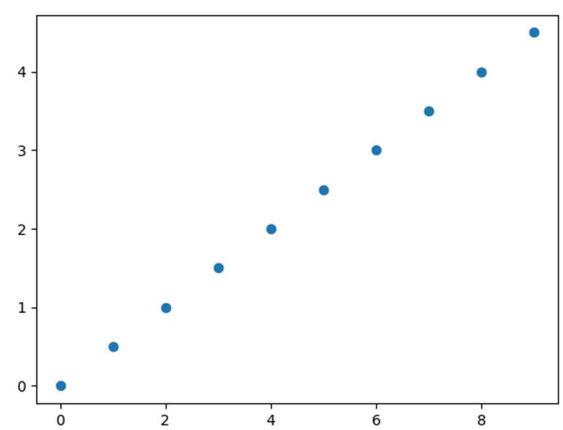
########################
Problem Set 5 | ML
Group Submission
Members:
#	Yasser Parambathkandy - (G01294910)
#	Indranil Pal - (G01235186)
Date: 11/18/2022

Linear Regression
(a) Use your closed form implementation to fit a model to the data in 1D-no-noise-lin.txt and 2D-noisy-lin.txt. What is the loss of the fit model in both cases? Include plots for both. Note: for some of these datasets the y-intercept will be non-zero, make sure you handle this case!
(b) What happens when you're using the closed form solution and one of the features (columns of X) is duplicated? Explain why. Note: you may want to test this out with your code as a useful first step, but you should think critically about what is happening and why.
(c) Does the same thing happen if one of the training points (rows of X) is duplicated? Explain why.
(d) Does the same thing happen with Gradient Descent? Explain why.

a) Closed form implementation :

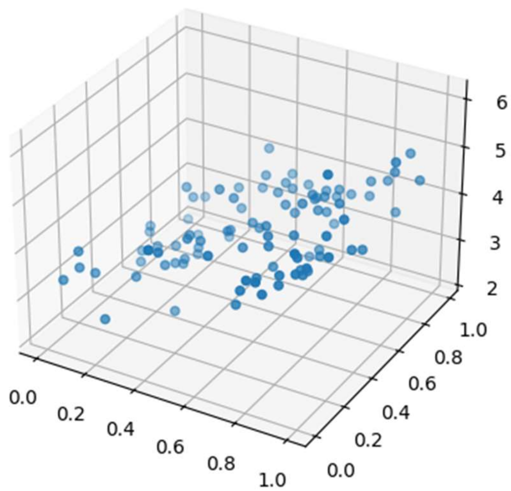
1D-no-noise-lin.txt

Theta : [[0.5]] , Mean squared error: 0.00

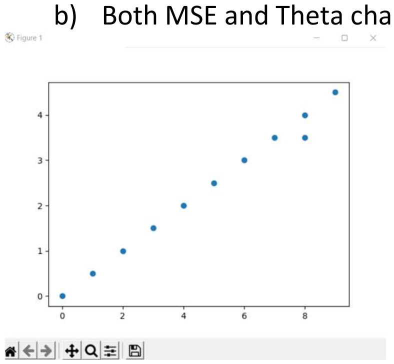


2D-noisy-lin.txt

Theta : [[4.47467363] [1.9625976]] , Mean squared error: 1.20



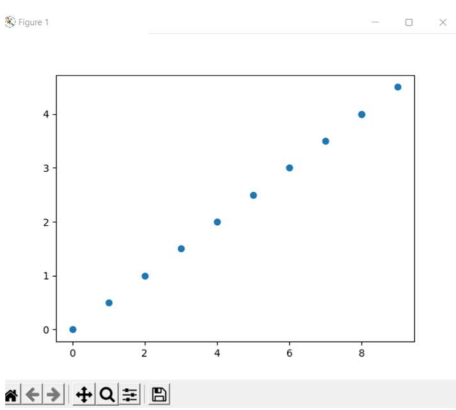
b) Both MSE and Theta changes .



Theta : [[0.48853868]] , Mean squared error: 0.02

c) When row is repeated, Theta and MSE don't change

Theta : [[0.5]] , Mean squared error: 0.00



d) With duplicate data(only X is repeated or row is duplicated), the loss is increased :

<u>Original :</u>

curr_theta [array([0.285]), array([0.045])], cur_loss 121.61812500000002 iteration 1
curr_theta [array([0.4035]), array([0.06345])], cur_loss 21.011072625000004 iteration 2
curr_theta [array([0.4527945]), array([0.070866])], cur_loss 3.6299309738062417 iteration 3
curr_theta [array([0.47332369]), array([0.07369717])], cur_loss 0.6271187447093247 iteration 4
curr_theta [array([0.48189644]), array([0.0746241])], cur_loss 0.10834490211962254 iteration 5
curr_theta [array([0.4854993]), array([0.07476094])], cur_loss 0.01872013017999265 iteration 6
curr_theta [array([0.48703622]), array([0.07457078])], cur_loss 0.003236291976632816 iteration 7
curr_theta [array([0.4877142]), array([0.07424611])], cur_loss 0.0005612387131172915 iteration 8
curr_theta [array([0.48803496]), array([0.07386691])], cur_loss 9.906543470552655e-05 iteration 9
curr_theta [array([0.48820701]), array([0.07346642])], cur_loss 1.9195450888155244e-05 iteration 10

