Marco Lin

A. Logistic Regression with Batch Gradient Descent

1. Errors in sample and out sample

With learning rate: 0.1

Classification error: in sample: 12

Out sample: 9

Likelihood error: in sample: 0.1

Out sample: 1.68

```
learning rate: 0.1
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Batch gradient descent completed in: 61
Classification Error in: 12.9762
Classification Error out: 9.9957
likehood Error in: 0.10304
likehood Error out: 1.6866
```

The number of the classification error in sample greater than out of sample because the training set was 4000 points, and the test set was only 1000. The size of the dataset will affect the value of error.

```
learning rate: 1
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Batch gradient descent completed in: 1000
Classification Error in: 86047.4589
Classification Error out: 38410.4571
likehood Error in: 0.70375
likehoodEin Error out: 122.7693
```

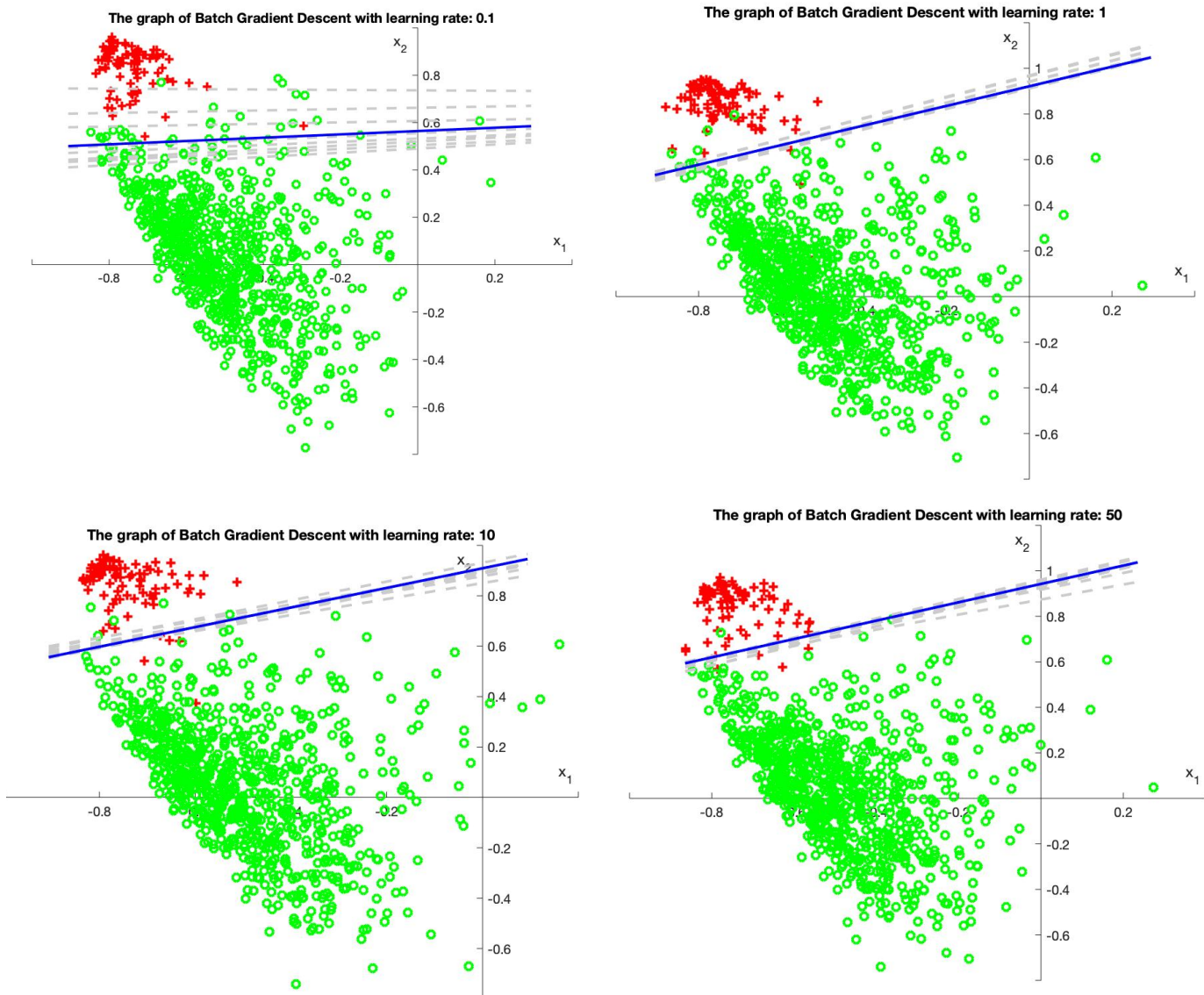when the learning rate, which is greater than 1, the iteration enhances. And the error also increases. In conclusion, the best learning rate is 0.1.

2. Stopping criteria

The stopping criteria of batch gradient is that, when the error has not changed significantly in the last 60 times, the loop will be ended. In addition, for avoiding the endless loop, when iteration is greater than 1000 times, the loop will stop.

3. Learning rate

When the learning rate greater than one, the number of the iteration will increase. it causes higher error, and unpredictable result. The following figure is the batch gradient with learning rate 0.1.

