Kernel Methods

FrankZhou-jun*

2019 年 9 月 10 日

目录

1 超平面

1.1 平面方程

首先理解一下什么叫平面方程，常用的有点法式、截距式、三点式[1]，

如果知道平面上的一个点，以及垂直于该平面的法线，我们就可以得到这平面啦，假设这个平面在三维空间下，依据点构成线，线构成面，面构成体，其实线、面、体都是由点构成构成的，我们求的平面方程，其实就是求满足条件下所有点。由于我们通常是在三维坐标下考虑的平面，所以平面类的点是三维的，当在大于三维空间的情况下，比如四维空间、五维空间、十维空间等，平面内的点对应维数为四维、五维、十维等。之前看到的某个帖子，当我们考虑的空间维数大于三维，所以叫超平面。

继续前面的例子哈, 现在我们在三维空间下考虑平面的问题，如果我们知道平面内的点 $\boldsymbol{P_1}(x_1, y_1, z_1)$ 和垂直于该平面的法线 $\vec{\boldsymbol{n}}$，我们就可以通过点法式求该平面啦。可以先看看图 1

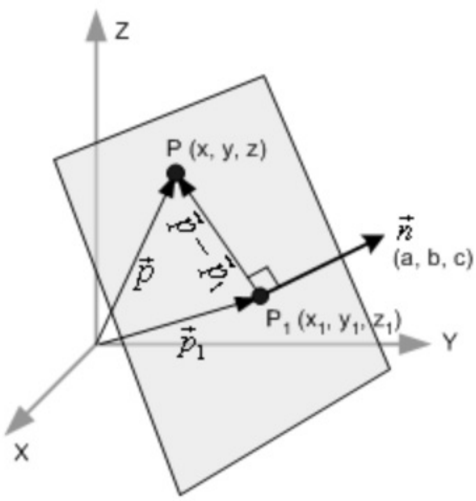

图 1: 点法式

*研究方向：信号处理，机械故障诊断，深度学习，强化学习，邮箱:zhoujun14@yeah.net

[1]高等数学

根据”垂直这一条件”，我们可以找到该平面内任意一点 $\boldsymbol{P}(x,y,z)$ 与点 $\boldsymbol{P_1}(x_1,y_1,z_1)$ 构成的向量 $\overrightarrow{\boldsymbol{P_1P}}$ 与法向量 $\vec{\boldsymbol{n}}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$ 相乘为零。

$$\begin{aligned}\overrightarrow{\boldsymbol{P_1P}}\cdot\vec{\boldsymbol{n}}&=0\\(x-x_1,y-y_1,z-z_1)\cdot(a,b,c)&=0\\a(x-x_1)+b(y-y_1)+c(z-z_1)&=0\\ax+by+cz-ax_1-by_1-cz_1&=0\end{aligned}\tag{1}$$

法向量 $\vec{\boldsymbol{n}}$ 与点 $\boldsymbol{P_1(x_1,y_1,z_1)}$ 已知，令 $D=-(ax_1+by_1+cz_1)$, 则得到公式如下：

$$ax+by+cz+D=0\tag{2}$$

这样就可以得到对应的平面方程啦，可以看到该平面是由满足条件的所有点 $\boldsymbol{P}(x,y,z)$ 组成。可以看大点构成面哈，我们一般叫求平面方程，其实求的是点。

核函数的意义就是，大白话来说就是将低维空间上的点映射到高维空间，是在低维空间不易区分的特征可以在高维上区分，如图 2

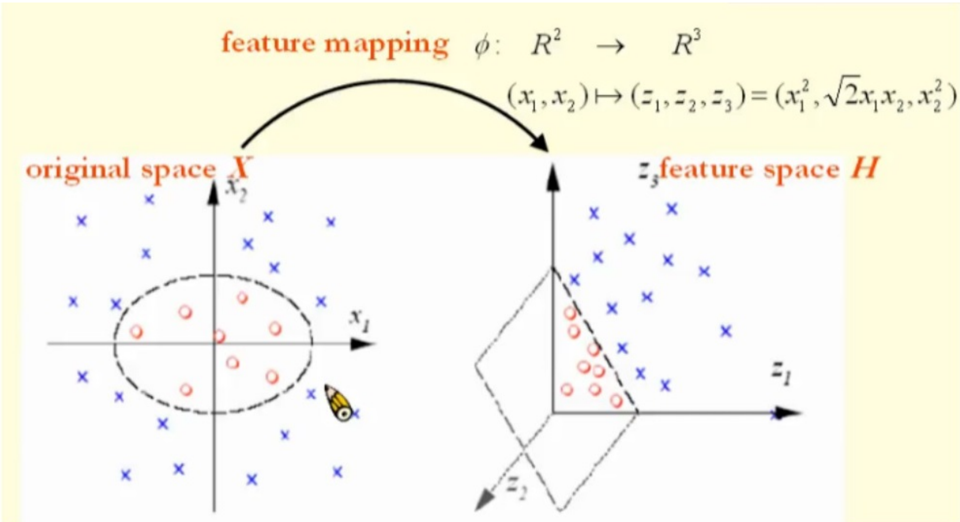

图 2: 超平面

可以看到在低维空间需要一个椭圆才能将两个类分开，实际情况可能比较复杂，简单的椭圆不可能完全分开，可能需要弯弯拐拐的才行。而在右图可以看到用以个超平面就可以将两类区分，所以核函数存在具有明显的价值。

