Erasmus University Rotterdam
ECONOMETRIC INSTITUTE

Report 7503/ES

A NOTE ON GOLDBERGER'S BEST LINEAR UNBIASED PREDICTOR IN THE GENERALIZED REGRESSION MODEL

by

S. Schim van der Loeff and L. Leclercq

January 1975

A NOTE ON GOLDBERGER'S BEST LINEAR UNBIASED PREDICTOR IN THE GENERALIZED REGRESSION MODEL

by

S. Schim van der Loeff and L. Leclercq *

This note serves to demonstrate that apart from the principle of best linear unbiased prediction, Goldberger's predictor in the generalized linear regression model [See Goldberger (1962)] may be derived in at least two other ways. The first one obtains by treating the values of the regressand in the predictive period as missing observations; the second one results from minimizing the generalized variance of the predictive distribution. At the same time Goldberger's predictor is generalized so as to apply to any arbitrary number of predictive periods, which may be important in the analysis of multiperiod decision problems.

Let us consider the model

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\beta + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{1}$$

where y_1 is the T×1 vector of regressand observations, y_2 is the s×1 vector of regressand values to be predicted or missing, x_1 is the T×K matrix of regressor observations, x_2 is the s×K matrix of values of the regressors in the predictive period, β is the K×1 vector of regression coefficients, u_1 is the T×1 vector of disturbances in the sample period and u_2 is the s×1 vector of disturbances in the predictive period. The disturbances are assumed to be normally distributed with

$$E\left[\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right] \equiv E[u] = 0$$

and (known) covariance matrix

$$E[uu'] = E\left[\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(u_1' \quad u_2')\right] = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \equiv \Omega \tag{2}$$

* The authors are indebted to Dr. R. Harkema for his many helpful suggestions. This note was written while the second author was visiting the Econometric Institute from CORE (University of Louvain).

Best Linear Unbiased Prediction

Consider the problem of finding the linear unbiased predictor

(3) $$p = Cy_1$$

C being of order s×T, which minimizes the mathematical expectation of an arbitrary quadratic form in the prediction errors, i.e., which minimizes

(4) $$E[(p - y_2)'A(p - y_2)]$$

where A is a positive-definite symmetric matrix, which can be viewed upon as giving weights to the prediction errors in different periods. Making use of the unbiasedness condition

(5) $$CX_1 \equiv X_2$$

the problem boils down to minimizing the following expression

(6) $$\mathrm{tr}\{C'AC\Omega_{11} - C'A\Omega_{21} - AC\Omega_{12} + A\Omega_{22}\} - 2\mathrm{tr}\{(CX_1 - X_2)\Lambda\}$$

where Λ is the K×s matrix of Lagrange multipliers. Differentiation with respect to C and Λ yields the first order conditions

(7.a) $$2AC\Omega_{11} - 2A\Omega_{21} - 2\Lambda'X_1' = 0$$

(7.b) $$CX_1 - X_2 \quad = 0$$

In order to solve this system, we first determine Λ' by postmultiplying (7.a) by $\frac{1}{2}\Omega_{11}^{-1} X_1(X_1'\Omega_{11}^{-1}X_1)^{-1}$ and using (7.b) to obtain

$$AX_2(X_1'\Omega_{11}^{-1}X_1)^{-1} - A\Omega_{21}\Omega_{11}^{-1}X_1(X_1'\Omega_{11}^{-1}X_1)^{-1} = \Lambda'$$

Insert this result in (7.a) so that

$$2A[C\Omega_{11} - \Omega_{21} - X_2(X_1'\Omega_{11}^{-1}X_1)^{-1}X_1' + \Omega_{21}\Omega_{11}^{-1}X_1(X_1'\Omega_{11}^{-1}X_1)^{-1}X_1'] = 0$$

This condition requires that the expression between square brackets be zero, which leads to the following solution for C

(8) $$C = \Omega_{21}\Omega_{11}^{-1} + (X_2 - \Omega_{21}\Omega_{11}^{-1}X_1)(X_1'\Omega_{11}^{-1}X_1)^{-1}X_1'\Omega_{11}^{-1}$$

Substituting (8) into (3) yields Goldberger's predictor for s predictive periods

(9) $$p = X_2\hat{\beta} + \Omega_{21}\Omega_{11}^{-1}\hat{u}_1$$

where

(10) $$\hat{\beta} = (X_1'\Omega_{11}^{-1}X_1)^{-1}X_1'\Omega_{11}^{-1}y_1, \text{ and } \hat{u}_1 = y_1 - X_1\hat{\beta}$$

The Missing Observations Approach

When treating the regressand values to be predicted as missing observations, we may benefit from a lemma that has recently been proved by Sargan and Drettakis [See Sargan and Drettakis (1974), p. 40-41]. It says that, apart from a factor that has to do with the covariance matrix, the result of maximizing the likelihood function of the observed data is equal to the result of maximizing the likelihood function for all data (observed as well as missing), treating the missing observations as parameters with respect to which the likelihood function is also to be maximized. As in the present case the covariance matrix is supposed to be known, we may thus simply maximize the likelihood function for all data with respect to β and y_2.

The log-likelihood function for all data is given by

(11) $$\mathcal{L} = k - \tfrac{1}{2}[(y_1 - X_1\beta)'(y_2 - X_2\beta)']\Omega^{-1}\begin{bmatrix}(y_1 - X_1\beta)\\(y_2 - X_2\beta)\end{bmatrix}$$

where k denotes an irrelevant constant number. Maximizing (11) with respect to β and y_2, yields the following first-order conditions

(12) $$\begin{bmatrix} X_1'\Omega^{11}X_1 + X_1'\Omega^{12}X_2 + X_2'\Omega^{21}X_1 + X_2'\Omega^{22}X_2 & -(X_1'\Omega^{12} + X_2'\Omega^{22}) \\ -(\Omega^{21}X_1 + \Omega^{22}X_2) & \Omega^{22}\end{bmatrix}\begin{bmatrix}\beta\\ y_2\end{bmatrix}$$

$$= \begin{bmatrix}(X_1'\Omega^{11}y_1 + X_2'\Omega^{21}y_1)\\ -\Omega^{21}y_1\end{bmatrix}$$

where Ω^{ij} (i,j = 1,2) denote the submatrices of the partitioned Ω^{-1}

matrix. Solving the first-order conditions yields the same estimates as given in (9) and (10).

