6.036 Pset 3

Kam Cho

1a) $X = U \times V = \begin{bmatrix} 6 \\ 2 \\ 3 \\ 3 \\ 5 \end{bmatrix} \begin{bmatrix} 4 & 1 & 5 \end{bmatrix}$

$$= \begin{bmatrix} 24 & 6 & 30 \\ 8 & \cancel{4}\,2 & 10 \\ 12 & 3 & 15 \\ 12 & 3 & 15 \\ 20 & 5 & 25 \end{bmatrix}$$

b) Squared error : $\sum_{(a,i)\in D} (Y_{ai} - [UV^T]_{ai}^2)/2$

$$[(5-24)^2 + (7-30)^2 + (2-2)^2 + (1-3)^2 + (4-15)^2 + (4-12)^2 + (5-3)^2 + (6-25)^2] \tfrac{1}{2}$$

$$= [(19)^2 + (23)^2 + 0^2 + (2)^2 + (11)^2 + (8)^2 + (2)^2 + (19)^2] \tfrac{1}{2}$$

$$= \boxed{722}$$

Regularization: $\frac{\lambda}{2}\left(\sum_{a=1}^{n}\sum_{j=1}^{k} U_{aj}^2 + \sum_{i=1}^{m}\sum_{j=1}^{k} V_{ij}^2\right)$

$k=1 \quad \lambda=1$

$$\rightarrow \frac{1}{2}\left(\sum_{a=1}^{n} U_a^2 + \sum_{i=1}^{m} V_i^2\right)$$

$$= \frac{1}{2}((36+4+9+9+25) + (16+1+25))$$

$$= \boxed{\frac{125}{2}}$$

c) Minimize $\sum_{i:(a,i)\in D} (Y_{ai} - u^{(a)} \cdot v^{(i)})^2/2 + \tfrac{1}{2}\|u^a\|^2$

Minimum an $u^{(a)}$ can go is zero

$\rightarrow (Y_{ai} - u^{(a)} v^{(i)})^2/2 + \tfrac{1}{2}\|u^a\|^2 = 0$

$\rightarrow$ solve for each $u^{(a)}$

$\Rightarrow$ $a=1$ $i=1$

$$\left(5-(u)(4)\right)^2/2+\tfrac{1}{2}\|u^2\|\geq 0$$

$$(25-40u+16u^2)+(u^2)\geq 0$$

combine.

$$17u^2\ 40u+25\geq 0$$

$$\cancel{\min(17u^2-40u+25) \text{ is when } u=\frac{-(-40)}{2(17)}=\frac{20}{17}}$$

$\Rightarrow$ $a=1$ $i=3$

$$(7-5u)^2+u^2+(5-4u)^2+u^2$$

$$= 49-70u+25u^2+25-40u+16u^2+2u^2$$

$$43u^2-110u+74=0$$

$$110^2<4(43)(74)$$

$\Rightarrow$ min $(43u^2-110u+74)$ is when

$$u=\frac{-(-110)}{2(43)}=\frac{55}{43}=u_1$$

u_2

$$\Rightarrow (2-2u_2)^2+(u_2^2)$$

$$4-8u_2+4u_2^2+u_2^2=4-8u_2+5u_2^2$$

$$\text{Min}(4-8u_2+5u_2^2)=\cancel{\frac{8}{2(5)}=\frac{4}{5}=u_2}$$

$$=\frac{4}{2(2)}=1=u_2$$

u_3

$$\Rightarrow (1-u_3)^2+(u_3^2)+(4-5u_3)^2+u_3^2$$

$$1-2u_3+u_3^2+u_3^2+16-40u_3+25u_3^2+u_3^2$$

$$28u_3^2-42u_3+17$$

$$\text{Min}(28u_3^2-42u_3+17) @ \frac{42}{2(28)}=\frac{21}{28}=\frac{3}{4}=u_3$$

