CS 760: Machine Learning Spring 2024

Homework 3: Logistic Regression

AUTHORS: Jed Pulley

DO NOT POLLUTE! AVOID PRINTING, OR PRINT 2-SIDED MULTIPAGE.

Problem 3.1

(a) I used good ol' fashioned guess and check. I set the iterations to 2000 and printed out the weights after every 100 iterations. Once the values started to level off, I considered that convergence

(b) It took me about 3000 iterations to converge to a respectable value

(c) I performed min-max normalization to avoid running into a RuntimeWarning (when the exponential got WAY too big), so my results for $\hat{\theta}$ were:

$$[-0.10924571, 0.03981236, -0.77742697, -0.01827487, 0.002808]$$

(d) I stored the likelihood for every iteration just to compare, which can be found in the likelihoods list in my *grad_ascent.py* file. With that said, the final value for the log-likelihood of $\hat{\theta}$ is: -0.6700727101571831

(e) From Theorem 6.2 in the Logistic Regression notes, we can see that $\hat{\theta} \longrightarrow \mathcal{N}(\theta^*, I_{\theta^*}^{-1})$ where the Fisher Information is shown as:

$$I_{\theta^*} = \sum_{i=1}^{N} \frac{e^{-\theta^{*T}\mathsf{x_i}}}{(1+e^{-\theta^{*T}\mathsf{x_i}})^2}\mathsf{x_i}\mathsf{x_i^T}$$

Problem 3.2

(a)

(b)

Problem 3.3

(a)

(b)

(c)

Problem 3.4

(a)

(b)

(c)