Annals of Pure and Applied Mathematics
Vol. 5, No.1, 2013, 64-70
ISSN: 2279-087X (P), 2279-0888(online)
Published on 13 November 2013
www.researchmathsci.org

Annals of
Pure and Applied Mathematics

Nonlinear Statistical Model and its Applications to Diffusion of Mobile Telephony in India

P. K. Das

Indian Institute of Foreign Trade Kolkata Campus
J-1/14 EP & GP Block, Sector V, Salt Lake City, Kolkata 700091
Email: pkdas@iift.ac.in

Received 28 September 2013; accepted 5 October2013

Abstract. In this article, nonlinear statistical models have been described along with its applications in diffusion of mobile telephony in India. In statistical regression sense, nonlinear statistical models are those in which at least one of the parameters appears nonlinearly. Iterative techniques are employed to obtain the parameter estimates of nonlinear models. Internal influence model was found to be appropriate for describing the growth trajectory of mobile telephony in India. The maximum possible GSM subscriber base was found to be 760 million, which is likely to be achieved beyond the year 2020.

Keywords: Nonlinear statistical model, Levenberg-Marquardt algorithm, innovation diffusion model.

AMS Mathematics Subject Classification (2010): 62J02

1. Introduction

Most of the processes of bio-physical and socio-economic domains are nonlinear in nature. However, to model such processes most of the times linear models are employed. Simple, elegant, easy-to-use linear models remained the choice of researchers for a long time. Linear models provide only an approximation of the system being studied, and may sometimes represents a distortion of the underlying process. Today, due to the availability of high speed low cost computers coupled with appropriate software, which is built on sound statistical theory, made it possible to use nonlinear models, which require high computing dexterity (Seber and Wild [7], Gallant [1]). Models whether deterministic or stochastic, empirical or mechanistic, may be linear or nonlinear in nature. Linearity and nonlinearity can be explained in a variety of ways:

1.1. Input output relationship

In linear model output is proportional to input, whereas in nonlinear model the output is not proportional to input.

1.2 System theory

A model is linear in the system theory sense (LST) if the principle of superposition holds, that is, given that y1, y2 are the outputs corresponding to inputs x1, x2, a model is LST if the output corresponding to input x1+x2 is y1+y2; else the system is nonlinear. Thus, if input X and output Y are related by the equation $Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_1^2 + \beta_4X^2_2+\beta_5X_1X_2$, or $Y=\beta_0+\beta_1X+\beta_2X^2$, then the models are said to be nonlinear in system theory sense.

1.3. Appearance of parameters in the model

The model is linear in the statistical regression sense (LSR) if at least one of the parameters of the model appears linearly. Thus, if input X and output Y are related by the equation $Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_1^2+\beta_4X^2_2+\beta_5X_1X_2$, or $Y=\beta_0+\beta_1X+\beta_2X^2$,
or $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k$, then the models are linear in the statistical regression sense. Therefore, linear regression models are ones in which parameters appear linearly, whereas nonlinear regression models have at least one parameter appearing nonlinearly. Examples of nonlinear models are: $Y = \exp(\beta_1 X + \beta_2 X^2)$ and $Y = \beta_1 X + \exp(-\beta_2 X)$.

1.4. Derivative of the function

If the derivatives of the function with respect to the independent variable do not depend upon the independent variable; the model is considered as linear. If the derivatives are functions of the parameter and independent variable, the model is said to be nonlinear.

1.5. Nonlinear measures

Models which results intrinsic (IN) and parameter effects (PE) curvature measures in the parameter space are nonlinear models. Linear models do not have such properties. In a data set having n observations, if different values obtained with different parameter values are plotted then the hyperplane so obtained will have a curvature. This curvature is known as intrinsic nonlinearity. The spacing of these points is unequal. This property is known as parameter effects (Ratkowsky[3], Bates and Watts[6]).

