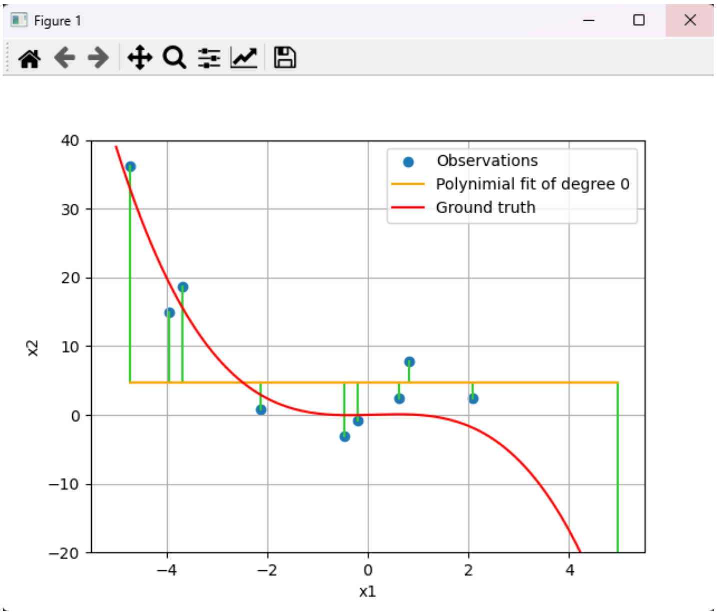
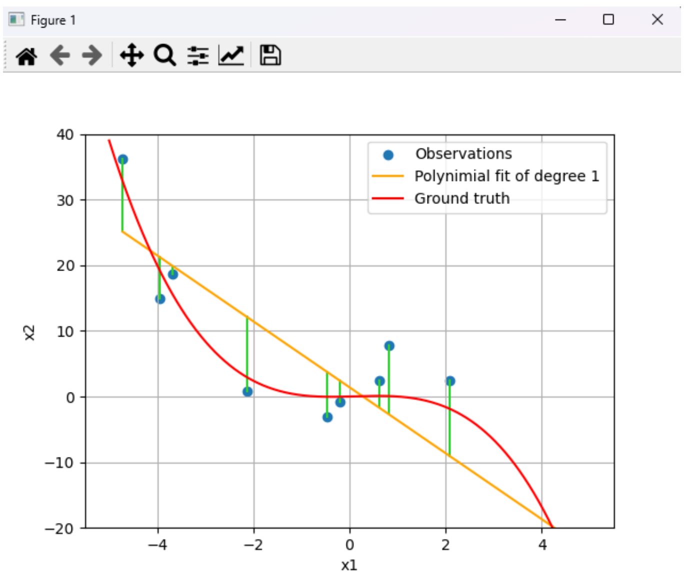
1. Experiment with different polynomial degrees. Randomly select 10 data points from the data set using 'np.random.seed(100)'. Fit polynomial curves of degrees 0, 1, 2, 3, 4, and 9. Plot the resulting curves and report the corresponding errors. Using 'x1' as the independent variable (x) and 'x2' as the dependent variable (y).

