第四章

线性判别函数

- 实际中，不去恢复类条件概率密度函数，而是利用样本集直接设计分类器。
- 线性判别函数形式简单，易分析。
- 线性判别函数往往不是最优分类器。

概念：

- 线性判别函数

$$g(x) = w^T x + w_o$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \qquad w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix}$$

两类： $g(x)=g_1(x)-g_2(x)$

$$\begin{cases} g(x)>0, & 则决策 x \in \omega_1 \\ g(x)<0, & 则决策 x \in \omega_2 \\ g(x)=0, & 可将x任意分到某一类，或拒绝 \end{cases}$$

$g(x)=0$定义了一个决策面，它把两个类别的样本点分割开来。

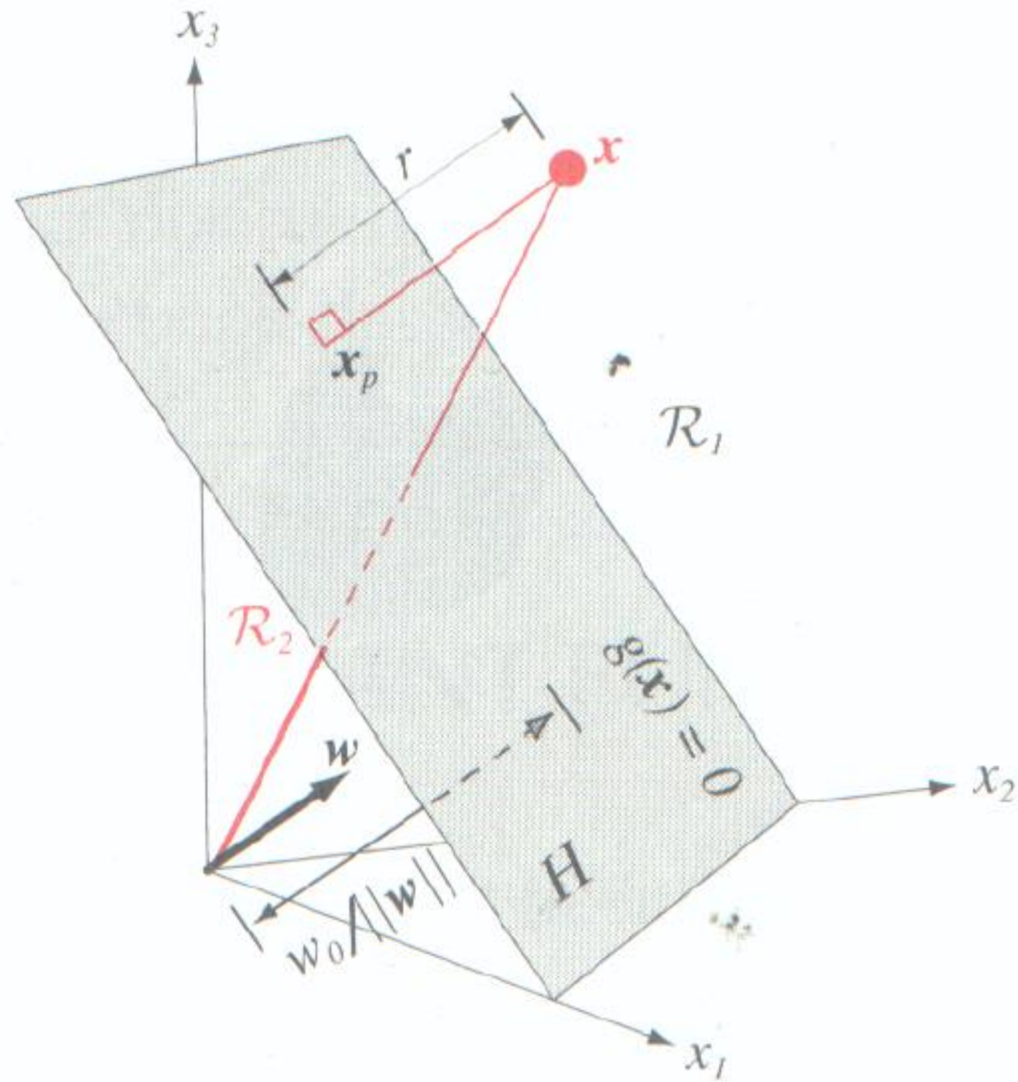
x_1, x_2都在决策面H上

$$w^T x_1 + w_0 = w^T x_2 + w_0$$

$$\Rightarrow w^T (x_1 - x_2) = 0$$

