The sampling properties of conditional independence graphs for structural vector autoregressions

Marco Reale* **and Granville Tunnicliffe Wilson**†

** Department of Mathematics & Statistics,*
University of Canterbury,
Private Bag 4800, Christchurch, New Zealand.

Report Number: UCDMS2000/3 April 2000

Keywords: Partial correlation, Moralization, Causality
† *Department of Mathematics & Statistics, Lancaster University, Lancaster LA1 4YF, UK.*

The sampling properties of conditional independence graphs for structural vector autoregressions.

Marco Reale* and Granville Tunnicliffe Wilson†

September 6, 2000

Abstract

Structural vector autoregressions allow contemporaneous series dependence and assume errors with no contemporaneous correlation. Such models having a recursive structure can be described by a directed acyclic graph. An important tool for identification of these models is the conditional independence graph constructed from the contemporaneous and lagged values of the process. We determine the large sample properties of statistics used to test for the presence of links in this graph. A simple example illustrates how these results may be applied.

KEY WORDS: Partial correlation, moralization, causality

*Department of Mathematics and Statistics, University of Canterbury, PB 4800 Christchurch, New Zealand

†Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK.

1 Introduction

We consider the structural p-th order vector autoregressive model (SVAR) of a stationary, m dimensional, time series $x_t = (x_{t,1}, x_{t,2}, \ldots, x_{t,m})'$, of the form

$$\Phi_0 x_t = d + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \cdots + \Phi_p x_{t-p} + a_t. \tag{1}$$

We require (a) that the variance matrix D of a_t is diagonal and (b) that Φ_0 represent a recursive (causal) dependence of each component of x_t on other contemporaneous components. This is equivalent to the existence of a re-ordering of the elements of x_t such that Φ_0 is triangular with unit diagonal. We also require that Φ_p has at least one non-zero element.

A given process x_t generated by such a model has in general many statistically equivalent representations of the same form, corresponding to different causal orderings of the contemporaneous dependence. We suppose, however, that there is one representation which is sparse in the sense that many of the elements of the coefficient matrices are zero. This sparseness may be represented by a directed acyclic graph (DAG) with nodes corresponding to the $m(p+1)$ series elements which appear in (1). The only links in this graph are to, and between, contemporaneous values of x_t, in the direction of causality, corresponding to the non-zero elements of Φ_i.

In a previous paper, Reale and Tunnicliffe Wilson (1999), we advocated an approach to identifying the model (1) using a conditional independence graph (CIG) on the same nodes, estimated using a sample of x_t of length N. In that paper we only considered tests for links to, and between, contemporaneous values of x_t. We made brief reference to standard results to justify these tests. In the next section of this paper we give a formal derivation of the sampling properties of the test for a link in the CIG between any pair of nodes $x_{t-r,i}$ and $x_{t-s,j}$ for lags r and s from zero to some maximum lag $k \geq p$. In general this test is non-standard, in that the variance of the statistic is not the same as that for the CIG of independent samples from a multivariate normal distribution. The exception is that in large samples the variance is the same if one node has zero lag.

In the final section we describe the structure of the CIG for model (1) and present a simple example illustrating the application of our results.

2 Sampling properties of the test statistic.

Following Whittaker (1990) the CIG of the Gaussian variables $x_{t-k,i}$ for lags $k = 0, 1, \ldots, K$ and series $i = 1, 2, \ldots, m$, is determined by the matrix of pair-wise partial correlations of the variables conditional upon all the remaining variables. The nodes corresponding to a pair of variables are linked if and only if their partial correlation is non-zero. The starting point for estimation of these partial correlations is the data matrix X consisting of the collection of contemporaneous and lagged data vectors $(x_{K+1-k,i}, \ldots, x_{N-k,i})'$. The sample values of the partial correlations may then be derived from the estimated covariance matrix of the variables, $\hat{V} = X'X/T$ where $T = N - K$. We assume here that the time series or data vectors have been mean corrected. Our test is to reject $\rho = 0$, for the partial autocorrelation ρ between a particular pair of variables, if $|r| > c$ where r is the sample value of ρ and c is some critical value. To determine c we use the relationship between r and the t coefficient of one member of the pair of variables, in the regression of the other member of the pair upon that and all the remaining variables. This is $r = t/\sqrt{(t^2 + \nu)}$ (see Greene, 1993, p. 180). We therefore proceed to determine the large sample properties of t. The important point here is that ordinary least squares regression does not in general furnish the correct standard error of the regression coefficient in this time series context. This is because the regression will, in general, be upon variables which are both in the past and the future and consequently the regression errors are not white noise.