Scanned by CamScanner

u_4

$(4-4u_4)^2+u_4^2$

$16-32u_4+16u_4^2+u_4^2$

$\text{Min}(16-32u_4+17u_4^2)$ @ $\frac{32}{2(17)} = \frac{16}{17} = u_4$

u_5

$(3-u_5)^2+u_5^2 + (6-5u_5)^2+u_5^2$

$9-6u_5+u_5^2+u_5^2+36-60u_5+25u_5^2+u_5^2$

$28u_5^2 - 66u_5 + 45$

$\text{Min}(28u_5^2-66u_5+45)$ @ $\frac{66}{2(28)} - \frac{33}{28} = u_5$

$\Rightarrow \boxed{U^{(1)} = [\,55/43,\ 1,\ 3/4,\ 16/17,\ 33/28\,]}$

2a) $x = \langle x_1, x_2\rangle \quad q = \langle q_1, q_2\rangle$

$\Rightarrow K(x,q) = (x_1 q_1 + x_2 q_2 + 1)^2$

$= (x_1q_1)^2 + (x_2q_2)^2 + 1 + 2(x_1q_1) + 2(x_2q_2)$

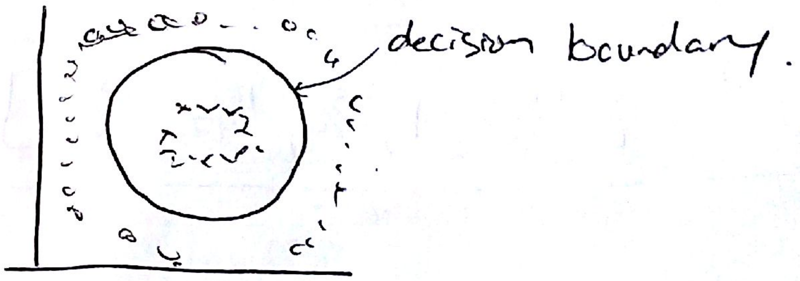
$+ 2(x_1q_1x_2q_2)$

$\Rightarrow$ see to make all x-terms in $\phi(x)$ vector.

$\Rightarrow \phi(x) = (x_1^2, x_2^2, x_1\sqrt{2}, x_2\sqrt{2}, x_1x_2\sqrt{2}, 1)$

b) Looking at the two graphs, choice 2 where $\phi(x)_3 = x_1^2 + x_2^2$ is the correct function this is because the smaller in magnitude both x_1 & x_2 were the closer to zero & the bigger → it was farther.

The decision boundary would be a plane cutting ~~down~~ across the middle part of the coin. in a 2-d graph it would be a circle.



3a) i) $\phi(x) = (x - .5)^2$ ii) $\phi(x) = x^3$

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

$$\nabla_\theta = \sum_{i=1}^{n} 2(y^t - \theta x^t - \theta_0)(-x^t) + 2\lambda\theta$$

$$\nabla_{\theta_0} = \sum_{i=1}^{n} 2(y^t - \theta x^t - \theta_0)(-1)$$

ii

$$0 = \sum_{i=1}^{n} y^t - \sum_{i=1}^{n} \theta x^t - \sum_{i=1}^{n} \theta_0$$

$$n\theta_0 = \sum_{i=1}^{n} y^t - \theta \sum_{i=1}^{n} x^t$$

$$\theta_0 = \frac{\sum_{i=1}^{n} y^t - \theta \sum_{i=1}^{n} x^t}{n}$$

ii) $0 = \sum_{i=1}^{n} y^t x^t - \theta \sum_{i=1}^{n} (x^t)^2 - \theta_0 \sum_{i=1}^{n} x^t - \lambda \theta$

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

$$\theta = \cancel{\frac{\sum_{i=1}^{n} (x^t)^2 - \left(\sum_{i=1}^{n} x^t\right)^2}{n}}$$

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

c) As λ increases, we have our linear regression be hurt in a negative way $\Rightarrow$ 2 will be worse than 1 & 4 worse than 3.

$\Rightarrow$ Plug in points (0,1), (1,1), (2,2) into above theta & theta_0

$\rightarrow$ get that

1) $\rightarrow$ c　　3) $\rightarrow$ d

2) $\rightarrow$ b　　d) $\rightarrow$ a

based on change in behaviour notice w/ first 2 pair.