Homework 3 for
MSAN 502, Review of Linear Algebra

Due: Wednesday, August 3, 11:59pm

Strang Problems (4th edition)

Problem Set 3.5: Do problems 2, 11, 20, 23, 41
Problem Set 3.6: Do problems 1, 4, 14, 24, 25,26
Problem Set 4.1: Do problems 3, 4, 7, 11, 21
Problem Set 4.2: Do problems 1a, 11b, 13, 21
Problem Set 4.3: Do problems 1, 12, 16,26

Python Problem

Write code in Python to find the best fit line in the sense of least squares to a set of data consisting of paired observations in the form $(x; y)$. Here's how your code should work:

1. Read a txt file into a pandas data frame and call it fr. (You can do this manually, or just for one of the 3 les in your code.)

2. Each file has a column named x and a column named y. Column x (resp. y) consists of the x values (resp y values) of the data points (in the same order, of course).

3. Output the slope and intercept of the best-fit line, as well as the sum of the squares of the distances from the observed y-values to the y-values on your line.

4. You should do this directly using linear algebra (i.e., the normal equations), and NOT using built-in regression functions like statsmodels.ols.

There are three data sets on the course webpage in canvas on which you can test your program. The data sets are TVlife.txt, population.txt, and nba.txt. Each file has several columns of data, but the ones we are interested (called 'x' and 'y') in each are:

(a) Life Expectancy (y) as a function of Televisions per Thousand People (x) (TVlife.txt)

(b) National Population (y) as a function of Year (x) (population.txt)

(c) Team Winning Percentage (WinPct) (y) as a function of PM (x) (the average point difference over all that team's games) (nba.txt)

It's also worth taking a moment to plot your data and make sure that a linear relationship between x and y is a reasonable assumption. If you want to check your program, run statsmodels.ols and compare with your answer.

Finally, find one more data set that is of interest to *you* and include it in your analysis report and upload it in your submission.