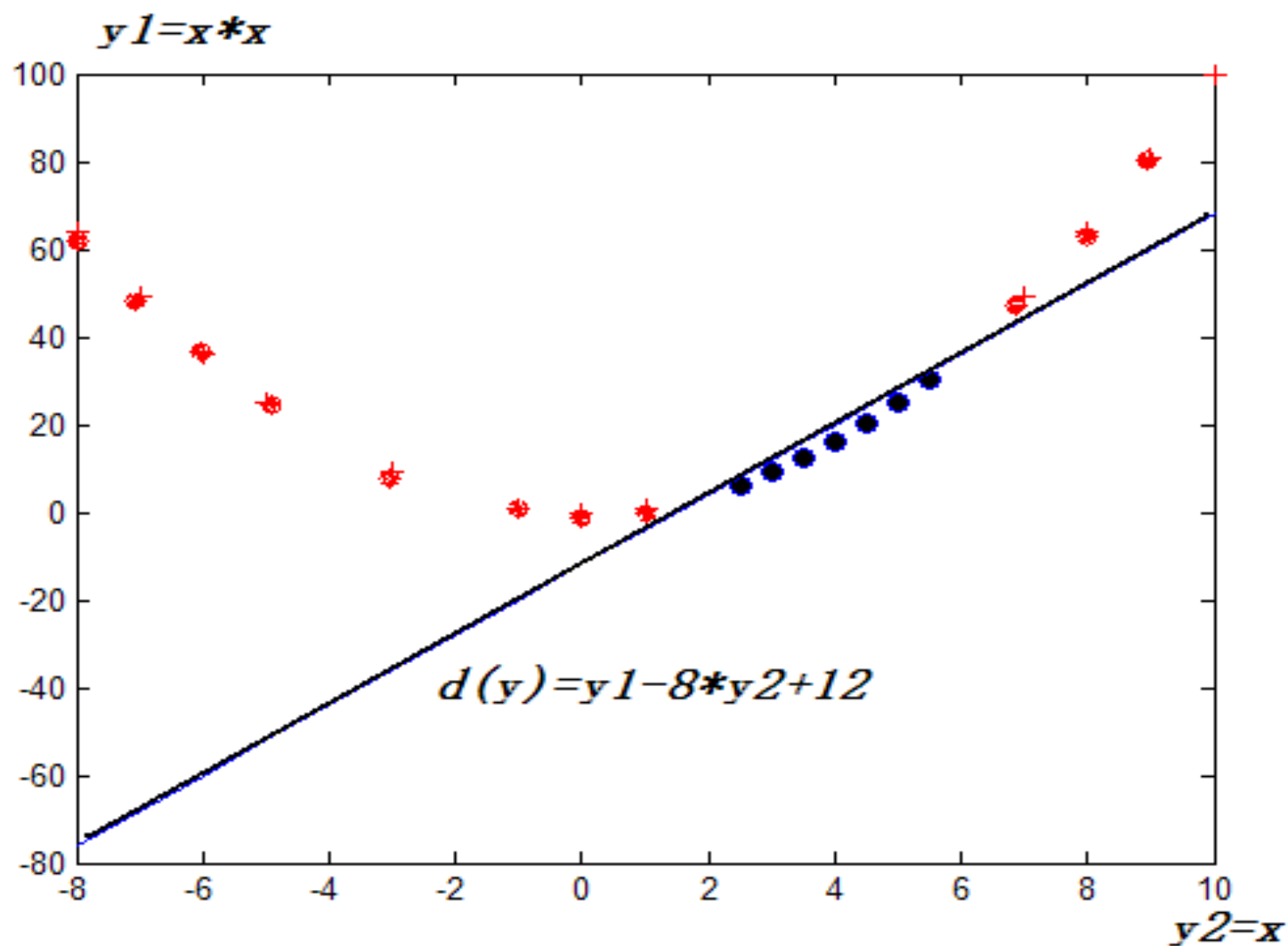
$$f_i(\vec{x}) \quad (i=1, 2, \dots, d)$$

使 $\{\vec{y}_j = (f_1(\vec{x}_j), f_2(\vec{x}_j), \dots, f_d(\vec{x}_j))'\}$ 在特征空间是线性可分的, 称其为广义线性判别函数。

