Bulletin of the Graduate School of International Relations I.U.J. No. 6. December 1986

(Research Notes)

A SAS Program for Calculating Prediction Variance Decomposition in VARMA Models

Noriyoshi Shiraishi

1. Introduction

After the influential papers by Sims (1972, 1980a, 1980b) on causal relationships among economic variables, the profession has devoted substantial research efforts to developing and extending time series methodologies for testing the presence and direction of Granger (1969) causality. These techniques include: in a bivariate system, a cross correlation method suggested by Haugh (1976) and Pierce (1977), a one-sided distributed lag approach implied by Granger (1969), and a two-sided distributed lag method developed by Sims (1972); in a multivariate system, a vector autoregression (VAR) proposed by Sims (1980a, 1980b) and a vector autoregressive moving average (VARMA) model developed by Tiao and Box (1981).

Among these methodologies for causality detection, perhaps the most popular is the VAR developed by Sims since it allows more than two variables to be considered and causal relationships among variables are easily detected via a prediction variance decomposition technique. As is well known, however, the Sims VAR often suffers from the overparameterization of the model. The number of observations available is too small for obtaining precise estimates of too many free parameters in the VAR model.

The VARMA model, as a more general case of the special VAR model, can be characterized by its parsimonious way of parameterization: that is, the model uses the smallest number of parameters required for adequately representing the data. Thus the VARMA in general is a more parsimonious representation of the process than the VAR. Although the prediction variance decomposition is theoretically applicable, VARMA's causality detection procedure has been based almost exclusively on hypothesis testing for parameters. This is due to the fact that the prediction variance decomposition technique has not been available in commercial statistical software packages. Since it is often very difficult to interpret the dynamic behavior of the process by looking at and testing the estimated parameters of the model, there is a definite need to develop a computer program for calculating the prediction variance decomposition in the VARMA model.

The purpose of this paper is to develop a SAS program for calculating the prediction variance decomposition in the VARMA model. The software package SAS (Statistical Analysis System) is selected because of its prevailing popularity in academic institutions. The SAS program facilitates the implementation of causality detection in the VARMA analysis. In section 2 the prediction variance decomposition is described, and a VARMA model as well as a VAR model are briefly reviewed. Section 3 reports a SAS program. Section 4 shows an example, and some concluding remarks are given in section 5.

2. Prediction Variance Decomposition

Let $\boldsymbol{z}_t=(z_{1t}, \ldots, z_{kt})'$ represent a suitably differenced (to achieve stationarity) k-dimensional vector of multiple time series. Quenouille (1957) indicated that a time series model for $\boldsymbol{z}_t$ can be written as a VARMA (p, q) process:

$$\boldsymbol{\phi}(B)\boldsymbol{z}_t=\boldsymbol{\theta}(B)\boldsymbol{a}_t, \tag{1}$$

where

$$\begin{aligned}\boldsymbol{\phi}(B)&=\boldsymbol{I}-\boldsymbol{\phi}_1B-\ldots-\boldsymbol{\phi}_pB^p,\\ \boldsymbol{\theta}(B)&=\boldsymbol{I}-\boldsymbol{\theta}_1B-\ldots-\boldsymbol{\theta}_qB^q\end{aligned} \tag{2}$$

are matrix polynomials in the backshift operator B (i.e., $B\boldsymbol{z}_t=\boldsymbol{z}_{t-1}$), the $\boldsymbol{\phi}$'s and $\boldsymbol{\theta}$'s are $k\times k$ matrices, and $\boldsymbol{a}_t=(a_{1t}, \ldots, a_{kt})'$ is a series of white noise vectors identically and independently distributed as multivariate normal $N(\mathbf{O}, \boldsymbol{\Sigma})$. If $q=0$, (1) is reduced to a VAR(p) process; if $p=0$, it is reduced to a VMA(q) process. Thus the VARMA process in (1) includes a wide range of multivariate time series models.

