International Journal of Engineering & Technology, 7 (4.10) (2018) 543-546

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

A Monograph on Nonlinear Regression Models

B. Mahaboob [1], B.Venkateswarlu[2]*, C. Narayana [3] , J. Ravi sankar [4] and P. Balasiddamuni [5]

[1]*Department of Mathematics, K.L.E.F(Deemed to be University), Vaddeshwaram, Vijayawada, Andhra Pradesh.*
[2,4]*Department of Mathematics, VIT University, Vellore, Tamilnadu.*
[3]*Department of Mathematics, Sri Harsha institute of P.G Studies, Nellore.*
[5]*Rtd. Professor, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh.*
**Corresponding author E-mail: venkatesh.reddy@vit.ac.in*

Abstract

This research article uses Matrix Calculus techniques to study least squares application of nonlinear regression model, sampling distributions of nonlinear least squares estimators of regression parametric vector and error variance and testing of general nonlinear hypothesis on parameters of nonlinear regression model. Arthipova Irina et.al [1], in this paper, discussed some examples of different nonlinear models and the application of OLS (Ordinary Least Squares). MA Tabati et.al (2), proposed a robust alternative technique to OLS nonlinear regression method which provide accurate parameter estimates when outliers and/or influential observations are present. Xu Zheng et.al [3] presented new parametric tests for heteroscedasticity in nonlinear and nonparametric models.

Keywords: *Nonlinear regression model, Residual form of Squares, error variance, Least Squares Estimator, Parametric vector, variance-Covariance matrix, OLS.*

1. Introduction

This One of the most common situations of statistical analysis in Business, Economics, Engineering, Physical Sciences is likely to lead to a nonlinear regression model. Generally, the models to be obtained as solutions of differential equations arising in Physical Sciences, Engineering, Ecology and Business Economics are some examples of nonlinear models. Several Nonlinear Regression models may be classified into three types namely, Univariate Nonlinear Regression Models, Multivariate Nonlinear Regression Models and Nonlinear Simultaneous Equations Regression Models. Among these nonlinear regression models, the Univariate Nonlinear Regression Models have a wide number of applications in practice than other types of nonlinear regression models.

2. Least Squares Application Of Nonlinear Regression Model

For a sample of 'n' observations, consider a univariate nonlinear regression model with an additive error as

$$Y_i = f(X_i, \beta) + \varepsilon_i, \qquad i = 1, 2, ..., n \tag{2.1}$$

where X_i is a k-component vector denoting the i^{th} observation on k independent variables; Y_i is i^{th} observation on dependent variable and β is a p-component parametric vector; $f(.)$ is a nonlinear function and the error ε_i is independently identically distributed with mean zero and unknown constant variance σ^2.

For instance, a response function with $k = 3$ and $p = 4$, is given by $f(X, \beta) = \beta_1 X_1 + \beta_2 X_2 + \beta_4 e^{\beta_3 X_3}$ Such that

$$Y_i = f(X_i, \beta) + \varepsilon_i, \qquad i = 1, 2, ..., n$$

Or $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_4 e^{\beta_3 X_{3i}} + \varepsilon_i, \qquad i = 1, 2, ..., n$

Here, $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \qquad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}$

The set of 'n' nonlinear regression equations in (2.1) can be written in a convenient vector form as

$$Y_{n\times 1} = f_{n\times 1}\left(X_{n\times k}, \beta_{p\times 1}\right) + \varepsilon_{n\times 1} \text{ or } \qquad Y = f(\beta) + \varepsilon$$

Where $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{bmatrix}_{n\times 1}, \quad f(\beta) = \begin{bmatrix} f(X_1, \beta) \\ f(X_2, \beta) \\ . \\ . \\ f(X_n, \beta) \end{bmatrix}_{n\times 1}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{bmatrix}_{n\times 1}$

