A Decomposition of Block Toeplitz Matrices with Applications to Vector Time Series

Fallaw Sowell

Graduate School of Industrial Administration

Carnegie Mellon University

November 1989

[1]This paper presents efficient procedures to calculate the inverse and the determinant of the covariance matrix of a stationary vector time series. If normal errors are assumed, these calculations are required to evaluate the unconditional likelihood function. It is also shown how matrices derived from these procedures can be used to simulate data for any vector covariance stationary model.

Three related results are presented. The first is a recursion to calculate the forward and backward partial correlation coefficients the ordinary least squares (OLS) coefficients for a stationary vector time series. This recursion makes it possible to quickly calculate the partial correlation coefficients. Given these coefficients the mean square error from the best linear prediction given n lagged valued can be derived.

The second result used either the forward the backward partial correlation coefficients and the variances of the prediction errors to calculate the inverse and the determinant of the covariance matrix of a sample from the time series. This result is motivated by the inversion of the covariance matrix of a stationary time series model, however this procedure can be used to calculate the inverse of any nonsingular block Toeplitz matrix. The number of calculations required to invert an arbitrary $kN \times kN$ matrix grows at the rate $(kN)^3$. For this procedure the calculations required to invert the block Toeplitz matrix grows at the rate k^3N^2.

The first and second results can be viewed more generally as a procedure to decompose a Block Toeplitz matrix. The autocovariances in the first result can be replaced with the corresponding elements of the first block row or block column of any (nonsingular) block

[1]This is an extension of Chapter 4 of Sowell(1986). Several mistakes in that earlier work have been corrected. I would like to thank John Geweke and Tony Smith for helpful suggestions and comments.

Toeplitz matrix. The matrices derived by the recursion in the first result have natural interpretations in a stationary time series framework, but in general can be used in the second result without this interpretation.

The final result concerns the simulation of data for a time series model. The Cholesky decomposition of the inverse of the covariance matrix can also be formed with the partial correlation coefficients and variances of the prediction errors. If this is written $CC' = \Sigma(n+1)^{-1}$ and $U_{n+1} \sim (0, I_{k(n+1)})$ then $(C')^{-1}U_{n+1}$ will be normally distributed with zero mean and covariance $\Sigma(n+1)$. This simulation procedure is not an approximation but is exact, i.e. the simulated sample was the same theoretical moments as the model. If the time series model under consideration does not have normal errors, the simulation is exact in the sense that the simulated data and the time series will have the same spectral density, i.e. they will be identical up to the second moments.

The results presented are not original. However, proofs of all the procedures have not been published. The proofs are original and show how and why the procedures work. The first and second result were first presented in the statistical literature by Peter Whittle (1963) and has been used by several econometricians in particular see Geweke (1986).

1 Notation

Let x_t be a $k \times 1$ covariance stationary vector time series. The $t-s$ autocovariance, $\mathrm{E}(x_t x_s')$, will be denoted $\gamma(t-s)$; note $\gamma(s-t) = \gamma(t-s)'$. The following notation will be used[2]

$$X(n) = [\ x_{t-1}' \quad x_{t-2}' \quad \dots \quad x_{t-n}'\ ]',$$

$$\bar{X}(n) = [\ x_{t+1}' \quad x_{t+2}' \quad \dots \quad x_{t+n}'\ ]',$$

$$\mathrm{E}[X(n)X(n)'] = \Sigma(n) = \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(n-1) \\ \gamma(-1) & \gamma(0) & \dots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(1-n) & \gamma(2-n) & \dots & \gamma(0) \end{bmatrix},$$

$$\mathrm{E}[\bar{X}(n)\bar{X}(n)'] = \bar{\Sigma}(n) = \begin{bmatrix} \gamma(0) & \gamma(-1) & \dots & \gamma(1-n) \\ \gamma(1) & \gamma(0) & \dots & \gamma(2-n) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \dots & \gamma(0) \end{bmatrix},$$

$$\mathrm{E}[X(n)x_t'] = \Gamma(n) = \begin{bmatrix} \gamma(-1) \\ \gamma(-2) \\ \vdots \\ \gamma(-n) \end{bmatrix}$$

[2] It is customary in time series analysis to stack a data in a column vector with the latest observation at the top and the most recent at the bottom. If this conventions is used the data vector corresponds to $\bar{X}(n)$ in the notation in this paper and the covariance matrix is $\bar{\Sigma}(n)$.

and

$$\mathrm{E}[\bar{X}(n)x_t'] = \bar{\Gamma}(n) = \begin{bmatrix} \gamma(1) \\ \gamma(2) \\ \vdots \\ \gamma(n) \end{bmatrix}.$$

