CS767

SPR23

Assignment 3

Yiduo Feng

02/14/2023

**Problem 1(100 points)**

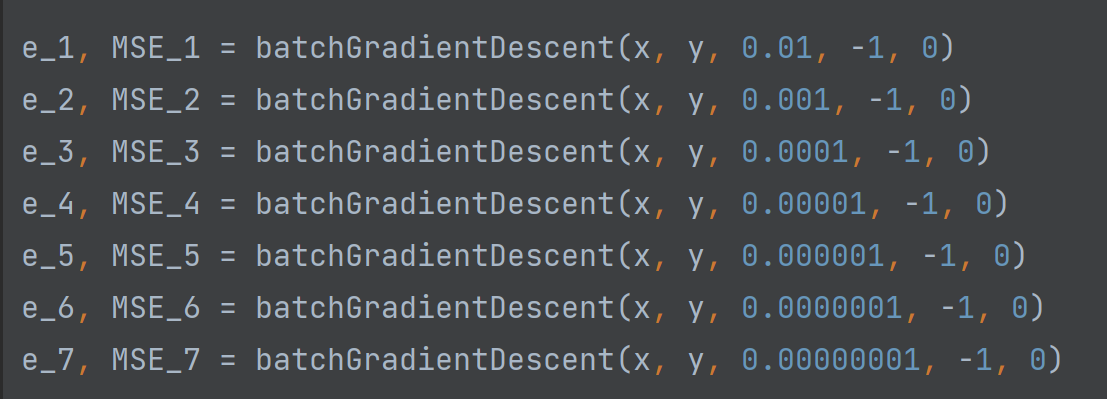
Adding Leaky ReLU activation and regularization to Assignment 2 – problem 1:

Using single layer Perceptron neural network which is connected to “Leaky ReLU” activation function with (a=0.05) to predict salary of baseball players using the data set “Assignment\_3\_Hitters.csv”. Use batch gradient descent to adjust the weights and predict salary with L2 regularization and Lasso with lambda = 0.01 and 10.

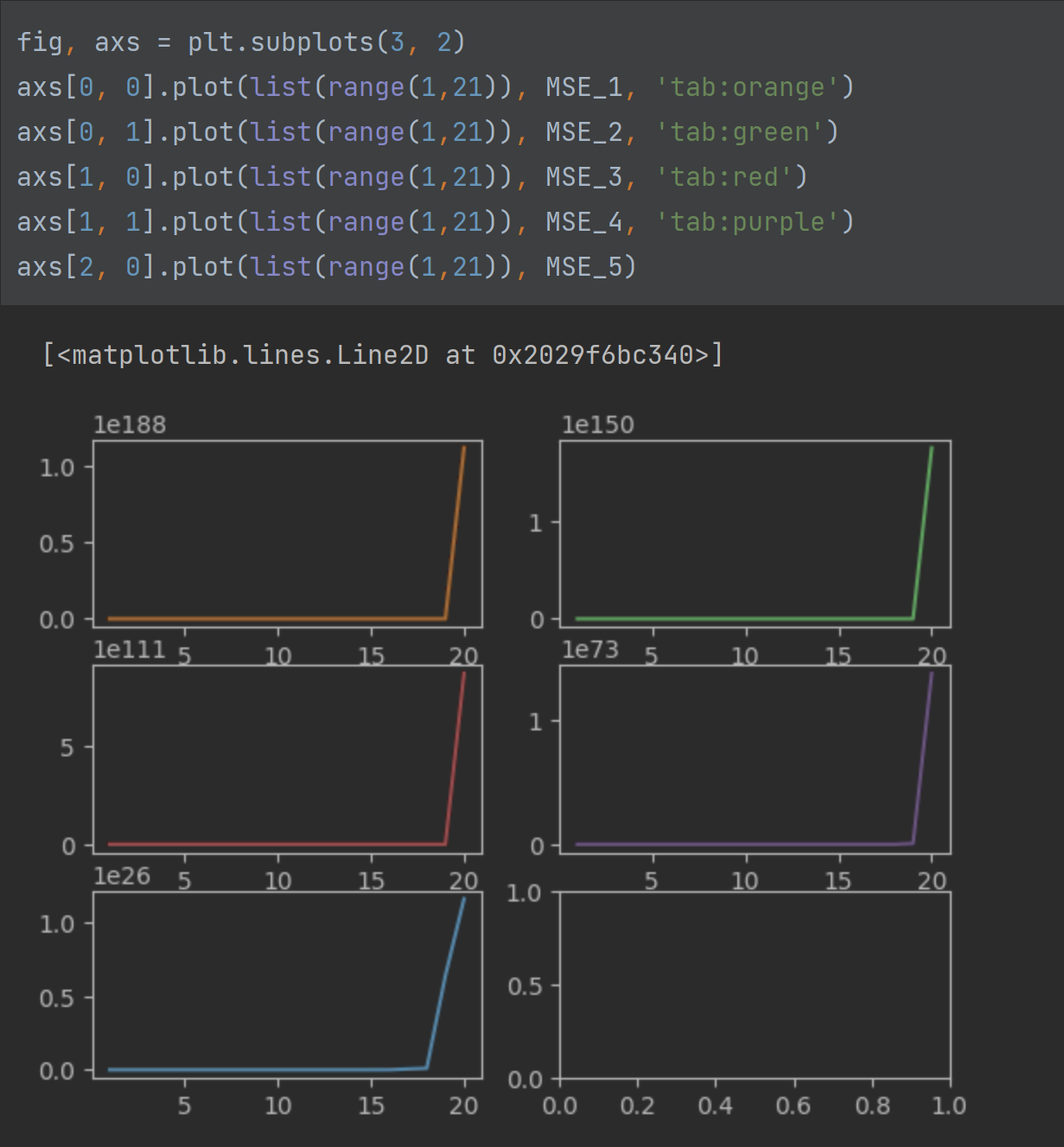
Questions:

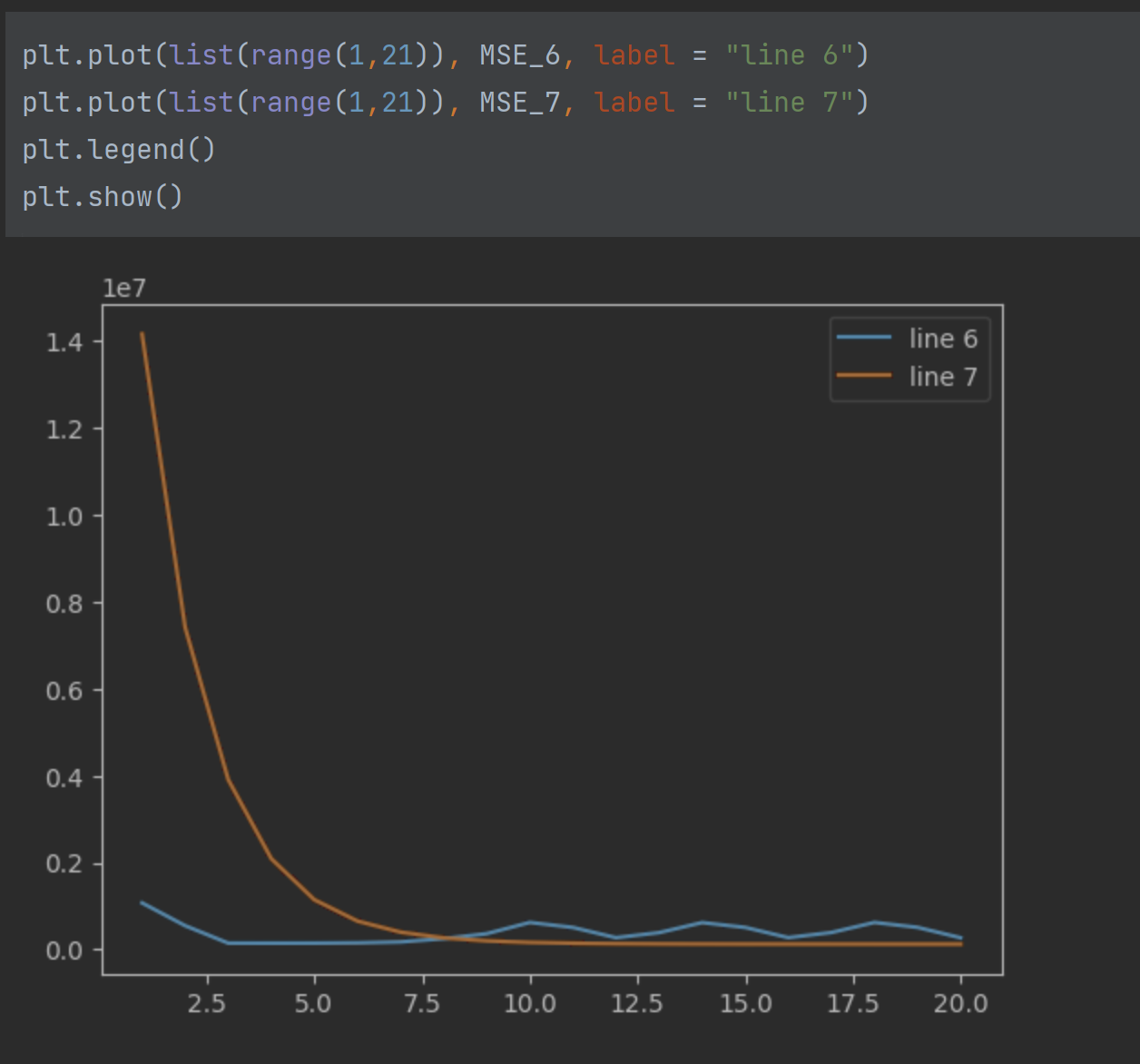
* 1. Choose a learning rate. Show details of your work and all the steps that you take to choose a suitable learning rate.

