Logistic Regression with Regularizer

L_2 regularizer

Regularization strength λ	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

L_1 regularizer

Regularization strength λ	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 L_1 regularizer, in our experiment, when $\lambda = 0.001, 0.01, \text{and } 0.05$, the classification errors on the test data are the same. When $\lambda = 0.1$, the classification error increases. Therefore, when λ reaches a certain threshold, our performance becomes worse (we impose a too strong regularization and are underfitting). When $\lambda = 0.001, 0.01$, number of zeros in the learned weight vector is 0. When λ increases, number of zeros in the learned weight vector increases, because we are imposing a stronger regularization, that is, more truncation happens, because $w'(t+1)$ and $w(t+1)$ have more chance to have different signs.