Lab Assignment 6

Linear and Nonlinear least Squares

ACS II

Fall 2016

Due 8/4/16

Background

Given a set of independent and dependent variables one often wishes to find some relationship between the values of the independent variables and the values of the dependent variables. In the real world, this is often hampered by the presence of noise in the data, which can obfuscate the underlying relation and prevent there from being a function of the desired form which exactly reproduces the data. In such circumstances, the goal shifts from being the exact recovery of the response function that describes the data to the similar goal of recovery of the optimal response function for the data. This is often case as attempting to minimize residuals between a trial response function and the data. In certain cases where the response appears or is expected to be linear, this can be as simple as solving a linear system, which can be performed in a single step. However, when the response is expected to be nonlinear, an iterative approach such as Gauss-Newton may be needed.

Linear Least Squares for Parameter fitting

Consider the case of a 1D linear model, where the response function consists of a linear combination of variables

$$Y = \sum_{i=1} a_i X_i + a_0$$

This can be write in matrix notation as

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}^T = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

In the case were error exists in the model, this linear system will not have an exact solution, and instead we seek the optimal solution; that which minimizes the residuals as computed in some norm. The minimization in the L2 norm will minimize the sum of the squared distances, giving us what is known as the linear least squares fit. The linear least squares fit can be computed a variety of ways, but one of the simplest of with is the use of the normal equations, which are given as

$$A^T A x = A^T b$$

Which can be solved a variety of ways.

Once you have parameter estimates for a_0 and a_1, you can perform hypothesis tests on them to determine if your linear model accurately represents them. Casting your hypotheses as

$$H_0: a_1 = a_t$$

$$H_1: a_1 \neq a_t$$

We can calculate a T-test statistic for this hypothesis, as the residuals should fall along a T-distribution with $n-2$ degrees of freedom. The test statistic for the slope parameter is given by

$$T_0 = (\widehat{a_1} - a_t)/se(\widehat{\alpha_1})$$

Where $se(\widehat{\alpha_1})$ is the standard error of the fitted $\widehat{\alpha_1}$ and is given by

$$se(\widehat{\alpha_1}) = \sqrt{\frac{\frac{\sum_{i=1}^{n} r_i^2}{n-2}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

If we let the theoretical parameter $a_t = 0$ and them perform the statistical hypothesis test with our least squares fit parameter $\widehat{a_1}$, a failure to reject the null hypothesis indicates that our data is not accurately represented by the fitted model.

Nonlinear least squares for parameter fitting

Often, you will wish to perform fits for nonlinear models; in this case, one must turn to nonlinear least squares approaches. Consider the problem of minimizing the residuals for a given model.

$$r(x_i, \beta) = y_i - M(x_i, \beta)$$

As the x, y pairs are knows, we can recast this merely as a function of the parameter set β.

$$S(\beta) = \sum_{i=1}^{m} r_i^2(\beta)$$

Starting with an initial guess for the model parameters, β^0, one can iterate successively until a minimum is reached, with each subsequent parameter set given by

$$\beta^{s+1} = \beta^s - (J_r^T J_r)^{-1} J_r^T r(\beta^s)$$

Where J is the Jacobian matrix of the residual functions and given by

$$(J_r)_{i,j} = \frac{\partial r_i \beta^s}{\beta_j}$$

This iteration for new parameter values can continue until some tolerance or failure criteria is met.

Experiment

Your tasks are as follows:

- Take the data is sample1.dat and solve for the parameters of a linear model for it using the normal equations. Then perform a T test on the slope parameter of this linear model. Plot the data along with your fit of the data, and report your T statistic as well as whether or not you accept the null hypothesis. Use $a_t = 0$ and discuss what this means.
- Take the data in sample2.dat and use the Gauss Newton method to fit a model of the form $y = a_3x^3 + a_2x^2 + a_1x + a_0$. Choose an appropriate tolerance criteria and report on the convergence of your algorithm from two different initial guesses for the a_i parameters. Plot the data as well as your fits.

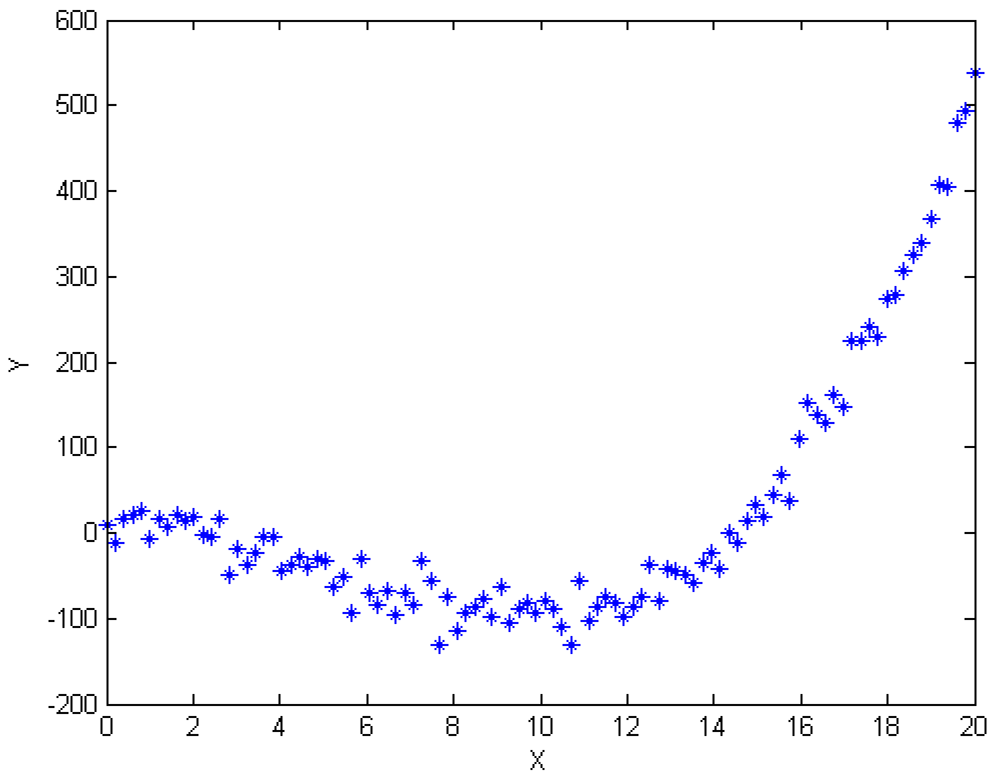

Figure 1:Scatterplot of the data from sample2.dat