2020.2.17 统计学习方法7.3 7.4 SVM

1. 如何判断函数是一个核函数？

定理7.5说当对应的Gram矩阵，，是半正定矩阵时，是一个正定核函数。

1. SMO第一个变量选择什么样的？

第一个变量是在训练样本中选取违反KKT 条件最严重的样本点。检验样本点是否满足KKT条件的公式如下：

该检验过程先检验支持向量，再检验整个训练集上的点。

1. SMO第二个变量如何选择？

第二个变量选择要选择使得有足够大变化的点，其选择是通过计算找到最大的所对应的，是预测和真实值的差，那么其减法就是找与第一个差值的差值最大的点。E1如果是正数，则选择最小的E2；E1是负数，就选择最大的E2，这样保证绝对值|E1-E2|最大。

1. 几何间隔是什么意思？

几何间隔就是点到超平面的距离。

在二维空间中，几何间隔就是点到直线的距离。在三维及以上空间中，就是点到超平面的距离。

当||w||=1几何间隔就是函数间隔.这个时候, 如果任意放大||w||,几何间隔是不会改变的, 因为||w||也会随着被放大。也就是说，当系数成倍的变化时，几何间隔不会改变。

1. 各种核函数的特点？
   1. 线性核函数，主要用于线性可分情况，一般情况下，对于特征数量相对于样本数量非常多时，适合线性核函数
   2. 多项式核函数，对于正交归一化后的数据，可以优先选择此函数
   3. 径向基核函数，灵活性强，在不确定使用什么核函数时，可以优选选择此核函数
   4. sigmoid核函数，SVM使用此核函数，相当于一个两层的感知机
2. SVM为什么不能处理缺失值？

因为其涉及到了距离度量，涉及到了计算点与点之间的距离，类似的还有KNN算法，但类似决策树模型就不存在这样的问题，它可以采用缺失值对应类别的中位数替换。

1. 松弛变量的作用？

因为如果有1万个点符合分类结果，但是有一个点不符合，我们没有必要为这个一个点而再次浪费计算资源，因此我们大概率把这一个点当成噪声来处理。但是在硬约束的SVM中，没有类似的噪声处理机制，因此我们需要加入松弛变量，让原本的计算式有容错的能力。也就是说，我们可以得到更大的几何间隔，允许某些样本的间隔比1小。