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 3000 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 2000 iterations to converge to a respectable value, although extending it to 3000 refined it just a bit more.

(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) The maximum log-likelihood of $\hat{\theta}$ is: -0.6700727101571831

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

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

Problem 3.2

(a) Borrowing from the Logistic Regression notes again, we can see that the MLE of the log-odds $\hat{\omega} := \hat{\theta}^\mathsf{T}\mathsf{x}$ where $\hat{\theta}$ are the true parameters, θ^*.

(b) Furthermore, the asymptotic distribution of $\hat{\omega}$ is defined as $\hat{\omega} \xrightarrow{\text{d}} \mathcal{N}(\theta^{*\mathsf{T}}\mathbf{x}, \mathbf{x}^\mathsf{T} I_{\theta^*}^{-1}\mathbf{x})$

Problem 3.3

(a)

(b)

(c)

Problem 3.4

(a)

(b)

(c)