International Journal of the Physical Sciences Vol. 7(31), pp. 5218-5223, 16 August, 2012
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS12.506
ISSN 1992 -1950 ©2012 Academic Journals

Full Length Research Paper

A new method for least squares identification of parameters of the transcendental equations

P. H. Kloppers, C. R. Kikawa* and M .Y. Shatalov

Department of Mathematics and Statistics, Tshwane University of Technology, Arcadia Campus, 175 Nelson Mandela Drive. Arcadia, Pretoria X680, 0001 South Africa.

Accepted 15 August, 2012

Transcendental models are often solved by using a different approach, which can be a derivative free, direct optimisation or iterative linearization method. All these approaches require guess values for the unknown parameters to start the iteration procedure. However, if the transcendental model involves several parameters, some of these methods become very cumbersome and computationally expensive. A new method for computing parameter estimates which are then used as initial values for the unknown model parameters to start the iteration process was proposed. Confidence intervals for the estimated parameters were constructed using the bootstrap method. We generated two randomised datasets that simulated the decay and growth processes. A three parameterized single exponential model $f(x) = \alpha \exp(\beta x) + \gamma$ was identified using the simulated datasets in each case. The absolute percentage errors were used as a measure of comparison between the proposed method and the current Levenberg-Marquardt (L-M) method. Tables and figures were used to present results from both methods. The proposed method appeared to produce better results than the current L-M method. The superiority of the proposed method over the current methods is that it does not require initial guess values and it guarantees convergences. Thus the proposed method could be adopted to solve real life problems.

Key words: Least-squares, parameter identification, transcendental models, confidence intervals, bootstrap method, initial guess values, Gaussian white noise, probability model.

INTRODUCTION

Many real physical problems can be modelled and analyzed using transcendental mathematical models. A problem of fundamental importance to the chemical engineer is the analysis of system or plant output data for the purpose of process modelling. Given output data, it may be desired to construct an adequate mathematical model of the system, often in the form of estimating parameters in the process model. For economic decision theory, the economic interactions may be described with models that take into account the geometry of the process. Economic forecasting will usually involve some nonlinearities for resource utilization (Fishburn, 1964).

Parameter estimation is an integral part of model identification and is a task which must be considered when modelling many physical situations or processes. A notion closely linked to parameter estimation is the concept of the identifiability of model parameters. There are usually two ways to consider when identifying model parameters: *Structural* or *priori* and *numerical or practical* identification (Nadja et al., 2008). Structural determines whether the parameters can be identified from a specified input-output experiment when reliable data are available. Numerical concerns the question of identifiability in the presence of real, noisy data and is essentially a problem of parameter estimation accuracy (Nadja et al., 2008). A number of methods have been developed for solving nonlinear least squares problems such as Gauss–Newton method, Levenberg–Marquardt (L-M) method, Powell's Dog Leg method, a hybrid method for L–M and Quasi–Newton, a secant version of the L–M method, and a secant version of the Dog Leg method (Maden et al., 2004). The Gauss- Newton method is the basis of the

*Corresponding author. E-mail: kikawacr@tut.ac.za or richard.kikawa@gmail.com. Tel: +27(012) 3826358. Fax: +27(012) 3826114.

very efficient methods. It is based on implemented first derivatives of the components of the vector function. In special cases it can give quadratic convergence as the Newton-method does for general optimization (Frandsen et al., 2004). Levenberg (1944) and later Marquardt (1963) suggested using a *damped* Gauss-Newton method (Maden et al., 2004).

Least squares (LS) problems are optimization problems in which the objective (error) function may be expressed as a sum of squares. All methods for non-linear optimization are iterative, requiring a starting point (initial value). This value is seldom a guess value which may not guarantee convergence, may lead to longer computation time if the guess value is far away from the optimal value and providing such a value needs high expertise. All these are shortcomings of the current methods in use. It is known that some form of trial and error is required to find initial values which can lead to rapid convergence of the current algorithms in use to the least-squares values (Marshall, 1948). Our aim is to demonstrate a method of finding the initial values to a broad range of transcendental problems, the problems which could be formulated in terms of ordinary differential equations that are linear with respect to unknown parameters. We consider two particular examples (the decay and growth models) of this formulation and solutions of the problems. Confidence intervals of the estimated parameters were computed using the percentile bootstrap technique, (Efron, 1987), a program in mathematica software was written to effect the percentile bootstrap method. A comparison of parameter estimates between our proposed method and the more robust Levenberg–Marquardt method which is implemented in mathematica was done (Tables 2 and 3 reveal the growth and decay model parameters and percentage absolute errors). For these particular examples considered, our method proved to be superior to the current L-M method in some cases and at some point the L-M method was superior in estimation of the known model parameters. We noticed that the variation in superiority of the two methods depended on the dataset being used.

