연구논문집 제10권 제 2 호 p. 125~127 1979

Note on Generalized Least Squares with Stepwise Estimation Procedure

Yeo, Sung-Chil
Dept. of General Education

〈Abstract〉

The generalized least squares (GLS) method for estimating the parameters of a linear regression model $y=X\beta+e$ with $E[e]=0$ and $E[ee']=V\sigma^2$ is commonly used. When the further regressors x_j's are introduced for the given GLS model $y=X\beta+e$, the GLS estimators for the parameters are obtained by direct calculation from the normal equation. While the GLS estimators may also be obtained by the stepwise estimation procedure.

단계적 추정방법으로 일반화한 최소자승추정에 관한 고찰

余 成 七
교 양 과 정 부

〈요 약〉

일반화한 선형회귀모형에서 새로운 변수가 첨가되었을때 회귀계수의 일반화한 최소자승 추정치를 단계적 추정치(stepwise estimator)로 나타내었고, 첨가된 새변수에 대한 검정기준을 제시하였다.

I. Introduction

Suppose after fitted the GLS model

$$E[y]=X\beta,\ \mathrm{Var}[y]=V\sigma^2 \qquad (1.1)$$

we decide to introduce further regressors x_j's into the model so that the model(1.1) is now enlarged to

$$E[y]=X\beta+Z\gamma$$
$$=W\delta^{(1)} \qquad (1.2)$$

where

- y is a $n\times1$ vector of observations on the dependent variable,
- X is a $n\times p$ matrix of observations on the first regressors $x_i(i=1,2,\cdots,p)$ with rank p,
- β is a $p\times1$ vector of parameters,
- Z is a $n\times q$ matrix of observations on the second regressors $x_j(j=1,2,\cdots,q)$ with rank q,
- γ is a q$\times$1 vector of parameter,
- W is a $n\times(p+q)$ matrix of observations on the whole regressors $x_k(k=1,2,\cdots,p-q)$ with rank $p-q$,
- δ is a $(p+q)\times1$ vector of parameters,
- V is the known positive definite $n\times n$ matrix.

Since V is positive definite, there exists an $n\times n$ nonsingular matrix P such that $V=PP'$.[2] If we pre-multiply both sides of the model (1.2) by a matrix P^{-1}, then the model (1.2) is transformed into

(1) Here we have set $W=[X,Z]$ and $\delta=\begin{bmatrix}\beta\\ \gamma\end{bmatrix}$
(2) See Seber [5], pp. 385—386.

$$E[y^*] = X^*\beta + Z^*\gamma = W^*\delta, \text{[3]} \qquad (1.3)$$
$$\mathrm{Var}[y^*] = \sigma^2 I$$

Thus the model (1.3) is reduced to an ordinary least squares (OLS) model. Hence we can apply the OLS method to the transformed model (1.3). And so the GLS estimators $\hat{\beta}_G$ and $\hat{\gamma}_G$ for parameters are obtained as

$$\hat{\delta}_G = [W^{*\prime}W^*]^{-1}W^{*\prime}y^* = [W'V^{-1}W]^{-1}W'V^{-1}y$$

, which may be partitioned as

$$\begin{bmatrix}\hat{\beta}_G \\ \hat{\gamma}_G\end{bmatrix} = \begin{bmatrix} X^{*\prime}X^* & X^{*\prime}Z^* \\ Z^{*\prime}X^* & Z^{*\prime}Z^*\end{bmatrix}^{-1}\begin{bmatrix}X^{*\prime}y^* \\ Z^{*\prime}y^*\end{bmatrix} = \begin{bmatrix} X'V^{-1}X & X'V^{-1}Z \\ Z'V^{-1}X & Z'V^{-1}Z\end{bmatrix}^{-1}\begin{bmatrix}X'V^{-1}y \\ Z'V^{-1}y\end{bmatrix}$$

