Exact joint forecast regions for vector autoregressive models
Wai-Sum Chan;Cheung, Siu Hung;Ka Ho Wu
Journal of Applied Statistics; Jan 1999; 26, 1; ProQuest
pg. 35

Journal of Applied Statistics, Vol. 26, No. 1, 1999, 35–44

Exact joint forecast regions for vector autoregressive models

WAI-SUM CHAN[1], SIU HUNG CHEUNG[2] & KA HO WU[3], [1]*Department of Statistics and Actuarial Science, The University of Hong Kong,* [2]*National University of Singapore and* [3]*Department of Statistics, Chinese University of Hong Kong*

SUMMARY *Assume that a k-element vector time series follows a vector autoregressive (VAR) model. Obtaining simultaneous forecasts of the k elements of the vector time series is an important problem. Based on the Bonferroni inequality, Lütkepohl (1991) derived the procedures which construct the conservative joint forecast regions for the VAR model. In this paper, we propose to use an exact method which provides shorter prediction intervals than does the Bonferroni method. Three illustrative examples are given for comparison of the various VAR forecasting procedures.*

1 Introduction

Methods of using vector autoregressive (VAR) models for analyzing the dynamics of economic systems have attracted considerable research interest in recent years. This is as a result of the influential work of Sim (1980). VAR order selection, estimation and model adequacy checking have been extensively discussed by Lütkepohl (1991, Chapters 2–5) and Hamilton (1994, Chapters 9–12).

The VAR model with order p for a k-element vector time series $\mathbf{Y}_t = (Y_{t,1}, \ldots, Y_{t,k})'$ can be written as

$$\mathbf{A}(L)\mathbf{Y}_t = \nu + \mathbf{u}_t \tag{1}$$

where

$$\mathbf{A}(L) = \mathbf{I}_k - \mathbf{A}_1 L - \ldots - \mathbf{A}_p L^p$$

is the matrix polynomial in the backward shift operator L, with $\mathbf{A}_i$ $(i = 1, \ldots, p)$

Correspondence: W.-S. Chan, Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong.

0266-4763/99/010035-10 $7.00 © 1999 Carfax Publishing Ltd

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

being the fixed $(k \times k)$ coefficient matrices. Also, $\boldsymbol{v} = (v_1, \ldots, v_k)'$ is a fixed $(k \times 1)$ vector of intercept terms. The vector of innovations is denoted by $\mathbf{u}_t$ and

$$\mathbf{u}_t \overset{\text{ind}}{\sim} N_k(\mathbf{0}, \boldsymbol{\Sigma}_u)$$

where $\boldsymbol{\Sigma}_u$ is non-singular. We also assume that the zeros of the determinantal polynomial $|\mathbf{A}(L)|$ all lie outside the unit circle, i.e. the process is stationary. The VAR(p) process in equation (1) has a moving average (MA) representation, where $\mathbf{Y}_t$ is expressed in terms of past and present vectors $\mathbf{u}_t$ and the intercept vector $\boldsymbol{v}$. The MA representation of equation (1) is given by

$$\mathbf{Y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Phi}_i \mathbf{u}_{t-i} \tag{2}$$

where $\boldsymbol{\mu}$ is the mean of the process and $\boldsymbol{v} = \mathbf{A}(1)\boldsymbol{\mu}$. Furthermore $\boldsymbol{\Phi}_i$ $(i = 1, 2, \ldots)$ denotes fixed $(k \times k)$ MA coefficient matrices, which can be computed recursively using

$$\boldsymbol{\Phi}_0 = \mathbf{I}_k \qquad \text{and} \qquad \boldsymbol{\Phi}_i = \sum_{j=1}^{i} \boldsymbol{\Phi}_{i-j} \mathbf{A}_j. \tag{3}$$

The h-step-ahead simultaneous forecast of the values of $Y_{t+h,m}$ $(1 \leqslant m \leqslant k)$ is one of the major objectives of analyzing vector time series, as argued in Lütkepohl (1991, p. 1). Joint prediction intervals are particularly important to decision-makers. Analogously to the arguments of Chatfield (1993), there are three reasons why joint forecast regions might be calculated:

(1) to assess future risk and uncertainty;
(2) to design plans and strategies that can accommodate different scenarios indicated by the forecast regions;
(3) to compare more thoroughly vector forecasts from different methods.

