核方法

Kernel Method

① 背景介绍

② 正定核

背景介绍（Introduction）

- Kernel Method 核方法（从思想角度）
- Kernel Trick 核技巧（从计算角度）
- $\boxed{\text{Kernel Function 核函数}}$

- 非线性带来高维转换（从模型角度）
- 对偶表示带来内积（从优化角度）

$\boxed{\text{① 非线性带来高维转换}}$

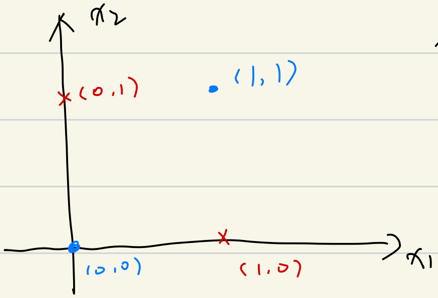
从最基本的线性分类入手：

线性可分
- 感知机（PLA）
- 硬间隔SVM

$\xrightarrow{\text{允许出错}}$

- Pocket Algorithm
- 软间隔SVM

当问题是严格非线性时，如下图（异或问题），引入一个 $\phi(x)$

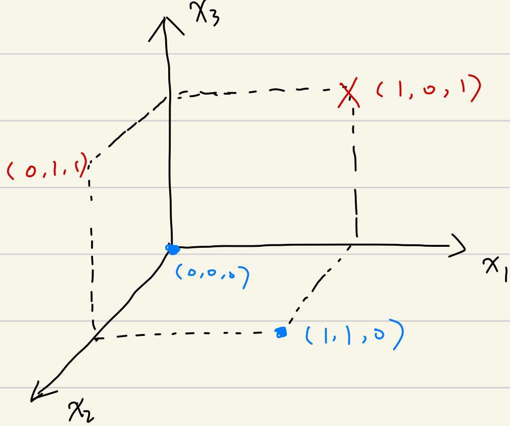


作用是：$\underset{\text{source space}}{X} \xrightarrow[\phi(x)]{} \underset{\text{feature space}}{Z}$

将数据从原始空间映射到特征空间。

针对异或问题，引入 $\phi(x)$，使得 $\underbrace{(x_1, x_2)}_{\text{二维}} \rightarrow \underbrace{(x_1, x_2, (x_1-x_2)^2)}_{\text{三维}}$

如下图：



显然此时，在中间取一水平面就可以将数据分开。

Cover Theorem：高维比低维更容易线性可分。

对严格非线性 $\begin{cases} \phi(x) + \text{PLA} \\ \phi(x) + \text{硬间隔SVM} \end{cases}$

② 对偶表示带来内积

频率角度，问题最终化为优化问题

Primal Problem

$$\begin{cases} \min\limits_{w,b} \frac{1}{2} w^T w \\ s.t.\ y_i(w^T x_i + b) \geq 1, \quad i=1,2,\cdots,N. \end{cases}$$

Dual Problem

$$\boxed{\phi(x_i)^T \phi(x_j)}$$

$$\begin{cases} \min \frac{1}{2} \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i^T x_j - \sum\limits_{i=1}^{N} \lambda_i \\ s.t.\ \lambda_i \geq 0, \quad i=1,2,\dots,N \\ \quad \sum\limits_{i=1}^{N} \lambda_i y_i = 0. \end{cases}$$

由于 $\phi(x)$ 未知，且一般情况下，$\phi(x)$ 十分难求，所以引入 $\boxed{\text{Kernel Function}}$，不求 $\phi(x)$，直接求 $\phi(x_i)^T \phi(x_j)$

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

定义：$\forall x, x' \in \mathcal{X}$，$\exists \phi: \mathcal{X} \longmapsto \mathcal{Z}$，$s.t.\ K(x, x') = \phi(x)^T \phi(x')$

则称 $K(x, x')$ 是一个核函数.

