CSCI3320 Programming Assignment 2
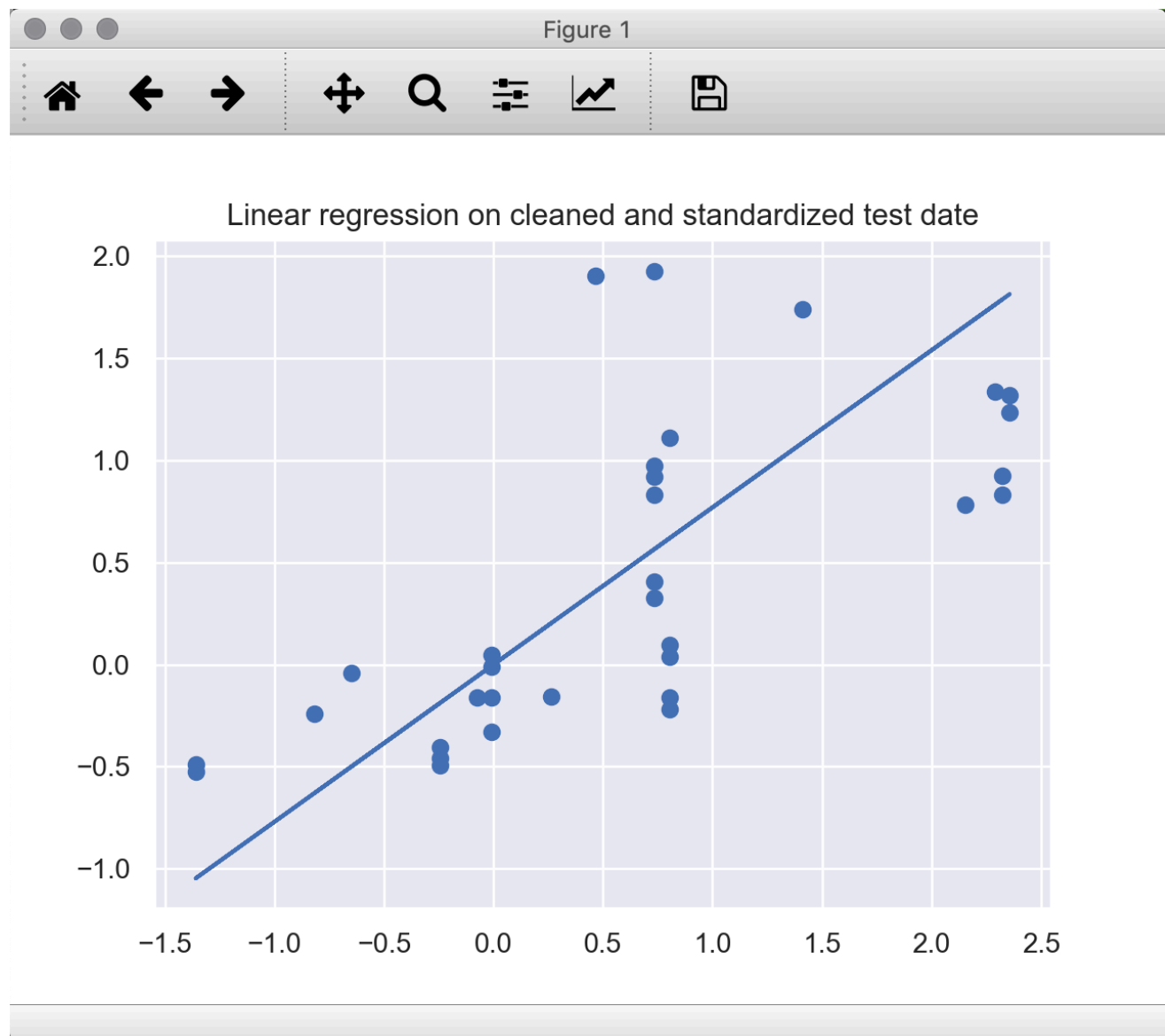