Estimation methods of nonlinear regression models

Cite as: AIP Conference Proceedings **2177**, 020081 (2019); https://doi.org/10.1063/1.5135256
Published Online: 04 December 2019

Ranadheer Donthi, S. Vijay Prasad, B. Mahaboob, J. Peter Praveen, and B. Venkateswarlu

Estimation Methods of Nonlinear Regression Models

Ranadheer Donthi[1, b)], S. Vijay Prasad[2], B.Mahaboob[2], J. Peter Praveen[2], B.Venkateswarlu[3, a)]

[1]*St. Martin's Engineering College, Dhulapally, Kompally, Hyderabad, Telangana, India 500100*
[2]*Department of Mathematics,Koneru Lakshmaih Education Foundation, Vaddeswaram, Guntur Dist.AP-522502-India*
[3]*Department of Mathematics, Vellore Institute of Technology, Vellore, Tamilnadu*

[a)]Corresponding author: venkatesh.reddy@vit.ac.in
[b)] ranadheer.donthi@gmail.com

Abstract. Nonlinear regression models have been a subject of an intensive investigation in recent years and they have wide uses in applied sciences namely Medicine, Forensic Science, Food Science, Information Science, Applied Ecology, Agronomy, Sports Science and Space Science. This research article mainly focuses on some important and innovative nonlinear estimation techniques of parameters of nonlinear regression models based on principles in matrix differentiation. The methods depicted here are principle of least squares, linear approximation method, and MLE estimation method. Oral Capps, Jr, (see [1]), in his research paper presented a theoretical discussion and some empirical results using maximum likely-hood (ML) method and iterative version of Zellner's seemingly unrelated regression (IZEF) method in the estimation of a non linear system of demand equations when the disturbance terms are both contemporaneously and serially correlated. S. Neal, (see [3]), in his research article considered a discrete real- time nonlinear estimation problem using a least square criterion and derived a sequential algorithm which follows consideration of second order nonlinearities in system measurements. Besides some alternative nonlinear estimation techniques were discussed and examples were given which compare the various estimation algorithms. In 2003, Thomas Schon (see [4]), in his thesis, made a discussion on how to use convex optimization for solving the estimation problem.

INTRODUCTION

Modelling is a cyclic process of creating and modifying models of empirical solutions to understand them better and improve decisions. The role of modelling and mathematical modelling have received increasing attention as generating authentic learning and revealing the ways of thinking that produced it. Regression analysis is nothing but predictive modelling which enquires the connection between the free variable and dependent variable and it has a number of practical applications in real life, e.g. the relationship between rash driving and number of road accidents by a driver is best studied through regression analysis. There are number of types of regression techniques in the literature and some of them are linear regression, logistic regression, polynomial regression, stepwise regression, ridge regression, lasso regression and elastic net regression.

In nonlinear regression analysis the dependent variables are modelled as a nonlinear functional models with unknown coefficients and one or more free variables. In literature there are a large number of nonlinear regression models. Nonlinear regression analysis provides particular technologies to under-stand itself better. There are some important nonlinear growth models which are very useful to know the growth behavior in a particular period namely Maltus model, Monomolecular model, Logistic model, Gompertz model and Richards model. Human beings need to understand naturally observed phenomena and predict them. Besides it helps us to

Recent Trends in Pure and Applied Mathematics
AIP Conf. Proc. 2177, 020081-1–020081-5; https://doi.org/10.1063/1.5135256
Published by AIP Publishing. 978-0-7354-1924-7/$30.00

manipulate the world around us using the several mathematical approaches namely complex systems, statistics, dynamical systems and chaos, numbers, geometry, calculus including differential equations, algebra humans reach the above goal. Casting the observed phenomenon into a mathematical description can be done using mathematical modelling. Some observed phenomenon may be too complicated to understand in which case mathematical modelling does its work hopefully. As model is a simplification of the reality there are always minute differences between expected and observed phenomenon. Now-a-days mathematical modelling is an indispensable tool to depict, estimate and manipulate the natural phenomenon. The success of mathematical model is that if one can form an exact mathematical model for some phenomenology then it is used for prediction of future behavior of that physical process. Moreover in many elds of applied sciences mathematical modelling is generally used to crack highly concrete practical problems.

NON LINEAR LEAST SQUARES ESTIMATION METHOD.

A nonlinear regression model can be put as

$$Z_j = g(Y_{1j,}Y_{2j,}.....Y_{lj};\alpha_1,\alpha_2......\alpha_r)+\varepsilon_j;\ j=1,2,3....m. \quad (1)$$

where Z is dependent variable,

$Y_1, Y_2, ...Y_l$ are independent variables,

$\alpha_1, \alpha_2.....\alpha_r$ are parameters,

ε is error random variable,

m= number of rows,

g (.)= functional which is not linear.

