Lecture7-8 作业

1，假设两个样本 $\{(\vec{V}_1, y_1) = ((v_1, v_2)^T, 1), (\vec{V}_2, y_2) = ((-v_1, -v_2)^T, -1)\}$, 假设 H 是这两个样本的最大间隔分类面，写出其表达式。

解：两个样本关于原点对称，最大间隔分类面会垂直于两个样本的连线，且穿过原点，即样本连线的斜率与分类面（分类线）斜率的乘积为-1，而样本连线的斜率为$\frac{v_2}{v_1}$，所以，分类面（线）的斜率为：$-\frac{v_1}{v_2}$，且 b=0。

所以，最大间隔分类面为：

$$x_2 = -\frac{v_1}{v_2} x_1$$

即：$v_1 x_1 + v_2 x_2 = 0$

2，假设三个样本为 $D = \{(\vec{x}_1, y_1) = ((3,0)^T, 1), (\vec{x}_2, y_2) = ((0,4)^T, 1), (\vec{x}_3, y_3) = ((0,0)^T, -1)\}$,计算这三个样本到平面：$x_1 + x_2 = 1$的距离。

解：$d = \frac{|\vec{w}^T \vec{x} + b|}{\|\vec{w}\|}$

$$x_1 + x_2 = 1 \rightarrow x_1 + x_2 - 1 = 0$$

$$d_1 = \frac{|\vec{w}^T \vec{x}_1 + b|}{\|\vec{w}\|} = \frac{\left|(1,1)\binom{3}{0} - 1\right|}{\sqrt{(1^2+1^2)}} = \sqrt{2}$$

$$d_2 = \frac{|\vec{w}^T \vec{x}_2 + b|}{\|\vec{w}\|} = \frac{\left|(1,1)\binom{0}{4} - 1\right|}{\sqrt{(1^2+1^2)}} = \frac{3}{2}\sqrt{2}$$

版权归SMART-AIA-HUST所有，严禁复制

$$d_3 = \frac{|\vec{w}^T \vec{x}_3 + b|}{\|\vec{w}\|} = \frac{\left|(1,1)\binom{0}{0} - 1\right|}{\sqrt{(1^2 + 1^2)}} = \frac{\sqrt{2}}{2}$$

3， 假设训练样本集为 $D = \{(\vec{x}_1, y_1) = ((0,0)^T, -1), (\vec{x}_2, y_2) = ((2,2)^T, -1), (\vec{x}_3, y_3) = ((2,0)^T, 1), (\vec{x}_4, y_4) = ((3,0)^T, 1)\}$,使用 QP 求解器时，$\vec{a}_n^T$(n=1,2,3,4)分别为多少？

解：$\vec{a}_1^T = (-1,0,0)$，$\vec{a}_2^T = (-1,-2,-2)$，$\vec{a}_3^T = (1,2,0)$，$\vec{a}_4^T = (1,3,0)$

4， 假设训练样本集为： $D = \{(\vec{x}_1, y_1) = ((1,1)^T, 1), (\vec{x}_2, y_2) = ((2,2)^T, 1), (\vec{x}_3, y_3) = ((2,0)^T, 1), (\vec{x}_4, y_4) = ((0,0)^T, -1), (\vec{x}_5, y_5) = ((1,0)^T, -1), (\vec{x}_6, y_6) = ((0,1)^T, -1)\}$, 请分别在 $y_n(\vec{w}^T \vec{x}_n + b) \geq 1$ 和 $y_n(\vec{w}^T \vec{x}_n + b) \geq 5$ 的条件下用 Primal SVM 方法来设计最优分类面 g($\vec{x}$)，判断两种情况下的分类面是否一致，指出哪些是候选的支撑向量，并回答如何确认哪些是支撑向量。

解：（1）对于条件 $y_n(\vec{w}^T \vec{x}_n + b) \geqslant 1$，可列出如下的式子

$$\min_{\boldsymbol{b},\boldsymbol{w}} \frac{1}{2} \boldsymbol{w}^T \boldsymbol{w}$$

$$s.t. \begin{cases} w_1 + w_2 + b \geqslant 1 \\ 2w_1 + 2w_2 + b \geqslant 1 \\ 2w_1 + b \geqslant 1 \\ -b \geqslant 1 \\ -w_1 - b \geqslant 1 \\ -w_2 - b \geqslant 1 \end{cases} \Longrightarrow \begin{cases} w_1 \geqslant 2 \\ w_2 \geqslant 2 \\ b \leqslant -3 \end{cases}$$

