1. **自己提出的问题的理解：**

**问题1：**为什么在选择α1的时候要选择违反KKT条件最严重的样本点，以及如何选择一个违反KKT条件最严重的样本点。

讨论后的理解：直观上理解，为了能够更好的优化模型，减小损失函数的大小，需要尽可能的将所有的分类错误的样本点都能够正确分类，而为了能够更加快速的缩小损失函数，可以选择分类错误最严重的样本点，即违反KKT条件最严重的样本点。

之后，为了能够选择到一个这样的样本点，需要将该样本点代入到损失函数中取值最大的那一个样本点，这样才能够最快速的优化整个模型。

**问题2：**为什么α2选择的标准是希望有足够大的变化，以及为什么可以根据|E1-E2|来判断？

讨论后的理解：书中有提到，选择第二个变量的标准是希望α2能够有足够大的变化，因为根据式（7.106），可以知道，只要（E1-E2）足够大，那么α的变化就会足够快，而α1和E1都是确定的，所以可以通过|E1-E2|来判断α2的变化最大。

1. **别人提出的问题的理解：**

**问题1：**SMO为什么需要优化两个变量α1和α2

讨论后的理解：根据KKT条件中的式（7.102），可以知道从3开始其他的所有的α和y的乘积的和为α1和α2的与其y的乘积的和，而他们的和又都是一个常数，所以只要优化的α1和α2，就能够优化3到其他的所有的α，从而使得所有样本点满足KKT条件。

**问题2：**p144页的y1不等于y2为什么能推出来a1-a2=k？

讨论后的理解：

因为根据第一二节中的推到可知，y的取值只能为+1或者-1，所以y1α1+y2α2只能有两种取值，即α1+α2和α1-α2.

**问题3：**P49页，SMO算法步骤3,若在精度 范围内....为什么满足这个条件就停机？

讨论后的理解：之所以需要满足KKT条件，是因为只要所有的样本点都能够满足KKT条件，那么此时模型收敛 ，达到我们想要的最优模型。而之所以需要在精度范围内，则是因为可以放宽精度的要求，加快收敛速度。

1. **读书计划**

1、本周完成的内容章节：《统计学习方法》第7章剩余部分

2、下周计划：《统计学习方法》第8章

1. **读书摘要及理解或伪代码的具体实现**
2. **读书摘要：**

1、二次规划是求一个二次函数的全局最小值，关于其具体说明：<https://blog.csdn.net/miscclp/article/details/7488318>

2、SMO 算法是一种启发式算法，其基本思路是如果所有变量的解都满足此最优化问题的KKT条件，那么这个最优化问题的解就得到了。

3、字符串核是定义在字符串集合上的核函数。 字符串 核函数在文本分类、信息检索 、 生物信息学等方面都有应用。

4、非线性问题往往不好求解，所以希望能用解线性分类问题的方法解决这个问题。所采取的方法是进行一个非线性变换，将非线性问题变换为线性问题，通过解变换后的线性问题的方法求解原来的非线性问题

1. **代码实现**

from sklearn.svm import SVC

from sklearn.datasets import load\_breast\_cancer

from sklearn.model\_selection import train\_test\_split

data = load\_breast\_cancer()

X = data['data']

y = data['target']

x\_train, x\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.3)

clf1 = SVC(kernel='linear')

clf2 = SVC(kernel='rbf', C=10, gamma=0.0001)

clf1.fit(x\_train, y\_train)

clf2.fit(x\_train, y\_train)

print('线性SVM的精度为{}：'.format(clf1.score(x\_test, y\_test)))

print('高斯核函数SVM的精度为{}：'.format(clf2.score(x\_test, y\_test)))