Question 4.

For the SVM that is applied to data, which is linearly separable, we have made the following derivations

$$\begin{aligned}\max_{\alpha} W(\alpha) &= \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} y_i y_j \alpha_i \alpha_j x_i^T x_j \\ \text{s.t.} \quad & \alpha_i \geq 0, \quad , i = 1, \ldots, m \\ & \sum_{i=1}^{m} \alpha_i y^{(i)} = 0.\end{aligned}$$

But, for some cases it is not clear to see a hyper plane separating the data, since it is highly susceptible to outliers etc.
Therefore to make the model more robust to deal with these issues we need to re-formulate the regularization.

$$\begin{aligned}\min_{w,b,\xi} \ & \frac{1}{2}||w||^2 + C\sum_{i=1}^{m} \xi_i \\ \text{s.t.} \quad & y_i(w^T x_i + b) \geq 1 - \xi_i, \quad i = 1, \ldots, m \\ & \xi_i \geq 0, \quad i = 1, \ldots, m\end{aligned}$$

In making the above change, the only affect in our dual problem is that our initial constraint of $\alpha >= 0$, now becomes $0 <= \alpha <= C$. And in this case, the training points are segregated into three types which are described as follows,

- $\alpha = 0$: non-interesting points
- $C > \alpha > 0$; $\beta = 0$: a support vector on the margin line, no slack variable.
- $\alpha = C$; $\beta > 0$: a support vector, inside the side (or even misclassified)

Question 5.

$y = wx + b$;

$w = -1$;
$b = 1.5$

Margin ~ 0.5