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$\{(\vec{x}_1,y_1),(\vec{x}_2,y_2),\ldots,(\vec{x}_N,y_N)\}$，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：

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

对其求梯度，得到：

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

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