当且仅当 $w_1 = 2, w_2 = 2, b = -3$，

版权归SMART-AIA-HUST所有，严禁复制

$$\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}=\frac{1}{2}(w_1^2+w_2^2)\geqslant\frac{1}{2}(2^2+2^2)=4\text{取得最小值。}$$

可以验证 constraints 均满足。

故此时的最优分类面为

$$\boldsymbol{w}_1^T\boldsymbol{x}+b_1=0$$

其中$\boldsymbol{w}_1=[2\quad 2]^T, b_1=-3$。

可以验证，将$\boldsymbol{w}_1=[2\quad 2]^T, b_1=-3$代入上述 constraints 中有第 1、3、5、6 是严格等式，故候选支撑向量为$\boldsymbol{x}_1,\boldsymbol{x}_3,\boldsymbol{x}_5,\boldsymbol{x}_6$。

由 Dual SVM 知识可知，当求解 Dual SVM 问题时，在如下式子中

$$\alpha_n(1-y_n(\boldsymbol{w}^T\boldsymbol{x}_n+b))=0$$

$\boldsymbol{x}_1,\boldsymbol{x}_3,\boldsymbol{x}_5,\boldsymbol{x}_6$满足$\alpha_n>0, y_n(\boldsymbol{w}^T\boldsymbol{x}_n+b)=1$所对应的样本即为支撑向量。

（2）对于条件$y_n(\vec{w}^T\vec{x}_n+b)\geqslant 5$，可列出如下的式子

$$\min_{\boldsymbol{b},\boldsymbol{w}}\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}$$

$$s.t.\begin{cases}w_1+w_2+b\geqslant 5\\2w_1+2w_2+b\geqslant 5\\2w_1+b\geqslant 5\\-b\geqslant 5\\-w_1-b\geqslant 5\\-w_2-b\geqslant 5\end{cases}\Longrightarrow\begin{cases}w_1\geqslant 10\\w_2\geqslant 10\\b\leqslant -15\end{cases}$$

当且仅当$w_1=10, w_2=10, b=-15$时有

$$\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}=\frac{1}{2}(w_1^2+w_2^2)\geqslant\frac{1}{2}(10^2+10^2)=100\text{取得最小值，}$$

可以验证 constraints 均满足。

版权归SMART-AIA-HUST所有，严禁复制

故此时的最优分类面为

$$\boldsymbol{w}_2^T\boldsymbol{x}+b_2=0, which\ is\ exactly\ equivalent\ to\ \boldsymbol{w}_1^T\boldsymbol{x}+b_1=0$$

其中$\boldsymbol{w}_2=[10\quad 10]^T, b_2=-15$。

可以验证，将$\boldsymbol{w}_2=[10\quad 10]^T, b_2=-15$代入上述constraints中有第1、3、5、6是严格等式，故候选支撑向量为$\boldsymbol{x}_1,\boldsymbol{x}_3,\boldsymbol{x}_5,\boldsymbol{x}_6$。

由 Dual SVM 知识可知，当求解 Dual SVM 问题时，在如下式子中

$$\alpha_n(1-y_n(\boldsymbol{w}^T\boldsymbol{x}_n+b))=0$$

$\boldsymbol{x}_1,\boldsymbol{x}_3,\boldsymbol{x}_5,\boldsymbol{x}_6$满足$\alpha_n>0, y_n(\boldsymbol{w}^T\boldsymbol{x}_n+b)=5$所对应的样本即为支撑向量。

5，Hinge Loss 是支撑向量机的误差函数，因此，除了用二次规划求解最佳分类面外，也能用梯度下降法求解，（1）请推导梯度并写出算法流程；（2）假设初始增广权向量$\vec{w}=(0,0,0)^T$，用第 4 题训练样本集去设计分类面，指出哪些向量在边界上？假设它们都是支撑向量的话，请问最佳权系数向量是否是这些支撑向量的线性组合？

解：（1）已知样本集合$\{(\vec{x}_1,y_1),(\vec{x}_2,y_2),\dots,(\vec{x}_N,y_N)\}$，每个样本的标签为$y_n\in\{+1，-1\}$，我们基于 Hinge Loss，对于每个样本定义其误差函数为：

$$err_{SVM}=\max(0,1-y_n(\vec{w}^T\vec{x_n}+b))$$

对其求梯度，得到：

当$1-y_n(\vec{w}^T\vec{x_n}+b)\geq 0$，$\frac{\partial E_{in}(\vec{w})}{\partial\vec{w}}=-y_n\vec{x_n}$