Let J_k denote the $k \times k$ matrix with ones along the minor diagonal (bottom left corner to top right corner) and zeros in all other positions. Premultiplying with the matrix J_k reverses the order of the elements of a column vector. If I_n denotes the $n \times n$ identity matrix, then $\Sigma(n)$ and $\bar{\Sigma}(n)$ are related by

$$\Sigma(n) = (J_n \otimes I_k)\bar{\Sigma}(n)(J_n \otimes I_k) \tag{1}$$

where $\otimes$ denotes the Kronecker product. Using this, $\Sigma(n)^{-1}$ can be calculated in two ways, either directly or by first calculating $\bar{\Sigma}(n)^{-1}$ and then using

$$\Sigma(n)^{-1} = (J_n \otimes I_k)\bar{\Sigma}(n)^{-1}(J_n \otimes I_k).$$

The matrix $(J_n \otimes I_k)$ will be denoted $\mathcal{J}_n$.

The partitioned inverse formula

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1}[I + A_{12}E^{-1}A_{21}A_{11}^{-1}] & -A_{11}^{-1}A_{12}E^{-1} \\ -E^{-1}A_{21}A_{11}^{-1} & E^{-1} \end{bmatrix} \tag{2}$$

where $E = A_{22} - A_{21}A_{11}^{-1}A_{12}$, will be used.

2 Calculation of the Coefficients and Variances

This section will be concerned with calculating the forward and backward partial correlation coefficients for the vector time series x_t. The backward partial correlation coefficients are defined as the coefficients which minimize the mean square error from the linear prediction of x_t given n lagged values of x_t. Similarly, the forward partial correlation coefficients are defined as the coefficients which minimize the mean square error from the linear prediction of x_t given n future values of x_t.

2.1 Backward partial correlations and variances

One way to think of the backward partial correlation coefficients is as the coefficients of the projection of x_t on n lagged values. This is equivalent to the solution of the problem

$$\min_{\alpha(n)} \qquad \mathrm{E}[x_t - \alpha(n)'X(n)][x_t - \alpha(n)'X(n)]'$$

where $\alpha(n)' = [\ \alpha(n,1) \quad \alpha(n,2) \quad \ldots \quad \alpha(n,n)\ ]$. The values of the coefficients $\alpha(n)$ which minimize the prediction variance, which will be denoted $A(n)$, are

$$A(n) = \Sigma(n)^{-1}\Gamma(n). \tag{3}$$

The minimum value of the prediction variance will be denoted $v(n)$ and can be written

$$v(n) = \gamma(0) - \Gamma(n)'\Sigma(n)^{-1}\Gamma(n) = \gamma(0) - A(n)'\Gamma(n). \tag{4}$$

This variance matrix is the minimum in the sense that the variance matrix associated with any other linear predictor will differ from $v(n)$ by a positive definite matrix.

2.2 Forward partial correlations and variances

Similarly the projection of x_t on n future values can be though of solving the problem

$$\min_{\bar{\alpha}(n)} \qquad \mathrm{E}[x_t - \bar{\alpha}(n)'\bar{X}(n)][x_t - \bar{\alpha}(n)'\bar{X}(n)]'.$$

The values of the coefficients $\bar{\alpha}(n)$ which minimize the prediction variance, which will be denoted $\bar{A}(n)$, are

$$\bar{A}(n) = \bar{\Sigma}(n)^{-1}\bar{\Gamma}(n). \tag{5}$$

The minimum value of the prediction variance will be denoted $\bar{v}(n)$ and written

$$\bar{v}(n) = \gamma(0) - \bar{\Gamma}(n)'\bar{\Sigma}(n)^{-1}\bar{\Gamma}(n) = \gamma(0) - \bar{A}(n)'\bar{\Gamma}(n). \tag{6}$$

2.3 The recursion

The coefficients associated with x_t projected on n lagged values, $A(n)$, will be calculated simultaneously with the coefficients associated with the projection of x_t on n future values, $\bar{A}(n)$. The n coefficients can be calculated recursively from the $n-1$ coefficients. Expanding equation (3) gives

$$A(n+1) = \begin{bmatrix} A(n+1,1)' \\ A(n+1,2)' \\ \vdots \\ A(n+1,n)' \\ A(n+1,n+1)' \end{bmatrix} = \begin{bmatrix} \Sigma(n) & \mathcal{J}_n\bar{\Gamma}(n) \\ \bar{\Gamma}(n)'\mathcal{J}_n & \gamma(0) \end{bmatrix}^{-1} \begin{bmatrix} \Gamma(n) \\ \gamma(-n-1) \end{bmatrix}$$

using the partitioned inverse formula and equations (5) and (6) this can be written

$$= \begin{bmatrix} \Sigma(n)^{-1} + \mathcal{J}_n\bar{A}(n)\bar{v}(n)^{-1}\bar{\Gamma}(n)'\mathcal{J}_n\Sigma(n)^{-1} & -\mathcal{J}_n\bar{A}(n)\bar{v}(n)^{-1} \\ -\bar{v}(n)^{-1}\bar{\Gamma}(n)'\mathcal{J}_n\Sigma(n)^{-1} & \bar{v}(n)^{-1} \end{bmatrix} \begin{bmatrix} \Gamma(n) \\ \gamma(-n-1) \end{bmatrix}.$$