Deg=0



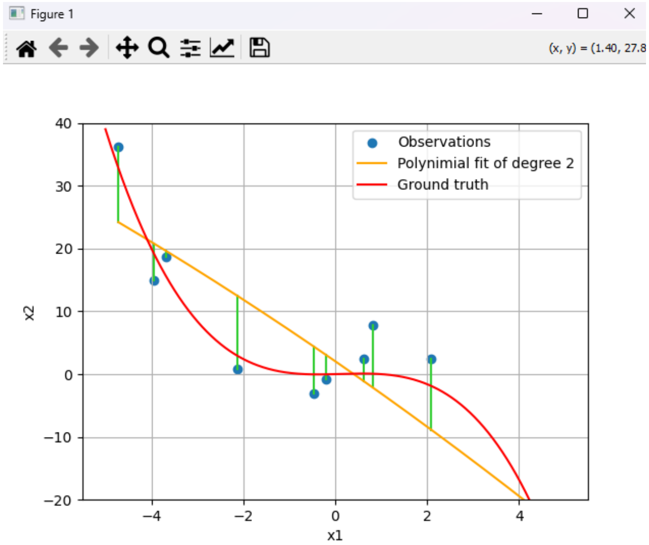
MSE = 273.876

Deg=1



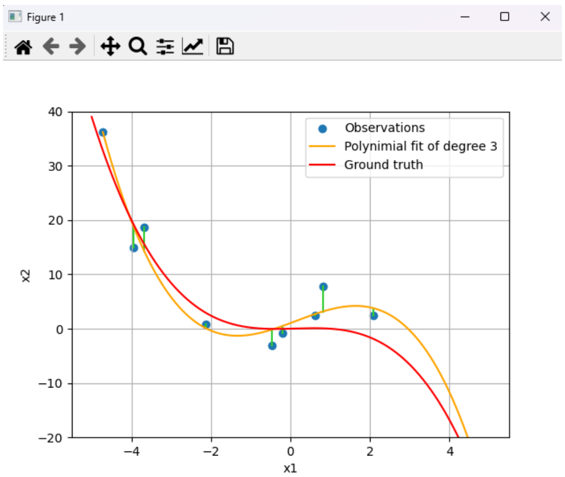
MSE = 67.753

Deg = 2



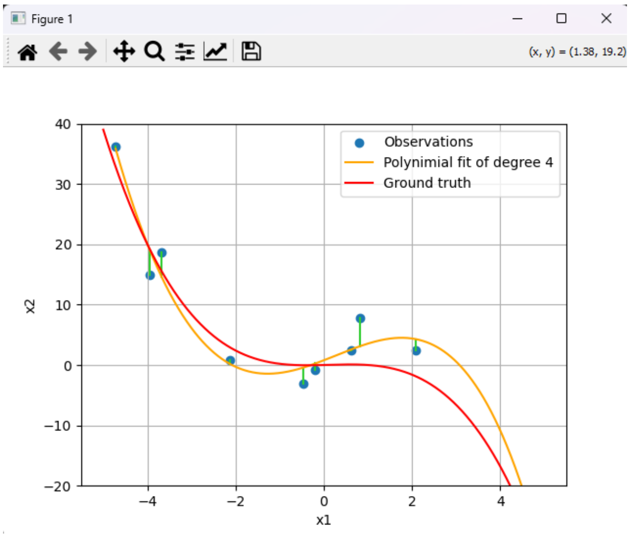
MSE = 67.27

Deg = 3



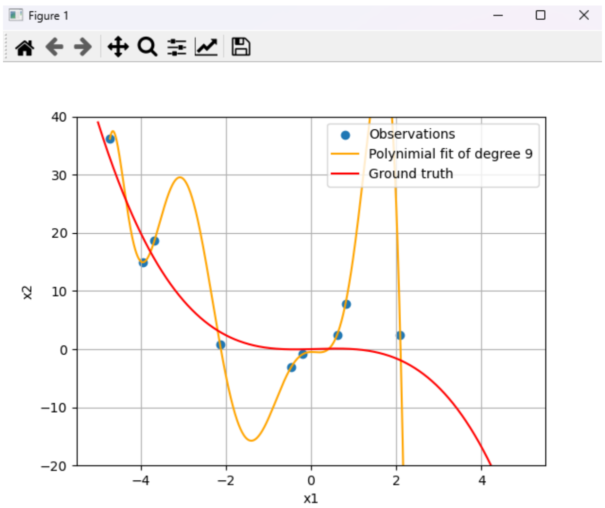
MSE = 6.793

Deg =4



MSE = 6.717

Deg =9

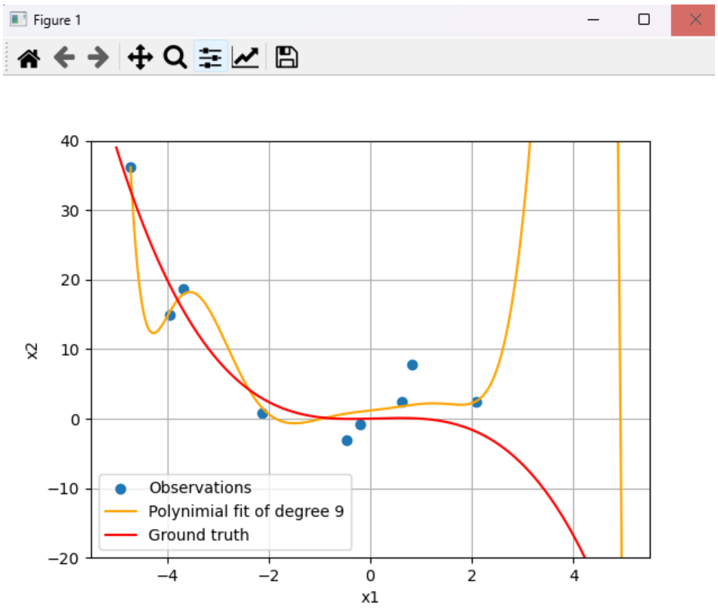


MSE = 0.0