w与超平面垂直

$$x = x_p + r\frac{w}{\|w\|}$$



$$g(x) = w^T\left(x_p + r\frac{w}{\|w\|}\right) + w_0$$

$$= w^T x_p + w_0 + r\frac{w^T w}{\|w\|} = r\|w\|$$

$$\Rightarrow r = \frac{g(x)}{\|w\|}$$

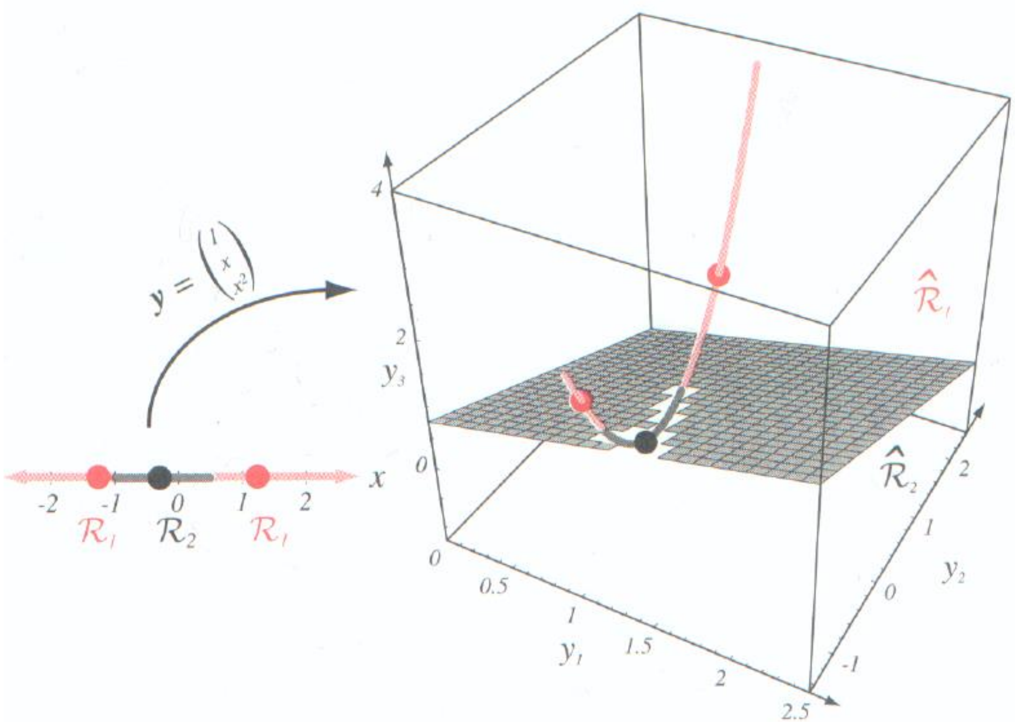
广义线性判别 图取自Duda教材

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}$$

$$g(x) = (x-a)(x-b)$$

$$g(x) = c_0 + c_1 x + c_2 x^2 = a^T y$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$



$$g(x) = w_0 + w^T x = a^T y$$

$$y = \begin{bmatrix} 1 \\ x \end{bmatrix} \quad , a = \begin{bmatrix} w_0 \\ w \end{bmatrix}$$

