International Journal of Applied Research 2024; 10(5): 184-188

ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor (RJIF): 8.4
IJAR 2024; 10(5): 184-188
www.allresearchjournal.com
Received: 15-02-2024
Accepted: 20-03-2024

Mohamed Jaber
Department of Statistics, College of Science, University of Misurata, Libya

Mohamed Muftah
Department of Mathematics, College Arts and Science, Al Abayar, University of Benghazi, Libya

Yusriyah Hamad
Department of Statistics, College of Arts and Science, Al Maraj, University of Benghazi, Libya

Farag Hamad
Department of Statistics, College of Arts and Science, Al Maraj, University of Benghazi, Libya

Corresponding Author:
Mohamed Jaber
Department of Statistics, College of Science, University of Misurata, Libya

Separable nonlinear least squares for estimating nonlinear regression model

Mohamed Jaber, Mohamed Muftah, Yusriyah Hamad and Farag Hamad

DOI: https://doi.org/10.22271/allresearch.2024.v10.i5c.11750

Abstract
Regression analysis is a statistical technique used to examine the relationship between (dependent and independent) variables. Regression analysis is typically used by academics to examine the impact of several independent factors, or explanatory variables, on a single variable, or response variable. The regression equation is used by the investigators to explain how the response and explanatory variables relate to one another. We need to meet many assumptions to estimate the relationship (model). Several techniques, including the ordinary least squares (OLS), and maximum likelihood approach (MLE) can be used to estimate the parametric regression model. Moreover, the Spline or Kernel methods can be used for estimating nonparametric regression. In this work, we attempt to demonstrate the significant and practical method for estimating the nonlinear model. Separable nonlinear least squares (SNLS) method is a special case of nonlinear least squares (NLS) method, for which the objective function is a mixture of linear and nonlinear functions. In this technique, the nonlinear function (model) can be linearized by applying special transformation or by using expanded Taylor expansion to linearize functions. The separable nonlinear least squares (SNLS) are a very flexible technique that is used to linearize the nonlinear functions. The SNLS can be used after linearizing the nonlinear function through the transformation of the variable of interest. Moreover, the SNLS can be used to approximate a wide variety of functional shapes. The results show that the SNLS performed very well in comparison with the NLS. We can observe from the model goodness residuals standard error, AIC, and BIC, that the SNLS method has provided an estimate equivalent to that NLS provided. Therefore, we can say that it is useful to estimate nonlinear model separable. Furthermore, we plan to apply the SNLS to a more complex model using different simulation studies to check the validity of the method.

Keywords: Linear, nonlinear, separable nonlinear least squares, Taylor expansion, gauss-newton method

Introduction
The statistical tool used to investigate the relationship between (dependent and independent) variables is called regression analysis. Usually, researchers use regression analysis to analyze the effect of some independent variables (explanatory variables) on one variable (response variable) [1, 2]. The investigators use the regression equation to describe the relationship between the response and explanatory variables. The regression model includes one or more hypothesized regression unknown parameters [3]. The regression model can be estimated using several methods such as ordinary least squares (OLS) and maximum likelihood method (MLE) for the parametric regression model [4]. Nonparametric regression can be estimated by using the Kernel method or Spline method [5].

The most well-known and classic estimators for regression coefficients are the ordinary least squares (OLS) estimators obtained by minimizing the sum of squared residuals [6]. The least squares method needs the error to be assumed as independent and identically distributed with mean zero and constant variance (the normality assumption). Under the Gauss-Markov theorem, the estimated parameters are the best linear unbiased estimators (BLUE). In practice, there are many problems caused when the assumptions are violated, e.g. non-normality, heteroscedasticity, and of particular interest is the nonlinear independence of regresses (independent variables) [7].

Unlike the linear regression model, the nonlinear regression model is not restricted to belonging to a specific relation. To apply the linear regression model, we need to satisfy several assumptions such as linearity, no multicollinearity, and normality. While to apply the nonlinear regression, we need a model to fit with data and an initial guess to start estimating the model parameters. The Gauss-Newton method (GNM) is the most popular method that is use to fit the nonlinear regression model [8].

