Numerical Algorithm: Assignment 1

Marco Busato 852074

June 2020

Contents

1 Part 1: Strain design

In the first part of the assignment we are asked to use the studied approximation techniques to solve the strain problem design.
In particular the use of these three functions:

- polynomial interpolation function (implemented through *polyfit* and *polyval* functions in Matlab)
- cubic spline function (implemented through *spline* function in Matlab)
- linear approximation function, custom linearization function implemented by me.

One needs to determine the length variation Δl_p of each mast, under a design stress $\sigma_p = 735\frac{kg}{cm^2}$
The interpolation results is shown in figure 1 and the value ϵ_p obtained is shown in table 1.
Moreover the table shows the maximum and minimum values of the considered functions in the interval.

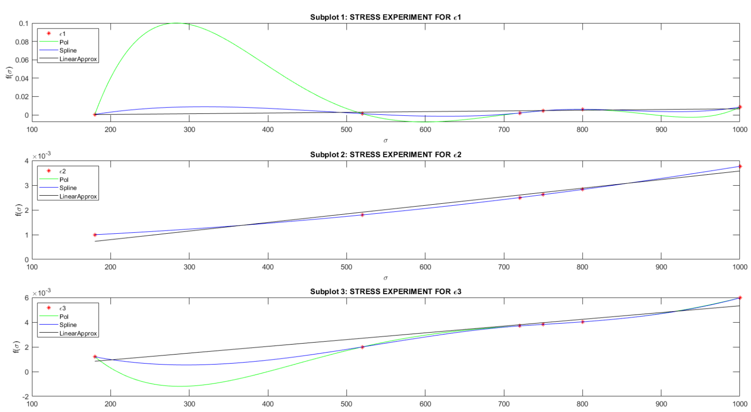

Figure 1: Interpolant functions for each stress experiments

	ϵ_1			ϵ_2			ϵ_3		
	min	ϵ_p	max	min	ϵ_p	max	min	ϵ_p	max
Polynomial	0.0020	0.0034	0.0045	0.0025	0.0026	0.0026	0.0038	0.0038	0.0038
Spline	0.0020	0.0033	0.0045	0.0025	0.0026	0.0026	0.0037	0.0038	0.0038
Linear Approx.	0.0044	0.0046	0.0047	0.0026	0.0027	0.0027	0.0038	0.0039	0.0040

Table 1: Min and max value in the interval $720 \leq \sigma_p \leq 750$

1.1 Results

As we can see from Figure 1 the polynomial interpolation for ϵ_1 and ϵ_2 is subjected to the Runge effect, thus is not an optimal interpolation. In order to avoid Runge effect one can consider the Spline function or still consider the polynomial function evaluated with the Chebyshev node. However the spline function pass through all the points in these experiment which is unlike to happen when the measurement are subjected to errors (such as incorrect instrument reading). Hence we can conclude that the best method that approximate the values is the linear approximation.

Finally the variation Δl_p for the selected methods are: $\Delta l_p^{\epsilon_1} = 4.6\text{x}10^{-2}$ $\Delta l_p^{\epsilon_2} = 2.7\text{x}10^{-2}$ $\Delta l_p^{\epsilon_3} = 3.9\text{x}10^{-2}$

2 Part 2: Time series of epidemiological data: data extrapolation

In this second part is asked to use the techniques for approximation/interpolation to perform data extrapolation, in particular applied to COVID-19 outbreak in Italy. The data considered refers to a particular region (Marche) and to the hospitalized individuals in the lapse of time from February 1^{st} to April 15^{th}. Hence this time lapse is subdivided into two groups and it's been considered a number N_e=6 of points for data extrapolation:

- data (t_i^{int}, y_i^{int}) with i=3,..,N-N_e use for interpolation/approximation[1]
  We refer to N^{int} as N^{int}=N-N_e
- data (t_i^{ext}, y_i^{ext}) with i=N-N_e,...,N use for extrapolation

2.1 Interpolation functions

For each n=1,...,20 is asked to compute:

- the n+1 nodes for the interpolation
- compute the polynomial interpolant function (polynomial and spline) on the n+1 nodes
- the relative errors on the internal and external point of the considered function (polynomial and spline)