It follows that

$$A(n+1,n+1)' = \bar{v}(n)^{-1}\left[\gamma(-n-1) - \bar{\Gamma}(n)'\mathcal{J}_n A(n)\right] \tag{7}$$

and

$$\begin{bmatrix} A(n+1,1)' \\ A(n+1,2)' \\ \vdots \\ A(n+1,n)' \end{bmatrix} = A(n) - \mathcal{J}_n \bar{A}(n) A(n+1,n+1)'. \tag{8}$$

This shows how to calculated $A(n+1,n+1)$ given $A(n)$ and $\bar{v}(n)$. Using $A(n)$, $\bar{A}(n)$ and $A(n+1,n+1)$ it is possible to obtain $A(n+1,k)$ for $k = 1,2,\ldots,n$. Once $A(n+1)$ is calculated, $v(n+1)$ follows by equation (4).

The same argument implies the coefficients of $\bar{A}(n+1,n+1)$ can be derived by the equations

$$\bar{A}(n+1,n+1)' = v(n)^{-1}\left[\gamma(n+1) - \Gamma(n)'\mathcal{J}_n\bar{A}(n)\right] \tag{9}$$

and

$$\begin{bmatrix} \bar{A}(n+1,1)' \\ \bar{A}(n+1,2)' \\ \vdots \\ \bar{A}(n+1,n)' \end{bmatrix} = \bar{A}(n) - \mathcal{J}_n A(n) \bar{A}(n+1,n+1)'. \tag{10}$$

These formulas can be summarized in the following recursion

Result 1 *Let* $v(0) = \bar{v}(0) = \gamma(0)$*,* $D(1) = \gamma(1)$ *and* $\bar{D}(1) = \gamma(-1)$*, then for* $n = 1,2,\ldots,T-1$

$$D(n) = \gamma(n) - \sum_{j=1}^{n-1} A(n-1,n-j)\gamma(j)$$

$$\bar{D}(n) = \gamma(-n) - \sum_{j=1}^{n-1} \bar{A}(n-1,n-j)\gamma(-j)$$

$$A(n,n) = D(n)\bar{v}(n-1)^{-1}$$

$$\bar{A}(n,n) = \bar{D}(n)v(n-1)^{-1}$$

and for $k = 1, 2, \ldots, n-1$

$$A(n,k) = A(n-1,k) - A(n,n)\bar{A}(n-1,n-k)$$

$$\bar{A}(n,k) = \bar{A}(n-1,k) - \bar{A}(n,n)A(n-1,n-k)$$

$$v(n) = \gamma(0) - \sum_{j=1}^{n} A(n,j)\gamma(-j)$$

and

$$\bar{v}(n) = \gamma(0) - \sum_{j=1}^{n} \bar{A}(n,j)\gamma(j).$$

3 Calculating The Inverse and Determinant of a Block Toeplitz Matrix

Using the partitioned inverse formula, $\Sigma(n+1)^{-1}$ can be written in terms of $\bar{A}(n)$, $\bar{v}(n)$ and $\Sigma(n)$

$$\begin{aligned}
\Sigma(n+1)^{-1} &= \begin{bmatrix} \Sigma(n)^{-1} + \mathcal{J}_n\bar{A}(n)\bar{v}(n)^{-1}\bar{A}(n)'\mathcal{J}_n & -\mathcal{J}_n\bar{A}(n)\bar{v}(n)^{-1} \\ -\bar{v}(n)^{-1}\bar{A}(n)'\mathcal{J}_n & \bar{v}(n)^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \Sigma(n)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \\
&+ \begin{bmatrix} 0 & -\mathcal{J}_n\bar{A}(n)\bar{v}(n)^{-1/2} \\ 0 & \bar{v}(n)^{-1/2} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ -\bar{v}(n)^{-1/2}\bar{A}(n)'\mathcal{J}_n & \bar{v}(n)^{-1/2} \end{bmatrix}.
\end{aligned} \tag{11}$$

