

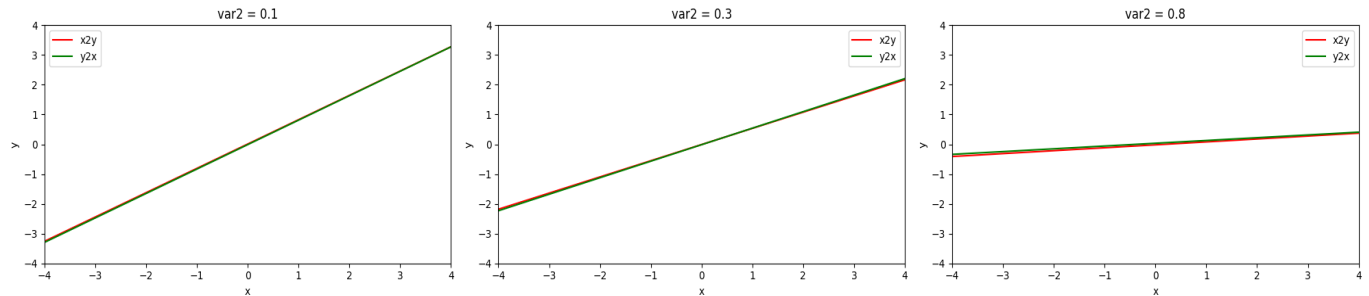
Problem 1

(1)

The result of Q1(1):

Predicting y from x (x2y): weight=0.5167201568567926 bias = 0.014639177929737807
Predicting x from y (y2x): weight=0.5219077671061749 bias = -0.029939410468034916

(2)

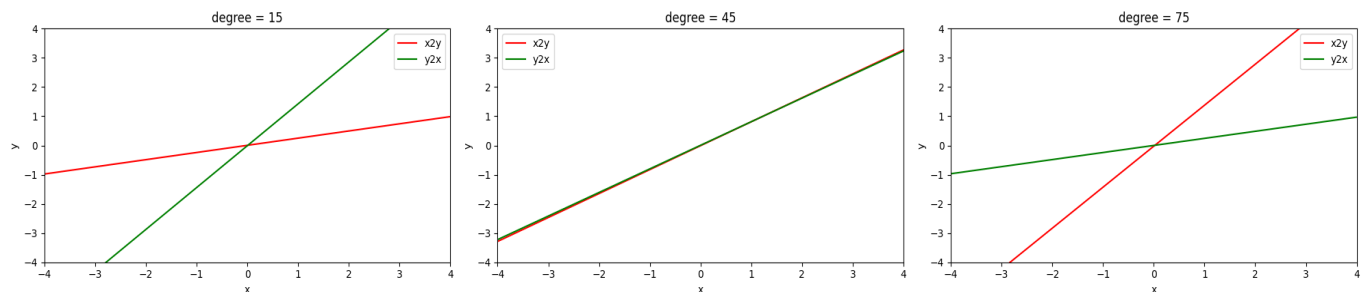


(3)

From the regression models in (2) we can conclude that when we change the value of var2 , $x2y$ and $y2x$ still nearly overlap, which means the parameters of the regression models will not be influenced by changing the value of var2 .

So the description of the phenomena found in (1) and (2) is that set different values of var2 can not affect the output of regression models.

(4)



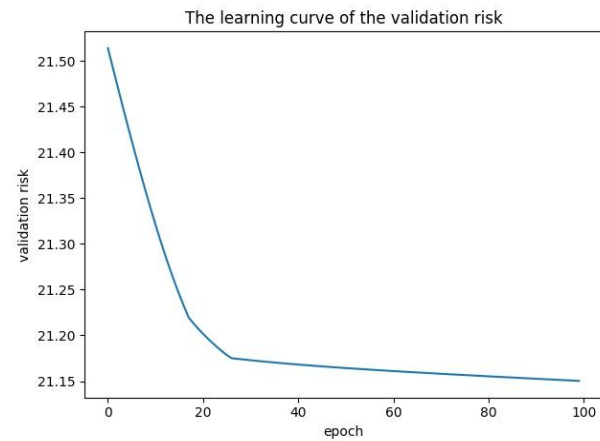
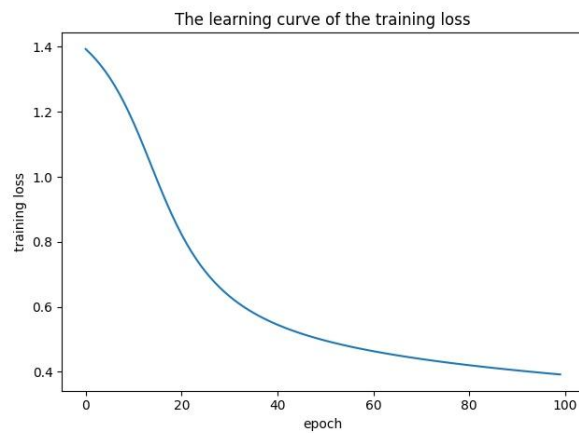
Since we design a controlled experimental protocol and set three different degree values 15, 45, 75. From the regression models in (3) we can conclude that when we change the value of degree , only when $\text{degree}=45$ can $x2y$ and $y2x$ overlap, which means the parameters of the regression models will be influenced by changing the value of degree .