Minimizing the Generalized Variance

Let us now consider the problem of finding the linear unbiased predictor defined in (3) which minimizes the generalized variance of the predictive distribution, i.e.,

$$\det\{E[(p - y_2)(p - y_2)']\} \tag{13}$$

Taking into account the restriction (5), the following expression has to be minimized

$$\det\{(C\Omega_{11}C' - C\Omega_{12} - \Omega_{21}C' + \Omega_{22})\} - 2\mathrm{tr}\{(CX_1 - X_2)\Lambda\} \tag{14}$$

Differentiation[1] with respect to C and Λ yields again the first-order conditions (7), where the matrix A now has become

$$A = \det\{(C\Omega_{11}C' - C\Omega_{12} - \Omega_{21}C' + \Omega_{22})\}(C\Omega_{11}C' - C\Omega_{12} - \Omega_{21}C' + \Omega_{22})^{-1} \tag{15}$$

, i.e., A is the matrix of cofactors of $(C\Omega_{11}C' - C\Omega_{12} - \Omega_{21}C' + \Omega_{22})$. Remembering that formula (8) holds good for any positive-definite symmetric matrix A, the solution to this problem is again the best linear unbiased predictor in (9) and (10).

---

[1] See the appendix for the necessary derivations.

REFERENCES

[1] Goldberger, A.S. (1962), "Best Linear Unbiased Prediction in the Generalized Regression Model", Journal of the American Statistical Association, Vol. 57, 369-75.

[2] Sargan, J.D., and E.G. Drettakis, (1974), "Missing Data in an Autoregressive Model", International Economic Review, Vol. 15, 39-58.

APPENDIX

In this appendix we shall differentiate the determinant of the $s\times s$ matrix V defined as

$$\text{(A.1)}\qquad V = C\Omega_{11}C' - C\Omega_{12} - \Omega_{21}C' + \Omega_{22}$$

with respect to the elements of C. Since V is symmetric there are $\frac{1}{2}s(s+1)$ independent elements of V. Thus

$$\text{(A.2)}\qquad \frac{\partial|V|}{\partial c_{ij}} = \sum_{k=1}^{s} \frac{\partial|V|}{\partial v_{kk}} \frac{\partial v_{kk}}{\partial c_{ij}} + \sum_{k=1}^{s-1} \sum_{\ell=k+1}^{s} \frac{\partial|V|}{\partial v_{k\ell}} \frac{\partial v_{k\ell}}{\partial c_{ij}}$$

where $v_{k\ell}$ is the (k,ℓ)-element of V. In terms of the elements of C and Ω, $v_{k\ell}$ can be written as

$$\text{(A.3)}\qquad v_{k\ell} = \sum_{n=1}^{T} \sum_{m=1}^{T} c_{km}\omega_{mn}c_{\ell n} - \sum_{m=1}^{T} c_{km}\omega_{m,T+\ell} - \sum_{m=1}^{T} \omega_{T+k,m}c_{\ell m} + \omega_{T+k,T+\ell}$$

$$(k,\ell = 1, \ldots, s)$$

If V is symmetric and $v^{k\ell}$ denotes the (k,ℓ)-element of V^{-1}, it is well known that[1]

$$\text{(A.4)}\qquad \frac{\partial|V|}{\partial v_{kk}} = v^{kk}|V|, \text{ and } \frac{\partial|V|}{\partial v_{k\ell}} = 2v^{k\ell}|V| \quad \text{for } k \neq \ell$$

The first term on the right-hand side of (A.2) yields, substituting (A.3) and (A.4)

$$\text{(A.5)}\qquad \sum_{k=1}^{s} \frac{\partial|V|}{\partial v_{kk}} \frac{\partial v_{kk}}{\partial c_{ij}} = 2|V|v^{ii}\left(\sum_{m=1}^{T} c_{im}\omega_{mj} - \omega_{T+i,j}\right)$$

Likewise

$$\text{(A.6)}\qquad \sum_{k=1}^{s-1} \sum_{\ell=k+1}^{s} \frac{\partial|V|}{\partial v_{k\ell}} \frac{\partial v_{k\ell}}{\partial c_{ij}} = 2|V| \sum_{k=1}^{i-1} v^{ki}\left(\sum_{m=1}^{T} c_{km}\omega_{mj} - \omega_{T+k,j}\right) +$$

$$+ 2|V| \sum_{\ell=i+1}^{s} v^{i\ell}\left(\sum_{m=1}^{T} c_{\ell m}\omega_{jm} - \omega_{j,T+\ell}\right)$$

---

1 See e.g. Theil, H. (1971), "Principles of Econometrics", John Wiley & Sons, New York, p.32.

Because of the symmetry of V and Ω, (A.2) becomes

(A.7) $$\frac{\partial|V|}{\partial c_{ij}} = 2|V| \sum_{k=1}^{s} v^{ki}\left(\sum_{m=1}^{T} c_{km}\omega_{mj} - \omega_{T+k,j}\right)$$

Or in matrix notation

(A.8) $$\frac{\partial|V|}{\partial C} = 2|V|V^{-1}(C\Omega_{11} - \Omega_{21})$$