Rev. Téc. Ing., Univ. Zulia
Vol. 5 No. 1, 1982.

RESUMEN
Tesis: M.S.

NON-LINEAR REGRESSION ANALYSIS FOR UNCONSTRAINED OPTIMIZATION USING NUMERICAL DIFFERENTIATION

(University of Oklahoma - 1978)

Rafael Romero Silva
Escuela de Petróleo
Facultad de Ingeniería
Universidad del Zulia.

Regression analysis is a method for establishing the functional relationship between two variables where the variation in one measurement is considered while the other is held fixed. In many engineering situations a straight-line relationship can express the dependence of one variable on another. This is referred to as the linear regression analysis. The relationship is

$$Y=A_0 + A_1X \quad \text{first-order equation}$$

where Y = dependent variable,which is measured at a given level of X X= independent variable, which is held fixed; A_0, A_1 = regression parameters to be determined from sample data.

This equation can be extended to the higher order (nonlinear regression analysis), such as

$$Y = A_0 + A_1X + A_2X^2 \quad \text{second-order equation}$$

$$Y = A_0 + A_1X + A_2X^2 + A_3X^3 \quad \text{third-order equation}$$

In each of these equations, one is interested in determining the values of the regression parameters A_0, A_1, A_2, A_3,. . . from the experimental data.

Non-linear regression analysis of observed pressure and production history to obtain reliable estimates of gas in place and other reservoir properties has been described in the oil industry literature.
The particular regression approach used in one of these studies employs a variation of the Gauss-Newton iterative method to reduce a least-squares objective function Q, defined by

$$Q = \sum_{i=1}^{n} Ri^2$$

the Ri are residuals of the reservoir model and are defined by:

$$Ri = Yi^{obs} - Yi^{calc}$$

where Yi^{obs}= observed volume of gas produced up to time level i, Yi^{calc} = volume of gas that would have been produced up to that time, given a particular set of reservoir and aquifer properties.

The Gauss-Newton iterative method used to calculate the regression variables uses the following equations:

$$\frac{dQ}{dX} = \frac{Q}{X_1 - X_2}$$

and

$$X_2 = X_1 - \frac{Q}{dQ/dX} \qquad (1)$$

When the curve Q vs X has an approximately parabolic shape there is a point where dQ/dX = 0, and the method will not work, because there will be a zero in the denominator of equation (1) and the whole term will be infinite.

So, in this case a regression approach different from Gauss-Newton should be used.

The primary objective of this thesis is the developing of a new technique of solving regression parameters using numerical differentiation to calculate the slope of the objective function Q (sum of squares) or partial derivatives of Q respective to the regression parameters (dQ/dX). Such a study involves writing a computer program to apply the new method of non-linear regression analysis. A subroutine to calculate the objective function values is provided. The main problem that is considered is the calculation of the least value of a given objective function Q(X) or sum of squares, where X is a vector of n real variables, and where there are no constraints on the values of the variables.

The method was applied to five different sets of data in order to test such a new method. The listings and results from the program output are shown in appendix A1of the thesis.