Unfortunately, VAR predictions are often only given as point forecasts (see, for example, Boero, 1990; Edlund & Karlsson, 1993; Otter, 1990; Webb, 1995). This might be because the computation of joint VAR forecast intervals requires the evaluation of high-dimensional multivariate normal probabilities, which was extremely difficult in the past. Lütkepohl (1991, p. 89) derived the approximate joint prediction regions using the conservative Bonferroni method. However, in recent years, we have witnessed a drastic improvement in the algorithms which can compute high-dimensional multivariate normal probabilities (Genz, 1992; Joe, 1995). In this paper, we apply these advanced algorithms to compute the exact joint prediction regions for VAR forecasts.

In the next section, the basic concepts of VAR interval forecasts will be outlined. The proposed method for calculating the exact joint prediction intervals is described in Section 3. In Section 4, we illustrate the VAR forecasting procedures using two examples. Finally, some possible extensions of the proposed procedure to other vector forecast models are discussed in Section 5.

2 VAR forecasts

Assume that the intercept vector $\boldsymbol{v}$, the coefficient matrices $\mathbf{A}_i$ and the variance covariance matrix $\boldsymbol{\Sigma}_u$ in equation (1) are known. Let $\mathbf{Y}_t(h)$ be the h-step-ahead

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

linear minimum mean squared error (MSE) forecast of $\mathbf{Y}_{t+h}$ at time t. It is given by Lütkepohl (1991, p. 29) as

$$\mathbf{Y}_t(h) = \boldsymbol{\nu} + \mathbf{A}_1\mathbf{Y}_t(h-1) + \ldots + \mathbf{A}_p\mathbf{Y}_t(h-p). \tag{4}$$

The corresponding forecast error is (see Lütkepohl, 1991, p. 32)

$$\mathbf{e}_t(h) = \mathbf{Y}_{t+h} - \mathbf{Y}_t(h) = \sum_{i=0}^{h-1} \boldsymbol{\Phi}_i \mathbf{u}_{t+h-i} \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}_y(h)) \tag{5}$$

where

$$\boldsymbol{\Sigma}_y(h) = \sum_{i=0}^{h-1} \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_u \boldsymbol{\Phi}_i' = (\sigma_{ij}(h), \qquad 1 \leqslant i,j \leqslant k). \tag{6}$$

Let t and h be fixed. For $1 \leqslant m \leqslant k$, define

$$S_m = \frac{Y_{t+h,m} + Y_{t,m}(h)}{[\sigma_{mm}(h)]^{1/2}} \tag{7}$$

where $Y_{t+h,m}$ and $Y_{t,m}(h)$ are the mth elements of $\mathbf{Y}_{t+h}$ and $\mathbf{Y}_t(h)$ respectively. Note that the random vector $\mathbf{S} = (S_1, \ldots, S_k)'$ has a multivariate normal distribution with mean vector zero and variance covariance matrix $\boldsymbol{\Sigma}_s = (\rho_{ij})$, where

$$\rho_{ij} = \frac{\sigma_{ij}(h)}{[\sigma_{ii}(h)\sigma_{jj}(h)]^{1/2}} \tag{8}$$

for $1 \leqslant i,j \leqslant k$. Hence the $100(1-\alpha)\%$ joint prediction intervals for $Y_{t+h,m}$ $(m = 1, \ldots, k)$ are

$$Y_{t,m}(h) \pm \xi_\alpha [\sigma_{mm}(h)]^{1/2} \tag{9}$$

where the constant ξ_α satisfies the implicit equation

$$\Pr(|S_m| \leqslant \xi_\alpha, m = 1, \ldots, k) = 1 - \alpha. \tag{10}$$

The determination of ξ_α in equation (10) requires the computation of k-variate normal probabilities, which was extremely difficult for $k \geqslant 6$ in the past (Schervish, 1984). Lütkepohl (1991, p. 34) considered a conservative method to construct the joint prediction intervals, which was based on the first-order Bonferroni inequality. The $100(1-\alpha)\%$ conservative joint prediction limits for $Y_{t+h,m}$ $(m = 1, \ldots, k)$ are

$$Y_{t,m}(h) \pm z_{(\alpha/2k)} [\sigma_{mm}(h)]^{1/2} \tag{11}$$

where $z_{(\alpha/2k)}$ is the upper $\alpha/(2k)$ percentage point of the standardized normal distribution. The intervals in equation (11) are conservative, in the sense that they provide a joint confidence level of at least $1-\alpha$. Furthermore, it is easy to show that $z_{(\alpha/2k)} \geqslant \xi_\alpha$, which implies that the widths of equation (11) are greater than the widths of equation (9). Hence, the forecast intervals in equation (11) are less precise than are the forecast intervals in equation (9). Therefore, the exact joint prediction intervals are preferred if the required computational effort is reasonable.

