

12-752: Data-Driven Energy Management of Buildings

Assignment #3

Mario Bergés
Assistant Professor
Department of Civil and Environmental Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
`marioberges@cmu.edu`

December 3, 2015

Some notes before you begin:

- Make sure you document everything you do, and not just write down the answer to the question. This will both help during grading as well as improving your learning process.
- Do not write down any solution or process that you do not understand. If you feel that you do not understand how to do something, seek some help: e-mail the TAs or the instructor
- To submit your assignment, please do so using Blackboard. Two files should be uploaded via Blackboard: (a) the IPython Notebook (i.e. a .ipynb file) documenting all the tasks found in the assignment and all of your answers (including the output of your code); and (b) a PDF copy of this notebook
- Please upload a single compressed ZIP file containing the above, and name it as follows: *andrewID_assignment-#.zip* (where *andrewID* is your AndrewID and *#* is the assignment number)

1 Preface

This assignment is a continuation of the tasks that you completed in Assignment #2. For this reason, you can continue working on the same Notebook and submit the whole of Assignment #2 plus the tasks in here when you submit. To make things easier, I have not re-started the task number count here (i.e., we will continue the task numbers where we left off in last assignment).

1.1 Evaluating Linear Regression Models

Task #13 [20%]: Calculate the coefficient of determinatino (R^2) value for your model on the test set and comment on what the results imply

You may want to read up on what this value means, but here's the formula (\bar{y} is a scalar containing the average of all values in Y):

$$R^2 = 1 - \frac{(Y - \mathbf{X}\hat{\beta})^T(Y - \mathbf{X}\hat{\beta})}{(Y - \bar{y})^T(Y - \bar{y})} \quad (1)$$

Task #14 [40%]: Compute the 95% confidence intervals for each of the coefficients in $\hat{\beta}$. Comment on what the results imply

Task #15 [40%]: Assume that we have a null hypothesis $H_0 : \hat{\beta}_k = 0$ for each one of the coefficients in $\hat{\beta}$. Can you reject the null hypothesis for some of these at any level of significance (i.e., at any α)?