[1] Pagadala Srivyshnavi

[2] Dr. D. Aju*

SURE Model as an Advanced Data Science Technique

Journal of Electrical Systems

Abstract: - Seemingly unrelated regression equations (SURE) model and it's associated inferential aspects have been generating substantial applications in various fields such as Statistics, Advanced Econometrics, Data Science Techniques; Business, Management and Marketing fields; physical sciences and Engineering etc. Among the regression based data science techniques, a few data engineers have applied SURE models in their data analysis. Researchers can use SURE techniques as advanced data science techniques in the fields of electrical systems, computer engineering and other areas of engineering and technology. The classical SURE model deals with the sets of linear regression equations by which establishing relationships among the sets of dependent variables and explanatory variables. Several advanced feasible estimation methods exist in the literature used either OLS or GLS residuals in their estimation. In the present research study, due to shortcomings of these residuals, new iterative feasible OLS and feasible GLS estimators have been proposed to estimate the parameters of SURE model with nonspherical first order vector autoregressive errors by using studentized residuals.

Keywords: Nonspherical VAR errors, Data science techniques, First order Autoregressive errors and Studentized residuals

1. Introduction

SURE model has both theory and practical applications in Applied Mathematics, Statistics, Econometrics, Engineering and various fields of Applied Sciences. There are mainly two types of SURE model namely, Linear SURE model and Nonlinear SURE model. Zellner introduced linear SURE model in which each of linear model is correlated with each other, even though superficially they may not seem to be. Statistical inference in these models is an advanced fertile area of research to the data engineers, statisticians and econometricians. At present, most of the regression based data science techniques are using basic applied regression analysis methods.

There is an urgent need of applications of both linear and nonlinear advanced SURE models to cater the needs of Data Scientists, Artificial Intelligence Engineers and Business Intelligence Engineers.

A few data scientists have been using certain types of basic linear SURE model. In the context of time series data, generally the errors are assumed to be generated by some autoregressive schemes. Further, the cross errors intertemporal correlations can be considered by using vector autoregressive schemes for errors in the SURE model.

An Iterative method of estimation of Linear SURE model with nonspherical Vector Autoregressive (VAR) errors has been developed in the proposed work. This method has various applications in electric networks of electrical systems.

2. Related Work

Zeller systematically discussed in his seminal research papers about linear SURE model. A feasible iterative generalized least squares (GLS) estimation was also proposed for model [1,2]. Mehta and Gibber have estimated model with vector autoregressive schemes for errors by using GLS estimation [3]. Srivastava and Giles have considered almost all problems of model together with different estimation methods [4]. Nagabhushana Rao presented various forms of SURE models [5].

Kakwani proved the unbiasedness and efficiency properties of Zellner's estimation [6,7]. Narayana discussed about the inferential aspects of sets of linear models [8]. Several researchers have developed different estimation methods for the model involving serially and contemporaneously correlated disturbances [9,10,11,12].

Certain problems of modelling such as non-normal disturbances and nonspherical errors respectively considered by Olamide and Sireesha with their proposed estimation methods [13,14]. Iterative efficient estimation methods for model

have discussed in research papers [15,16]. Margolin defined studentized residuals with their distribution via Laplas transform inversion [17]. Philips derived exact distribution of SURE estimator [18]. Hemmingren and Hamann have given a software package 'system fit' for estimating the system of simultaneous equations in R-Language [19]. Grigoras Gheorghe and Bogdan Neagu have proposed regression based load modelling for electric distribution networks [20].