Finally, when the parameters in equation (1) are not known, a multivariate least-squares procedure can be used to estimate them. The formulae to compute the estimates of $\boldsymbol{\nu}$, $\mathbf{A}_i$ $(i = 1, \ldots, p)$, $\boldsymbol{\Sigma}_u$, $\mathbf{Y}_t(h)$, and $\boldsymbol{\Sigma}_y(h)$ are provided in Lütkepohl (1991, pp. 63–68, 85–89).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3 Exact prediction limits for VAR forecasts

In this section, the method which computes the exact joint prediction limits in equation (9) for VAR forecasts is examined. When α and $\boldsymbol{\Sigma}_s$ are given, we have

$$\Pr(|S_m| \leqslant \xi_\alpha, m = 1, \ldots, k) = \int_{-\xi_\alpha}^{\xi_\alpha} \cdots \int_{-\xi_\alpha}^{\xi_\alpha} f(s_1, \ldots, s_k) \mathrm{d}s_1 \ldots \mathrm{d}s_k \qquad (12)$$

where $f(s_1, \ldots, s_k)$ is the k-dimensional multivariate normal density function of $(S_1, \ldots, S_k)'$. Numerically, the evaluation of the multivariate normal probabilities of equation (12) has long been a very difficult task, which makes the conservative Bonferroni method in equation (11) very attractive in the construction of joint prediction limits for VAR forecasts. However, as the computation techniques and computing power have drastically improved recently, the evaluation of equation (12) becomes feasible even for k as large as 20 (Genz, 1992; Joe, 1995). The following steps outline an algorithm which computes the exact prediction intervals in equation (9) for VAR forecasts of $Y_{t+h,m}$ $(m = 1, \ldots, k)$.

Step 1. Input v, $\mathbf{A}_i$ $(i = 1, \ldots, p)$, $\boldsymbol{\Sigma}_u$, h and α.
Step 2. Compute $\mathbf{Y}_t(h)$ using equation (4).
Step 3. Evaluate $\boldsymbol{\Sigma}_y(h)$ using equations (3) and (6).
Step 4. Standardize $\boldsymbol{\Sigma}_y(h)$ using equation (8) and obtain the matrix $\boldsymbol{\Sigma}_s$.
Step 5. Use the Secant method (Burden & Faires, 1989) to evaluate ξ_α in equation (10).
  (a) Input two initial guesses to start the Secant algorithm. The suggestions for these two guesses are $z_{(\alpha/2)}$ and $z_{(\alpha/2k)}$.
  (b) Compute equation (12) using the Genz (1992) algorithm, which is explained in Appendix A.
Step 6. Use equation (9) to generate the required prediction intervals.

If the parameters in equation (1) are unknown then v, $\mathbf{A}_i$ $(i = 1, \ldots, p)$, $\boldsymbol{\Sigma}_u$, $\mathbf{Y}_t(h)$, $\boldsymbol{\Sigma}_y(h)$, and $\boldsymbol{\Sigma}_s$ in the algorithm described are replaced by $\hat{v}$, $\hat{\mathbf{A}}_i$ $(i = 1, \ldots, p)$, $\hat{\boldsymbol{\Sigma}}_u$, $\hat{\mathbf{Y}}_t(h)$, $\hat{\boldsymbol{\Sigma}}_y(h)$ and $\hat{\boldsymbol{\Sigma}}_s$ respectively (refer to Section 2 for the computation of these estimators).

With reference to Step 5(a), other numerical methods can be applied to solve equation (10). However, the Secant method converges rapidly and requires no derivative evaluation. For Step 5(b), the algorithm of Joe (1995) is an alternative to that of Genz (1992). Both algorithms are efficient and highly accurate. (For comparisons of these two algorithms, see Joe (1995).)