On the other hand, the first stepwise estimator $\tilde{\beta}$ is obtained when one ignores γ , and the second estimator $\tilde{\gamma}$ is obtained by regressing the estimated residuals $\tilde{y}(=y^* - X^*\tilde{\beta})$ upon the variable Z. Therfore we have $\tilde{\gamma}$ and $\tilde{\beta}$ such that

$$\tilde{\beta} = [X^{*\prime}X^*]^{-1}X^{*\prime}\tilde{y}^* = [X'V^{-1}X]^{-1}X'V^{-1}y$$

and

$$\tilde{\gamma} = [Z^{*\prime}Z^*]^{-1}Z^{*\prime}\tilde{y}^* = [Z^{*\prime}Z^*]^{-1}Z^*[y^* - X^*\tilde{\beta}] = [Z'V^{-1}Z]^{-1}Z'V^{-1}[y - X\tilde{\beta}]$$

Although $\hat{\beta}_G$ and $\hat{\gamma}_G$ are unbiased estimators of β and γ, the stepwise estimators $\tilde{\beta}$ and $\tilde{\gamma}$ are biased. Our purpose here is to give the expressions of the GLS estimators $\hat{\beta}_G$ and $\hat{\gamma}_G$ by the stepwise estimators $\tilde{\beta}$ and $\tilde{\gamma}$. In addition, we give the test criterion to test the hypothesis $\gamma = 0$ for parameters of the newly introduced variables in the enlarged GLS model. The test criterion is given by assuming the normality of the model.

We have given here the results for a two stepwise procedure, the results for a multistepwise procedure follow on similar fashion.

II. The connection between the GLS and the stepwise estimators

From the equation (1.4) we have that

$$\begin{bmatrix}X'V^{-1}y \\ Z'V^{-1}y\end{bmatrix} = \begin{bmatrix} X'V^{-1}X & X'V^{-1}Z \\ Z'V^{-1}X & Z'V^{-1}Z\end{bmatrix}\begin{bmatrix}\hat{\beta}_G \\ \hat{\gamma}_G\end{bmatrix}$$

, which we write as

$$\begin{bmatrix}X'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y \\ Z'V^{-1}Z(Z'V^{-1}Z)^{-1}Z'V^{-1}(y - X\tilde{\beta} + X\tilde{\beta})\end{bmatrix} = \begin{bmatrix} X'V^{-1}X & X'V^{-1}Z \\ Z'V^{-1}X & Z'V^{-1}Z\end{bmatrix}\begin{bmatrix}\hat{\beta}_G \\ \hat{\gamma}_G\end{bmatrix}$$

It follows that

$$\begin{bmatrix}X'V^{-1}X\tilde{\beta} \\ Z'V^{-1}Z\tilde{\gamma} + Z'V^{-1}X\tilde{\beta}\end{bmatrix} = \begin{bmatrix}X'V^{-1}X\hat{\beta}_G + X'V^{-1}Z\hat{\gamma}_G \\ Z'V^{-1}X\hat{\beta}_G + Z'V^{-1}Z\hat{\gamma}_G\end{bmatrix}$$

From the equation(2.1)

$$\tilde{\beta} = \hat{\beta}_G + [X'V^{-1}X]^{-1}X'V^{-1}Z\hat{\gamma}_G$$

and

$$\tilde{\gamma} = [I - R^2]\hat{\gamma}_G$$

and where $R^2 = [Z'V^{-1}Z]^{-1}Z'V^{-1}X \times [X'V^{-}X]^{-1}X'V^{-1}Z$[4]

Or equivalently we have

$$\hat{\beta}_G = \tilde{\beta} - [X'V^{-1}X]^{-1}X'V^{-1}Z\hat{\gamma}_G$$
$$\hat{\gamma}_G = [I - R^2]^{-1}\tilde{\gamma} \qquad (2.2)$$

Therefore the equation (2.2) gives the expressions by the stepwise estimators $\tilde{\beta}$ and $\tilde{\gamma}$ for the GLS estimators $\hat{\beta}_G$ and $\hat{\gamma}_G$.

