Homework 2: Regression
class in "Machine Learning", Fall 2016/17

Marco Favorito
Master of Science in Engineering in Computer Science
Department of Computer, Control, and Management Engineering
University of Rome "La Sapienza'
favorito.1609890@studenti.uniroma1.it

25 October 2016

1 Regression

1.1 Iteration from 1 to 10 degree polynomials

In the following I will show plots representing regression fitting on a given dataset using different with different degree.

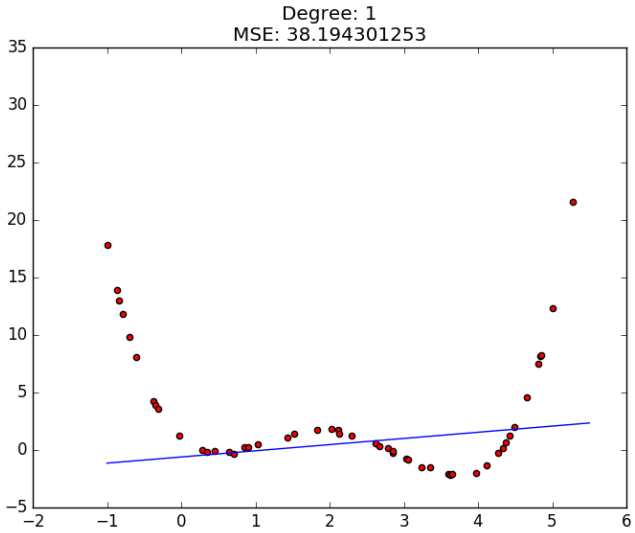

Figure 1: Regression model after mapping with degree 1

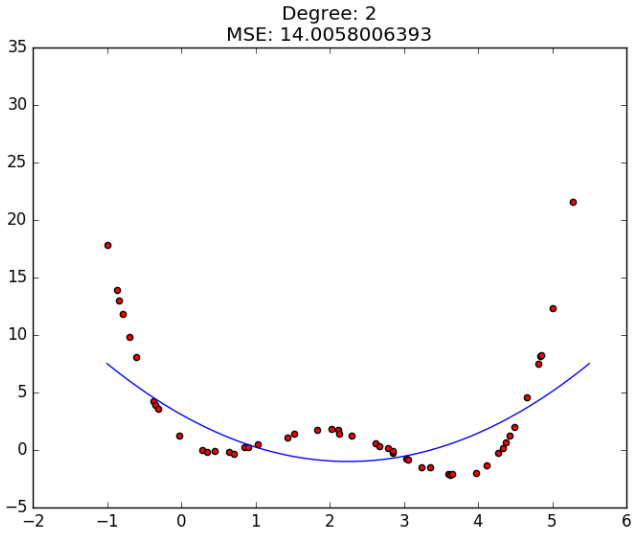

Figure 2: Regression model after mapping with degree 2

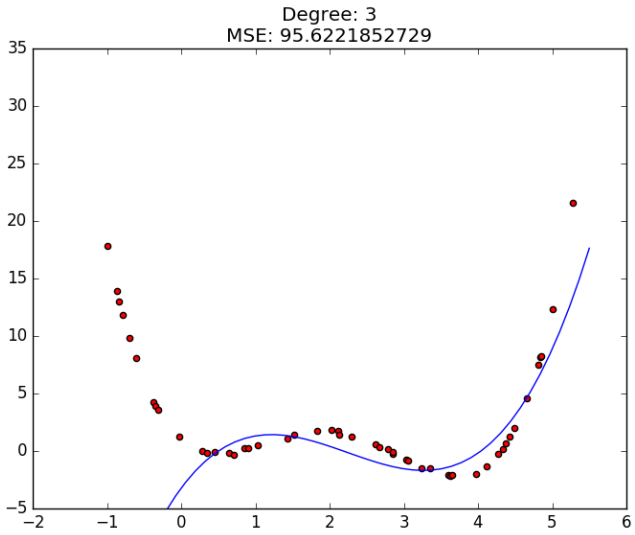

Figure 3: Regression model after mapping with degree 3

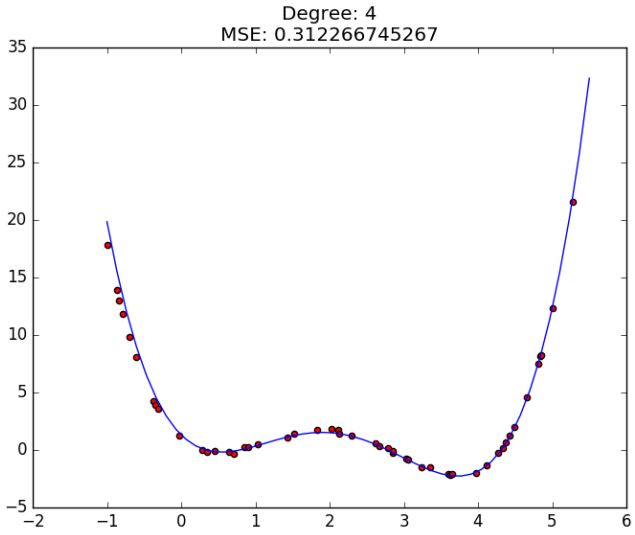

