AOSC 652: Analysis Methods in AOSC

Assignment #5: Least Squares Fitting
Due: Monday, 3 October 2016 (at start of class)

Name: ____________________________
120 points total
Late penalty: 10 pts per day

This assignment involves writing a FORTRAN code that solves for the coefficients of a quadratic least squares fit ***and*** the use of least squares fitting to examine the trends in global average temperature time series that was used in last week's assignment.

1 (60 points). In class on Wednesday, we presented the nomenclature for solving for the three coefficients *a*, *b*, and *c* that define the quadratic least squares fit of the equation:

$$y = a + b\,x + c\,x^2$$

where y & x represent a series of data points.

a) Copy file ~rjs/aosc652/week_05/data_fit.f to your work area and **add appropriate lines of code to subroutine quad_lst_sqr to compute the coefficients *a*, *b*, and *c* of a quadratic least squares fit** to the input data. Material presented in class on Wednesday is essential for completion of this task.

Please take your time to "understand" how to proceed before you start typing away ☺. ***Please ask for help if you are not certain how to proceed***. If you are not able to complete this task, please let one of the instructors know. We will provide working code, <u>under the provision that you may not share this code with any other student</u>. You will not get full credit if you choose this option, but the questions below require working code (as well as considerable thought) is needed to answer the questions below.

To test that the quadratic least squares fit is working properly, copy file:

~rjs/aosc652/week_05/quad_function.dat

to your work area and test your program data_fit.f using this file as input data for a quadratic fit.

To understand what values of *a*, *b*, and *c* your code should supply, take a look at program ~rjs/aosc652/week_05/quad_gen.f that was used to generate file quad_function.dat.

What values of *a*, *b*, and *c* ***should*** be produced for a quadratic fit to data in quad_function.dat?

____________ ____________ ____________

Now, fit data in column 2 vs column 1. What values for *a*, *b*, and *c* ***does*** your code produce?

Note: these values may not exactly match the ideal values due to numerical round off issues.

____________ ____________ ____________

The data in column 3 of file quad_function.dat was generated by adding "random noise" to the function. Now conduct a quadratic fit of data in column 3 vs column 1. What values for *a*, *b*, and *c* does your code produce for this fit?

____________ ____________ ____________

b) Prepare a plot of the data in column 2 vs column 1 (plot these as points) ***and*** the quadratic fit to column 2 versus column 1 (plot the fit as a line).

Prepare a second plot of the data in column 3 vs column 1 (plot these as points) ***and*** the quadratic fit to column 3 versus column 1 (plot the fit as a line, just as above).

Please turn in these two plots, along brief answers to the following:

i) If the coefficients differ, do the fits still look similar? Or, do you get markedly different fits?

ii) Most importantly, state whether your plots "make sense" supported by a sentence or two.

2. (60 points) Now we will apply the linear and quadratic least squares fit routine to the 5 year running mean of the global average temperature anomaly that we computed in Homework #4b.

Construct a plot, starting at **1880** and ending at **2060**, that shows:

i) the 5 year running mean of the global average temperature anomaly ($\Delta T_{GMST}^{5yr\text{-}RM}$)

ii) a linear, least squares fit to $\Delta T_{GMST}^{5yr\text{-}RM}$

iii) a quadratic least squares fit to $\Delta T_{GMST}^{5yr\text{-}RM}$

For this plot, we would like to see the fits ***extended to year 2060***, even though the 5 year running mean ends in year 2013 !

Note: you will either have to modify program data_fit.f to extend the fits out to year 2060 OR write another small code to evaluate the fits out to 2060. We would prefer you use FORTRAN or some other linux routine for this task, rather than a calculator ☺. However, do whatever you need to do, to get the job done correctly. Clearly it is much easier to "extend" a linear fit (could evaluate only for end point of 2060 and connect with a straight line) than the quadratic fit (this could be evaluated for years 2015, 2020, 2025, 2030, 2035, etc. if not every year between 2014 and 2060).

Now the fun part ☺. Compute how much temperature changed, from **1900** to **2013**, for both the linear least squares and the quadratic fit to $\Delta T_{GMST}^{5yr\text{-}RM}$:

$\Delta T_{1900 \text{ to } 2013}$ (linear) =

$\Delta T_{1900 \text{ to } 2013}$ (quadratic) =

Comment on the similarity, or lack thereof, of these two numerical values:

Now, compute how much temperature is "predicted to change", from 1900 to 2060, for the linear least squares and quadratic fits to this data record:

$\Delta T_{1900 \text{ to } 2060}$ (linear) =

$\Delta T_{1900 \text{ to } 2060}$ (quadratic) =

Comment on the similarity, or lack thereof, of these two numerical values:

Finally: is the functional form of the fit to past temperature important for the projection of future temperature? Answer with either "yes" or "no", and a supporting sentence.

As always, please turn in a hard copy of all code used to complete this assignment, printed using enscript and the full pathname of the code.