**《人工智能与模式识别》平时作业**

**第三章作业**：(共10分)

（单选（0.5分）×4，多选（1分）×3，简答题（0.5分）×2，大题（4分）×1）

1.【单选】用一对多的构建方式来求解三分类问题时，在什么情况下是可以做出分类决策的？**（B）**

A. 三个判别函数值都小于0 B. 仅有一个判别函数值大于0

C. 有两个判别函数值大于0 D. 三个判别函数值都大于0

2.【单选】用一对一的构建方式来求解三分类问题时，做出多分类决策的原则是？ **(C)**

A. 没有冲突的原则 B. 最大值的原则

C. 多数投票的原则 D. 三个判别函数值都大于0

3.【单选】梯度下降算法中，学习步长的取值情况为： **(D)**

A.可以取小于0的数 B.可以取值为0

C.取值任意 D.只能取正数

4.【单选】感知器算法用什么方法求解准则函数的最优值？ **(B)**

A.最小均方误差 B.梯度下降法

C.最大均方误差 D.平均值法

5.【多选】下列选项中，属于感知机的特点的是？**（ACD）**

A. 感知机将输入信号加权求和与阈值比较，当信号加权和大于阈值时输出1，当信号加权和小于阈值时输出-1

B.单个感知机可以解决非线性分类问题

C.感知机具有多路输入、单路输出

D.单个感知机只能解决二分类问题

6.【多选】下列选项中，哪些因素会影响到感知器算法中随机梯度下降法的求解结果？**(BCD)**

A.学习规则 B.学习步长

C.初始权向量设置 D.样本处理顺序不同

7.【多选】通过构建多个线性判别函数可以解决多分类问题，以下哪种表述是正确的？**(ACD)**

1. 一对多的构建方式存在不可识别区域较多的问题

B.一对一的构建方式中不可识别的区域比一对多方式的更多

C.一对一的构建方式需要训练的线性判别函数的数目较多

D.最大值组合方式不存在不可识别区域，其构成的多分类器的整体性能比较好

1. 简述线性分类器训练的一般思路是什么？

**答：**(1) 设计一个准则函数(损失函数)，其值能够代表解的优劣程度，准则函数值越小，说明解越符合要求，越好; (2) 通过寻找准则函数的极小值，找到最优的一个解。

1. 简述广义线性化的思想。

**答：**通过一个非线性变换，将一个在低维特征空间中线性不可分的分类问题映射到高维特征空间中变成线性可分的问题， 这种方法就称为“广义线性化”。

10.给定一个训练集，其中正例样本为，，负例样本为，请用感知机算法求出能正确分类该数据集的分类模型：

，其中， .