Setting $Z=\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ z_m \end{bmatrix}$, $Y=\begin{bmatrix} Y_{11} & Y_{12} & .. & Y_{1l} \\ Y_{21} & Y_{22} & .. & Y_{2l} \\ .. & .. & .. & .. \\ Y_{1m} & Y_{2m} & .. & Y_{lm} \end{bmatrix}$, $\alpha=\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ .. \\ \alpha_r \end{bmatrix}$, $\varepsilon=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ .. \\ \varepsilon_m \end{bmatrix}$, and

$g(Y,\alpha)=\begin{bmatrix} g(Y_{p1},\alpha) \\ g(Y_{p2},\alpha) \\ .. \\ g(Y_{pm},\alpha) \end{bmatrix}$, $p=1,2,...l.$ gives the matrix notation of nonlinear regression model as

$$Z = g(Y,\alpha)+\varepsilon. \quad (2)$$

The error are taken as identically independent distributed random variables with $\varepsilon \approx (o,\sigma^2 I_n)$ and the exact distribution is not known.

RSS is defined as

$$\left(g\left(Y,\hat{\alpha}\right)-Z\right)'\left(g\left(Y,\hat{\alpha}\right)-Z\right). \quad (3)$$

The NLLS estimator $\hat{\alpha}$ and α is got by minimizing $R(\hat{\alpha})$ with respect to $\hat{\alpha}$ and solving the normal equations

$$R(\hat{\alpha})=\sum_{i=1}^{m}(Z_j - g(Y_j,\hat{\alpha}))^2$$

The normal equations are obtained by

$$\frac{\partial R(\hat{\alpha})}{\partial \hat{\alpha}} = 0$$

$$\sum_{j=1}^{m} (Z_j - g(Y_j, \hat{\alpha})) \left(\frac{\partial g(Y_j, \alpha)}{\partial \alpha} \right)_{\alpha = \hat{\alpha}} = 0$$

Since $g(Y_j, \alpha)$ is nonlinear in $\alpha's$, the normal equations will be nonlinear in Y's and $\alpha's$. Generally $\hat{\alpha}$ will not be a linear function of Z and optimally principles cannot be obtained for NLLS estimators. It is tedious to solve these normal equations to that iterative techniques should be implemented. Under certain regularity conditions the NLLS estimator $\hat{\alpha}$ is consistent and follows asymptotical normal distribution and reasonable estimator of σ^2 is $\hat{\sigma}^2 = \frac{R(\hat{\alpha})}{m-1}$.

LINEAR APPROXIMATION MODEL

In matrix nonlinear model is given by

$$Z = g(Y, \alpha) + \varepsilon. \tag{4}$$

where $\varepsilon : ^{iid} (o, \sigma^2 I_m)$.

Using Taylors series in Multivariate Calculus and neglecting the terms from second and higher order derivatives the normal functional $g(Y, \alpha)$ as a linear functional like

$$g(\alpha, Y) \approx g(\alpha^*, Y) - \left(g(Y, \alpha)\right)'_{\alpha = \alpha^*} (\alpha^* - \alpha)$$

$$g(\alpha, Y) \approx g(\alpha^*, Y)$$

Now

$$Z \approx g(Y, \alpha^*) + Q(\alpha^*)(\alpha - \alpha^*) + \varepsilon$$

$$Z^* \approx Q(\alpha^*)\alpha + \varepsilon. \tag{5}$$

Where $Z^* \approx Z - g(Y, \alpha^*) + Q(\alpha^*)\alpha^*$.

Since (4) and (5) are identical for $\alpha = \alpha^*$, a consistent estimator $\hat{\alpha}_{NL}$ say for α^* in (4) will almost coincide with a corresponding consistent estimator $\hat{\alpha}_{LS}$ say for α^* in (5) in the limit and thus these properties of $\hat{\alpha}_{NL}$ and $\hat{\alpha}_{LS}$ are similar at least large samples, (5) is called linear Pseudo model. The LS estimator $\hat{\alpha}_{LS}$ for α is got by

$$\hat{\alpha}_{LS} = [Q(\alpha^*)^1 Q(\alpha^*)]^{-1} [Q(\alpha^*)^1 Z^*]. \tag{6}$$

The variance – covariance matrix of ε is assumed as

$$\psi_\varepsilon = \sigma^{*2} I_m$$

Then that of $\hat{\alpha}_{LS}$ is

$$(\psi_\alpha)_{LS} = \sigma^{*2} [Q(\alpha^*)^1 Q(\alpha^*)]^{-1}. \tag{7}$$

The covariance matrix of $\hat{\alpha}_{NL}$ is approximated as $(\psi_\alpha)_{NL} = \sigma^{*2} [Q(\alpha^*)^1 Q(\alpha^*)]^{-1}$ is large samples.