When the relationships in data are not linear, additional flexibility is needed to apply the traditional approach of the Nonlinear least squares. However, recent advances in statistical techniques help to analyze data where questions of nonlinearity arise [9]. Smoothing splines and semi-parametric regression, which allow more flexibility than the nonlinear regression models, are examples of these techniques. This work will provide a review of nonlinear least squares and separable nonlinear least squares with applications of these techniques. Separable nonlinear least squares (SNLS) method is a special case of nonlinear least squares (NLS) method, for which the objective function is a mixture of linear and nonlinear functions [10].

Moreover, in this work, we will introduce basic concepts of the separable nonlinear least square technique, where the basic idea will be clarified by an example. Whereas, the separable least squares regression is concerned with the flexible incorporation of nonlinear functional relationships in regression analyses. It has many applications in many different areas, especially in operations research, and industry engineering [11]. Furthermore, the separable nonlinear least square can be used in many practical cases such as fuzz regression model [12, 13]. In details, the advantages and disadvantages of these advanced regression techniques will be evaluated and discussed for a partly linear regression model and partly nonlinear regression model.

Methodology

The nonlinear models are used to describe a more complicated relationship between the response and explanatory variable. Moreover, when the relationships between the response and explanatory are not linear relation more flexibility is needed to apply the traditional approach of the Nonlinear least squares. Unlike the linear regression model, the parameters may not linear function in the predictors. Therefore, the assumptions of applying ordinary least squares are violated [14]. The basic form for a nonlinear model between the response y and a predictor x is given as,

$$Y = f(X;\ \theta) + \varepsilon$$

where, f is a nonlinear function involving the predictor and the parameter vector θ, relating E(Y), and θ are the vector of P parameters [15]. Also, the error term is assumed to have the same properties as in the linear regression models. In the nonlinear regression models, at least one of the derivatives of the expectation function f with respect to the parameters will have at least one of the parameters [2].

Estimation of Nonlinear Regression model

Nonlinear Least Squares Estimates (NLS)

The least squares method is used to estimate the parameters of the nonlinear models. To estimate the parameters using nonlinear least squares, like in linear least squares case. The nonlinear regression model, given by,

$$Y = f(X;\ \theta) + \varepsilon$$

Where

$Y = (y_1, y_2, \ldots, y_n)^T, f(X;\theta) = (f(X_1,\theta), f(X_2,\theta), \ldots, f(X_n,\theta))^T,\quad X = (X_1, X_2, \ldots, X_p)^T, \theta = (\theta_1, \theta_2, \ldots, \theta_p)^T$, and $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T;\ \varepsilon \sim NID(0, \sigma^2 I)$

The least squares estimate of θ, labeled by $\hat{\theta}$, is the choice of parameters that minimizes the sum of squared residuals

$$ss(\theta) = \sum_{i=1}^{n} [Y_i - f(X_i;\theta)]^2, i = 1,2, \ldots n$$

Or, it can be written as:

$$ss(\theta) = \varepsilon^T\varepsilon = [Y - f(X;\theta)]^T[Y - f(X;\theta)]$$

The partial derivatives of $ss(\theta)$, with respect to each θ_j in turn, set equal to zero to obtain the p normal equations [16]. Each normal equation has the general form

$$\frac{\partial}{\partial\theta_j}[ss(\theta)]_{\theta=\hat{\theta}} = 0$$

$$\sum_{i=1}^{n} [Y_i - f(X_i;\theta)][\frac{\partial}{\partial\theta_j} f(X_i;\theta)]_{\theta=\hat{\theta}} = 0, j = 1,2, \ldots, p$$

Where, $\frac{\partial}{\partial\theta_j} f(X_i;\theta)$ the partial derivatives of a nonlinear model are functions of the parameters. A major difficulty with nonlinear least squares arises in trying to solve the normal equations for $\hat{\theta}$, since clear solutions cannot be obtained, iterative numerical methods are used. These methods require initial guesses, or starting values. For the starting value parameters are labeled as θ^0 and find $\theta^1, \theta^2, \ldots$ until we obtain a sufficiently small adjustment being made at each step, when this happens, the process is said to have converged to a solution [17].

