9/22/2021 In Class Problem 2

Prove that the cost function $\|X\theta - Y\|_2^2$ is a convex function.

Proof : Given : $J(\theta) = \|X\theta - Y\|_2^2$

To prove that $J(\theta)$ is a convex function, we can take a double derivative of $J(\theta)$ to get the Hessian. If the hessian H is Positive Semi Definite, we can say that $J(\theta)$ is convex.

So, $\frac{\partial J(\theta)}{\partial \theta} = \frac{\partial(\|X\theta - Y\|_2^2)}{\partial \theta}$

$\therefore \frac{\partial J(\theta)}{\partial \theta} = 2X^T(X\theta - Y)$ [utilizing the property $\frac{\partial \|X\|_2^2}{\partial X} = 2X$, note the X^T is due to the chain rule]

taking derivative second time,

$H = \frac{\partial^2 J(\theta)}{\partial \theta} = 2X^TX$

Scanned with CamScanner

~~now, since X is a $R^{n\times d}$ matrix, X^TX will be a $d\times d$ square matrix.~~

~~now, utilizing the property that if A is any sq~~

Now X^TX is a symmetric matrix of dimension $d \times d$, We have to prove that X^TX is PSD.

Note that for a $R^{n\times n}$ matrix A to be PSD,

$$U^T A U \geq 0 \quad \text{for} \quad U \neq 0 \in R^n$$

$$\Rightarrow U^T(X^TX)U \geq 0$$

$$\Rightarrow (U^TX^T)(XU) \geq 0$$

$$\Rightarrow (XU)^T(XU) \geq 0$$

$$\Rightarrow \|XU\|_2 \geq 0 \qquad (\text{since } A^TA = \|A\| \text{ for any vector } A)$$

now, we know that the property of L-2 Norm is that it is always greater than or equal to zero.

hence we can say that

$\therefore \; u^T(X^TX)u \geq 0$

which means that X^TX is Positive-Semi Definite

Therefore $H = \frac{\partial^2(J(\theta))}{\partial^2(\theta)} = PSD$.

$$\boxed{\begin{array}{l}\text{Therefore, } J(\theta) = ||X\theta - Y||_2^2 \text{ is a convex function} \\ \text{Hence Proved.}\end{array}}$$

Scanned with CamScanner