As we known, the default fixed learning rate is 0.01, so I started with 0.01.



According to the graph below, when the learning rate is from 0.01 to 0.000001, it will make the cost becomes larger and larger.

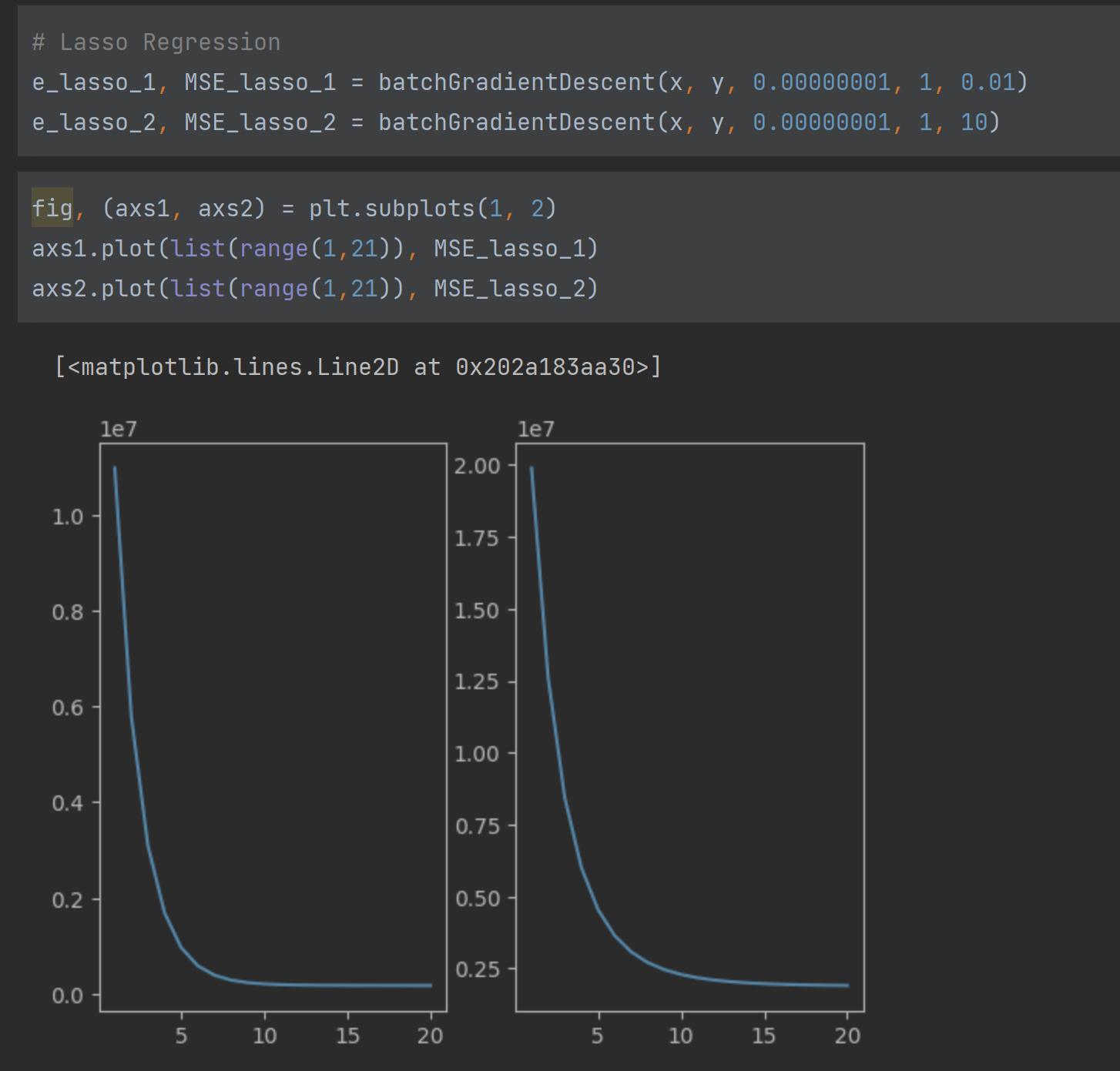