PROBLEM STATEMENT

Consider a set of data points, $(x_1, y_1),...,(x_m, y_m)$ and a transcendental model in parameters, $y = f(x,\theta)$ that in addition to the variable x it also depends on n parameters, $\theta = (\theta_1,...,\theta_n)$ with $m \geq n$. It is desired to find θ of parameters such that the curve fits the given data in the least squares sense, which is the sum of squares:

$$\psi = \sum_{i=1}^{m} u_i^2 \quad (1)$$

Equation (1) is minimized, where the residuals u_i are given by;

$$u_i = y_i - f(x_i, \hat{\theta}) \text{ for } i = 1,2,...,m \quad (2)$$

In this paper, we focus on the method proposed to estimate the unknown parameters $\theta = (\theta_1,...,\theta_n)$ in the least squares sense and also construct their confidence bounds. The estimates can be used as initial guess values for the unknown parameters to start the iteration process for any of the methods in current use.

THEORETICAL BASIS OF THE METHOD

The theoretical understanding of the method is contained in two theorems (Marshall, 1948) for the Weierstrass theorem of continuous function approximation and for the linearly independent functions (Alvin and Bruce, 2008).

We consider a three parameterized transcendental model of the form,

$$f(x) = \alpha e^{\beta x} + \gamma \quad (3)$$

Where α, β and γ are the unknown parameters to be estimated.

$$f'(x) = \alpha\beta e^{\beta x} \quad (4)$$

rewriting Equation (4) we obtain

$$f'(x) = \eta e^{\beta x} \quad \text{With } \eta = \alpha\beta \quad (5)$$

$$f''(x) = \eta\beta e^{\beta x} = \beta f'(x) \quad (6)$$

Equation (6) can be rearranged as

$$f''(x) - \beta f'(x) = 0. \quad (7)$$

Integrating Equation (7) we obtain

$$\int_a^x f''(x)dx - \beta \int_a^x f'(x)dx = 0 \quad (8)$$

$$f'(x) - \beta f(x) - f'(a) + \beta f(a) = 0 \quad (9)$$

Let $\beta f(a) - f'(a) = \lambda$ in Equation (9); on rewriting it we obtain

$$f'(x) - \beta f(x) + \lambda = 0 \quad (10)$$

Integrating Equation (10),

$$\int_a^x f'(\xi)d\xi - \beta \int_a^x f(\xi)d\xi + \lambda \int_a^x d\xi = 0 \quad (11)$$

we obtain,

$$f(x) = \beta I(x) - \lambda x + f(a), \text{ with } I(x) = \int_a^x f(\xi) d\xi \quad (12)$$

Integration is performed using numerical methods.
Equation (12) can be viewed as:

$$Y = \beta X_1 - \lambda X_2 + C \quad (13)$$

Where $Y = f(x), X_1 = I(x), X_2 = x$ and $C = f(a)$; parameters β, λ and C are estimated using standard regression techniques. However, λ and C are here considered as nuisance parameters and subsequently ignored in the process.
On estimating β from Equation (13), we can now reformulate Equation (3) as a linear regression equation.

$$f(x) = \alpha X_3 + C_1 \text{ for } X_3 = e^{\beta x} \text{ and } C_1 = \gamma \quad (14)$$

Equation (14) is as well estimated as Equation (13) by standard least squares methods. From Equations (13 and 14) we are able to estimate all the parameters in the original transcendental model, Equation (3).

Hence the transcendental parameters α, β and γ have been estimated and can be used as starting values for the iteration process. However, confidence intervals of these parameters need to be constructed.

BOOTSTRAP METHOD

Bootstrap is a re-sampling technique widely used to compute accurate statistical inferences. The technique is computer based and intended to assign measures of accuracy to statistical estimates. It allows us to compute almost all statistics without knowledge of their parental distributions. It is on this basis that the bootstrap method is preferred for our work. The bootstrap method substitutes the more mathematically involving Monte Carlo techniques with an increase of several orders of magnitude in the computing needed for a statistical analysis.