请使用随机梯度下降法更新，初始化为零向量，取学习步长η=1，写出详细的计算过程。

**解答**：①对训练数据集进行规范化得到：

x\_1=(3, 3,1), y\_1=+1; x\_2=(5,2,1), y\_2=+1; x\_3=(1,1,1), y\_3=−1; **(0.5分)**

②利用随机梯度下降法求解：初始化w=0； 取学习步长η=1；

对于x\_1，因为y\_1 x\_1=0，修正权向量: w = w + η y\_1 x\_1 = (3,3,1) **(0.5分)**

对于x\_2，无需修正权向量；

对于x\_3，因为y\_3 x\_3 <0，需要修正：w = w + η y\_3 x\_3=(2, 2, 0) **(0.5分)**

对于x\_1，无需修正；对于x\_2，无需修正；

对于x\_3，y\_3 x\_3 <0，需要修正；w = w + η y\_3 x\_3 = (1, 1, -1) **(0.5分)**

对于x\_1，无需修正；对于x\_2，无需修正；

对于x\_3，y\_3 x\_3 < 0，需要修正；w = w + η y\_3 x\_3 = (0, 0, -2) **(0.5分)**

对于x\_1，y\_1 x\_1 < 0，需要修正：w = w + η y\_1 x\_1= (3, 3, -1) **(0.5分)**

对于x\_2，无需修正；

对于x\_3，y\_3x\_3 < 0，需要修正：w = w + η y\_3 x\_3 = (2, 2, -2) **(0.5分)**

对于x\_1和x\_2，无需修正；

对于x\_3，y\_3x\_3 < 0，需要修正；w = w + η y\_3 x\_3 = (1, 1, -3) **(0.5分)**

接下来，对于x\_1、x\_2，x\_3均无需修正，算法终止。

所以，求得的模型为：

**第四章作业**：(共12分)

（单选（0.5分）×4，多选（1分）×3，大题（7分）×1）

1.【单选】硬间隔最大化思想中的“间隔”指的是什么？ **(D)**

A. 训练数据集中所有样本点到决策超平面的垂直距离的平均距离

B. 训练数据集中所有样本点到决策超平面的垂直距离中的最大距离

C. 样本点到决策超平面的垂直距离

D. 训练数据集中所有样本点到决策超平面的垂直距离中的最小距离

2.【单选】关于“间隔最大化”的动机，以下最准备的描述是哪一个？ **(A)**

A. 找到一个超平面，以足够大的置信度将所有训练样本正确分开

B. 找到一个将训练样本正确分开的超平面

C. 将训练样本正确分开

D. 无

3.【单选】线性支持向量机对误分类的容忍程度与其惩罚系数有什么关系？**(C)**

A. 二者没有联系

B. 二者之间存在联系，但是不能找到联系间存在的规律

C. 惩罚系数越大，容忍度越低

D. 惩罚系数越大，容忍度越高

4.【单选】在使用非线性SVM分类算法前，需要提前确定（ **D** ）

A. 将输入数据映射到多少维的高维空间中

B. 非线性映射

C. 数据映射到高维空间中后是否是线性可分的

D. 核函数和惩罚参数

5.【多选】原始问题和对偶问题的最优解一定满足KKT条件，请问KKT条件包括哪些条件？(**ABCD** )