Let y, in the set $x_{t-k,i}$, be the variable selected as the regressor, and w be the remaining variables in this set. Let Y and W be the corresponding subsets of the data matrix X. Let $M = \mathrm{Var}\,(w)$, $P = \mathrm{Cov}\,(w, y)$, $\hat{M} = T^{-1}W'W$ and $\hat{P} = T^{-1}W'Y$, the elements of $\hat{M}$ and $\hat{P}$ being sample covariances of x_t. We assume that x_t does follow the model (1) so that the necessary conditions apply (Priestley, 1994, p. 324–330) for consistency of these sample values: $\mathrm{plim}\,\hat{M} = M$ and $\mathrm{plim}\,\hat{P} = P$. The solution of the least squares equations $W'W\hat{\beta} = W'Y$ for the vector of regression coefficients therefore satisfies $\mathrm{plim}\,\hat{\beta} = M^{-1}P = \beta$. Use this to define the asymptotic error series of the regression, $e_t = y - w\beta$, which may be expressed in the form

$$e_t = \sum_{i,r} \psi_{i,r} x_{t-r,i}. \tag{2}$$

By construction, $\mathrm{Cov}\,(w, e) = 0$ and we may re-formulate the least squares equations as

$$\hat{M}\left[T^{\frac{1}{2}}(\hat{\beta} - \beta)\right] = T^{\frac{1}{2}}\left[T^{-1}W'E\right] \tag{3}$$

where E is the data vector corresponding to e_t. The elements of $D = T^{-1}W'E$ are the sample values of $\mathrm{Cov}\,(w, e)$. Being finite linear combinations of the sample covariances of x_t, they therefore satisfy (Priestley, 1994, p. 337–339).

$$T^{\frac{1}{2}}D \xrightarrow{d} \mathrm{MVN}\,(0, Q), \tag{4}$$

where we use $\mathrm{E}\,(D) = 0$ and Q is $\lim_{T\to\infty} \mathrm{Var}\,(T^{\frac{1}{2}}D)$. The elements of Q are

$$\begin{aligned} &\lim\nolimits_{T\to\infty} T^{-1}\mathrm{Cov}\,\left(\textstyle\sum_t x_{j,t-u}e_t, \sum_s x_{l,s-w}e_s\right) = \\ = \;& \textstyle\sum_{i,r}\sum_{k,v}\psi_{i,r}\psi_{k,v}\lim_{T\to\infty} T^{-1}\mathrm{Cov}\,\left(\sum_t x_{j,t-u}x_{i,t-r}, \sum_s x_{l,s-w}x_{k,s-v}\right) = \\ = \;& \textstyle\sum_{i,r}\sum_{k,v}\psi_{i,r}\psi_{k,v}\sum_m\left(\gamma_{i,k,m-r+v}\gamma_{j,l,m-u+w} + \gamma_{i,l,m-r+w}\gamma_{m-u+v}\right), \end{aligned} \tag{5}$$

where $\gamma_{i,j,k} = \mathrm{Cov}\,(x_{i,t}, x_{j,t-k})$.

We require construction of a consistent estimate $\hat{Q}$ of Q and propose the following approach which performs well. We first express (5) using the components of the spectrum

S of x_t, as

$$\int_{f=-\frac{1}{2}}^{\frac{1}{2}} \left\{ \psi(B) S \psi(B^{-1})' B^{w-u} S_{l,j} + [\psi(B)S]_l \, [\psi(B)S]_j \, B^{-w-u} \right\} df. \tag{6}$$

Here we use the backward shift operator B to aid interpretation, substituting $B = \exp(2\pi i f)$ to evaluate the integral. The operator $\psi(B)$ is defined by $e_t = \psi(B)x_t$ and is consistently estimated using the regression coefficients $\hat{\beta}$. Writing (1) as $x_t = \Phi(B)a_t$ the spectrum S is given by:

$$S = \Phi(B)^{-1} D [\Phi(B^{-1})^{-1}]', \tag{7}$$

which is consistently estimated by OLS estimation of (1). The integral may be evaluated numerically with no difficulty. An important practical point is that the estimated model must be stationary. This may be ensured by padding the observed series x_t, following mean correction, with K zeros at the start and finish. The regression estimates then become the multivariate Yule-Walker estimates, which are known to be stationary (Reinsel, 1993, p. 89–91).

Again using $\operatorname{plim} \hat{M} = M$ we have from (3) that

$$T^{\frac{1}{2}}(\hat{\beta} - \beta) \xrightarrow{d} \text{MVN}\,(0, M^{-1} Q M^{-1}). \tag{8}$$

where of course M is consistently estimated by $\hat{M}$. For comparison, the limiting variance matrix from the OLS regression is $M^{-1}\hat{\sigma}_e^2$.

In general these variance matrices differ. However, in the particular case that the chosen regressor is a contemporaneous variable, they agree, so that the regression standard errors may be used. This may be seen from (6) in which $\psi(B)S\psi(B^{-1})'$ reduces to σ_e^2 and $\psi(B)S$ to an operator with no positive powers of B, so that the second component of the integral becomes zero. Then Q reduces to $M\sigma_e^2$. In practice the estimates of these variance matrices are also identical in finite samples if the order p used to estimate (1) is set equal to the maximum lag K. This is a consequence of Yule-Walker estimation which ensures that the model autocovariances, derived here via the model spectrum, are equal to the sample autocovariances up to lag K.

