Date ___/___/___
Page ___

In Class Prob 8

10/8/2021 Prove if negative-log liklihood for Logistic Regression is convex.

=> The negative log-likelihood function is:

$$J(\theta) = -L(\theta) = \sum_{i=1}^{N} \log\left(1+e^{\theta^T x_i}\right) - y_i \theta^T x_i$$

Now to prove that this is convex, we can take double derivative of this funct to get the Hession and prove that it is PSD.

$$\therefore \frac{\partial J(\theta)}{\partial \theta} = \sum_{i=1}^{N} \frac{x_i e^{\theta^T x_i}}{1+e^{\theta^T x_i}} - y_i x_i$$

$$\frac{\partial J(\theta)}{\partial \theta} = \sum_{i=1}^{N} \left(\frac{1}{1+e^{-\theta^T x_i}} - 1\right) x_i - y_i x_i$$

$$H = \frac{\partial^2 J(\theta)}{\partial \theta} = \sum \frac{e^{-\theta^T x_i}}{(1+e^{-\theta^T x_i})^2} \cdot x x^T$$

$$= \sum (\hat{y}_i)(1-\hat{y}_i)\, x_i x_i^T$$

Now, H is PSD if for $\vec{v}$,

$$v^T H v \geq 0$$

$$v^T \left((\hat{y}_i)(1-\hat{y}_i)\, x_i x_i^T \right) v \geq 0$$

$$\therefore (\hat{y}_i)(1-\hat{y}_i)(v^T x^T)(x v) \geq 0.$$

$$\therefore (\hat{y}_i)(1-\hat{y}_i)\, \|x v\|_2^2 \geq 0$$

Now, $\hat{y}_i$, $(1-y_i)$ will always be ≥ 0,

Also $\|xv\|_2^2$ is norm-2 and will always be ≥ 0,

Date / /
Page

So $u^T H u \geq 0$, therefore H is PSD.

Hence the negative log-liklihood function is convex.