下图所示两类模式是线性不可分的。



经过非线性变换，两类模式是线性可分的。





3.7 非线性判别函数

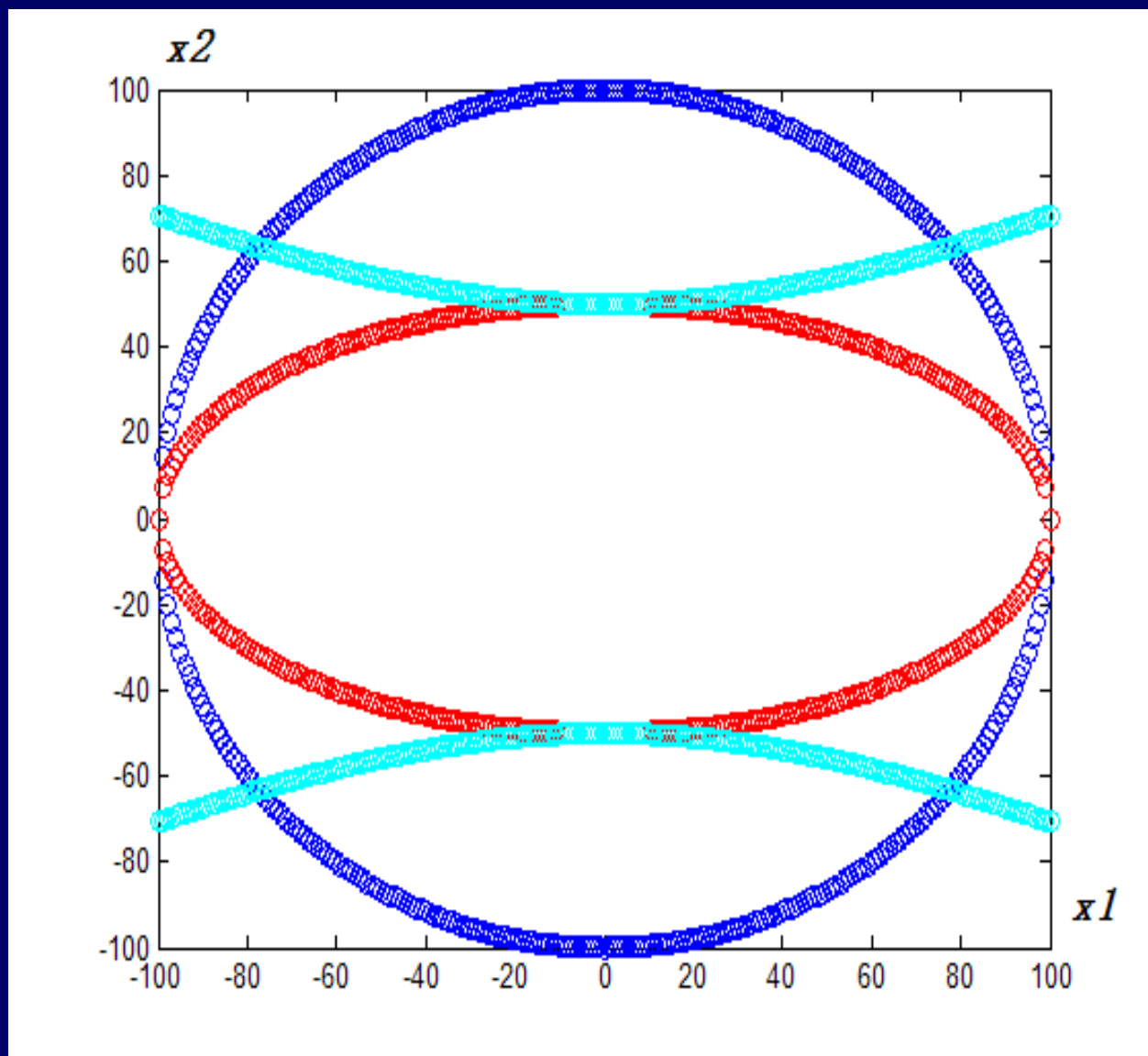
3.7.2 二次判别函数

二次判别函数是一种常用的非线性判别函数，函数构造比较简单，适用面比线性判别函数要广。

在 n 维特征空间中，二次判别函数的一般表示式为：

$$\begin{aligned} d(\vec{x}) &= \vec{x}'W\vec{x} + \vec{w}'\vec{x} + w_{n+1} \\ &= \sum_{k=1}^n w_{kk} x_k^2 + 2 \sum_{j=1}^{n-1} \sum_{k=j+1}^n w_{jk} x_j x_k + \sum_{j=1}^n w_j x_j + w_{n+1} \end{aligned}$$