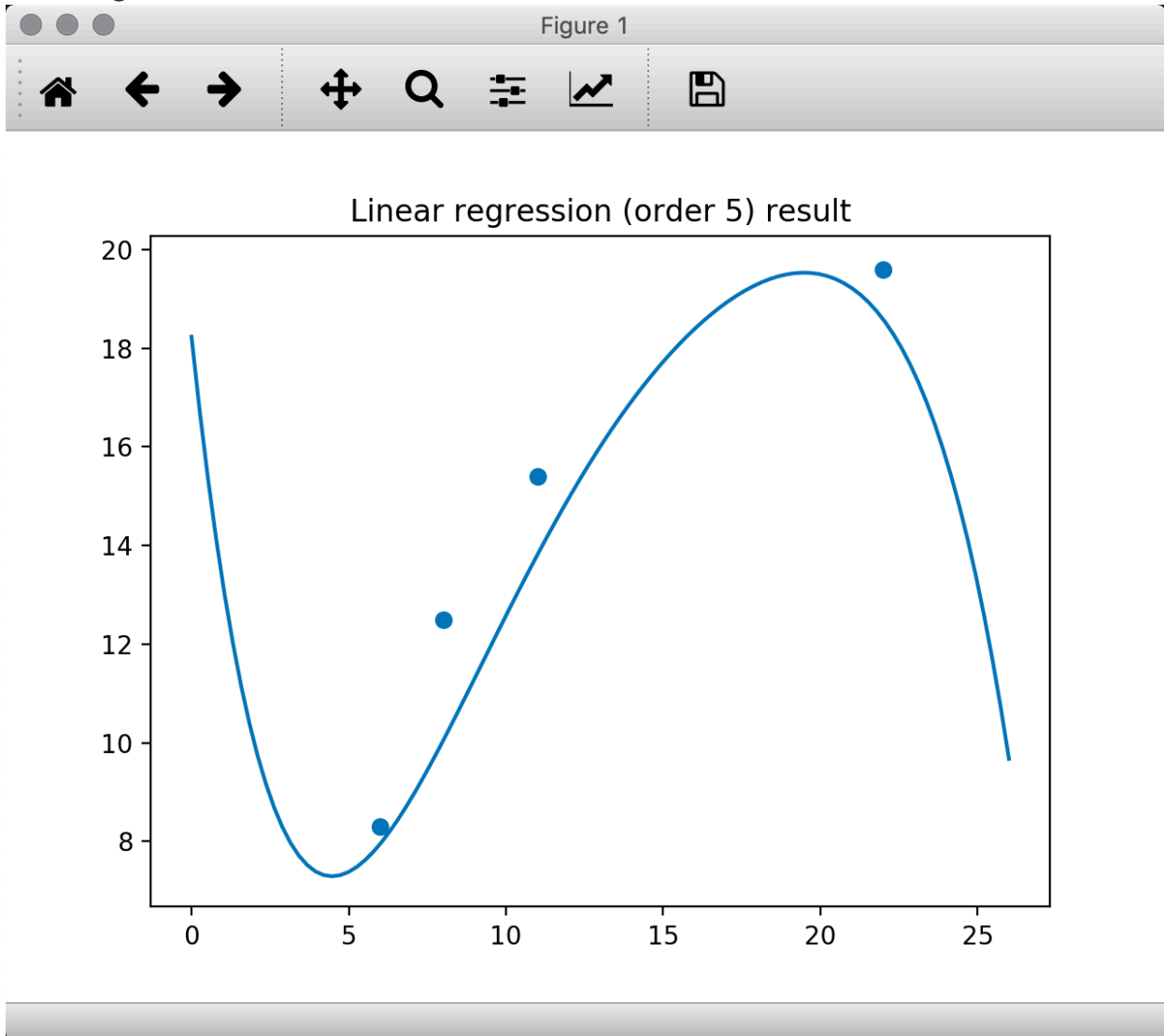
LI Yunxiang