In this exercise we see that the higher Polynomial degree that is used, the smaller the Mean Square Error(MSE) and it approaches 0, but with that high a degree we see that there is overfitting

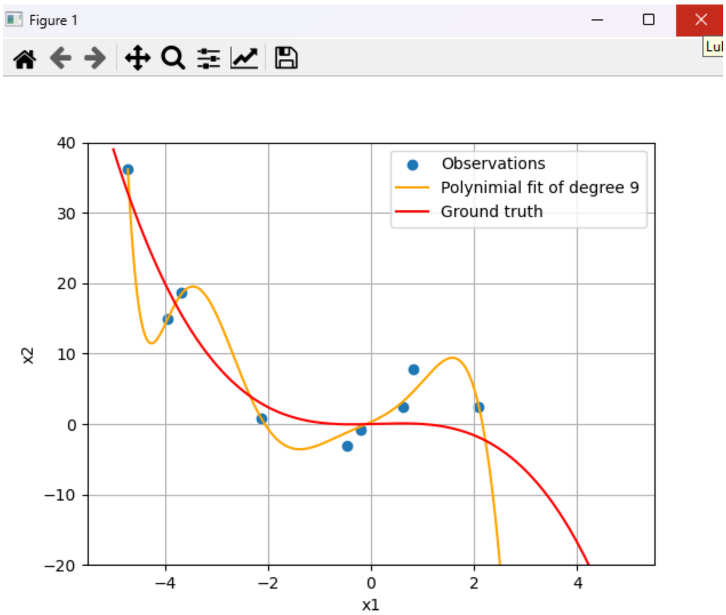
2. Explore L2 regularization for a polynomial of degree 9 by choosing regularization parameters of 10, 1, 0.1 and 0.01. Plot the curves and report the errors.