Suppose $p\neq 0$ and $q\neq 0$. If the VARMA(p, q) process (1) is invertible, all roots of the determinantal polynomial $|\boldsymbol{\theta}(B)|$ are outside the unit circle, (1) can be written as a VAR(∞) process:

$$\boldsymbol{\theta}^{-1}(B)\boldsymbol{\phi}(B)\boldsymbol{z}_t=\boldsymbol{a}_t \tag{3}$$

or

$$\boldsymbol{\pi}(B)\boldsymbol{z}_t=\boldsymbol{a}_t, \tag{4}$$

where

$$\boldsymbol{\theta}^{-1}(B)\boldsymbol{\phi}(B)=\boldsymbol{\pi}(B)=\boldsymbol{I}-\boldsymbol{\pi}_1 B-\ldots-\boldsymbol{\pi}_n B^n-\ldots \tag{5}$$

and the $\boldsymbol{\pi}$'s are $k\times k$ matrices. Thus the VARMA representation of the process in (1) in general is more parsimonious than the VAR representation in (4). It is noted, however, that a VARMA model should be estimated with the conditional or exact likelihood ratio method (Tiao and Box 1981) while a VAR model can be estimated by the unconstrained least square method (Sims 1980a, 1980b).

Similarly, if the series $\boldsymbol{z}_t$ in (1) is stationary, all roots of the determinantal polynomial $|\boldsymbol{\phi}(B)|$ are outside the unit circle, (1) can be written as a linear combination of current and past innovations (or one-step ahead forecast errors) termed the VMA(∞) process:

$$\boldsymbol{z}_t=\boldsymbol{\phi}^{-1}(B)\boldsymbol{\theta}(B)\boldsymbol{a}_t \tag{6}$$

or

$$\boldsymbol{z}_t=\boldsymbol{\psi}(B)\boldsymbol{a}_t, \tag{7}$$

where

$$\boldsymbol{\phi}^{-1}(B)\boldsymbol{\theta}(B)=\boldsymbol{\psi}(B)=\boldsymbol{I}-\boldsymbol{\psi}_1 B-\ldots-\boldsymbol{\psi}_n B^n-\ldots \tag{8}$$

and the $\boldsymbol{\psi}$'s are $k\times k$ matrices. The i, j-th element of the coefficient matrix $\boldsymbol{\psi}_m$ represents the response of the i-th series of $\boldsymbol{z}_t$, z_{it}, after m periods to an initial unit shock in the j-th component of $\boldsymbol{a}_t$, a_{jt}. The covariance matrix $\boldsymbol{\Sigma}=E\boldsymbol{a}_t\boldsymbol{a}_t'$ in general is not diagonal. However, the following theorem allows the alternative moving average representation of the process having the diagonal covariance matrix.

Theorem: Given the positive definite matrix $\boldsymbol{\Sigma}$ there exists a unique nonsingular matrix $\boldsymbol{C}$, which is lower triangular with one's on the diagonal, such that $\boldsymbol{C\Sigma C}'=\boldsymbol{\Sigma}_d$, where $\boldsymbol{\Sigma}_d$ is diagonal. (For the proof see Graybill 1983, pp 207–20.)

From the theorem we rewrite (7) as

$$\boldsymbol{z}_t=\boldsymbol{\psi}(B)\boldsymbol{C}^{-1}\boldsymbol{C}\boldsymbol{a}_t \tag{9}$$

or

$$\boldsymbol{z}_t=\boldsymbol{\nu}(B)\boldsymbol{u}_t, \tag{10}$$

where

$$\boldsymbol{\nu}(B)=\boldsymbol{\psi}(B)\boldsymbol{C}^{-1}=\boldsymbol{\nu}_0-\boldsymbol{\nu}_1 B-\ldots-\boldsymbol{\nu}_n B^n-\ldots \tag{11}$$

and

$$\boldsymbol{u}_t = \boldsymbol{C}\boldsymbol{a}_t. \tag{12}$$

It is noted that the covariance matrix of the innovation $\boldsymbol{u}_t$, $E\boldsymbol{u}_t\boldsymbol{u}_t' = \boldsymbol{\Sigma}_d$, is diagonal and (12)

$$\begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ \vdots \\ u_{kt} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{21} & 1 & 0 & \dots & 0 \\ c_{31} & c_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \dots & 1 \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \\ a_{3t} \\ \vdots \\ a_{kt} \end{bmatrix}$$

shows that $\boldsymbol{u}_t$ is recursively determined by $\boldsymbol{a}_t$. The i, j-th element of the coefficient matrix $\boldsymbol{\nu}_m$ also displays the response of the i-th component of $\boldsymbol{z}_t$, z_{it}, after m periods to an initial shock in the j-th component of $\boldsymbol{u}_t$, u_{jt}.