二次判别函数图例





3.7 非线性判别函数

3.7.2 二次判别函数

一般判别规则

计算训练模式 $\vec{x}_j \in \omega_1 (j=1, 2, \dots, N_1)$

$$\hat{m}_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} \vec{x}_j \quad \hat{C}_1 = \frac{1}{N_1 - 1} \sum_{j=1}^{N_1} (\vec{x}_j - \hat{m}_1)(\vec{x}_j - \hat{m}_1)'$$

构造判别函数: $d(\vec{x}) = K^2 - (\vec{x} - \hat{m}_1)' \hat{C}_1^{-1} (\vec{x} - \hat{m}_1)$

对未知模式:

$$\begin{cases} d(x) > 0 \text{ 当 } x \in \omega_1 \\ d(x) < 0 \text{ 当 } x \in \omega_2 \end{cases}$$



3.7 非线性判别函数

3.7.3 位势函数分类法

特点：（1）可直接确定判决函数
（2）适用于非线性和线性可分的情况

位势函数的概念：

位势为0的等位线——判决界面（判别函数）

对于两类问题 ω_1 , ω_2 , 认为

如果 $x \in \omega_1$, 则 x 带正电荷

如果 $x \in \omega_2$, 则 x 带负电荷



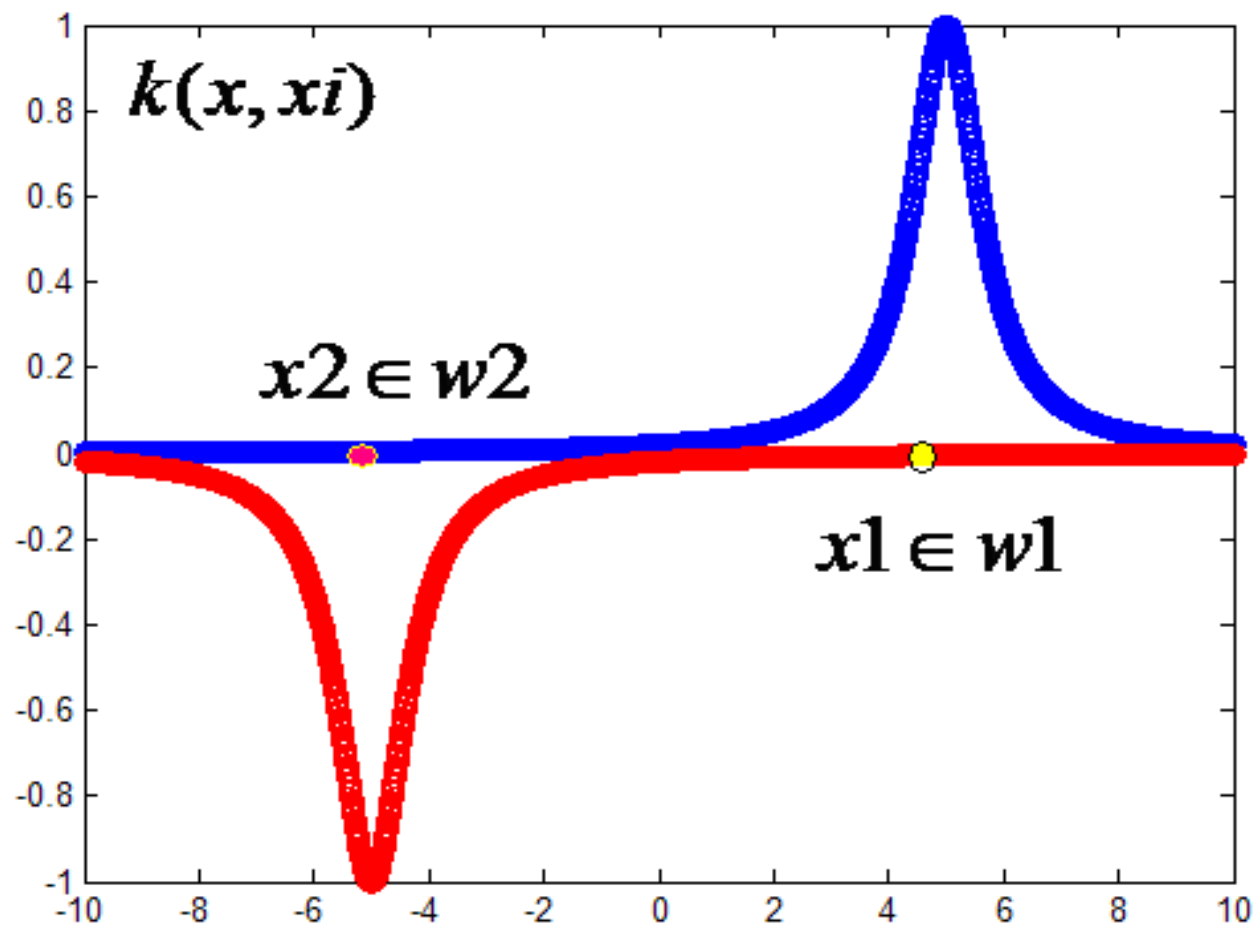
3.7 非线性判别函数

3.7.3 位势函数分类法

为了将上述分类思想数学化，我们需定义点 \vec{x}_j 处的位势函数 $K(\vec{x}, \vec{x}_j)$ ，它应满足：

- (1) $K(\vec{x}, \vec{x}_j) = K(\vec{x}_j, \vec{x})$
- (2) $K(\vec{x}, \vec{x}_j)$ 是连续光滑函数；
- (3) $K(\vec{x}, \vec{x}_j)$ 是 \vec{x} 与 \vec{x}_j 间距离的单值单调下降函数。当且仅当 $\vec{x} = \vec{x}_j$ 时， $K(\vec{x}, \vec{x}_j)$ 达其最大值。当 \vec{x} 与 \vec{x}_j 之间的距离趋于无穷大时， $K(\vec{x}, \vec{x}_j)$ 趋于零。

位势函数图例





3.7 非线性判别函数

3.7.3 位势函数分类法

下面介绍两类位势函数：

第一类位势函数

设 $\{\varphi_i(\vec{x}), i=1, 2, \dots\}$ 是 \vec{x} 定义域中的完备正交函数集，位势函数取为 $K(\vec{x}, \vec{x}_j) = \sum_{i=1}^m \varphi_i(\vec{x})\varphi_i(\vec{x}_j)$
这是对称的有限项多项式。