Let $F(\beta) = \frac{\partial}{\partial \beta'} f(\beta) = \left(\left(\frac{\partial}{\partial \beta_j} f(X_i, \beta) \right) \right)_{n\times p}$, where i is the row index and j is the column index, i=1,2,…,n and j=1,2,…,p.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Assuming that $f(X,\beta)$ is twice continuously differentiable on Θ, where $\beta \in \Theta$, by applying Taylor series expansion to linearize the nonlinear function $f(X,\beta)$, one may obtain for $\beta = \beta^0$ as

$$f(\beta) = f(\beta^0) + (\beta - \beta^0) F(\beta^0)$$

Where

$$f(\beta^0) = \begin{bmatrix} f_1(\beta^0) \\ f_2(\beta^0) \\ . \\ . \\ f_n(\beta^0) \end{bmatrix}_{n\times 1} \text{and} \qquad F(\beta^0) = \frac{\partial}{\partial \beta'} f(\beta^0)$$

By the least squares estimation, the estimator $\hat{\beta}$ minimizes residual sum of squares $R(\hat{\beta})$ over the parametric space Θ, Consider the least squares residual sum of squares as $R(\hat{\beta}) = \left[Y - f(\hat{\beta})\right]' \left[Y - f(\hat{\beta})\right]$. If $R(\hat{\beta})$ is once continuously differentiable on some open set Θ^0 with $\beta \in \Theta^0 \subset \Theta$, then $\hat{\beta}$ satisfies the set of normal equations $F'(\hat{\beta})\left[Y - f(\hat{\beta})\right] = 0$

$$\left[\because \frac{\partial}{\partial \hat{\beta}} R(\hat{\beta}) = 0 \text{ at any local optimum}\right]$$

In the nonlinear regression analysis, the least squares residuals are orthogonal to the columns of the Jocobian of $f(\beta)$ evaluated at $\beta = \hat{\beta}$, namely, $F'(\hat{\beta})\left[Y - f(\hat{\beta})\right] = 0$

Here, $f(\beta)$ is an $(n\times 1)$ vector valued function or a p-dimensional argument β. The Jacobian of $f(\beta)$ evaluated at $\beta = \hat{\beta}$ is $(n\times p)$ matrix, which is given by,

$$\frac{\partial}{\partial \beta'} f(\hat{\beta}) = \begin{bmatrix} \frac{\partial}{\partial \beta_1} f_1(\hat{\beta}) & \frac{\partial}{\partial \beta_2} f_1(\hat{\beta}) & . & . & \frac{\partial}{\partial \beta_p} f_1(\hat{\beta}) \\ \frac{\partial}{\partial \beta_1} f_2(\hat{\beta}) & \frac{\partial}{\partial \beta_2} f_2(\hat{\beta}) & . & . & \frac{\partial}{\partial \beta_p} f_2(\hat{\beta}) \\ . & . & & & . \\ . & . & & & . \\ \frac{\partial}{\partial \beta_1} f_n(\hat{\beta}) & \frac{\partial}{\partial \beta_2} f_n(\hat{\beta}) & . & . & \frac{\partial}{\partial \beta_p} f_n(\hat{\beta}) \end{bmatrix}$$

Thus, the least squares estimator of the unknown parameter β in the nonlinear model $Y = f(\beta) + \varepsilon$ is the $(p\times 1)$ vector $\hat{\beta}$ that will be obtained by solving the set of nonlinear normal equations, $F'(\hat{\beta})\left[Y - f(\hat{\beta})\right] = 0$. Further, the estimate of the error variance corresponding to the least squares estimator $\hat{\beta}$ is

$$\hat{\sigma^2} = S^2 = \frac{R(\hat{\beta})}{n-p}, \quad \text{where } R(\hat{\beta}) = \left[Y - f(\hat{\beta})\right]' \left[Y - f(\hat{\beta})\right]$$

3. Sampling Distributions of Nonlinear Least Squares Estimators of Regression Parametric Vector β and Error Variance σ^2

Consider the nonlinear model

$$Y_{n\times 1} = f_{n\times 1}(\beta_{p\times 1}) + \varepsilon_{n\times 1} \tag{3.1}$$

