支持向量机

2.3.1 线性分类

$f: X \subseteq R^n \to R$ （n 为输入维数，R 为实数）

$$f(x) = \langle w \cdot x \rangle + b = \sum_{i=1}^{n} w_i x_i + b$$

$\langle w \cdot x \rangle + b$ 定义一个超平面

最优超平面：以最大间隔分开数据的超平面

2.3.2 线性不可分

将样本 x 映射到某个高维特征空间

$$f(x) = \sum_{i=1}^{N} w_i \varphi_i(x) + b$$

$\varphi: X \to F$　输入空间 $X \to$ 特征空间 F

分类函数：$f(x) = \sum_{i=1}^{l} \underbrace{\alpha_i}_{\text{通过学习得到}} \overset{\text{类别标记}}{y_i} \underbrace{\langle \varphi(x_i) \cdot \varphi(x) \rangle}_{\text{内积}} + b$

2.3.3 构造核函数

定义：核是一个函数 K，对所有 $x, z \in X$，满足

$$K(x, z) = \langle \varphi(x) \cdot \varphi(z) \rangle$$

分类函数（决策规则）：$f(x) = \sum_{i=1}^{l} \alpha_i y_i K(x_i, x) + b$.

核函数要求：构造的某个特征空间必须是对称的

$$K(x, z) = \langle \varphi(x) \cdot \varphi(z) \rangle = \langle \varphi(z) \cdot \varphi(x) \rangle = K(z, x)$$

② 并且满足：$K(x, z)^2 = \langle \varphi(x) \cdot \varphi(z) \rangle^2 \le \|\varphi(x)\|^2 \|\varphi(z)\|^2 = K(x, x) K(z, z)$

$\|\cdot\|$ 是范数函数.

③ Mercer定理，对 X 的任意有限子集，相应的矩阵是半正定的（特征值非负）

（注意：只要一种核函数满足Mercer条件，它就对应某一空间中的内积.）