AMATH 301 - Spring 2018

Homework #5

Due on Thursday, May 3, 2018

1. The file `SeaPopData.mat`, which is included with the homework, contains the following population data for the city of Seattle.

Year	Population
1860	188
1870	1151
1880	3533
1890	42837
1900	80671
1910	237194
1920	315312
1930	365583
1940	368302
1950	467591
1960	557087
1970	530831
1980	493846
1990	516259
2000	563374
2010	608660

You should load this data into Matlab using the `load` command. Be sure that the file `SeaPopData.mat` has been downloaded into the same directory as your script file. **You do NOT need to upload this file to Scorelator. Scorelator has its own copy.** If the `load` command is successful, you will have two new vectors in your workspace, `t` and `Seattle_Pop`. The values of the vector `t` are the number of years since 1860. Therefore, $t = 0$ is 1860 and $t = 150$ is 2010. The vector `Seattle_Pop` has the corresponding populations from the table above.

Data from:

1. `https://en.wikipedia.org/wiki/Demographics_of_Seattle`
2. `https://www.seattle.gov/opcd/population-and-demographics`

As you work through this assignment, it is a good idea to plot the data points and the best fit curves or interpolants to make sure your code is working correctly. However, you should remove all plot commands before submitting to Scorelator.

(a) Find the line of best fit for the data. That is, find a line $P = mt + b$ where t is the number of years since 1860 and P is the population of Seattle. Save the slope of the line in `A1.dat`. Calculate the root-mean-square error and save it in `A2.dat`. The most recent population estimate for Seattle is that the 2017 population was 713,700. Use the equation of the best fit line to predict the population in 2017 (i.e. plug in $t = 157$). Save this value in `A3.dat`.

Things to think about: What is the meaning of the slope of the line of best fit? What does it tell us about how the population is changing?

(b) Find the best fit quadratic function for the data. Use this curve to predict the population in 2017, and save the prediction in `A4.dat`. Repeat this process for the best fit polynomials of degree 3 (cubic) and degree 9. Save the predictions in `A5.dat` and `A6.dat`, respectively.

(c) Figure out the degree of the polynomial interpolant of the data. Save the degree in `A7.dat`. Then find the polynomial interpolant, and use it to predict the population in 2017. Save the prediction in `A8.dat`.

Things to think about: How accurate were the predictions given by the different polynomial fits? Try calculating more best fit polynomials to see how well they predict the 2017 population. In particular, try using a degree 5 polynomial. It seems to do a good job. Would you trust it to predict the population in 2060? Now try adding the 2017 population to the data set (add a 157 at the end of the `t` vector and 713,700 at the end of the `Seattle_Pop` vector). How much does this change each of your best fit curves? Are some more resilient to new information than others? You can also try changing just one of the data points to the wrong number to see how it changes the best fit curves. Which will make the biggest mistake if given "bad" data?

(d) Fit an exponential function ($P = ae^{rt}$) to the data by taking a natural logarithm of the population values and using a linear fit. Make a 1×2 row vector with the values of the parameters with a being the first component and r being the second (i.e. `[a r]`). Save this vector in `A9.dat`. Then use this exponential function to predict the population in 2017, and store the prediction in `A10.dat`.

Things to think about: You can also do an exponential fit by using `fminsearch`, but it gives a very different answer. Why? Which method is better?