International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2018; 3(2): 581-587

www.mathsjournal.com
Received: 02-01-2018
Accepted: 04-02-2018

A Chinna kesavulu
Research Scholar, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

C Mani
Head, Department of Statistics, S.V. Arts College, Tirupati, Andhra Pradesh, India

SS Rajkumar
Block Health Statistician, Govt primary health centre, Kolathur, Salem, Tamil nadu, India

B Mahaboob
Department of Mathematics, Swetha Institute of Technology and Science, Tirupati, Andhra Pradesh, India

P Sri Vyshnavi
Senior Assistant Professor, Department of CSE, SPMVV Engineering College, Tirupati, Andhra Pradesh, India

P Manohar
Dept of Statistics, Sri Vidyanikethan, A. Rangampeta, Tirupati, Andhra Pradesh, India

P Balasiddamuni
Rtd. Professor, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

Correspondence
A Chinna kesavulu
Research Scholar, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

Statistical inference in nonlinear sure model

A Chinna kesavulu, C Mani, SS Rajkumar, B Mahaboob, P Sri Vyshnavi, P Manohar and P Balasiddamuni

Abstract
The seemingly unrelated Regression equations (SURE) model is a generalization of a regression model that consists of several regression equations, each having its an dependent variable and potentially different sets of exogenous explanatory variables. The SURE model approach for estimating system of linear regression equations in which the errors are contemporaneously corrected across equations across equations but not auto correlated. The SURE model is containing to receive wide spread attention in terms of both theoretical developments and empirical applications. In the present study, a Nonlinear SURE model has been specified and a feasible generalized least squares (GLS) estimator for the parametric vector has been developed along with its asymptotic variance covariance matrix. Further a consistent estimator of the dispersion matrix has been derived and it will be the nonlinear maximum likelihood estimator.

Keywords: Statistical inference, nonlinear sure model

1. Introduction
In Econometrics, the seemingly unrelated regression equations (SURE) model is a generalization of a linear regression model that consists of several regression equations, each having its own dependent variable and potentially different sets of exogenous explanatory variables. The Seemingly Unrelated Regression Equations (SURE) model approach for estimating a system of linear regression equations in which the disturbances are contemporaneously correlated across equations but not autocorrelated. Finally, a common situation which may suggest a SURE specification is where regression equations explaining a certain economic activity in different geographical locations are to be estimated. The SURE model is continuing to receive widespread attention, in terms of both theoretical developments and empirical applications.
Further, many Non-linear models can often be rearranged to be in a linear form. A Non-linear model is one in which at Least one of the parameters appears non-linearly. More formally, in a nonlinear model, at Least one derivative with respect to a parameter should involve that parameter.

2. Non-Linear Statistical Model
The general Non-linear regression model can be written as

$$y_i = f\left(x_i, \theta\right) + z_i, i = 1, 2, \ldots, \mathrm{N} \qquad (2.1)$$

Where f is the expectation function and x_i is a vector of associated regressor variables for the i^{th} case and z_i the error terms follow N (0, σ^2). This model is of exactly the same form as linear regression model except that the expected responses are nonlinear function of the parameters.
Consider the vectors x_i, i= 1, 2… N as fixed (observations) and concentrate on the dependence of the expected response on θ. We create the N-vector $\eta\left(\theta\right)$ with the i^{th} element

$$\eta(\theta) = f(x_i, \theta), i = 1, 2, \ldots, \text{N}$$

The model can be written as

$$Y = \eta(\theta) + z$$

With z assumed to have a spherical normal distribution. That is

$$E(z) = 0 \ and$$
$$Var(z) = E(zz^I) = \sigma^2 I$$

3. Estimating in the Parameter θ

The Least square principle is used to estimate the parameters in non-linear regression model. The residual sum of squares of $\hat{\theta}$ denoted in the matrix notation is

$$\text{SS (Res (}\hat{\theta}\text{))} = \left(Y - \eta\left(\hat{\theta}\right)\right)^I \left(Y - \eta\left(\hat{\theta}\right)\right),$$