[1] Research Scholar Email Id: vyshu.sri@gmail.com

[2]*Associate Professor, Department of Computer Science & Engineering VIT, Vellore, India – 632014
Email Id: daju@vit.ac.in

3. SURE model

With usual matrix notation, a linear system containing m equations which may be expressed in the compact form:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix} = \begin{bmatrix} X_1 & 0 & .... & 0 \\ 0 & X_2 & .... & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & X_m \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix} \tag{3.1}$$

$$\Rightarrow Y_{mnx1} = X_{mnxk^*}\beta_{k^*x1} + \epsilon_{mnx1} \tag{3.2}$$

Where β is k^*x1 regression coefficients;

$\in$ is mnx1 unobservable disturbances and $k^* = \sum_{i=1}^{m} K_i$

Also, ϵ_i is such that $E(\epsilon_i) = 0, (\epsilon_i \epsilon_j^T) = \sigma_{ij} I_n, \forall i, j = 1, 2, ...m$ (3.3)

$$\Rightarrow E(\epsilon) = 0 \text{ and } E(\epsilon\epsilon^T) = \begin{bmatrix} \sigma_{11}I_n & \sigma_{12}I_n & \cdots & \sigma_{1m}I_n \\ \sigma_{21}I_n & \sigma_{22}I_n & \cdots & \sigma_{2m}I_n \\ \vdots & & & \\ \sigma_{m1}I_n & \sigma_{m2}I_n & \cdots & \sigma_{mm}I_n \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & & & \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mm} \end{bmatrix} \otimes I_n$$

or $E(\epsilon\epsilon^T) = \sum \otimes I_n = \Psi_{mnxmn}$ and $\sum$ is mxm symmetric positive definite matrix. (3.4)

Here, $E(\epsilon_i \epsilon_i^T) = \sigma_{ij} I_n$ implies the disturbance in any single regression equation as homoscolastic and non autocorrelated.

$E(\epsilon_i \epsilon_j^T) = \sigma_{ij} I_n, \forall i \neq j$ gives a non zero correlation between contemporaneous disturbances in the i^{th} and j^{th} equations but all lagged disturbances are uncorrelated with each other; and the symbol $\otimes$ denotes Kronecker product which gives that each element in $\sum$ is multiplied by Identity matrix I_n.

The system of linear regression equations with the aforementioned assumptions is known as “Seemingly Unrelated Regression Equations (SURE) model”.

Remark:

If $X_1 = X_2 = ...X_m$ and Rank (X_i) = K, $\forall$ i=1,2,….m then the SURE model will be coincide with the standard multivariate linear regression model.

4. Some important estimators of parametric vector of SURE model with the variance-covariance matrices

Consider the SURE model (3.2) as $Y_{mnx1} = X_{mnxk^*}\beta_{k^*x1} + \epsilon_{mnx1}$

Such that $E(\epsilon) = 0$ and $E(\epsilon\epsilon^T) = \Psi = \Sigma \otimes I_n$ (4.1)

$$\text{Where } \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \cdots & \sigma_{mm} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & & & \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mm} \end{bmatrix} \tag{4.2}$$

(a) The ordinary least squares (OLS) estimator:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y \tag{4.3}$$

$$Var(\hat{\beta}_{OLS}) = (X^T X)^{-1}[X^1\left(\Sigma^{-1} \otimes I_n\right)X](X^T X)^{-1} \tag{4.4}$$

(b) Aitken’s estimator or GLS estimator

$$\tilde{\beta}_{GLS} = \left[X^T\left(\Sigma^{-1} \otimes I_n\right)X\right]X^T\left(\Sigma^{-1} \otimes I_n\right)Y \tag{4.5}$$

$$Var(\tilde{\beta}_{GLS}) = [X^T(\Sigma^{-1} \otimes I_n)x]^{-1} \tag{4.6}$$

Remarks:

(i) If Σ is diagonal matrix, then $\tilde{\beta}_{GLS}$ reduces to $\hat{\beta}_{OLS}$

(ii) If $X_1 = X_2 = ...X_m$ =X for $K_i = K, \forall i = 1, 2, ...m$ and Rank(X) =K, then the $\tilde{\beta}_{GLS}$ reduces to $\hat{\beta}_{OLS}$

(iii) From the Generalized Gauss-Markoff Theorem (Aitken's theorem) when Σ is known non-stochastic matrix, it

follows that $\tilde{\beta}_{GLS}$ is b.l.u.e. for β.