<u>in case of only X is repeated in gradient_descent</u>

curr_theta [array([0.31]), array([0.04727273])], cur_loss 193.14274999999998 iteration 1
curr_theta [array([0.41873554]), array([0.06372727])], cur_loss 23.754529531930864 iteration 2
curr_theta [array([0.45688781]), array([0.06937458])], cur_loss 2.921558072820383 iteration 3
curr_theta [array([0.47028654]), array([0.07123244])], cur_loss 0.35932145792087383 iteration 4
curr_theta [array([0.47500414]), array([0.071762])], cur_loss 0.044193268888371676 iteration 5
curr_theta [array([0.47667717]), array([0.07182636])], cur_loss 0.005435807226777534 iteration 6
curr_theta [array([0.47728239]), array([0.07172821])], cur_loss 0.0006690414198957546 iteration 7
curr_theta [array([0.47751303]), array([0.07157371])], cur_loss 8.277413724558271e-05 iteration 8
curr_theta [array([0.4776122]), array([0.07140007])], cur_loss 1.0664305081282643e-05 iteration 9
curr_theta [array([0.47766518]), array([0.07122035])], cur_loss 1.7904908881373216e-06 iteration 10

<u>Whole row is repeated :</u>

curr_theta [array([0.31727273]), array([0.04818182])], cur_loss 202.27331818181818 iteration 1
curr_theta [array([0.42857851]), array([0.06482645])], cur_loss 24.877495132982713 iteration 2
curr_theta [array([0.46765178]), array([0.07041235])], cur_loss 3.0596726704796433 iteration 3
curr_theta [array([0.48139301]), array([0.0721213])], cur_loss 0.3763097124016626 iteration 4
curr_theta [array([0.48625012]), array([0.07247191])], cur_loss 0.046284238920407506 iteration 5
curr_theta [array([0.48799139]), array([0.07234746])], cur_loss 0.005694551122802042 iteration 6
curr_theta [array([0.48863973]), array([0.0720577])], cur_loss 0.0007024242597145351 iteration 7
curr_theta [array([0.4889046]), array([0.07171127])], cur_loss 8.842361578517254e-05 iteration 8
curr_theta [array([0.48903478]), array([0.07134623])], cur_loss 1.2886723164859169e-05 iteration 9
curr_theta [array([0.48911753]), array([0.07097596])], cur_loss 3.575437066274882e-06 iteration 10

2. Now use your Gradient Descent implementation to fit a model to 1D-no-noise-lin.txt with alpha=0.05, num iters=10, and initial Theta set to the zero vector. What is the output (the full list of (theta, loss) tuples for each of the 10 iterations)?
(b) Using the default parameters for alpha and num iters, and with initial Theta set to 0, do you get the same model parameters as you did with the closed form solution? the same loss? Report this for both 1D-no-noise-lin.txt and 2D-noisy-lin.txt.
(c) Find a set of values of the learning rate and iterations where you get the same answers and a set of values where the answers are noticeably different. Explain what's going on in both cases, and for both datasets 1D-no-noise-lin.txt and 2D-noisy-lin.txt.

a) 1D-no-noise-lin.txt

curr_theta [array([1.425]), array([0.225])], cur_loss 3040.4531249999995 iteration 1
curr_theta [array([-1.3125]), array([-0.21375])], cur_loss 11228.926640625003 iteration 2
curr_theta [array([3.9493125]), array([0.62325])], cur_loss 41470.39599697267 iteration 3
curr_theta [array([-6.16169063]), array([-0.99126563])], cur_loss 153157.4481496576 iteration 4
curr_theta [array([13.27019719]), array([2.10562172])], cur_loss 565637.3270448203 iteration 5
curr_theta [array([-24.07239457]), array([-3.85152919])], cur_loss 2088997.8886419649 iteration 6
curr_theta [array([47.69211809]), array([7.59120129])], cur_loss 7715035.7134086555 iteration 7
curr_theta [array([-90.22145905]), array([-14.40437198])], cur_loss 28492980.47816256 iteration 8
curr_theta [array([174.81666662]), array([27.86072179])], cur_loss 105229575.94620734 iteration 9
curr_theta [array([-334.52315806]), array([-53.36785037])], cur_loss 388631286.3027927 iteration 10

b) 1D-no-noise-lin.txt

With default parameter , the loss increases with each iteration in gradient_descent, and initial loss is also higher than closed form

linreg_grad_desc(init_theta, data_X, data_Y)