1155092144

2.1.4

2.2

- Parameter theta calculated by normal equation:

    [[6.17672579e-17] [-1.45788344e-01] [1.12758691e-01] [6.46914750e-01] [-3.81560961e-03]]

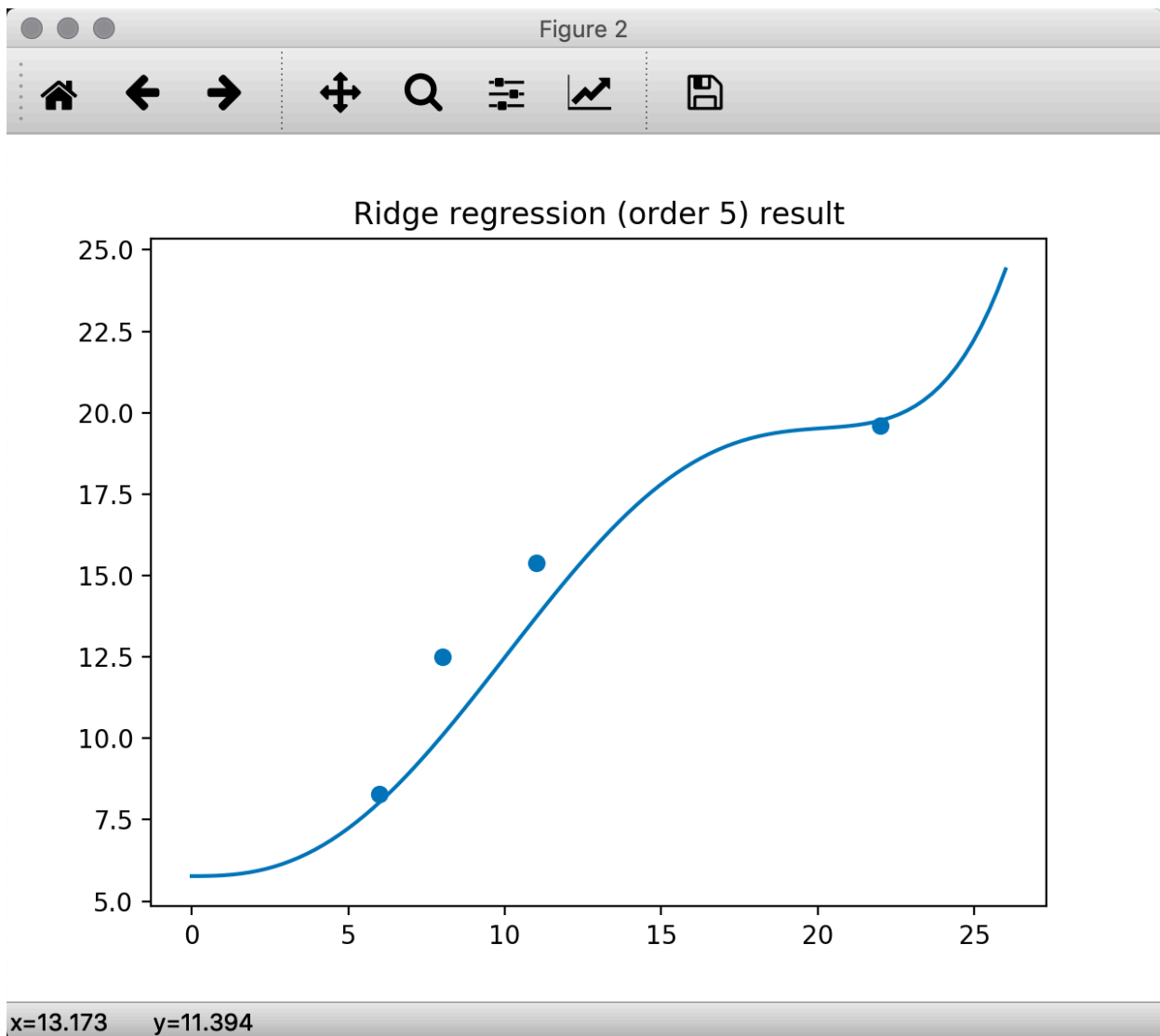- Parameter theta calculated by SGD: ([-0.0001467] [-0.14517694 0.11251498 0.64656019 -0.00401849])

2.3.1

- y1 = 18.2339283 + (-6.05120366)x + 1.11188554x*x + (-0.07.98839339e-02)x*x*x + (2.73075958e-03)x*x*x*x + (-3.75981662e-05)x*x*x*x*x
- Linear regression (order 5) score is: 0.8567469917016008

