

Assignment #1

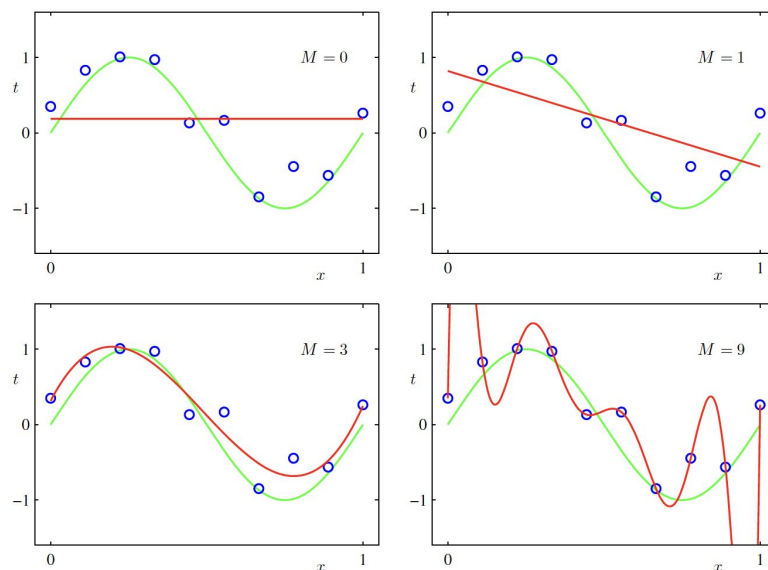
The goal of this assignment is to learn about the concept of overfitting using the Polynomial regression.

- You will post a complete assignment using Jupyter notebook in your homepage.
- You can use scikit-learn to get weights
- Below is the process
 - Generate 20 data pairs (X, Y) using $y = \sin(2\pi X) + N$
 - Use uniform distribution between 0 and 1 for X
 - Sample N from the normal gaussian distribution
 - Use 10 for train and 10 for test
 - Using root mean square error, find weights of polynomial regression for order is 0, 1, 3, 9
 - Display weights in table

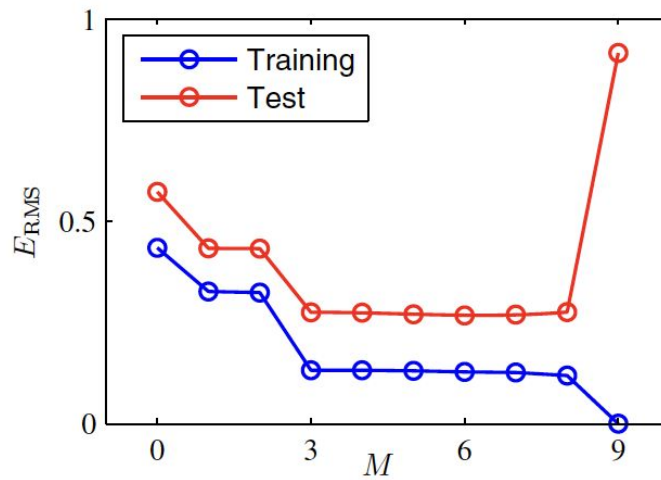
Table of the coefficients w^* for polynomials of various order. Observe how the typical magnitude of the coefficients increases dramatically as the order of the polynomial increases.

	$M = 0$	$M = 1$	$M = 6$	$M = 9$
w_0^*	0.19	0.82	0.31	0.35
w_1^*		-1.27	7.99	232.37
w_2^*			-25.43	-5321.83
w_3^*			17.37	48568.31
w_4^*				-231639.30
w_5^*				640042.26
w_6^*				-1061800.52
w_7^*				1042400.18
w_8^*				-557682.99
w_9^*				125201.43

-
- Draw a chart of fit data



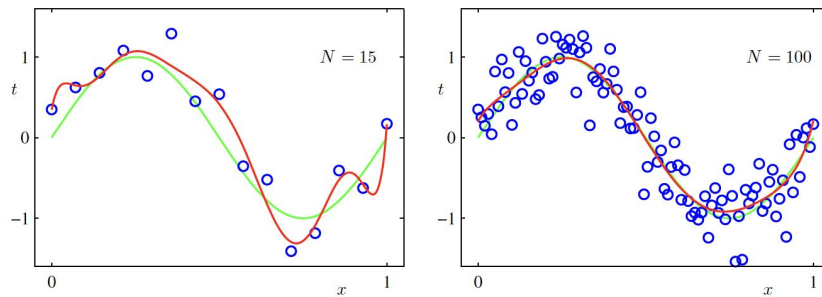
-
- Draw train error vs test error



■

■ To get this chart, you need to use all order from 0 to 9

f. Now generate 100 more data and fit 9th order model and draw fit



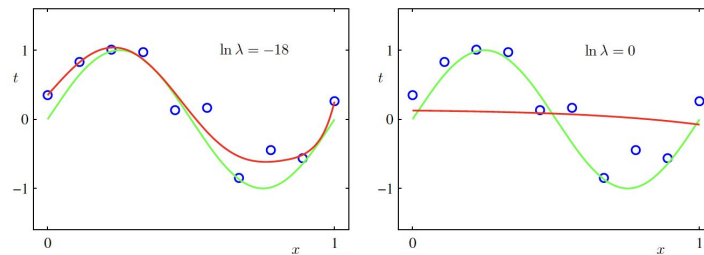
■

g. Now we will regularize using the sum of weights.

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

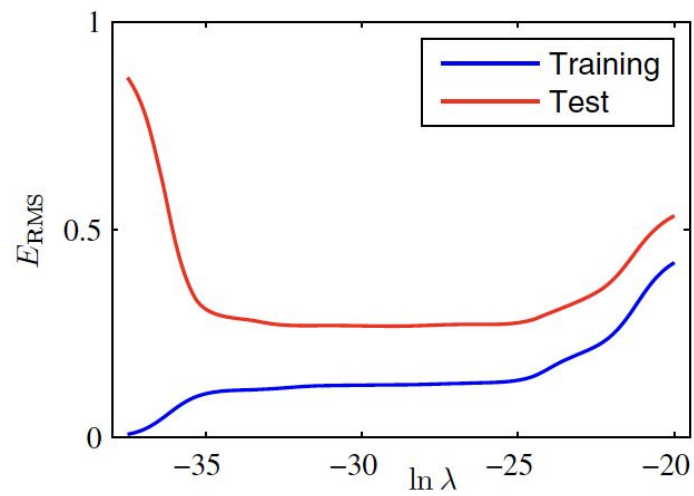
■

h. Draw chart for lambda is 1, 1/10, 1/100, 1/1000, 1/10000, 1/100000



■

i. Now draw test and train error according to lambda



■

- j. Based on the best test performance, what is your model?
- k. Submit your jupyter notebook and name the file in the following format-
 - l. **lastName_nn.ipynb**
 - Where, lastName = your last name, nn = 2 digit assignment no, starting from 01.