

# CLL 788 – Assignment 1 Solution

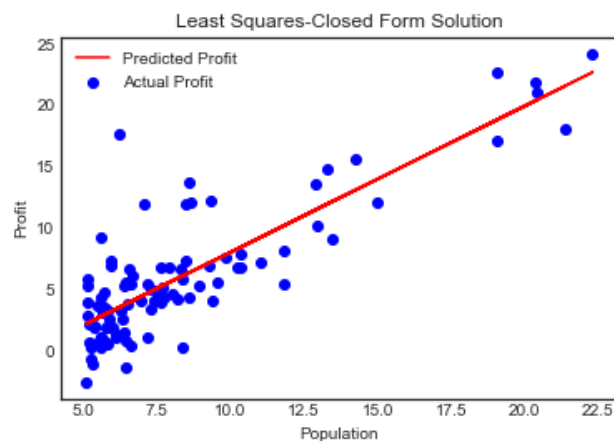
## Q1

Codefile – q1\_linreg.ipynb at : [https://github.com/bhanumagotra/IITD\\_.git](https://github.com/bhanumagotra/IITD_.git)

a. After applying the various regression methods, following plots and results were obtained.

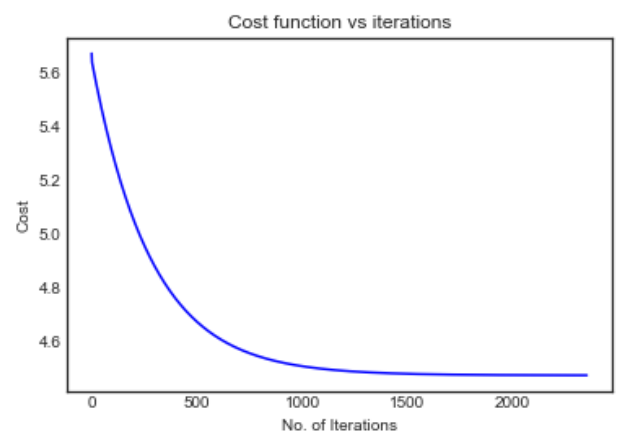
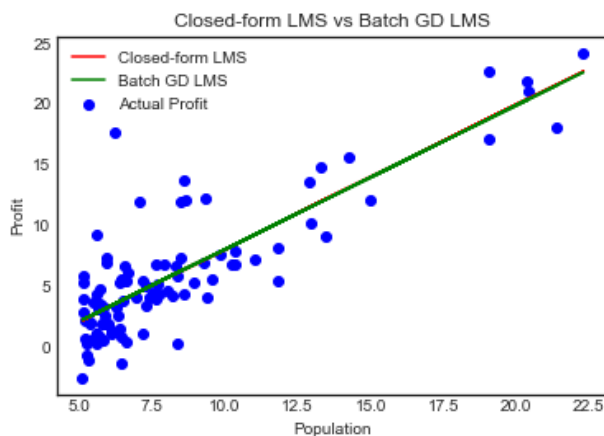
### 1. Closed Form Solution

- Intercept = -3.9150842427308126
- Slope = 1.1930336441895943
- MSE = 4.476971375975179
- Time Taken = 0.01367 seconds



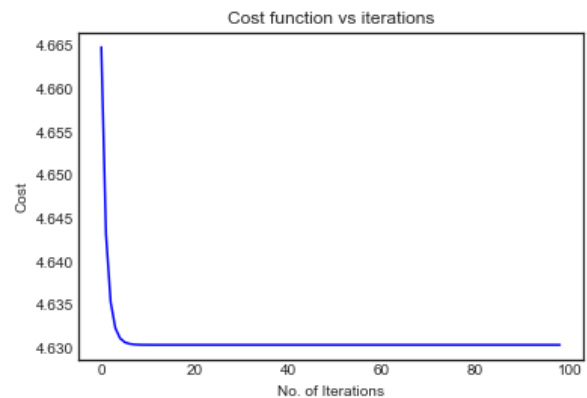
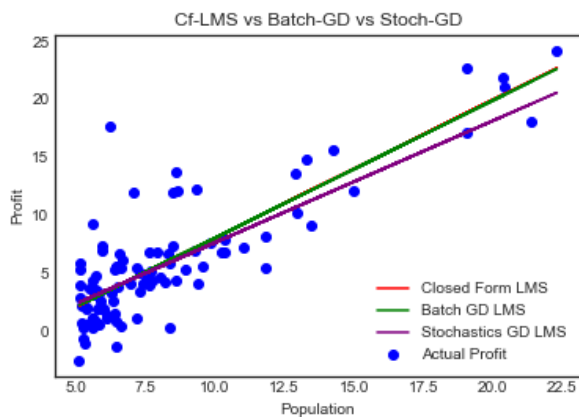
### 2. Batch Gradient Descent

- The learning rate was set to 0.01
- Intercept = -3.8590406875857317
- Slope = 1.1874476282864326
- MSE = 4.47252659230787
- Time Taken = 0.01110 seconds
- No. of Iterations: 2358
- It can be seen that the regression lines are almost coniciding.



