Anusandhanvallari
Vol 2023, No.1
May 2023
ISSN 2229-3388

Estimation of the Parameters Under Linear and Nonlinear Least Squares Methods

Gajraj Singh

Discipline of Statistics, School of Sciences,

Indira Gandhi National Open University, Delhi-110068 India

gajrajsingh@ignou.ac.in

ORCID Id: 0000-0003-0870-921X

Abstract: Estimating parameters using linear and nonlinear least squares algorithms is a crucial method in statistical modeling and data analysis. Under certain presumptions, the least squares approach minimizes the sum of squared variances between actual and expected values to produce accurate and objective estimations. The linear least squares method shows the model's parameters linearly, and analytical solutions can be produced using normal equations. The nonlinear least squares method, on the other hand, works with models where parameters enter nonlinearly and requires iterative numerical approaches, such as the Gauss-Newton or Levenberg-Marquardt algorithms. These methods are widely used for model fitting and prediction in the social, biological, engineering, and economic sciences. The accuracy and dependability of parameter estimate depend on a number of factors, including model definition, convergence criteria, and data quality. Thus, least squares estimation remains a dependable and adaptable method for the exploration of both linear and nonlinear interactions.

Keywords: Linear least squares, nonlinear least squares, Optimization, Gauss-Newton method.

1. Introduction

Quantities, structures, space, and change are all studied in mathematics. Another perspective on mathematics is knowledge, where we use the principles' logical reasoning to arrive at exact and novel results. Despite not being a natural science, mathematics is frequently very helpful for describing occurrences and addressing issues in the natural sciences because the specific structures that mathematicians typically examine are found in natural science, particularly physics. Although mathematics is essential to natural science, engineering, economics, and medicine, mathematicians frequently develop methods and structures separately from the sciences that use them. Hamedani(2002).

2. Methods of curve fitting

In many fields, it is crucial to examine the relationships between quantities. If we have data, we may wish to use a mathematical formula to express the relationship. Curve fitting work starts with determining the appropriate parameter values for such a formula. Typically, two variables behave differently in relation to one another. The relationship between two variables can follow a linear, logarithmic, exponential, partial, or similar pattern. For these behaviors, the curve fitting procedure falls into two broad types. There is a linear curve fitting group and a non-linear curve fitting group. Finding a non-linear model for and the relationship between a dependent variable and a group of independent variables can be done with this kind of curve fitting. It should be noted that although

this approach will provide an approximate match to the data,

When we study curve fitting methods, we will divide them into two categories

1. Linear curve fitting

2. Non-linear curve fitting

Linear curve fitting, which is restricted to fitting linear models, is the most popular type of curve fitting. Non-linear curve fitting, on the other hand, can fit any connection between independent and dependent variables. Compared to linear curve fitting, non-linear curve fitting is more computationally challenging and relies on numerical techniques.

2.1 Linear curve fitting: Suppose we have a set of points (t_1, y_1), (t_2, y_2),….. (t_n, y_n) and a mathematical function $m = (\vec{x}, t)$ whose shape is defined by the values in $\vec{x}$. A linear least squares problem is an unconstrained minimization problem of the form

$$\min_{\vec{x}} f(\vec{x}) = \min_{\vec{x}} \sum_{i}^{n} r_i(\vec{x})^2 \qquad (1)$$

Where $r_i(\vec{x}) = y_i - m(\vec{x}, t_i)$ and the function $m(\vec{x}, t)$ depends linearly on *x*. If the number of independent variables in this relationship is more than on one curve fitting model is called multiple linear. For example, *m*(*x*, *t*) could be a straight line $m(\vec{x}, t) = x_1 t + x_2$. We can also have a linear combination of several functions of *t*,

For example- $m(\vec{x}, t) = x_1 f_1(t) + x_2 f_2(t) + \ldots\ldots\ldots\ldots x_k f_k(t)$

For a linear least square problem (1) can be rewritten using vectors as follows:

$$\min_{\vec{x}} f(\vec{x}) = \min_{\vec{x}} \sum_{i}^{n} r_i(\vec{x})^2$$

$$= F(\vec{x})^T F(\vec{x})$$

$$F(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ . \\ . \\ f_k(\vec{x}) \end{pmatrix}^T$$

2.2 Non-Linear curve fitting: There are several similarities between linear and non-linear curve fitting. The definition of a linear least square issue and the definition of a non-linear one are extremely similar. To make it clear, we shall write it out completely.