Bootstrapping is performed on the linear regression models. The regression model is bootstrapped by applying two methods depending on the explanatory variables. If the explanatory variables are fixed as in the designed experiment, then the bootstrap method must retain that structure. For that case each bootstrap sample must have the same explanatory variables. On the other hand, regression models built from survey data typically have explanatory variables that are as random as the explained variable, and bootstrap samples should also possess this additional variation (Robert, 1989). In our proposed method we consider the randomised case of the explanatory variables and compare the confidence intervals for both the normal theory and the bootstrap techniques. It was prudent to consider the examples from a randomised set of data as it is more practically used in real life and more realistic; typical of survey studies.

The statistical inference of our model parameters is considered at the linearised stage of the proposed method. The linearised equations are estimated as regression models by the method of ordinary least squares (OLS).

Percentile bootstrap confidence intervals

Mathematical statisticians never talk of reliable confidence intervals, but only a reliable confidence interval method, that is one that gives intervals whose actual coverage probability is close to the normal value (DiCiccio and Efron, 1996).

With an assumption that the explanatory variables are stochastic, Equations (13 and 14) are estimated via OLS, and the residuals are stored. Then, *N* different "dummy-data" Y^* are generated in a manner coherent with the estimated model (Equation 15).

$$Y^* = \hat{\beta} X_1^* - \hat{\lambda} X_2^* + \hat{C} + U \quad (15)$$

Where $\hat{\beta}, \hat{\lambda}$ and $\hat{C}$ are OLS estimates and the *u* dummy disturbance terms are drawn randomly with replacement from the estimated residuals $\hat{\varepsilon}$. For each of the *N* generated series, a regression coefficient $\hat{\beta}$ is calculated, and the $\theta - th$ percentile of the resulting distribution is used as the $1-\theta$ percent critical value for this procedure. Then the percentile intervals of the respective estimates can be represented as:

$$[\hat{\beta}_N^{*(\theta)}, \hat{\beta}_N^{*(1-\theta)}], [\hat{\alpha}_N^{*(\theta)}, \hat{\alpha}_N^{*(1-\theta)}] \text{ and } [\hat{\gamma}_N^{*(\theta)}, \hat{\gamma}_N^{*(1-\theta)}] \quad (16)$$

These are the approximate $1-2\theta$ percentile intervals for the model parameters as presented in Equation (3).

Equation (15) has a nuisance parameter λ, but this work ignores the accuracy of the confidence intervals in their coverage aspect in the presence of nuisance parameters. However, we must state that this is rather a complicated case and ignoring the presence of the nuisance parameter is only at the programming stage when we consider only the parameter of interest for confidence interval estimation. Efron (1987) discusses confidence intervals in the presence of nuisance parameters. Robert (1989) comprehensively discusses methods of bootstrapping random and fixed explanatory variables. In the present study we only put our focus on the model with variable (stochastic) explanatory variables, for both the decay and the growth models.

PRACTICAL IMPLEMENTATION

The proposed method was implemented on two known exponential models, the growth and decay models. Using Mathematica software, two datasets were then generated to simulate both models. A randomly distributed noise (Gaussian white noise) of mean *0* and variance *5* was added to the dataset of the growth process while noise of mean *0* and variance *0.05* was added to the decay process to simulate a real world survey dataset which can be modelled by a probability model in each case, since Gaussian white noise is a good approximation of many real-world situations and generates mathematically tractable models (Athanasios, 1991). We then estimated the known parameters of Equations (17 and 17a) using the simulated dataset from both the proposed method and the already existing, but more robust Levenberg-Marquardt method. This was done to compare the results (estimates) from both the proposed and the L-M methods to establish which of the two gives better results or estimates close to the exact (known) parameters. We showed examples of the problems for the simple cases of the growth model, Equation (17) and the decay model, Equation (17a).

.

$$f(x) = 0.2\exp(1.1x) + 15 \quad (17)$$

Table 1. Exact and approximate central 95% confidence intervals for the estimated model parameters n =1000.