正定核（Positive Definite Kernel）

通常情况下，核函数指的是正定核函数。

两个定义

定义1：$K: \mathcal{X} \times \mathcal{X} \longmapsto \mathbb{R}$，$\forall\, x, z \in \mathcal{X}$，有 $K(x, z)$

若 $\exists\, \phi: \mathcal{X} \longmapsto \mathbb{R}$，$\phi \in \mathcal{H}$，s.t. $K(x, z) = \langle \phi(x), \phi(z) \rangle$

希尔伯特空间

那么称 $K(x, z)$ 为正定核函数。

定义2：$K: \mathcal{X} \times \mathcal{X} \longmapsto \mathbb{R}$，$\forall\, x, z \in \mathcal{X}$，有 $K(x, z)$

如果 $K(x, z)$ 满足：① 对称性 ② 正定性

则称 $K(x, z)$ 为正定核函数。

① 对称性：$K(x, z) = K(z, x)$

② 正定性：任取 N 个元素，$x_1, x_2, \cdots, x_N \in \mathcal{X}$，

对应的 Gram Matrix 是半正定的

$K = [K(x_i, x_j)]_{N \times N}$

要证：$K(x, z) = \langle \phi(x), \phi(z) \rangle \Longleftrightarrow$ Gram Matrix 半正定

（对称性太简单，这里就不说了）

希尔伯特空间：完备的，可能是无限维的，被赋予内积的线性空间。

$\boxed{\text{完备的}}$：对极限是封闭的。假设若存在序列 $\{K_n\}$

$\lim_{n\to\infty} K_n = K$，收敛。

$\boxed{\text{内积}}$：

- 对称性：$\langle f, g\rangle = \langle g, f\rangle$
- 正定性：$\langle f, f\rangle \geq 0$，当且仅当 $f=0$ 时，$\langle f, f\rangle = 0$
- 线性：$\langle r_1 f_1 + r_2 f_2, g\rangle = r_1\langle f_1, g\rangle + r_2\langle f_2, g\rangle$

$\boxed{\text{线性空间}}$：向量空间（加法和数乘是封闭的，即经过加法和数乘之后，仍属于该空间）

完整的定义有8条，这里只作简单说明。

证明：$K(x,z)=\langle\phi(x),\phi(z)\rangle \Longleftrightarrow$ Gram Matrix 半正定

必要性证明：($\Longrightarrow$)　（这里只证明必要性）.

已知：$K(x,z)=\langle\phi(x),\phi(z)\rangle$，证 Gram Matrix 半正定，且 $K(x,z)$ 对称

证：对称性证明：

$$K(x,z)=\langle\phi(x),\phi(z)\rangle$$

$$K(z,x)=\langle\phi(z),\phi(x)\rangle$$

由于内积运算本身具有对称性

则 $K(x,z)=K(z,x)$，

$K(x,z)$ 满足对称性

证明一个矩阵 A 是半正定的方法有以下两种：

① 特征值 ≥ 0

② $\forall \alpha \in \mathbb{R}^n$，$\alpha^T A \alpha \geq 0$　（充要条件）

由于 Gram Matrix：$K=[K(x_i,x_j)]_{N\times N}$

欲证 K 半正定，即证：$\forall \alpha \in \mathbb{R}^N$，$\alpha^T K \alpha \geq 0$.

$$\alpha^T\cdot K\cdot\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_N)\begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1N}\\ K_{21} & K_{22} & \cdots & K_{2N}\\ \vdots & \vdots & & \vdots\\ K_{N1} & K_{N2} & \cdots & K_{NN}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_N\end{pmatrix}$$

$$=\sum_{i=1}^N\sum_{j=1}^N \alpha_i\alpha_j K_{ij}$$　$K_{ij}=K(x_i,x_j)$

$$=\sum_{i=1}^N\sum_{j=1}^N \alpha_i\alpha_j\langle\phi(x_i),\phi(x_j)\rangle$$

$$=\left[\sum_{i=1}^N \alpha_i\phi(x_i)\right]^T\cdot\sum_{j=1}^N \alpha_j\phi(x_j)$$

$$= \left\langle \sum_{i=1}^{N} \alpha_i \phi(x_i), \sum_{j=1}^{N} \alpha_j \phi(x_j) \right\rangle$$

这两个元素值是两同元。其实就是同一个元素。

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

$$\geqslant 0.$$

故. K 半正定.

对必要性 $K(x,z) = \langle \phi(x), \phi(z) \rangle \Rightarrow$ Gram Matrix 半正定

且 $K(x,z)$ 对称

得证