1.6. Examples of nonlinear models

Let us consider some examples of nonlinear models, the volume v and pressure p of gas satisfy the relationship $pv^{\gamma}=k$. Writing y=p and $x=v^{-1}$, we get $y=kx^{\gamma}=f(x;k,\gamma)$, γ is a parameter for each gas, which is to be estimated from the data. This model is nonlinear in γ but linear in k (Draper and Smith[10]). Let us take one more example, in an irreversible chemical reaction, substance A changes into substance B, which in turn changes into substance C. Let A(t) denotes the quantity of substance A at time t. The differential equations governing the changes are

$$\frac{dA(t)}{dt} = -\theta_1 A(t), \quad \frac{dB(t)}{dt} = \theta_1 A(t) - \theta_2 B(t), \quad \frac{dC(t)}{dt} = \theta_2 B(t)$$

where θ_1 and θ_2 are unknown parameters. Assuming A(0)=1, B(t) has the solution

$$B(t) = \frac{\theta_1}{\theta_1 - \theta_2}[\exp(-\theta_2 t) - \exp(-\theta_1 t)].$$

This model is nonlinear in θ_1 and θ_2 (Draper and Smith[10]). Cobb-Douglas production function given by $Y=\beta_0 X_1^{\beta 1} X_2^{\beta 2}$ is an example of nonlinear model, where Y is output and X_1 and X_2 are capital and labour input, respectively. Another well-known model is constant elasticity of substitution (CES) production function given by $Y=\beta_0[\delta K^{-\beta}+(1-\delta)L^{-\beta}]^{-1/\beta}$, where Y is output, K and L are capital and labour input, respectively. β ($\beta \geq 1$), β_0 and δ ($0<\delta<1$) are substitution parameter, scale parameter, and distribution parameter, respectively (Intriligator, Bodkin and Hsiao[9]). Nonlinear models which can be transformed into linear form are known as *intrinsically linear nonlinear models*. Nonlinear models, which cannot be transformed into linear form, are known as *intrinsically nonlinear models.*

1.7. Objectives

(i) To briefly describe nonlinear models along with estimation technique.
(ii) To apply nonlinear model to understand the growth trajectory of mobile telephony in India.

2. Estimation of parameters of nonlinear models

All the models described in Section 1 appeared deterministically as if data never deviates from the model, which is unrealistic predominantly in biological and socio-economic sciences and also in physical sciences. To make the model realistic, independently, identically normally (iidN) distributed stochastic error term is added to the right hand side of the mathematical model, which results a nonlinear regression model. Nonlinear regression model differ greatly in their estimation properties from linear regression models in that, given the usual assumption of an independently and identically distributed normal stochastic error term, linear model give rise to unbiased, normally distributed minimum variance estimators. Nonlinear regression models tend generally to do so as the sample size becomes very large (asymptotically) (Gallant [1], Ross [8]).

2.1. Least square method

Let us consider the following nonlinear model

$$Y_t = X_t^{\theta} + \varepsilon_t$$

where θ is the parameter to be estimated. Least square method is used to estimate the parameters of the model. In least square technique θ is estimated by minimizing the error sum of squares

$$S(\theta) = \sum_{t=1}^{n}(Y_t - X_t^{\theta})^2, \; \frac{\delta S}{\delta \theta} = -2\sum_{t=1}^{n}(Y_t - X_t^{\theta})(\log X_t)(X_t^{\theta})$$

Denoting least square estimate of θ by $\hat{\theta}$, we get

$$\sum_{t=1}^{n} Y_t (\log X_t) X_t^{\hat{\theta}} = \sum_{t=1}^{n} X_t^{2\hat{\theta}} \log X_t$$

No explicit solution of θ is possible. Least square estimate of θ can be obtained iteratively starting from some assumed value of $\hat{\theta}$. The least square estimator $\hat{\theta}$ of θ in the above equation does not have properties possessed by least square estimator of linear model. Also, unlike linear model least square estimate of θ is not a linear function of observed Y–values (Ratkowsky [3]). One of the important issues in fitting nonlinear models is that model fitting generally requires the iterative optimization of functions. Unfortunately, the iterative process often does not converge easily to the desired solution. For finite samples, the general statement may be made that even though Y_t may be normally distributed about its mean X_t^{θ} with some finite unknown variance σ^2 for all t, t=1,2,...,n, $\hat{\theta}$ is not a linear combination of the Y_t and hence, in general, is not normally distributed, hence it is neither unbiased for θ, nor is it a minimum variance estimator. Thus, unlike a least square estimator of a parameter in a linear model, a least square estimator of a parameter in a nonlinear model has essentially unknown properties for finite sample sizes. A number of powerful algorithms for fitting nonlinear models are now available. These have been designed to handle complex models and to allow for the various contingencies that can arise in iterative optimization. Three main methods are: (i) Taylor series (or linearization) (ii) Steepest descent (or gradient) and (iii) Levenberg-Marquardt (Marquardt [5]). The details of these methods along with their merits and demerits are given in Draper and Smith [10]. Levenberg-Marquardt algorithm represents a compromise between the linearization method and the gradient method, and combines successfully the best features of both and avoids the serious disadvantages of both the methods. It is good in the sense that it almost always converges and does not ‘slow down’ at the latter part of the iterative process.

