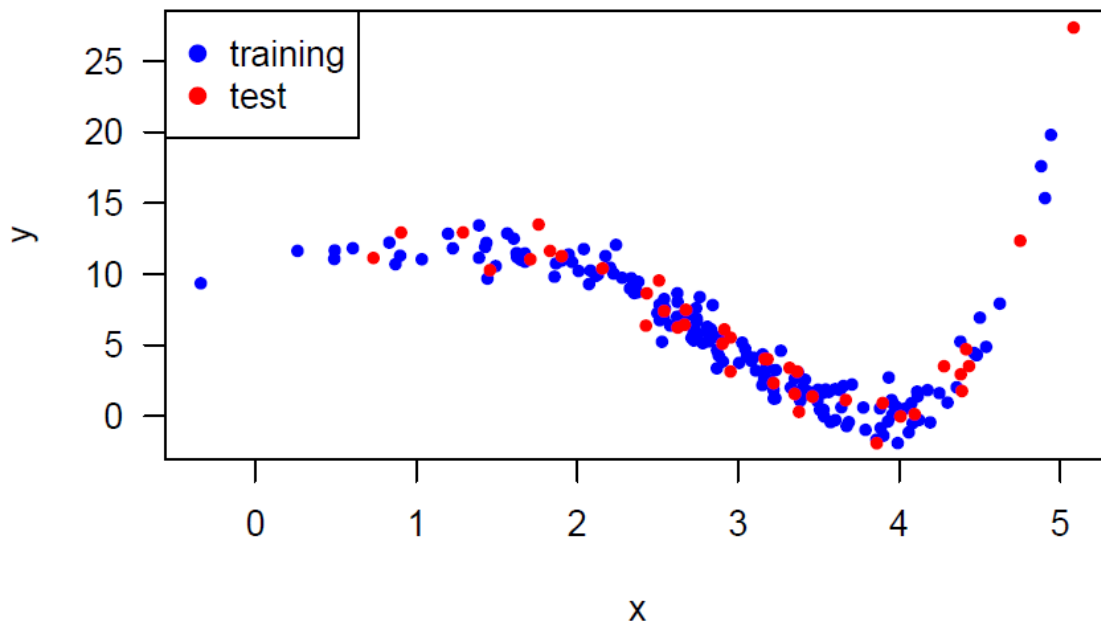


COMP421X HW01: Multivariate Regression

Deadline: March 8, 2019, 11:59 PM

In this homework, you will implement a multivariate parametric regression algorithm in R, Matlab, or Python. Here are the steps you need to follow:

1. Read Chapter 5 from the textbook.
2. You are given a univariate regression data set, which contains 200 data points, in the file named hw01_data_set.csv. Divide the data set into two parts by assigning the first 160 data points to the training set and the remaining 40 data points to the test set.
3. Draw training data points and test data points on the same figure. Your figure should be similar to the following figure.



4. The model that we want to learn is $y = w_0 + w_1 x + w_2 x^3 + w_3 \exp(x) + w_4 \sin(x)$. Learn the parameters of the linear regression algorithm w_0, w_1, w_2, w_3 , and w_4 . Your parameter estimation should be similar to following output.

```
##          [,1]
## [1,]  9.3342591
## [2,]  0.6412598
## [3,] -0.8257511
## [4,]  0.7650880
## [5,]  0.9005130
```

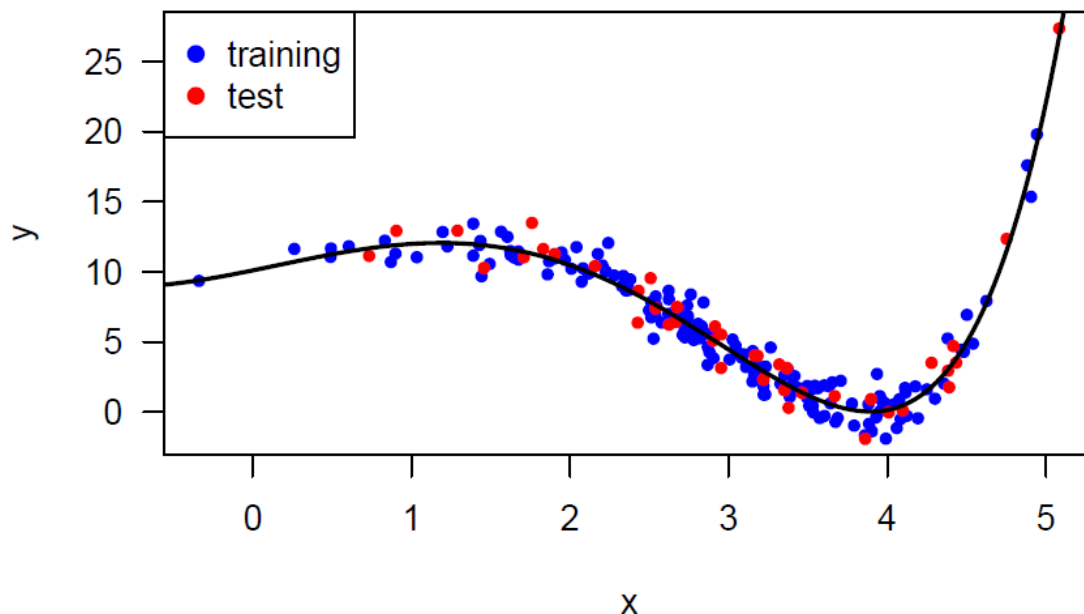
5. Test your algorithm with the learned parameters on the test data set. Calculate the root mean square error for test data points. The formula for RMSE can be written as:

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^{N_{test}} (y_i - \hat{y}_i)^2}{N_{test}}\right)}.$$

Your result should be similar to the following output.

```
## [1] "RMSE is 1.037429"
```

6. Run your algorithm on the data interval $[-1, 6]$ and draw training data points, test data points, and fitted curve for this interval on the same figure. Your figure should be similar to the following figure.



What to submit: You need to submit your source code in a single file (.R file if you are using R, .m file if you are using Matlab, or .py file if you are using Python) and a short report explaining your approach including your results but not any part of your code (.doc, .docx, or .pdf file). You will put these two files in a single zip file named as **STUDENTID.zip**, where **STUDENTID** should be replaced with your 7-digit student number.

How to submit: Upload the zip file you created to the corresponding Blackboard Assignment. Please follow the exact style mentioned and upload a zip file named as **STUDENTID.zip (Your student ID.zip)**. Submissions that do not follow these guidelines will not be graded.

If you have any questions please email cak14@ku.edu.tr with the subject line ***Intro2MachineLearningHW01***. You could also use Blackboard forum for your questions.

Late submission policy: Late submissions will not be graded.

Cheating policy: Very similar submissions will not be graded.