Theta : [[0.5]] , Mean squared error: 0.00

curr_theta [array([1.425]), array([0.225])], cur_loss 3040.4531249999995 iteration 1
....
curr_theta [array([1.79617243e+141]), array([2.86444118e+140])], cur_loss 1.1171417917815495e+286 iteration 499
curr_theta [array([-3.45181885e+141]), array([-5.50477887e+140])], cur_loss 4.125800637500844e+286 iteration 500

2D-noisy-lin.txt

Theta : [[4.52484452]
[1.92893988]] , Mean squared error: 1.26

curr_theta [array([0.22628725, 0.19552166]), array([0.38288361, 0.38288361])], cur_loss 2549.8183809417246 iteration 1
....
curr_theta [array([2.06429188, -0.62765551]), array([2.70519279, 4.1563301])], cur_loss 0.00014230196655771327 iteration 499
curr_theta [array([2.06438754, -0.62820643]), array([2.70513707, 4.15664106])], cur_loss 0.00014040878095893393 iteration 500

c) Similar result

1D-no-noise-lin.txt : learning rate : 0.01 , num_iteration : 8

Theta : [[0.5]] , Mean squared error: 0.00
curr_theta [array([0.285]), array([0.045])], cur_loss 121.61812500000002 iteration 1
curr_theta [array([0.4035]), array([0.06345])], cur_loss 21.011072625000004 iteration 2
curr_theta [array([0.4527945]), array([0.070866])], cur_loss 3.6299309738062417 iteration 3
curr_theta [array([0.47332369]), array([0.07369717])], cur_loss 0.6271187447093247 iteration 4
curr_theta [array([0.48189644]), array([0.0746241])], cur_loss 0.10834490211962254 iteration 5
curr_theta [array([0.4854993]), array([0.07476094])], cur_loss 0.01872013017999265 iteration 6
curr_theta [array([0.48703622]), array([0.07457078])], cur_loss 0.0032362919766328116 iteration 7
curr_theta [array([0.4877142]), array([0.07424611])], cur_loss 0.0005612387131172915 iteration 8
curr_theta [array([0.48803496]), array([0.07386691])], cur_loss 9.906543470552655e-05 iteration 9
curr_theta [array([0.48820701]), array([0.07346642])], cur_loss 1.9195450888155244e-05 iteration 10