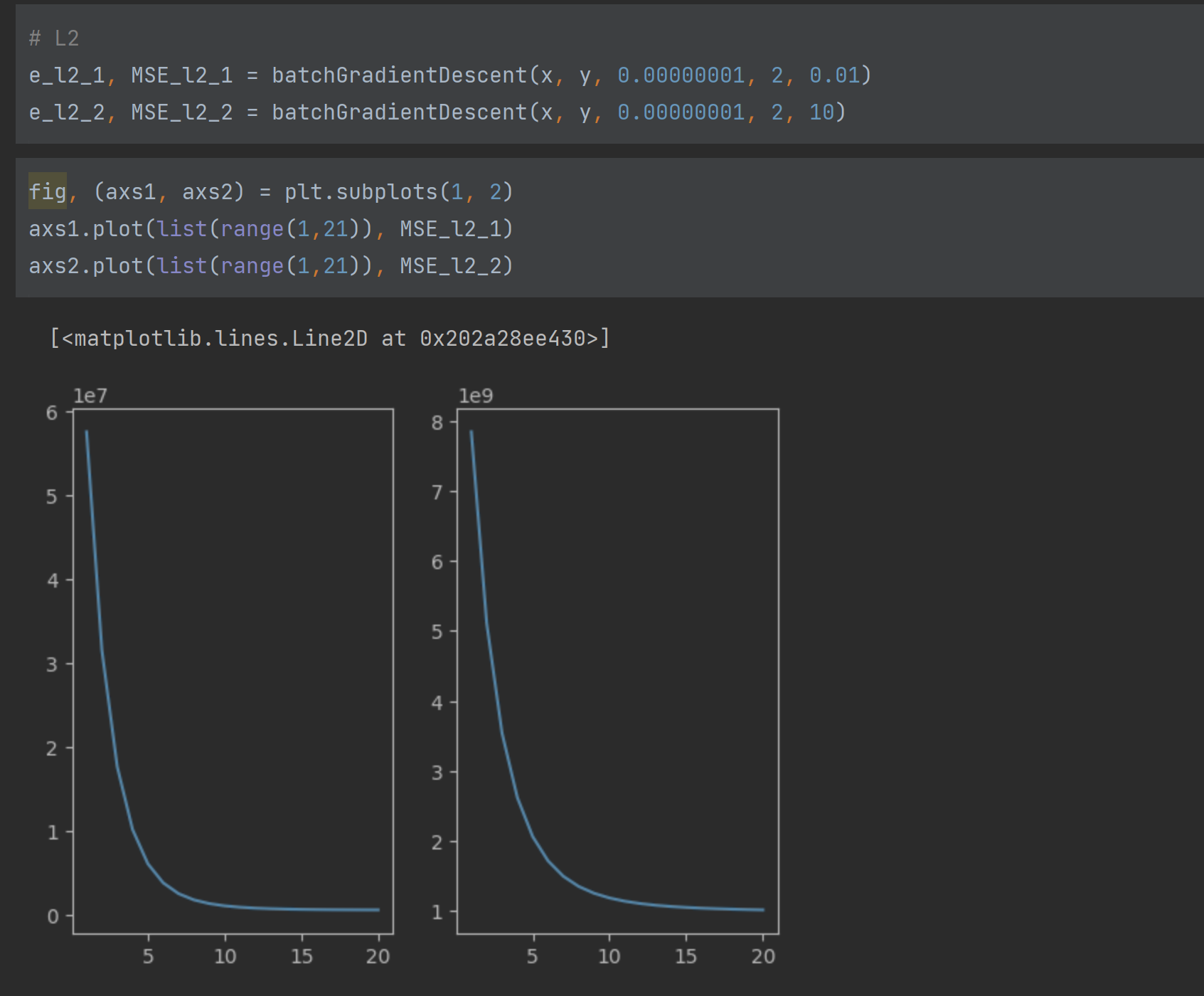
Even though 0.0000001 doesn’t make the cost larger, but the trend of line is not stable. Learning rate 0.00000001 makes a perfect decreasing line for the MSE, so I choose 0.00000001 as the learning rate.