The least squares estimator of the unknown parametric vector β is $\hat{\beta}$ that minimizes the residual sum of squares

$$R(\hat{\beta}) = \left[Y - f(\beta)\right]' \left[Y - f(\beta)\right] \tag{3.2}$$

Under asymptotic normality theory, it can be easily shown that

$$\hat{\beta} = \beta + (F'F)^{-1} F'\varepsilon + O_p\left(\frac{1}{\sqrt{n}}\right) \tag{3.3}$$

And

$$\hat{\sigma^2} = S^2 = \frac{\varepsilon'\left[I - F(F'F)^{-1}F'\right]\varepsilon}{n-p} + O_p\left(\frac{1}{n}\right) \tag{3.4}$$

where, $F = F(\beta) = \frac{\partial}{\partial \beta'} f(\beta)$ is the matrix with typical row $\frac{\partial}{\partial \beta'} f(X_i, \beta)$.

Here, $O_p\left(\frac{1}{\sqrt{n}}\right)$ and $O_p\left(\frac{1}{n}\right)$ denote matrix valued random variables $X_n = O_p\left(\frac{1}{\sqrt{n}}\right)$ and $X_n = O_p\left(\frac{1}{n}\right)$ with the property that each element X_{ijn} satisfies $\lim_{n\to\infty} p\left[\left|\frac{X_{ijn}}{a_n}\right| > \varepsilon\right] = 0$, for any $\varepsilon > 0$; and $\{a_n\}$ is some sequence of real numbers such that $a_n = \frac{1}{\sqrt{n}}$ or $a_n = \frac{1}{n}$.

These expressions suggest that a good approximation to the joint distribution of $(\hat{\beta}, S^2)$ can be obtained by ignoring the terms $O_p\left(\frac{1}{\sqrt{n}}\right)$ and $O_p\left(\frac{1}{n}\right)$.

Thus, one may obtain,

$$\hat{\beta} = \beta + (F'F)^{-1} F'\varepsilon \tag{3.5}$$

And

$$S^2 = \frac{\varepsilon'\left[I - F(F'F)^{-1}F'\right]\varepsilon}{n-p} \tag{3.6}$$

By assuming normality of the errors ε, one may have approximately the multivariate normal distribution for $\hat{\beta}$ as

$$\hat{\beta} \overset{asy}{\sim} N_p\left[\beta, \sigma^2 (F'F)^{-1}\right] \tag{3.7}$$

i.e., $\hat{\beta}$ has the asymptotically p-dimensional multivariate normal distribution with mean vector β and variance – covariance matrix $\sigma^2 (F'F)^{-1}$.

Further $\left[\frac{(n-p)S^2}{\sigma^2}\right] \overset{asy}{\sim} \chi^2_{(n-p)}$ (3.8)

Here, $\hat{\beta}$ and S^2 are independent, so that the joint distribution of $(\hat{\beta}, S^2)$ is the product of marginal distributions. Since, $\hat{\beta}$ converges almost surely to β; S^2 converges almost surely to σ^2; $\frac{1}{n}\left[F'(\hat{\beta})\right]' \left[F(\hat{\beta})\right]$ converges almost surely to a matrix Q; and

$\sqrt{n}\left(\hat{\beta}-\beta\right)$ is asymptotically distributed as the p-dimensional multivariate normal distribution with mean zero and variance - covariance matrix $\sigma^2 Q^{-1}$

i.e., $\sqrt{n}\left(\hat{\beta}-\beta\right) \overset{asy}{\sim} N_p\left(O,\sigma^2 Q^{-1}\right)$ (3.9)

In the applications, $(F'F)^{-1}$ can be approximated by the matrix

$$\hat{\phi}=\left\{\left[F\left(\hat{\beta}\right)\right]'\left[F\left(\hat{\beta}\right)\right]\right\}^{-1} \quad (3.10)$$

and $$\hat{Q}=\frac{1}{n}\left[F\left(\hat{\beta}\right)\right]'\left[F\left(\hat{\beta}\right)\right] \quad (3.11)$$