2.3.2

- y1 = 5.76871657 + (-2.93846420e-03)x + (1.23508737e-02)x*x + (1.43333276e-02)x*x*x + (-1.10345880e-03)x*x*x*x + (2.21101288e-05)x*x*x*x*x
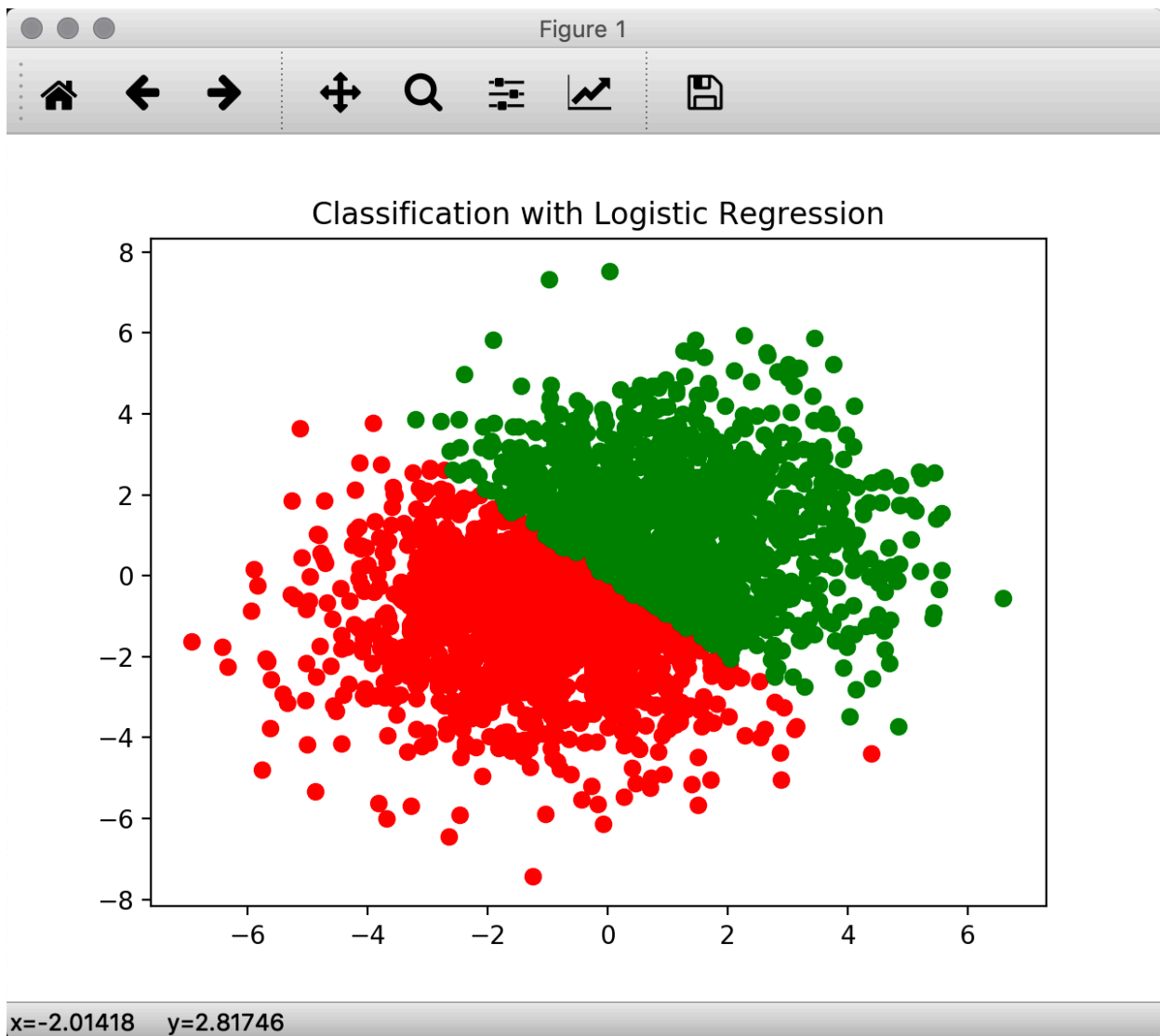- Ridge regression (order 5) score is: 0.8706689045766896

Ridge regression (order 5) result

x=13.173   y=11.394

2.3.3

- Q1. Ridge Regression has the highest score
- Q2. Yes

3.1

- Q: No

Classification with Logistic Regression

x=-2.01418   y=2.81746

3.2

- Q: Number of wrong predictions is: 550