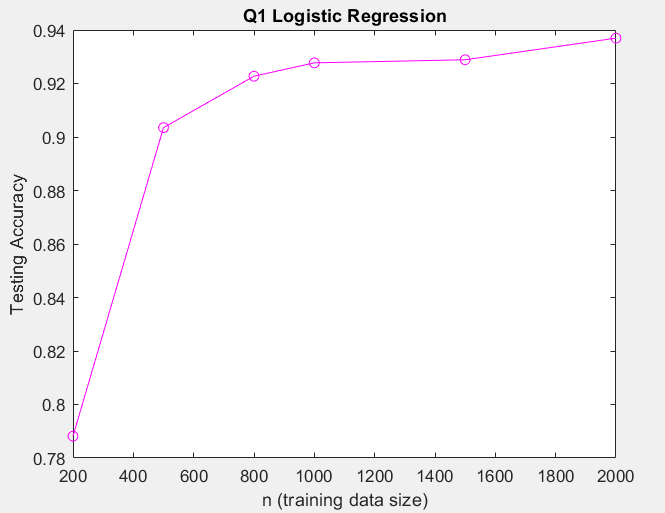
CSE 847 – HW4

Julia Zheng

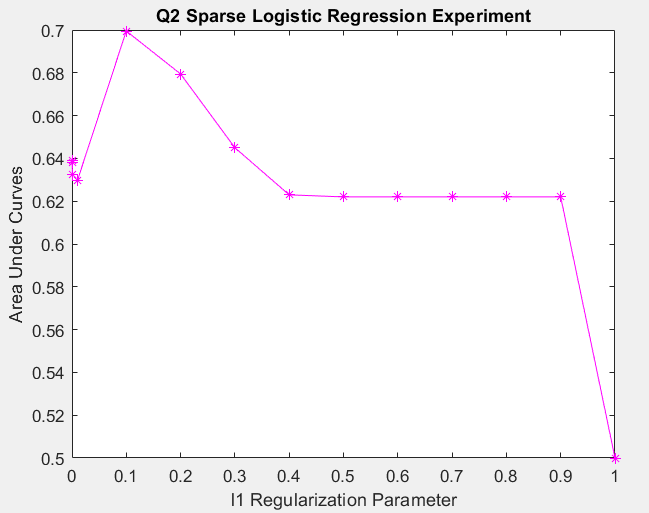
March 20, 2021

## Section 1: Logistic Regression Experiment



From figure 1, we see that logistic regression accuracy increases with number of training samples, assuming the samples are balanced, follow iid, and are otherwise well-chosen samples. Accuracy increases because the model is refined on more data samples and is thus more generalizable.

## Section 2: Sparse Logistic Regression Experiment



The above figure 2 looks are area under curves (AUC) comparing different l1 regularization parameters on the sparse logistic regression. The parameters chosen were [0, 0.0001, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0].

Between 0 and 0.01 the AUC steadily decreased from 0.64 to 0.63. The global optimum occurs at par = 0.1 with AUC reaching 0.7. From that point onwards, par values larger than 0.1 cause AUC to decrease until reaching a plateau between par equals 0.4 to 0.9 reach AUC of 0.62. When par = 1, the AUC reaches 0.5, demonstrating too much regularization leads to underfitting and is again not generalizable on the testing set.

Notice that when par = 0, the logistic regression doesn’t have regularization and has AUC = 0.64, which is higher than all other regularization parameters tested except for 0.1, 0.2, and 0.3. This demonstrates that choosing the regularization parameter well is important to the logistic regression’s generalizability and performance on testing sets.

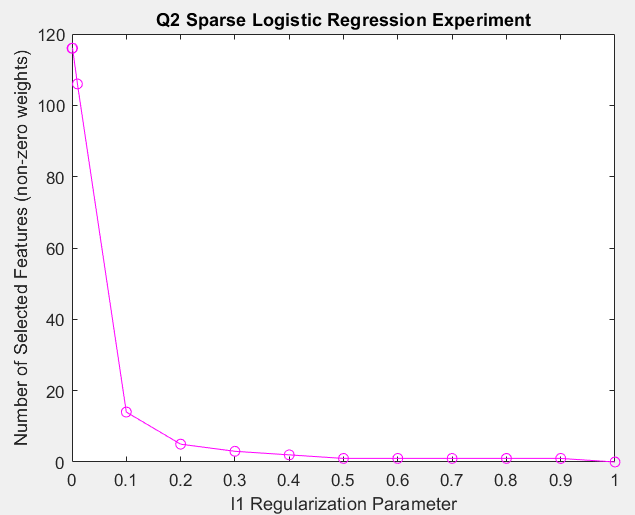


Figure 3 shows feature selection from l1 regularization of the logistic regression. This image shows a steady decline of number of features across the range of regularization parameters tested. Starting with par = 0, the full set of features started at 116 non-zero weights. Corresponding with increasing regularization parameter, the number of features selected decreases because the larger the regularization parameter, the smaller the weights w, such that when w is negligibly close to 0 the feature is considered not selected and the whole model attains measurable levels of sparsity. Additionally, we see that when regularization parameter par = 1, features selected = 0, and AUC = 0.5 because the model essentially becomes a coin flipper since there are no features selected so the binary label model is simply guessing at random instead of classifying.