Model	for $\hat{\alpha}$	for $\hat{\beta}$	for $\hat{\gamma}$
Exact (normal theory)	[0.205065; 0.205351] $\hat{\alpha}$ =0.205208	[1.09558; 1.09558] $\hat{\beta}$ =1.09558	[14.90234; 14.96398] $\hat{\gamma}$ =14.93316
Percentile bootstrap	[0.0675918; 0.233158] $\hat{\alpha}$ =0.123389	[1.08202; 1.13606] $\hat{\beta}$ =1.12028	[15.1519; 19.6884] $\hat{\gamma}$ =17.4967

$$f(x) = 0.2\exp(-1.1x) + 15 \quad (17a)$$

Equations (17) and (17a) are estimated according to the proposed method as the systematic procedure from Equation (3) through Equation (14) is followed. Equations (18) and (18a) shows the estimated models (growth and decay) respectively, using our proposed method:

$$f(x) = 0.177825\exp(1.11473x) + 16.4341 \quad (18)$$

$$f(x) = 0.225935\exp(-2.23258x) + 15.0108 \quad (18a)$$

The confidence intervals of the estimated parameters are computed using the bootstrap method discussed in "Percentile bootstrap confidence intervals".

RESULTS AND DISCUSSION

We presented results of the estimated model parameters and the corresponding confidence intervals derived by the non parametric bootstrap percentile method and normal theory assumptions.

Table 1 shows the parameter estimates and their respective confidence intervals for the growth model (Equation 17). Only the growth model was presented since the same method was followed to compute results for the decay model (Equation 17a). The estimated parameters were computed using the proposed method under normal theory assumptions (OLS) and non parametric bootstrap techniques in each case. Confidence intervals of the individual estimates were computed for each method (normal theory and bootstrap) (Table 1). Bootstrapping was done 1000 times implying that the model was estimated on n=1000 data points.

Comparison of proposed and Levenberg-Marquardt methods

We compared the parameter estimates computed from our proposed method and those computed using the existing Levenberg-Marquardt method with the known (exact) model parameters. The Levenberg-Marquardt method is a standard technique used to solve nonlinear least squares problems. Least squares problems arise when fitting a parameterised function to a set of measured data points by minimising the sum of the squares of the errors between the data points and the function. Nonlinear least squares problems arise when the function is not linear in parameters. Non linear least squares methods involve an iterative improvement of the parameter values in order to reduce the sum of the squares of the errors between the function and the measured data points. The Levenberg-Marquardt curve fitting method is actually a combination of two minimisation methods: the gradient descent and the Gauss-Newton method (Henri, 2011). The "exact" parameter values P_{Exact}, the estimated parameter values using the proposed method $P_{\text{prop-method}}$, the estimated parameters values using the Levenberg-Marquardt (L-M) method $P_{\text{L-M method}}$, the absolute error between the "exact" values and estimates by the L-M method $P_{\text{Exact}} - P_{\text{L-M method}}$, and the absolute error between the "exact" values and our proposed method P_{Exact}-$P_{\text{prop-method}}$, were all shown in Table 2. The percentage absolute errors are in parentheses in each case. All computations were done using mathematica software. Table 2, shows results of the growth model (Equation 17) and fitted curves to compare the performance of the two methods in estimating the original model is shown (Figure 1). We observed that our proposed method outperformed the Levenberg-Marquardt method as seen from the absolute percentage errors in parentheses (Table 2). We compared how each of the methods performed on estimating each of the parameters by computing the absolute percentage error in parentheses. We preferred the Levenberg-Marquardt method for our comparison due to its robustness and being a combination of two minimisation methods: the gradient descent and the Gauss-Newton method.

From Figure 1, it is almost impossible to distinguish between the "exact" curve from the fitted curves of both the proposed and Levenberg-Marquardt methods. However, we observed that the fitted (green) curve using the proposed method appeared to be closer to the "exact" (red) curve than the fitted (blue) curve using the L-M method. This was justified by the absolute percentage errors in parentheses (Table 2).

Table 3 shows parameter estimates of the decay model (Equation 17a). The computational methods were the same as those for the growth model and the results

Table 2. Parameter values of the proposed and L-M methods and their respective absolute and percentage errors (values in parentheses) for the growth model. n=30.

