VASISHT DUDDU
20151137

Theory Questions - Assignment 2

1. SVM can prevent overfitting by tuning the regularization parameter. The SVM is an approximate implementation of a bound of generalized error & decision boundary & margin depend on support vectors.
SVM is independent of dimensionality of feature space & hence using kernel methods to move to a higher dimension won't cause overfitting as SVM boundary is determined by support vectors & independent of dimension of space.

2. Given a data set of two pts x_1, x_2, the hyperplane can be obtined by

$$\arg\min \frac{1}{2}\|\omega\|^2$$

$$\text{s.t} \quad \omega^T x_1 + \omega_0 = 1$$
$$\omega^T x_2 + \omega_0 = -1$$

Using Laggrange's multipliers,

$$L(\omega) = \arg\min \frac{1}{2}\|\omega\|^2 + \alpha_1(\omega^T x_1 + \omega_0 - 1)$$
$$+ \alpha_2(\omega^T x_2 + \omega_0 + 1)$$

$$\frac{\partial L(\omega)}{\partial \omega} = 0$$

$$\frac{\partial L(\omega)}{\partial \omega_0} = 0$$

$$\Rightarrow \quad 0 = \omega + \alpha_1 x_1 + \alpha_2 x_2$$
$$0 = \alpha_1 + \alpha_2$$

$$\Rightarrow \quad \omega_0 = -\frac{\omega(x_1 + x_2)}{2}$$

We can determine the parameters without knowing α_i. Hence, using 2pts we can determine boundary

3. Class 1

$$\left\{ \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix} \right\}$$

Class 2

$$\left\{ \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ -2 \end{pmatrix}, \begin{pmatrix} -8 \\ -2 \end{pmatrix} \right\}$$

Since, data is linearly seperable we can use linear kernel, where mapping $\phi()$ is identity function.

Following are the support vectors,

$$\left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ -2 \end{pmatrix} \right\}$$

$$\quad x_3 \qquad x_5 \qquad x_7$$

$$\alpha_1 \phi(x_3) \phi(x_3) + \alpha_2 \phi(x_5) \phi(x_3) + \alpha_3 \phi(x_7) \phi(x_3) = -1$$

$$\alpha_1 \phi(x_3) \phi(x_5) + \alpha_2 \phi(x_5) \phi(x_5) + \alpha_3 \phi(7) \phi(x_5) = 1$$

$\alpha_1 \phi(x_3) \phi(x_7) + \alpha_2 \phi(x_5) \phi(x_7) + \alpha_3 \phi(7) \phi(x_7)$
$= 1$

$\Rightarrow \alpha_1 x_3 \cdot x_3 + \alpha_2 x_5 \cdot x_3 + \alpha_3 x_7 \cdot x_3 = -1$

$\alpha_1 x_3 x_5 + \alpha_2 x_5 x_5 + \alpha_3 x_7 x_5 = 1$

$\alpha_1 x_3 x_7 + \alpha_2 x_5 x_7 + \alpha_3 x_7 x_7 = 1$

$4\alpha_1 + 10\alpha_2 + -10\alpha_3 = -1$

$10\alpha_1 + 14\alpha_2 \quad -29\alpha_3 = 1$

$-10\alpha_1 - 29\alpha_2 \quad + 14\alpha_2 = 1$

We can find $\alpha_1, \alpha_2, \alpha_3$ & hence find the parameter:

$\bar{\omega} = \Sigma \alpha_i s_i$

$y = \omega x + b$

4.

x_1	x_2	y
-1	-1	-1
-1	$+1$	$+1$
$+1$	-1	$+1$
$+1$	$+1$	-1

This is linearly inseperable & hence to use SVM we have to map it to a higher dimension space.

We find basis which is
$1, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2, x_1^2, x_2^2$

In this space, the optimal hyperplane is at $x_1x_2 = 0$.

Hence, by moving to higher dimension space we can ~~find the~~ model XOR for SVM.