Figure 4: Regression model after mapping with degree 4

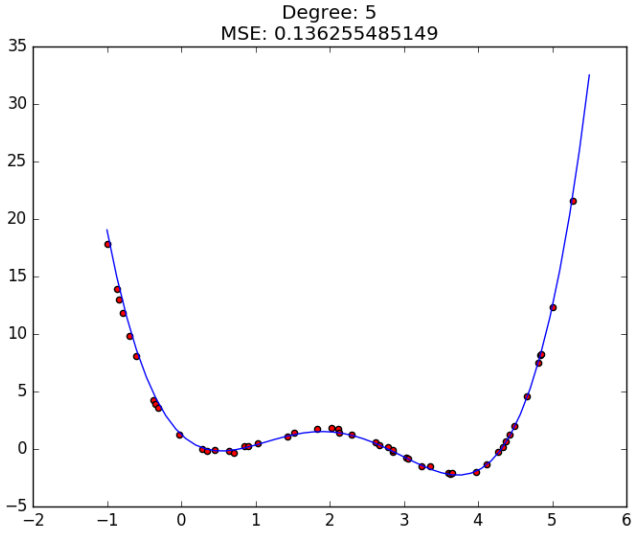

Figure 5: Regression model after mapping with degree 5

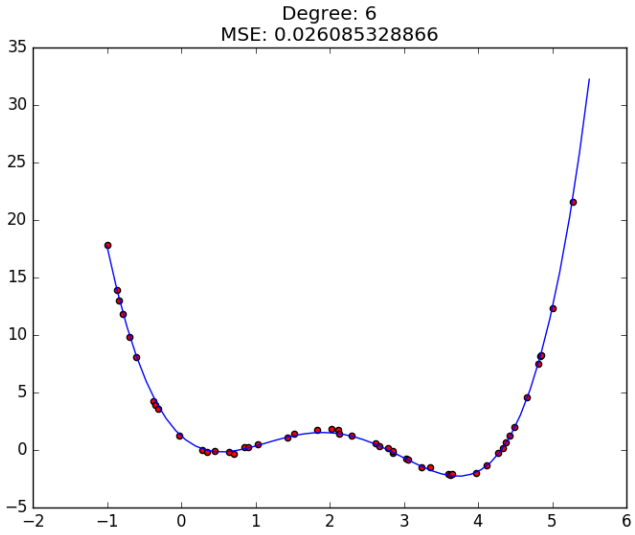

Figure 6: Regression model after mapping with degree 6

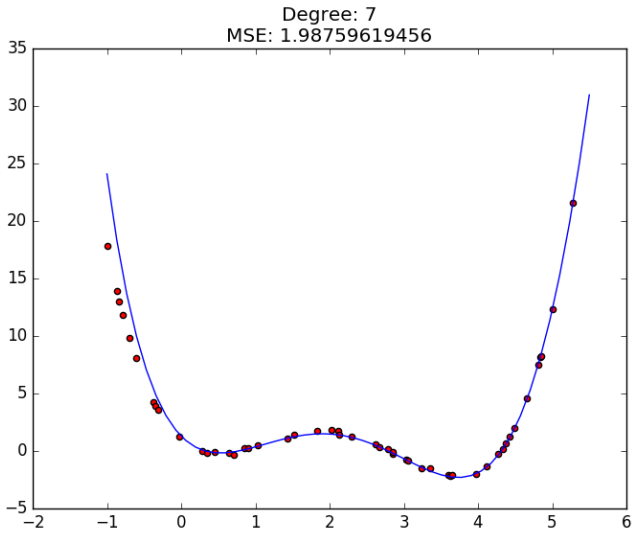

Figure 7: Regression model after mapping with degree 7

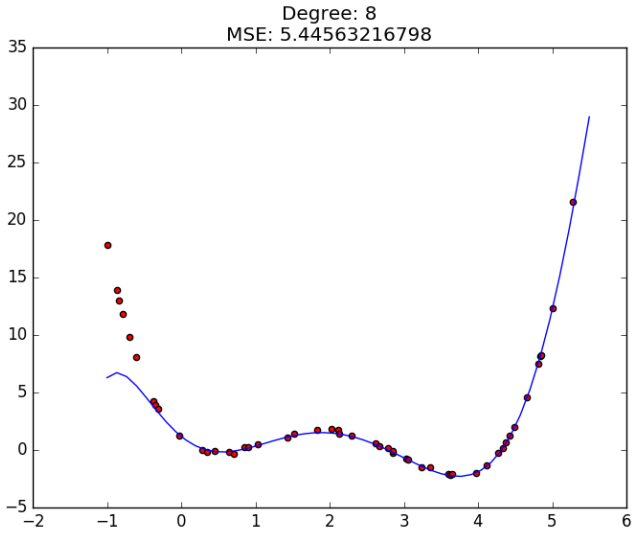

Figure 8: Regression model after mapping with degree 8

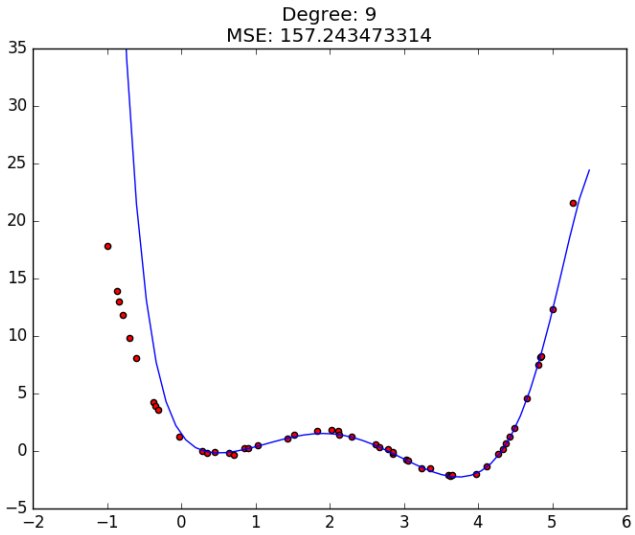

Figure 9: Regression model after mapping with degree 9

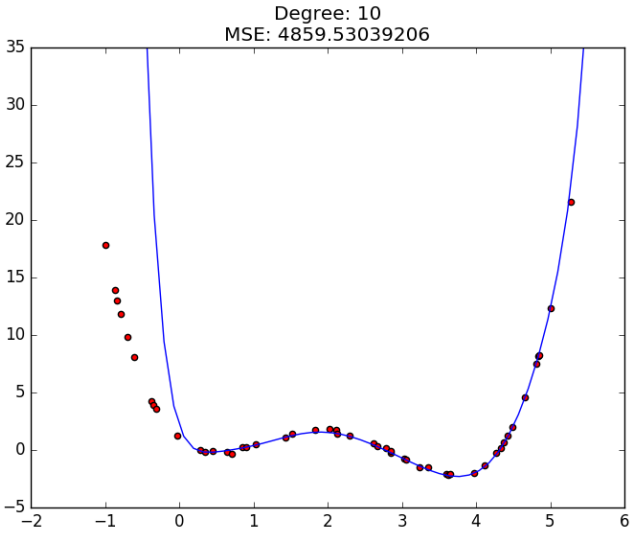

Figure 10: Regression model after mapping with degree 10

Table 1: Summary table on degree of the model and MSE over testing data

Degree	MSE
1	38.194301253
2	14.0058006393
3	95.6221852729
4	0.312266745267
5	0.136255485149
6	0.026085328866
