Assignment 1
The relationship between solar radio flux F10.7 and sunspot number

Performance - Thursday, September 3, 2020
Due to submit a performance report – Wednesday, September 9, 2020, 23:59 p.m.

The objective of this Assignment is to understand the relationship between main indicators of solar activity, sunspot number and the solar radio flux at 10.7 cm (2800 MHz) by applying multi-dimensional linear regression technique. This will bring about a deeper understanding of the Least-Square method (LSM), acquaintance with main indicators of solar activity, and more detailed qualitative and quantitative understanding of the important interactions. This exercise is also important for the following topics of the course.

Brief information on sunspot number and solar radio flux at 10.7 cm.
Sunspot number and solar radio flux at 10.7 cm are main indicators of solar activity. Sunspot number is daily measured by observatories over the globe and is determined as $R = k(n + 10g)$, here n – number of observed sunspots, g – number of observed sunspot groups, k – coefficient of a telescope. Solar radio flux at 10.7 cm is a measurement of radio emission at a wavelength of 10.7 cm (2800 MHz) from all sources present on the solar disk (sunspots, solar flares, solar proton events, white light faculae fields).

This Assignment is performed in the class by students as in teams of 3-4 and the team will submit one document reporting about the performance till the given deadline. Within your group, you may discuss all issues openly, and discuss and debate until you reach a consensus.

Here is the recommended procedure:

1. The code (Matlab/Python) should be commented.
   It should include:
   - Title of the Assignment, for example
     %Relationship between solar radio flux F10.7 and sunspot number
   - The names of a team, indication of Skoltech, and date, for example,
     %Tatiana Podladchikova, Skoltech, 2020
     Main procedures also should be commented, for example
     %13-month running mean
     …here comes the code

2. Download monthly mean sunspot number and solar radio flux F10.7cm measurements from Canvas. Every team will have different files to work with. You can download either Matlab or txt data (for Python):
   Group 1 - data_group1.mat
   Group 2 - data_group2.mat
   Group 3 - data_group3.mat
   Group 4 - data_group4.mat
   Group 5 - data_group5.mat
   Group 6 - data_group6.mat
   Group 7 - data_group7.mat
   Group 8 - data_group8.mat
   Group 9 - data_group1.mat
   Group 10 - data_group2.mat
   Group 11 - data_group3.mat
   Group 12 - data_group4.mat
   Group 13 - data_group5.mat

Group 14 - data_group6.mat
Group 15 - data_group7.mat
Group 18 - data_group8.mat
Group 19 - data_group1.mat
Group 20 - data_group2.mat

Format of data:
1 column - year
2 column – month
3 column – monthly solar radio flux at 10.7 cm
4 column – monthly sunspot number

3. Make scatter plot between monthly mean sunspot number and solar radio flux F10.7cm. **Every plot should contain: title, title of x axis, title of y axis, legend of lines on plot.** Make a conclusion if relationship between solar activity indicators is observed.

4. Make smoothing of monthly mean data (sunspot number and solar radio flux F10.7) by 13-month running mean. Plot results.
13-month running mean $\bar{R}$

$$\bar{R} = \frac{1}{24}R_{i-6} + \frac{1}{12}(R_{i-5} + R_{i-4} + \cdots + R_{i-1} + R_i + R_{i+1} + \cdots + R_{i+5}) + \frac{1}{24}R_{i+6}$$

Comment:
First six months in the available data are averaged to get the smoothed estimates.
The same with last six months of data.

5. Construction of multi-dimensional linear regression.
Multi-dimensional linear regression
$F_i = \beta_0 + \beta_1 R_i + \beta_2 R_i^2 + \beta_3 R_i^3 + \varepsilon_i$ (1)

Here

$F = \begin{vmatrix} f_1 \\ f_2 \\ \dots \\ f_N \end{vmatrix}$ – Vector of dependent variables, regressand, solar radio flux at 10.7 cm,

$f_1,\ f_{2,\dots,}f_N$ –flux at different times.

$R = \begin{vmatrix} 1 & r_1 & r_1^2 & r_1^3 \\ 1 & r_2 & r_2^2 & r_2^3 \\ \dots & \dots & \dots & \dots \\ 1 & r_N & r_N^2 & r_N^3 \end{vmatrix}$ – Matrix of independent variables, regressors,

$r_1, r_2, \dots, r_N$ - sunspot number at different times.

$\beta = \begin{vmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{vmatrix}$ - vector of coefficients

6. Determine vector of regressands (dependent variable), matrix of regressors vector (independent variables), and vector of coefficients.

7. Determine vector of coefficients by LSM

The vector of coefficients is determined according to the given equation

$$\beta = (R^T R)^{-1} R^T F$$

8. Reconstruct solar radio flux at 10.7 cm on the basis of sunspot number using Equation (1)

9. Determine the variance of estimation error of solar radio flux at 10.7

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(f_i - \hat{f}_i\right)^2$$

10. Make conclusions to the Assignment.
    Conclusions should be done in a form of a learning log. **A learning log** is a journal which evidences your **own learning and skills development.** It is not just a diary or record of **"What you have done"** but a record of **what you have learnt, tried and critically reflected upon.**

11. Prepare performance report and submit to Canvas:
    Performance report should include 2 documents:
    1) A report (PDF) with performance of all the items listed above
    2) Code (PDF), or a code and report together

Notes:

- PDF report should contain the names of team members, number of the assignment
- All questions of the assignment should be addressed
- All figures should have a caption, all axes should have labels, a legend to curves should be given, and short conclusions/discussions/results related to figures should be provided.
- The overall conclusion to the assignment should be provided in a form of a learning log.