THEORY ASSIGNMENT

PRABHAT KUMAR
MT17036.

Ans 2:

Let there be 2 datapoints such that

$x_1 \in C_1 (y = +1)$ and $x_2 \in C_2 (y = -1)$, the max margin is determined by solving

$$\min_{\theta_0} \frac{1}{2}||\theta||^2$$

subject to,

$$\theta x_1 + \theta_0 = +1$$

$$\theta x_2 + \theta_2 = -1$$

By converting to Lagrange's form we get

$$\min_{\theta} \frac{1}{2}||\theta||^2 + \alpha_1(\theta x_1 + \theta_0 - 1) + \alpha_2(\theta x_2 + \theta + 1)$$

By taking derivative wrt θ. we get

$$0 = \theta + \alpha_1 x_1 + \alpha_2 x_2$$

$$0 = \alpha_1 + \alpha_2$$

$\therefore$ $\alpha_1 = -\alpha_2$ and $\alpha \cdot \theta = \alpha(x_1 - x_2)$

$\therefore$ we solve,

$$2\theta_0 = -\theta(x_1 + x_2)$$

Hence lagrange multiplier are endependent of θ.

* Refered from Andrew Ng Notes.

Ans 2: When we map data to higher dimension using RBF Kernel, we maxemise the margin of decision boundary using regularization which avoids overfitting of data

Ans 4 :

By using the kernel

$$K(x, x_i) = (c + x^T x_i)^2$$

with $c = 1$

$x = [x_1, x_2]^T$ $x_i = [x_{i1}, x_{i2}]$

$\rightarrow K(x, x_i) = 1 + x_1^2 x_{i1}^2 + x_2 x_{i2}^2 + x_2^2 x_{i2}^2 +$

$2x_1 x_{i1} + 2x_2 x_{i2} + 2x_1 x_2 x_{i1} x_{i2}$

$\therefore$ ~~Obtain~~ $\phi(x)$ can be obtained

$$\phi(x) = [1, x_i^2, x_2^2, \sqrt{x_1 x_2}, \sqrt{2}x_1, \sqrt{2x_2}]^T$$

Also $\phi(x_i) = [1, x_{i1}^2, x_{i2}^2, \sqrt{x_{i1}, x_{i2}}, \sqrt{2x_{i1}},$

$\sqrt{2x_{i2}}]^2$

Using the given kernel XOR fn can be successfully made linearly seperable.

~~Kernel based~~

$K = (1 + x_T x)^2$ is kernel needed for SVM

as XOR fn is not linearly seperable

~~For~~

Refered to Andrew Ng Notes for

$K(1 + x^T x)^2$