A copy of the FORTRAN-coded subroutine for the algorithm is available from the first author on request.

4 Illustrative examples

Example 1

We consider the 95% forecast intervals for the individual components of a known VAR(1) process with

$$\mathbf{A}_1 = \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_u = \begin{pmatrix} 2.25 & 0.75 & 1.05 \\ 0.75 & 1.00 & 0.50 \\ 1.05 & 0.50 & 0.75 \end{pmatrix} \qquad (13)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 1. Joint forecast regions (95%) for forecasts of the VAR(1) process in equation (13)

		Bonferroni		Exact		Improvement
h	m	Interval	Length	Interval	Length	(%)
1	1	(−6.591, 0.591)	7.182	(−6.463, 0.463)	6.926	3.6
	2	(0.806, 5.594)	4.788	(0.891, 5.509)	4.618	3.6
	3	(1.027, 5.173)	4.146	(1.100, 5.100)	4.000	3.6
2	1	(−5.515, 2.515)	8.030	(−5.358, 2.358)	7.716	3.9
	2	(0.319, 5.581)	5.262	(0.422, 5.478)	5.056	3.9
	3	(0.277, 4.863)	4.586	(0.366, 4.774)	4.408	3.9

Assume that $\mathbf{Y}_t(1) = (-3.0, 3.2, 3.1)'$ and $\mathbf{Y}_t(2) = (-1.50, 2.95, 2.57)'$. Using equations (3) and (6), the forecast MSE matrices for $h = 1$ and 2 are calculated as

$$\boldsymbol{\Sigma}_y(1) = \boldsymbol{\Sigma}_u \tag{14}$$

$$\boldsymbol{\Sigma}_y(2) = \boldsymbol{\Sigma}_u + \boldsymbol{\Phi}_1 \boldsymbol{\Sigma}_u \boldsymbol{\Phi}_1' = \begin{pmatrix} 2.8125 & 1.0575 & 1.2825 \\ 1.0575 & 1.2080 & 0.6790 \\ 1.2825 & 0.6790 & 0.9175 \end{pmatrix}. \tag{15}$$

Following the algorithm in Section 3, $\xi_{0.05} = 2.309$ for $h = 1$ and $\xi_{0.05} = 2.301$ for $h = 2$ are obtained. For the Bonferroni method, $z_{(\alpha/2k)} = 2.394$. The results are summarized in Table 1. It should be noted that all the lengths of the component intervals in the exact regions are shorter than the respective lengths of the intervals obtained by the conservative method. In fact, the improvements in length are reported in the last column of the table.

Example 2

Here, four historical UK macro-economic time series are selected from Feinstein (1972). They are GNP, unemployment rate (UNE), retail price index (RPI) and wage index (WAG). All the series are taken from Feinstein (1972): GNP, column 9 of Table 1; UNE, column 6 of Table 57; RPI, column 1 of Table 65; and WAG, column 3 of Table 65. The data set consists of 111 yearly observations from 1885 to 1965. To fit the model with the stationarity assumption, we apply the first differencing to the logarithmic transformed data, except for the unemployment rate. Only logarithmic transformation is used for the variable UNE. The resulting series are denoted by $\mathbf{Y}_{t,1}$, $\mathbf{Y}_{t,2}$, $\mathbf{Y}_{t,3}$ and $\mathbf{Y}_{t,4}$ respectively.

A Sims (1980) style of macro-economic VAR model with lag order $p = 4$ is considered. Following Lütkepohl (1991, Section 3.2), the mulivariate least-squares estimation results of the system are given in Table 2. The one-period- and two-period-ahead forecasts are obtained. Their corresponding approximate $\hat{\boldsymbol{\Sigma}}_y(h)$ matrices are computed. The results are summarized in Table 3. Applying the algorithm as described in Section 3, the constant ξ_α in equation (9) is found to have values of 2.399 and 2.380 for $h = 1$ and $h = 2$ respectively. The Bonferroni coefficient $z_{(\alpha/2k)} = 2.498$ for $\alpha = 0.05$ and $k = 4$. The 95% joint prediction regions for $\hat{\mathbf{Y}}_{110}(1)$ and $\hat{\mathbf{Y}}_{110}(2)$ are computed in Table 4 using the exact and the conservative methods. Finally, additional results for time horizons of $h = 1$–6 are displayed in Table 5. The relative interval width losses from using the Bonferroni method range

