1、拉格朗日函数

$$L(x,\alpha)=c^Tx+\alpha^T(Ax-b)$$

拉格朗日对偶函数

$$G(\alpha)=\inf_x L(x,\alpha)$$

由KKT条件得

$$\nabla_x L(x,\alpha)=0 \Rightarrow c+\alpha^T A=0$$

$\therefore$ $G(\alpha)=-\alpha^T b$

$\therefore$ 对偶问题为

$$\max\ G(\alpha)=-\alpha^T b$$

$$s.t.\quad \alpha\geq 0$$

2.1 $\forall w_1,w_2\in R^n$，$f(\frac{w_1+w_2}{2})=\frac{1}{4}(w_1+w_2)^T(w_1+w_2)$

$$=\frac{1}{4}(w_1^Tw_1+w_1^Tw_2+w_2^Tw_1+w_2^Tw_2)$$

$$\frac{1}{2}[f(w_1)+f(w_2)]=\frac{1}{2}(w_1^Tw_1+w_2^Tw_2)$$

$\because$ $f(\frac{w_1+w_2}{2})-\frac{1}{2}[f(w_1)+f(w_2)]=\frac{1}{4}(w_1^Tw_2+w_2^Tw_1-w_1^Tw_1-w_2^Tw_2)$

$$=\frac{1}{4}(w_1^T-w_2^T)(w_2-w_1)$$

$$=-\frac{1}{4}(w_2-w_1)^T(w_2-w_1)\leq 0$$

$\therefore$ $f(\frac{w_1+w_2}{2})\leq\frac{1}{2}[f(w_1)+f(w_2)]$，$f(w)$是凸的。

2.2. ① 正确。由于数据集是线性可分的，则一定存在两个边界

$$w^Tx+b=\pm 1$$

则所有点均在这两个边界的两侧，即

$$y^{(i)}(w^Tx^{(i)}+b)\geq 1$$

此时 $\xi_i=0$

② 有时软边距的SVM更好，因为硬边距的SVM会不可避免地产生过拟合的现象。而软边距的SVM增加了惩罚项，减少了数据集噪声的影响，在测试集上的表现可能更好。

2.3 ① 由KKT条件得

$$\alpha_i+\gamma_i=C,\ \forall i$$

对于 in-bound SVs， $0<\alpha_i<C$

$\therefore \gamma_i\neq 0$

又 $\because \gamma_i\cdot\xi_i=0,\ \forall i$ （KKT条件）

$\therefore \xi_i=0$

$\therefore$ 该向量满足 $y^{(i)}(w^T x^{(i)}+b)=1$，在边界上。

② 对于 bound SVs， $\alpha_i=C$

由KKT条件 $\alpha_i+\gamma_i=C,\ \forall i$ 得

$$\gamma_i=0$$

又 $\because \begin{cases}\xi_i\cdot\gamma_i=0,\ \forall i\\ \xi_i\geq 0,\ \forall i\end{cases}$

$\therefore \xi_i\geq 0$

又 $\because \alpha_i\left(y^{(i)}(w^T x^{(i)}+b)+\xi_i-1\right)=0$

$\Rightarrow y^{(i)}(w^T x^{(i)}+b)=1-\xi_i\leq 1$

$\therefore$ 对于该向量来说，它在边界上或边界内部。