CET 513 Transportation Networks and Optimization

Homework #4

Out: 11/12/2019   **Due: 11/21/2019**

Problem 1

Use the KKT conditions to solve the following problem:

$$\min z(x_1, x_2) = 4(x_1 - 2)^2 + 3(x_2 - 4)^2$$

subject to

$$x_1 + x_2 \geq 5$$
$$x_1 \geq 1$$
$$x_2 \geq 2$$

Plot the feasible region. Is the solution unique?

Problem 2

Use the Frank-Wolf Method to solve the following program:

$$\min f(x_1, x_2) = 4(x_1 - 10)^2 + (x_2 - 4)^2$$

subject to

$$x_1 - x_2 \leq 10$$
$$\frac{1}{3}x_1 - x_2 \geq 3$$
$$x_1 \geq 0$$

Plot the feasible region and show the intermediate solution ($x(k)$ for the kth interation) at each iteration graphically.

Note: 1) You only need to present a few steps of the procedure; 2) When solving the linearized program for finding a feasible descent direction, you may use Python or Excel solver or solve it by inspection.

NOTE: You only need to complete Problem 3 or Problem 4 below. There will be NO bonus point even if you finish both.

Problem 3:

Data from Amazon shows the monthly sales of a product for 5 years (60 months) together with the Price ($), Ad Cost (in million $), and (total) Discounts offered (in million $) for each month. Build a linear regression model to explore the relationship between these key factors (price, ad cost, discounts) and the monthly sales. You may try both the OLS and RLM methods in Python. Which method works better and why? You should also show the performances of your methods using the testing dataset.

The data are in the file: "Sales.xlsx" with two sheets: "Training" for the training dataset and "Testing" for the testing dataset.

Problem 4

There are 3,700 travel analysis zones (TAZ) in the Puget Sound region; see the illustration on the next page. TAZs are geographic areas defined by Metropolitan Planning Organizations (MPO) to conduct transportation planning and other related analysis. The figure below shows the TAZs in King County in the Puget Sound region. Puget Sound Regional Council (PSRC), the MPO for the Puget Sound region, has conducted surveys, data collection, and modeling research to come up with the total number of trips generated from each TAZ in the region.

A group of researchers were interested in understanding whether "big data" may be used to infer the same information (i.e., total number of trips generated from each TAZ). They collected data from mobile devices and conducted analysis to derive such information.

The PSRC trips and the big data based estimates are shown in the Excel file ("Trip_originated-rev.csv"). PSRC trips are in the column "PSRC_trips" and big data generated trips are in the column "App_data_estimated_trips" respectively. Noticed that the PSRC trips are perturbed by adding slight random noise.

To test whether the two results are consistent, you will need to conduct a correlation analysis. The first step is to check if the following linear relationship exists between the two results.

$$y = ax + b$$

Here x, y are the PSRC trips and big data estimates respectively, and a and b are the coefficients that need to be estimated. You will need to perform the following tasks:

(a) Formulate an optimization problem to solve for a and b.
(b) Solve it using the steepest decent method in Python or Matlab. Plot the objective value and step size vs. the iteration #.
(c) Solve the same problem using the ordinary least square (OLS) method and the Robust LS method (RLM) in Pandas in Python. Did you get the same or similar solutions with that from (b)? Discuss the solutions you obtained.