2D-noisy-lin.txt: learning rate : 0.01 , num_iteration : 84
Theta : [1.92893988]] , Mean squared error: 1.26
curr_theta [array([0.04525745, 0.03910433]), array([0.07657672, 0.07657672])], cur_loss 101.99273523766898 iteration 1
curr_theta [array([0.08933501, 0.07711721]), array([0.15112897, 0.15120996])], cur_loss 96.7278732277214 iteration 2
curr_theta [array([0.13226376, 0.11406704]), array([0.22371009, 0.22395009])], cur_loss 91.73480923699269 iteration 3
curr_theta [array([0.17407393, 0.14998153]), array([0.29437201, 0.29484616])], cur_loss 86.99951090281193 iteration 4
curr_theta [array([0.21479498, 0.18488762]), array([0.3631653 , 0.36394597])], cur_loss 82.50867036463995 iteration 5
curr_theta [array([0.25445558, 0.2188116]), array([0.43013919, 0.43129607])], cur_loss 78.24966685535233 iteration 6
curr_theta [array([0.29308366, 0.25177904]), array([0.49534163, 0.49694178])], cur_loss 74.21053122417909 iteration 7
curr_theta [array([0.3307064 , 0.28381486]), array([0.55881929, 0.56092729])], cur_loss 70.37991229155382 iteration 8
curr_theta [array([0.36735026, 0.31494334]), array([0.62061761, 0.6232956])], cur_loss 66.74704494127161 iteration 9
curr_theta [array([0.40304102, 0.34518812]), array([0.68078085, 0.68408862])], cur_loss 63.301719860244326 iteration 10
curr_theta [array([0.43780376, 0.37457223]), array([0.73935208, 0.74334717])], cur_loss 60.03425484077127 iteration 11
curr_theta [array([0.47166293, 0.4031181]), array([0.79637326, 0.801111])], cur_loss 56.93546756464123 iteration 12
curr_theta [array([0.5046423 , 0.43084759]), array([0.85188522, 0.85741883])], cur_loss 53.99664979254289 iteration 13
curr_theta [array([0.53676505, 0.45778198]), array([0.90592774, 0.91230839])], cur_loss 51.20954288621784 iteration 14
curr_theta [array([0.56805371, 0.48394199]), array([0.95853953, 0.96581641])], cur_loss 48.56631459453661 iteration 15
curr_theta [array([0.59853025, 0.50934782]), array([1.00975829, 1.01797869])], cur_loss 46.059537038226296 iteration 16
curr_theta [array([0.62821604, 0.53401914]), array([1.05962073, 1.06883009])], cur_loss 43.682165831359804 iteration 17
curr_theta [array([0.6571319 , 0.55797511]), array([1.10816258, 1.11840455])], cur_loss 41.42752028090186 iteration 18

```
curr_theta [array([0.68529809, 0.58123438]), array([1.15541866, 1.16673516])], cur_loss 39.289264608646675 iteration 19
curr_theta [array([0.71273436, 0.60381514]), array([1.20142282, 1.21385414])], cur_loss 37.26139014275077 iteration 20
curr_theta [array([0.73945991, 0.62573509]), array([1.24620808, 1.25979285])], cur_loss 35.338198428795806 iteration 21
curr_theta [array([0.76549344, 0.64701148]), array([1.28980653, 1.30458187])], cur_loss 33.51428521289881 iteration 22
curr_theta [array([0.79085318, 0.6676611 ]), array([1.33224946, 1.34825098])], cur_loss 31.784525251837945 iteration 23
curr_theta [array([0.81555686, 0.68770031]), array([1.37356732, 1.39082917])], cur_loss 30.14405790749261 iteration 24
curr_theta [array([0.83962174, 0.70714506]), array([1.41378975, 1.43234469])], cur_loss 28.588273485095343 iteration 25
curr_theta [array([0.86306464, 0.72601086]), array([1.45294562, 1.47282506])], cur_loss 27.112800276889455 iteration 26
curr_theta [array([0.88590193, 0.74431284]), array([1.49106304, 1.51229708])], cur_loss 25.713492274766942 iteration 27
curr_theta [array([0.90814954, 0.76206573]), array([1.52816937, 1.55078685])], cur_loss 24.386417517342828 iteration 28
curr_theta [array([0.92982299, 0.77928387]), array([1.56429125, 1.58831982])], cur_loss 23.12784703870735 iteration 29
curr_theta [array([0.9509374 , 0.79598123]), array([1.59945463, 1.62492073])], cur_loss 21.934244387785522 iteration 30
curr_theta [array([0.97150748, 0.81217143]), array([1.63368477, 1.66061373])], cur_loss 20.80225568884277 iteration 31
curr_theta [array([0.99154754, 0.82786772]), array([1.66700626, 1.69542231])], cur_loss 19.728700215192355 iteration 32
curr_theta [array([1.01107154, 0.84308302]), array([1.69944304, 1.72936937])], cur_loss 18.71056144960564 iteration 33
curr_theta [array([1.03009305, 0.8578299 ]), array([1.73101844, 1.7624772 ])], cur_loss 17.744978606294925 iteration 34
curr_theta [array([1.04862531, 0.87212062]), array([1.76175515, 1.79476752])], cur_loss 16.82923859063415 iteration 35
curr_theta [array([1.06668117, 0.88596711]), array([1.79167527, 1.82626149])], cur_loss 15.960768374015117 iteration 36
curr_theta [array([1.08427318, 0.899381  ]), array([1.82080033, 1.85697971])], cur_loss 15.137127762405807 iteration 37
curr_theta [array([1.10141354, 0.91237361]), array([1.84915128, 1.88694226])], cur_loss 14.356002538278128 iteration 38