(c) Zellner's Feasible GLS estimator:

Generally, Σ is unobservable. Zellner replaced Σ with an observable sample variance-covariance matrix S_{mxm} and proposed the Zellner's feasible GLS estimator for β as

$$\tilde{\beta}_{FGLS} = [X^T(S^{-1} \otimes I_n)x]^{-1} X^T(S^{-1} \otimes I_n)Y \tag{4.7}$$

$$\operatorname{Var}(\tilde{\beta}_{FGLS}) = [X^T(S^{-1} \otimes I_n)x]^{-1} \tag{4.8}$$

Where S= $((S_{ij}))$ is a non-singular sample variance-covariance matrix such that S_{ij} is some estimator of the corresponding element σ_{ij} of Σ. For many possible choices of S, various types of feasible GLS estimators have been proposed by several statisticians, Mathematicians and Econometricians. Most of these feasible GLS estimators are based on OLS and GLS residuals.

Some important feasible GLS estimators for β in SURE model are given by: FGLS estimators, Momentless FGLS estimators, The Telser-Conniffe estimator, Iterative Zellner's FGLS estimator, Iterative OLS estimator, Maximum likelihood estimator (MLE), Pretest and Stein-Rule estimators, corrected FGLS estimators and restricted least squares estimators,

Remark: By choosing S= I_m, then Zellner's FGLS estimator. $\tilde{\beta}_{FGLS}$ reduces to OLS estimator $\hat{\beta}_{OLS}$ for β in SURE model.

5. Estimation of SURE model with nonspherical first order vector autoregressive errors

Suppose that the SURE model containing m linear regression equations with usual matrix notation is given by

$$Y_{mn\times 1} = X_{mn\times k}\beta_{k\times 1} + \in_{mn\times 1} \quad \text{such that } E(\in) = 0 \tag{5.1}$$

We assume that the elements $\in_{si}$'s of error vector $\in$ are contemporaneously correlated. Here, $\in_{si}$ is s^{th} element of i^{th} error vector $\in_i$ in the i^{th} equation of the SURE model.

Define the first order stationary autoregressive error process as

$$\in_{si} = \rho_i \in_{(s-1)i} + u_{si}, \qquad s = 1,2,\ldots,n \text{ and } i = 1,2,\ldots,m \tag{5.2}$$

Here $|\rho_i| <|$ is a constant and u_{si} is an error random variable such that

$$E(u_{si}) = 0. \forall \text{ s and i}, \quad E(u_{si}\, u_{tj}) = \sigma_{ij} \text{ for s=t and } \forall i, j,$$
$$= 0, \text{ for } s \neq t \text{ and } \forall \text{ i, j.} \tag{5.3}$$

Equation (4.2) reveals that $\in_{si}$ depends stochastically upon only the preceding error term $\in_{(si-1)i}$. Also $\in_{si}$'s and u_{si}'s are contemporaneously correlated but the $\in_{si}$'s are no longer temporarily independent. It should be noted that u_{si}'s are temporarily independent. By defining U as an mnx1 vector as $\in$ is defined,

we have, E(U)=0 and $E(UU^T) = \Sigma \otimes I_n$ (5.4)

and hence $E(\in) = 0$ and $E(\in\in^T) = \Omega = ((\Omega_{ij}))_{m\times n}$ (5.5)

$$\text{Where } \Omega_{ij} = \frac{\sigma_{ij}}{(1-\rho_i\rho_j)} \begin{bmatrix} 1 & \rho_j & \cdots & \rho_j^{n-1} \\ \rho_i & 1 & \cdots & \rho_j^{n-2} \\ \vdots & & & \\ \rho_i^{n-1} & \rho_i^{n-2} & \cdots & 1 \end{bmatrix} \tag{5.6}$$