当$1-y_n(\vec{w}^T\vec{x_n}+b)<0$，$\frac{\partial E_{in}(\vec{w})}{\partial\vec{w}}=0$

版权归SMART-AIA-HUST所有，严禁复制

利用随机梯度下降法得到新的$\vec{w}$

$$\vec{w}^{(t+1)}=\vec{w}^{(t)}-\eta\frac{\partial E_{in}\left(\vec{w}^{(t)}\right)}{\partial\vec{w}^{(t)}}=\vec{w}^{(t)}+\eta[\![1-y_n(\vec{w}^T\vec{x_n}+b)\geq 0]\!]y_n\vec{x_n}$$

其中，$[\![\cdot]\!]=\begin{cases}1,\ if\ condition\ is\ satisfied\\0,\ otherwise\end{cases}$

（2） 初始增广权向量$\vec{w}^{(0)}=(0,0,0)^T$

$\vec{x_1}=(1,1,1)^T,\vec{x_2}=(1,2,2)^T,\vec{x_3}=(1,2,0)^T,$

$\vec{x_4}=(1,0,0)^T,\vec{x_5}=(1,1,0)^T,\vec{x_6}=(1,0,1)^T$

$y_1=1,y_2=1,y_3=1,y_4=-1,y_5=-1,y_6=-1$

取学习率$\eta=1$

第一轮迭代

$$\max\left(0,1-y_1\left(\vec{w}^{(0)^T}\vec{x_1}\right)\right)=\max(0,1)=1$$

$$\frac{\partial E_{in}(\vec{w}^{(0)})}{\partial\vec{w}^{(0)}}=-y_1\vec{x_1}=(-1,-1,-1)^T$$

$$\vec{w}^{(1)}=\vec{w}^{(0)}-\eta\frac{\partial E_{in}\left(\vec{w}^{(1)}\right)}{\partial\vec{w}^{(1)}}=\vec{w}^{(0)}+y_1\vec{x_1}=(1,1,1)^T$$

第二轮迭代

$$\max\left(0,1-y_2\left(\vec{w}^{(1)^T}\vec{x_2}\right)\right)=\max(0,-4)=0$$

$$\vec{w}^{(2)}=\vec{w}^{(1)}=(1,1,1)^T$$

第三轮迭代

$$\max\left(0,1-y_3\left(\vec{w}^{(2)^T}\vec{x_3}\right)\right)=\max(0,-2)=0$$

$$\vec{w}^{(3)}=\vec{w}^{(2)}=(1,1,1)^T$$

第四轮迭代

版权归SMART-AIA-HUST所有，严禁复制

$$\max\left(0,1-y_4\left(\overrightarrow{w}^{(3)^T}\overrightarrow{x_4}\right)\right)=\max(0,2)=2$$

$$\frac{\partial E_{in}(\overrightarrow{w}^{(3)})}{\partial\overrightarrow{w}^{(3)}}=-y_4\overrightarrow{x_4}=(1,0,0)^T$$

$$\overrightarrow{w}^{(4)}=\overrightarrow{w}^{(3)}+y_4\overrightarrow{x_4}=(0,1,1)^T$$

第五轮迭代

$$\max\left(0,1-y_5\left(\overrightarrow{w}^{(4)^T}\overrightarrow{x_5}\right)\right)=\max(0,2)=2$$

$$\frac{\partial E_{in}(\overrightarrow{w}^{(4)})}{\partial\overrightarrow{w}^{(4)}}=-y_5\overrightarrow{x_5}=(1,1,0)^T$$

$$\overrightarrow{w}^{(5)}=\overrightarrow{w}^{(4)}+y_5\overrightarrow{x_5}=(-1,0,1)^T$$

第六轮迭代

$$\max\left(0,1-y_6\left(\overrightarrow{w}^{(5)^T}\overrightarrow{x_6}\right)\right)=\max(0,1)=1$$

$$\frac{\partial E_{in}(\overrightarrow{w}^{(5)})}{\partial\overrightarrow{w}^{(5)}}=-y_6\overrightarrow{x_6}=(1,0,1)^T$$

$$\overrightarrow{w}^{(6)}=\overrightarrow{w}^{(5)}+y_6\overrightarrow{x_6}=(-2,0,0)^T$$

