1\. a.

$$X = UV^T = [6\ 0\ 3\ 6]^T\ [4, 2, 1]$$

$$= \begin{bmatrix} 24 & 12 & 6 \\ 0 & 0 & 0 \\ 12 & 6 & 3 \\ 24 & 12 & 6 \end{bmatrix}$$

b.

Square Error $= \frac{1}{2} \sum (X_{ij} - y_{ij})^2$

$$= \frac{1}{2} \left[(24-5)^2 + (7-6)^2 + 2^2 + (12-4)^2 + (12-3)^2 + (6 \right.$$

$$= \frac{1}{2} \cdot 255.5 \times 2 = 255.5$$

c. Let $U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}$ Then, $UV^T = \begin{bmatrix} 4u_1 & 2u_1 & u_1 \\ 4u_2 & 2u_2 & u_2 \\ 4u_3 & 2u_3 & u_3 \\ 4u_4 & 2u_4 & u_4 \end{bmatrix}$

$$J(u_1) = \frac{1}{2} \left[(5-4u_1)^2 + (7-2u_1)^2 \right] + \frac{\lambda}{2} u_1^2$$

$$\frac{\partial J(u_1)}{\partial u_1} = (5-4u_1)(-4) + (7-2u_1)(-2) + 1 \cdot u_1$$

For $\frac{\partial}{\partial u_1} J(u_1) = 0$;

$$-20 + 16u_1 - 14 + 4u_1 + u_1 = 0$$

$$\Rightarrow -34 =$$

d. The value of u_1 will decrease to very low value.

1. c.

$$U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}, \quad UV^T = \begin{bmatrix} 4u_1 & 2u_1 & u_1 \\ 4u_2 & 2u_2 & u_2 \\ 4u_3 & 2u_3 & u_3 \\ 4u_4 & 2u_4 & u_4 \end{bmatrix}$$