Let $\hat{\boldsymbol{z}}_t(h)$ be the minimum mean squared error forecast of $\boldsymbol{z}_{t+h}$ made at time origin t and $\boldsymbol{e}_t(h) = \boldsymbol{z}_{t+h} - \hat{\boldsymbol{z}}_t(h)$ be the corresponding forecast error. It is well known that $\hat{\boldsymbol{z}}_t(h)$ is the conditional expectation of $\boldsymbol{z}_{t+h}$ given all past history up to time t, or equivalently

$$\hat{\boldsymbol{z}}_t(h) = E(\boldsymbol{z}_{t+h} | \boldsymbol{z}_t, \boldsymbol{z}_{t-1}, \dots). \tag{13}$$

Also, the error vector $\boldsymbol{e}_t(h)$ is normally distributed with zero mean and covariance matrix

$$\mathrm{Cov}[\boldsymbol{e}_t(h)] = \sum_{m=0}^{h-1} \boldsymbol{\nu}_m \boldsymbol{\Sigma}_d \boldsymbol{\nu}_m', \tag{14}$$

where $\boldsymbol{\nu}_m$ is the coefficient matrix in (11).

The covariance matrix $\boldsymbol{\Sigma}_d$ of the innovation vector $\boldsymbol{u}_t$ is diagonal and we write the diagonal entries of $\boldsymbol{\Sigma}_d$ as $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$. Then the i-th diagonal element of $\mathrm{Cov}[\boldsymbol{e}_t(h)]$ represents the variance of the i-th component in the h-step ahead forecast error $\boldsymbol{e}_t(h)$, $e_{it}(h) = z_{i,t+h} - \hat{z}_{it}(h)$, and is written as

$$\mathrm{Var}[e_{it}(h)] = \sum_{m=0}^{h-1} \sum_{n=1}^{k} \sigma_n^2 \nu_m^2(i, n), \tag{15}$$

where $\nu_m(i, j)$ denotes the i, j-th element of the coefficient matrix $\boldsymbol{\nu}_m$ in (11). It is noted that equation (15) shows that the error variance of each element in the h-step ahead forecast $\hat{\boldsymbol{z}}_t(h)$ is decomposed into k variances of the innovation vector $\boldsymbol{u}_t$, i.e., $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$. Therefore, the relative variance contribution (RVC) of the j-th component of the innovation vector, u_{jt}, to the i-th component of the h-step ahead forecast error, $e_{it}(h)$, can be calculated as

$$\mathrm{RVC}_{j\to i}=\frac{\sum_{m=0}^{h-1}\sigma_j^2\nu_m^2(i,j)}{\sum_{m=0}^{h-1}\sum_{n=1}^{k}\sigma_n^2\nu_m^2(i,n)} \tag{16}$$

and is called the prediction variance decomposition (Sims 1980a, 1980b).

3. SAS Program for Calculating Prediction Variance Decomposition in VARMA Models

A SAS computer program has been developed for calculating the prediction variance decomposition in VARMA models. The program uses the PROC MATRIX procedure described in the SAS User's Guide: Statistics (1982). The following example, which will be discussed in the next section, is used to illustrate the computer program. Suppose we have the estimated VARMA model

$$(\boldsymbol{I}-\boldsymbol{\phi}_1B-\boldsymbol{\phi}_4B^4-\boldsymbol{\phi}_5B^5)\boldsymbol{z}_t=\boldsymbol{a}_t, \tag{17}$$

where

$$\boldsymbol{\phi}_1=\begin{bmatrix} -.242 & -.471 & .239 \\ (.118) & (.157) & (.168) \\ .050 & .467 & -.044 \\ (.077) & (.102) & (.109) \\ -.027 & -.184 & .308 \\ (.085) & (.113) & (.121) \end{bmatrix},\quad \boldsymbol{\phi}_4=\begin{bmatrix} -.184 & -.043 & .049 \\ (.120) & (.171) & (.160) \\ .148 & -.291 & .132 \\ (.077) & (.110) & (.103) \\ -.052 & -.092 & -.456 \\ (.086) & (.123) & (.115) \end{bmatrix},$$

$$\boldsymbol{\phi}_5=\begin{bmatrix} -.093 & -.303 & -.097 \\ (.113) & (.173) & (.176) \\ -.281 & .237 & .111 \\ (.073) & (.112) & (.113) \\ .001 & -.087 & .226 \\ (.081) & (.124) & (.126) \end{bmatrix},$$

$$\boldsymbol{\Sigma}=\begin{bmatrix} .000197 & & \\ -.000027 & .000082 & \\ .000006 & -.000026 & .000102 \end{bmatrix},\ \text{and}$$

standard errors are in parentheses. Although this example uses the VAR model as a special case of the VARMA model, the computer program described in this section is applicable for a general class of VARMA models.