Where $\eta\left(\hat{\theta}\right)$ the nx1 vector of $f\left(x_i, \hat{\theta}\right)$ is evaluated at that n values of x_n. The solution to the normal equation gives the least square estimate of θ. Then the General normal equation is written as

$$\frac{\partial\left(SS\left(\operatorname{Re} s\left(\hat{\theta}\right)\right)\right)}{\partial\hat{\theta}_j} = -\sum_{i=1}^{n}\left(y_i - \eta\left(\hat{\theta}\right)\right)\left[\frac{\partial\eta\left(\hat{\theta}\right)}{\partial\hat{\theta}_j}\right] = 0 \qquad (2.2)$$

Where $\eta(\theta) = f(x_i, \theta)$, i = 1, 2… N and $\left[\frac{\partial\eta\left(\hat{\theta}\right)}{\partial\hat{\theta}_j}\right]$ is the partial derivative of the functional form of the model.

The five important computational methods given in the literature for finding a solution to the system of normal equations are:

1) Newton-Raphson method
2) Steepest descent and ascent methods
3) Gauss-Newton method
4) Marquardt-Levenberg method
5) Derivative-free methods.

The first four methods require the partial derivatives of the model with respect to each of the parameters. The Fifth method, a Derivative-free method can be used in which numerical estimates of the derivatives are computed from observed shifts in $\hat{y}$ as the values of the $\hat{\theta}_j$ are changed.

4. Inference in Sure Model

SURE is based on the idea of a set of equations of the form:

$$y_i = X_i\beta_i + \varepsilon_i, i = 1, 2, \ldots, N \qquad (3.1)$$

Where y_i and ε_i are T-dimensional vectors, X_i is T x k_i and β_i is a k_i-dimensional vector.

The Set of equations may be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix} = \begin{bmatrix} X_1 & 0 & . & . & . & 0 \\ 0 & X_2 & . & . & . & 0 \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ 0 & 0 & . & . & . & X_N \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_N \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_N \end{bmatrix}$$ Or

$$y = X\beta + \varepsilon \quad (3.2)$$

Where β is a K-dimensional vector of unknown parameter that needs to be estimated and $k = \sum_{i=1}^{N} k_i$.

Assumptions:

1. E (ε) = 0
2. E ($\varepsilon_i \varepsilon_j^I$) = $\sigma_{ij} I_T$, i,j = 1,2, …, N

Where I_T is a T x T identity matrix. The assumption i = j gives the disturbance in any equation as homoscedastic and non autocorrelated. When $i \neq j$ the assumption gives a non-zero correlation between contemporaneous disturbance in the ith and jth equations. But zero, correlation between all lagged disturbances. Thus the full covariance matrix of ε is given by $\Omega = \sum \otimes I_T$ where $\sum = \left[\sigma_{ij} \right]$ is the N x N contemporaneous covariance matrix and $\otimes$ denotes the kronecker product. Each of the N equations is individually assumed to satisfy the classical assumption associated with the linear regression model and can be estimated separately.

5. Estimation of Sure Model

Generalized Least square estimator is defined as

$$\hat{\beta}(\Sigma) = \left[X^I (\Sigma^{-1} \otimes I_T) X \right]^{-1} X^I \left(\Sigma^{-I} \otimes I_T \right) y$$

Where

$$\Sigma^{-I} \otimes I_T = \begin{bmatrix} \sigma_{11} I & . & . & . & \sigma_{1N} I \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ \sigma_{N1} & . & . & . & \sigma_{NN} I \end{bmatrix} \quad (3.3)$$

Here $\otimes$ denotes Kronecker's product. The covariance matrix of $\hat{\beta}(\Sigma)$ given by

$$V\left(\hat{\beta}(\Sigma)\right) = \left(X^I (\Sigma^{-I} \otimes I_T) X \right)^{-I} \quad (3.4)$$

To construct a feasible GLS estimator by estimating each of the N equations separately by Ordinary Least squares when:

1) There is an absence of contemporaneous correlation $\left(\sigma_{ij} = 0, i \neq j \right)$ (or)

2) The same set of explanatory variables are included in each equation $\left(X_1 = X_2 = \ldots = X_N \right)$.