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 2. Least-squares estimation for the UK macro-economic data

$\hat{\nu}$	$\hat{\mathbf{A}}_1$			$\hat{\mathbf{A}}_2$			$\hat{\mathbf{A}}_3$			$\hat{\mathbf{A}}_4$			$\hat{\boldsymbol{\Sigma}}_u \times 10^{-2}$		
0.03	0.38	0.00	0.49	−0.43	0.01	−0.01	0.25	−0.26	0.05	0.03	0.32	−0.34	−0.10	−0.02	0.09	0.02	0.19	−1.30	0.13	0.08
0.41	−2.02	0.90	−3.46	6.75	−2.54	−0.05	2.26	−2.95	−1.35	−0.47	−4.36	5.80	−1.75	0.38	2.46	−2.01	−1.30	18.57	−0.78	−0.30
0.02	0.33	0.01	0.69	−0.61	0.01	−0.02	−0.19	0.10	0.16	0.02	0.48	−0.58	0.01	−0.02	0.21	−0.11	0.13	−0.78	0.16	0.08
0.01	0.49	0.01	0.14	0.16	−0.02	−0.01	0.02	−0.38	0.20	0.00	0.36	−0.33	0.02	−0.01	0.23	−0.20	0.08	−0.30	0.08	0.08

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 3. Point forecasts results for the UK macro-economic data

$\hat{\mathbf{Y}}_{110}(1)$	$\hat{\boldsymbol{\Sigma}}_{\hat{y}}(1)$				$\hat{\mathbf{Y}}_{110}(2)$	$\hat{\boldsymbol{\Sigma}}_{\hat{y}}(2)$			
0.06	0.22	−1.51	0.16	0.10	0.05	0.30	−2.55	0.23	0.17
0.22	−1.51	21.55	−0.90	−0.35	0.42	−2.55	48.84	−1.94	−1.19
0.05	0.16	−0.90	0.18	0.10	0.03	0.23	−1.94	0.27	0.17
0.05	0.10	−0.35	0.10	0.09	0.05	0.17	−1.19	0.17	0.18

TABLE 4. Joint forecast regions (95%) for forecasts of the UK macro-economic data

		Bonferroni		Exact		Improvement
h	m	Interval	Length	Interval	Length	(%)
	1	(−1.122, 1.245)	2.367	(−1.075, 1.198)	2.273	4.0
1	2	(−11.373, 11.816)	23.189	(−10.914, 11.357)	22.271	4.0
	3	(−1.026, 1.122)	2.148	(−0.984, 1.079)	2.063	4.0
	4	(−0.718, 0.818)	1.536	(−0.688, 0.787)	1.475	4.0
	1	(−1.314, 1.408)	2.722	(−1.250, 1.343)	2.593	4.7
2	2	(−17.035, 17.877)	34.912	(−16.211, 17.053)	33.264	4.7
	3	(−1.268, 1.327)	2.595	(−1.206, 1.266)	2.472	4.7
	4	(−1.018, 1.109)	2.127	(−0.968, 1.059)	2.027	4.7

TABLE 5. Coefficients for constructing joint prediction regions (95%) for the UK macro-economic data

h	Bonferroni $z_{(\alpha/2k)}$	Exact ζ_α	Improvement (%)
1	2.498	2.399	4.0
2	2.498	2.380	4.7
3	2.498	2.370	5.1
4	2.498	2.358	5.6
5	2.498	2.358	5.6
6	2.498	2.362	5.4

from 4.0% to 5.6%, and are potentially significant from an economic point of view. In summary, the overall performance of the exact method is superior in this example.

5 Extensions

The procedure described in Section 3 can easily be extended to other Gaussian vector forecasting models. In this section, we will briefly discuss its extension to cointegrated processes and dynamic simultaneous equations systems, which are commonly used by economic forecasters.