Table 1. Program Listing

```
*----------------------------------------------------------------------*
|  PREDICTION VARIANCE DECOMPOSITION TECHNIQUE FOR VARMA               |
|  MODELS                                                              |
|    NOTE: ASSUME THAT VARMA(P,Q) AND ITS VARIANCE-COVARIANCE          |
|          MATRIX HAVE BEEN ALREADY ESTIMATED.                         |
*----------------------------------------------------------------------*;
PROC MATRIX;

MAORDER=1; ORDER OF MA POLYNOMIAL INCLUDING THE IDENTITY
                    MATRIX IS SPECIFIED.;
PSIORDER=16; ORDER OF PSI POLYNOMIAL TO BE COMPUTED INCLUDING
                    THE IDENTITY MATRIX IS SPECIFIED.;
K=3;              * DIMENSION K OF THE VECTOR TIME SERIES IS SPEC-
                    IFIED.;

                  * ESTIMATED AR POLYNOMINAL IS SPECIFIED.;
AR=1 0 0  -.242  -.471   .239  0 0 0  0 0 0  -.184  -.043   .049
          -.093  -.303  -.097/
   0 1 0   .050   .467  -.044  0 0 0  0 0 0   .148  -.291   .132
           .281   .237   .111/
   0 0 1  -.027  -.184   .308  0 0 0  0 0 0  -.052  -.092  -.456
           .001  -.087   .226;
                  * ESTIMATED MA POLYNOMINAL IS SPECIFIED.;
MA=1 0 0/
   0 1 0/
   0 1 0;
                  * ESTIMATED VARIANCE-COVARIANCE MATRIX IS SPEC-
                    IFIED.;
SIGMA=  .000197  -.000027   .000006/
       -.000027   .000082  -.000026/
        .000006  -.000026   .000102;

 * PSI POLYNOMINAL IS COMPUTED.;
PSI=MRATIO (AR, MA, MAORDER, PSIORDER);

 * FIND THE UPPER TRIANGULAR (NONSINGULAR) MATRIX C SUCH THAT
     C'*SIGMA*C=DSIGMA.;
UPPER=HALF(SIGMA);
C=XMULT(INV(UPPER),DIAG(UPPER));
DSIGMA=XMULT(DIAG(UPPER),DIAG(UPPER));

 * PRINT THE MATRICES TO BE USED.;
PRINT AR MA PSI;
LOWERC=C';
PRINT SIGMA LOWERC DSIGMA;

 * COMPUTE RELATIVE VARIANCE CONTRIBUTIONS (RVC).;
VV=J(K,K,0);
ROW=1:K;
DO STEP=0 TO PSIORDER-1;
   VV=VV+XMULT(PSI(ROW,STEP*K+1:(STEP+1)*K),INV(C'))##2;
   NV=XMULT(VV,DIAG(DSIGMA)); * NV: NOISE VARIANCE;
   FEV=XMULT(NV,J(K,K,1)); * FEV: FORCAST ERROR VARIANCE;
   RVC=NV#/FEV; * RVC: RELATIVE VARIANCE CONTRIBUTION;
```

Table 1 (continued).

```
    * PRINT THE COMPUTED RVC'S.;
   FCSTEP–STEP | 1;
   PRINT FCSTEP RVC;
END;
*-----------------------------------------------------------------------*
 RVC(I,J) MEASURES THE RELATIVE VARIANCE CONTRIBUTION OF THE
 J-TH COMPONENT OF THE INNOVATION VECTOR TO THE I-TH COM-
 PONENT OF THE H(=PSIORDER)-STEP AHEAD FORECAST ERROR.   OR
 SYMBOLICALLY,

                        1   2 ... K
                  1  [  ←   ← ... ← ]
                  2  [  ←   ← ... ← ]
 RVC(I, J):       :  [  :   :  ·. : ]
                  K  [  ←   ← ... ← ]
*-----------------------------------------------------------------------*;
```

We use the MRATIO function (SAS/ETS Econometrics and Time Series Library) to transform a VARMA model into a pure moving average form. The HALF function is used to obtain the Cholesky decomposition of a matrix. Table 1 exhibits the annotated program listing.

