秦旭-200209

一、 自己提出的问题的理解：

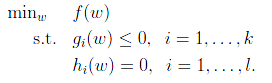
1、 支持向量机原始问题是不是凸优化问题，凸优化问题求解的优势：

(1): 线性可分支持向量机学习的最优化问题是一个凸二次优化问题；

(2): 凸优化可以采用梯度下降法进行求解，方便求解的继续，故很多方法努力讲问题转化成凸二次优化的问题，然后用梯度下降的方法解决；凸优化的性质很好：对于带约束的非凸连续优化问题，其对偶问题一定为凸并且该对偶问题的解作为原问题的一个下界；大部分凸优化问题解起来比较快，可以在多项式时间内求解。

2、 拉格朗日乘子在不等式约束情况下大于等于0 的解释：

原问题为：



转化为：

https://images.cnblogs.com/cnblogs_com/jerrylead/201103/201103131234524417.png

需要求解这个问题的最大值，假设g(w)>0或者h(w)!=0，总可以调整alpha和beta使上式最大值为无穷大，只有g(w)或h(w)满足条件，才能使上式得到常规情况下的最大值，所以要设置alpha>=0;

另外一种解释从几何角度解释：

参见链接： <https://zhuanlan.zhihu.com/p/26514613>

3、 SVM的对偶形式和感知机对偶形式的比较：

感知机对偶形式是针对每一个标记点选取优化变量a，根据误分类的情况选择该点变量的大小；同时将每次都要进行的内积运算通过GRAM矩阵提前算好，从而优化计算量；同时依次选取点，判断其误分类情况，从而改变对应优化变量a；这种方法的收敛性得到了证明。

SVM的对偶形式同样采取相同的方法：针对每个标记点，设置单独的优化变量，决定该点的w,b参数在最终分类向量参数中的占比；同样对目标函数进行优化，不需要每次都进行内积运算，而是可以通过提前计算得到；略有不同地方是，在这里没有直接提及依次选取点从而使得全局最优，而是直接提及全局最优。在7.4中，SMO是通过依次优化不满足KKT条件的点从而达到最优条件的启发式算法，可以理解为对偶的依次求解方法。

二、 别人提出的问题的理解：

4、 公式7.13中的1/2怎么理解：

后续有对w的求导，前面加一个1/2正好将此系数消掉，不影响问题整体的求解；

5、 SVM的对偶形式为什么能够降低计算量：

第一点是遍历时候不需要每次都用做内积，参见问题3；

第二点是优化函数变化：原始问题的约束方程数对应于对偶问题的变量数, 而原始问题的变量数对应于对偶问题的约束方程数, 而约束方程数目越少, 优化问题求解的复杂度越低；在线性SVM的原问题中，样本量为N；但在优化目标函数中只有1个w ，有 N个限制条件；

在线性SVM的对偶问题中，只有1个限制条件；当训练样本总数不大，特征空间的维度d>>n时，选择在对偶问题中求解将有效减少计算量

6 如何直观理解核方法

将样本点非线性转为线性：

SVM是对特征空间进行划分的方法，如果原方法的特征空间不是线性可分支持向量机或者线性向量机划分的情况，就需要对原始参数进行变化；另外一方面，在支持向量机原始问题的对偶形式中，有Xi和Xj的乘积，这个地方能够方便核方法的使用。

三、 （必填）读书计划

1、本周完成的内容章节：7.1-7.3

2、下周计划：7.3-7.4，细节证明