curr_theta [array([1.11811412, 0.92495596]), array([1.87674851, 1.91616869])], cur_loss 13.615197956630203 iteration 39
curr_theta [array([1.13438652, 0.9371388 ]), array([1.90361191, 1.94467804])], cur_loss 12.91263257681719 iteration 40
curr_theta [array([1.15024197, 0.94893259]), array([1.92976079, 1.97248886])], cur_loss 12.246332412853008 iteration 41
curr_theta [array([1.16569146, 0.96034752]), array([1.95521401, 1.99961922])], cur_loss 11.614425385737325 iteration 42
curr_theta [array([1.18074567, 0.97139353]), array([1.97998988, 2.02608672])], cur_loss 11.015136062215296 iteration 43
curr_theta [array([1.19541499, 0.98208027]), array([2.00410627, 2.05190851])], cur_loss 10.44678066517781 iteration 44
curr_theta [array([1.20970954, 0.99241717]), array([2.02758056, 2.07710128])], cur_loss 9.907762341678321 iteration 45
curr_theta [array([1.22363919, 1.00241341]), array([2.05042966, 2.1016813 ])], cur_loss 9.396566675265356 iteration 46
curr_theta [array([1.23721353, 1.01207791]), array([2.07267007, 2.12566442])], cur_loss 8.91175743001197 iteration 47
curr_theta [array([1.2504419 , 1.02141938]), array([2.09431781, 2.14906606])], cur_loss 8.451972514283986 iteration 48
curr_theta [array([1.26333341, 1.03044631]), array([2.11538852, 2.17190126])], cur_loss 8.015920152897303 iteration 49
curr_theta [array([1.27589691, 1.03916694]), array([2.1358974 , 2.19418467])], cur_loss 7.602375256906629 iteration 50
curr_theta [array([1.28814103, 1.04758932]), array([2.15585927, 2.21593053])], cur_loss 7.210175980821147 iteration 51
curr_theta [array([1.30007417, 1.05572129]), array([2.17528853, 2.23715276])], cur_loss 6.838220457568839 iteration 52
curr_theta [array([1.31170451, 1.06357048]), array([2.19419924, 2.25786487])], cur_loss 6.485463702031657 iteration 53
curr_theta [array([1.32304   , 1.07114431]), array([2.21260505, 2.27808005])], cur_loss 6.150914674451372 iteration 54
curr_theta [array([1.3340884 , 1.07845003]), array([2.23051929, 2.29781114])], cur_loss 5.83363349544598 iteration 55
curr_theta [array([1.34485726, 1.0854947 ]), array([2.2479549 , 2.31707065])], cur_loss 5.532728804813244 iteration 56
curr_theta [array([1.35535392, 1.09228517]), array([2.26492451, 2.33587076])], cur_loss 5.24735525669549 iteration 57
curr_theta [array([1.36558555, 1.09882815]), array([2.28144039, 2.35422332])], cur_loss 4.9767111440656535 iteration 58
curr_theta [array([1.37555911, 1.10513015]), array([2.29751452, 2.37213991])], cur_loss 4.720036145857095 iteration 59
curr_theta [array([1.38528137, 1.11119751]), array([2.31315853, 2.38963178])], cur_loss 4.476609190404846 iteration 60
curr_theta [array([1.39475896, 1.11703642]), array([2.32838377, 2.4067099 ])], cur_loss 4.245746429194059 iteration 61
curr_theta [array([1.40399829, 1.1226529 ]), array([2.34320128, 2.42338494])], cur_loss 4.0267993152190344 iteration 62
curr_theta [array([1.41300564, 1.12805281]), array([2.3576218 , 2.43966731])], cur_loss 3.819152780552476 iteration 63
curr_theta [array([1.42178709, 1.13324188]), array([2.37165581, 2.45556715])], cur_loss 3.622223508002656 iteration 64
curr_theta [array([1.43034858, 1.13822566]), array([2.38531348, 2.47109433])], cur_loss 3.4354582920003085 iteration 65
curr_theta [array([1.43869589, 1.14300958]), array([2.39860476, 2.48625847])], cur_loss 3.2583324841086094 iteration 66
curr_theta [array([1.44683466, 1.14759892]), array([2.41153929, 2.50106892])], cur_loss 3.090348518786913 iteration 67
curr_theta [array([1.45477035, 1.15199881]), array([2.42412648, 2.51553482])], cur_loss 2.931034515265277 iteration 68
curr_theta [array([1.46250831, 1.15621428]), array([2.4363755 , 2.52966505])], cur_loss 2.779942951599188 iteration 69
curr_theta [array([1.47005372, 1.16025019]), array([2.44829525, 2.54346827])], cur_loss 2.636649407178976 iteration 70
curr_theta [array([1.47741165, 1.1641113 ]), array([2.45989443, 2.5569529 ])], cur_loss 2.5007513701585955 iteration 71
curr_theta [array([1.48458702, 1.16780223]), array([2.47118148, 2.57012717])], cur_loss 2.3718671064529233 iteration 72
curr_theta [array([1.49158463, 1.17132751]), array([2.48216463, 2.58299907])], cur_loss 2.2496345871244503 iteration 73
curr_theta [array([1.49840913, 1.17469152]), array([2.49285191, 2.5955764 ])], cur_loss 2.1337104711441466 iteration 74
curr_theta [array([1.50506509, 1.17789853]), array([2.50325111, 2.60786674])], cur_loss 2.0237691406688514 iteration 75
curr_theta [array([1.51155691, 1.18095271]), array([2.51336983, 2.61987749])], cur_loss 1.919501786123156 iteration 76
```