7	1.98759619456
8	5.44563216798
9	157.243473314
10	4859.53039206

And this is the summary table of degree and relative MSE: Its plot is: It is easy to see that the best fit is when degree is equal to 6, since the mean square error (MSE) is the lowest.
High degree polynomials have a high error because they *overfit* the model, i.e. give too much importance to noisy data. In other words, remainding that, in case of n-degree polynomial regression, we have:

$$y = a_0 + a_1 x + \cdots + a_n x^n + \epsilon$$

It may happen that a too much high degree of the polynomial don't allow to *minimize* effectively the Least Square Error when try to estimate $\mathbf{a}$ over data $\mathbf{x}$, due to high elements in the polynomial.

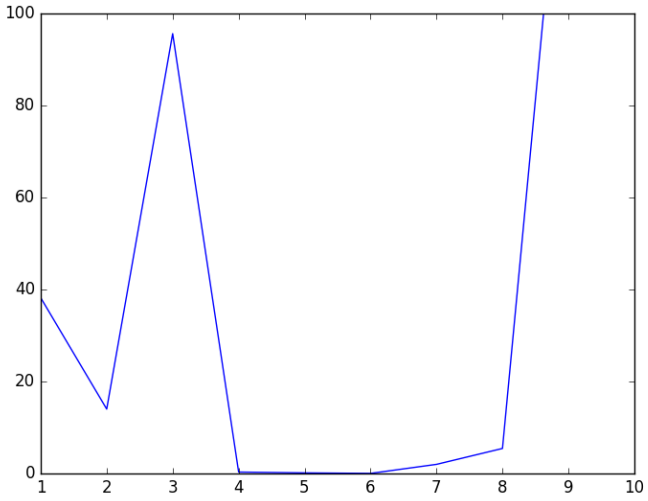

Figure 11: Plot of MSE in function of degree