So we can find that setting different values of degree can affect the output of regression models. The greater the deviation between the acquired angle and 45° , the less fitting.

Problem 2

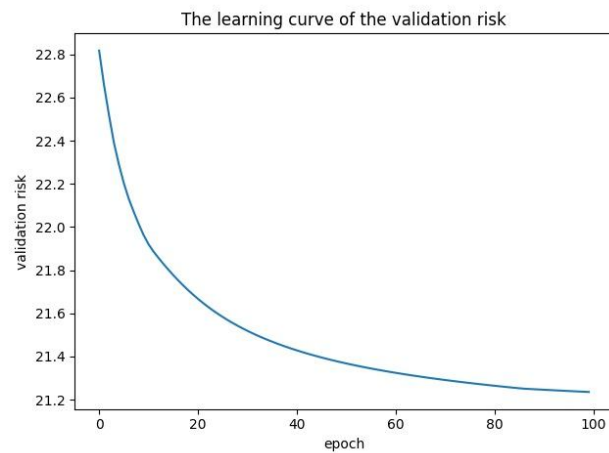
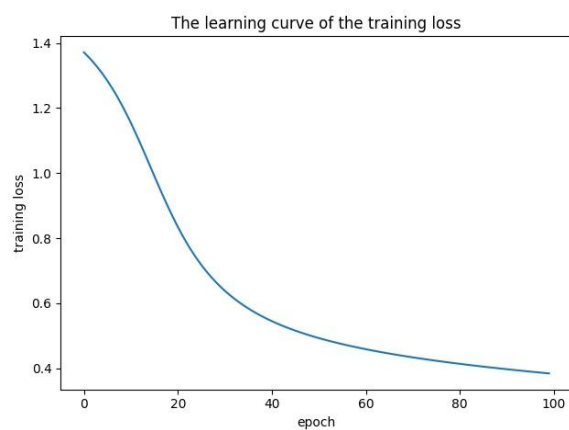
(a)

1. The number of epoch that yields the best validation performance: 100
2. The validation performance (risk) in that epoch: 21.15027084533993
3. The test performance (risk) in that epoch: 21.585162329099735



(b)

1. the best hyperparameter, which is the one yields the best performance: 0.01
2. The number of epoch that yields the best validation performance: 100
3. The validation performance (risk) in that epoch: 21.235882001489603
4. The test performance (risk) in that epoch: 21.463675573856772



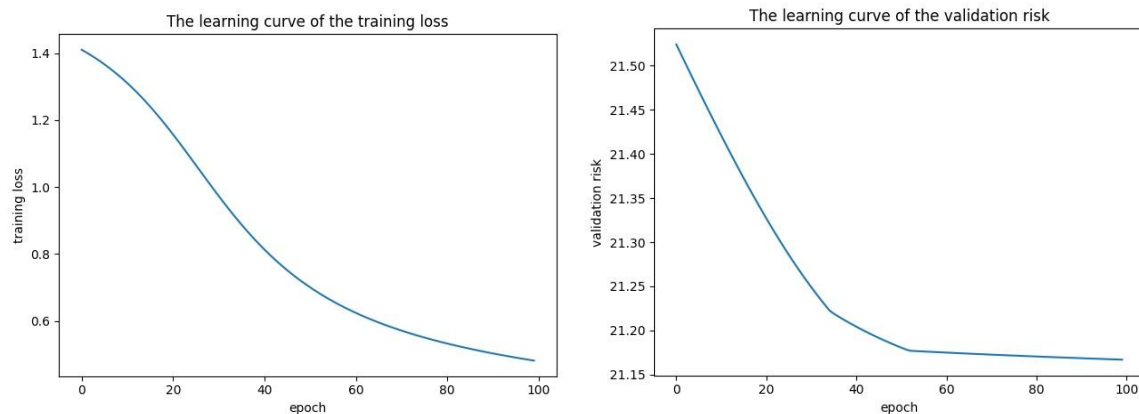
(c)

Question: We can explore when the value of `batch_size` increases, if it will affect mini-batch gradient descent and the performance of linear regression.

Design: We can let `batch_size` increase to 20, then report three numbers and two plots.

Results:

1. The number of epoch that yields the best validation performance: 100
2. The validation performance (risk) in that epoch: 21.166918372418042
3. The test performance (risk) in that epoch: 21.58596806195461



Conclusion: Comparing the results to 2(a), we can find that the three numbers did not change a lot. And the slope of the two plots has slowed down. So we can conclude that when the value of `batch_size` increases, it will affect mini-batch gradient descent, but will not affect the performance of linear regression.