3.7 非线性判别函数

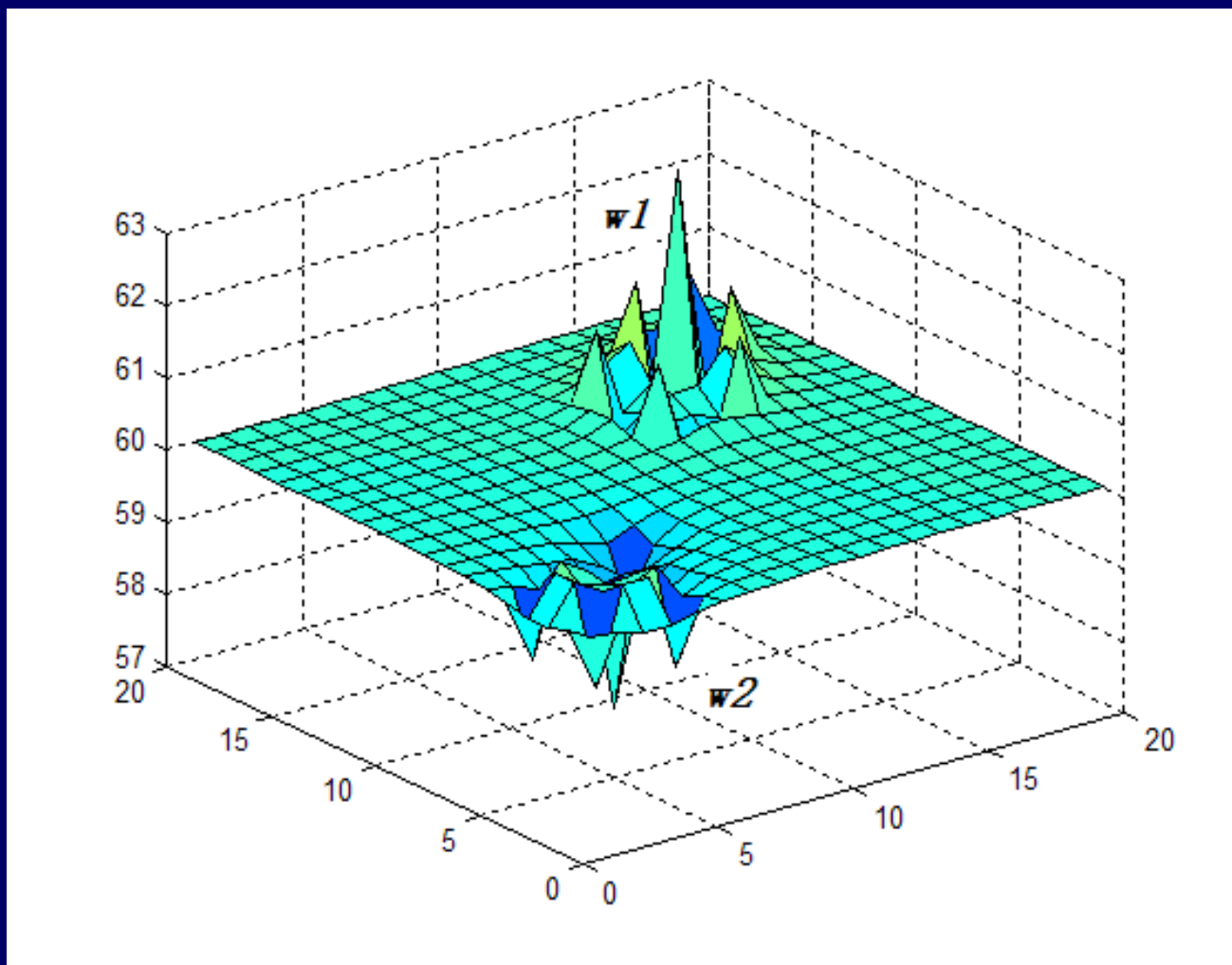
3.7.3 位势函数分类法

第二类位势函数

取关于 \vec{x} 和 \vec{x}_j 的距离的对称函数作位势函数，例如

$$K(\vec{x}, \vec{x}_j) = \exp[-\alpha \|\vec{x} - \vec{x}_j\|^2]$$

$$K(\vec{x}, \vec{x}_j) = \frac{1}{1 + \alpha \|\vec{x} - \vec{x}_j\|^2}$$



位势函数图例



3.7 非线性判别函数

3.7.3 位势函数分类法

由位势函数产生判别函数的训练算法及分类规则：

设训练模式集为 $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N\}$ ，它们分属 ω_1 和 ω_2 类。定义一个位势函数 $K(\vec{x}, \vec{x}_j)$ 。

(1) 初始化。令特征空间中各点处的初始积累位势函数 $K_0(\vec{x}) = 0$ ，判错计数 $m = 0$ ；

(2) 令 $j = 1$ ，输入训练模式 \vec{x}_1 ，使积累位势函数

$$K_1(\vec{x}) = \begin{cases} K_0(\vec{x}) + K(\vec{x}, \vec{x}_1) > 0, & \text{若 } \vec{x}_1 \in \omega_1 \\ K_0(\vec{x}) - K(\vec{x}, \vec{x}_1) < 0, & \text{若 } \vec{x}_1 \in \omega_2 \end{cases}$$

(3) 令 $j = j + 1$, 输入 \vec{x}_j , 计算积累位势函数

$$K_j(\vec{x}) = K_{j-1}(\vec{x}) + \alpha_j K(\vec{x}, \vec{x}_j)$$

