

Numerical Calculus - 2020/2021 - Practice 1 Applied Mathematics and Computation 121

Name and Lastname:

Group:

Hand in before noon on Tuesday the 27th of October. All the solutions to the exercises must appear in a single `.m` file. Answers to questions must appear displayed on screen when running your code. You have plenty of time, but do not leave it for the last minute. Also, you may divide the task among the members of the team, since it contains several subtasks (writing the code, checking that the code does what expected, explaining the results,...), but I recommend that you understand what others did. Name your file `practice01teamXX.m` where `XX` is the number of your team. On the first few lines of code write as comment the names and NIAs of all the members of the team.

These weeks we will work on understanding our nemesis: the spread of coronavirus cases. For the moment we will try to study the quality of the data that the governments produce, and how that differs on different days of the week for each country. Each team will work with 4 countries. All teams will work with Spain and 3 more countries, and the list of countries for each team is at the end of this statement. Make sure you know which team you are part of!!

You can find attached a spreadsheet called `datapr1` containing a table with the data of new cases and deaths for each day of some periods of the epidemic in some countries with high incidence. Save it in the same folder you will write your MATLAB script on and you may assume that I will have the data available as well on the same folder (no need to submit the spreadsheet!). The source of the data is clear on the spreadsheet.

You may find it useful to introduce *anonymous functions* although this is not obligatory. Check the description in the link below:

https://es.mathworks.com/help/matlab/matlab_prog/anonymous-functions.html

You can also make use of the MATLAB function `csape`, which is related with splines.

Problem 1. Write a program that imports the data needed from the spreadsheet and asks you for a 3 letter code. If that code represents one of the countries you were assigned ('ESP' for Spain,...), the program should give back the information and graphs relevant about that country. If the code is 'ALL', the program will deliver your written response to the questions below.

What information about each country? The one resulting from the following process:

- (1) Construct your data set: Make averages of cases and deaths on each country for the last 7 days up to each particular date (the days at the beginning are afterwards ignored: we just use them to be able to compute the averages on a particular first day, d_0). That will give you a table with data of dates d_k , cases C_k and deaths D_k on average on the 7 days up to that day, for $k = 0, \dots, n$. If there are gaps between two consecutive dates, say d_{k+1} and d_k of at least 18 days, add 3 more lines to the table with values $(d, C, D) = (d_k + 2, 0, 0)$, $(d, C, D) = (d_k + 10, 0, 0)$ and $(d, C, D) = (d_{k+1} - 2, 0, 0)$.

- (2) Prepare the problem: Now, divide your data into two kinds of triples (d, C, D) : one is for constructing interpolation polynomials and it takes the values of the table on days $d_0, d_2, d_4, \dots, d_{2t}, \dots, d_n$ (or d_{n-1} depending on how many values there are). We will call this sample the *working sample*. The rest of the values corresponding to days $d_1, d_3, d_5, \dots, d_{n-1}$ (or d_{n-2}) will be the *control sample*.
- (3) Interpolate: Construct a natural cubic spline function that interpolates the data for dates and cases for the working sample. Do the same for the dates and deaths.
- (4) Measure: Measure the errors on each day of the control sample (error = difference between the spline and the control sample value, for each day) and the relative total error on cases and on deaths.
- (5) Compare: Repeat the process with the original data without averaging.
- (6) View: Produce a graph with the evolution in time of both the average cases and the real cases on each day, and the corresponding spline functions (4 functions in 1 graph). Do the same for the deaths in a different graph.

Questions to be answered about the group of countries (case = 'ALL'):

- (Q1) Make a reasoned ranking of the 4 countries with regards to what you judge to be the quality of their reports.
- (Q2) What is more predictable in relative terms? Cases or deaths? How do you interpret that?
- (Q3) For the 1 week averages, are the errors for some country much larger on some particular day of the week? Is there an effect of the day of the week on the quality of the report?

Teams and countries

Team1	Spain	USA	Russia	Argentina
Team2	Spain	Brazil	Colombia	Mexico
Team3	Spain	India	Colombia	SouthAfrica
Team4	Spain	USA	Colombia	France
Team5	Spain	Brazil	Peru	Argentina
Team6	Spain	India	Russia	Mexico
Team7	Spain	USA	Peru	Mexico
Team8	Spain	Brazil	SouthAfrica	France
Team9	Spain	Russia	Peru	France
Team10	Spain	USA	Brazil	India