CS325: Group Assignment 4 – Linear programming

Due: Wednesday, March 9th at 11:59PM

You are required to work in groups of two or three students. See the provided group guidelines for how to form your groups on canvas. Each group should only submit one copy of the work to canvas, including the **project report as a pdf** and the code that implements the algorithms. The report should have **all member's names included**. You may use Java, Python, C/C++, Matlab, or R. The submission to Canvas can include multiple files. Acceptable file types include pdf, zip, gzip. So you should zip up your code for submission. It is also required that the report is submitted as a separate pdf file so that the TA can directly comment on it via canvas.

For this project, you will model a problem as a linear program and solve them using a language and linear programming solver of your choice. For a (non-comprehensive) list of freely available LP solvers, see this wikipedia page: `http://en.wikipedia.org/wiki/Linear_programming` These problems were tested using Matlab and Matlab's linear programming solver, `linprog` (which you have access to through the College of Engineering) as well as GLPK version via the Python PuLP package (see `https://projects.coin-or.org/PuLP`). The latter was faster and less cumbersome to use.

Local temperature change

The daily average temperature at a given location can be modelled by a linear function plus two sinusoidal functions; the first sinusoidal function has a period of one year (modelling the rise and fall offind temperature with the seasons) and the second sinusoidal function has a period of 10.7 years (modelling the solar cycle). That is, a decent model of average temperature T on day d is given by:

$$T(d) = \underbrace{x_0 + x_1 \cdot d}_{\text{linear trend}} + \underbrace{x_2 \cdot \cos\left(\frac{2\pi d}{365.25}\right) + x_3 \cdot \sin\left(\frac{2\pi d}{365.25}\right)}_{\text{seasonal pattern}} + \underbrace{x_4 \cdot \cos\left(\frac{2\pi d}{365.25 \times 10.7}\right) + x_5 \cdot \sin\left(\frac{2\pi d}{365.25 \times 10.7}\right)}_{\text{solar cycle}}$$

The values of $x_0, x_1, \ldots, x_5$ depend on the location. For example, the amplitude of the seasonal change (x_2 and x_3) would be much greater in Chicago, IL than in San Diego, CA. Given daily temperature recordings (pairs (d_i, T_i): the average temperature T_i on day d_i), your goal is to find values for $x_0, x_1, \ldots, x_5$ that result in an equation $T(d)$ that best fits the data. You will use what you learned in class and in the practice assignment to find the curve $T(d)$ of best fit that minimizes the maximum absolute deviation for a given set of daily average temperatures.

Data from Corvallis can be found on canvas `Corvallis_data.csv`. The first four columns are the raw data downloaded from NOAA. Raw minimum and maximum temperature recordings are given in *tenths* of degrees Celcius. Average daily temperature (column *average*) is in degrees Celcius and is simply the average of the maximum and minimum temperatures on a given day. The number of days since May 1, 1952 is given in the last column (*day*). Note that you need to take the day number into account because several days (and a few entire months) were missing from the data set. These last two columns give you the (d_i, T_i) data pairs.

Your best fit curve (defined by the linear program you use to find the values of $x_0, x_1, \ldots, x_5$ that minimize the maximum absolute deviation of $T(d)$ from your data points), gives you a value x_1 that describes the linear drift of the temperature as degrees per day.

Your report must include:

1. A description for a linear program for finding the best fit curve for temperature data.
2. The values of all of the variables to your linear program in the optimal solution that your linear program solver finds for the Corvallis data. Solving this LP may take a while depending on your computer. Be patient. You may want to do testing on a small part of the data set. Include the output of the LP solver that you use (showing that an optimal solution was found).
3. A single plot that contains:
   - the raw data plotted as points,
   - your best fit curve, and
   - the linear part of the curve $x_0 + x_1 \cdot d$.
4. Based on the value x_1 how many degrees Celcius per century is Corvallis changing and is it a warming or cooling trend?