Northeastern University
College of Engineering
Group member name(s): Yanming Liu
Group member UID(s): 002199402

May 20, 2022
Assignment 2

IE 7374: Machine Learning

Prove the cost function in linear regression, $\|X\theta - Y\|_2^2$ is a convex function.

Answer to Question

let $J(\theta) = \|X\theta - Y\|_2^2$, to prove the statement, i.e., hessian of $J(\theta)$, $H(J)$, is Positive Semi Definite.

Get $H(j)$

$$
\begin{aligned}
\nabla J(\theta) &= \frac{\partial((X\theta - Y)^T(X\theta - Y))}{\partial\theta} \\
&= \frac{\partial((\theta^T X^T - Y^T)(X\theta - Y))}{\partial\theta} \\
&= (\theta^T X^T - Y^T)\frac{\partial(X\theta - Y)}{\partial\theta} + (X\theta - Y)^T\frac{\partial(\theta^T X^T - Y^T)^T}{\partial\theta} \\
&= (\theta^T X^T - Y^T)X + (X\theta - Y)^T\frac{\partial(X\theta - Y)}{\partial\theta} \\
&= (\theta^T X^T - Y^T)X + (X\theta - Y)^T X \\
&= \theta^T X^T X - Y^T X + \theta^T X^T X - Y^T X \\
&= 2(\theta^T X^T - Y^T)X \\
\nabla^2 J(\theta) &= \frac{\partial(2(\theta^T X^T - Y^T)X)}{\partial\theta} \\
&= 2X^T X \\
&= H(J)
\end{aligned}
$$

Prove $H(J)$ **is PSD**
$\exists z \neq 0 \in \mathbb{R}^n$, let $z^T H(J) z = 2z^T X^T X z = 2\|Xz\|_2^2 >= 0$, thus $H(J)$ is PSD.

$J(\theta)$ is a convex function.