

Assignment 2

Anurag Behera

Uttam Singh

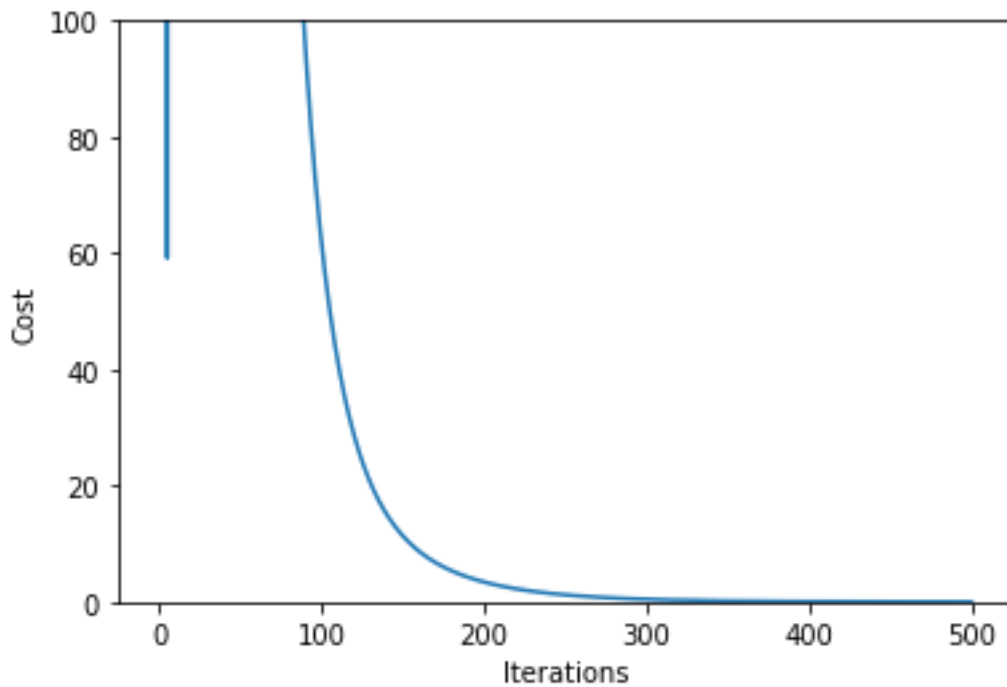
POLYNOMIAL REGRESSION

Gradient Descent Method

Data is divided into training and validation sets. Training data contains 80 % of the original data whereas validation data that is used for testing contains the other 20%. R2 and RMSE score are calculated on the testing/validation set

Degree 1 Parameters

Learning Rate 0.1



(Note: Here cost ranges from 0 to 100)