4. Example

The SAS program described in the previous section was used to examine the relationship between money supply, real income, and price data on the Japanese economy. We first estimated the VARMA model over the period 1965I—1984III using the model building method proposed by Tiao and Box (1981). The money supply used is M2CD which is the sum of M2 (the total of the cash currency, deposit money, and quasi-money) and Certificates of Deposit. The money supply outstanding at the last month of the quarter is used as the quarterly figure. For both money supply and real income (RGNP) the data before seasonal adjustments are used. The price (PRICE) used is the GNP deflator. These series were logged and suitably differenced to achieve stationarity as $z_t=[(1-B)(1-B^4)\text{lnRGNP},\ (1-B)(1-B^4)\text{lnPRICE},\ (1-B)(1-B^4)\text{lnM2CD}]$. The estimated VARMA model is shown in (17) in Section 3.

Given the estimated VARMA model, we used the SAS program to examine the relationship between macroeconomic variables. We set the forecast horizon to be 16 quarters, i.e., the variable PSIORDER in the program

Table 2. Proportions of *k*-month Ahead Forecast Error Variance Explained by Each Innovation

Forecast Error in:	Forecast Horizon	By Innovation in: RGNP	PRICE	M2CD
RGNP	1	1.00	0	0
	4	0.87	0.11	0.02
	8	0.75	0.22	0.03
	12	0.69	0.27	0.04
	16	0.65	0.31	0.04
PRICE	1	0.05	0.95	0
	4	0.04	0.96	0
	8	0.11	0.88	0.01
	12	0.09	0.88	0.03
	16	0.08	0.88	0.04
M2CD	1	0	0.08	0.92
	4	0	0.14	0.86
	8	0.01	0.22	0.77
	12	0.01	0.28	0.71
	16	0.01	0.34	0.65

is 16. The SAS output summarized in Table 2 shows the allocation of forecast error variance to innovations of RGNP, PRICE, and M2CD.

Table 2 shows that the proportion of forecast error variance in real GNP accounted for by price innovations at the 16-quarter horizon is 31 percent. Price innovations account for 34 percent of forecast error variance in money supply at the 16-quarter horizon. 88 percent of forecast error variance in price at the 16-quarter horizon is explained by its own innovations. Since the purpose of this section is to provide an example of the SAS program developed in Section 3, we leave aside questions that may rise from an economics perspective.

5. Concluding Remarks

We have presented a SAS computer program for calculating the prediction variance decomposition in VARMA models. The motivation for developing this computer program stems from the fact that the decomposition technique has not been available in commercial software packages and thus

VARMA's causality detection has been based only on the hypothesis testing for parameters. Since it is often difficult to interpret the dynamic behavior of the time series process by looking at and testing the estimated parameters, there is a need to develop a computer program for calculation the prediction variance decomposition in VARMA models. It is hoped that the computer program presented here facilitates the implementation of causality detection in VARMA analysis.

Acknowledgement

The author used the IBM 4331 at the IUJ Computing Center for executing the computations.

References

Granger, C. W. J. (1969) "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica* 37, 424–438.

Graybill, F. A. (1983) *Matrices with Applications in Statistics*, 2nd Edition, Wadsworth, Inc.

Haugh, L. D. (1976) "Checking the Independence of Two Covariance Stationary Time Series: A Univariate Residual Cross-Correlation Approach," *Journal of the American Statistical Association* 71, 378–385.

Pierce, D. A. (1977) "Relationships—and the Lack of Thereof—between Economic Time Series with Special Reference to Money and Interest Rates," *Journal of the American Statistical Association* 72, 11–22.

Quenouille, M. H. (1957) *The Analysis of Multiple Time Series*, Griffin.

SAS Institute. (1982) *SAS User's Guide: Statistics*, 1982 Edition, SAS Institute, Inc.

SAS Institute. (1982) *SAS/ETS User's Guide: Econometrics and Time Series Library*, 1982 Edition, SAS Institute, Inc.

Sims, C. A. (1972) "Money, Income and Causality," *American Economic Review* 62, 540–552.

——— (1980a) "Macroeconomics and Reality," *Econometrica* 48, 1–48.

——— (1980b) "Comparison of Interwar and Postwar Cycles: Monetarism Reconsidered," *American Economic Review* 70, 250–257.

Tiao, G. C. and Box, G. E. P. (1981) "Modeling Multiple Time Series with Applications," *Journal of the American Statistical Association* 76, 802–816.