The Gauss-Newton method (GNM) uses a linearization based on a Taylor expansion in the parameter space to estimate parameter values. For the Taylor expansion of $f(X_i;\theta)$ around the starting value θ^0, to obtain a linear approximation of the model in the region nears the starting values. If θ is close to θ^0, the following approximation holds:

$$f(X_i;\ \theta) \approx f(X_i;\ \theta^0) + \sum_{j=0}^{p} \left[\frac{\partial}{\partial\theta_j} f(X_i;\ \theta^j)\right]_{\theta=\theta^0} (\theta_j - \theta_j^0)$$

Where, $\theta = \left(\theta_1, \theta_2, \ldots, \theta_p\right)^T,\ \theta^0 = \left(\theta_0^0, \theta_1^0, \ldots, \theta_p^0\right)^T, f_i^0 = f(X_I;\ \theta^0), \beta_j^0 = \left(\theta_j - \theta_j^0\right)$, and $F_{ij}^0 = \left[\frac{\partial}{\partial\theta_j} f(X_i;\ \theta)\right]_{\theta=\theta^0}$

Therefore, the following equation can be written as

$$f(X_i;\ \theta) = f_i^0 + \sum_{j=0}^{p} F_{ij}^0\, \beta_j^0$$

Also, the nonlinear model can be written as

$$Y_i^0 = \sum_{j=0}^{p} F_{ij}^0\, \beta_j^0 + \varepsilon_i, \text{where } Y_i^0 = Y_i - f_i^0$$

Which is of the similar form of the multiple linear regression model. Using the matrix notation can be written as

$$Y^0 = F^0\beta^0 + \varepsilon$$

Where,

$$Y^0 = \begin{bmatrix} Y_1^0 \\ Y_2^0 \\ \vdots \\ Y_n^0 \end{bmatrix}, F^0 = \begin{bmatrix} F_{10}^0 & F_{11}^0 & \cdots & F_{1p}^0 \\ F_{20}^0 & F_{21}^0 & \cdots & F_{2p}^0 \\ \vdots & \vdots & \ddots & \vdots \\ F_{n0}^0 & F_{n1}^0 & \cdots & F_{np}^0 \end{bmatrix}, \beta^0 = \begin{bmatrix} \beta_0^0 \\ \beta_1^0 \\ \vdots \\ \beta_p^0 \end{bmatrix}, \text{and } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

The least squares estimate of the parameters β^0 are obtained as

$$\delta^0 = \left(F^{0^T} F^0\right)^{-1} F^0 Y^0$$

New values of the parameters are obtained by adding the estimated shift to the initial values using Gauss-Newton increment by

$$\theta^1 = \theta^0 + \delta^0$$
$$\theta^2 = \theta^1 + \delta^1$$
$$\vdots$$
$$\theta^a = \theta^{a-1} + \delta^{a-1}$$

Where, δ^a is called the Gauss-Newton increment. The model is then linearized about the new values of the parameters and linear least squares are again applied to find the second set of adjustments, and so forth until the desired degree of convergence is attained [18]. The adjustments obtained from the Gauss-Newton method can be too large and bypass the solution, in which case the residual sum of squares may increase at that step rather than decrease. Moreover, the Gauss-Newton algorithm will converge only with a good initial guess when the matrix F^0 is a full rank matrix in a neighborhood of the least square's solution [19]. Otherwise, there is no guarantee that the Gauss-Newton algorithm will converge.
In practice, the previous technique can be used only when the function $f(X_i;\ \theta)$ is continuously and differentiable respect to the parameter θ. If the function $f(X_i;\ \theta)$ is not continuous and differentiable with respect to the parameter θ, it is usually necessary to modify the model or use another technique to estimate the nonlinear model. To apply the Gauss-Newton Algorithm for the nonlinear model, we need to find the Taylor expansion for the model.

$$f(X_i, \theta) = \theta_1\left(1 - \theta_2 e^{-\theta_3 X}\right) + \varepsilon_i$$

Gauss-Newton Algorithm for (NLS)
Taylor expansion for $f(X_i;\ \theta)$ is

$$f(X_i;\ \theta) \approx f(X_i;\ \theta^0) + \sum_{j=0}^{3} \left[\frac{\partial}{\partial\theta_j} f(X_i;\ \theta^j)\right]_{\theta=\theta^0} (\theta_j - \theta_j^0)$$

$$f(X_i;\ \theta) \approx f(X_i;\ \theta^0) + \frac{\partial}{\partial\theta_1}\left[\theta_1\left(1 - \theta_2 e^{-\theta_3 X}\right)\right](\theta_1 - \theta_1^0)$$

$$+\frac{\partial}{\partial\theta_2}\left[\theta_1(1 - \theta_2 e^{-\theta_3 X})\right](\theta_2 - \theta_2^0) + \frac{\partial}{\partial\theta_3}\left[\theta_1(1 - \theta_2 e^{-\theta_3 X})\right](\theta_3 - \theta_3^0)$$

$$f(X_i;\ \theta) \approx f(X_i;\ \theta^0) + \left(1 - \theta_2 e^{-\theta_3 X}\right)(\theta_1 - \theta_1^0) - \theta_1 e^{-\theta_3 X}(\theta_2 - \theta_2^0) + \theta_1\theta_2 X e^{-\theta_3 X}\,(\theta_3 - \theta_3^0)$$

