Thomas Chung
Fang
12 February 2018

Homework 3

Directions:

Perform linear regression directly using the closed form solution. Compute the RMSE value on the training data and test data, respectively.

Perform ridge regression directly using the closed form solution. Use k-fold cross validate (k=5) to select the optimal λ parameter. Compute the RMSE value on the test data.

You can begin by running the solver with λ = 400. Then, cut λ down by a factor of 2 and run again. Continue the process of cutting λ by a factor of 2 until you have models for 10 values of λ in total.

Perform linear regression using the gradient descent algorithm. Compute the RMSE value on the training data and test data, respectively.
For the initial weights, you can just use Gaussian N(0, 1) random variables. Define "converging" as the change in any coefficient between one iteration and the next is no larger than 10–5.

Perform ridge regression using the gradient descent algorithm. Compute the RMSE value on the test data.

Output:

The following output is what happens when I run my implementation of linear regression, ridge regression, and the gradient descent versions of the regression models.

Linear Regression RSME: 0.14583464490949063
Linear Regression Gradient Descent RSME: 0.4539722799005858
Ridge Regression RSME: 0.14574650707058057
Ridge Regression Gradient Descent RSME: 0.3416995574819495

This output was achieved by initializing my gradient descent models with a weight vector of *Gaussian N(0,1)*.

The **next output** is what happens when we lower the tolerance 'epsilon' (from 1e-5 to 1e-7) of our gradient descent models in an attempt to achieve a better RSME value.

Linear Regression RSME: 0.14583464490949063
Linear Regression Gradient Descent RSME: 0.165161997358132
Ridge Regression RSME: 0.14574650707058057
Ridge Regression Gradient Descent RSME: 0.14664612886453104

And the next **final output** is what happens when we lower the tolerance 'epsilon' (from 11e-5 to 1e-7) and raise our step size 'alpha' (from 0.001 to 0.01) in order to achieve a reasonable calculation speed:

Linear Regression RSME: 0.14583464490949063
Linear Regression Gradient Descent RSME: 0.148927834276187
Ridge Regression RSME: 0.14574650707058057
Ridge Regression Gradient Descent RSME: 0.1455565675016494

As we can see, modifying our step size and tolerance parameters allowed our gradient model to approach the RSME value of the close form solutions, but at the cost of our processing speed.