The graph of Batch Gradient Descent with learning rate: 0.1

The graph of Batch Gradient Descent with learning rate: 1

The graph of Batch Gradient Descent with learning rate: 10

The graph of Batch Gradient Descent with learning rate: 50

B. Logistic Regression with Stochastic Gradient Descent

1. Errors in sample and out sample

With learning rate: 0.1

Classification error: in sample: 0.631

Out sample: 0.6983

Likelihood error: in sample: 0.745

Out sample: 0.77522

```
learning rate: 0.1
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Classification Error in: 0.63163
Classification Error out: 0.69834
likehood Error in: 0.74596
likehood Error out: 0.77522
```

With learning rate:10

Classification error: in sample: 17

Out sample: 0.4

Likelihood error: in sample: 0.72

Out sample: 0.43

```
learning rate: 10
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Stochastic Gradient descent completed in: 2
Classification Error in: 17.6908
Classification Error out: 0.49855
likehood Error in: 0.72365
likehood Error out: 0.43427
```

With learning rate:50

Classification error: in sample: 457

Out sample: 8.7

Likelihood error: in sample: 3.0

Out sample: 0.452

```
learning rate: 50
Stochastic Gradient descent completed in: 6
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Stochastic Gradient descent completed in: 4
Classification Error in: 457.0209
Classification Error out: 8.7124
likehood Error in: 3.0441
likehoodEin Error out: 0.45274
```

2. Stopping criteria

I chose when the gradient is small enough, or it has not changed significantly in the last several times

3. Learning rate

Comparing with batch gradient descent, with learning rate which are 0.1, 1, 10, the results of the error are still acceptable. The following are the figures with different learning rates.
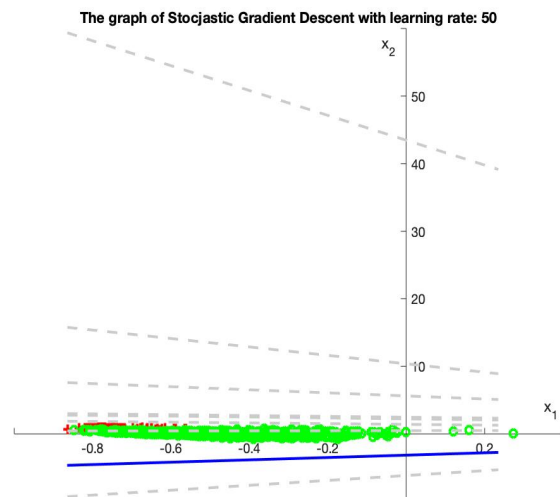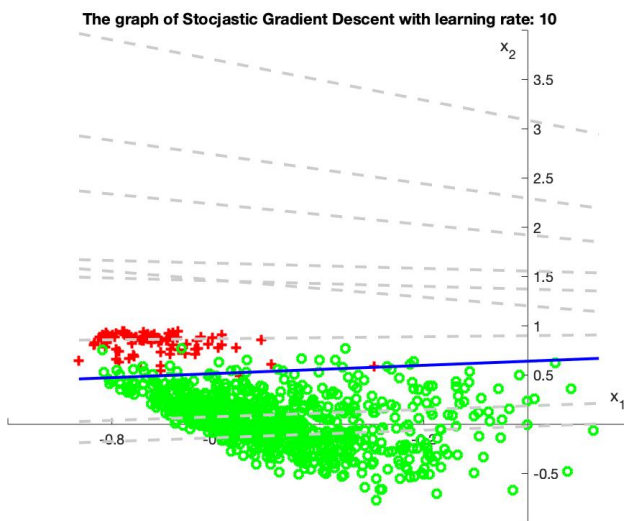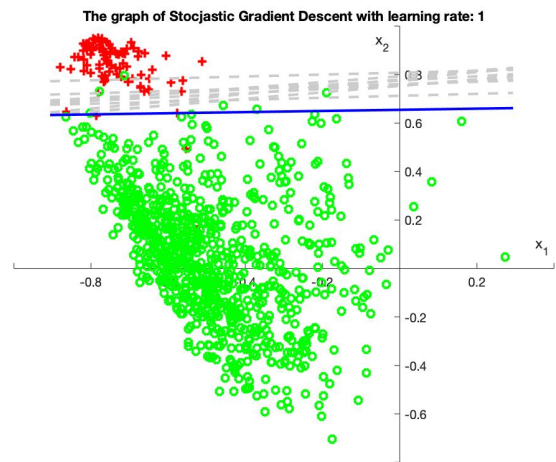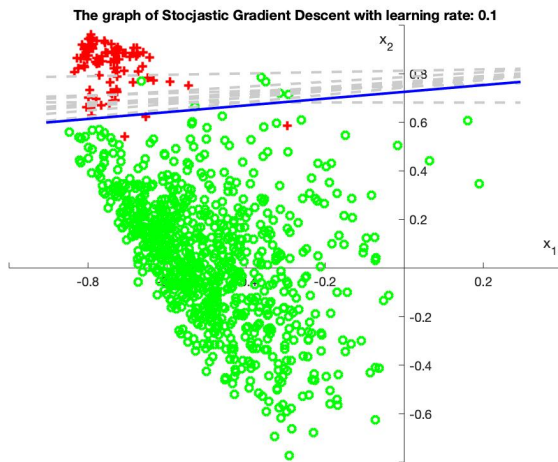
It shows that with the learning rate 50, the gradient was moving too far that makes harder to fine the correct line.