Suppose we have a set of points (t_1, y_1), (t_2, y_2),….. (t_n, y_n) and a mathematical function $m = (\vec{x}, t)$ whose

shape is defined by the values in $\vec{x}$. A linear least squares problem is an unconstrained minimization problem of the form

$$\min_{\vec{x}} f(\vec{x}) = \min_{\vec{x}} \sum_{i}^{n} r_i(\vec{x})^2$$

Where $r_i(\vec{x}) = y_i - m(\vec{x}, t_i)$ and the function $m(\vec{x}, t)$ depends non-linearly on $\vec{x}$. This means that there are many more types of functions that could be used in non-linear curve fitting. That is why we often need to use numerical methods to find $\vec{x}$.

3. Gradient, Jacobian and Hessian:

In the coming sections we will discuss optimization problems and algorithms for curve fitting and in all of these subjects we will use the gradient, the Jacobian and the Hessian so we will dene these for convenience. We will use the notation $\vec{x} = (x_1, x_2, ..., x_n)$ to denote a row vector.

Definition 3.1: The gradient ∇ of a function f is a vector consisting of the function's partial derivatives is

$$\nabla f(x) = \nabla f(x_1, x_2 .....x_n) = \left(\frac{\delta f}{\delta x_1}, \frac{\delta f}{\delta x_2}, ..., \frac{\delta f}{\delta x_n} \right)^T$$

Definition 3.2: The Jacobian *J(x)* is a matrix that contains the gradients of the functions $r_1, r_2 .....r_m$.

$$J(x) = \frac{\delta r_j}{\delta x_i} = \begin{pmatrix} \nabla r_1(x) \\ \nabla r_2(x) \\ . \\ . \\ \nabla r_m(x) \end{pmatrix}, j = 1, ..., m, \; i = 1, ..., n$$

Definition 3.3: The Hessian matrix ∇^2 of a function f is the square matrix of the second partial derivatives.

$$G(x) = \nabla^2 f(x_1, x_2 .....x_n) = \begin{pmatrix} \frac{\delta^2 f}{\delta x_1^2} & \frac{\delta^2 f}{\delta x_1 \delta x_2} ..... & \frac{\delta^2 f}{\delta x_1 \delta x_n} \\ \frac{\delta^2 f}{\delta x_2 \delta x_1} & ... & \frac{\delta^2 f}{\delta x_2 \delta x_n} \\ ... & & ... \\ \frac{\delta^2 f}{\delta x_n \delta x_1} & \frac{\delta^2 f}{\delta x_n \delta x_2} & \frac{\delta^2 f}{\delta x_n^2} \end{pmatrix}$$

We only consider function that has continuous second order partial derivatives so that the order of the second partial derivative is not important by Schwarz's theorem, which means that this matrix is symmetrical.

4. Linear and Non-linear Optimization

4.1 Linear equation

A linear equation in n unknowns $x_1, x_2,...,x_n$ is an equation of the form

$a_1 x_1 + a_2 x_2 + ... + a_n x_n = b$, where $a_1\ a_2\ ...\ a_n$ and b are given real numbers.

A system of m linear equations in n unknowns $x_1, x_2,...,x_n$ is a system of linear equations

$$a_{11}x_1 + a_{12}x_2 + ...a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + ...a_{2n}x_n = b_2$$
$$.....$$
$$.....$$
$$a_{m1}x_1 + a_{m2}x_2 + ...a_{mn}x_n = b_m$$

We want determine if such a system has a solution, that is to find out if there exist numbers $x_1, x_2,...,x_n$ which satisfy all of the equations simultaneously. The system is consistent if it has a solution, otherwise the system is called inconsistent.