2.2. Computation of starting values

All the iterative procedures require starting values of the parameters. The choice of good starting values can spell the difference between success and failure in locating the fitted value or between rapid and slow convergence to the solution. However, there is no standard procedure for computing starting values of the parameters. Some common techniques are: (i) Plotting of data (ii) Prior information (iii) Previous experiments (iv) Known values for similar systems (v) Values computed from theoretical considerations (vi) Linearization (vii) Solving a system of equations (viii) Using properties of the model. Sometimes a combination of two or three methods result good starting values.

2.3. Software packages for nonlinear regression

Commercial software like Eviews (Quick>Estimate Equation>LS-Least squares (NLS and ARMA)), IBM SPSS (Analysis>Regression>Nonlinear), Minitab (Stat> Regression >Nonlinear Regression), SAS (NLIN) and Stata (Statistics>Linear models and related>Nonlinear least-squares estimation) has options to compute parameter estimates of nonlinear regression model. Free software R also has tools (NLS2) to compute parameter estimate of nonlinear model.

2.4. Goodness of fit of nonlinear model

This is generally assessed by the coefficient of determination, R^2. However, as pointed out by Kvalseth [13], eight different expressions for R^2 appear in the literature. One of the most frequent mistakes occurs when the fits of a linear and a nonlinear model are compared by using the same R^2 expression but different variables. Thus, for example, a power model or an exponential model may first be linearized by using a logarithmic transformation and then fitted to data by using ordinary least squares method. The R^2-value is then often calculated using the log of observed and log of predicted data points. The R^2 is generally interpreted as a measure of goodness of fit of even the original nonlinear model, which is incorrect. Scott and Wild[2] have given a real example where two models are identical for all practical purposes and yet have very different values of R^2 calculated on the transformed scales. Kvalseth [13] has emphasized that, although R square given by

$$R^2 = 1 - \frac{RegSS}{TotalSS}$$

where RegSS is the regression sum of squares and TotalSS is the total sum of squares, is quite appropriate even for nonlinear models. Other summary statistics like

$$\text{Mean Absolute Error (MAE)} = \frac{\sum | y_i - \hat{y}_i |}{n}$$

and

$$\text{Mean Squared Error (MSE)} = \frac{\Sigma (y_i - \hat{y}_i)^2}{(n - p)}$$

should also be computed. Here n is the total number of observed values and p denotes the number of model parameters.

2.5. Model diagnostics

Coefficient of determination, MAE and MSE are important parameters of goodness of fit. However, sole dependency on these statistics may fail to reveal important data characteristics and model inadequacies. It is strongly recommended to carry out detailed analysis of the residuals to find out the suitability of a model. Following two important assumptions (i) errors are independent and (ii) errors are normally distributed can be tested using run test and Kolmogorov-Smirnov and Shapiro-Wilk test, respectively.

3. Important nonlinear models and applications

There are several important nonlinear models described in the literature under growth models (Seber and Wild [7], Ratkowsky [4], Prajneshu and Das [11,12]), yield-density models, enzyme kinetics and innovation diffusion models (Mahajan and Peterson [14]). In this Section an attempt is made to develop an appropriate nonlinear model for describing the diffusion path of mobile telephony in India.

3.1. Innovation Diffusion Model

Let n(t) and N(t) denotes the number of subscribers and cumulative number of subscribers, respectively to GSM service at time t in India. Also consider that total number of potential adopters in the social system (all India) is denoted by K, then [K-N(t)] is the number of individuals yet to adopt the technology. Let, the coefficient of diffusion is denoted by g(t). The rate of diffusion of new technology (mobile telephony) is assumed to be proportional to the individuals yet to adopt the technology. Following differential equation is proposed to model the growth trajectory of the mobile telephony in India.

$$\frac{dN(t)}{dt} = g(t)[K - N(t)]$$

with the boundary condition

$N(t = t_0) = N_0 =$ cumulative number of adopters at time t_0

$N(t) = \int_{t_0}^{t} n(t)dt$, $n(t)$ being the commutative number of adopters at time t.