III. A test of the hypothesis $H_0 : \gamma = 0$

The model(1.2) may be expressed in the form

$$y = X\beta + Z\gamma + e = W\delta + e \qquad (3.1)$$

, where e is distributed $N(0, V\sigma^2)$.[5]

And from the expression (1.3) the GLS model (3.1) may also be written as

$$y^* = X^*\beta + Z^*\gamma + e^* = W^*\delta + e \qquad (3.2)$$

(3) Here we have set $y^* = P^{-1}y$, $X^* = P^{-1}X$, $Z^* = P^{-1}Z$ and $W^* = [X^*, Z^*]$.

(4) Here R^2 may be interpreted as the multiple correlation coefficient between the values of the matrix X and the values of Z in the model (1.2).

(5) Here we have assumed the normality of the disturbance terms e to give the the test criterion for hypothesis $r = 0$.

, where e^* is distributed $N(0, \sigma^2 I)$.

Since the model (3.2) satisfies the usual assumptions of an OLS model, we can apply the theory of an OLS model into the transformed model (3.2). Therefore we have the following theorem obtained from the model (3.2) by reducing the theory into the original model (3.1).(6)

Theorem: Let the model (3.1) be given as

$$y = X\beta + Z\gamma + e$$
$$= W\delta + e$$

, where $e \sim N(0, V\sigma^2)$ and where V is the known positive definite $n \times n$ matrix. If we set

$$S_E = y'V^{-1}y - \hat{\delta}_G'W'V^{-1}y$$

and $$S_2 = \hat{\delta}_G'W'V^{-1}y - \tilde{\beta}'X'V^{-1}y$$

, then the quantity u given by

$$u = \frac{n-p-q}{q}\ \frac{S_2}{S_E}$$

is noncentral F distributed as $F'(q, n-p-q, \lambda)$

, where $$\lambda = \frac{1}{2\sigma^2}\gamma' M \gamma$$

and where

$$M = X'V^{-1}X - X'V^{-1}Z[Z'V^{-1}Z]^{-1}Z'V^{-1}X$$

Since M is positive definite, u is distributed as $F(q,\ n-p-q)$ when the hypothesis $H_0 : \gamma = 0$ is true.

On the other hand, from the results of the theorem we may have the following the analysis of variance table for testing the hypothesis $\gamma = 0$.

Analysis of Variance for testing $H_0 : \gamma = 0$

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Total	n	$y'V^{-1}y$		
Due to β, γ	$p+q$	$\hat{\delta}_G'W'V^{-1}y$		
Due to β (unadj)	p	$\tilde{\beta}'X'V^{-1}y$		
Due to γ (adj)	q	$\hat{\delta}_G'W'V^{-1}y - \tilde{\beta}'X'V^{-1}y = S_2$	$\frac{S_2}{q}$	$\frac{n-p-q}{q}\ \frac{S_2}{S_E}$
Error	$n-p-q$	$y'V^{-1}y - \hat{\delta}_G'W'V^{-1}y = S_E$	$\frac{S_E}{n-p-q}$	

Consequently the above results give the test criterion to test the hypothesis $\gamma = 0$. In other words, if we have

$$\frac{n-p-q}{q}\ \frac{S_2}{S_E} > F_\alpha(q,\ n-p-q)^{(7)}$$

, then we reject the null hypothesis $H_0 : \gamma = 0$ and conclude that the addition of Z is significant with α level of significance.

References

1. A.S. Goldberger, "Stepwise Least Squares: residual analysis and specification error," Journal of the American Statistical Association, 56(1961), pp.998—1000.
2. F.A. Graybill, "An Introduction to Linear Statistical Model," Vol.1. McGraw-Hill: New York, 1961.
3. J. Johnston, "Econometric Methods," Second Edition, Tokyo: McGraw-Hill Kogakusha, Ltd., 1972.
4. D.G. Kabe, "Stepwise multivariate linear regression," Journal of the American Statistical Association, 58(1963), pp.770—773.
5. G.A.F.Seber, "Linear RegressionAnalysis," John Wiley & Sons, 1977.

(6) For the proof of the theorem in an OLS model, see Graybill[2] pp.135—139.
(7) Here $F_\alpha(q,\ n-p-q)$ means the critical value for the given α level of significance.