Non-linear Least Squares

Consider the non-linear model $f(c,t)$ where $c \in \mathbf{R}^N$ with $N = 6$ given by

$$f(c,t) = c_1 \sin\big(c_2 + c_4 \sin(c_5 + c_6 t) + c_3 t\big). \tag{1}$$

This model represents a sine wave with amplitude c_1, phase c_2 and frequency c_3 modulated by another sine wave with amplitude c_4, phase c_5 and frequency c_6. Suppose this model is used to generate the data

$$y_i = f(c,t_i) + \eta_i \qquad \text{for} \qquad i = 1,\ldots,M$$

where $c \in \mathbf{R}^N$ is a vector of parameters, t_i is a sequence of times and η_i a sequence of independent normally distributed random variables with mean zero and variance σ^2. The goal of this programming project is to find the value of c for which the sum of squares

$$E(c) = \sum_{i=1}^{M} \big|f(c,t_i) - y_i\big|^2$$

is minimized using the Gauss–Newton method, see Solomon [1].

Recall that the Gauss–Newton method approximates a solution to the non-linear least squares problem by successively solving a sequence of linear least squares problems. Let c be an initial guess for the parameters. First, linearize f about the estimate c as

$$f(c+\xi,t) \approx f(c,t) + \nabla_c f(c,t)\cdot\xi.$$

Next, define $A \in \mathbf{R}^{M\times N}$ to be the matrix with entries a_{ij} and $B \in \mathbf{R}^M$ to be the vector with entries b_i given by

$$a_{ij} = \frac{\partial f(c,t_i)}{\partial c_j} \qquad \text{and} \qquad b_i = y_i - f(c,t_i) \tag{2}$$

where $i = 1,\ldots,M$ and $j = 1,\ldots,N$. Finally, minimize the the linear least squares problem

$$L_c(\xi) = \sum_{i=1}^{M} \big|\nabla_c f(c,t_i)\cdot\xi - y_i + f(c,t_i)\big|^2$$

using by solving the over-determined matrix equation $A\xi = B$ in the least squares sense so that $\|A\xi - B\|_2$ is minimized. At this point we have found a new approximation of the parameters as $c+\xi$. To obtain a sequence of approximations using this procedure replace c in equation (2) by $c+\xi$ and iterate.

For each student a random vector $c \in \mathbf{R}^6$ was chosen such that

$$c_j \in [1,3] \qquad \text{for} \qquad j = 1,\ldots,6.$$

Then 120 data points (t_i, y_i) were written to a file using the model given in (1) where the noise term η_i has $\sigma = 0.05$. Please download the data file corresponding to your name from our website `http://fractal.math.unr.edu/~ejolson/702/` and answer the questions on the following page.

References

1. Justin Solomon, *Numerical Algorithms*, Chapter 12, CRC Press, 2015.

Your answers should be presented in the form of a written report with source code, graphs, tables and program output where appropriate. Style of presentation counts as well as spelling, punctuation and grammar. Please work independently; however, it is fine to visit the UNR Writing Center for help with writing style. If you have any difficulties please talk with me in my office hours or set up an appointment.

1. Plot the points in the data file. Describe the qualitative behavior of the frequency modulated wave and try to guess the frequencies c_3 and c_6.

2. Make other guesses for c_1, c_2, c_4 and c_5 and perform the Gauss–Newton non-linear optimization algorithm. Does it converge? If so, to what?

3. Write a program that randomly chooses $c \in \mathbf{R}^6$ such that $c_j \in [1, 3]$ for $j = 1, \ldots, 6$ and for each random choice of c performs the Gauss–Newton optimization algorithm. You may use any programming language you prefer. Please include a listing of the code in your report.

4. Run the code as many times as needed to find ten different guesses for c that converge to values such that $c_j \in [1, 3]$ for all j. Print those initial guesses along with the resulting vectors they converge to.

5. Denote by p^k where $k = 1, \ldots, 10$ the ten resulting vectors obtained by Gauss–Newton optimization in the previous step. Are all the p^k equal or are some different? Note that different limits indicate the presence of local minima in the for the non-linear optimization problem.

6. For each distinct choice of parameters $c = p^k$ compute $E(c)$ and arrange the results in a well-formatted table. From these results suggest the choice of parameters c from which the data in the file was most likely to have been generated.

7. [Extra Credit] Try different ways to increase the efficiency of the Gauss–Newton non-linear optimization routine. Consider including checks and corrections to to make sure the values of c don't exit the range $c_j \in [1, 3]$. You may also want to try the Levenberg–Marquardt algorithm which uses Tikhonov regularization as given in [1] and other references to limit the size of the steps in the Gauss–Newton iterations.