One may get, $\hat{\beta}\dot{\sim} N_p\left(\beta, S^2\hat{\phi}\right)$ or $\hat{\beta}\dot{\sim} N_p\left(\beta, S^2\left(\frac{1}{n}\right)\hat{Q}^{-1}\right)$

and $\sqrt{n}\left(\hat{\beta}-\beta\right)\dot{\sim} N_p\left(O, S^2\hat{Q}^{-1}\right)$ or $\sqrt{n}\left(\hat{\beta}-\beta\right)\dot{\sim} N_p\left(O, S^2 n\hat{\phi}\right)$

Sometimes, $\overset{*}{\sigma^2}=\frac{R\left(\hat{\beta}\right)}{n}$ may be used as an estimate of σ^2 instead of S^2.

4. Testing General Nonlinear Hypothesis on Parameters of Nonlinear Regression Model

Consider the nonlinear regression model in vector notation as

$$Y=f(\beta)+\varepsilon \quad (4.1)$$

such that $\varepsilon \sim N_n\left(O,\sigma^2 I_n\right)$. One may wish to test the hypothesis as $H_0: g(\beta)=0 \sim H_1: g(\beta)\neq 0$ .where $g(\beta)$ is continuously differentiable function mapping $\square^p$ into $\square^q$ with Jacobian $G(\beta)=\frac{\partial}{\partial\beta'}g(\beta)$ which is $(q\times p)$ matrix. When $G(\beta)$ is evaluated at $\beta=\hat{\beta}$, one may denote $\hat{G}=G\left(\hat{\beta}\right)$ and $G=G(\beta)$. Under asymptotic normality theory $g\left(\hat{\beta}\right)$ may be characterized as

$$g\left(\hat{\beta}\right)=g(\beta)+G(F'F)^{-1}F'\varepsilon+O_p\left(\frac{1}{\sqrt{n}}\right) \quad (4.2)$$

Where $F=\frac{\partial}{\partial\beta'}f(\beta)$. Ignoring the remainder term, an approximate distribution of $g\left(\hat{\beta}\right)$ is given by

$$g\left(\hat{\beta}\right)\overset{\square}{\sim} N_q\left[g(\beta),\sigma^2 G(F'F)^{-1}G'\right] \quad (4.3)$$

One may obtain approximately the distribution as

$$\left\{\frac{\left[g\left(\hat{\beta}\right)\right]'\left[G(F'F)^{-1}G'\right]^{-1}\left[g\left(\hat{\beta}\right)\right]}{\sigma^2}\right\}\overset{\square}{\sim}\text{Noncentral}\ \chi^2_q \quad (4.4)$$

with non -centrality parameter

$$\lambda=\frac{\left[g(\beta)\right]'\left[G(F'F)^{-1}G'\right]^{-1}\left[g(\beta)\right]}{2\sigma^2} \quad (4.5)$$

Since, $\left[\frac{(n-p)S^2}{\sigma^2}\right]$ is distributed independently with $\hat{\beta}$ as the χ^2_{n-p}, one may have,

$$\left\{\frac{\left[g\left(\hat{\beta}\right)\right]'\left[G(F'F)^{-1}G'\right]^{-1}\left[g\left(\hat{\beta}\right)\right]/q\sigma^2}{(n-p)S^2/\left[(n-p)\sigma^2\right]}\right\}\overset{\square}{\sim}\text{Noncentral}\ F_{[q,n-p]} \quad (4.6)$$

with non-centrality parameter λ.

$$\Rightarrow\left\{\frac{\left[g\left(\hat{\beta}\right)\right]'\left[G(F'F)^{-1}G'\right]^{-1}\left[g\left(\hat{\beta}\right)\right]}{qS^2}\right\}\overset{\square}{\sim}\text{Noncentral}\ F_{[q,n-p]} \quad (4.7)$$

with non-centrality parameter λ .In applications, G and $(F'F)^{-1}$ can be replaced with the corresponding estimations G' and $\hat{\phi}$ where $\hat{\phi}=\left\{\left(F\left(\hat{\beta}\right)\right)'\left(F\left(\hat{\beta}\right)\right)\right\}^{-1}$