curr_theta [array([1.51788893, 1.18385813]), array([2.52321546, 2.63161585])], cur_loss 1.8206155385139797 iteration 77
curr_theta [array([1.52406531, 1.18661873]), array([2.53279522, 2.64308884])], cur_loss 1.7268326465400679 iteration 78
curr_theta [array([1.53009017, 1.18923837]), array([2.54211611, 2.65430328])], cur_loss 1.6378896961833829 iteration 79
curr_theta [array([1.53596746, 1.19172079]), array([2.55118496, 2.66526584])], cur_loss 1.553536870589461 iteration 80
curr_theta [array([1.54170107, 1.19406965]), array([2.56000841, 2.675983])], cur_loss 1.4735372481558349 iteration 81
curr_theta [array([1.54729476, 1.1962885]), array([2.56859295, 2.68646107])], cur_loss 1.3976661368567855 iteration 82
curr_theta [array([1.55275221, 1.19838082]), array([2.57694487, 2.6967062])], cur_loss 1.3257104429335291 iteration 83
curr_theta [array([1.55807698, 1.20034997]), array([2.58507031, 2.70672439])], cur_loss 1.2574680721749767 iteration 84
curr_theta [array([1.56327257, 1.20219924]), array([2.59297525, 2.71652147])], cur_loss 1.1927473621075089 iteration 85

