核方法: ①非线性带来高维转换 ②对偶表示带来内积

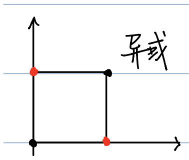
线性可分	一点点错误	严格非线性
PLA 感知机	Pocket Algorithm	PLA+$\phi(x)$
Hard-margin SVM	Soft-margin SVM	Hard-margin SVM+$\phi(x)$

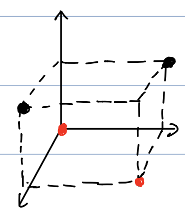
非线性学习的两种思路:

① 多层感知机 → 深度学习

hidden layer ≥ 1, 逼近任一个连续函数

② 非线性可分 $\xrightarrow{\phi(x)}$ 线性可分





Cover Theorem:

高维空间比低维空间更容易线性可分

Primal Hard-margin SVM
$$\begin{cases} \min\limits_{w,b} \frac{1}{2} w^T w \\ s.t.\ y_i(w^T x_i + b) \geq 1 \quad N\text{个约束} \end{cases}$$

Dual:
$$\begin{cases} \min\limits_{\lambda} \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j \boxed{x_i^T x_j} - \sum_{i=1}^{N} \lambda_i \\ s.t.\ \lambda_i \geq 0 \text{ for all } i=1,2,\cdots,N \\ \sum_{i=1}^{N} \lambda_i y_i = 0 \end{cases}$$

$\phi(x_i)^T \phi(x_j)$

核函数: 隐含了一次非线性转换和内积

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

$\forall x,x' \in \mathcal{X}$ $\exists \phi: \mathcal{X} \to \mathcal{Z}$ s.t. $K(x,x') = \phi(x)^T \phi(x')$ 则称$K(x,x')$是一个核函数

正定核的两个定义

函数$K: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ 满足 $\forall x,z \in \mathcal{X}$, 有$K(x,z)$, 则称$K(x,z)$为核函数

希尔伯特空间

正定核: $K: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ $\forall x,z \in \mathcal{X}$ 有$K(x,z)$. 如果 $\exists$: $\phi: \mathcal{X} \to \mathbb{R}$, $\phi \in \mathcal{H}$ s.t. $K(x,z) = \langle \phi(x), \phi(z) \rangle$, 则称$K(x,z)$为正定核

如果$K(x,z)$满足①对称性 ②正定性, 则称$K(x,z)$为正定核

对称性 $\Leftrightarrow$ $K(x,z)=K(z,x)$

正定性 $\Leftrightarrow$ 任取N个元素，$x_1 x_2 \cdots x_N \in \mathcal{X}$，对应的Gram矩阵 $K=[K(x_i,x_j)]$ 是半正定的

Gram矩阵

Hilbert Space: $\boxed{\text{完备的}}$，可能是无限维的，被赋予$\boxed{\text{内积}}$运算的$\boxed{\text{线性空间}}$ ①加法封闭 ②数乘封闭

对极限是封闭的 $\longrightarrow \lim\limits_{n\to\infty} K_n = K \in \mathcal{H}$

①对称性 $\langle f,g\rangle=\langle g,f\rangle$

②正定性 $\langle f,f\rangle \geq 0$ "=" $\Leftrightarrow f=0$

③线性 $\langle r_1f_1+r_2f_2, g\rangle = r_1\langle f_1,g\rangle + r_2\langle f_2,g\rangle$

正定核的必要性证明

证明 $K(x,z)=\langle\phi(x),\phi(z)\rangle \Leftrightarrow$ Gram矩阵半正定

必要性证明 ($\Longrightarrow$):

$A_{n\times n}$半正定：①特征值 ≥ 0 ② $\forall$向量 $\alpha\in\mathbb{R}^n$，$\alpha^T A\alpha \geq 0$

已知 $K(x,z)=\langle\phi(x),\phi(z)\rangle$ 证明 $[K(x_i,x_j)]_{N\times N}$ 半正定

即证 $\forall a\in\mathbb{R}^N$，$a^TKa\geq 0$.

$$a^TKa=(a_1\ a_2\ \cdots\ a_N)\begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & & & \\ \vdots & & & \\ K_{N1} & & & \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}$$

$$=\sum_{i=1}^N\sum_{j=1}^N a_i a_j K_{ij}$$

$$=\sum_i\sum_j a_i a_j \langle\phi(x_i),\phi(x_j)\rangle \quad = \sum_i\sum_j a_i a_j\, \phi(x_i)^T\phi(x_j) = \left(\sum_i a_i\phi(x_i)\right)^T\left(\sum_j a_j\phi(x_j)\right)$$

$$=\left\langle \sum_i\phi(x_i),\ \sum_j\phi(x_j)\right\rangle$$

$$=\left\|\sum_{i=1}^N a_i\phi(x_i)\right\|^2$$

≥ 0 证毕，必要性得证