Results

Training cost after 500 iteration 0.17856538647442677

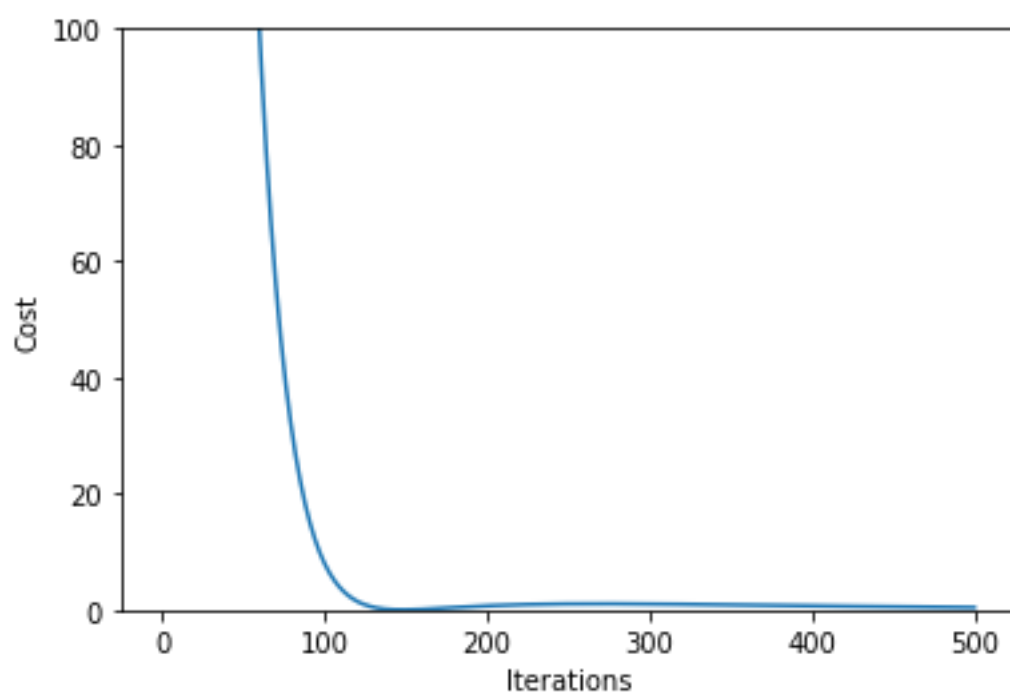
R2 score	0.12245364
RMSE	0.8676320845527065

Degree 2

Parameters

Learning Rate 0.1

Plots



Results

Training cost after 500 iteration 0.0108685791227500445

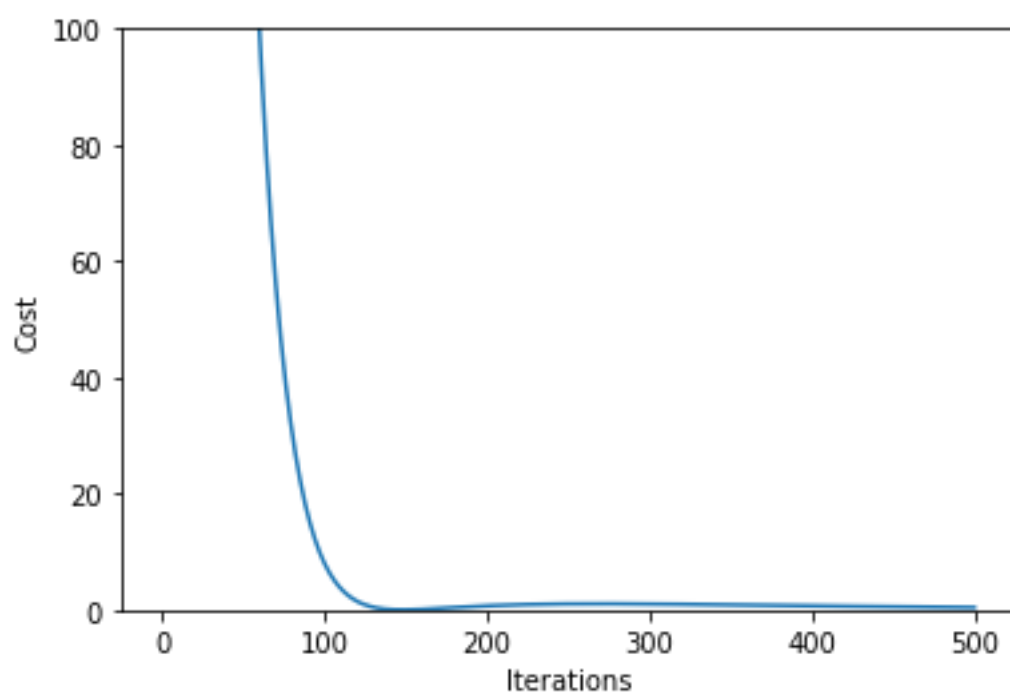
R2 score	0.12520146
RMSE	0.8662726312928897