A. 拉格朗日函数关于原问题变量的梯度在最优解处为零向量

B. 对偶问题的约束

C. 对偶互补条件

D. 原始问题的等式和不等式约束

6.【多选】关于线性可分的SVM中的支持向量，下列描述正确的是 **(ABCD)**

A. 将除支持向量以外的样本点删除，也不影响分类决策边界

B. 支持向量离决策超平面的距离最近

C. 支持向量的数目一般比较少

D. 在决定分类决策超平面的位置时，只有支持向量起作用，其它样本点不起作用

7.【多选】关于软间隔SVM，以下描述正确的有：**(ABD)**

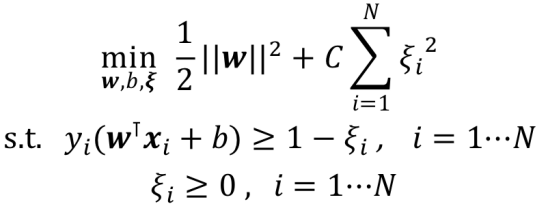
A. 能解决由于少量的噪声或异常点引起的线性不可分问题

B. 不强迫所有的训练样本都被正确分类，允许一些样本被错误分类

C. 能解决真正的非线性分类问题

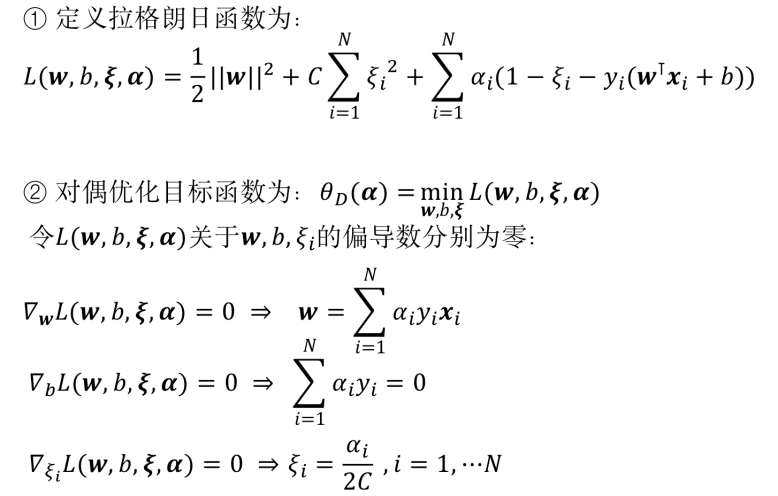
D. 比硬间隔最大化的SVM适用范围更广

8. 线性支持向量机的优化问题还可以表示为如下形式：



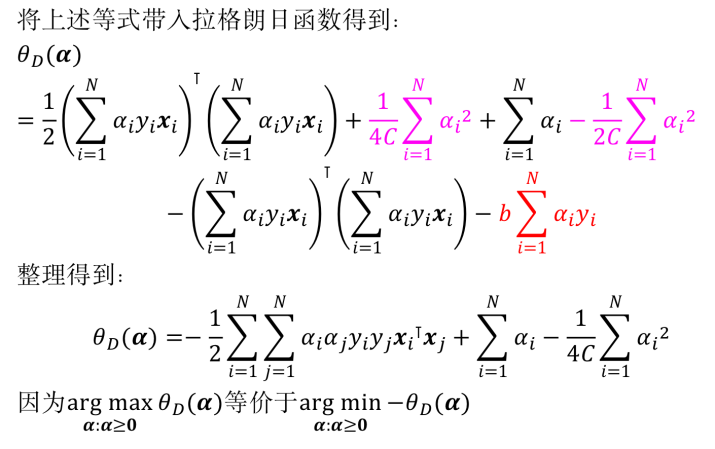
1. 请推导出该优化问题的对偶问题。
2. 假设对偶问题最优解为，请利用KKT条件构造出原始问题的最优解。

答：（1）

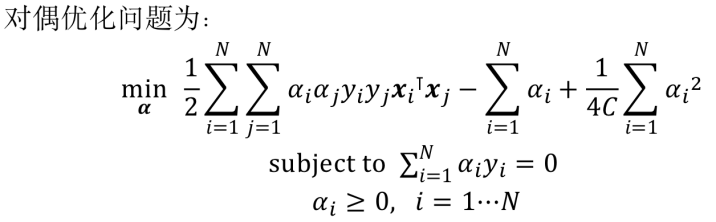


（1分）

（1分）

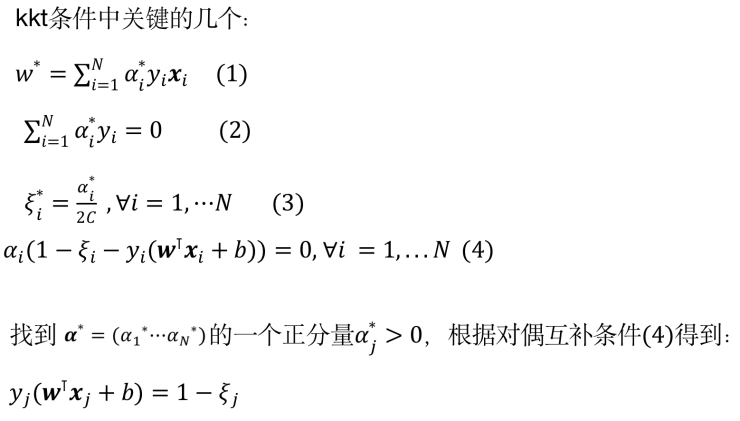


（1分）

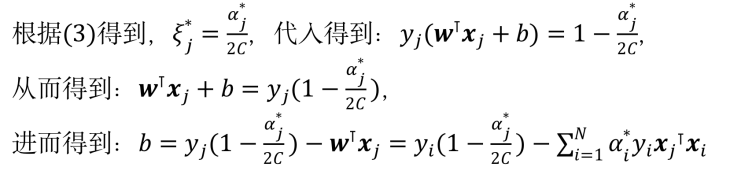


（1分）

（2）



（2分）



（1分）