Z5305320

Dong AO

(a):

Note: means the size of

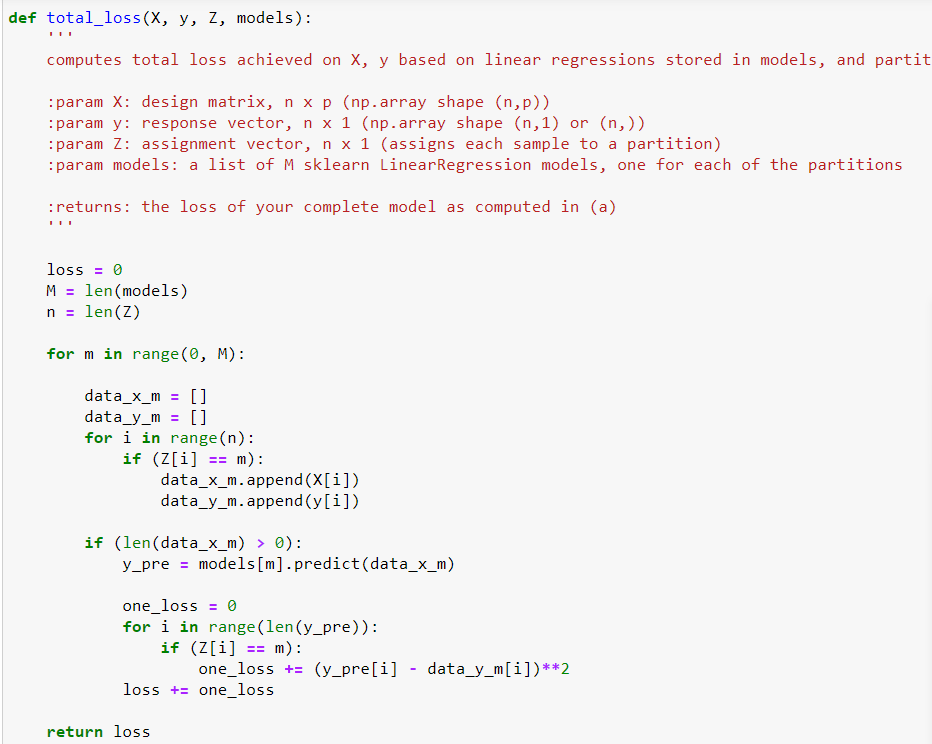
(b):

1. If are fixed: For any data point in . We calculate the m values: and we find the m which makes the least . We assign the point to .

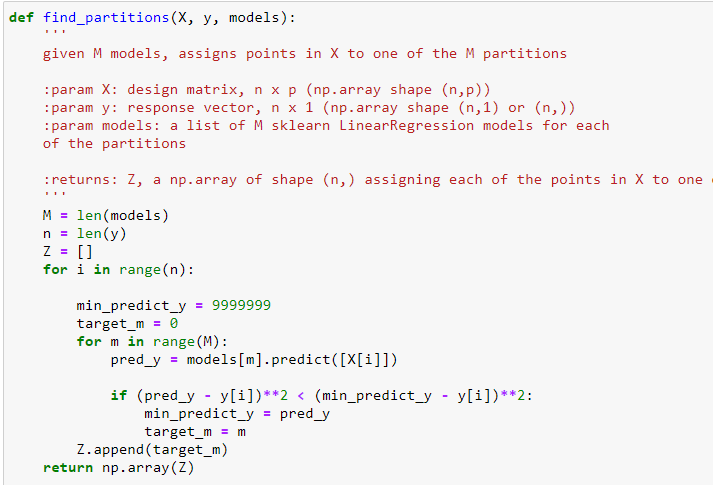
2. If are fixed: For every data point , we will try to minimize , which is to learn here.

3. Return to process 1 and repeat until betas are not changed any more compared to last iteration.

(c)

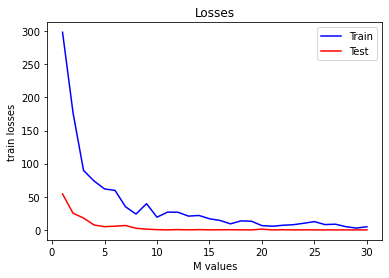


(d):



(e):

I choose 20 since from the plot we can see it is the only one that train loss is not increasing at that point and test loss is small enough.



|  |  |  |
| --- | --- | --- |
| M | Train loss | Test loss |
| 5 | 61.86355743 | 4.917106440992632 |
| 10 | 19.21842497 | 0.3965816078482734 |
| 15 | 16.82499839 | 0.128809950913761 |
| 20 | 6.47424378 | 1.1235300096590044 |
| 25 | 12.5374972 | 0.0096851771840851 |
| 30 | 4.83253453 | 0.0411392935295462 |

