# 第四章 支持向量机

## 一、填空题

1. 50 个样本训练集上的线性 SVM 原始优化问题，具有（50）个（线性不等式）约束条件.
2. 线性 SVM 问题经过拉格朗日对偶处理后的优化问题的自变量是（拉格朗日乘子$\alpha_i$）
3. 线性不可分问题可以使用（软间隔）SVM 或（核化）SVM 进行求解。
4. 常见的核函数有（多项式）核函数、（高斯）核函数和（Sigmoid）核函数

## 二、判断题

5. 对于一组线性可分的训练集，所有能够将它们正确分类的线性分类器都具有相同的分类性能。(×)
6. 所有支撑向量均在间隔区域的边界上。(×)
7. 间隔区域越宽意味着模型的经验风险越小。(×)
8. 线性 SVM 求解必须依靠拉格朗日对偶技巧。(×)

## 三、选择题

9. 对于包含 100 个样本的训练集，线性 SVM 原始问题的待优化变量数量为(　)

A. 2 个　　B. 3 个　　C. 101 个　　D. 无法确定

10. 原始线性 SVM 问题经过拉格朗日对偶处理后转化为一个：(　)

A. 无约束最大化问题　　B. 无约束最小化问题

C. 有约束最大化问题　　D. 有约束最小化问题

11. 利用 SMO 算法求解 SVM 问题时，每轮迭代选取（ ）个系数变量进行更新

A. 1 个　　B. 2 个　　C. 样本数量　　D. 无法确定

## 四、简答题

1. 简述支持向量的定义，并画图说明。
2. 简述 SVM 优化问题的三要素，并辅助变量/表达式形式进行说明。
3. 写出样本空间中的点$\boldsymbol{x}$到超平面$(\boldsymbol{\omega},\boldsymbol{\gamma})$的距离公式，并对符号进行解释。
4. 推导线性 SVM 约束条件 $y_i(\boldsymbol{\omega}^T\boldsymbol{x}_i+\gamma)\geq 1,\ \forall\boldsymbol{x}_i$。
5. 推导线性 SVM 的间隔宽度为$\frac{2}{\|\omega\|}$。
6. 写出线性 SVM 原始问题的 KKT 条件。

7. 写出经过拉格朗日对偶变换后线性 SVM 的最优化问题的标准形式。
8. 试分析，在软间隔 SVM 中，当松弛变量$\xi_i$对应的系数 C 增大时，最优解对应间隔宽度的变化趋势，并解释原因。
9. 在一个松弛项系数$C=0.3$的软间隔 SVM 模型中，如果一个样本$\boldsymbol{x}_i$对应的系数$\alpha_i=0.1$，试分析该样本与最优决策面和间隔区域的关系。

   答：样本$\boldsymbol{x}_i$对应的系数$\alpha_i=0.1$，即有$\alpha_i>0$,

   根据 KKT 条件$\alpha_i(1-\xi_i-y_if(x_i))=0$，推出$1-\xi_i-y_if(x_i)=0$，因此$\boldsymbol{x}_i$是一个支撑向量。

   其次，由于$0.1=\alpha_i<C=0.3$，根据 KKT 条件，$C=\alpha_i+\beta_i$推出$\beta_i>0$,

   再根据 KKT 条件$\beta_i\xi_i=0$，推出$\xi_i=0$。即松弛变量为 0，说明样本$\boldsymbol{x}_i$正好处在间隔区域的边界上。
10. 在非线性 SVM 的推导过程中，为什么要使用非线性映射的策略而不使用构造非线性决策面的策略呢?
11. 试说明，如何判断一个函数是不是核函数

