# 10-601: Machine Learning Homework 3

Jiashi Cao(Andrew ID: jiashic)

Carnegie Mellon University

## I. POLICY ON COLLABORATION

(1) Did you receive any help whatsoever from anyone in solving this assignment? Yes / No. If you answered yes, give full details: (e.g. Jane explained to me what is asked in Question 3.4)

Answer: No.

(2) Did you give any help whatsoever to anyone in solving this assignment? Yes / No. If you answered yes, give full details: (e.g. I pointed Joe to section 2.3 to help him with Question 2.

Answer: No.

## II. LOGISTIC REGRESSION WITH $L_2$ REGULARIZATION

(1) Find $\lambda$ Based on 10-fold Cross-Validation

In order to find the best regularization coefficient $\lambda$, I implemented 10-fold Cross-Validation(CV) on training dataset. To do this, I tried $\lambda = 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}$, 0.01, 0.1, 0.2, 0.5, and 1 respectively. Run 10-fold Cross-Validation(CV) using each $\lambda$ respectively, and got the averaged results of each fold in CV. And then plot the relationship between accuracy and $\lambda$s, results are shown in Fig. 1 and Fig. 2(using $log_{10}$ of $\lambda$ as X-axis ).

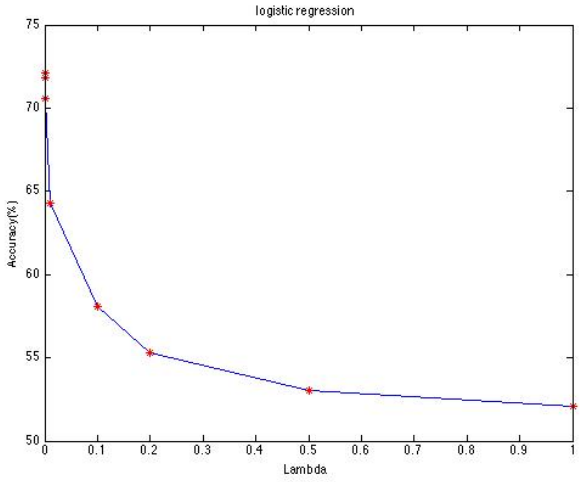


Fig. 1. Relationship between Accuracy and Regularization Coefficient $\lambda$

As it shown in Fig. 1 and Fig. 2, we can clearly see that accuracy is decreasing as the value of $\lambda$ increases when $\lambda$ larger than $10^{-4}$ based on training dataset. But the value of accuracy doesn't change obviously when $\lambda$ is less than $10^{-4}$. And it will consume more time using a very small $\lambda$. So in this

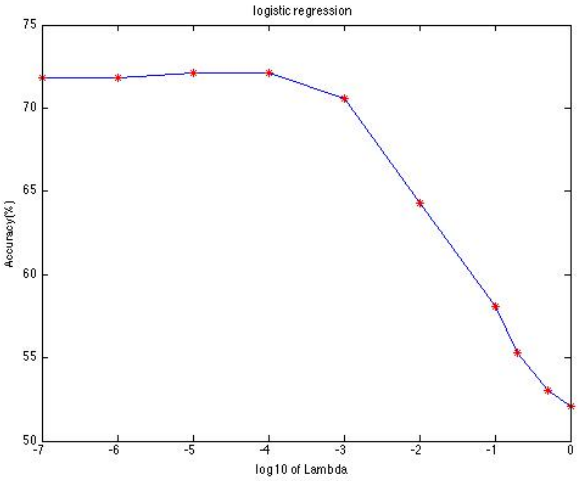


Fig. 2. Relationship between Accuracy and $log_{10}\lambda$

case, $\lambda = 10^{-4}$ seems to be the best regularization coefficient as long as I have tried. Therefore, I chose $10^{-4}$ as $\lambda$ in lr_train function.

(2)Test Logistic Regression Algorithm

By using $\lambda = 10^{-4}$, I tested my model with Xtest and Ytest. The result is :

P:0.762, R:0.713, A:0.727

Since the algorithm ramdomly shuffles training set everytime, the results of each time running the algorithm will be slightly different. However, all the test results of accuracy are around 72% when $\lambda$ is $10^{-4}$.