Note that the above system can be written concisely as

$$\sum_{j=1}^{n} a_{ij}x_j = b_i = 1,2,.....,m \qquad (2)$$

We can write the system of equation in the matrix form

$$\begin{pmatrix} a_{11}\ a_{12}.....a_{1n} \\ a_{21}\ a_{22}.....a_{2n} \\ ..... \\ ..... \\ a_{m1}\ a_{m2}.....a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_m \end{pmatrix} \qquad (3)$$

The matrix in (3) is called the coefficient matrix of the system. Solving a linear equation system in two or three geometric dimensions is equivalent to determining whether or not a family of lines or planes, has a common point of intersection [Igor Griva, Stephen G. Nash, Ariela Sofer(2008)]. Systems of linear equations are central to almost all optimization algorithms and are part of many optimization models. Linear equation systems can be used to represent constraints in a model. Finally solving systems of linear equation is an important step in the simplex method for linear programming and Newtons method for non-linear optimization, and is a technique used to determine dual variables (Lagrange multipliers) [Igor Griva, Stephen G. Nash, Ariela Sofer (2008)].

4.2 Linear Least Squares for data fitting

Linear least squares data can be used to fit linear models. We start with an example where we fit a quadratic function. For example: If we have the same number of points as unknown parameters we solve the problem exactly without using curve-fitting. We can fit a quadratic function to the data given in the table:

t 2 3 5

b 1 6 4

Quadratic function: $b(t) = x_1 + x_2 t + x_3 t^2$, x_1, x_2, x_3 are unknown parameters

Solution: $b(t_i) = b_i$ (4)

$$x_1 + x_2.2 + x_3.2^2 = 1 \Rightarrow x_1 + 2x_2 + 4x_3 = 1 \quad (5)$$

$$x_1 + x_2.3 + x_3.3^2 = 6 \Rightarrow x_1 + 3x_2 + 9x_3 = 6 \quad (6)$$

$$x_1 + x_2.5 + x_3.5^2 = 4 \Rightarrow x_1 + 5x_2 + 25x_3 = 4 \quad (7)$$

Now we can write the system of equations in the matrix form as:

$$\begin{pmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \\ 4 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -21 \\ 14.97 \\ -1.98 \end{pmatrix}$$

$$(x_1, x_2, x_3) = (-21, 14.97, -1.98)$$

$$b_i = -21 + 14.97t - 1.98t^2$$

We can also describe fitting of polynomials in a more general way.

Form: $y = a + bx + cx^2$ and data: $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$; here we still have the same number of points as unknown parameters so do not need to use least-square fitting.

Least square fitting of a polynomial:

If we have more data points of unknown parameters we need to use least squares fitting. If there are n data points and k < n parameters the model of the form is:

$$y(x) = a_1 + a_2 x + ... + a_n x^{k-1}$$

The system of the equation of the form:

$$\begin{pmatrix} 1 & x_1 .....x_1^{k-1} \\ 1 & x_2 .....x_2^{k-1} \\ ..... & \\ ..... & \\ 1 & x_n ..... \; x_n^{k-1} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix}$$

By using the principal of least square, we want to minimize S (the sum of squares of the residuals) is given by:

$$S = \sum_{I-1}^{n} e_i^2 = e_1^2 + e_2^2 + e_3^2 ..... + e_n^2 \quad (8)$$

Where $e_i = y_i - (a - bx_i + cx_i^2)$

$$S = \sum_{i=1}^{n} (y_i - a - bx_i + cx_i^2)^2$$

To find the minimum values of our unknown parameters a, b and c by the partial derivatives a, b and c respectively, $\frac{\partial S}{\partial a} = 0, \frac{\partial S}{\partial b} = 0$ and $\frac{\partial S}{\partial c} = 0$

We can take the partial derivatives of S given by (2.8), set them to zero and then obtain the following formulas:

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^2 \quad (9)$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3 \quad (10)$$

$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4 \quad (11)$$

Equation (9), (10) and (11) are called the normal equations. Note that (9), (10) and (11) together form a linear equations with unknowns parameters a, b and c. Solving this equations we find the values a, b and c. This is the solution to the least squares problem.

Example: Find best fit in the least-square sense for values a, b for $y = a + bx$ the given data

Sr No.	x	y	xy	x^2
1	0	1	0	0
2	1	1.8	1.8	1
3	2	3.3	6.6	4
4	3	4.5	13.5	9
5	4	6.3	25.2	16

$\sum x$=10 $\sum y$=16.9 $\sum xy$=47.1 $\sum x^2$=30

Solution: $y = a + bx$ (12)

$$\sum y = na + b\sum x \quad (13)$$

$$\sum xy = a\sum x + b\sum x^2 \quad (14)$$

Here, n=5, $\sum x$=10, $\sum y$=16.9, $\sum xy$=47.1 and $\sum x^2$ =30

Now solve the equations (13) and (14), we get the values of a and b are i.e. a=0.72 and b=1.33

$$y = 0.72 + 1.33x$$

The expression we derived does not serve as an exact solution to the initial system of linear equations. There will be discrepancies, but they are reduced in the least squares framework.