Degree 3

Parameters

Learning Rate 0.1

Plots



Results

Training cost after 500 iteration 0.4110392503308762484

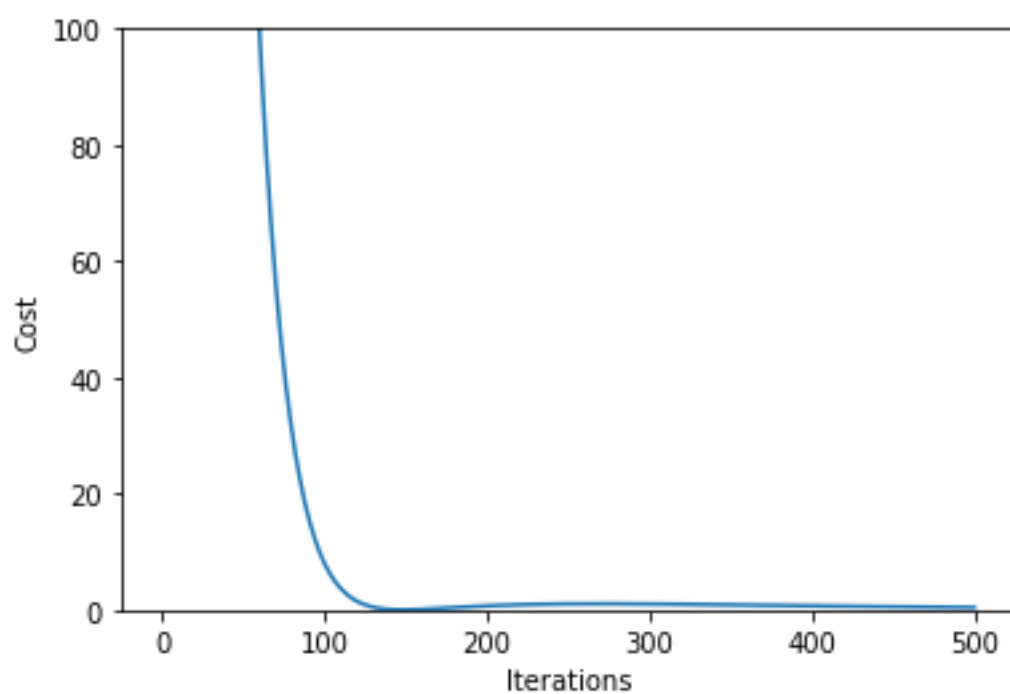
R2 score	0.1197764
RMSE	0.8689545765293664