Parameter	P_{Exact}	$P_{prop\text{-}method}$	$P_{L\text{-}M\ method}$	Absolute error (P_{Exact} - $P_{L\text{-}M\ method}$)	Absolute error (P_{Exact}-$P_{prop\text{-}method}$)
α	0.2	0.177825	0.156527	0.0434729 (21.7)	0.0221752 (11.1)
β	1.1	1.11473	1.13597	0.0359724 (3.3)	0.0147301 (1.3)
γ	15	16.4341	17.0395	2.03955 (13.6)	1.43406 (9.7)

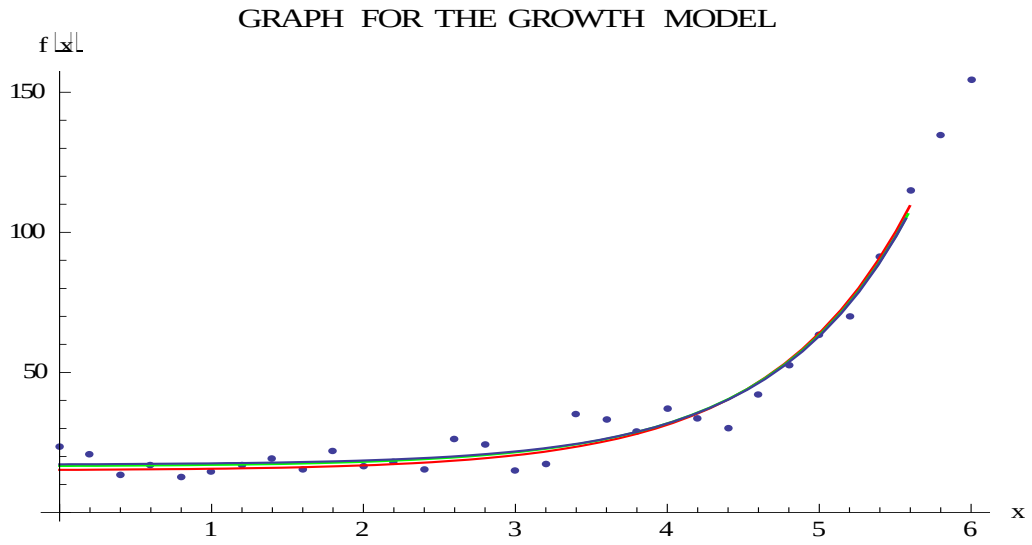

Figure 1. Comparison of the fitted curves between the proposed (green) and L-M (blue) methods with the "exact" (red) curve.

Table 3. Parameter values of the proposed and L-M methods and their respective absolute and percentage errors (values in parentheses) for the decay model (n=30).

Parameter	P_{Exact}	$P_{prop\text{-}method}$	$P_{L\text{-}M\ method}$	Absolute error (P_{Exact} - $P_{L\text{-}M\ method}$)	Absolute error (P_{Exact}-$P_{prop\text{-}method}$)
α	0.2	0.22594	0.226496	0.026496 (13.2)	0.025935 (13.0)
β	-1.1	-2.23258	-2.60767	1.50767 (137.1)	1.13258 (102.1)
γ	15	15.0108	15.0127	0.0126603 (0.08)	0.0108018 (0.07)

were compared with those of the Levenberg-Marquardt method still the same way it was done for the growth model (Table 2). The figures in parentheses are the absolute percentage errors due to the proposed and Levenberg-Marquardt methods on estimating the "exact" parameter, $P_{Exact.}$ The proposed method appeared to produce better estimates than the existing Levenberg-Marquardt method. The superiority of the proposed method over the L-M method was supported by the magnitude of the absolute percentage errors between the two methods on estimating the "exact" individual parameters (Table 3 show values in parentheses).

The decay graph model (Figure 2) was constructed to compare the fitted curves from both methods with the "exact" curve. It was assumed that the "exact" (red) curve was the best fit to the randomised dataset of the simulated decay process while the green and blue curves represented fitted curves using the proposed and the existing Levenberg-Marquardt methods respectively. We observed that the green curve was closer to the red "exact" curve than the blue curve throughout the entire domain of approximation. This clearly showed that the proposed method was more superior to the existing L-M method for this particular problem.