积累位势函数的调整规则为:

$$\alpha_j = \begin{cases} 0, & \text{若 } \vec{x}_j \in \omega_1 \text{ 和 } K_{j-1}(\vec{x}_j) > 0 \\ 0, & \text{若 } \vec{x}_j \in \omega_2 \text{ 和 } K_{j-1}(\vec{x}_j) < 0 \\ 1, & \text{若 } \vec{x}_j \in \omega_1 \text{ 和 } K_{j-1}(\vec{x}_j) \leq 0 \\ -1, & \text{若 } \vec{x}_j \in \omega_2 \text{ 和 } K_{j-1}(\vec{x}_j) \geq 0 \end{cases} \quad m = m + |\alpha_j|$$

(4) 如果 $j < N$, 返至(3)。如果 $j = N$, 检查是否有模式判错, 即 $m = 0$, 则结束 $d(\vec{x}) = K_j(\vec{x})$;
若 $m \neq 0$, 令 $j = 0$, $m = 0$, 返至(3)。

与感知器算法类似，位势函数训练算法也可以用于多类问题，其技术要点是：

(1) 设初始积累位势函数 $K_0^{(i)}(\vec{x}) = 0$

这里 i 表示类别， $i = 1, 2, \dots, c$ 。

(2) 当 $\vec{x}_{j+1} \in \omega_i$ 时，迭代规则是：

如果 $K_j^{(i)}(\vec{x}_{j+1}) > K_j^{(l)}(\vec{x}_{j+1}) \quad (\forall l \neq i)$ 则 $K_{j+1}^{(i)}(\vec{x}) = K_j^{(i)}(\vec{x}) \quad (\forall i)$