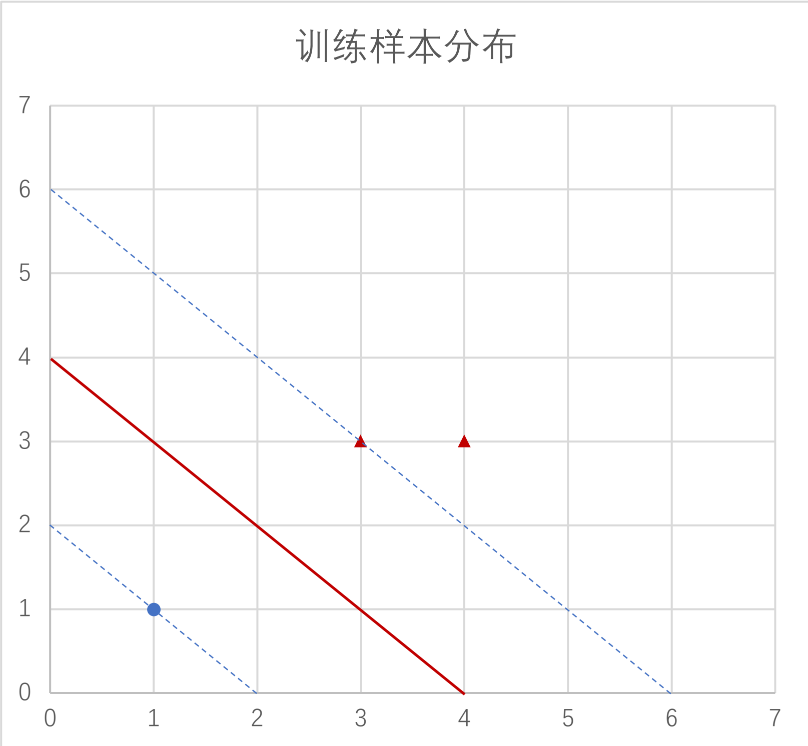
## 五、计算（画图）题

12. 求解有约束优化问题，并给出求解过程。

$$\min_{\boldsymbol{x}}[(x_1-2)^2+x_2^2]$$

$$\begin{aligned}\text{s.t. } & x_2-x_1+1=0\\ & {x_1}^2+x_2^2-1\le 0\end{aligned}$$

13. 已知一个如下图所示的训练数据集，其正类样本为$\boldsymbol{x_1}=(\mathbf{3},\mathbf{3})^{\boldsymbol{T}},\boldsymbol{x_2}=(\mathbf{4},\mathbf{3})^{\boldsymbol{T}}$，负类样本是$\boldsymbol{x_3}=(\mathbf{1},\mathbf{1})^{\boldsymbol{T}}$。




1） 试写出上述问题的线性 SVM 原始优化问题的数学形式，包含目标函数与约束条件；

答：

$$\min \frac{1}{2}\|\boldsymbol{\omega}\|^2$$
$$\text{s.t.} \quad 3\omega_1 + 3\omega_2 + \gamma \geq 1$$
$$4\omega_1 + 3\omega_2 + \gamma \geq 1$$
$$\omega_1 + \omega_2 + \gamma \leq -1$$

2） 试写出经过拉格朗日对偶处理后的优化问题的数学形式，包含目标函数与约束条件；

答：

$$\max \sum_{i=1}^{3} \alpha_i - \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} \alpha_i\alpha_j y_i y_j \boldsymbol{x}_i^T \boldsymbol{x}_j$$
$$s.t., \quad \sum_{i=1}^{3} \alpha_i y_i = 0$$
$$\alpha_i \geq 0, \quad i = 1,2,3$$

代入$y_1 = 1, y_2 = 1, y_3 = -1; \boldsymbol{x}_1 = [3,3]^T, \boldsymbol{x}_2 = [4,3]^T, \boldsymbol{x}_3 = [1,1]^T$

$$
\begin{aligned}
&\max(\alpha_1+\alpha_2+\alpha_3-9\alpha_1^2-12.5\alpha_2^2-\alpha_3^2-21\alpha_1\alpha_2+7\alpha_2\alpha_3+6\alpha_3\alpha_1)\\
&s.t., \quad \alpha_1+\alpha_2-\alpha_3=0\\
&\qquad\qquad \alpha_i\geq 0, \quad i=1,2,3
\end{aligned}
$$