Conclusion

When the parameter values are estimated their confidence intervals can be computed using the non-

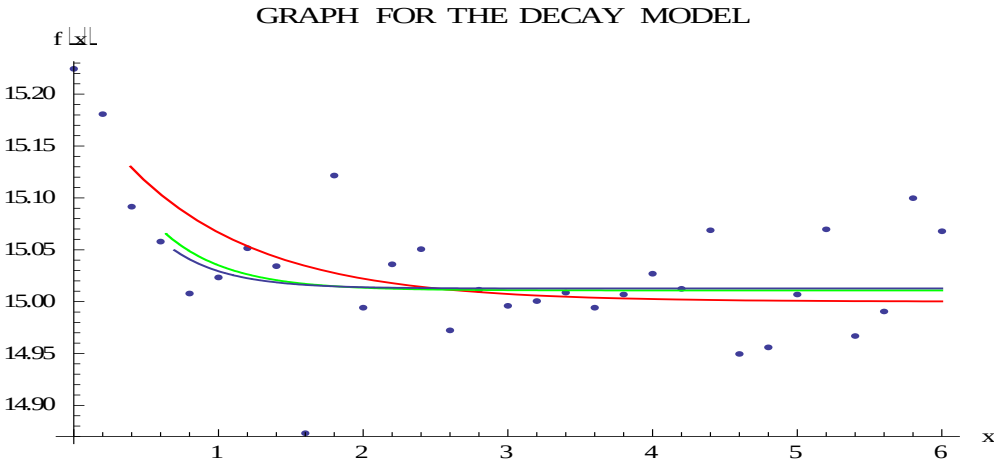

Figure 2. Comparison of the fitted curves between the proposed (green) and L-M (blue) methods with the "exact" (red) curve.

parametric bootstrap percentile method (Table 1). The bootstrap method was preferred, because it does not require asymptotic assumptions of the underlying data set. Since these confidence intervals are not based on asymptotic theory, they may provide a good guide in any actual case.

We demonstrated that the proposed method performed better than the current Levenberg-Marquardt method for this particular model. Firstly, the superiority of the proposed method was justified by the relatively smaller magnitudes of the absolute percentage errors (Tables 2 and 3 reveals the growth and decay models, respectively; values in parentheses). Secondly, looking at the fitted curves for both processes (Figures 1 and 2), it was observed that the green curve, fitted using our proposed method was closer to the "exact" red curve than the blue curve, fitted using the existing Levenberg-Marquardt method. This implied that the proposed method computed estimates that were more less the same as the known "exact" model values (Equations 17 and 17a) than the existing Levenberg-Marquardt method for these particular problems.

The method appeared to work well for the single exponential model studied in this paper, however more research is required that may probably be extended to more complicated models which can be reformulated in terms of ordinary differential equations that are linear with respect to the unknown parameters. The proposed method has the advantage over the current methods of not requiring initial guess values. The current methods require initial guess values for the parameters to start the iteration process. These initial values may not guarantee convergence, may lead to long computation time if the provided guess is far from the optimal value and providing a "good" guess requires high experience.

Thus, we believe that the technique proposed in this paper can be adopted and will often be useful in real practice.

REFERENCES

Alvin CR, Bruce GS (2008). Linear Models in Statistics, 2nd ed. New Jersey: John Wiley and Sons.
Athanasios P (1991). Probability, Random Variables, and Stochastic Processes, 3rd ed. WCB/McGraw-Hill
DiCiccio TJ, Efron B (1996). Bootstrap Confidence Intervals. J. Statist. Sci. 11(3):189-228.
Efron B (1987). Better Bootstrap Confidence Intervals. J. Am. Statist. Assoc. 82(397):171-185.
Fishburn PC (1964). Decision and Value Theory, John Wiley & Sons, New York.
Frandsen PE, Jonasson K, Nielsen HB, Tingleff O (2004). Unconstrained Optimization. 3rd ed. IMM, Technical University of Denmark.
Henri G (2011). The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. Department of Civil and Environmental Engineering, Duke University.
Levenberg K (1944). A Method for the Solution of Certain Non-linear Problems in Least Squares. Quart. Appl. Math. 2:164-168
Maden K, Nielsen HB, Tingleff O (2004). Methods for nonlinear least squares, 3rd ed. Informatics and Mathematical Modelling: Technical University of Denmark.
Marquardt D (1963). An Algorithm for Least Squares Estimation on Nonlinear. Parameters. Siam J. Appl. Math. 11:431–441.
Marshall H, Stone HS (1948). The generalised Weierstrass Approximation Theorem. Math. Magazine 21(5):237-254.
Nadja H, Mario Z, Stanko S (2008). Nonlinear model parameter estimation-estimating a feasible parameter set with respect to model use, Math. Comput. Model L. Dynam. Syst. 14(6):587-605.
Robert S (1989). An Introduction to Bootstrap Methods, Examples and Ideas. University of Pennsylvania.