* 1. Plot total cost (MSE - Mean Square Error + regularization cost) as a function of iterations for both regularizations.

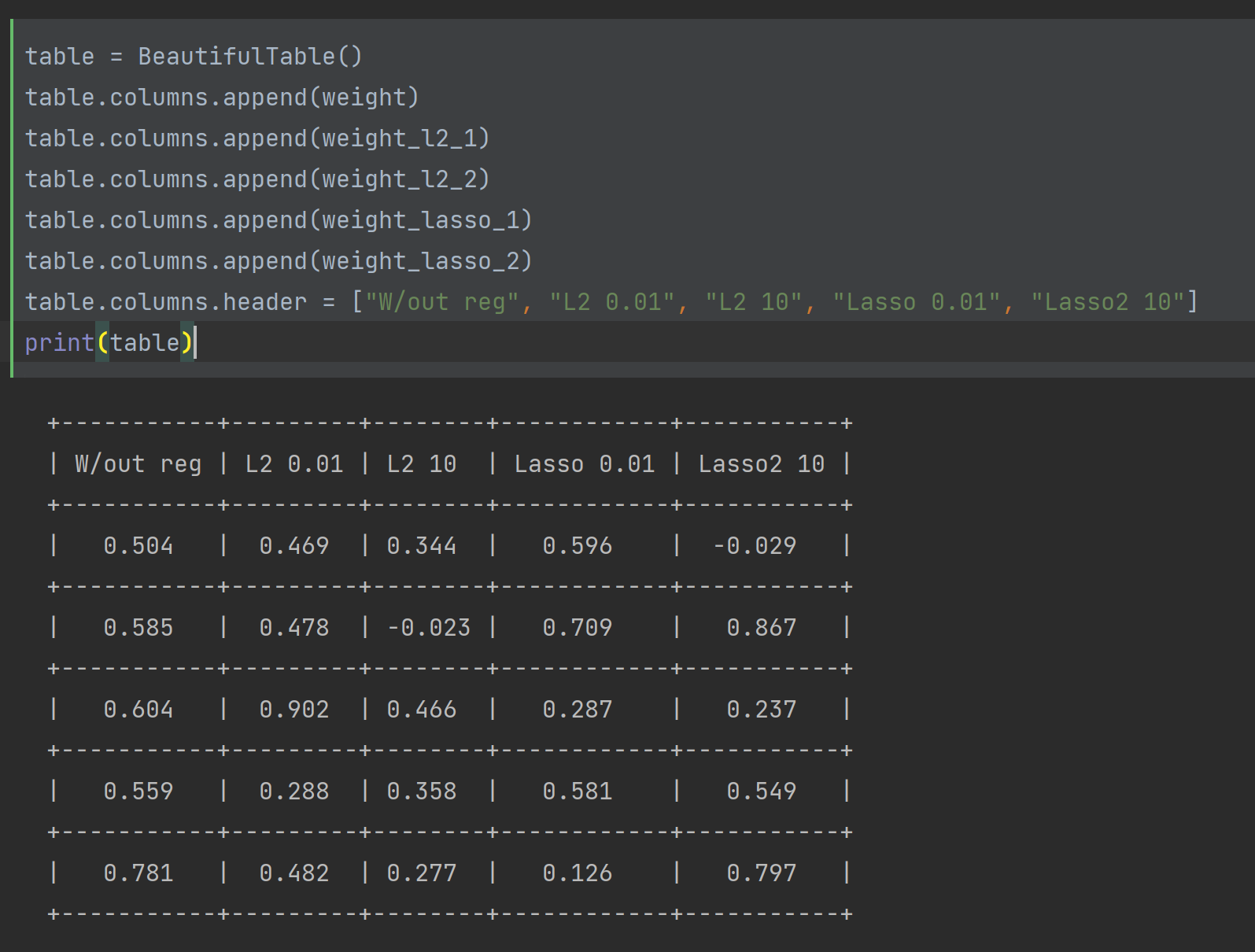
Lasso



L2 regularization



* 1. Create a table of the weights and show the final weights of the solution
     1. Without regularization
     2. With L2 regularization and two different lambdas
     3. With Lasso regularization and two different lambdas



+-----------+---------+--------+------------+-----------+

| W/out reg | L2 0.01 | L2 10 | Lasso 0.01 | Lasso2 10 |

+-----------+---------+--------+------------+-----------+

| 0.504 | 0.469 | 0.344 | 0.596 | -0.029 |

+-----------+---------+--------+------------+-----------+

| 0.585 | 0.478 | -0.023 | 0.709 | 0.867 |

+-----------+---------+--------+------------+-----------+

| 0.604 | 0.902 | 0.466 | 0.287 | 0.237 |

+-----------+---------+--------+------------+-----------+

| 0.559 | 0.288 | 0.358 | 0.581 | 0.549 |

+-----------+---------+--------+------------+-----------+

| 0.781 | 0.482 | 0.277 | 0.126 | 0.797 |

+-----------+---------+--------+------------+-----------+

| 0.266 | 0.995 | 0.403 | 0.591 | 0.456 |

+-----------+---------+--------+------------+-----------+

| 0.179 | 0.219 | 0.642 | 0.127 | 0.831 |

+-----------+---------+--------+------------+-----------+

| -0.423 | -0.333 | -0.264 | -0.347 | -0.035 |