### 3. Stochastic Gradient Descent

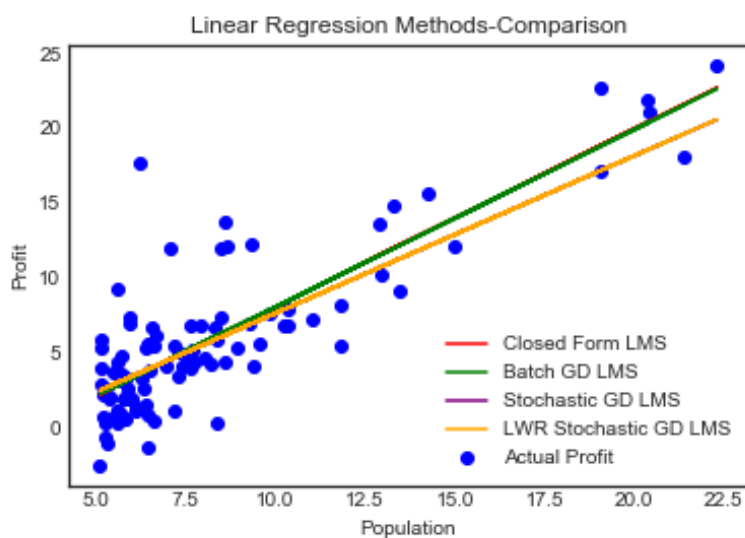
- The learning rate was kept at 0.01 and theta was given a value of -3,1
- The while loop was run for 100 iterations to smoothen out the convergence.
- Intercept = -2.9456365099712047
- Slope = 1.0543634900288048
- MSE = 4.630440350364714
- Time Taken = 0.01071 seconds
- It can be seen that the convergence in case of Stochastic GD happens much earlier as compared to Batch GD.



b. Locally weighted regression was applied by using a query point of 7.576 and bandwidth parameter of 0.5.

#### Locally Weighted Regression using Stochastic Gradient Descent

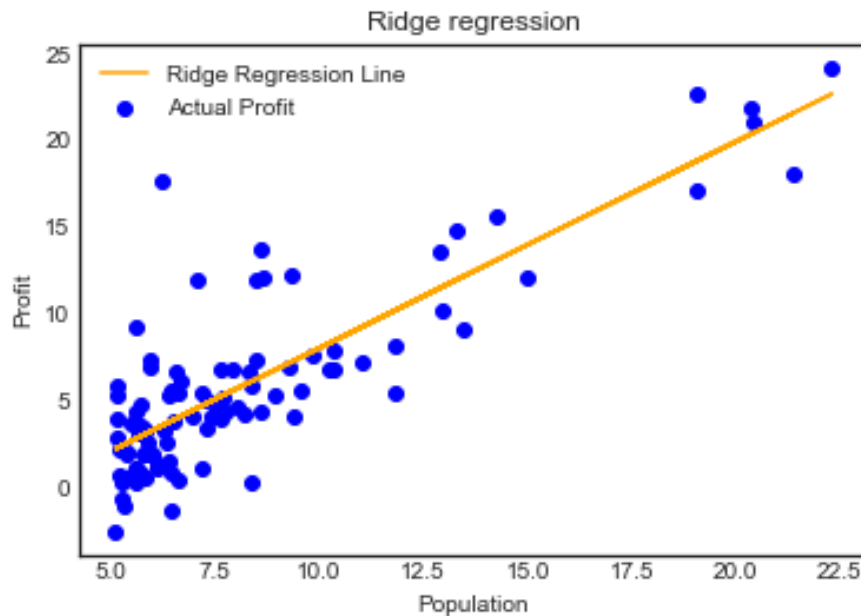
- Intercept = -2.9773598194925075
- Slope = 1.0226401805074927
- MSE = 4.8024393086163855
- Time Taken = 0.01075 seconds
- It can be seen that LWR and Stochastic GD lines have almost coincided.



c. Ridge, lasso and elastic net were applied on the same dataset and results were compared with earlier methods.