DisSimilar result
1D-no-noise-lin.txt : learning rate : 0.01 , num_iteration : 8

Theta : [[0.5]] , Mean squared error: 0.00
curr_theta [array([0.285]), array([0.045])], cur_loss 121.61812500000002 iteration 1
curr_theta [array([0.4035]), array([0.06345])], cur_loss 21.011072625000004 iteration 2
curr_theta [array([0.4527945]), array([0.070866])], cur_loss 3.6299309738062417 iteration 3
curr_theta [array([0.47332369]), array([0.07369717])], cur_loss 0.6271187447093247 iteration 4
curr_theta [array([0.48189644]), array([0.0746241])], cur_loss 0.10834490211962254 iteration 5
curr_theta [array([0.4854993]), array([0.07476094])], cur_loss 0.01872013017999265 iteration 6
curr_theta [array([0.48703622]), array([0.07457078])], cur_loss 0.0032362919766328116 iteration 7
curr_theta [array([0.4877142]), array([0.07424611])], cur_loss 0.0005612387131172915 iteration 8
curr_theta [array([0.48803496]), array([0.07386691])], cur_loss 9.906543470552655e-05 iteration 9
curr_theta [array([0.48820701]), array([0.07346642])], cur_loss 1.9195450888155244e-05 iteration 10

2D-noisy-lin.txt: learning rate : 0.01 , num_iteration : 84
Theta : [1.92893988]] , Mean squared error: 1.26
curr_theta [array([0.04525745, 0.03910433]), array([0.07657672, 0.07657672])], cur_loss 101.99273523766898 iteration 1
curr_theta [array([0.08933501, 0.07711721]), array([0.15112897, 0.15120996])], cur_loss 96.7278732277214 iteration 2
curr_theta [array([0.13226376, 0.11406704]), array([0.22371009, 0.22395009])], cur_loss 91.73480923699269 iteration 3
curr_theta [array([0.17407393, 0.14998153]), array([0.29437201, 0.29484616])], cur_loss 86.99951090281193 iteration 4
curr_theta [array([0.21479498, 0.18488762]), array([0.3631653 , 0.36394597])], cur_loss 82.50867036463995 iteration 5
curr_theta [array([0.25445558, 0.2188116]), array([0.43013919, 0.43129607])], cur_loss 78.24966685535233 iteration 6
curr_theta [array([0.29308366, 0.25177904]), array([0.49534163, 0.49694178])], cur_loss 74.21053122417909 iteration 7
curr_theta [array([0.3307064 , 0.28381486]), array([0.55881929, 0.56092729])], cur_loss 70.37991229155382 iteration 8
curr_theta [array([0.36735026, 0.31494334]), array([0.62061761, 0.6232956])], cur_loss 66.74704494127161 iteration 9
curr_theta [array([0.40304102, 0.34518812]), array([0.68078085, 0.68408862])], cur_loss 63.301719860244326 iteration 10
curr_theta [array([0.43780376, 0.37457223]), array([0.73935208, 0.74334717])], cur_loss 60.03425484077127 iteration 11
curr_theta [array([0.47166293, 0.4031181]), array([0.79637326, 0.801111])], cur_loss 56.93546756464123 iteration 12
curr_theta [array([0.5046423 , 0.43084759]), array([0.85188522, 0.85741883])], cur_loss 53.99664979254289 iteration 13
curr_theta [array([0.53676505, 0.45778198]), array([0.90592774, 0.91230839])], cur_loss 51.20954288621784 iteration 14
curr_theta [array([0.56805371, 0.48394199]), array([0.95853953, 0.96581641])], cur_loss 48.56631459453661 iteration 15
curr_theta [array([0.59853025, 0.50934782]), array([1.00975829, 1.01797869])], cur_loss 46.059537038226296 iteration 16
curr_theta [array([0.62821604, 0.53401914]), array([1.05962073, 1.06883009])], cur_loss 43.682165831359804 iteration 17
curr_theta [array([0.6571319 , 0.55797511]), array([1.10816258, 1.11840455])], cur_loss 41.42752028090186 iteration 18
curr_theta [array([0.68529809, 0.58123438]), array([1.15541866, 1.16673516])], cur_loss 39.289264608646675 iteration 19
curr_theta [array([0.71273436, 0.60381514]), array([1.20142282, 1.21385414])], cur_loss 37.26139014275077 iteration 20
curr_theta [array([0.73945991, 0.62573509]), array([1.24620808, 1.25979285])], cur_loss 35.338198428795806 iteration 21
curr_theta [array([0.76549344, 0.64701148]), array([1.28980653, 1.30458187])], cur_loss 33.51428521289881 iteration 22
curr_theta [array([0.79085318, 0.6676611]), array([1.33224946, 1.34825098])], cur_loss 31.784525251837945 iteration 23
curr_theta [array([0.81555686, 0.68770031]), array([1.37356732, 1.39082917])], cur_loss 30.14405790749261 iteration 24
curr_theta [array([0.83962174, 0.70714506]), array([1.41378975, 1.43234469])], cur_loss 28.588273485095343 iteration 25
curr_theta [array([0.86306464, 0.72601086]), array([1.45294562, 1.47282506])], cur_loss 27.112800276889455 iteration 26
curr_theta [array([0.88590193, 0.74431284]), array([1.49106304, 1.51229708])], cur_loss 25.713492274766942 iteration 27
curr_theta [array([0.90814954, 0.76206573]), array([1.52816937, 1.55078685])], cur_loss 24.386417517342828 iteration 28
curr_theta [array([0.92982299, 0.77928387]), array([1.56429125, 1.58831982])], cur_loss 23.12784703870735 iteration 29
curr_theta [array([0.9509374 , 0.79598123]), array([1.59945463, 1.62492073])], cur_loss 21.934244387785522 iteration 30