An estimate σ^{*^2} is got by

$$\sigma^{*^2} = \frac{[Z^* - Q(\alpha^*)\hat{\alpha}_{NL}]^1[Z^* - Q(\alpha^*)\hat{\alpha}_{NL}]}{m-s}. \tag{8}$$

Under certain regularity conditions $\hat{\alpha}_{NL}$ follows asymptotically normal distribution with mean α and covariance

$$\sigma^{*^2}[Q(\alpha^*)^1 Q(\alpha^*)]^{-1}. \tag{9}$$

Since $Q(\alpha^*) = \frac{\partial g(Y,\alpha)}{\partial \alpha}$ at $\alpha = \alpha^*$ is still unknown it can be approximated by

$$\hat{Q}(\alpha^*) = \left(\frac{\partial g(Y,\alpha_{NL})}{\partial \alpha_{NL}}\right)_{\alpha=\hat{\alpha}_{NL}}. \tag{10}$$

$[Q(\alpha^*)^1 Q(\alpha^*)]^{-1}$ is approximated by

$$\hat{Q} = [\hat{Q}(\alpha^*)^1 \hat{Q}(\alpha^*)]^{-1} = (\phi_{ij}). \tag{11}$$

The test statistic applied to test the null hypothesis $H_0 = \alpha_j = \alpha_{j0},$ one can use

$$t - \frac{\hat{\alpha}^j_{NL} - \alpha_{j0}}{\hat{\sigma}^* \sqrt{\hat{\phi}_{ii}}} : t_{n-p}. \tag{12}$$

MAXIMUM LIKELIHOOD METHOD OF ESTIMATIONLINEAR APPROXIMATION MODEL

Under certain regularity conditions if the observations are independent the ML estimators posses some optimal principles. The nonlinear regression model in specified form is

$$Z = g(Y,\alpha) + \varepsilon. \tag{13}$$

where $\varepsilon : N(0,\eta)$

The likelihood function Z is given by

$$L(\alpha,\eta) = \frac{1}{(\sqrt{2\pi})^m(\sqrt{|\eta|})^m}\exp(-(0.5)(g(Y,\alpha)-Z)^1 1/\eta(g(Y,\alpha)-Z). \tag{14}$$

$$\log L(\alpha,\eta) = \frac{m}{2}[\log 2 + \log \pi] + \frac{m}{2}\log|\eta| + \frac{1}{2}(g(Y,\alpha)-Z)^1 1/\eta(g(Y,\alpha)-Z). \tag{15}$$

The MLE's α and η are obtained by maximizing $\log L(\alpha,\eta)$ or minimizing $\log(L(\alpha\eta)^{-1})$ with respect to α and η

Put $L^* = -\log L(\hat{\alpha},\hat{\eta})$

Set $\eta = \sigma^2 I_m.$

Then one can achieve

$$L^* = -\log L = m(0.5)[\log 2 + \log \pi] + m(0.5)\log \hat{\sigma}^2 + (0.5)\hat{\sigma}^{-2}(g(Y,\hat{\alpha})-Z)^1(g(Y,\hat{\alpha})-Z). \tag{16}$$

Now one can minimize $(g(Y,\alpha)-Z)^1(g(Y,\alpha)-z)$ with respect to $\hat{\alpha}$.

Thus the MLE of α say $\hat{\alpha}_{ML} = OLS$ estimator of $\alpha = \hat{\alpha}_{OLS}$.

Moreover $\hat{\alpha}^2_{ML} = (m)^{-1}(g(Y,\hat{\alpha}_{ML})-Z)^1(g(Y,\hat{\alpha}_{ML})-Z).$

Under certain regularity conditions the MLE's are consistent, sufficient. Further they are asymptotically efficient and follow asymptotic normal distribution.

CONCLUSIONS

In the above research article some significant nonlinear estimation methods of parameters of nonlinear regression models based on matrix calculus viz. Nonlinear least squares estimation, Taylor's series expansion method and MLE method are described. In context of future research one can derive Newton Rapson method, Steepest Descent/Ascent method, Gauss Newton method, Scoring method, Quadratic hill climbing method and Conjugate gradient method.

REFERENCES

1. Oral Gpps Jr, Western Journal of Agriculture Economics. 50-63(1983).
2. Raja Manish, Prdue University, West Lafaye , USA. (2014).
3. S. Neal, IEE digital explore, IEE Digital Library., **13** (1968).
4. Thomas Schon, Department of Electrical Engineering, Sweden (2003).
5. G.E. Hovland, Science Direct, Elsevier **13**, 1341-1355(2005).
6. B. Mahaboob, IOP Conference series: Material Science and Engineering. **263** (2017).
7. S. J. Julier, J.K. uhimann, Proceedings of IEE., **92** (2004).
8. Ratowsky. D.A, Willey Publications**., 2** (1990).
9. Ross, G.J.S, Journal of Applied Statistics. **17,** 299-306(1990).
10. Popli, G.K, Journal of Quantitative Economics. **16**,), 31-36, (2000).