**Logistic Regression with Regularizer**

L2 regularizer

|  |  |  |
| --- | --- | --- |
| Regularization strength | Classification error on test data | Number of 0’s in learned weight vector |
| 0.001 | 0.1172 | 0 |
| 0.01 | 0.1241 | 0 |
| 0.05 | 0.1241 | 0 |
| 0.1 | 0.1172 | 0 |

L1 regularizer

|  |  |  |
| --- | --- | --- |
| Regularization strength | Classification error on test data | Number of 0’s in learned weight vector |
| 0.001 | 0.1241 | 0 |
| 0.01 | 0.1241 | 0 |
| 0.05 | 0.1241 | 1 |
| 0.1 | 0.1655 | 3 |

For L1 regularizer, in our experiment, when , the classification errors on the test data are the same. When , the classification error increases. Therefore, when reaches a certain threshold, our performance becomes worse (we impose a too strong regularization and are underfitting). When , number of zeros in the learned weight vector is 0. When increases, number of zeros in the learned weight vector increases, because we are imposing a stronger regularization, that is, more truncation happens, because and have more chance to have different signs.