curr_theta [array([0.97150748, 0.81217143]), array([1.63368477, 1.66061373])], cur_loss 20.80225568884277 iteration 31
curr_theta [array([0.99154754, 0.82786772]), array([1.66700626, 1.69542231])], cur_loss 19.728700215192355 iteration 32
curr_theta [array([1.01107154, 0.84308302]), array([1.69944304, 1.72936937])], cur_loss 18.71056144960564 iteration 33
curr_theta [array([1.03009305, 0.8578299]), array([1.73101844, 1.7624772])], cur_loss 17.744978606294925 iteration 34
curr_theta [array([1.04862531, 0.87212062]), array([1.76175515, 1.79476752])], cur_loss 16.82923859063415 iteration 35
curr_theta [array([1.06668117, 0.88596711]), array([1.79167527, 1.82626149])], cur_loss 15.960768374015117 iteration 36
curr_theta [array([1.08427318, 0.899381]), array([1.82080033, 1.85697971])], cur_loss 15.137127762405807 iteration 37
curr_theta [array([1.10141354, 0.91237361]), array([1.84915128, 1.88694226])], cur_loss 14.356002538278128 iteration 38
curr_theta [array([1.11811412, 0.92495596]), array([1.87674851, 1.91616869])], cur_loss 13.615197956630203 iteration 39
curr_theta [array([1.13438652, 0.9371388]), array([1.90361191, 1.94467804])], cur_loss 12.91263257681719 iteration 40
curr_theta [array([1.15024197, 0.94893259]), array([1.92976079, 1.97248886])], cur_loss 12.246332412853008 iteration 41
curr_theta [array([1.16569146, 0.96034752]), array([1.95521401, 1.99961922])], cur_loss 11.614425385737325 iteration 42
curr_theta [array([1.18074567, 0.97139353]), array([1.97998988, 2.02608672])], cur_loss 11.015136062215296 iteration 43
curr_theta [array([1.19541499, 0.98208027]), array([2.00410627, 2.05190851])], cur_loss 10.44678066517781 iteration 44
curr_theta [array([1.20970954, 0.99241717]), array([2.02758056, 2.07710128])], cur_loss 9.907762341678321 iteration 45
curr_theta [array([1.22363919, 1.00241341]), array([2.05042966, 2.1016813])], cur_loss 9.396566675265356 iteration 46
curr_theta [array([1.23721353, 1.01207791]), array([2.07267007, 2.12566442])], cur_loss 8.91175743001197 iteration 47
curr_theta [array([1.2504419 , 1.02141938]), array([2.09431781, 2.14906606])], cur_loss 8.451972514283986 iteration 48
curr_theta [array([1.26333341, 1.03044631]), array([2.11538852, 2.17190126])], cur_loss 8.015920152897303 iteration 49
curr_theta [array([1.27589691, 1.03916694]), array([2.1358974 , 2.19418467])], cur_loss 7.602375256906629 iteration 50
curr_theta [array([1.28814103, 1.04758932]), array([2.15585927, 2.21593053])], cur_loss 7.210175980821147 iteration 51
curr_theta [array([1.30007417, 1.05572129]), array([2.17528853, 2.23715276])], cur_loss 6.838220457568839 iteration 52
curr_theta [array([1.31170451, 1.06357048]), array([2.19419924, 2.25786487])], cur_loss 6.485463702031657 iteration 53
curr_theta [array([1.32304   , 1.07114431]), array([2.21260505, 2.27808005])], cur_loss 6.150914674451372 iteration 54
curr_theta [array([1.3340884 , 1.07845003]), array([2.23051929, 2.29781114])], cur_loss 5.83363349544598 iteration 55
curr_theta [array([1.34485726, 1.0854947]), array([2.2479549 , 2.31707065])], cur_loss 5.532728804813244 iteration 56
curr_theta [array([1.35535392, 1.09228517]), array([2.26492451, 2.33587076])], cur_loss 5.24735525669549 iteration 57
curr_theta [array([1.36558555, 1.09882815]), array([2.28144039, 2.35422332])], cur_loss 4.9767111440656535 iteration 58
curr_theta [array([1.37555911, 1.10513015]), array([2.29751452, 2.37213991])], cur_loss 4.720036145857095 iteration 59
curr_theta [array([1.38528137, 1.11119751]), array([2.31315853, 2.38963178])], cur_loss 4.476609190404846 iteration 60
curr_theta [array([1.39475896, 1.11703642]), array([2.32838377, 2.4067099])], cur_loss 4.245746429194059 iteration 61
curr_theta [array([1.40399829, 1.1226529]), array([2.34320128, 2.42338494])], cur_loss 4.0267993152190344 iteration 62
curr_theta [array([1.41300564, 1.12805281]), array([2.3576218 , 2.43966731])], cur_loss 3.819152780552476 iteration 63
curr_theta [array([1.42178709, 1.13324188]), array([2.37165581, 2.45556715])], cur_loss 3.622223508002656 iteration 64
curr_theta [array([1.43034858, 1.13822566]), array([2.38531348, 2.47109433])], cur_loss 3.4354582920003085 iteration 65
curr_theta [array([1.43869589, 1.14300958]), array([2.39860476, 2.48625847])], cur_loss 3.2583324841086094 iteration 66
curr_theta [array([1.44683466, 1.14759892]), array([2.41153929, 2.50106892])], cur_loss 3.090348518786913 iteration 67
curr_theta [array([1.45477035, 1.15199881]), array([2.42412648, 2.51553482])], cur_loss 2.931034515265277 iteration 68
curr_theta [array([1.46250831, 1.15621428]), array([2.4363755 , 2.52966505])], cur_loss 2.779942951599188 iteration 69
curr_theta [array([1.47005372, 1.16025019]), array([2.44829525, 2.54346827])], cur_loss 2.636649407178976 iteration 70
curr_theta [array([1.47741165, 1.1641113]), array([2.45989443, 2.5569529])], cur_loss 2.5007513701585955 iteration 71
curr_theta [array([1.48458702, 1.16780223]), array([2.47118148, 2.57012717])], cur_loss 2.3718671064529233 iteration 72
curr_theta [array([1.49158463, 1.17132751]), array([2.48216463, 2.58299907])], cur_loss 2.2496345871244503 iteration 73
curr_theta [array([1.49840913, 1.17469152]), array([2.49285191, 2.5955764])], cur_loss 2.1337104711441466 iteration 74
curr_theta [array([1.50506509, 1.17789853]), array([2.50325111, 2.60786674])], cur_loss 2.0237691406688514 iteration 75
curr_theta [array([1.51155691, 1.18095271]), array([2.51336983, 2.61987749])], cur_loss 1.919501786123156 iteration 76
curr_theta [array([1.51788893, 1.18385813]), array([2.52321546, 2.63161585])], cur_loss 1.8206155385139797 iteration 77
curr_theta [array([1.52406531, 1.18661873]), array([2.53279522, 2.64308884])], cur_loss 1.7268326465400679 iteration 78
curr_theta [array([1.53009017, 1.18923837]), array([2.54211611, 2.65430328])], cur_loss 1.6378896961833829 iteration 79
curr_theta [array([1.53596746, 1.19172079]), array([2.55118496, 2.66526584])], cur_loss 1.553536870589461 iteration 80
curr_theta [array([1.54170107, 1.19406965]), array([2.56000841, 2.675983])], cur_loss 1.4735372481558349 iteration 81
curr_theta [array([1.54729476, 1.1962885]), array([2.56859295, 2.68646107])], cur_loss 1.3976661368567855 iteration 82
curr_theta [array([1.55275221, 1.19838082]), array([2.57694487, 2.6967062])], cur_loss 1.3257104429335291 iteration 83
curr_theta [array([1.55807698, 1.20034997]), array([2.58507031, 2.70672439])], cur_loss 1.2574680721749767 iteration 84
curr_theta [array([1.56327257, 1.20219924]), array([2.59297525, 2.71652147])], cur_loss 1.1927473621075089 iteration 85