第七轮迭代

$$\max\left(0,1-y_1\left(\overrightarrow{w}^{(6)^T}\overrightarrow{x_1}\right)\right)=\max(0,3)=3$$

$$\frac{\partial E_{in}(\overrightarrow{w}^{(7)})}{\partial\overrightarrow{w}^{(7)}}=-y_1\overrightarrow{x_1}=(-1,-1,-1)^T$$

$$\overrightarrow{w}^{(7)}=\overrightarrow{w}^{(6)}+y_1\overrightarrow{x_1}=(-1,1,1)^T$$

第八轮迭代

$$\max\left(0,1-y_2\left(\overrightarrow{w}^{(7)^T}\overrightarrow{x_2}\right)\right)=\max(0,-2)=0$$

$$\overrightarrow{w}^{(8)}=\overrightarrow{w}^{(7)}=(-1,1,1)^T$$

第九轮迭代

$$\max\left(0,1-y_3\left(\overrightarrow{w}^{(8)^T}\overrightarrow{x_3}\right)\right)=\max(0,0)=0$$

版权归SMART-AIA-HUST所有，严禁复制

$$\frac{\partial E_{in}(\vec{w}^{(8)})}{\partial \vec{w}^{(8)}} = -y_3\overrightarrow{x_3} = (-1,-2,0)^T$$

$$\vec{w}^{(9)} = \vec{w}^{(8)} + y_3\overrightarrow{x_3} = (0,3,1)^T$$

第十轮迭代

$$\max\left(0,1 - y_4\left({\vec{w}^{(9)}}^T\overrightarrow{x_4}\right)\right) = \max(0,1) = 1$$

$$\frac{\partial E_{in}(\vec{w}^{(9)})}{\partial \vec{w}^{(9)}} = -y_4\overrightarrow{x_4} = (1,0,0)^T$$

$$\vec{w}^{(10)} = \vec{w}^{(9)} + y_4\overrightarrow{x_4} = (-1,3,1)^T$$

第十一轮迭代

$$\max\left(0,1 - y_5\left({\vec{w}^{(10)}}^T\overrightarrow{x_5}\right)\right) = \max(0,1) = 1$$

$$\frac{\partial E_{in}(\vec{w}^{(4)})}{\partial \vec{w}^{(4)}} = -y_5\overrightarrow{x_5} = (1,1,0)^T$$

$$\vec{w}^{(11)} = \vec{w}^{(10)} + y_5\overrightarrow{x_5} = (-2,2,1)^T$$

第十二轮迭代

$$\max\left(0,1 - y_6\left({\vec{w}^{(11)}}^T\overrightarrow{x_6}\right)\right) = \max(0,0) = 0$$

$$\frac{\partial E_{in}(\vec{w}^{(11)})}{\partial \vec{w}^{(11)}} = -y_6\overrightarrow{x_6} = (1,0,1)^T$$

$$\vec{w}^{(12)} = \vec{w}^{(11)} + y_6\overrightarrow{x_6} = (-3,2,0)^T$$

第十三轮迭代

$$\max\left(0,1 - y_1\left({\vec{w}^{(12)}}^T\overrightarrow{x_1}\right)\right) = \max(0,2) = 2$$

$$\frac{\partial E_{in}(\vec{w}^{(12)})}{\partial \vec{w}^{(12)}} = -y_1\overrightarrow{x_1} = (-1,-1,-1)^T$$

$$\vec{w}^{(7)} = \vec{w}^{(6)} + y_1\overrightarrow{x_1} = (-2,3,1)^T$$

第十四轮迭代

版权归SMART-AIA-HUST所有，严禁复制

对于$\overrightarrow{x_2}$、$\overrightarrow{x_3}$、$\overrightarrow{x_4}$满足$1-y_n\left(\vec{w}^{(13)^T}\overrightarrow{x_n}\right)<0$

$\max\left(0,1-y_5\left(\vec{w}^{(13)^T}\overrightarrow{x_5}\right)\right)=\max(0,2)=2$

$\vec{w}^{(14)}=\vec{w}^{(13)}+y_5\overrightarrow{x_5}=(-3,2,1)^T$

第十五轮迭代

对于$\overrightarrow{x_6}$满足$1-y_n\left(\vec{w}^{(13)^T}\overrightarrow{x_n}\right)<0$

$\max\left(0,1-y_1\left(\vec{w}^{(14)^T}\overrightarrow{x_1}\right)\right)=\max(0,1)=1$