6. Feasible Generalized Least Squares Estimator

Consider the class of feasible GLS estimators that differ only in the choice of estimator used for the contemporaneous covariance matrix,

$$\hat{\beta}\left(\hat{\Sigma}\right)=\left[X^{I}(\hat{\Sigma}^{-1}\otimes I_{T})X\right]^{-1}X^{I}\left(\hat{\Sigma}^{-I}\otimes I_{T}\right)y \tag{3.5}$$

And inference is based on the estimator of the asymptotic covariance matrix of $\hat{\beta}\left(\hat{\Sigma}\right)$ given by

$$V\left(\hat{\beta}\left(\hat{\Sigma}\right)\right)=\left(X^{I}(\hat{\Sigma}^{-I}\otimes I_{T})X\right)^{-I} \tag{3.6}$$

An estimated covariance matrix obtained by using Ordinary Least squares residuals is $S=\left(S_{ij}\right)$ where $S_{ij}=\frac{u_i u_j^{\ I}}{T}$ and $u_i u_j^{\ I}=\left(y_i-X_i b_i\right)^{I}\left(y_i-X_j b_j\right)$ and b_i is the OLS estimator of β_i. S Has been referred to as the restricted estimator of $\sum$, but estimator can also be based on the unrestricted residuals derived from OLS regression which include all explanatory variables from the SUR system.

7. Inference in Nonlinear Sure Model Using Student zed Residuals

Generally, the specification of Linear Seemingly Unrelated Regression Equations (SURE) models may be unduly restrictive. In the estimation of the System of equations model underlying certain optimization problems, the Non-Linearities may be a feature of the model. Nonlinearities in the variables often may be eliminated by an appropriate transformation of the model or renaming the variables. In this situation, the model is inherently linear.

Consider the Nonlinear models which are inherently Nonlinear and the errors are additive as

$$\left.\begin{aligned} Y_{ti} = f_i\left(X_{ti},\theta_i\right)+\varepsilon_{ti}, i=1,2,\cdots,M \\ t=1,2,\cdots,n \end{aligned}\right\} \tag{4.1}$$

Where X_{ti} is a $\left(K_i \times 1\right)$ vector containing the t^th observation on each of the independent variables in the i^th equation; $\theta_i \in \Theta_i$ is a $\left(q_i \times 1\right)$ vector of unknown parameters in the i^th equation; and $f_i\left(\cdot\ ,\ \cdot\right)$ refers to the i^th dependent variable function. Thus, the Nonlinearities behavior may differ from equation to equation in the model.

In the matrix, more compactly, the model may be expressed as

$$Y=f\left(\theta\right)+\varepsilon \tag{4.2}$$

Such that $\mathrm{E}\left(\varepsilon\right)=0$ (4.3)

And $\mathrm{E}\left(\varepsilon\varepsilon^{I}\right)=\sum\otimes I_n$ (4.4)

Where, $Y=\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_M \end{bmatrix},\quad f\left(\theta\right)=\begin{bmatrix} f_1\left(\theta_1\right) \\ f_2\left(\theta_2\right) \\ . \\ . \\ . \\ f_M\left(\theta_M\right) \end{bmatrix},\quad \varepsilon=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_M \end{bmatrix}$

Here, $f_i\left(\theta_i\right)$ is a (nx1) vector with an element $f_i\left(X_{ti},\theta_i\right)$, θ is the $\left[\sum_{i=1}^{M} q_i \times 1\right]$ vector formed by stacking the elements of the θ_i's.