4.3 Non-linear Least Squares curve fitting

$$m(x_1, x_2, t) = x_1 e^{-x_2 t} \quad (15)$$

$$\frac{\partial m}{\partial x_1} = e^{-x_2 t} = 0 \text{ and } \frac{\partial m}{\partial x_2} = -t x_1 e^{-x_2 t} = 0$$

From the partial derivatives, we can see that $m(x_1, x_2, t)$ depends linearly on x_1 but non-linearly on x_2. The partial derivative with respect to x_1, is independent of x_1, but the partial derivative with respect to x_2 depends on variable x_2 and therefore $m(x_1, x_2, t)$) depends non-linearly on x_2. Here we have a definition of the differences of the linear equation with the non-linear equation.

Linear Equation-

1. The diagram is a straight line.
2. The graph gradient is constant.
3. Solving linear equations is simple and easy.

Non-linear Equation-

1. All parameters of the linear least squares are linear, but in non-linear terms, at least one parameter is a non-linear equation.
2. Parameters that exist in linear equations are limited, but the parameters that exist in non-linear equations are extensive.
3. Numerous processes in scientific and engineering fields can be represented using linear or other straightforward models; however, there are many processes that cannot be captured by linear models due to their inherently non-linear nature. For instance, in construction, concrete strength experiences a rapid initial increase, but after a brief period, it transitions into a curve that is not adequately represented by a linear model. Therefore, a non-linear model must be utilized instead.

Convergence rate- When employing an iterative numerical approach to solve an equation, if each iteration brings us closer to the solution, we refer to the method as convergent. The rate at which the sequence approaches its true value is known as the convergence rate. If we have a sequence x_0, x_1, x_2... x_n, they go to the main root x or, in mathematical language, they say, they are converging.

$$x_n \rightarrow x \text{ as } n \rightarrow \infty$$

That's true, if $$\lim_{n\to\infty} \frac{|x_{n-1} - x|}{|x_{n-1} - x|^b} = \lambda < \infty$$

with, b≥1, then we call b the rate of convergence.

1. If b=1, the sequence converges linearly.
2. If b=2, the sequence converges quadratically.
3. If 1< b < 2, the sequence converges super-linearly.

Definition 4.3.1 Let $x^* \in R$, and consider a sequence x_k for k = 0, 1, 2,.... The sequence x_k is said to converge to x^* if $\lim_{n\to\infty} |x^k - x^*| = 0$

1. Linear if ∃ c ∈ [0,1) and there is an integer K > 0 such that, for k ≥ K
$$|x_{k+1} - x^*| \le c\,|x_k - x^*|$$
2. Super linear if $c_k \to 0$ and there is an integer K > 0, such that for, k ≥ K
$$|x_{k+1} - x^*| \le c_k\,|x_k - x^*|$$
3. Quadratic if ∃ c ∈ [0,1) and there is an integer K > 0 such that; for k ≥ K
$$|x_{k+1} - x^*| \le c\,|x_k - x^*|^2$$

5. There are three methods based on Non-linear least squares:-

1. Newton method
2. Gauss Newton Method
3. Levenberg Marquardet method

5.1 Newton method

The Newton method is based on taking a second-order Taylor's approximation of a function $f(\vec{x})$. Now apply the Newton method to solve a non-linear system of equations of the form [Pereyra, (1967)].

$$\nabla f(\vec{x}) = 0$$

To find an approximation of the root of a function $f(\vec{x})$ we choose a point and draw the tangent line to $f(\vec{x})$ at that point. The point where this tangent line crosses the x-axis gives a new approximation and if this procedure is repeated many times you usually get a good approximation of the root.