2.1.1 Polynomial function

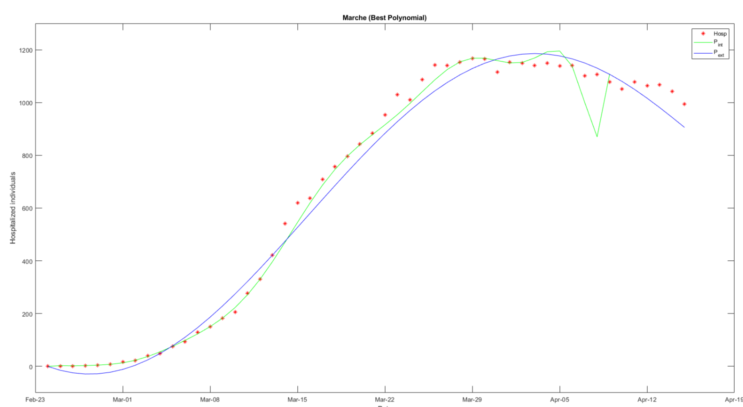

Figure 2: Best interpolant polynomial functions that minimize the errors

[1]For the first two days the Marche region registered 0 cases

n	e_n^{int}	e_n^{ext}	min Pn	max Pn
1	1,41165E+02	8,30199E-01	2,27374E-13	1,25460E+03
2	2,08289E+02	5,49839E-01	-2,27374E-13	1,16861E+03
3	1,24183E+02	1,03403E+00	-7,86296E+01	1,23739E+03
4	1,02855E+02	6,34750E-01	-6,27761E+01	1,20821E+03
5	5,23380E+01	3,60293E-01	-2,88244E+01	1,18571E+03
6	1,27762E+02	2,20970E+00	-5,68434E-13	1,17629E+03
7	2,85102E+02	1,71822E+01	-3,48463E+01	8,20945E+03
8	7,84366E+01	1,30752E+01	-4,80935E+03	1,21513E+03
9	7,01232E+02	7,61483E+01	-3,51290E+04	1,38966E+03
10	1,86012E+02	7,77558E+01	-3,65772E+04	1,33541E+03
11	9,64407E+02	4,80238E+02	-1,43409E+02	2,49678E+05
12	2,62587E+02	3,81693E+02	-1,99317E+05	1,63647E+03
13	2,98182E+00	3,78117E+02	1,10845E-10	2,09571E+05
14	1,28265E+03	2,28659E+03	-2,44492E+03	1,33419E+06
15	1,44993E+02	6,45243E+03	-8,98102E+02	3,80452E+06
16	1,06597E+03	4,34844E+03	-6,13316E+02	2,48617E+06
17	1,50887E+04	2,92685E+05	-1,83907E+08	3,78298E+04
18	3,53112E+04	4,46367E+05	-1,03367E+05	2,92177E+08
19	7,11103E+03	2,15475E+06	-1,43011E+09	7,48645E+04
20	1,83763E+03	1,38418E+06	-9,41807E+08	2,86959E+04

Table 2: Values of polynomial function and the errors in the interval (t_0, t_N)

As we can see in the Table 2 the best interpolant polynomial for the internal nodes that minimize the error is the one that have the degree 13 and the best interpolant polynomial for the external nodes that minimize the error is the one that have the degree 5.
Notice that the higher is the polynomial's degree, the higher is the error due to the Runge effect.

2.1.2 Spline function

As we can see in the Table 3 the best interpolant spline for the internal nodes that minimize the error is the one computed on 19 nodes and the best interpolant spline function for the external nodes that minimize the error is the one computed on 6 nodes.
As we expected the spline functions perform a better job than the polynomial, in fact if we take, for example, the ev_n^{sn} column the error oscillate between 0.2 and 1.6.