如果 $K_j^{(i)}(\vec{x}_{j+1}) < K_j^{(m)}(\vec{x}_{j+1}) \quad (m \neq i)$ 则

$$K_{j+1}^{(i)}(\vec{x}) = K_j^{(i)}(\vec{x}) + \alpha_j K(\vec{x}, \vec{x}_{j+1})$$

$$K_{j+1}^{(m)}(\vec{x}) = K_j^{(m)}(\vec{x}) - \alpha_j K(\vec{x}, \vec{x}_{j+1})$$

$$K_{j+1}^{(l)}(\vec{x}) = K_j^{(l)}(\vec{x}) \quad (\forall l \neq i, m)$$



3.7 非线性判别函数

3.7.3 位势函数分类法

(3) 在全部训练模式均满足当 $\vec{x} \in \omega_i$ 时, 有

$$K_{j+1}^{(i)}(\vec{x}) > K_{j+1}^{(l)}(\vec{x}) \quad (\forall l \neq i)$$

算法结束。

这里 $K_{j+1}^{(i)}(\vec{x})$ 就是解多类问题第三个途径中的 $d_i(\vec{x})$ 。

例：已知如图所示两类训练样本：

$$\vec{x}_1 = (0,0)', \vec{x}_2 = (2,0)' \in \omega_1; \vec{x}_3 = (1,1)', \vec{x}_4 = (1,-1)' \in \omega_2$$

试用势函数法进行分类器训练。

解：选用第二类势函数，令 $\alpha=1$ ，在二维情况下，

$$K(\vec{x}, \vec{x}_j) = \exp\left[-\|\vec{x} - \vec{x}_j\|^2\right] = \exp\left[-((x_1 - x_{j1})^2 + (x_2 - x_{j2})^2)\right]$$

$K(\vec{x})$ 为积累位势函数

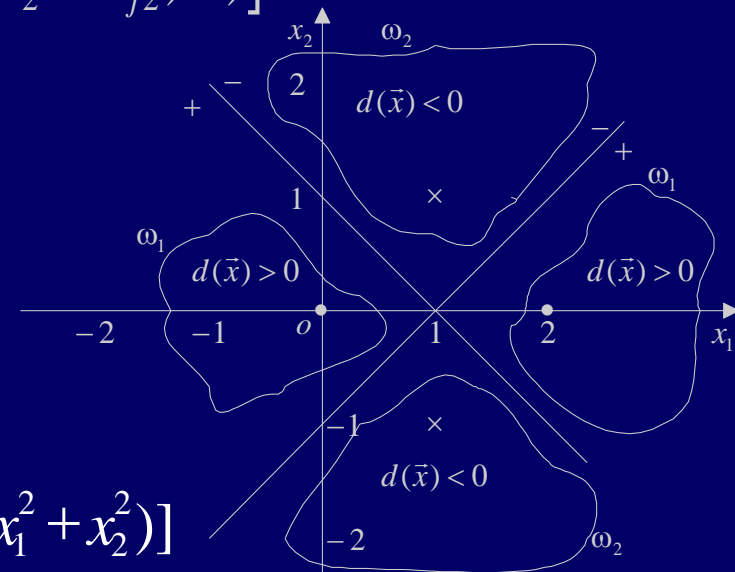
$$j=0, K_0(\vec{x}) = 0$$

$$j=1, \vec{x}_1 \in \omega_1, K_1(\vec{x}) = K(\vec{x}, \vec{x}_1)$$

$$= \exp\left[-((x_1 - 0)^2 + (x_2 - 0)^2)\right] = \exp[-(x_1^2 + x_2^2)]$$

$$j=2, \vec{x}_2 \in \omega_1, K_1(\vec{x}_2) = \exp[-(x_1^2 + x_2^2)] = \exp[-2^2] = e^{-4} > 0$$

$$K_2(\vec{x}) = K_1(\vec{x}) = \exp[-(x_1^2 + x_2^2)]$$





$$\vec{x}_1 = (0,0)', \vec{x}_2 = (2,0)' \in \omega_1; \vec{x}_3 = (1,1)', \vec{x}_4 = (1,-1)' \in \omega_2$$

$$j = 3, \vec{x}_3 \in \omega_2, K_2(\vec{x}_3) = \exp[-(1^2 + 1^2)] = e^{-2} > 0$$