Algorithm for the Newton method [H. B. Keller and V. Pereyra(1976)] and its notations:

$$\nabla f(\vec{x}) = J(\vec{x})^T r(\vec{x})$$

$$\nabla^2 f(\vec{x}) = \left(J(\vec{x}_k)^T\right) J(\vec{x}_k) + S(\vec{x}_k) \qquad (16)$$

$$S(\vec{x}_k) = \sum_i^m r_i(\vec{x}_k)\nabla^2 r_i(\vec{x}_k)$$

1. Starting from an initial guess $\vec{x}_0$ Note: (if we do not have $\vec{x}_0$, we choose one manually)

2. Compute new approximation $\vec{x}_{k+1} = \vec{x}_k - \nabla^2 f(\vec{x}_k)^{-1} \nabla f(\vec{x}_k)$
3. Compute $\Delta\vec{x}_k = \vec{x}_{k+1} - \vec{x}_k$. If the method is converging then $|\Delta \vec{x}_k|$ is expected to get smaller and smaller. If $|\Delta \vec{x}_k|$ is much smaller than the required accuracy the algorithm have found a good approximation. If $|\Delta \vec{x}_k|$ is still large, repeat step 2 again.

The Newton method is locally quadratically convergent [P.C.Hansen, V.Pereyra and G. Schere (2013)].

5.2 Damped Gauss-Newton Method-

If $S(\vec{x}_k)$ in (16) is small it can be interpreted as the problem being only slightly non-linear near $\vec{x}_k$. $S(\vec{x}_k)$ can be ignored and still have quadratic convergence. If $S(\vec{x}_k)$ is large we can get very poor convergence and instead of using the Newton method described before a better alternative is the damped Gauss-Newton method [P.C.Hansen, V.Pereyra and G. Schere (2013)].

1. Start initial guess $\vec{x}_0$ for k = 0, 1, 2,.....
2. $\Delta\vec{x}_k = \dfrac{J(\vec{x}_k^T) r(\vec{x})}{J(\vec{x}_k^T) J(\vec{x})}$
3. Choose a step length c_k so that there is enough descent $c_k \in (0,1)$ or $0 < c_k < 1$.

   There are several ways to determine appropriate c_k, a common method is to choose a

   c_k, such that $f(\vec{x}_k + c_k\Delta\vec{x}_k) < f(\vec{x}_k) + c_k \nabla f(\vec{x}_k)^T \Delta\vec{x}_k = f(\vec{x}_k) + c_k r(\vec{x}_k)^T J(\vec{x}_k)^T \Delta\vec{x}_k$
4. New $\vec{x}_{k+1} = \vec{x}_k + c_k\Delta\vec{x}_k$
5. Check for convergence similarly to the Newton method described in the previous section.

5.3 Levenberg-Marquardt Method

If the Newton method does not converge, or converges too slowly, we use the damped Gauss-Newton method, and if we do get the right convergence we go to Newton's method. Although we use Gauss-Newton's method for ill-conditioned problem mode, it might not give us satisfactory results since it tends to be slow [Netlib Repository at UTK and ORNL [Online] http://www.netlib.org]. The Levenberg-Marquardt method is based on solving the following equation.

$$\left(J(\vec{x}_k)^T J(\vec{x}_k) + \lambda_k I\right) = -J(\vec{x}_k)^T r(\vec{x}_k)$$

$$\Delta\vec{x}_k = \left(J(\vec{x}_k)^T J(\vec{x}_k) + \lambda_k I\right)^{-1} \left(-J(\vec{x}_k)_k^T r(\vec{x}_k)\right)$$

Method for solving:

1. Initial guess $\vec{x}_0$ for iteration k = 0, 1, 2, 3....

2. At each step *k* choose the Lagrange multiplier λ_k

3. Compute $\Delta\vec{x}_k$ using the formula above.

4. $\vec{x}_{k+1} = \vec{x}_k + \Delta\vec{x}_k$

5. Check for convergence similarly to the Newton method. There are different methods for choosing λ_k, below a commonly used one is described:

   1. The initial value λ_0 as $\left\|J(\vec{x}_0)^T J(\vec{x}_0)\right\|_2 \Rightarrow \lambda_0 \cong \left\|J/(\vec{x}_0)^T J(\vec{x}_0)\right\|_2$

   2. Subsequent steps

$$p_k = \frac{\text{actual reduction}}{\text{predicted reduction}} = \frac{f(\vec{x}_{k+1}) - f(\vec{x}_k)}{1/2\Delta\vec{x}_k^T (J(\vec{x}_k)^T r(\vec{x}_k) + \lambda_k \vec{x}_k)}$$

If p_k is large then λ_k can be decreased and if p_k is small then λ_k should be increased. Gauss Newton search direction is obtained by solving the linear system

$$\nabla^2 f\left(\vec{x}^*\right) P = -\nabla f\left(\vec{x}\right) \nabla F\left(\vec{x}\right) \nabla F(\vec{x})^T P \ = -\nabla f\left(\vec{x}\right) f(\vec{x})$$

If p_k is a large value then Gauss-Newton is good.

