

Linear Regression And Polynomial Regression

Wang Zongwei

Abstract—This report is for the research on Linear regression and polynomial regression to simulate model,we need to use Python to achieve.DataSet is NACA0012 airfoils.

Index Terms—linear regression,polynomial regression

1 INTRODUCTION

For the data set NACA0012 airfoils,we use the linear regression,polynomial regression respectively to deal with,simulating a good enough model.

1.1 Linear Regression

We use the gradient descent method to achieve linear regression. Gradient descent method is an optimization algorithm, usually called the steepest descent method. The steepest descent method is one of the simplest and oldest methods for solving unconstrained optimization problems. Although it is not practical now, many effective algorithms are based on it and are improved and corrected. The steepest descent method is to use the negative gradient direction for the search direction, steepest descent method closer to the target value, the smaller the step, the slower the progress.

1.2 Polynomial Regression

We use the Particularity that the polynomial regression have only unique solution,and add regularization term to punish to prevent over fitting.

2 ALGORITHM

2.1 Linear Regression

Use mainly on Gradient descent algorithm:

enter:training set,y,alpha,loopmax (train set storage 5 feature value,y is result,alpha is the steps,loopmax is the cycles)

output:forecast and actual square error

(1)Read the data from file,and storage

(2)Gradient reduction of the data.(use the random gradient reduction) A.generate a theta,initialize it.Generate a random number,get the train set value and the y value,calculate its gradient. B.use the theta subtract gradient,get the new theta. C.cycle a,b,still to the end.return the last theta

(3)use theta predict the result,calcuete the Square difference.Return the Square difference

2.2 Polynomial Regression

Polunomial Regression is also use the gradient descent algprithm,but it add the regularization term to publish its overfitting:

enter:training set,loopmax,the first paradigm coefficient,the second paradigm coefficient,rate,degree (train set storage 5 feature value,y is result,alpha is the steps,loopmax is the cycles,degree is the polynomial's order)

output:forecast and actual square error

(1)Read the data from file,and storage

(2)Scale the matrix of trainset

(3)use the formula,to calculate the x of polynomial from the 5 feature value.

$$X_i = \sum_{k=1}^n A_k^{degree}$$

(4)Then use gradient reduction with regularization term to calculate the coefficient,w0,w1,w2...

(5)use the w to calculate the formula that adds the regularization term to calculate the square error,as follows:

$$\min \frac{1}{N} \sum_{i=1}^N L(y_i, f(x_i)) + \lambda J(f)$$

3 CONCLUSION

If you really use the method of mine to calculate.We could find the result's deviation is too large.So i think the algorithm could be optimization.However,we could be familiar with the Gradient descent method and regularization term use for the Linear Regression and Polynomial Regression.