Problem 1:

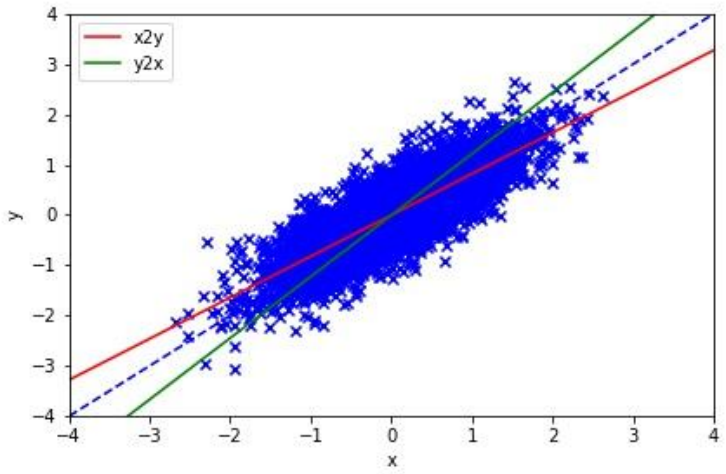
a)

W_x2y = 0.52666 b_x2y = -0.02335

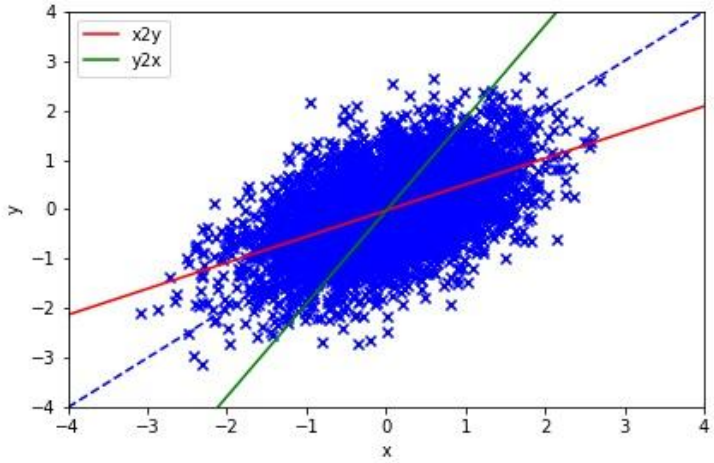
W_y2x=0.53370   b_y2x = 0.01818

b)

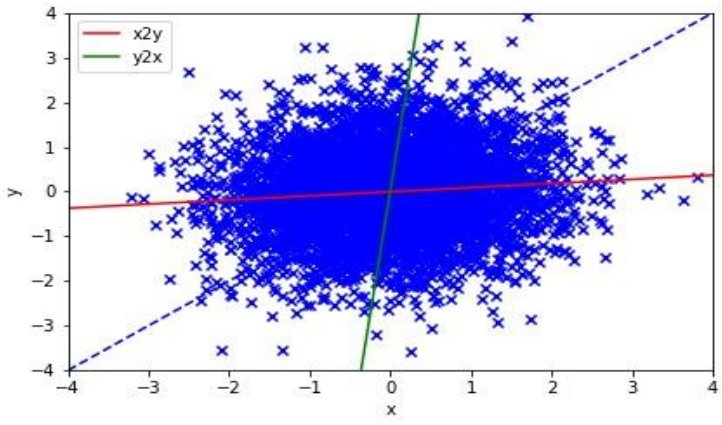
Var2 = 0.1



Var2 = 0.3



Var3 = 0.8

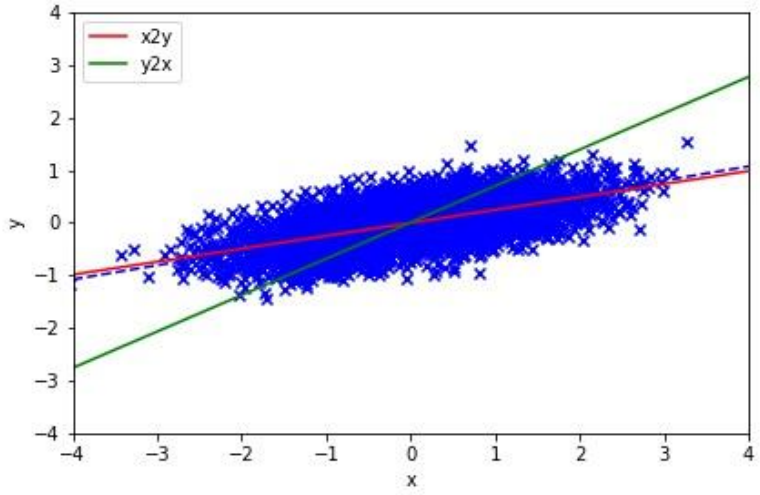


c).