Without loss of generality, the first order vector autoregressive process for the SURE model's errors in matrix notation as

$$\begin{bmatrix} \in_{s1} \\ \in_{s2} \\ \vdots \\ \in_{sm} \end{bmatrix} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2m} \\ \vdots & & & \\ \rho_{m1} & \rho_{m2} & \cdots & \rho_{mm} \end{bmatrix} \begin{bmatrix} \in_{(s-1)1} \\ \in_{(s-1)2} \\ \vdots \\ \in_{(s-1)m} \end{bmatrix} + \begin{bmatrix} u_{s1} \\ u_{s2} \\ \vdots \\ u_{sm} \end{bmatrix} \tag{5.7}$$

or $\in_{(s)} = \Gamma \in_{(s-1)} + u_{(s).}$

Here, Γ is an m x m matrix of unknown autoregressive parameters and the u_{si} 's are errors assumed to have the same properties as in (5.3).
If the absolute values of all the characteristic roots of Γ are les than unity, then the vector autoregressive process is a stationary process.
Further, we have, $E(\in_{(s)}\in_{(s)}^{T}) = \Psi$, which can be written as $\Psi = (\Gamma\Psi\Gamma^{T}) + \Sigma$ (5.8)
Here, Ψ is symmetric matrix.
If $\in(\in\in^{T}) = \Omega$, then there exists a non-singular matrix G_{mnxmn} such that
$\Omega^{-1} = G^{T}(\Sigma^{-1}\otimes I_n)G$
By solving equations (5.8), the elements of Ψ can be obtained in terms of Γ and Σ .
Using transforming matrix G, the transformed SURE model can be written as

$$GY = Gx\beta + G\in \text{ or } Y^{*} = X^{*}\beta + \in^{*} \quad (5.9)$$

such that $E(\in^{*}) = 0$ and $E\left(\in^{*}\in^{*T}\right) = \sum\otimes I_n$

Where, Y^*=GY, X^*=GX and $\in^{*} = G\in$
Now, for a choice of G = G^*, we can directly obtain estimators for β by using OLS and FGLS estimators:

(i) $\hat{\beta}^{*}_{OLS} = (X^{*T}X^{*})^{-1}X^{*T}Y$ or $\hat{\beta}^{*}_{OLS} = (X^{T}G^{*^{T}}G^{*}X)^{-1}X^{T}G^{T}G^{*}Y$ (5.10)

(ii) $\tilde{\beta}^{*}_{FGLS} = [X^{*T}(S^{-1}\otimes I_n)x^{*}]^{-1}X^{*T}(S^{-1}\otimes I_n)Y^{*}$ or

$\tilde{\beta}^{*}_{FGLS} = [X^{T}G^{*T}(S^{-1}\otimes I_n)G^{*}X]^{-1}X^{T}G^{T}(S^{-1}\otimes\sum)G^{*}Y$ (5.11)

where $S = ((S_{ij}))_{mxm}$ which is sample variance-covariance matrix based on OLS residual vectors.
Here, G^* is a particular choice of G, a block-diagonal matrix which was suggested by parks [12] as

$$G^{*} = \begin{bmatrix} G^{*}_{11} & 0 & .... & 0 \\ 0 & G^{*}_{22} & .... & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & G^{*}_{mm} \end{bmatrix} \text{ and } G^{*}_{ii} = \begin{bmatrix} \frac{1}{\sqrt{1-\rho_i^2}} & 0 & 0 & ... & 0 & 0 \\ -\rho_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho_i & 1 \end{bmatrix} \quad i=1,2,\ldots,m \quad (5.12)$$

where ρ_i 's are unknown autoregressive parameters.
Since, ρ_{ij}'s of matrix G^* and σ_{ij}'s of matrix $\sum$ are unknown, several researchers have used some consistent estimators for ρ_{ij}'s and σ_{ij}'s ; hence obtained the Feasible GLS estimators for β in the SURE model. Most of them have used either OLS or GLS residuals in estimating ρ_{ij}'s and σ_{ij}'s .
In the present study, an Iterative estimation procedure has been proposed by using any of studentized residuals or predicted residuals in estimating ρ_{ij}'s and σ_{ij}'s to obtain the Feasible GLS estimators for β in SURE Model.
Step (1): First obtain the OLS residuals ρ_{ij}'s by estimating the original SURE model and then obtain the corresponding Internally and Externally studentized residuals $e_{ij}(I)$'s and $e_{ij}(E)$'s respectively by using the following usual formulae given by Margolin:

(i) $$e_{si}(I) = \frac{e_{si}}{\hat{\sigma}^{i}_{s}\sqrt{1-h^{i}_{ss}}}, \quad s= 1,2,\ldots, n \quad (5.13)$$

$$\text{and } \hat{\sigma}^{i}_{s} = \sqrt{\frac{\sum_{s=1}^{n} e^{2}_{si}}{n-k}}, \quad i=1,2,\ldots, m \quad (5.14)$$

Here, h^{i}_{ss} is the s^{th} diagonal element of HAT matrix for the i^{th} linear regression equation of SURE model and

(ii) $$e_{si}(E) = e_{si}(I)\left[\frac{n-k_i-1}{n-k_i-e^{2}_{si}(I)}\right]^{1/2}, \quad i=1,2,\ldots m \quad (5.15)$$

Step (2): Substitute the values of the estimates $e_{si}(E)$'s and $e_{(s-1)i}(E)$'s respectively for $\in_{(s)}$ and $\in_{(s-1)}$ in the equation (5.7) and then the OLS estimates of the elements ρ_{ij}'s of the matrix Γ can be obtained.

Step (3): The values of OLS estimates of ρ_{ij}'s can be substituted and then the new transformation matrix $\hat{G}^*$ can be obtained. Replace G by $\hat{G}^*$ in the transformed model (5.9) and then the transformed SURE model can be estimated by using OLS estimation.

Step (4): Repeating the steps (1) and (2), the second Iteration estimators of the elements of Transformation matrix G can be obtained. Later, the second Iteration estimated G say $\hat{\hat{G}}^*$ can be substituted in the model (5.9) and perform step (3). This Iterative process will be continued. It will be stopped after obtaining successive estimates of the $\rho_{ij}^{'s}$ of G differ by less than some prescribed small quantity.

The final iteration estimated G say $G^*(F)$ can be substituted in the model (5.9) and write the transformed model as

$$Y^{(F)}_{nm\times 1} = X^{(F)}_{mn\times k}\beta + \in^{(F)}_{mn\times 1} \tag{5.16}$$

where, $Y^{(F)} = G^*(F)Y, X^{(F)} = G^*(F)X$ and $\in^{(F)} = G^*(F)\in$ is the final SURE model.

Further, sample variance-covariance matrix $S = ((S_{ij}))$ can be modified by using the externally studentized residuals sum of squares in the place of the usual OLS residual sum of squares. Let the modified S be denoted by S(F).
Now, the proposed Iterative OLS (IOLS) and Iterative FGLS (IFGLS) estimators for β in the SURE model are given by:

(i) $$\hat{\beta}^{(F)}_{IOLS} = \left(X^T G^{*T}(F) G^*(F) X\right)^{-1} X^T G^{*T}(F) G^*(F) Y \tag{5.17}$$

and (ii) $$\tilde{\beta}^{(F)}_{IFGLS} = \left[X^T G^{*T}(F)\left(S^{-1}(F)\otimes I_n\right)G^*(F)X\right]^{-1} X^T G^{*^T}(F)\left(S^{-1}(F)\otimes I_n\right)G^*(F)Y \tag{5.18}$$

Under certain regularity conditions, these iterative estimators will be consistent for β in the SURE model.