$\vec{w}^{(15)}=\vec{w}^{(14)}+y_1\overrightarrow{x_1}=(-2,3,2)^T$

第十六轮迭代

对于$\overrightarrow{x_2}$、$\overrightarrow{x_3}$、$\overrightarrow{x_4}$满足$1-y_n\left(\vec{w}^{(15)^T}\overrightarrow{x_n}\right)<0$

$\max\left(0,1-y_5\left(\vec{w}^{(15)^T}\overrightarrow{x_5}\right)\right)=\max(0,2)=2$

$\vec{w}^{(16)}=\vec{w}^{(15)}+y_5\overrightarrow{x_5}=(-3,2,2)^T$

检验对任意$\overrightarrow{x_n}$满足$1-y_n\left(\vec{w}^{(15)^T}\overrightarrow{x_n}\right)<0$，迭代结束

得到分类面为$2x_1+2x_2-3=0$

$\vec{w}=(-3,2,2)^T$

将 x_1、x_3、x_5、x_6 代 入 $1-y_n(\vec{w}^T\overrightarrow{x_n}+b)$均为 0，说明这四个样本在边界上，均为候选的支撑向量。

为简单起见（不用求解对偶 SVM），按照本题题意候选的支撑向量均为支撑向量，则：$\vec{w}=7x_1+0x_3-5x_5-5x_6$，即最佳权系数向量为支撑向量的线性组合。

版权归SMART-AIA-HUST所有，严禁复制

6，假如做了非线性变换后的两个训练样本为：$\{(\vec{Z}_1,+1)=(\vec{z},1),(\vec{Z}_2,-1)=(-\vec{z},-1)\}$，请写出用于设计硬间隔 SVM 时的拉格朗日函数$L(\overrightarrow{w},b,\alpha)$。

解：根据定义：

$$L(\overrightarrow{w},b,\alpha)=\frac{1}{2}\overrightarrow{w}^T\overrightarrow{w}+\alpha_1(1-y_1(\overrightarrow{w}^T\vec{z}_1+b)+\alpha_2(1-y_2(\overrightarrow{w}^T\vec{z}_2+b))$$

将两个样本代入，得到：

$$L(\overrightarrow{w},b,\alpha)=\frac{1}{2}\overrightarrow{w}^T\overrightarrow{w}+\alpha_1(1-(\overrightarrow{w}^T\vec{z}+b)+\alpha_2(1+(-\overrightarrow{w}^T\vec{z}+b))$$

7，对于一个单变量w，假设要在$w\geq 1$和$w\leq 3$这两个线性约束条件下，求$\frac{1}{2}w^2$的最小值，请写出其拉格朗日函数$L(w,\alpha)$以及这个最优问题的 KKT 条件。

解：由于是单变量，根据定义及约束条件：

$$L(w,\alpha)=\frac{1}{2}w^2+\alpha_1(1-w)+\alpha_2(w-3)$$

KKT 条件为：

$\alpha_1\geq 0$，$\alpha_2\geq 0$，

$w=\alpha_1-\alpha_2$，（通过$\frac{\partial L(w,\alpha)}{\partial w}=0$得到）

$\alpha_1(1-w)=0$，$\alpha_2(w-3)=0.$

8，假如做了非线性变换后的两个训练样本为：$\{(\vec{Z}_1,+1)=(\vec{z},1),(\vec{Z}_2,-1)=(-\vec{z},-1)\}$，在求解硬间隔 SVM 的对偶问题时，假定得到的最佳$\alpha_1>0$，最佳$\alpha_2>0$，请问最佳 b 为多少？

版权归SMART-AIA-HUST所有，严禁复制

解：由于$\alpha_1 > 0$，$\alpha_2 > 0$，所以：$\vec{Z}_1$和$\vec{Z}_2$为支撑向量，根据定义：

$$b = y_1 - \vec{w}^T\vec{Z}_1 = y_2 - \vec{w}^T\vec{Z}_2 = 1 - \vec{w}^T\vec{z} = -1 + \vec{w}^T\vec{z}$$

得到：$\vec{w}^T\vec{z} = 1$，$b = 0$

9，假设有 5566 个样本用以训练对偶硬间隔 SVM 时得到 1126 个支撑向量，请问落在分类面边界上的样本数（也就是候选的支撑向量）有可能是：（a）0；（b）1024；（c）1234；（d）9999。

