

Coding Assignment 1

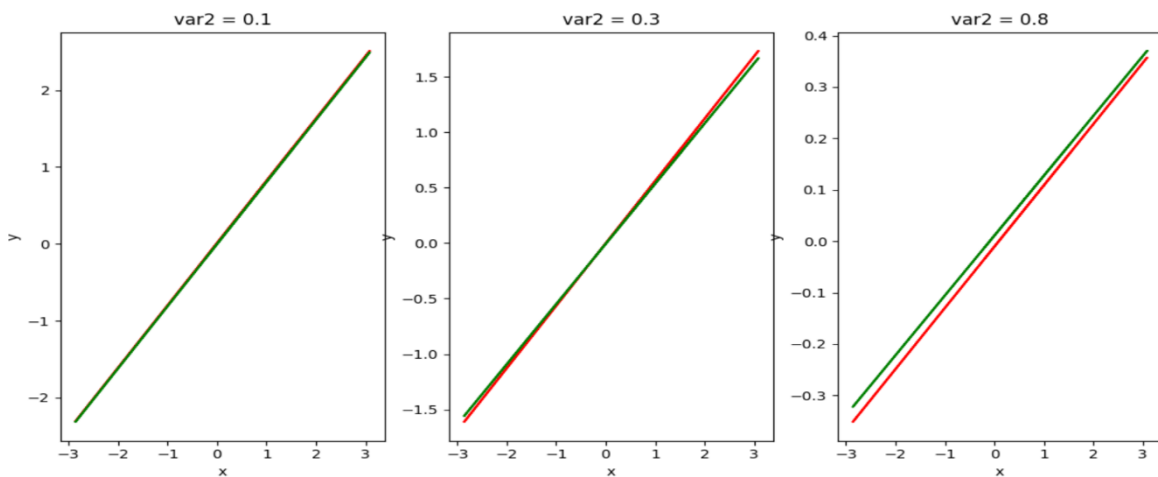
Yulong Zhou

Problem 1:

1)

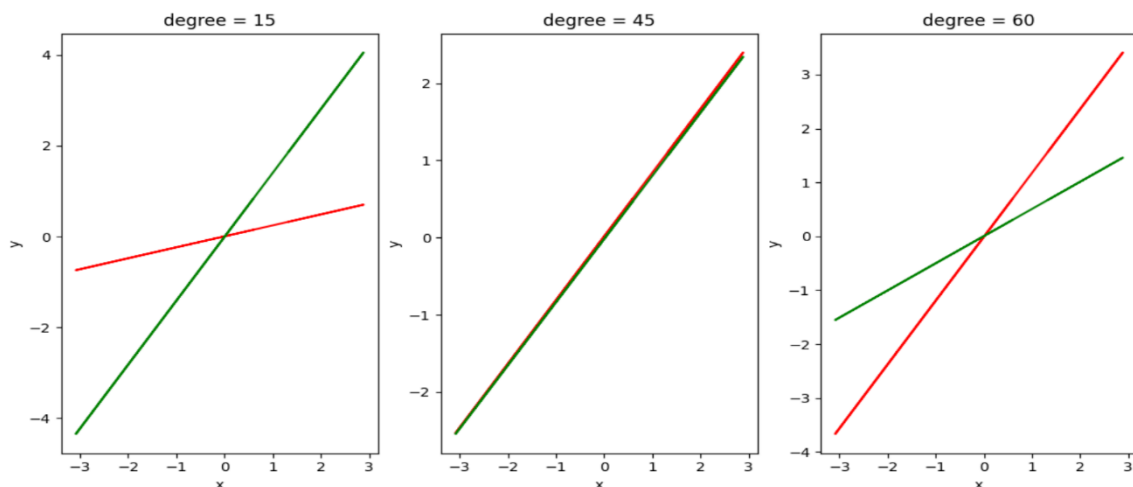
```
Predicting y from x (x2y): w_x2y=0.5301259601642684 b_x2y = -0.011106422874893652  
Predicting x from y (y2x): w_y2x=0.5380043392947025 b_y2x = 0.022663729411151
```