5.1 Cointegrated VAR processes

We call the k-dimensional VAR(p) process in equation (1) cointegrated of rank r if the matrix

$$\boldsymbol{\Pi} = \mathbf{I}_k - \mathbf{A}_1 - \ldots - \mathbf{A}_p \tag{16}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 6. Joint forecast regions (95%) for the Germany deutschmark exchange rate data

		Bonferroni		Exact		Improvement
h	m	Interval	Length	Interval	Length	(%)
1	1	(423.119, 436.942)	13.823	(423.224, 436.838)	13.614	1.5
	2	(528.593, 533.372)	4.779	(528.630, 533.337)	4.707	1.5
	3	(526.097, 531.862)	5.765	(526.141, 531.819)	5.678	1.5
2	1	(419.486, 438.810)	19.624	(419.380, 438.617)	19.237	2.0
	2	(527.886, 536.101)	8.215	(527.967, 536.020)	8.053	2.0
	3	(525.184, 534.788)	9.604	(525.279, 534.694)	9.415	2.0

has rank $r < k$. The system is non-stationary and has $(k - r)$ unit roots. Methods for testing, estimating and forecasting cointegrated models are discussed in Lütkepohl (1991, Chapter 11) and Hamilton (1994, Chapters 19–20).

When the parameters of the cointegrated VAR model are known, Lütkepohl (1991, p. 375) showed that equations (4)–(6) are still valid for calculating interval forecasts for the cointegrated system. Furthermore, he also proved that the forecast errors of the estimated process and the process with known parameters are asymptotically equivalent. Therefore, for large samples, the unknown coefficient matrices in equation (6) can be substituted by their estimates. For small-sample situations, Reimers (1995) added a correction term to equation (6) to take into account the effects of estimation. The approximate forecast MSE matrix for the cointegrated VAR system is

$$\hat{\boldsymbol{\Sigma}}_{\hat{y}}(h) = \sum_{i=0}^{h-1} \hat{\boldsymbol{\Phi}}_i \hat{\boldsymbol{\Sigma}}_u \hat{\boldsymbol{\Phi}}_i' + \frac{1}{T} \hat{\boldsymbol{\Omega}}(h) \tag{17}$$

where T is the sample size and the formulae for calculating $\hat{\boldsymbol{\Omega}}(h)$ are derived in Reimers (1995). Therefore, the exact joint prediction regions can be generated using the algorithm in Section 3. For comparison purposes, we recalculate the example of interval forecasts for the Swiss franc exchange via the deutschmark (D), the price deflator of final demand in Germany (P^{G}) and the price deflator of final demand in Switzerland (P^{CH}) in Reimers (1995). Let $Y_{t,1} = 100[\ln(D)]$, $Y_{t,2} = 100[\ln(P^{\mathrm{G}})]$ and $Y_{t,3} = 100[\ln(P^{\mathrm{CH}})]$. The results are summarized in Table 6 and, as expected, the Bonferroni interval widths are larger. However, the differences between the Bonferroni and exact interval widths are less drastic when compared with previous examples.

5.2 *Dynamic simultaneous equations models*

A dynamic simultaneous equations system may have the form

$$\mathbf{A}(L)\mathbf{Y}_t = \mathbf{B}(L)\mathbf{X}_t + \mathbf{w}_t \tag{18}$$

where

$$\mathbf{A}(L) = \mathbf{A}_0 - \mathbf{A}_1 L - \ldots - \mathbf{A}_p L^p, \qquad \mathbf{B}(L) = \mathbf{B}_0 - \mathbf{B}_1 L - \ldots - \mathbf{B}_q L^q \tag{19}$$

and $\mathbf{Y}_t = (Y_{t,1}, \ldots, Y_{t,k})'$ is a vector of endogenous variables, $\mathbf{X}_t = (X_{t,1}, \ldots, X_{t,v})'$ is a vector of exogenous variables, $\mathbf{A}_i$ and $\mathbf{B}_j$ are $(k \times k)$ and $(k \times v)$ coefficient matrices, respectively, and $\mathbf{w}_t$ is a k-dimensional error vector.

The optimal h-step-ahead forecast of $\mathbf{Y}_{t+h}$ and its forecast MSE matrix are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

discussed in Lütkepohl (1991, pp. 334–338). Similarly, the corresponding joint prediction regions can be obtained using the algorithm in Section 3.

Acknowledgements

The German exchange rate data in Section 5 were kindly provided by Professor Hans-Eggert Reimers. The authors are grateful to the referee for providing several helpful comments. This research was supported by the Hong Kong RGC Earmarked Grant CUHK4081/97H.