$\frac{dN(t)}{dt} =$ rate of diffusion at time t

$g(t) =$ coefficient of diffusion.

g(t)*[K-N(t)] represents the expected number of adopters at time t.
Considering g(t)=bN(t), the differential equation can be solved, which results the following model

$$N(t) = \frac{K}{1 + \frac{K - N(0)}{N(0)} e - bK(t - t_0)}$$

Reparameterizing, it can be written as

$$N(t) = \frac{K}{1 + Be^{-At}}$$

This model is known as logistic model and in innovation diffusion literature as internal influence model, where potential adopters are assumed to be interacting with existing adopters before making adoption decision. It has a shape of elongated S and it can be divided into four distinct phases. These are: inception, fast growth, slow growth, and plateau. By changing the form of the expression g(t), other variants of the model can also be derived. However, these variants were not found appropriate in this case, hence not presented here. An stochastic error term was added to the above model to make it more realistic. The assumptions about the error term are that it is iid normal. The resulting model is

$$N(t) = \frac{K}{1 + Be^{-At}} + \varepsilon_t$$

3.2. Results

Monthly data of GSM subscribers collected from Cellular Operators Association from March 1997 to August 2013 was used to estimate the parameters of the above model. Parameters were estimated using Levenberg-Marquardt algorithm in IBM SPSS software. For development of the model data from March 1997 to August 2013 were used. One Step Ahead Forecast (OSAF) of September 2013 was made to check the suitability of the developed model. The models with g(t)=a and a+bN(t) failed to converge or converged to undesirable local minima having no logical interpretation. The final model is given by

$$\hat{N}(t) = \frac{760.874}{1 + 5108.903 e^{-0.056t}}.$$

R^2 of the model was found to be 0.997. OSAF was found to be very encouraging, which is within 4% of actual number. The maximum potential of GSM markets in India was found to be 760.874 million, which is likely to be achieved beyond 2020. The lower and upper limits of 95% confidence interval of this estimate are 748.351 and 773.397, respectively. The point of inflexion of this model is at K/2, which is found to be 380.437 and occurring in December 2009. The second growth phase *i.e.*, the fast growth phase appeared to have ended in December 2009. In the year 2020, model predicted number of GSM subscribers is found to be 760.032. The third phase of the growth process appeared to have started in 2009 and is likely to be continued till January 2020. Indian telecom industry is like to experience plateau in its growth trajectory beyond 2020.

REFERENCES

1. Ronald Gallant, *Nonlinear Statistical Models*, (1987), John Wiley & Sons, New York.
2. A. Scott, and C. J. Wild, Transformations and R^2, *Amer. Statistician,* 45, (1991). 127-28.
3. David A. Ratkowsky, *Nonlinear Regression Modeling a Unified Practical Approach*, (1983), Marcel Dekker, New York.
4. David A. Ratkowsky, *Handbook of nonlinear regression models*, (1990), Marcel Dekker, New York.
5. D. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters (doi: 10.1137/0111030), *Journal of the Society for Industrial and Applied Mathematics***,** 11(2) (1963), 431-441.
6. Douglas M. Bates and Donald G. Watts, *Nonlinear Regression Analysis and Its Applications*, (1988), John Wiley & Sons, New York.
7. G.A.F. Seber, and C.J. Wild, *Nonlinear regression*, (1989), John Wiley.
8. Gavin J. S. Ross, *Nonlinear Estimation*, (1990), Springer-Verlag, New York.
9. Michael D. Intriligator, Ronald Bodkin, and Cheng Hsiao, *Econometric Models, Techniques, and Applications*, (1996), Prentice Hall.
10. N. R. Draper and H. Smith, *Applied Regression Analysis*, (1998). 3rd ed. John Wiley.
11. Prajneshu and P. K. Das, Modelling wheat production at state level in post-green revolution era. *Indian Jr. of Agricultural Sciences*, 68(1), (1998), 23-6.
12. Prajneshu and P. K. Das, Growth models for describing state-wise wheat productivity, *Indian Jr. of Agricultural Research*, 34(3), (2000), 179-81.
13. T. O. Kvalseth, Cautionary note about R^2, *Amer. Statistician*, 39, (1985), 279-85.
14. Vijay Mahajan and Robert A. Peterson, *Models for Innovation Diffusion*, (1985), Sage Publications.