Now let $\hat{\beta}$ be the estimate of the coefficient relating to a particular link in the CIG, and let $\hat{\sigma}_\beta$ be the consistent estimate of its true standard error. We therefore reject the presence of a link if $|t| = |\hat{\beta}|/\hat{\sigma}_\beta$ exceeds the appropriate critical value of a standard normal variable. We evaluate the performance of this 'correct' test in the next section for a simple illustrative example and compare it with the 'incorrect' test using the wrong standard error from the OLS regression.

3 Application to a simple example

Consider the bivariate structural autoregression of order two given by

$$\begin{aligned} x_{t,1} &= 0.7x_{t-1,1} + a_{t,1} \\ x_{t,2} &= 1.5x_{t,1} - 0.5x_{t-2,2} + a_{t,2}. \end{aligned} \tag{9}$$

This model is purposefully chosen to be simple to illustrate the point of the paper. It is represented in Figure 1a by a DAG linking nodes up to lag 2. Figure 2a shows the CIG which may be deduced from this. It contains a new link corresponding to the moralization rules (Whittaker, 1990, p. 75–77). From this CIG alone we are not able to rule out the possibility that the original model contained a directed link, $x_{t,1} \leftarrow x_{t-2,2}$, in the place of this new link.

Consideration of the CIG with nodes up to lag 3 does rule this out. Figure 2a shows the DAG extended to this lag by simply shifting the structure backwards. The solid lines of Figure 2b shows the CIG derived from this; the broken line on this figure is that of a link $x_{t-1,1} - x_{t-2,2}$ which would not arise if the link $x_{t,1} \leftarrow x_{t-2,2}$ were not present in the original model. A test which confirms the absence of the broken line therefore identifies the correct model.

We have simulated samples of length 200 from (9) and applied the tests of section 2 for the presence of the link $x_{t-1,1} - x_{t-2,2}$. Using 1000 samples and a nominal size of 5%, we found the correct test rejected the link for 4.8% of samples. The incorrect test rejected for 10.9% . In this example the variance of the t statistic using the correct value for $\hat{\sigma}_\beta$ was 0.96, and that using the incorrect variance was 1.47. The importance of using the correct test is clearly demonstrated.

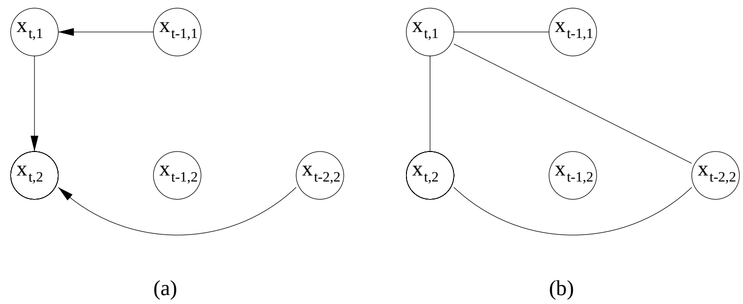

Figure 1: (a) The DAG representation of the model 9 to lag 2, and (b) its corresponding CIG

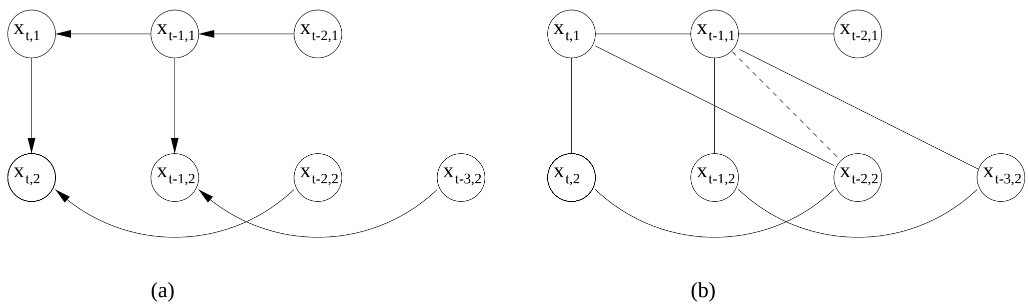

Figure 2: (a) The DAG representation of the model 9 to lag 3, and (b) its corresponding CIG with the broken line showing the link to be tested.

References

Greene, W.H. (1993), *Econometric Analysis*, second edition, Prentice-Hall: Englewood Cliffs.

Priestley, M.B. (1994), *Spectral Analysis and Time Series*, Academic Press: London.

Reale, M. and G. Tunnicliffe Wilson, "Identification of vector AR and ARMA models with recursive structural errors using conditional independence graphs", *submitted.*

Reinsel, G.C. (1993), *Elements of Multivariate Time Series Analysis*, Springer-Verlag: New York.

Whittaker, J.C. (1990), *Graphical Models in Applied Multivariate Statistics*, Wiley: Chichester.