REFERENCES

BOERO, G. (1990) Comparing ex-ante forecasts from a SEM and VAR model: an application to the Italian economy, *Journal of Forecasting*, 9, pp. 13–24.

BURDEN, R. L. & FAIRES, J. D. (1989) *Numerical Analysis* (Boston, MA, PWS-Kent).

CHATFIELD, C. (1993) Calculating interval forecasts, *Journal of Business and Economic Statistics*, 11, pp. 121–144.

EDLUND, P. O. & KARLSSON, S. (1993) Forecasting the Swedish unemployment rate VAR vs transfer function modelling, *International Journal of Forecasting*, 9, pp. 61–76.

FEINSTEIN, C. H. (1972) *National Income, Expenditure and Output of the United Kingdom 1855–1965* (Cambridge, Cambridge University Press).

GENZ, A. (1992) Numerical computation of multivariate normal probabilities, *Journal of Computational and Graphical Statistics*, 1, pp. 141–149.

HAMILTON, J. D. (1994) *Time Series Analysis* (Englewood Cliffs, NJ, Princeton University Press).

JOE, H. (1995) Approximations to multivariate normal rectangle probabilities based on conditional expectations, *Journal of the American Statistical Association*, 90, pp. 957–964.

LÜTKEPOHL, H. (1991) *Introduction to Multiple Time Series Analysis* (Berlin, Springer).

OTTER, P. W. (1990) Canonical correlation in multivariate time series analysis with an application to one year-ahead and multiyear-ahead macroeconomic forecasting, *Journal of Business and Economic Statistics*, 8, pp. 453–457.

REIMERS, H. E. (1995) Interval forecasting in cointegrated systems, *Statistical Papers*, 36, pp. 349–369.

SCHERVISH, M. (1984) Multivariate normal probabilities with error bound, *Applied Statistics*, 33, pp. 81–87.

SIMS, C. A. (1980) Macroeconomics and reality, *Econometrica*, 40, pp. 1–48.

WEBB, R. H. (1995) Forecasts of inflation from VAR models, *Journal of Forecasting*, 14, pp. 267–285.

Appendix A: evaluation of the multivariate normal probability in equation (12)

Let $\mathbf{Z} = (Z_1, \ldots, Z_k)' \sim N_k(\mathbf{0}, \boldsymbol{\Sigma})$ and let $\boldsymbol{\Sigma}$ be positive definite, to compute

$$\begin{aligned} G(\xi) &= P(|Z_m| \leqslant \xi, m = 1, \ldots, k) \\ &= [(2\pi)^k |\boldsymbol{\Sigma}|]^{-1/2} \int_{-\xi}^{\xi} \cdots \int_{-\xi}^{\xi} \exp(-\frac{1}{2}\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z})\, d\mathbf{Z} \end{aligned} \tag{A1}$$

Genz (1992) proposed a transformation technique to evaluate equation (A1). Let $\mathbf{C} = \{c_{ij}\}$ be a lower triangular matrix, such that $\mathbf{CC}'$ is the Cholesky decomposition of $\boldsymbol{\Sigma}$. Then, equation (A1) can be transformed to

$$G(\xi) = (e_1 - d_1) \int_0^1 (e_2 - d_2) \ldots \int_0^1 (e_k - d_k) \int_0^1 dw_1 \ldots dw_k \tag{A2}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where

$$d_1 = \Phi(-\xi/c_{11}), \qquad e_1 = \Phi(\xi/c_{11})$$

$$y_i = \Phi^{-1}(d_i + w_i(e_i - d_i)), \qquad i = 1, \ldots, k-1$$

$$d_i = \Phi\left(\left(-\xi - \sum_{j=1}^{i-1} c_{ij} y_j\right) \Big/ c_{ii}\right), \qquad i = 2, \ldots, k$$

$$e_i = \Phi\left(\left(\xi - \sum_{j=1}^{i-1} c_{ij} y_j\right) \Big/ c_{ii}\right), \qquad i = 2, \ldots, k$$

and Φ is the standard normal distribution function. The transformation in equation (A2) enables us to compute equation (A1) efficiently. Genz (1992) reported that even a simple Monte Carlo algorithm is very effective; other details of the algorithm can be found in Genz (1992).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.