2.2 Approximation functions

In the last part of the assignment we were asked to study the data following the SIR model, were the initial population is subdivided into three categories:

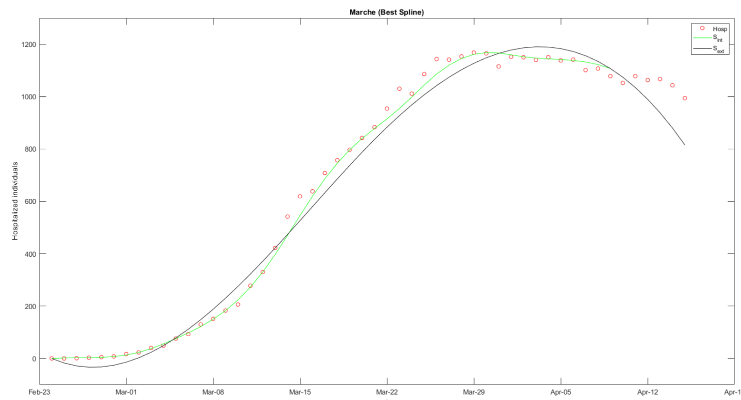

Figure 3: Best interpolant spline functions that minimize the errors

n	ee_n^{sn}	ev_n^{sn}	min Sn	max Sn
1	1,41165E+02	8,30199E-01	0,00000E+00	1,25460E+03
2	2,08289E+02	5,49839E-01	0,00000E+00	1,16861E+03
3	1,24183E+02	1,03403E+00	-7,86296E+01	1,23739E+03
4	9,48911E+01	7,36885E-01	-5,75605E+01	1,21195E+03
5	5,94948E+01	5,87657E-01	-3,32881E+01	1,18990E+03
6	8,58303E+00	2,48950E-01	0,00000E+00	1,15555E+03
7	2,93722E+01	1,12545E+00	0,00000E+00	1,41555E+03
8	4,64383E+00	4,17620E-01	0,00000E+00	1,15453E+03
9	2,17788E+01	1,11340E+00	-1,21049E+01	1,41686E+03
10	2,18831E+01	3,83735E-01	-1,18414E+01	1,17058E+03
11	2,78814E+00	1,62596E+00	0,00000E+00	1,15584E+03
12	1,30899E+01	9,51845E-01	0,00000E+00	1,16433E+03
13	3,33544E+00	5,12591E-01	0,00000E+00	1,16800E+03
14	5,73819E+00	3,82899E-01	-2,38633E+00	1,16500E+03
15	3,04179E+00	4,26199E-01	-8,89130E-01	1,16649E+03
16	2,56084E+00	6,68956E-01	0,00000E+00	1,27788E+03
17	5,23114E+00	5,98739E-01	0,00000E+00	1,25448E+03
18	8,27691E+00	1,51483E+00	-4,15689E+00	1,16800E+03
19	2,19645E+00	1,52621E+00	0,00000E+00	1,16630E+03
20	4,10679E+00	7,44629E-01	0,00000E+00	1,16500E+03

Table 3: Values of Spline function and the errors in the interval (t_0, t_N)

Susceptible, Infected and Recovered. Finally it has been assumed that the hospitalized individuals follow an exponential function:

$$H(t) = H(t_0)e^{\alpha(t-t_0)} \tag{1}$$

Therefore, it has been asked to approximate α by doing a linear regression on the logarithm of the function

$$log(H(t)) = log(H(t_0))\alpha(t - t_0) \tag{2}$$

2.2.1 Assignment 1

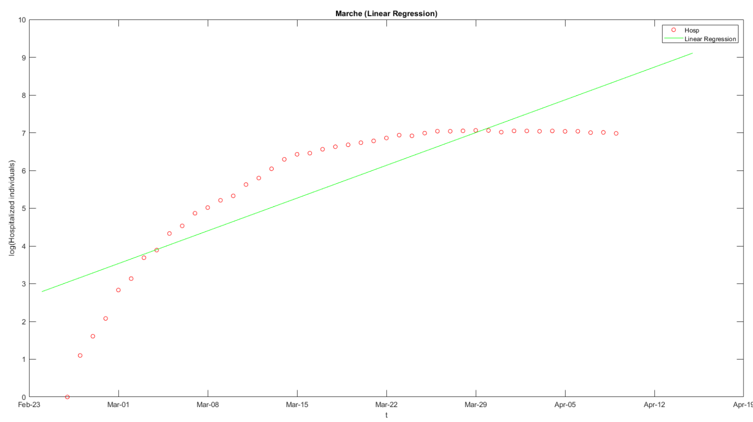

Figure 4: Linear regression on whole set of data

In Figure 4 a linear regression on whole set of data is presented. The slope of linear regression $\alpha = 1.5917$.
This is not a significant result because the increase of hospitalized individuals is not always an exponential function, so in the next part we consider a linear regression on windows of six days.