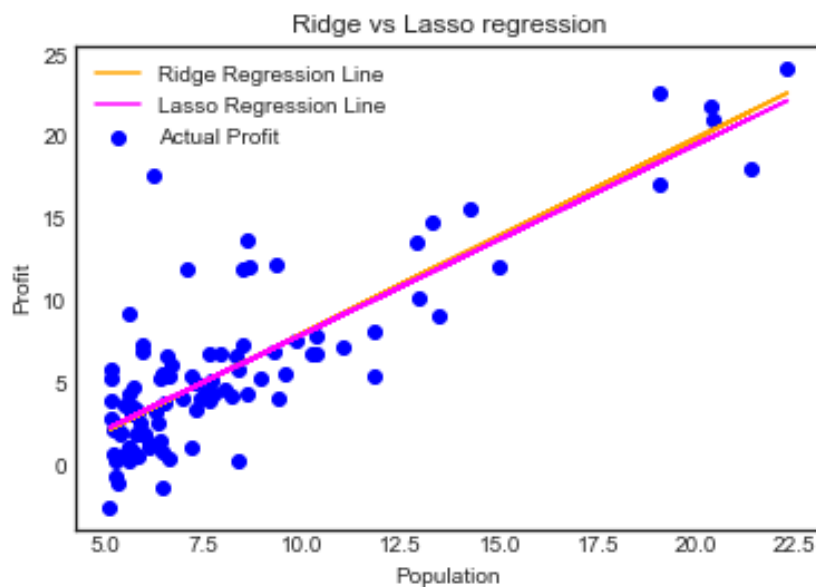
### 1.. Ridge Regression

- It was performed using Inbuilt packages from sklearn
- Alpha was taken as 0.5
- Intercept = [-3.91165835]
- Slope = [[1.19261888]]
- MSE = 4.476972650870759
- Time Taken = 0.01088 seconds



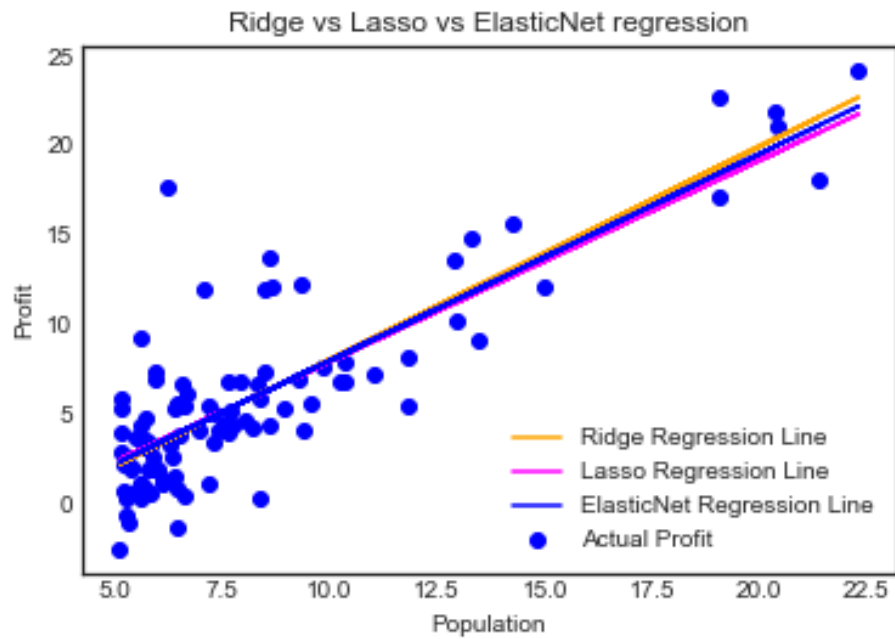
### 2. Lasso Regression

- Alpha was taken as 0.5
- Intercept = [-3.63644374]
- Slope = [1.15929911]
- MSE = 4.485405009510194
- Time Taken = 0.01109 seconds



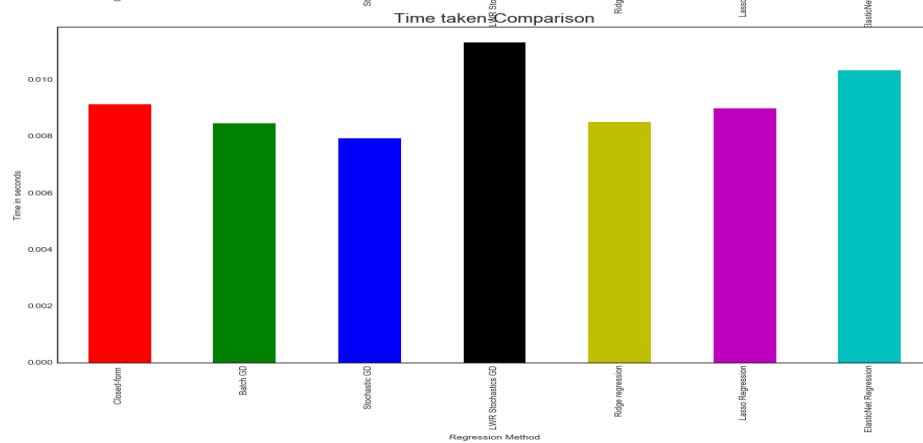
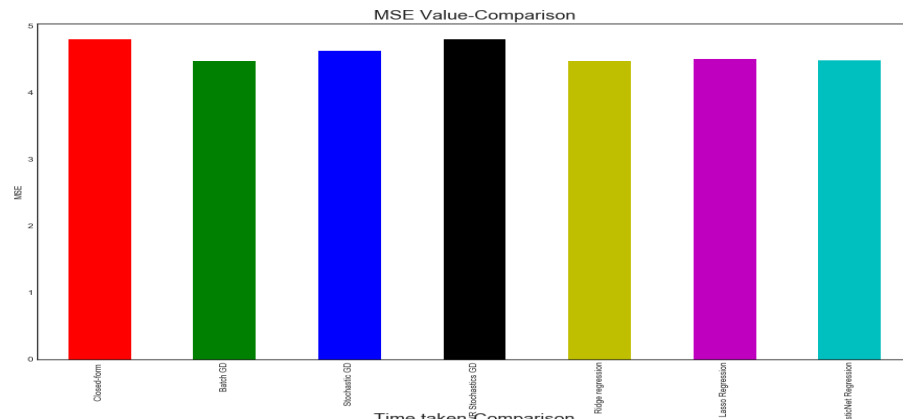
### 3. Elastic-net Regression