If $p_k \le 0$ the λ_{k+1} is increased, therefore used Levenberg-Marquardt's updating.

Levenberg-Marquardt's updated work by checking the following cases:

If $p_k > 0.75$ then $\lambda_{k+1} = \frac{\lambda_k}{3}$ and If $p_k > 0.25$ then $\lambda_k = 2\lambda_k$

Otherwise $\lambda_{k+1} = \lambda_k$. The update step should only be performed if $P_k > 0$.

6. Conclusion

We discussed some of the most popular techniques for non-linear curve fitting in this paper. The majority of processes in engineering and other sciences are either non-linear or curve linear. We have shown that we can obtain accurate and correct answers by applying the Newton, Gauss-Newton, or Levenberg Marquardt methods. A good solution is not always provided by Newton's approach. Depending on the initial guess that is manually selected, the approach may provide a variety of answers. The Gauss-Newton approach can be used in place of the Newton method when the latter fails to provide a satisfactory root. However, the Gaussian-Newton approach can be slow when used to solve problems. We can utilize the Levenberg-Marquardt approach, which is dependable and provides a decent result after a reasonable amount of iteration, to avoid these weaknesses.

7. Future Work

Non-linear curve fitting is covered in all of the several approaches this thesis presents. This subject can be regarded as crucial for further study in the study of non-linear processes. These techniques can be applied to a wide range of mathematical data in a variety of scientific fields. Numerous alternative techniques exist for fitting non-linear curves, such as Powell's method [1964] and the Nelder-Mead method [1965]. It could also be quite beneficial to compare the characteristics of the study's procedures with those of other approaches that were discussed.

References

[1] Gh. S. Hamedani. Encyclopedia Academic Mathematics. Mathematical publishing,Tehran, 2002.

[2] Igor Griva, Stephen G. Nash, Ariela Sofer. Linear and nonlinear optimization. SIAM, 2008.

[3] N.P. Bali. K.L. Sai Prasad. A textbook of engineering mathematics-III, University Science Press, 2017

[4] E.Malczewski. The Population of the United States University of Illinois (mste.illinois.edu/malcz/ExpFit/data.html), accessed on 2019-05-01.

[5] J. Dennis and R. Schnabel. Numerical Methods for Unconstrained Optimization and Non- linear Equations. SIAM, Philadelphia, 1996.

[6] H. B. Keller and V. Pereyra. Finite Dierence Solution of Two-Point BVP. Preprint 69. Department of Mathematics, University of Southern California, 1976.

[7] A. Griewank and A. Walther. Principles and Techniques of Algorithmic Differentiation Evaluating Derivatives: Other Titles in Applied Mathematics105 Second edition, SIAM, Philadelphia, 2008

[8] P. C. Hansen, V. Pereyra and G. Scherer Least Squares Data Fitting with applications. The John Hopkins University Press, Baltimore, 2013.

[9] V. Pereyra. Iterative methods for solving non-linear least squares problems. SIAM J. Numer. Anal. 4:2736, 1967.

[10] Netlib Repository at UTK and ORNL [Online] http://www.netlib.org

[11] P. C. Hansen. Inverse Problems Insight and Algorithms Discrete SIAM, Philadelphia, 2010. 46

[12] S. J. Wright, J. N. Holt. Algorithms for with linear non-linear least squares inequality constraints SIAM J. Sci. and Stat. Comput 6:10331048, 1985.

[13] J. A. Nelder, R. Mead. A Simplex method for function of optimization Computer Journal 7(4): 308-313, 1965.

[14] M. J. D. Powell. An efficient method for finding the minimum of a function of several Variables without calculating derivatives Computer Journal 7(2): 155-162, 1964.