Now the Wald test statistic for testing H_0 against H_1, is given by

$$W=\frac{\left[g\left(\hat{\beta}\right)\right]'\left[\hat{G}\hat{\phi}\hat{G}'\right]^{-1}\left[g\left(\hat{\beta}\right)\right]}{qS^2}\square F(q,(n-p)) \quad (4.8)$$

Remark: Suppose that one may fit the model as $e=\hat{F}\delta+\nu$ by the least squares and one may wish to test the hypothesis

$$H_0:\hat{G}\delta=g\left(\hat{\beta}\right)\square\ H_1:\hat{G}\delta\neq g\left(\hat{\beta}\right)$$

To test the H_0, the F – test statistic may be used as

$$F=\frac{\left[\hat{G}\hat{\delta}-g\left(\hat{\beta}\right)\right]'\left[\hat{G}\left(\hat{F}'\hat{F}\right)^{-1}\hat{G}'\right]^{-1}\left[\hat{G}\hat{\delta}-g\left(\hat{\beta}\right)\right]/q}{\left[e-\hat{F}\hat{\delta}\right]'\left[e-\hat{F}\hat{\delta}\right]/n-p} \quad (4.9)$$

Since, $\frac{\partial}{\partial\beta}R\left(\hat{\beta}\right)=-2\,\hat{F}e'$, one may have $\left(\hat{F}'\hat{F}\right)^{-1}\hat{F}'e=\hat{\delta}$ and F – test statistic is given by

$$W=\frac{\left[g\left(\hat{\beta}\right)\right]'\left[\hat{G}\,\hat{\phi}\,\hat{G}'\right]^{-1}\left[g\left(\hat{\beta}\right)\right]}{qS^2}\square F_{[q,(n-p)]} \quad (4.10)$$

5. Conclusion

In the above discussion some applications of nonlinear regression models has been presented. Sampling distribution of nonlinear least squares estimator has been given for testing general nonlinear hypothesis on parameters of nonlinear regression model. Using principles of matrix calculus least squares estimator of unknown parameter of a nonlinear regression model has been obtained.

References

[1] Arthipova Irina, Arthipovs Sergejs, "Application of Ordinary Least Square Method in Nonlinear models", International statistical Institute, 56th session, (2007).

[2] MA Tabatabai, JJ Kengwoung Keuomo and K. P. Singh, "A new Robust method for nonlinear regression", Journal of biometrics and biostatistics, Vol. (5). (2014).

[3] Xu Zheng, "Testing heteroscedasticity in nonlinear and nonparametric regression", Canadian Journal of statistics, Vol. (37), issue 2, (2009), Pp: 282-300.

[4] Gordon K. Smyth, "Nonlinear regression, Encyclopedia of Environ metrics Vol. (3), (2002), Pp: 1405-1411

[5] Gurleen K. Popli , "A note on the instrumental variable estimators in the nonlinear models", Journal of Quantitative Economics Vol.16. no.2, (2000), Pp: 31-36.
[6] Davidian M. and Giltinon D.M, "Nonlinear models for repeated measurement Data: An overview and update", Journal of Agricultural, Biological and Environmental statistics (JABES), Vol. (8), (2003), Pp: 387-419.
[7] Vasilyev D.M, "Theoretical and Practical Aspects of linear and nonlinear models order reduction Techniques", MIT, USA, (2008).
[8] E. Grafarent and J. Awange, "Application of linear and nonlinear models", Springer Geophysics, (2012).
[9] Bates D.M and Walts D.G, "Nonlinear regression: Iterative Estimation and Linear Approximations in Nonlinear regression Analysis and its Applications", John Wiley and sons Inc Hobeken, NJ, USA, (2008).