增广样本向量　augmented sample vector

增广权向量

感知准则函数

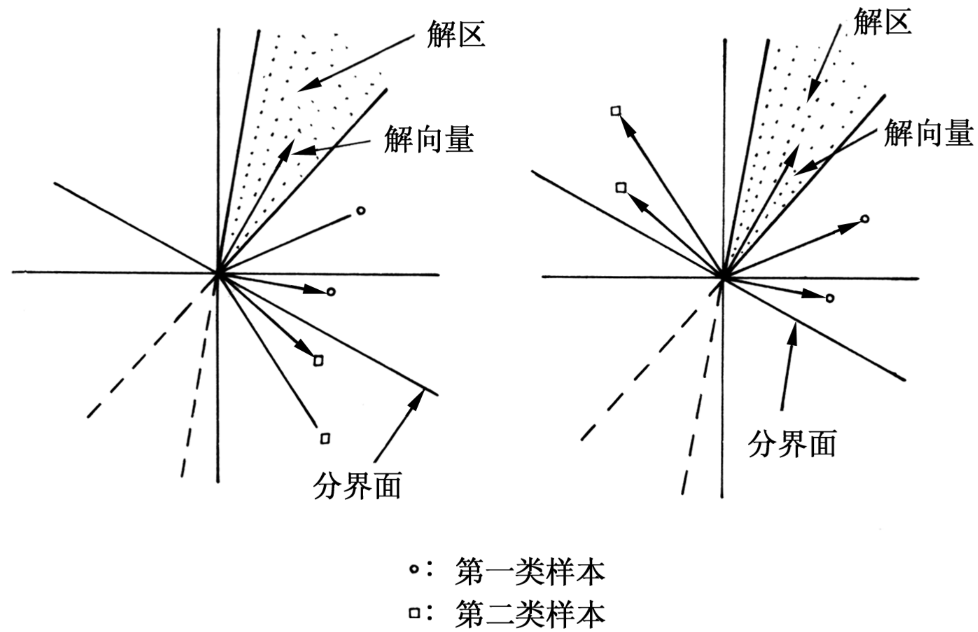
假设样本线性可分：

$a^T y$ 能正确分类每个样本

$$\begin{cases} a^T y_i > 0, \text{对一切} y_i \in \omega_1 \\ a^T y_i < 0, \text{对一切} y_i \in \omega_2 \end{cases}$$

$$y_n' = \begin{cases} y_i, \text{对一切} y_i \in \omega_1 \\ -y_i, \text{对一切} y_i \in \omega_2 \end{cases}$$

$\Rightarrow a^T y_n' > 0$ 规范化



(a) 未规范化 (b) 规范化

对线性可分样本，找a，$a^T y_n > 0, n = 1,2,\cdots,N.$

目标函数 $J_P(a) = \sum_{y \in Y^k} (-a^T y)$

Y^k：错分样本集合

梯度下降法：$\nabla J_P(a) = \frac{\partial J_P(a)}{\partial a} = \sum_{y \in Y^k} (-y)$

$$a(k+1) = a(k) - \rho_k \nabla J$$

$$a(k+1) = a(k) + \rho_k \sum_{y \in Y^k} y$$

Algorithm

Step 1: initialize $a(0), \rho_k, t = 0$

Step 2: calculate $\nabla J_P(a) = \dfrac{\partial J_P(a)}{\partial a} = \sum_{y \in Y^k} (-y)$

Step 3: Update $a(k+1) = a(k) + \rho_k \sum_{y \in Y^k} y$

Step 4: if vector does not change, stop.