Initial guesses or starting values are required for estimating the model parameters using Gauss-Newton algorithm. Moreover, we can inference about the model parameters by applying several assumptions around the estimated parameters $\hat{\theta}$ [20].

1. The estimated parameters $\hat{\theta}$ has approximate normal distribution with approximate mean θ and approximate covariance matrix $\sigma^2(F'F)^{-1}$.
2. An approximate $(1-\alpha)100\%$ joint confidence region for θ, which is an ellipsoid is given by:
3. $(\theta - \hat{\theta})'\hat{F}'\hat{F}(\theta - \hat{\theta}) \le ps^2 F_{(p,n-p,\alpha)}$
4. An approximate $(1-\alpha)100\%$ marginal confidence interval for θ_i is
5. $\hat{\theta}_i \pm t_{\left(n-p,\frac{\alpha}{2}\right)} se(\hat{\theta})$
6. An approximate $(1-\alpha)100\%$ confidence interval for the expected response variable at X_0 is
7. $f\left(\hat{\theta}, X_0\right) \pm t_{\left(n-p,\frac{\alpha}{2}\right)} s\sqrt{V_0'(\hat{F}'\hat{F})^{-1}V_0}$
8. An approximate $(1-\alpha)100\%$ confidence interval for the predicted mean of the response variable at X_0 is:

$$f\left(\hat{\theta}, X_0\right) \pm t_{\left(n-p,\frac{\alpha}{2}\right)} s\sqrt{1 + V_0'(\hat{F}'\hat{F})^{-1}V_0}\ ;\ V_0 = \left[\frac{\partial f\left(\hat{\theta}, X_0\right)}{\partial\theta}\right]_{\theta=\theta^0}$$

Separable of Nonlinear Least Squares (SNLS)
In the separable least squares, the objective function is a mixture of two components (linear and nonlinear functions) [21]. The separable of nonlinear least squares is a special case of nonlinear least squares in which the function can be derived into two parts [22]. The method can be used in many applications such as numerical analysis, neural networks, and Environmental Sciences. However, the SNLS is an invalid method when there are some constrains on the linear part of variables [23, 24]. Here, we proposed SNLS to solve a function that was solved by NSL.

$$f(X_i, \theta) = \theta_1\left(1 - \theta_2 e^{-\theta_3 X}\right) + \varepsilon_i = \theta_1 - \theta_1\theta_2 e^{-\theta_3 X} = \beta_0 + \beta_1\tilde{X} + \varepsilon_i,$$

$$\beta_0 = \theta_1, \beta_1 = -\theta_1\theta_2, \text{and } \tilde{X} = e^{-\theta_3 X}$$

Taylor expansion for $f\left(\tilde{X}_i;\ \theta_3\right)$ is

$$f\left(\tilde{X}_i;\ \theta_3\right) \approx f\left(\tilde{X}_i;\ \theta_3^0\right) + \sum_{j=1}^{\infty} \frac{(-\theta_3 X)^j}{j!}$$

$$f(\widetilde{X}_i;\ \theta_3) \approx\ f(\widetilde{X}_i;\ \theta_3^0) - \theta_3\widetilde{X} + \frac{(\theta_3\widetilde{X})^2}{2!} - \frac{(\theta_3\widetilde{X})^3}{3!} + \cdots$$

For any given value for θ_3, the θ_1 and θ_2 can be estimated by applying linear least squares method as:

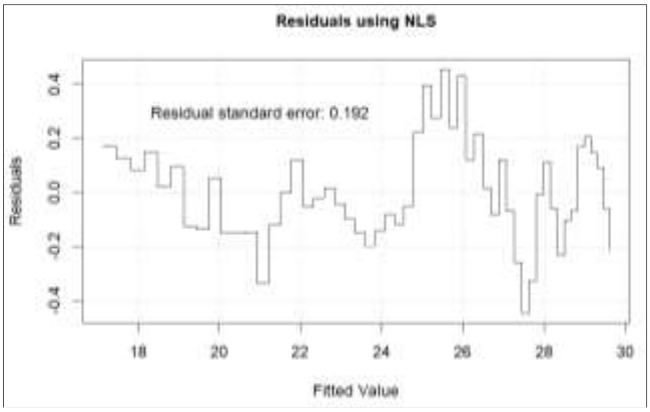
Estimating Nonlinear Model

The study implemented to comparison between the estimated model using nonlinear least squares (Gauss-Newton algorithm) and separable nonlinear least squares. The study was carried out to estimate the nonlinear model parameters $(\theta_1, \theta_2, \text{and}\ \theta_3)$. We used RStudio to generate the estimated parameters of the $f(X_i, \theta) = \theta_1(1 - \theta_2 e^{-\theta_3 X}) + \varepsilon_i$ using data science about chloride ion transport through blood cell walls the data set includes two factors (y donates the chloride concentration (in percent) and x donates to the time (in minutes)). For more details, see [25]. The review study was performed to compare between the nonlinear least squares and separable nonlinear least square. In this study, we give a short application on the separable nonlinear least square method unseparated scheme for NLS. The results, of estimating nonlinear model using NLS and SNLS are demonstrated in Table 1:

Table 1: Estimated model parameters using nonlinear least squares and separable nonlinear least squares

Parameters	NLS				SNLS			
	Estimate	**St. Error**	**T**	**P-value**	**Estimate**	**St. Error**	**T**	**P-value**
θ_1	39.09	0.974	40.12	<2e-16***	28.835	1.0952	26.328	<2e-16***
θ_2	0.828	0.008	99.80	<2e-16***	0.638	0.1132	5.634	7.64e-07***
θ_3	0.158	0.010	15.18	<2e-16***	0.227	1.5085	-15.113	< 2e-16***
	Residuals standard error=1.92 AIC=-20.09 & BIC=-12.13				Residuals standard error=1.95 AIC=-18.12 & BIC=-10.17			

From the results, we can see that the estimated model using SNLS is comparable with the estimated model using NLS. Moreover, based on the model goodness of fits both methods NLS and SNLS performed well with the data. The estimated model using SNLS still performed well even though the estimated value of the parameter θ_1 (28.835) was slightly different from the estimated value of θ_1 using NLS (39.09).



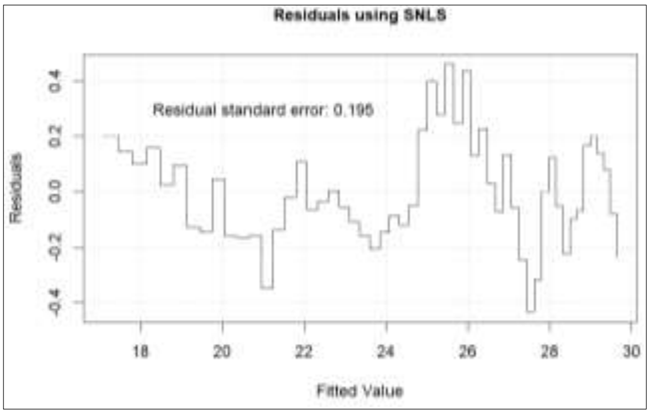


Fig 1: The typical residuals of nonlinear regression model with estimated model using NLS and SNLS

The above plots, show the mathematical function that explains the relationship between the dependent variable y and the response variable x throughout the nonlinear relationship. It can be observed from the above figures, that the estimated model using SNLS is close to the estimated model using NLS. However, the relationship looks linear which can be easily estimated by OLS but the linearity assumption for the model parameters is violated.

Conclusion

The statistical tool used to investigate the relationship between (dependent and independent) variables is called regression analysis. Usually, researchers use regression analysis to analyze the effect of some independent variables (explanatory variables) on one variable (response variable. The investigators use the regression equation to describe the relationship between the response and explanatory variables. The relationship might be linear and might be a nonlinear relationship. To estimate the relationship (model), we need to satisfy several assumptions. The parametric regression model can be estimated using several methods such as ordinary least squares (OLS) and maximum likelihood method (MLE). While the nonparametric regression can be estimated by using the Kernel method or Spline method.

