Results for Ridge Regression

RESULTS for ridge regression using Closed Form

    1) optimal Theta

n = 1: theta = [[24.83562613 14.1254215]]

n = 2: theta = [[30.58703919 13.4976292 -0.32367572]]

n = 3: theta = [[9.84129188 -0.8605458   1.40263187 0.51037625]]

n = 5: theta = [[1.06879498e+01   1.02301402e-01   1.13584026e+00   4.15127537e-01   9.05648400e-03
1.99464784e-03]]

    2) errors using Closed Form

n = 1:	training: 894.298458863	test: 1464.25842751
n = 2:	training:868.208697501	test: 1746.52302338
n = 3:	training: 44.2518755788	test: 210.528764077
n = 5	training:42.3717141141	test: 42.3717141141

RESULTS for ridge regression using Stochastic Gradient Descent

1) optimal Theta for training set

size of batch = 5:

Loss iteration 0: 530.439672393	Loss iteration 50: 568.59490209
Loss iteration 100: 568.59490209	Loss iteration 150: 568.59490209
Loss iteration 200: 568.59490209	Loss iteration 250: 568.59490209
Loss iteration 300: 568.59490209	Loss iteration 350: 568.59490209
Loss iteration 400: 568.59490209	Loss iteration 450: 568.59490209
Loss iteration 500: 568.59490209	Loss iteration 550: 568.59490209
Loss iteration 600: 568.59490209	Loss iteration 650: 568.59490209
Loss iteration 700: 568.59490209	Loss iteration 750: 568.59490209
Loss iteration 800: 568.59490209	Loss iteration 850: 568.59490209
Loss iteration 900: 568.59490209	Loss iteration 950: 568.59490209

theta = [[23.63874516] [17.52631251]]

2)optimal Theta for test set :

size of batch = 5:

Loss iteration 0: 1202.41479538   Loss iteration 50: 892.741483051

Loss iteration 100: 567.045346643   Loss iteration 150: 536.824609517

Loss iteration 200: 694.091881553   Loss iteration 250: 963.676817026

Loss iteration 300: 1293.93871141   Loss iteration 350: 1650.02246793

Loss iteration 400: 2008.97274869   Loss iteration 450: 2356.20415642

Loss iteration 500: 2682.95965009   Loss iteration 550: 2984.48799283

Loss iteration 600: 3258.7433574   Loss iteration 650: 3505.46324798

Loss iteration 700: 3725.51976453   Loss iteration 750: 3920.46770475

Loss iteration 800: 4092.2338361   Loss iteration 850: 4242.90690979

Loss iteration 900: 4374.59912358   Loss iteration 950: 4489.3578661

theta = [[107.24810412] [-9.56858099]]

Errors:

1)-Training data error: 568.59490209

2)Testing data error: 4587.25265916

Using k- fold cross validation:

Degree of Basis Function = 2

optimal Value for lambda: 0.1   Obtained Training Error = 46081.9238303

Choosing fold: Training with 2 folds   Optimal weight vector = [[27.6947757   14.50069733]]

Reported Testing error = 10527.1107464

optimal Value for lambda: 0.1   Obtained Training Error = 9359.53653512

Choosing fold: Training with 10 folds   Optimal weight vector = [[25.65949065 13.89902635]]

Reported Testing error = 11651.7630755

optimal Value for lambda: 0.1   Obtained Training Error = 938.725088318

Choosing fold: Training with 100 folds Optimal weight vector = [[24.55809178 14.18222542]]

Reported Testing error = 11755.7876427

Degree of Basis Function = 3

optimal Value for lambda: 0.1

Choosing fold: Training with 2 folds

Reported Testing error = 10527.1107464

Obtained Training Error = 46081.9238303

Optimal weight vector =[[27.6947757 14.50069733]]

optimal Value for lambda: 0.1

Choosing fold: Training with 10 folds

Reported Testing error = 11651.7630755

Obtained Training Error = 9359.53653512

Optimal weight vector = [[25.65949065 13.89902635]]

optimal Value for lambda: 0.1

Choosing fold: Training with 100 folds

Reported Testing error = 11755.7876427

Obtained Training Error = 938.725088318

Optimal weight vector = [[24.55809178 14.18222542]]

Degree of Basis Function = 5

optimal Value for lambda: 0.1

Choosing fold: Training with 2 folds

Reported Testing error = 10527.1107464

Obtained Training Error = 46081.9238303

Optimal weight vector = [[27.6947757   14.50069733]]

optimal Value for lambda: 0.1

Choosing fold: Training with 10 folds

Reported Testing error = 11651.7630755

Obtained Training Error = 9359.53653512

Optimal weight vector = [[25.65949065 13.89902635]]

optimal Value for lambda: 0.1

Choosing fold: Training with 100 folds

Reported Testing error = 11755.7876427

Obtained Training Error = 938.725088318

Optimal weight vector = [[24.55809178 14.18222542]]

Test error is a convex function of both Lambda and n. Therefore, it can run into underfitting and overfitting scenarios as visible.