Further, f is a function of the data as well as of θ. Using the Nonlinear Ordinary Least Squares (NOLS) estimation and treating each equation of the system separately, the NOLS estimators for θ_i's ay be obtained by minimizing the quantity

$$\frac{1}{n}\left[Y_i - f_i\left(\theta_i\right)\right]^I\left[Y_i - f_i\left(\theta_i\right)\right], \quad \forall \theta_i \in \Theta_i \tag{4.5}$$

Let the NOLS estimator of θ_i as $\hat{\theta}_i, i = 1, 2, \cdots, M$. If the errors ε_i's follow the normal distribution then $\hat{\theta}_i$ will be the Maximum Likelihood Estimator of θ_i. By stacking the estimator $\hat{\theta}_i$'s the NOLS estimator of θ may be obtained as $\hat{\theta}$. Defining the NOLS residuals as $e_i = \left[Y_i - f_i\left(\hat{\theta}_i\right)\right]$ and hence the Studentized residuals for each equation as $e_i^*, i = 1, 2, \cdots, M$, One may obtain the estimates of the elements of $\sum$ as

$$\sum^* = \left(\left(\hat{\Sigma}_{ij}^*\right)\right) = \left(\left(S_{ij}^*\right)\right) = S^* \tag{4.6}$$

Where $S_{ij}^* = \dfrac{e_i^{*I} e_j^*}{n} \quad \forall i, j = 1, 2, \cdots, M$ (4.7)

The Feasible Generalized Least Squares (FGLS) estimator for θ can be obtained by minimizing

$$\frac{1}{n}\left[Y - f\left(\theta\right)\right]^I\left[S^{*^{-1}} \otimes I_n\right]\left[Y - f\left(\theta\right)\right] \tag{4.8}$$

With respect to θ over the parametric space

$$\Theta = \left[\Theta_1 \times \Theta_2 \times \cdots \times \Theta_M\right]$$

The minimization of (4.8) or (4.5) can be performed by using Numerical analysis methods.

Let the FGLS estimator of θ be $\hat{\theta}_F^*$. It can be easily shown that $\hat{\theta}$ and $\hat{\theta}_F^*$ will the strong consistent estimators for θ; and $\sqrt{n}\left(\hat{\theta} - \theta\right)$ and $\sqrt{n}\left(\hat{\theta}_F^* - \theta\right)$ follow asymptotically the multivariate normal distributions with the null mean vectors and the asymptotic variance-covariance matrices Ψ and Ω respectively, by using the methods given by Gallant (1975) [7].

Here, $\Psi = \begin{bmatrix} \Psi_{11} & . & . & . & \Psi_{1M} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ \Psi_{M1} & . & . & . & \Psi_{MM} \end{bmatrix}$ (4.9)

With $\Psi_{ij} = \sigma_{ij}\left[V_{ii}^{-1}V_{ij}V_{jj}^{-1}\right], \quad \forall i, j = 1, 2, \cdots, M$ (4.10)

And $\Omega = \begin{bmatrix} \sigma^{11}V_{11} & . & . & . & \sigma^{1M}V_{1M} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ \sigma^{M1}V_{M1} & . & . & . & \sigma^{MM}V_{MM} \end{bmatrix}$ (4.11)

Where V_{ij} is the $\left(q_i \times q_j\right)$ matrix which is given by

$$V_{ij} = \int \left[\frac{\partial f_i\left(X_i, \theta_i\right)}{\partial \theta_i} \right] \left[\frac{\partial f_j\left(X_j, \theta_j\right)}{\partial \theta_j} \right] d\mu_X \qquad (4.12)$$

If $\sum$ is diagonal and/or $f_i = f_j, \ \forall i, j = 1, 2, \cdots, M$ then the two estimators $\hat{\theta}$ and $\hat{\theta}_F$ are equally asymptotically efficient.