2.2.2 Assignment 2

As mentioned before in this part is considered a linear regression on windows of six days.
The first 2 days are omitted because the number of the hospitalized individuals is 0, so the week 8 is composed only of 2 points[2].
Thus it has been asked to use the linear regression of α_8 to approximate the data in the extrapolation window.
The results of the approximation can be seen in Figure 5 and the values of α and the relative errors are reported in the Table 4.
As we can see in the Table 4 the value of α is decreasing during the last week, that is a sign that the restrictive measures imposed by the Italian government have been helpful

[2]Since the interpolation point are 44 and it is not a multiple of 6

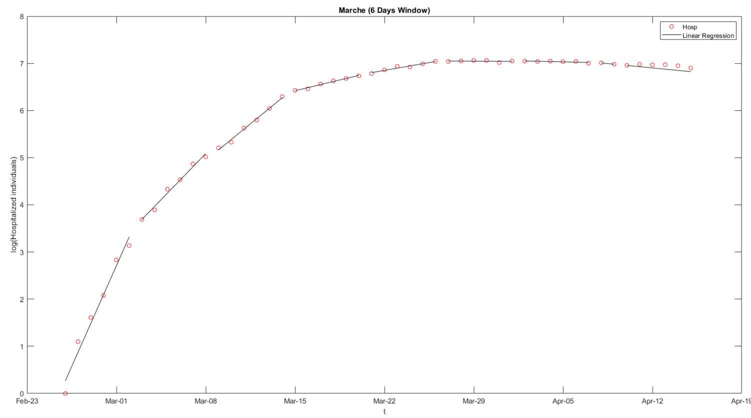

Figure 5: Linear regression on six days window

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
α	1.1413	0.5223	0.4143	0.1215	0.0888	-0.0029	-0.0121	-0.0188	-0.0188
re_w	0.8557	0.9663	0.9884	0.9920	0.9938	2.9814	3.9750	1.9598	14.9274

Table 4: Values of α and relative errors in the interval (t_0, t_N) subdivided into 9 subinterval

2.2.3 Assignment 3

Without the restriction imposed by the Italian government the hospitalized individuals up to the April 15^{th} would have been 1.2533e+13. However is impossible to have 1.2533e+13 hospitalized individuals in a region of 1,525,271 inhabitants[3], therefore all region's population would have been hospitalized around March 20^{th}.
The results can be seen in Figure 6.

[3]Data from ISTAT site available here

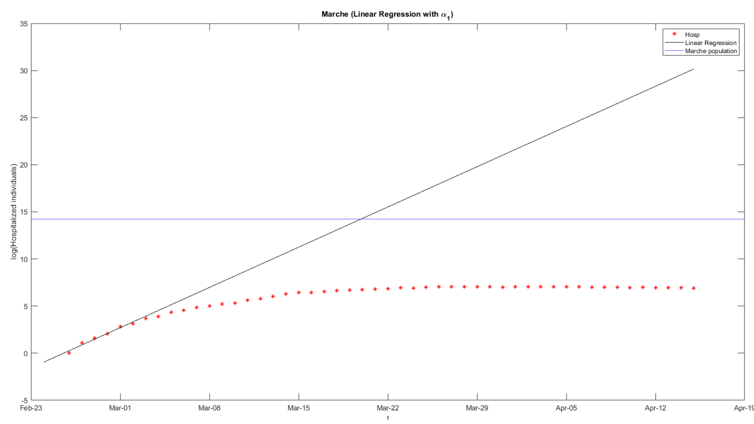

Figure 6: Linear regression on whole set of data using α_1