核方法

Kernel Method 从思想角度

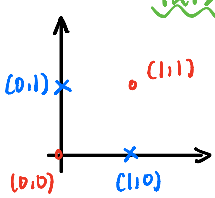
Kernel Trick 从计算角度

Kernel Function
- 非线性带来高维转换 (从模型角度) $x \rightarrow \phi(x)$
- 对偶表示带来内积 (从优化角度). $x_i^T \cdot x_j$

线性可分	一点点错误	严格非线性
PLA	Pocket Algorithm	$\phi(x)$ + PLA
Hard-Margin SVM	Soft-Margin SVM	$\underbrace{\phi(x) + \text{Hard-Margin}}$ Kernel SVM.

解决非线性问题有两种思路
- ① PLA → 多层感知机(神经网络) → Deep Learning
- ② 非线性可分 $\xrightarrow[\phi(x)]{\text{非线性转换}}$ 线性可分.

异或: 线性不可分

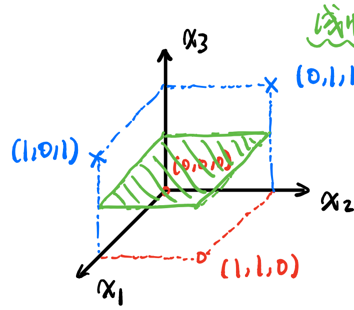


$x = (x_1, x_2)$ 二维

$\downarrow \phi(x)$

$z = (x_1, x_2, (x_1 - x_2)^2)$ 三维.

线性可分



Cover Theorem: 高维比低维更易线性可分.

Hard-Margin SVM

Primal Problem: $\min\limits_{w,b} \frac{1}{2}\|w\|^2$

Dual Problem: $\min \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j y_i y_j \underbrace{x_i^T x_j}_{\phi(x_i^T)\cdot\phi(x_j)} - \sum_{i=1}^{N}\lambda_i$

Kernel Function:

$$K(x,x') = \phi(x)^T \cdot \phi(x') = \langle \phi(x), \phi(x') \rangle$$

$\forall x, x' \in \mathcal{X}, \exists \phi: x \mapsto z$ s.t. $K(x,x') = \phi(x)^T \cdot \phi(x')$

则称$K(x,x')$是一个核函数.

$$K(x,x') = \exp\left(-\frac{(x-x')^2}{2\sigma^2}\right)$$