$$K_3(\vec{x}) = K_2(\vec{x}) - K(\vec{x}, \vec{x}_3) = \exp[-(x_1^2 + x_2^2)] - \exp[-((x_1 - 1)^2 + (x_2 - 1)^2)]$$

$$j = 4, \vec{x}_4 \in \omega_2, K_3(\vec{x}_4) = \exp[-2] - \exp[-4] = e^{-2} - e^{-4} > 0$$

$$\begin{aligned} K_4(\vec{x}) &= K_3(\vec{x}) - K(\vec{x}, \vec{x}_4) \\ &= \exp[-(x_1^2 + x_2^2)] - \exp[-((x_1 - 1)^2 + (x_2 - 1)^2)] \\ &\quad - \exp[-((x_1 - 1)^2 + (x_2 + 1)^2)] \end{aligned}$$

$$j = 5, \vec{x}_1 \in \omega_1, K_4(\vec{x}_1) = \exp[0] - \exp[-2] - \exp[-2] = e^0 - 2e^{-2} > 0$$

$$\begin{aligned} K_5(\vec{x}) &= K_4(\vec{x}) \\ &= \exp[-(x_1^2 + x_2^2)] - \exp[-((x_1 - 1)^2 + (x_2 - 1)^2)] \\ &\quad - \exp[-((x_1 - 1)^2 + (x_2 + 1)^2)] \end{aligned}$$



$$6, \vec{x}_2 \in \omega_1, K_5(\vec{x}_2) = \exp[-4] - \exp[-2] - \exp[-2] = e^{-4} - 2e^{-2} < 0$$

$$\begin{aligned} K_6(\vec{x}) &= K_5(\vec{x}) + K(\vec{x}, \vec{x}_2) = K_5(\vec{x}) + \exp\left[-((x_1 - 2)^2 + x_2^2)\right] \\ &= \exp\left[-(x_1^2 + x_2^2)\right] - \exp\left[-((x_1 - 1)^2 + (x_2 - 1)^2)\right] \\ &\quad - \exp\left[-((x_1 - 1)^2 + (x_2 + 1)^2)\right] + \exp\left[-((x_1 - 2)^2 + x_2^2)\right] \end{aligned}$$

$$j = 7, \vec{x}_3 \in \omega_2, K_6(\vec{x}_3) = e^{-2} - e^0 - e^{-4} + e^{-2} < 0$$

$$K_7(\vec{x}) = K_6(\vec{x})$$

$$j = 8, \vec{x}_4 \in \omega_2, K_7(\vec{x}_4) = e^{-2} - e^{-4} - e^0 + e^{-2} < 0$$

$$K_8(\vec{x}) = K_6(\vec{x})$$

$$j = 9, \vec{x}_1 \in \omega_1, K_8(\vec{x}_1) = e^0 - e^{-2} - e^{-2} + e^{-4} > 0$$

$$K_9(\vec{x}) = K_6(\vec{x})$$

$$j = 10, \vec{x}_2 \in \omega_1, K_9(\vec{x}_2) = e^{-4} - e^{-2} - e^{-2} + e^0 > 0$$

$$K_{10}(\vec{x}) = K_6(\vec{x})$$



$$d(\vec{x}) = K_6(\vec{x})$$

$$d(\vec{x}) = \exp[-(x_1^2 + x_2^2)] - \exp[-((x_1 - 1)^2 + (x_2 - 1)^2)] \\ - \exp[-((x_1 - 1)^2 + (x_2 + 1)^2)] + \exp[-((x_1 - 2)^2 + x_2^2)]$$

$$\text{令 } d(\vec{x}) = 0$$

$$\exp[-(x_1^2 + x_2^2)] - \exp[-((x_1 - 1)^2 + (x_2 - 1)^2)] - \exp[-((x_1 - 1)^2 + (x_2 + 1)^2)] + \exp[-((x_1 - 2)^2 + x_2^2)] \\ = \exp[-(x_1^2 + x_2^2)] \{1 - \exp[-2(1 - x_1 - x_2)] - \exp[-2(1 - x_1 + x_2)] + \exp[-4(1 - x_1)]\} \\ = \exp[-(x_1^2 + x_2^2)] \exp[-2(1 - x_1)] \{ \exp[2(1 - x_1)] + \exp[-2(1 - x_1)] - \exp[2x_2] - \exp[-2x_2] \} = 0$$