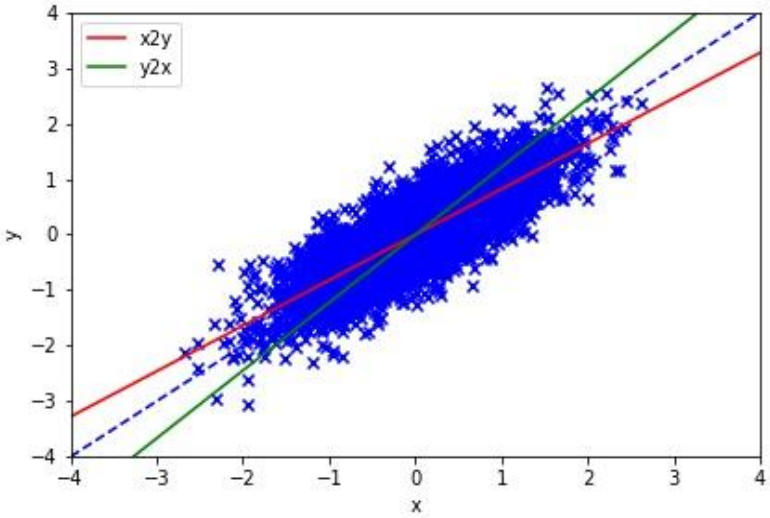
1. When the variance is small, the slope of x2y and y2x are nearly identical. However, when the variance is large, the slope of x2y is nearly horizontal, and the slope of y2x is nearly vertical.
2. When the variance is small, the data points are concatenated together. When the variance is large, the data points are scattered around the graph.

d).

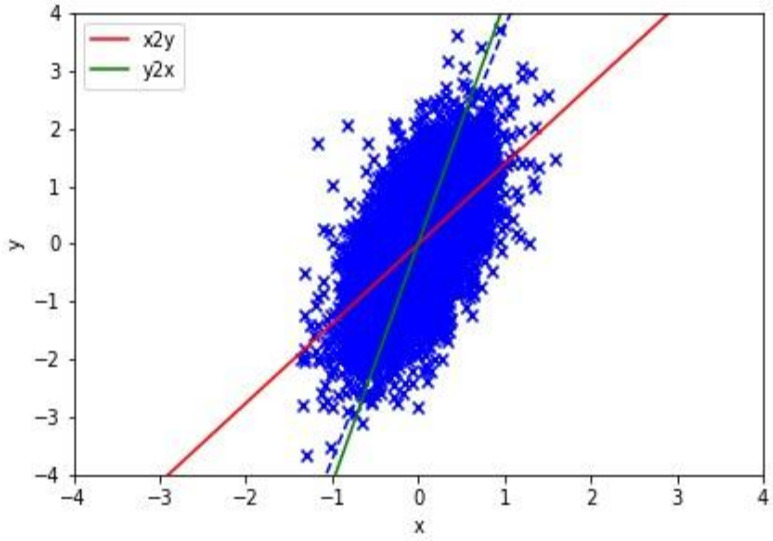
degree = 15°



degree = 45°



degree = 75°



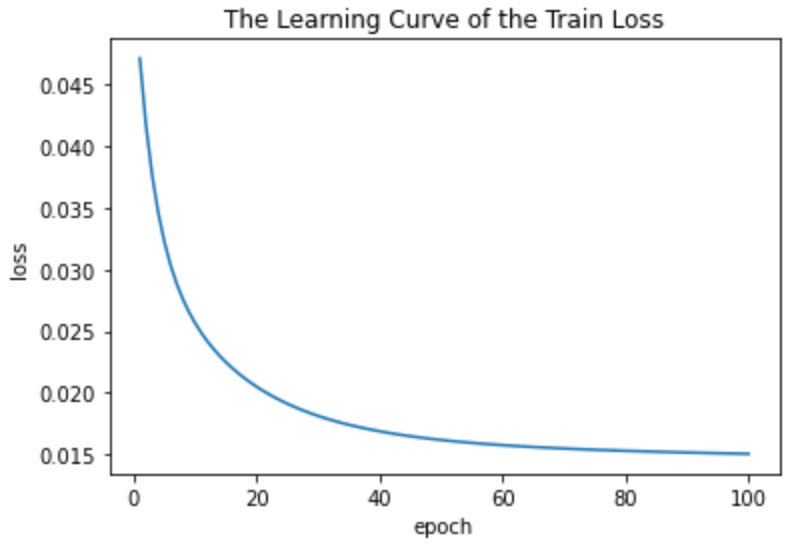
Experiment protocol:
Adjust rotation degree with 15, 45, 75 respectively, and keep all other hyperparameters unchanged (M = 5000, var1 = 1, var2 = 0.1).