解：　因为：支撑向量数≤候选的支撑向量数≤样本总数

　　所以选择（c）

10，如果两个样本$\vec{x}$和$\vec{x}'$的内积$\vec{x}^T\vec{x}' = 10$，计算其ϕ_2核函数$K_{\phi_2}(\vec{x},\vec{x}')$等于多少？

解：因为：$K_{\phi_2}(\vec{x},\vec{x}') = 1 + \vec{x}^T\vec{x}' + (\vec{x}^T\vec{x}')^2$

所以：　$K_{\phi_2}(\vec{x},\vec{x}') = 1 + 10 + 100 = 111$

11，假设训练样本集为：$\mathrm{D} = \{(\vec{x}_1, y_1) = ((1,0)^T, 1), (\vec{x}_2, y_2) = ((-1,0)^T, 1), (\vec{x}_3, y_3) = ((0,1)^T, -1), (\vec{x}_4, y_4) = ((0,-1)^T, -1)\}$，请用 Dual SVM 来设计最优分类面$\mathrm{g}(\vec{x})$，并指出哪些是支撑向量。

解：样本为非线性分布，所以，需要首先进行非线性变换：

令$\phi_2(\vec{x}) = \{1, x_1, x_2, x_1x_2, x_1^2, x_2^2\}$

则：$(\vec{x}_1, y_1) \rightarrow (\vec{z}_1, y_1)$: $\{(1,0)^T, 1\} \rightarrow \{(1,1,0,0,1,0)^T, 1\}$

$(\vec{x}_2, y_2) \rightarrow (\vec{z}_2, y_2)$: $\{(-1,0)^T, 1\} \rightarrow \{(1,-1,0,0,1,0)^T, 1\}$

版权归SMART-AIA-HUST所有，严禁复制

$$(\vec{x}_3, y_3) \to (\vec{z}_3, y_3)\colon \{(0,1)^T, -1\} \to \{(1,0,1,0,0,1)^T, -1\}$$

$$(\vec{x}_4, y_4) \to (\vec{z}_4, y_4)\colon \{(0,-1)^T, -1\} \to \{(1,0,-1,0,0,1)^T, -1\}$$

令$\alpha_1 \geq 0$、$\alpha_2 \geq 0$、$\alpha_3 \geq 0$、$\alpha_4 \geq 0$

由 SVM 对偶模型得到：

$$\begin{cases} L(\vec{w}, b, \alpha) = \dfrac{1}{2}\sum_{n=1}^{4}\sum_{m=1}^{4} \alpha_n \alpha_m y_n y_m \vec{z}_n^T \vec{z}_m - \sum_{n=1}^{4} \alpha_n \\ \sum_{n=1}^{4} y_n\, \alpha_n = 0 \end{cases}$$

求$L(\vec{w}, b, \alpha)$对α的梯度：$\frac{\partial L}{\partial \alpha_n} = \sum_{m=1}^{4} \alpha_m y_n y_m \vec{z}_n^T \vec{z}_m - 1$

且：$\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 = 0$

代入训练样本，

$$\frac{\partial L}{\partial \alpha_1} = 3\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 - 1 = 0 \to 2\alpha_1 - 1 = 0$$
$$\frac{\partial L}{\partial \alpha_2} = 3\alpha_2 + \alpha_1 - \alpha_3 - \alpha_4 - 1 = 0 \to 2\alpha_2 - 1 = 0$$
$$\frac{\partial L}{\partial \alpha_3} = 3\alpha_3 - \alpha_1 - \alpha_2 + \alpha_4 - 1 = 0 \to 2\alpha_3 - 1 = 0$$
$$\frac{\partial L}{\partial \alpha_4} = \alpha_4 - \alpha_1 - \alpha_2 + \alpha_3 - 1 = 0 \to 2\alpha_4 - 1 = 0$$

求解得到：$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{1}{2}$

$$\therefore\ \vec{w} = \sum_{n=1}^{4} \alpha_n y_n \vec{z}_n = \frac{1}{2}(\vec{z}_1 + \vec{z}_2 - \vec{z}_3 - \vec{z}_4) = (0,0,0,0,1,-1)^T$$

$$b = y_1 - \vec{w}^T \vec{z}_1 = 1 - (0,0,0,0,1,-1)(1,1,0,0,1,0)^T = 0$$

$$\therefore\ g_{SVM} = sign(\vec{w}^T \phi_2(\vec{x}) + b) = sign(x_1^2 - x_2^2)$$

且四个样本均为支撑向量。

版权归SMART-AIA-HUST所有，严禁复制