6. Conclusion

In the present research article, Zellner's Linear SURE model has been specified in a systematic matrix notation and some important estimators such as OLS, GLS and Zeller's Feasible GLS estimators for the parametric vector of SURE model have been given along with their variance – covariance matrices. The SURE model with nonspherical first order vector autorregressvie errors has been difined and an Iterative method has been developed and obtained Iterative FOLS and Iterative FGLS estimators by using externally studentized residuals. Under certain regularity conditions, these iterative estimators will give consistent estimators. The proposed iterative method can be extended by applying it to the nonlinear SURE model. Several statistical package programs such as R, SAS, Stata, Limdep, Python and Greti etc., are available to implement SURE estimation methods by data Scientists, AI and BI engineers. The Linear SURE model can be specified by considering multiple electrical outputs and multiple electrical inputs in electrical systems. Then the model can be estimated by using the proposed Iterative FGLS estimation, applying the statistical package programmes.

References

[1] A.Zellner, An efficient method of estimating seemingly unrelated regression equations and test for aggregation bias, Journal of Statistical Association, 57 (1962) 348-368.
[2] Z.Zellner, Estimators for SUREs : some exact finite sample results, Journal of American Statistical Association, 58 (1963) 977-992.
[3] J. K.Menta and R.F. Gibber, Estimation of Seemingly unclated regressions with autoregressive disturbances, Journal of American Statistical Association, 65 (1970) 186-197.
[4] Virendra K. Srivastava and David E.A. Giles, Seemingly unrelated regression equations models, Research Gale, ISBN: 9781000105728, August 2020.
[5] R.V.S.S. Nagabhushana Rao, P. Balasiddamuni and B. Ramana Murthu, Inference in SURE models, LAP Publishers, New Germany, 2013.
[6] N.C. Kakwani, The unbiasedness of Zellner's seemingly unrelated regression equations estimators, Journal of the American Statistical Association, 62 (1967), 141-142.
[7] N.C. Kakwani, A Note on the efficiency of Zellner's seemingly unrelated regressions estimators, Annals of the Institute of Statistical Mathematics, 26 (1974) 361-362.
[8] P. Narayana, P. Balasiddamuni and C.S. Reddy, Statistical inference in sets of linear regression models, LAP publishers, New German, 2013.
[9] P. Foschi, Estimating seemingly unrelated regression models for vector autoregressive disturbances, Journal of Economic Dynamics and Control, (2003) 27-44.
[10] Ye Chen, Jian Li and Quyuan Li, SUR estimation for VAR models with explosive roorts, Oxfored Bulletin of Economics and Statistics, 84(4) (2023), DOI: 10.1111/obes.12551.

[11] D.K. Guilkey and P. Schmidt, Estimation of seemingly unrelated regression with vector autoregressive errors, Journal of American Statistical Association, 68 (1973), 642-647.

[12] R.W. Parks, Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated, Journal of American Statistical Association, 62 (1967) 500-509

[13] E.I. Olamide, Estimation of Seemingly unrelated regression equations with non-normal disturbances, International Journal of Social Relevance of concern, 6(11) (2018) 1-5, DOI<10.26821/itsrc.6.11.2018.61 101.

[14] B. Sireesha, Statistical inference in sets of linear regression models with Nonspherical Disturbances, unpublished Ph.D., Thesis, Dept. of Statistics, S.V. University, Andhra Pradesh, India, 2016.

[15] L.G. Telser, Iterative estimation of a set of linear regression equations, Journal of American Statistical Association, 59 (1964) 845-862.

[16] L.Wang; H. Lian and R.S. Singh, On efficient estimators of seemingly unrelated regressions, Statist. Prob. Letters, 81 (2011) 563-570.

[17] B.H. Margolin, The distribution of internally studentized statistics via Laplas transform inversion, Biometrica, 64 (1977) 573-582.

[18] P.C.B. Philips, The exact distribution of the SUR estimator, Econometrica, 53 (1985) 745-756.

[19] A. Henningren and J.D. Hamann, System fit: A package for estimating system of simultaneous equations in R, Journal of Statistical Software, 23(4) (2007) 1-40. URI. http://www. Istatsoff.org/v23/104/.

[20] Grigoras Gheorghe and Bogdan Neagu, Regression analysis – based load modelling for electric distribution networks, in book : Numerical methods for energy applications, Edition I, Springer, Chapter 28, (2021) 769-794.