- Alpha was taken as 0.5 and L1 ratio was kept as 0.5 too.
- Intercept = [-3.61461828]
- Slope = [1.15665674]
- MSE = 4.48677793819175
- Time Taken = 0.02056 seconds



## A comparative analysis of all linear regression methods

Parameter	Closed-form	Batch GD	Stochastic GD	LWR Stochastics GD	Ridge regression	Lasso Regression	ElasticNet Regression
Intercept	-3.91508	-3.859	-2.94564	-2.97736	-3.91166	-3.3578	-3.61462
Slope	1.19303	1.18744	1.05436	1.02264	1.19262	1.12556	1.15666
MSE	4.80244	4.47725	4.63044	4.80244	4.47697	4.51071	4.48678
TimeTaken(s)	0.00913	0.00845	0.00793	0.0113	0.00851	0.00899	0.01033

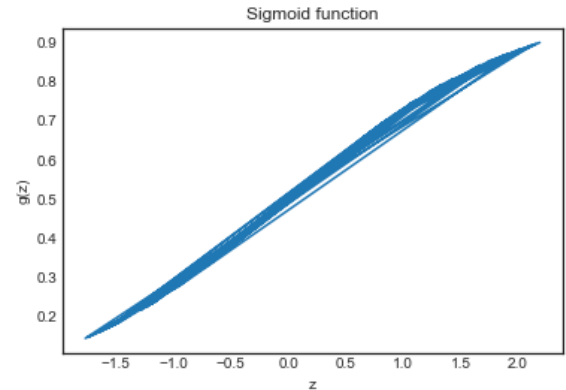


## Q2

Codefile – q2\_logreg.ipynb at : [https://github.com/bhanumagotra/IITD\\_.git](https://github.com/bhanumagotra/IITD_.git)

a. The logistic regression was applied to training data set with following parameters :

- Learning rate – 0.01
- Iterations – 100000
- As it can be seen from the figure, the result obtained after 100000 iterations tend to smoothen out at around 1 and 0.
- But due to less number of sample points, the actual shape of sigmoid couldn't be achieved even after varying the learning rate and number of iterations.
- The trained function was run on testing data set and a value =1 was assigned to predicted values greater than 0.7. The rest of the points were assigned 0.
- The results can be seen as presented in following table :



S.No	Aptitude	Verbal	Label
1	32.82283304	43.40717306	0
2	64.13932042	78.13168802	0
3	72.44649423	96.32759297	0
4	60.55788574	73.1949981	0
5	58.94095622	75.95844831	0
6	99.9278578	72.46925193	1
7	47.36426911	88.575865	0
8	50.5581598	75.90985953	0
9	60.55555629	42.60840944	0
10	82.32666158	42.81987854	1
11	89.01389642	69.9037889	1
12	94.93450672	45.7943068	1
13	67.41925747	66.68935318	0
14	57.33870632	59.61428198	0
15	80.466756	91.0601479	1
16	68.56852179	85.6943071	0
17	42.17545454	78.944786	0
18	75.57770201	90.524539	0
19	78.73542435	96.74742717	1
20	52.44800399	60.86950526	0
21	94.19433113	77.25910509	1
22	90.54855097	87.60879176	1
23	55.58216114	35.67070347	0
24	74.59269242	84.94513685	0
25	89.94580671	45.45828361	1
26	83.58916274	48.4802858	1
27	42.36170081	87.20385094	0
28	99.41500881	68.87540947	1
29	55.44001756	65.03193801	0
30	74.875893	89.6298129	0