South China University of Technology

The Experiment Report of Machine Learning

School: School of Software Engineering

Subject: Software Engineering

Author:
Yuehao Xiao

Supervisor:
Mingkui Tan

Student ID:
201630719984

Grade:
Undergraduate

October 21, 2018

Linear Regression and Stochastic Gradient Descent

Abstract—Two simple methods of linear regression, closed-form solution and stochastic gradient descent.

I. Introduction

IN this report, we will try two methods to train the program to get the unknown value y according to x, based on linear regression.

II. Methods and Theory

We use two methods for linear regression, closed-form solution and stochastic gradient descent. In closed-form solution, we try to write down the differential equation of the loss function, and then calculate the optimal parameters ω making the loss minimize. In stochastic gradient descent, we use a descent parameter ω and stochastic batch of x to make the result approaching the right answer.

III. Experiments

A. Dataset

Linear Regression uses Housing in LIBSVM Data, including 506 samples and each sample has 13 features. We divide it into training set, validation set as a partition of 0.33.

B. Closed-form solution

With samples:
$$x_1, x_2 \ldots x_n$$

Desired outputs:
$$y_i, y_2 \ldots y_n$$

Hypothesis:
$$h(x) = X\omega$$

Loss function:
$$\mathcal{L}(\omega) = \frac{1}{2}(y - X\omega)^T(y - X\omega)$$

Differential:
$$\frac{\partial \mathcal{L}(\omega)}{\partial \omega} = -X^T y + X^T X\omega$$

Suppose:
$$\frac{\partial \mathcal{L}(\omega)}{\partial \omega} = 0$$

Optimal parameter ω:
$$\omega = (X^T X)^{-1} X^T y$$

Result:

TABLE I
Loss

Loss	29.0297386
Loss_train	9.8499180
Loss_test	9.0891477

C. Stochastic gradient descent

With samples:
$$x_1, x_2 \ldots x_n$$

Desired outputs:
$$y_i, y_2 \ldots y_n$$

Initial weights:
$$\omega_1, \omega_2, \ldots \omega_b \text{ set to } [0, 0 \ldots 0]$$

Hypothesis:
$$h(x) = X\omega$$

Loss function:
$$\mathcal{L}(\omega) = \frac{1}{2}(y - X\omega)^T(y - X\omega)$$

Learning rate:
$$\eta = 0.0004$$

Descent:
$$d = \frac{1}{2}\omega + X^T(X\omega - y)$$

Weights:
$$\omega = \omega - \eta d$$

Result:

TABLE II
Loss

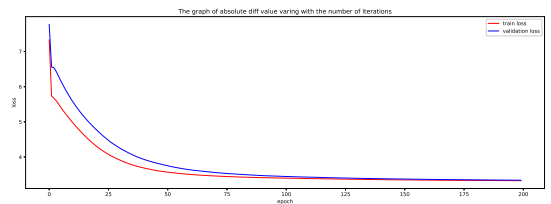| Loss_train	3.2926055
Loss_val	3.4676981

Fig. 1. Result of training and validation

IV. Conclusion

In this report, we conduct some experiments under small scale datasets, realize the process of optimization and adjusting parameters. Finally, we have further understand of linear regression, closed-form solution and stochastic gradient descent.