Degree 4

Parameters

Learning Rate 0.1

Plot



Results

Training cost after 500 iteration 0.056108048572672517

R2 score	0.12520064
RMSE	0.8662730396539436

Gradient Descent with regularization

Degree 6

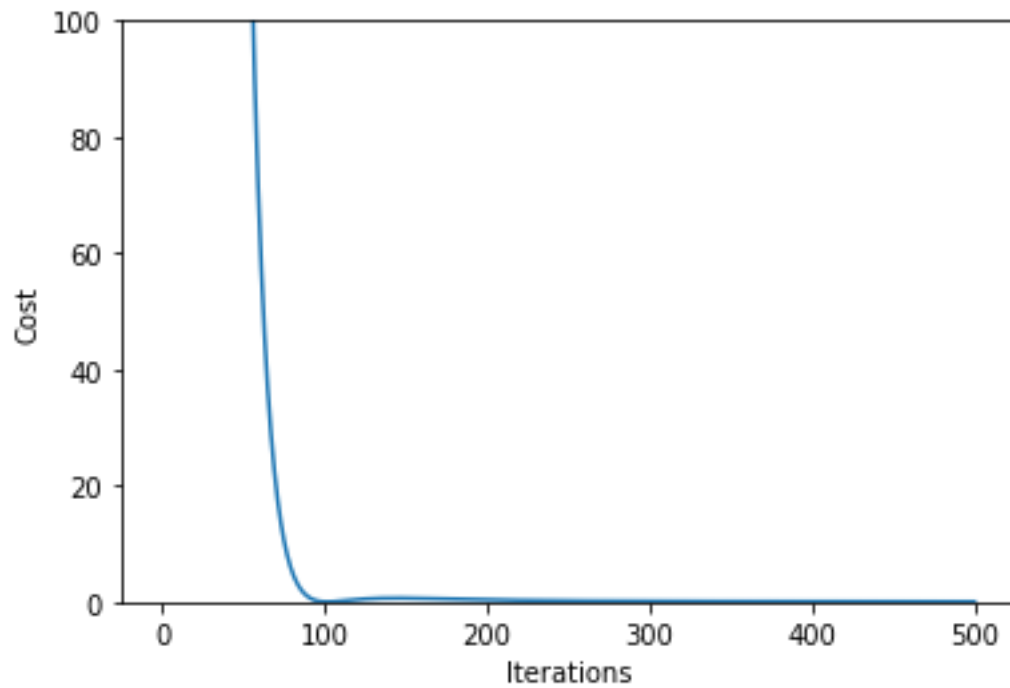
Model 1

Parameters

Learning Rate 0.1

No Regularization

Plot



Results

Training cost after 500 iteration Cost: 0.034950378009566574

R2 score	0.12539375
RMSE	0.8661774198945846

Model 2

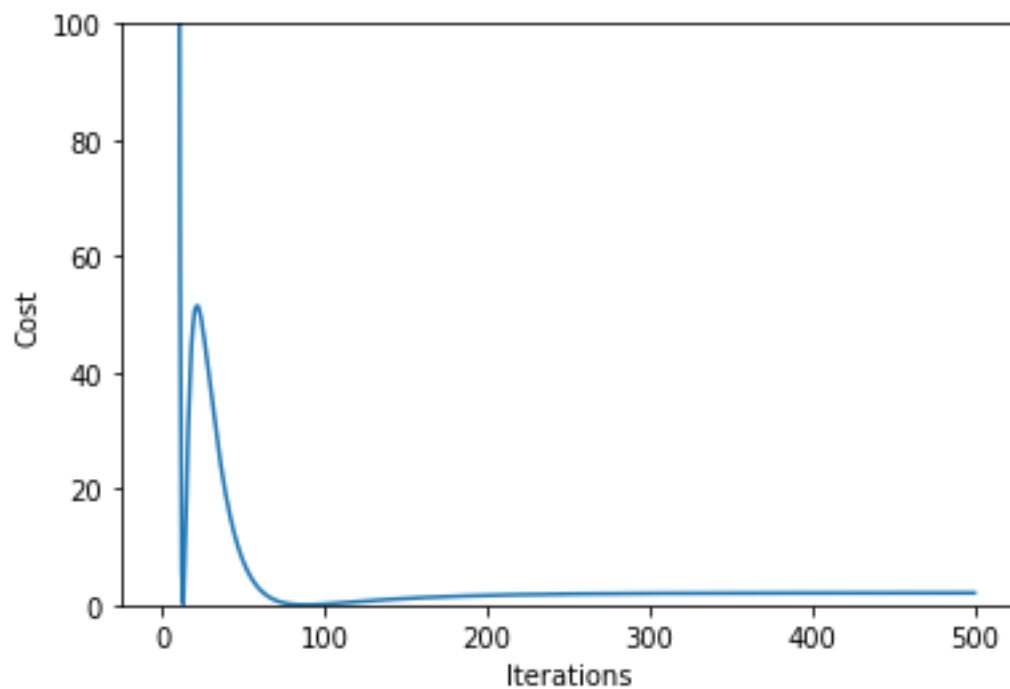
Parameters

Learning Rate 0.1

L2 Regularization

Regularization factor 0.1

Plot



Results

Training error (Mean Squared Error) after 500 iteration

Cost	2.06025262694515734
R2 score	0.13519338
RMSE	0.8613111555689186

Model 3

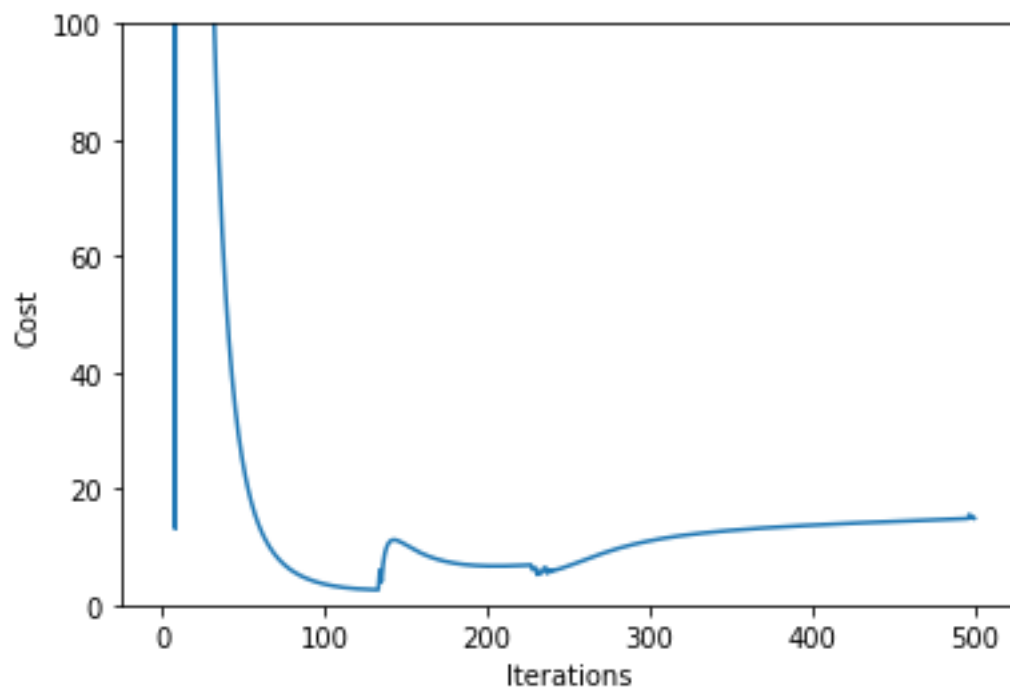
Parameters

Learning Rate 0.1

L1 Regularization

Regularization factor 0.01

Plot



Results

Training cost after 500 iteration Cost	14.903790058748408
R2 score	0.13663046
RMSE	0.8605952243076914

Conclusions

Effects of regularization

On overfitting

We observe clearly that polynomial gradient descent without regularization overfits, giving very less training error

No Regularization	0.034950378009566574
l2	0.8613111555689186
l1	0.8605952243076914

Adding regularization to the model increase the cost significantly however it helps us to perform a lot better on testing/ validation set

On R2 score

We observe that adding a regularization factor increase the training Mean Squared Error however we see significant (around 10% improvement) in R2 values thus helping us perform better on the unknown test/validation set.

No Regularization	0.12539375
L2 regularization	2.06025262694515734
L1 regularization	14.903790058748408

On test set loss

Surprisingly, adding regularization factor didn't help in decreasing the training error as can be observed from the above table

Degree 4

with regularization

Cost	1.7310348238609297671
R2 score	0.1344565
RMSE	0.8616780283389387

Without regularization

Training cost after 500 iteration 0.056108048572672517

R2 score	0.12520064
RMSE	0.8662730396539436

Degree 6

With l2 regularization

Cost	2.06025262694515734
R2 score	0.13519338
RMSE	0.8613111555689186

	Degree 6(with l2 regularization)	Degree 4(without regularization)
Cost	2.06025262694515734	0.056108048572672517
R2 score	0.13519338	0.12520064
RMSE	0.8613111555689186	0.8662730396539436