In this work, we try to show the important and useful technique for estimating the nonlinear model. In this technique, the nonlinear function (model) can be linearized by applying special transformation or by expanding using Taylor expansion to linearize functions. The separable nonlinear least squares are a very flexible technique that used to linearize the nonlinear functions. The SNLS can be used after linearizing the nonlinear function through the transformation of the variable of interest and the explanatory variables. Moreover, the SNLS can be used to approximate a wide variety of functional shapes. The results show that the SNLS performed very well in comparison with the NLS. We can observe from the model goodness residuals standard error, AIC, and BIC that the SNLS method has provided an estimate equivalent to that NLS provided. Therefore, we can say that it is useful to estimate nonlinear model separable.

Moreover, we plan to apply the SNLS to a more complex model using different simulation studies to check the validity of the method.

References

1. Sykes AO. An introduction to regression analysis; c1993.
2. Seber GA, Lee AJ. Polynomial regression. Linear Regres Anal; c2003. p. 165-185.
3. Weisberg S. Applied linear regression. Vol. 528. John Wiley & Sons; c2005.
4. Bro R, Sidiropoulos ND, Smilde AK. Maximum likelihood fitting using ordinary least squares algorithms. J Chemom. 2002;16(8-10):387-400.
5. Opsomer JD, Breidt FJ. Nonparametric Regression Using Kernel and Spline Methods; c2011.
6. Jaber M. Comparisons between Nonlinear Regression Models and Semi-parametric Regression Models; c2009 Jul 20.
7. Thinh R. A Comparison of Linear Regression Models for Heteroscedastic and Non-Normal Data; c2018.
8. Turner H, Firth D. Generalized nonlinear models in R: An overview of the GNM package; c2007.
9. Hamad F, Kachouie NN. A hybrid method to estimate the full parametric hazard model. Commun Stat-Theor Methods. 2019;48(22):5477-5491.
10. Sherrod PH. Nonlinear regression analysis program. Nashv. TN USA; c2005.
11. Eissa TEA. Effect of Industry Clusters and Commodity Flow Characteristics on the Selection of Freight Transportation Modes in the United States with Advanced Econometric Approaches; c2019.
12. Omar E, Jaber M, Hamad F, Kyamo M, Eissa T. Fuzzy Based Clustering Algorithm For Manets. Data Acquis Process. 2023;38:5416.
13. Omara E, *et al*. An Improved Fuzzy Based Clustering Algorithm for MANETs. Scand J Inf. Syst. 2022;34(2):33-38.
14. Hamad F, Younus N, Muftah MM, Jaber M. Specify underlining distribution for clustering linearly separable data: Normal and uniform distribution case. J Data Acquis Process. 2023;38:4675.
15. Walkowiak R, Kala R. Two-phase nonlinear regression with smooth transition. Commun Stat-Simul Comput. 2000;29(2):385-397.
16. Argyros IK, Hilout S. On the Gauss-Newton method. J Appl Math Comput. 2011;35:537-550.
17. Jaber M, Hamad F, Breininger RD, Kachouie NN. An Enhanced Spatial Capture Model for Population Analysis Using Unidentified Counts through Camera Encounters. Axioms. 2023;12(12):1094.
18. Milliken GA. Nonlinear regression analysis and its applications. 1990.
19. Jaber M. A Spatiotemporal Bayesian Model for Population Analysis; c2022.
20. Bates DM, Watts DG. Nonlinear regression analysis and its applications. Wiley Ser. Probab. Stat; c1988.
21. Smyth GK. Nonlinear regression. Encycl Environmetrics. 2002;3:1405-1411.
22. Gharibi W, Al-Mushayt OS. A note on separable nonlinear least squares problem. Presented at: 2011 International Conference on Future Computer Sciences and Application, IEEE; c2011. p. 54-56.
23. Ng MP, Grunwald GK. Nonlinear regression analysis of the joint-regression model. Biometrics; c1997. p. 1366-1372.
24. Baty F, Ritz C, Charles S, Brutsche M, Flandrois JP, Delignette-Muller ML. A toolbox for nonlinear regression in R: The package nlstools. J Stat Softw. 2015;66:11-21.
25. Ruckstuhl A. Introduction to nonlinear regression. IDP Inst Datenanalyse Prozessdesign Zurcher Hochsch Angew Wiss; c2010, 365.