Reg =10



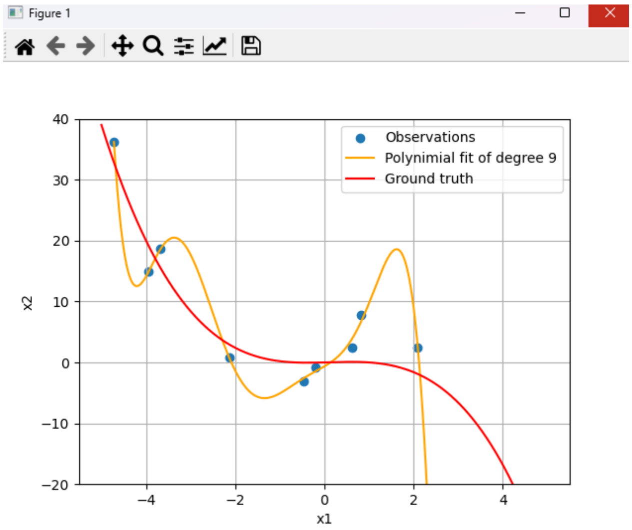
MSE = 5.534

Reg =1



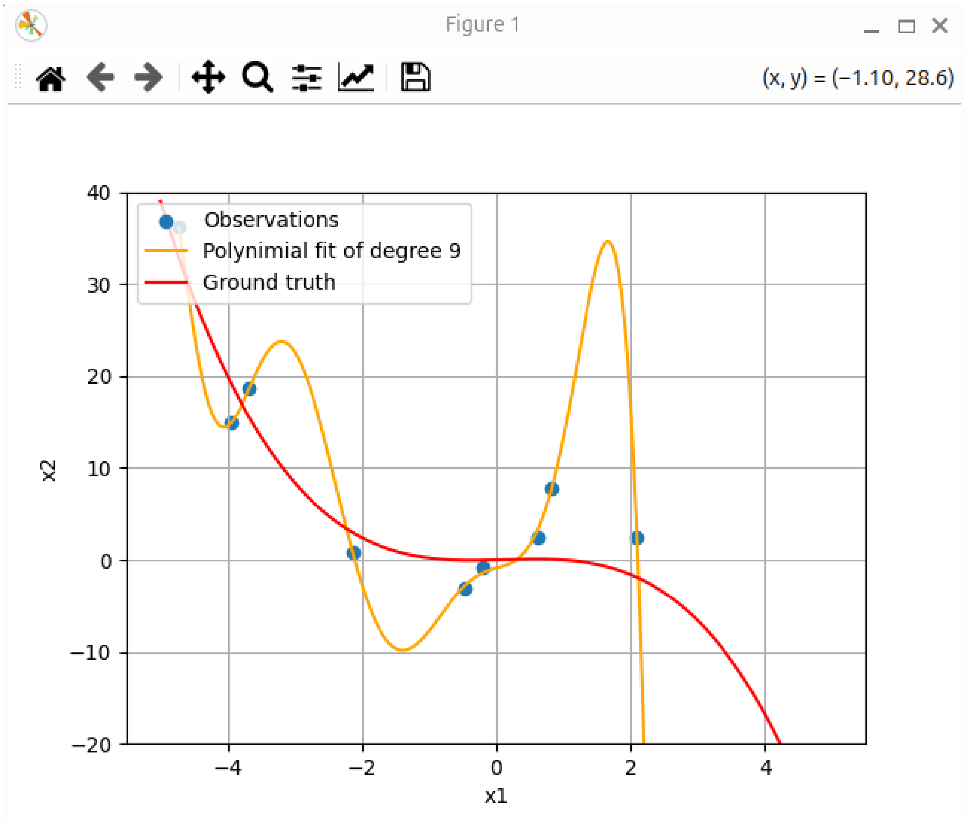
MSE = 1.565

Reg= 0.1



MSE = 0.426

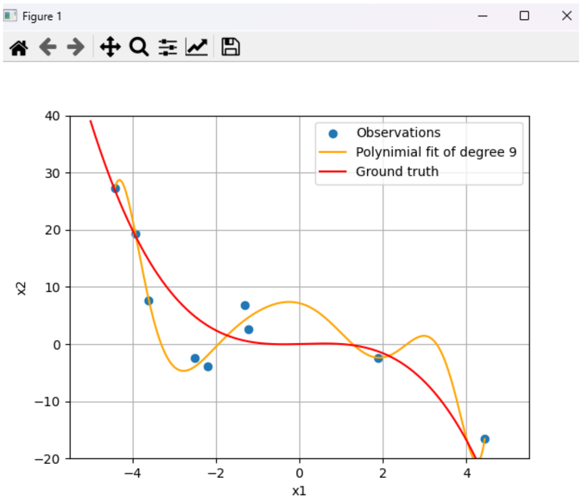
Reg = 0.01



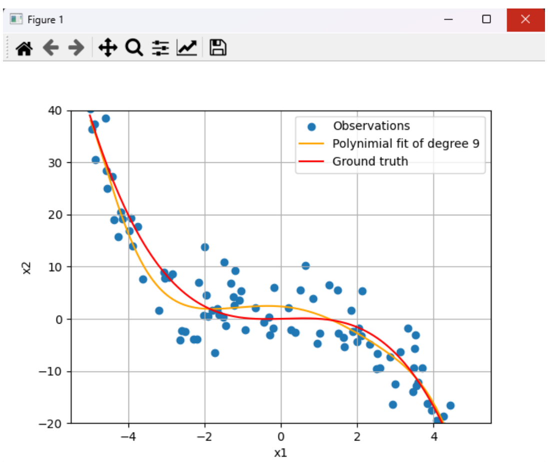
MSE = 0.109

We see that a higher regularization causes a higher MSE, and if it is too high there is a risk of underfitting, and with a lower regularization there is a lower MSE but there can still be too much overfitting with a low regularization

3. Investigate the effect of sample size. Now randomly select 100 data points from the data set using 'np.random.seed(100)'. Fit a polynomial of degree 9. Plot the curve and report the error.



MSE = 1.746



MSE = 18.91

With a higher sample count There is a larger MSE, but with a larger sample size we get a model much closer to the ground truth, and the higher the sample size the closer it will get to the ground truth