2)



3)

When we only change the “var2”, the linear regression models’ result will not be influenced. (red line is x2y, green line is y2x)



4)

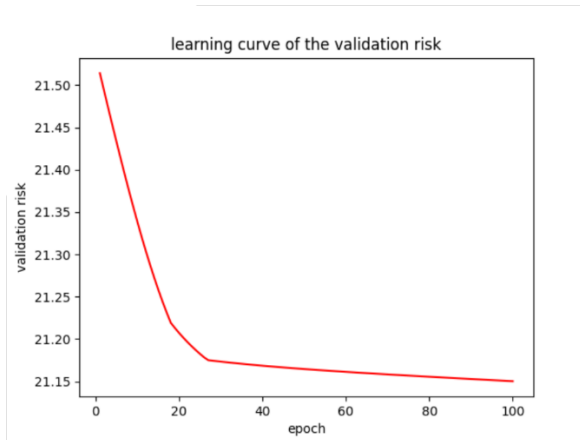
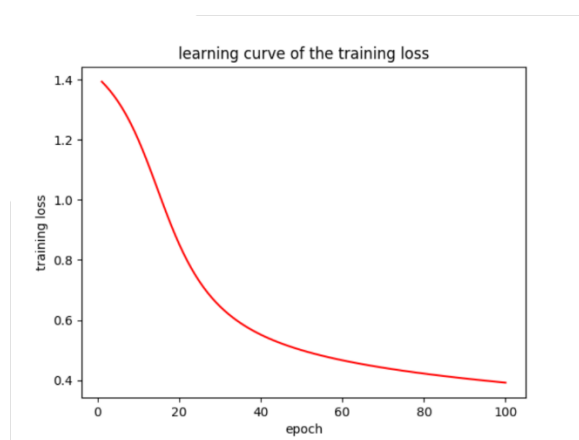
When we only change the “degree”, the linear regression models’ result will be

influenced. When the degree is much smaller or bigger, the output will be more flatter. If the degree is smaller than 45° , the x_2y (red line) will be influenced, and if the degree is larger than 45° , the y_2x (green line) will be influenced.

Problem 2:

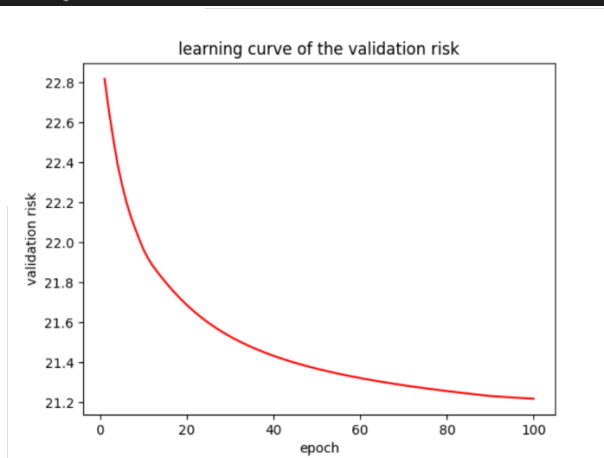
a)

```
The number of epoch that yields the best validation performance is : 100
The validation performance (risk) in that epoch is : 21.15027084533993
The test performance (risk) in that epoch is : 21.585162329099735
```



b)

```
The best hyperparameter weight decay is : 0.01
The number of epoch that yields the best validation performance is : 100
The validation performance (risk) in that epoch is : 21.218459636480148
The test performance (risk) in that epoch is : 21.42843558711758
```



c)

Q: What will happen if the learning rate is getting larger?

A: Suppose the learning rate is 1 which is 1,000 times larger than the original setting.

The gradient may oscillate back and forth near the minimum, and may even fail to converge. From the plot the model converges too quickly to a suboptimal solution. Therefore, we should be careful when we setting the learning rate.

```
The number of epoch that yields the best validation performance is : 100
The validation performance (risk) in that epoch is : 7.620198868737317
The test performance (risk) in that epoch is : 7.045814471442854
```

