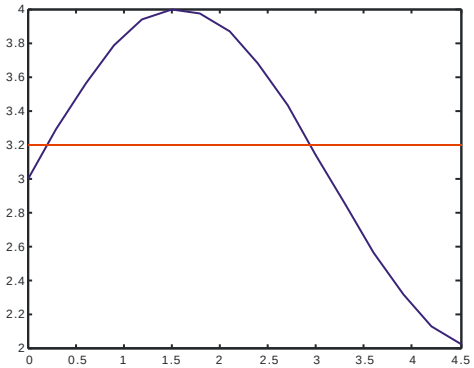
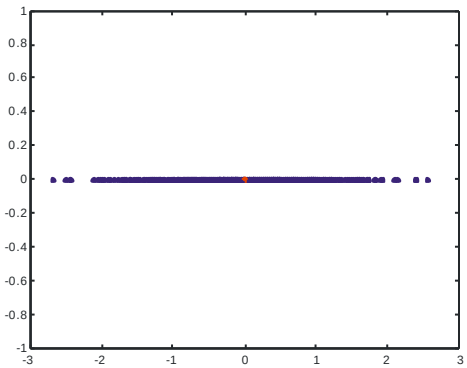
Exercise 4 (regularization - python code):





Consider the data set given in the attachment file (the code for reading from python is also given). Specifically, it consists of 10 data pairs of the form (y_i,x_i), i=1,…,10. All y_i's are accumulated in the vector y while all x_i's are accumulated in the vector x.

The aim is to unravel the relation between x_i's and y_i's.

(a) Plot the data.

(b) Fit a 8^{th} degree polynomial on the data using the LS estimator and plot the results (data points and the curve resulting from the fit). Output also the estimates of the parameters of the polynomial.

(c) Fit a 8^{th} degree polynomial on the data using the ridge regression estimator and plot the results (data points and the curve resulting from the fit). Output also the estimates of the parameters of the polynomial. Experiment with various values of λ.

(d) Fit a 8^{th} degree polynomial on the data using the lasso estimator and plot the results (data points and the curve resulting from the fit). Output also the estimates of the parameters of the polynomial. Experiment with various values of λ.

(e) Discuss briefly on the results.

Hint for ***(b), (c)***: The X matrix that needs to be constructed will contain 10 rows, one for each x_i. Each row will have the form [1, x_i, x_i^2, x_i^3, x_i^4, x_i^5, x_i^6, x_i^7, x_i^8].

Exercise 5 (python code + text):

Consider the set-up of exercise 5 of Homework 3. Consider also the ridge regression estimators resulting from eq. (A) in the exercise 1 above, for $\lambda = 0, 0.1, 0.2, \dots, 10000$. For each one of these values of λ, apply the steps (a), (b), (c1) of exercise 5 of Homework3, in order to compute the MSE. Plot MSE versus λ and

(i) determine the range of values of λ where the MSE is smaller than that of the unbiased LS estimator,

(ii) Comment on the results.