$$J(u_1) = \frac{1}{2}[(5-4u_1)^2 + (7-u_1)^2] + \frac{1}{2}u_1^2$$

$$\frac{\partial J(u_1)}{\partial u} = (5-4u_1)(-4) + (7-u_1)(-1) + u_1$$

$$= -20 + 16u_1 - 7 + u_1 + u_1$$

$$18u_1 = 27$$

$$\Rightarrow u_1 = \frac{3}{2}$$

$$J(u_2) = \frac{1}{2}[(2u_2-2)^2] + \frac{u_2^2}{2}$$

$$= (2u_2-2)\cdot 2 + u_2$$

$$0 = 5u_2 - 4$$

$$\Rightarrow u_2 = 4/5$$

$$J(u_3) = \frac{1}{2}[(4u_3-4)^2] + \frac{u_3^2}{2}$$

$$\Rightarrow \frac{\partial J(u_3)}{\partial u_3} = (4u_3-4)4 + u_3$$

$$\Rightarrow u_3 = \frac{16}{17}$$

$$J(u_4) = \frac{1}{2}(2u_4-3)^2 + \frac{u_4^2}{2} + \frac{1}{2}(u_4-6)^2$$

$$= 2(2u_4-3) + u_4$$

$$\quad + u_4 - 6$$

$$\Rightarrow u_4 = 2$$

$$u = \begin{bmatrix} 3/2 \\ 4/5 \\ 16/17 \\ 2 \end{bmatrix}$$

Scanned by CamScanner

$$K(x,q) = (x^T q + 1)^2$$

$$= \left(\sum x_i q_i + 1\right)^2$$

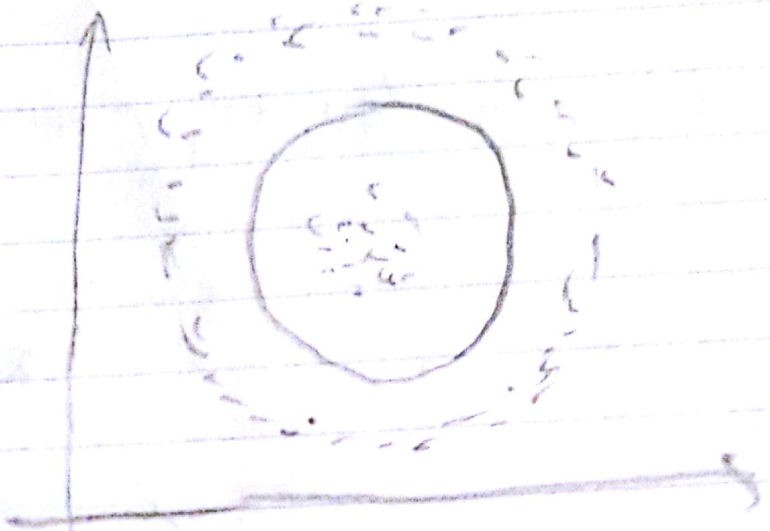
$$= (x_1 q_1 + x_2 q_2 + 1)^2$$

$$= x_1^2 q_1^2 + x_2^2 q_2^2 + 2x_1 x_2 q_1 q_2 + 2x_1 q_1 + 2x_2 q_2 + 1$$

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

$\sqrt{2}x_1x_2$ represents the measure when both x_1, x_2 appear simultaneously in bag of words vector.
ie if x_1, x_2 represents 2 words,
$\sqrt{2}x_1x_2 \neq 0$ when both are present.

$\phi(x)_3 = x_1^2 + x_2^2$ could have been used find x_3.



2

(1)

(a) Quadratic kernel. → hyperbolic

(b) Gaussian kernel → over fitted

(c) Linear kernel → straight line

(d) 3rd order kernel. → 3 turns in classifier

For value of λ, increasing the value of λ starts to underfit the data. Since (b) is over fitted (b) will increase it's performance on test d

3 Linear Regression

(a) (i) works like a eponential dataset $\phi(x) = e^x$

Exponential kernel will be best model.

(ii) piecewise function if $x > 0$, $\phi(x) = x - 1$

if $x < 0$, $\phi(x) = x + 1$

(b) (i) $L(\theta, \theta_0) = \sum_{i=1}^{n} (y^{(i)} - \theta \cdot x^{(i)} - \theta_0)^2 + \lambda \theta^2$

$$\frac{\partial L}{\partial \theta} = -x^{(t)} \sum_{t=1}^{n} (y^{(t)} - \theta \cdot x^{(t)} - \theta_0)^2 + 2\lambda\theta$$

$$\frac{\partial L}{\partial \theta_0} = -2\sum_{t=1}^{n} (y^{(t)} - \theta x^{(t)} - \theta_0)$$

(ii) For, $\frac{\partial L}{\partial \theta_0} = 0$.

$$-2\left(\sum_{t=1}^{n} y^{(t)} - \theta \cdot x^{(t)}\right) + 2n\theta_0 = 0$$

$$\Rightarrow \theta_0 = \frac{1}{n} \sum_{t=1}^{n} (y^{(t)} - \theta \cdot x^{(t)}$$

$$\frac{\partial L}{\partial \theta} = 2\lambda\theta - 2\sum_{t=1}^{n}\left(y^{(t)} - \theta x^{(t)} - \theta_0\right)x^{(t)}$$

$$= 2\lambda\theta - 2\sum_{t=1}^{n}\left(y^{(t)} - \theta x^{(t)} - \left[\frac{1}{n}\sum\left[y^{(t)} - \theta x^{(t)}\right]\right]\right)x^{(t)}$$

$$= 0$$

$$= \lambda\theta - \sum_{t=1}^{n} y^{(t)} \cdot x^{(t)} - \sum_{t=1}^{n} \theta x^{(t)} \cdot x^{(t)}$$

$$+ \sum_{t=1}^{n} \frac{1}{n}\sum\left(y^{(t)} - \theta x^{(t)}\right) x^{(t)} = 0$$

$$\lambda\theta - \frac{1}{n}\sum_{t=1}^{n}\sum_{t=1}^{n} \theta \cdot x^{(t)} \cdot x^{(t)} + \frac{1}{n}\sum_{t=1}^{n}\sum_{t=1}^{n} y^{(t)} x^{(t)}$$

$$- \sum_{t=1}^{n} y^{(t)} x^{(t)} + \theta\sum_{t=1}^{n}\left(x^{(t)}\right)^2 = 0$$

$$\Rightarrow \quad \boxed{\theta = \frac{\sum_{t=1}^{n} y^{(t)} \cdot x^{(t)} - \frac{1}{n}\sum_{t=1}^{n}\sum_{t=1}^{n} y^{(t)} \cdot x^{(t)}}{\lambda - \frac{1}{n}\sum_{t=1}^{n}\sum_{t=1}^{n} \theta \cdot x^{(t)} x^{(t)} + \sum_{t=1}^{n} (x^{(t)})^2}}$$