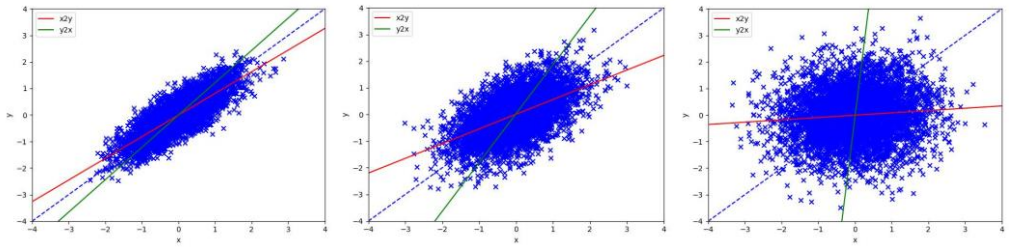
CMPUT 466 Coding Assignment 1

Jiannan Lu

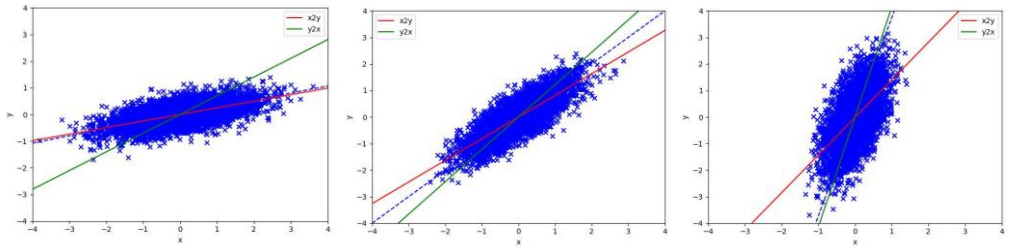
Problem 1

1) Predicting y from x (x2y): weight = 0.5521163600650316 bias = 0.009152652241253464
   Predicting x from y (y2x): weight = 0.5456829379944118 bias = -0.016634133136861566
2) var2 = 0.1 var2 = 0.3 var2 = 0.8



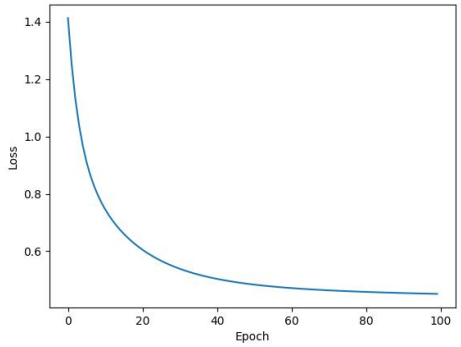
3) We made changes on the variance on b axis, the result is when the variance gets lower, the model fits the data better. The x2y line gets closers to the y2x line, and the true regression line as well.
4) Since the experiment is aim to find the effects of changing the degree of rotation, we take three samples to observe the effect at degrees 15, 45, and 75 of rotation respectively. As the plots shown below, with a low degree of rotation, the x2y model performs better. With the increase of the degree of rotation, the performance of x2y model decreases and the one of y2x model increases.

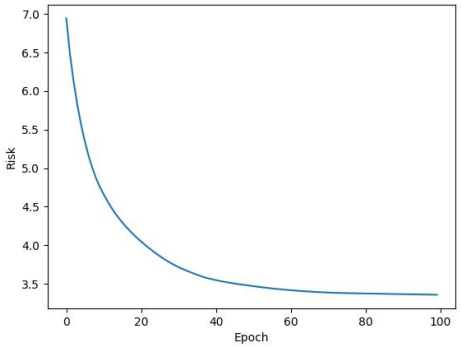
   degree = 15 degree = 45 degree = 75



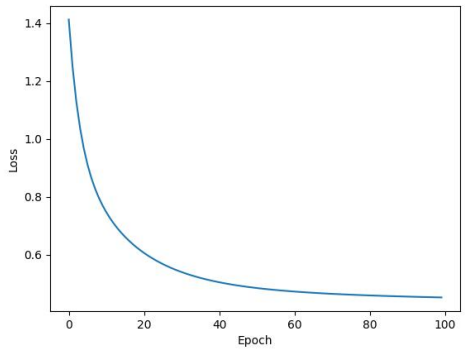
Problem 2

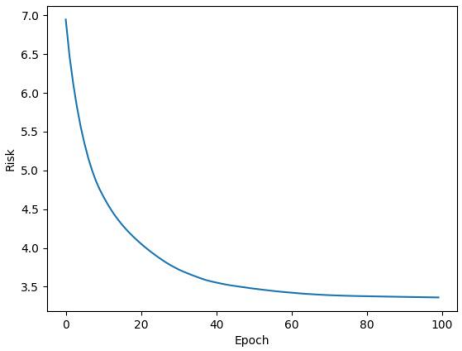
a) Best Epoch: 99
   Vali Risk: 3.3580816862888305
   Test Risk: 3.2370463070784026





b) Best Lambda: 0.01
Best Epoch: 97
Vali Risk: 3.361206850296914
Test Risk: 3.230761371999215





c) How will the change in step size affect the performance in practice? Is the final performance strictly better with the lower step size (alpha)? Apply λ=0.01, alphas = (0.05, 0.01, 0.005, 0.001, 0.0005), other parameters remain the same.

Result:

Alpha = 0.05
Best Epoch: 4
Vali Risk: 3.5178046733067374
Test Risk: 3.435732844095366

Alpha = 0.01
Best Epoch: 31
Vali Risk: 3.324587875046614
Test Risk: 3.3081407120415434

Alpha = 0.005
Best Epoch: 62
Vali Risk: 3.320181998491911
Test Risk: 3.303011305766671

Alpha = 0.001
Best Epoch: 99
Vali Risk: 3.361206850296914
Test Risk: 3.230761371999215

Alpha = 0.0005

Best Epoch: 99

Vali Risk: 3.4815731191892763

Test Risk: 3.164405949995249

From the data, we can observe that with the decreasing in alpha from 0.05 to 0.005, final performance gets better, and gradient descent reaches its optimum with epoch number less than 99. And with the decreasing in alpha from 0.005 to 0.0005, final performance gets worse.

The reason is that, in general, performance gets better with a decreasing alpha, however, the MaxIter is fixed for this practice. Too small of a step size may increase the training loss as we observes because it takes more than MaxIter epochs for the gradient descent to reach its optimum, hence lower the final performances (higher the Risks).