Note Title 19-11-2020

COL 774. Minor Solution
Sem I, 2020-21

Q2. (a) Locally weighted linear Regression

$$J_x(\theta) = \frac{1}{2m}\sum_{i=1}^{m} w_i\left[y^{(i)} - h_\theta(x^{(i)})\right]^2$$

where

$$w_i = e^{-\frac{(x^{(i)}-x)^T\Sigma^{-1}(x^{(i)}-x)}{2}}$$

$$X = \begin{bmatrix} \text{---} & x^{(1)T} & \text{---} \\ & \vdots & \\ \text{---} & x^{(i)T} & \text{---} \\ & \vdots & \\ \text{---} & x^{(m)T} & \text{---} \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \vdots \\ \theta_n \end{bmatrix}$$

Then $$X\theta - Y = \begin{bmatrix} \theta^T x^{(1)} - y^{(1)} \\ \vdots \\ \theta^T x^{(i)} - y^{(i)} \\ \vdots \\ \theta^T x^{(m)} - y^{(m)} \end{bmatrix}$$

Define $W = \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_m \end{bmatrix}$

$\in \mathbb{R}^{m \times m}$

Then consider the expression

$$(X\theta - Y)^T \; W \; (X\theta - Y)$$

$$= \begin{bmatrix} | \\ \theta^T x^{(i)} - y^{(i)} \\ | \end{bmatrix}^T \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_m \end{bmatrix} \begin{bmatrix} | \\ \theta^T x^{(i)} - y^{(i)} \\ | \end{bmatrix}$$

$$= \sum_{i=1}^{m} w_i \left[\theta^T x^{(i)} - y^{(i)}\right]^2 = \left[J_x(\theta)\right] 2m$$

for locally weighted linear regression $\Rightarrow$ $$J_x(\theta) = \frac{1}{2m}(X\theta - Y)^T W (X\theta - Y)$$

(b) $J_x(\theta) = \frac{1}{2m} (X\theta - Y)^T W (X\theta - Y)$

$$= \frac{1}{2m}\left[(X\theta)^T W X\theta - Y^T W X\theta - (X\theta)^T W Y + Y^T W Y \right]$$

$$= \frac{1}{2m}\left[\theta^T X^T W X\theta - (X\theta)^T (Y^T W)^T - (X\theta)^T W Y + Y^T W Y \right]$$

$$= \frac{1}{2m}\left[\theta^T (X^T W X)\theta - \theta^T (X^T W^T Y) + Y^T W Y \right]$$

$$= \frac{1}{2m}\Big[\underbrace{\theta^T (X^T W X)}_{\text{symmetric matrix}}\theta - \theta^T (X^T W^T y) + y^T W y\Big]$$

$$\Rightarrow \nabla_\theta J_x(\theta) = \frac{1}{2m}\left[2(X^T W X)\theta - X^T W^T y + 0\right]$$

Note:-

A:- Symmetric

$\nabla_\theta \theta^T A \theta = 2A\theta$

$\nabla^2_\theta \theta^T A \theta = 2A$

And :-

$$\nabla^2_\theta J_x(\theta) = \frac{1}{\cancel{2}m}\left[\cancel{2}(X^T W X)\right] + 0$$

$$= \frac{1}{m}\left[X^T W X\right] = \text{Hessian matrix}$$

We need to show that Hessian matrix

is positive semi-definite for convexity of $J_x(\theta)$
(Strong)

$$\Rightarrow Z^T H Z \geq 0 \quad \forall Z \in \mathbb{R}^{(n+1)}$$

$$\Rightarrow \frac{1}{m} Z^T X^T W X Z \geq 0$$

$$\Rightarrow Z^T X^T W X Z \geq 0$$

$$\Rightarrow Z^T X^T W^{1/2\,T} W^{1/2} X Z \geq 0$$

m sized vector

$$\Rightarrow \left(W^{1/2} X Z\right)^T \left(W^{1/2} X Z\right) \geq 0$$

Aside I

$$W = \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & w_m \end{bmatrix}$$

$w_i \geq 0 \quad \forall i$

$$\Rightarrow W^{1/2} = \begin{bmatrix} \sqrt{w_1} & & \\ & \ddots & \\ & & \sqrt{w_m} \end{bmatrix} = W^{1/2\,T}$$

which is true. (inner product of two vectors)

Hence:- H matrix is +ve semi-definite

$\Rightarrow J_x(\theta)$ is convex ($\forall x$ since we did not make any assumptions on x)

Note:- Part (b) can also be solved by directly expanding

$$J_x(\theta) = \frac{1}{2m}\left[\sum_{i=1}^{m} \left[y^{(i)} - h_\theta(x^{(i)}) \right] w_i \right]$$

& computing second order derivatives component wise (rather than explicitly using matrix form) & thus showing the semi-definiteness.