

Assignment 2

1. Generate 400 data points in R^2 from the uniform distribution $U [0,1]$. Construct the training set $T = \{ (x_{11}, x_{21}, y_1), (x_{21}, x_{22}, y_2), \dots, (x_{100\ 1}, x_{100\ 2}, y_{100}) \}$ using the relation

$$Y_i = \sin(2 \pi (x_{i1}^2 + x_{i2}^2)) + \epsilon_i \text{ where } \epsilon_i \sim N(0, 0.25).$$

In the similar way, construct a testing set of size 50

$$\text{i.e. Test} = \{ (x'_{11}, x'_{12}, y'_1), (x'_{21}, x'_{22}, y'_2), \dots, (x'_{50\ 1}, x'_{50\ 2}, y'_{50}) \}.$$

Estimate the regularized polynomial regression of order 6 with direct method and obtain the 3d plot on test set along with test data points. Find the NMSE, RMSE, MAE and R^2 .

2. Consider the dataset 1. You will find the only one independent variable (Income in thousand dollars) and one target variable (Card Balance in hundred dollars). Train the polynomial regression model with $M = 1, 2$ and 5 using the gradient descent method and obtain the plots of predictions upon training set and test. Compare the predictions obtained by gradient descent method and direct method with in terms of RMSEs.
3. Consider the motorcycle dataset. Estimate a regularized least square regression model (Also called Ridge Regression model) with Gaussian basis functions. Obtain the plot of estimated functions along with data points. Also obtain the RMSE, MAE, NMSE and R^2 for evaluating the quality of fit.