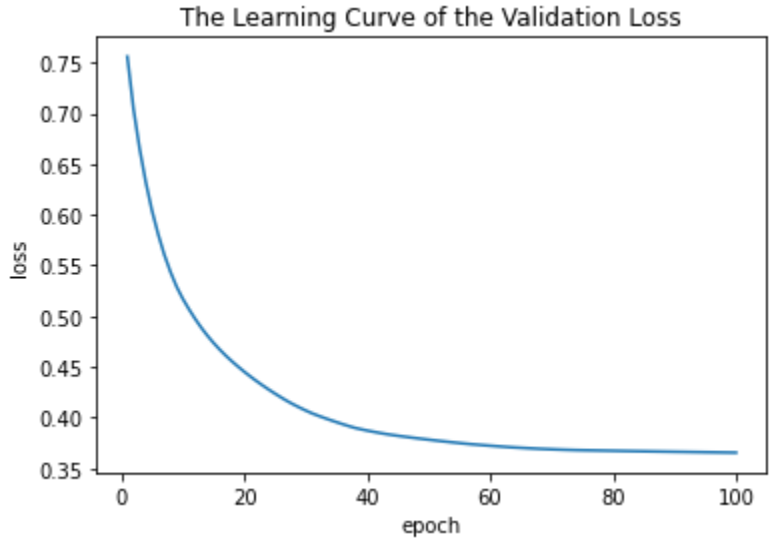
Findings:

1. The larger the rotation degree, the steeper the slope of x2y and y2x.
2. When the degree = 45, the difference between the slope of x2y (w_x2y) and the slope of y2x(w_y2x) is the smallest.
   When the degree gets larger/smaller, and all other params are the same, the difference between the slope of x2y (w_x2y) and the slope of y2x(w_y2x) would become larger.

Problem 2:
a).

1. The number of epoch that yields the best validation performance: 99
2. The validation performance (risk) in epoch 99: 0.36549
3. The test performance (risk) in epoch 99: 0.35231

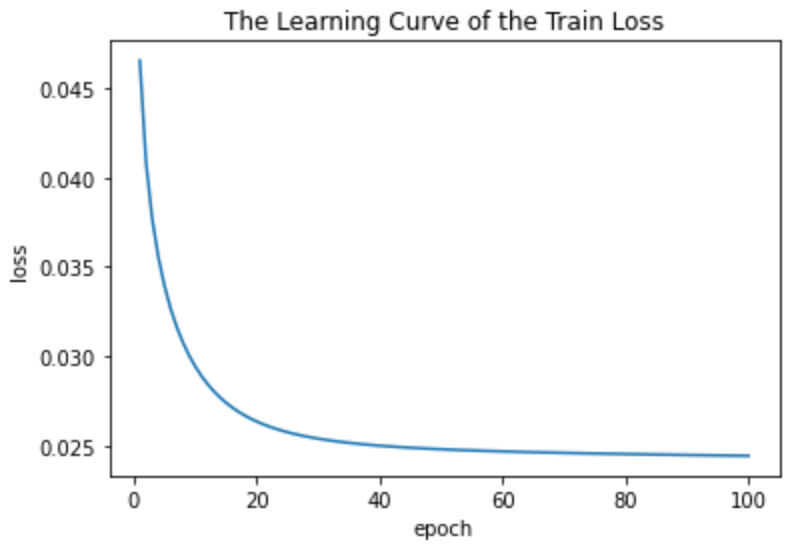


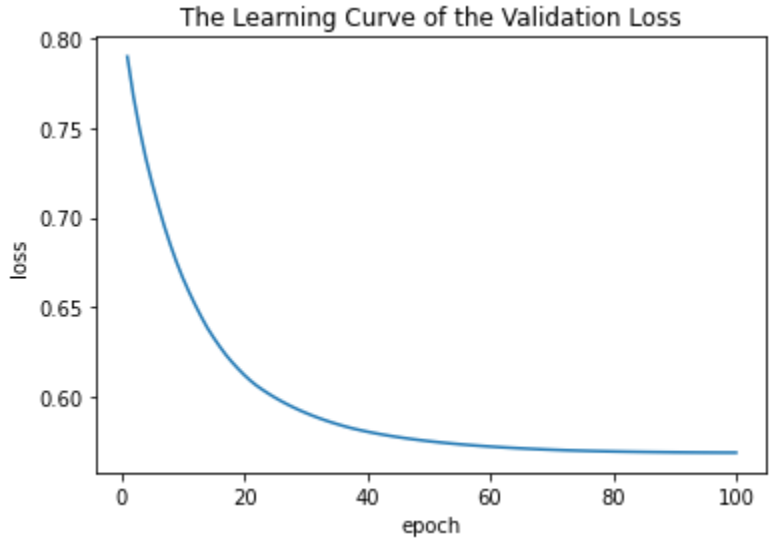


b).

The best lambda decay: 0.01

1. The number of epoch that yields the best validation performance: 99
2. The validation performance (risk) in epoch 99: 0.56852
3. The test performance (risk) in epoch 99: 0.45057





c).
Question: would different learning rates(alpha) affect the model's performance during training?
Experiment: Tune the learning rate from the set (0.1, 0.001, 0.00001), and keep other parameters unchanged. Record the test risk for different learning rate.

Observation:
For alpha=0.1, best epoch = 0, validation risk on epoch 0 = 0.39946, test risk = 0.39035
For alpha=0.001, best epoch=99, validation risk on epoch 99=0.36549, test risk=0.35231
For alpha=0.00001,best epoch=99, validation risk on epoch 99=0.75612, test risk=0.59515

Conclusion:
The learning rate would affect the model's performance. If the learning rate is too large or too small, the model may not converge to optimum.