$$\because \exp[-(x_1^2 + x_2^2)] \neq 0, \exp[-2(1 - x_1)] \neq 0$$

$$\therefore \exp[2x_2] + \exp[-2x_2] = \exp[2(1 - x_1)] + \exp[-2(1 - x_1)]$$

$$\text{即 } ch[2x_2] = ch[2 - 2x_1], (\text{双曲余弦函数: } ch(x) = \cosh(x) = \frac{e^x + e^{-x}}{2})$$

$$\left\{ \begin{array}{l} 2x_2 = 2 - 2x_1 \\ 2x_2 = -(2 - 2x_1) \end{array} \right\}$$

$$\left\{ \begin{array}{l} x_2 = 1 - x_1 \\ x_2 = x_1 - 1 \end{array} \right.$$

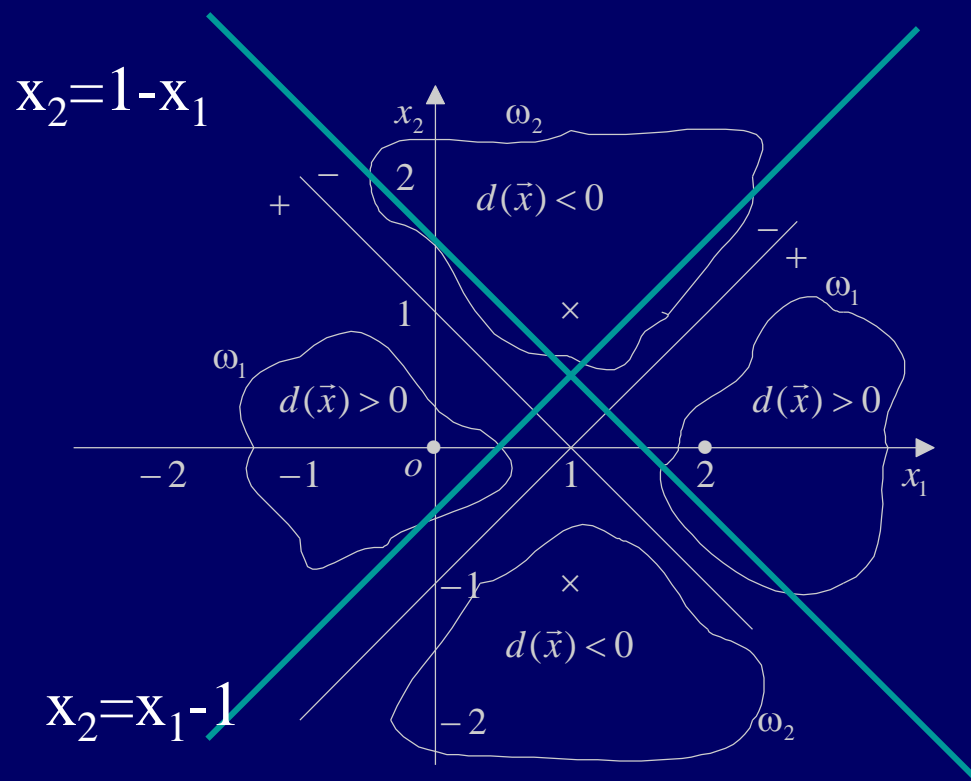


3.7 非线性判别函数

3.7.3 位势函数分类法

判别界面

$$\begin{cases} x_2 = 1 - x_1 \\ x_2 = x_1 - 1 \end{cases}$$





100元B面与100元D面的势函数分类

[illegible]



势函数法编程实例

100元B面与新100元B面的势函数分类

1,	0,	1,	0,	0,	1,	0,	0,	1,	0,
0,	1,	0,	0,	1,	0,	0,	0,	1,	0,
1,	0,	1,	0,	0,	0,	0,	1,	0,	0,
0,	0,	0,	0,	0,	0,	0,	1,	0,	1,
0,	1,	0,	1,	0,	0,	0,	0,	0,	0,
0,	1,	0,	0,	0,	0,	1,	0,	1,	0,
-1,	0,	0,	0,	0,	-1,	0,	0,	-1,	0,
0,	-1,	0,	-1,	0,	0,	0,	-1,	-1,	-1,
-1,	0,	0,	-1,	0,	0,	0,	0,	0,	0,
-1,	0,	0,	0,	0,	0,	0,	0,	-1,	0,
0,	0,	-1,	0,	-1,	0,	0,	0,	-1,	0,
-1,	0,	0,	0,	-1,	0,	0,	0,	0,	0,



谢谢！