3) 试写出经过拉格朗日对偶处理后的 3 条 KKT 条件。

答：对于每个样本$x_i$,应满足以下三条 KKT 条件

$$
\begin{cases}
\alpha_i\geq 0\\
y_i(\boldsymbol{\omega}^T\boldsymbol{x}_i+\gamma)-1\geq 0\\
\alpha_i(y_i(\boldsymbol{\omega}^T\boldsymbol{x}_i+\gamma)-1)=0
\end{cases}, \quad i=1,2,3
$$

代入$y_1=1,y_2=1,y_3=-1;\boldsymbol{x}_1=[3,3]^T,\boldsymbol{x}_2=[4,3]^T,\boldsymbol{x}_3=[1,1]^T$

得到三组 KKT 条件如下:

$$
\begin{cases}
\alpha_1\geq 0\\
3\omega_1+3\omega_2+\gamma-1\geq 0\\
\alpha_1(3\omega_1+3\omega_2+\gamma-1)=0
\end{cases},
\begin{cases}
\alpha_2\geq 0\\
4\omega_1+3\omega_2+\gamma-1\geq 0\\
\alpha_2(4\omega_1+3\omega_2+\gamma-1)=0
\end{cases},
\begin{cases}
\alpha_3\geq 0\\
-\omega_1-\omega_2-\gamma-1\geq 0\\
\alpha_3(\omega_1+\omega_2+\gamma+1)=0
\end{cases}
$$

4) 对照图中的决策面和间隔区域，试推算样本$\boldsymbol{x}_i,i=1,2,3$对应的系数$\alpha_i,i=1,2,3$。

答：首先根据图中的决策面位置，可以得出此时的决策面方程为:

$$x_1+x_2=4$$

因此有：$\omega_1=1,\omega_2=1,\gamma=-4;$

根据间隔宽度，可知样本$x_2$位于间隔之外，所以有

$$4\omega_1+3\omega_2+\gamma-1>0$$

根据（3）中的第二组 KKT 条件$\alpha_2(4\omega_1+3\omega_2+\gamma-1)=0$，可得$\alpha_2=0$

将$\alpha_2=0$代入（2）中的对偶问题中，消掉所有含有$\alpha_2$的项，得到:

$$
\begin{aligned}
&\max(\alpha_1+\alpha_3-9\alpha_1^2-\alpha_3^2+6\alpha_3\alpha_1)\\
&s.t., \quad \alpha_1-\alpha_3=0\\
&\alpha_i\geq 0, \quad i=1,3
\end{aligned}
$$

进一步的，由于$x_1,x_3$在间隔边界上，是支撑向量，所以有$\alpha_1=\alpha_3>0$

根据拉格朗日对偶过程中，广义拉格朗日函数对$\alpha_i$求导为 0，推出:

$$\boldsymbol{\omega}=\sum_{i=1}^{m}\alpha_i y_i\boldsymbol{x}_i$$

代入$\alpha_1=\alpha_3>0,\alpha_2=0$，有:

$$\begin{bmatrix}\mathbf{1}\\\mathbf{1}\end{bmatrix}=\alpha_1\begin{bmatrix}3\\3\end{bmatrix}+\alpha_1\begin{bmatrix}1\\1\end{bmatrix}$$

所以解得$\alpha_1=\alpha_3=\frac{1}{4},\alpha_2=0$

14. 对于 1 个标准的 XOR 问题，$\boldsymbol{x}_1=[1,0],\boldsymbol{x}_2=[0,1],\boldsymbol{x}_3=[0,0],\boldsymbol{x}_4=[1,1]$。其中$\boldsymbol{x}_1,\boldsymbol{x}_2\in\omega_1$类，$\boldsymbol{x}_3,\boldsymbol{x}_4\in\omega_2$类。请设计一个非线性矢量映射函数$\phi(\boldsymbol{x})$，将样本从 2 维空间映射到 3 位空间，使得数据在 3 维空间中可分，并给出对应的核函数。