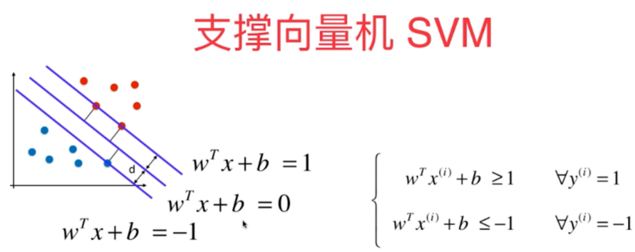
SVM优化建模方法调研

计试001 王嘉禾 2193211079

2022 年 10 月 6 日

1　经典SVM

支持向量机是1992年由Bell实验室的vladimir Vapnik和他的同事首次提出的。对于一个二分类问题，对每一个待分类结点进行一个向量表示，并寻找一个超平面，似的两个分类的结点分别处于超平面两边。为了使超平面具有鲁棒性，我们倾向于寻找一个距离所有结点的距离整体最远的超平面，使得二分类的依据更加明显。公式化的表达是，使得超平面距离所有结点的距离中最小的一个最大，也就是超平面上的点距离结点的最小距离中的最小值最大。



为达到将两类结点尽量分开的目的，我们可以令两种标签的点分别在超平面$\mathbf{wx}+b>1$之上和超平面$\mathbf{wx}+b<-1$之下，并通过调整比例求出$\mathbf{w}$和b的值，从直觉的角度来说，我们用一个最厚的“超平板”将两类结点分开，公式化的表达如下所示

$$\begin{aligned} \mathbf{wx_i}+b>1, y_i=1 \\ \mathbf{wx_i}+b<-1, y_i=-1 \end{aligned} \tag{1}$$

而令超平面距离所有结点距离的最小值最大的过程则相当于令这个“超平板”最厚，也就是超平面$\mathbf{wx}+b>1$和超平面$\mathbf{wx}+b<-1$之间的距离最大，并且这里直接将y_i的信息与不等式进行合并，可得

$$\begin{gathered} \min_{\mathbf{w}\in\mathbb{R}^d, b\in\mathbb{R}} \{\frac{2}{|\mathbf{w}|}\} \\ y_i(\mathbf{wx_i}+b)>1 \end{gathered} \tag{2}$$

2　带松弛条件的SVM

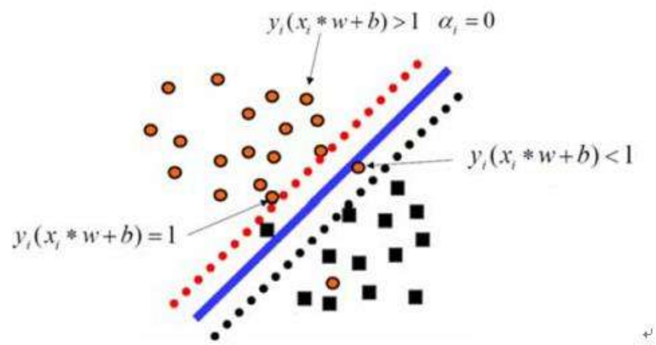
然而，由于超平面的分类能力有限，并不是所有的二分类问题都可以做到完全线性可分。此时，若遇到在超平面附近有小规模接触的情况，整体仍然是不严格的“线性可分”的，则可引入松弛因子ξ_i,令

$$\begin{aligned} \mathbf{w}\mathbf{x_i} + b > 1 - \xi_i, y_i = 1 \\ \mathbf{w}\mathbf{x_i} + b < -1 + \xi_i, y_i = -1 \end{aligned} \tag{3}$$

合并后的问题变为

$$\begin{aligned} \min_{\mathbf{w}\in\mathbb{R}^d, b\in\mathbb{R}} \{\frac{2}{|\mathbf{w}|} + c\sum_{i=1}^{n} \xi^i\} \\ y_i(\mathbf{w}\mathbf{x_i} + b) > 1 - \xi_i \end{aligned} \tag{4}$$

我们仍可以用刚才“超平板”的角度来解释松弛因子，这相当于分类并不是将所有结点都分在超平板的两侧，而是超平板两侧明确分类，而处在超平板中的结点不做要求，并且，由于每个点各对应一个松弛因子，因此这实际上是一个趋近于超平面的超曲面



3 支持向量机的核函数

这是用支持向量机解决线性不可分问题的另一种方式，线性不可分有时是因为维数不足或特征的表示方式无法利用，可通过扩展至更高的维度或进行变换以实现线性可分。例如，对与一维空间中的点x_i，我们构造映射函数

$$\phi(x_i) = \begin{bmatrix} x_i \\ x_i^2 \\ x_i^3 \end{bmatrix} \tag{5}$$

有或者对于三维空间中的点$\mathbf{x_i}$,构造映射函数

$$\phi(\mathbf{x_i}) = \begin{bmatrix} x_1x_2 \\ x_2x_3 \\ x_3^3 \end{bmatrix} \tag{6}$$

则我们分别将一维空间映射至一个三维空间，和将一个三维空间中的结点的特征进行了同维度上的重新表示。在这个(新的)三维空间上可能实现线性可分。我们希望将得到的特征映射后的特征应用于SVM分类，而不是最初的特征。下面将核函数形式化定义，假设原始的内积是$<\mathbf{x}, \mathbf{z}>$,则映射之后的内积为$<\phi(\mathbf{x}), \phi(\mathbf{z})>$,这里计算复杂度会提高很多，不仅是内积双方的维数变高，单纯计算映射也有不小的复杂度。但由于所有的分

类都是计算一次内积，于是我们把这个内积定义出来，令其为核函数$K(\mathbf{x},\mathbf{z}) = \phi(x)^T\phi(z)$，将核函数进行展开重合并或者代换等，可以实现算法复杂度的降低。