1.2 Hyperplane

讲了这么久终于到超平面了，图 2可以看到在三维空间下的超平面，超平面是 n 维欧氏空间中余维度等于一的线性子空间[2]，也就是必须是 (n-1) 维度。在维度大于三时，划分类的那个东东就不仅仅是二维的一个平面的，可能是三维，也可能是四维的，所以叫超平面，其维度等于环境空间的维度-1。

为啥超平面的自由度比空间维度小 1，举个例子，三维空间的超平面 $ax+by+cz+D=0$,a,b,z 中至少一个不为零，假设 $b\neq0$, 则

$$y=-\frac{a}{b}x-\frac{c}{b}z-\frac{d}{b}\tag{3}$$

可以看到在三维空间中的超平面，当有两个维度是自由的时候，另一个维度就不自由了，即大小确定，所以二维空间的超平面的维度为 2.

n 维空间下的超平面定义如下：

$$\boldsymbol{W^Tx}+d=0\tag{4}$$

[2]百度

其中 $\boldsymbol{W}, \boldsymbol{x}$ 的维度都为 n

$$\boldsymbol{w} = (w_1, w_2, \ldots, w_n)^T \tag{5}$$
$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \tag{6}$$

根据前面的推理可以得到，若 $\vec{\boldsymbol{n}}$ 为超平面的法向量，则超平面方程可以表示为

$$\vec{\boldsymbol{n}} \cdot (\boldsymbol{P} - \boldsymbol{P_1}) = 0 \tag{7}$$

超平面是怎么分类的呢，很简单的一个知识点，就是就点到平面的距离，哎，高中时候就学过了，这会到忘了，没事，我们在复习一遍。

空间中 $\boldsymbol{P_2}$ 到超平面的距离 D 等于超平面上的点 $\boldsymbol{P_1}$ 指向 $\boldsymbol{P_2}$ 构成的向量 $\overrightarrow{\boldsymbol{P_1 P_2}}$ 在该平面法向量 $\vec{n}$ 上投影的距离长度。参考图 3

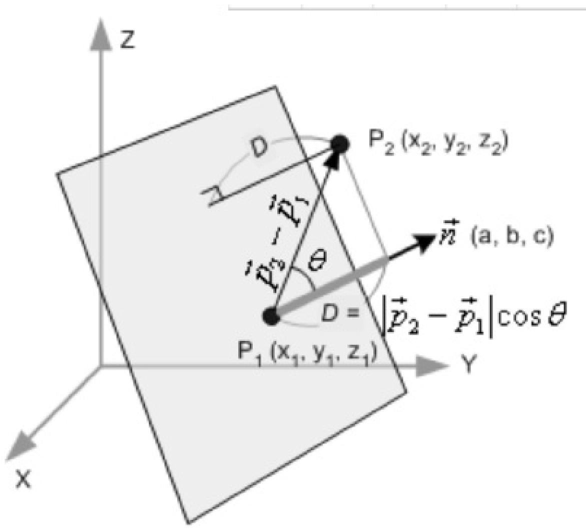

图 3: 点到平面的距离

距离 D 的大小表示为

$$D = |\overrightarrow{\boldsymbol{P_1 P_2}}| \cos\theta \tag{8}$$

而我们右知道两向量的夹角公式，利用点 $\boldsymbol{P_1}$ 和点 $\boldsymbol{P_2}$ 构成的向量 $\overrightarrow{\boldsymbol{P_1 P_2}}$ 与法向量 $\vec{\boldsymbol{n}}$ 计算角度 $\cos\theta$

$$\cos\theta = \frac{\vec{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{P_1 P_2}}}{|\vec{\boldsymbol{n}}||\overrightarrow{\boldsymbol{P_1 P_2}}|} \tag{9}$$

$$\begin{aligned} D =& |\overrightarrow{\boldsymbol{P_1 P_2}}| \cos\theta \\ =& |\overrightarrow{\boldsymbol{P_1 P_2}}| * \frac{\vec{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{P_1 P_2}}}{|\vec{\boldsymbol{n}}||\overrightarrow{\boldsymbol{P_1 P_2}}|} \\ =& \frac{\vec{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{P_1 P_2}}}{|\vec{\boldsymbol{n}}|} \end{aligned} \tag{10}$$

可以看到化简后方程的分子部分就为环境空间的超平面方程，则该环境空间中的任意一点的距离到超平面的距离表示为

$$D = \frac{\boldsymbol{W^T x} + d}{|\vec{\boldsymbol{n}}|} \tag{11}$$

$$T_\sigma = \begin{cases} above \quad plane, \boldsymbol{W^T x} + d > 0 \\ in \quad plane, \boldsymbol{W^T x} + d = 0 \\ under \quad plane, \boldsymbol{W^T x} + d < 0 \end{cases} \tag{12}$$

2 核函数

支持向量机通过某非线性变换 (x) ，将输入空间映射到高维特征空间。特征空间的维数可能非常高。如果支持向量机的求解只用到内积运算，而在低维输入空间又存在某个函数 K(x, x) ，它恰好等于在高维空间中这个内积，即 K(x, x) =< (x) (x) > 。那么支持向量机就不用计算复杂的非线性变换，而由这个函数 K(x, x) 直接得到非线性变换的内积，使大大简化了计算。这样的函数 K(x, x) 称为核函数