else goto Step 2

Single Sample Correction Algorithm

$$y_1, y_2, \cdots, y_n, y_1, y_2, \cdots, y_n, \cdots\cdots$$

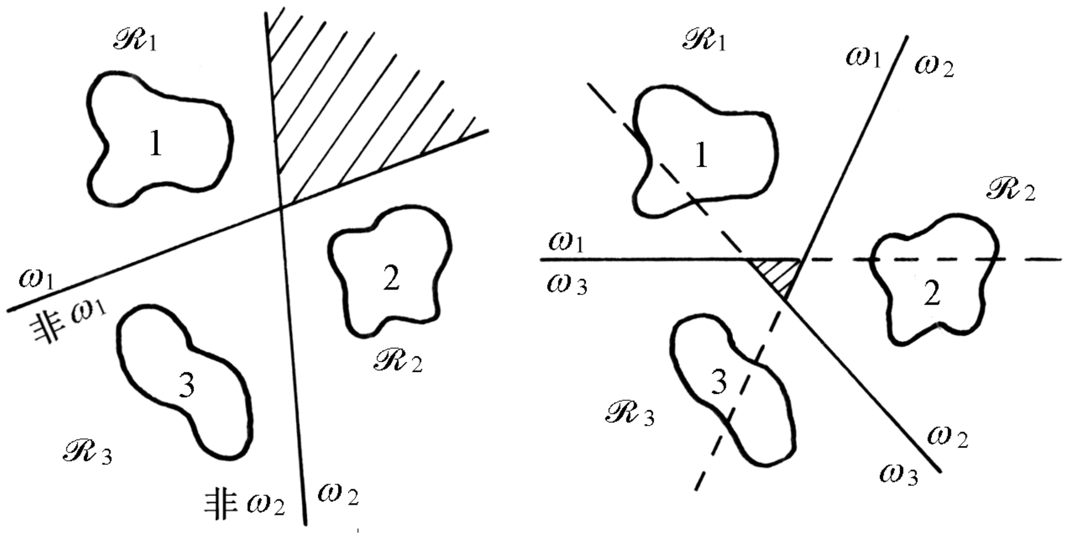
$$a(k+1) = a(k) + y^k$$

$$a^T(k+1)y^k = a^T(k)y^k + y^{k^T} \cdot y^k$$

多类问题

- → c-1个两类问题

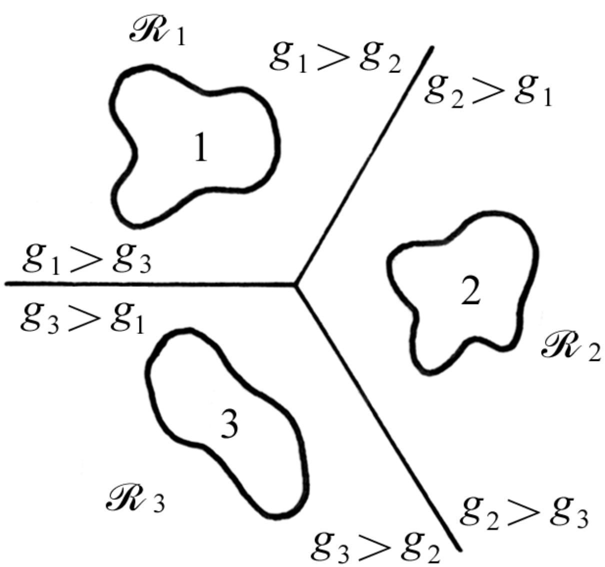
- → $\frac{c(c-1)}{2}$ 个线性判别函数



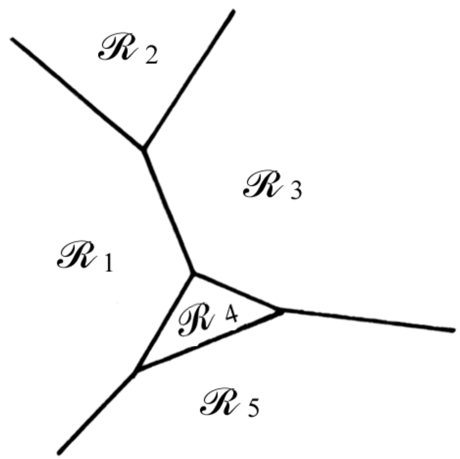
(a) ω_i/ 非 ω_i (b) ω_i/ ω_j

多类问题

理想状况



(a) 三类



(b) 五类