Otherwise, it can be shown that $\hat{\theta}_F$ is asymptotically efficient relative to $\hat{\theta}$. For the practical purposes, the consistent estimator of Ω is given by

$$\tilde{\Omega}^* = \frac{1}{n}\left[\hat{F}^I \left(S^{*^{-1}} \otimes I_n \right) \hat{F} \right]^{-1} \qquad (4.13)$$

Where, $\hat{F} = \text{diag}\left(\hat{F}_1, \hat{F}_2, \cdots, \hat{F}_M\right)$ (4.14)

And $$\hat{F}_i = \begin{bmatrix} \frac{\partial f_i\left(X_{1i}, \theta_i\right)}{\partial \theta_i^I} \\ . \\ . \\ . \\ \frac{\partial f_i\left(X_{ni}, \theta_i\right)}{\partial \theta_i^I} \end{bmatrix}_{\theta = \hat{\theta}_F} \qquad (4.15)$$

Under the assumption of normality distribution of errors $\varepsilon_i\text{'s}$, it can be shown that $\hat{\theta}_F^*$ is the Nonlinear Maximum Likelihood estimator of θ

Remarks: 1. the tests of General Nonlinear restrictions on the parameters can be carried out by generalizing the applications of usual tests on linear restrictions about the parameters such as Wald, Lagrange Multiplier and Likelihood Ratio tests.

2. Since, the estimators $\hat{\theta}$ and $\hat{\theta}_F^*$ cannot be expressed in closed form, generally their finite sample properties are unknown. Gallant (1975) [7] exhibited a limited amount of Monte Carlo evidence for the finite sample properties of the estimators.

8. Conclusions

The Classical Statistical models and its various extensions have perhaps been the most widely used technique in analyzing the research problems in different fields of Applied Science. Inferential problem in statistical model is an old and important problem in Statistics as well as in any other field that uses applied regression analysis.

In the present research study, the nonlinear SURE model has been specified and a feasible GLS estimator for the parametric vector has been developed along with its asymptotic variance covariance matrix. Further, the consistent estimator of the variance covariance matrix has been derived and it will be the nonlinear maximum likelihood estimator.

9. Reference

1. Agostinelli C. Robust model Selection in Regression via weighted Likelihood methodology, Statistics and Probability Letter. 2002; 56:289-300.
2. Balasiddamuni P. Advanced Tools for Mathematical and Stochastic Modelling", Proceeding of the International Conference on Statistical Modelling and Simulation, Allied Publishers, 2011, 21-26.
3. Baltagi BH. Econometric Analyses of Panel Data, John Wiley and Sons, 2nd edition, 2001.
4. Baraud Y, Huet S, Laurent B. Adaptive Tests of Linear Hypotheses by Model Selection, the Annals of Statistics. 2003; 31(1):225-251.

5. Durga Prasad S, Balasiddamuni P, Ramesh Mummineni. Statistical Inference in Time Series Regression Models, LAP, Lambert Academic Publishers, Germany, 2013.
6. Efron B, Johnstone I, Hastie T, Tibshirani R. Least angle regression, Ann. Statist. 2004; 32:407-499.
7. Gallant AR. Seemingly Unrelated Nonlinear Regressions, Journal of Econometrics. 1975; 3:35-50.
8. Gallant AR. Nonlinear Statistical Models, Wiley, New York, 1987.
9. Hari Babu O, Balasiddamuni P, Ramana Murthy B. Inferential Aspects of Regression Models, LAP, Lambert Academic Publishers, Germany, 2013.
10. Kadane JB, Lazar NA. Methods and Criteria for Model Selection, Journal of the American Statistical Association. 2004; 99(465):279-290.
11. Narayana P, Balasiddamuni P, Subbarami Reddy C. Statistical Inference in Sets of Linear Regression Models, LAP, Lambert Academic Publishers, Germany, 2013.