HOMEWORK

Name: Guoqing Zhang Date: 2018/10/14 Student ID:2018214231
I did this homework entirely by myself.

1.1. (a)We knew that the optimal $b^*$ and $w^*$ match those conditions:

$$y_i(w^{*T}x_i + b^*) \geq 1.$$

And the support vector are the points which are the closest to the hyperplane ($y = w^{*T}x + b^*$), those points match the condition:$y(w^{*T}x + b^*) = 1$.
Now we suppose a point $(x_i, y_i)$.

If $y_i = 1$: $w^{*T}x_i + b^* \geq 1$. And $min_{y_i=1}(w^{*T}x_i + b^*) = 1$. so
$b^* = 1 - min_{y_i=1}w^{*T}x_i$ ...1.

If $y_i = -1$: $w^{*T}x_i + b^* \leq 1$. And $max_{y_i=-1}(w^{*T}x_i + b^*) = -1$. so
$b^* = -1 - max_{y_i=-1}w^{*T}x_i$...2.

Add both sides of the equation 1 and 2,we get:

$$b = -\frac{1}{2}(min_{y_i=1}w^{*T}x_i + max_{y_i=-1}w^{*T}x_i).$$

(b) Firstly,we know $w^* = \sum_{i=1}^{l} \alpha_i^* y_i x_i$,so:

$$\frac{1}{2}||w^*||^2 = \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i^*\alpha_j^* y_i y_j < x_i, x_j > .$$

In fact, consider the condition:

$$\sum_{i=0}^{l} \alpha_i^*[y_i(w^{*T}x_i + b*) - 1] = 0$$

Because of all $\alpha_i \geq 0$ and $y_i(w^{*T}x_i + b*)ge1$,so the factual condition can be even tighter:

$$\alpha_i^*[y_i(w^{*T}x_i + b*) - 1] = 0, i = 1, 2, ..., l$$

But I still cannot solve this problem.

1.2. (a) We know that $\ell(y, z) = max\{1 - yz, 0\}$,
so the question can be also described as:

minimize $\frac{1}{2}||w||_2^2 + C\sum_{i=1}^{l} \varepsilon_i$
subject to

$$\varepsilon_i = \left\{ \begin{array}{c} 0, 1 - y_i(w^T x + b) \leq 0 \\ 1 - y_i(w^T x + b), 1 - y_i(w^T x + b) \geq 0 \end{array} \right. , i = 1, 2, 3..., l$$

The old question is:

minimize $\frac{1}{2}||w||_2^2 + C\sum_{i=1}^{l} \varepsilon_i$
subject to

$$\varepsilon_i \geq 0, i = 1, 2, 3..., l$$

$$y_i(w^T x_i + b) \geq 1 - \varepsilon_i, i - 1, 2, ..., l$$

So $\varepsilon_i \geq 1 - y_i(w^T x + b)$, but our goal is to minimize, so we can be sure the $\varepsilon_i = 1 - y_i(w^T x + b)$ when $1 - y_i(w^T x + b) \geq 0$,otherwise $\varepsilon_i = 0$ .
So the solution of the one can also solve the other question.They are equivalent.

(b) My proof is as follows:

$$f\left(\theta w_1 + (1-\theta) w_2, \theta b_1 + (1-\theta) b_2\right)$$

$= 0.5\ \|\theta w_1 + (1-\theta)w_2\|^2 + C\sum_{i=1}^{l} \ell(y_i, (\theta w_1 + (1-\theta)w_2)_i^x + b)$
$= 0.5\ (\theta^2\|w_1\|^2 + 2\theta(1-\theta)\|w_1\|\|w_2\| + (1-\theta)^2\|w_2\|^2) + C\sum_{i=1}^{l} \ell(y_i, (\theta w_1 + (1-\theta)w_2)_i^x + b)$
$\leq 0.5(\theta^2\|w_1\|^2 + \theta(1-\theta)(\|w_1\|^2 + \|w_2\|^2) + (1-\theta)^2\|w_2\|^2) + C\sum_{i=1}^{l} \ell(y_i, (\theta w_1 + (1-\theta)w_2)_i^x + b)$
$= 0.5\ (\theta\|w_1\|^2 + (1-\theta)\|w_2\|^2) + C\sum_{i=1}^{l} \ell(y_i, (\theta w_1 + (1-\theta)w_2)_i^x + b)$
$\leq 0.5(\theta\|w_1\|^2 + (1-\theta)\|w_2\|^2) + C\sum_{i=1}^{l}[\ell(y_i, (\theta w_1)_i^x + b) + \ell(y_i, (1-\theta)w_2)]$

$$= \theta f(w_1, b_1) + (1-\theta) f(w_2, b_2)$$