+-----------+---------+--------+------------+-----------+

| 0.652 | 0.819 | 0.146 | 0.685 | -0.082 |

+-----------+---------+--------+------------+-----------+

| 0.427 | 0.445 | 0.463 | 0.168 | 0.124 |

+-----------+---------+--------+------------+-----------+

| 0.361 | 0.448 | 0.585 | 0.743 | 0.489 |

+-----------+---------+--------+------------+-----------+

| 0.8 | 0.428 | 0.559 | 0.393 | 0.423 |

+-----------+---------+--------+------------+-----------+

| 0.572 | 0.043 | 0.712 | -0.03 | 0.02 |

+-----------+---------+--------+------------+-----------+

| 0.398 | 0.092 | 0.06 | 0.518 | 0.179 |

+-----------+---------+--------+------------+-----------+

| 0.822 | 0.039 | 0.815 | 0.809 | 0.254 |

+-----------+---------+--------+------------+-----------+

| 0.83 | 0.589 | 0.865 | 0.267 | 0.247 |

+-----------+---------+--------+------------+-----------+

| 0.663 | 0.225 | 0.806 | 0.904 | 0.82 |

+-----------+---------+--------+------------+-----------+

| 0.849 | 0.52 | 0.91 | 0.051 | 0.269 |

+-----------+---------+--------+------------+-----------+

| 0.304 | 0.456 | 0.485 | 0.741 | 0.266 |

+-----------+---------+--------+------------+-----------+

So the final weight for w/out regularization is 0.304, for L2 0.01 is 0.456, for L2 10 is 0.485, for lasso 0.01 is 0.741, for lasso 10 is 0.266.