③线性不可分支持向量机.

▲ 软间隔支持向量机.

这种支持向量机允许一些样本出错.

即 ①中的(3)式变成下面的形式

(4) $$\min_{\vec{w},b} \frac{1}{2}\|\vec{w}\|^2 + C\sum_{i=1}^{m} l_{0/1}((\vec{w}^T\vec{x}_i+b)y_i-1)$$

$l_{0/1}(z)$ 是0/1损失函数. $l_{0/1}(z)=\begin{cases}1 & z<0\\ 0 & z>0\end{cases}$

▲ (4)中的式子一般化的形式是 $\min_{f} \Omega(f) + C\sum_{i=1}^{m} l(f(\vec{x}_i), y_i)$

- $\sum_{i=1}^{m} l(f(\vec{x}_i), y_i)$ 是经验风险. 表示在样本集中的误差.
- $\Omega(f)$ 是结构风险, 描述了模型的某些性质(比如复杂程度), 可作为正则化项来防止过拟合.

▲ (4)中的损失函数有多种选择

- 对率损失 $l_{log}(z) = \log(1+\exp(-z))$
- hinge损失 $l_{hinge}(z) = \max(0, 1-z)$

▲ 如果用对率损失. (4)式会成为形式类似于logistic回归的极大似然估计

用hinge损失函数. (4)式变为:

$$\min_{\vec{w},b} \frac{1}{2}\|\vec{w}\|^2 + C\sum_{i=1}^{m} \max(0, 1-y_i(\vec{w}_i^T\vec{x}_i+b))$$