Now, for $j = 0, 1, \ldots, n$, define the $(n+1) \times (n+1)$ partitioned matrix

$$\bar{B}(j) = \begin{bmatrix} 0 & -\mathcal{J}_j \bar{A}(j) \bar{v}(j)^{-1/2} & 0 \\ 0 & \bar{v}(j)^{-1/2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where $\bar{A}(0) = I_k$. Using this notation and repeated applications of equation (11), the inverse of the covariance matrix can be written

$$\Sigma(n+1)^{-1} = \sum_{j=0}^{n} \bar{B}(j)\bar{B}(j)'$$

Result 2 *Because* $\bar{B}(j)\bar{B}(i)' = 0$ *if* $j \neq i$, *the inverse of the covariance matrix can be written*

$$\Sigma(n+1)^{-1} = \left(\sum_{j=0}^{n} \bar{B}(j)\right)\left(\sum_{i=0}^{n} \bar{B}(i)\right)' = \bar{\beta}_{n+1}\bar{\beta}'_{n+1}$$

where

$$\bar{\beta}_{n+1} = \begin{bmatrix} I_k & -\bar{A}(1,1)' & -\bar{A}(2,2)' & \ldots & -\bar{A}(n,n)' \\ 0 & I_k & -\bar{A}(2,1)' & \ldots & -\bar{A}(n,n-1)' \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \ldots & & 0 & I_k \end{bmatrix} \begin{bmatrix} \bar{v}(0) & 0 & 0 & \ldots & 0 \\ 0 & \bar{v}(1) & 0 & \ldots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & & 0 & \bar{v}(n) \end{bmatrix}^{-1/2}$$

The same procedure applied to $\bar{\Sigma}(n+1)^{-1}$ shows that

$$\bar{\Sigma}(n+1)^{-1} = \beta_{n+1}\beta'_{n+1}$$

where

$$\beta_{n+1} = \begin{bmatrix} I_k & -A(1,1)' & -A(2,2)' & \ldots & -A(n,n)' \\ 0 & I_k & -A(2,1)' & \ldots & -A(n,n-1)' \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \ldots & & 0 & I_k \end{bmatrix} \begin{bmatrix} v(0) & 0 & 0 & \ldots & 0 \\ 0 & v(1) & 0 & \ldots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \ldots & & 0 & v(n) \end{bmatrix}^{-1/2}.$$

The determinant of $\Sigma(n+1)$ is obtained from either β_{n+1} or $\bar{\beta}_{n+1}$ by their triangular form

$$|\Sigma(n+1)| = \prod_{j=0}^{n} |v(j)| = \prod_{j=0}^{n} |\bar{v}(j)|.$$

4 Data Simulation

From equation (1) it is clear that

$$\Sigma(n)^{-1} = \bar{\beta}_n \bar{\beta}'_n$$

and that

$$\Sigma(n) = (\bar{\beta}'_n)^{-1} \bar{\beta}_n^{-1}$$

hence, to simulate data for $X(n)$ it is necessary to calculate $(\bar{\beta}'_n)^{-1}$. The simulated data will be achieved by starting with a column vector of n multivariate iid observations, $U_n \sim (0, I_{kn})$. The simulated sample $Z_n = (\bar{\beta}'_n)^{-1} U_n$ is then created. Using the partitioned inverse formula this can be written

$$Z_n = (\bar{\beta}'_n)^{-1} U_n = \begin{bmatrix} \bar{\beta}'_{n-1} & 0 \\ -\bar{v}(n-1)^{-1/2} \bar{A}(n-1)' \mathcal{J}_{n-1} & v(n-1)^{-1/2} \end{bmatrix}^{-1} \begin{bmatrix} U_{n-1} \\ u_n \end{bmatrix}$$

$$= \begin{bmatrix} (\beta'_{n-1})^{-1} & 0 \\ \bar{A}(n-1)' \mathcal{J}_{n-1} (\beta'_{n-1})^{-1} & \bar{v}(n-1)^{1/2} \end{bmatrix} \begin{bmatrix} U_{n-1} \\ u_n \end{bmatrix}$$

The last equation shows how z_n can be evaluated from Z_{n-1} by the formula

$$z_n = \bar{A}(n-1)' \mathcal{J}_{n-1} Z_{n-1} + \bar{v}(n-1)^{1/2} u_n.$$

Repeated applications of this proves the following recursive procedure to simulated data.

Result 3 *A synthetic sample for $X(n)$, which will be denoted Z_n, can be created recursively by defining $z_1 = \bar{v}(0)^{1/2}u_1$ and for t=2,3,...*

$$z_n = \sum_{j=1}^{t-1} \bar{A}(t-1, t-j)z_j + \bar{v}(t-1)^{1/2}u_t.$$

Similarly, a sample $\bar{Z}_n = (\beta_n')^{-1}U_n$ can be calculated recursively, define $\bar{z}_1 = v(0)^{1/2}u_1$ and for $t = 2, 3, \ldots$

$$\bar{z}_n = \sum_{j=1}^{t-1} A(t-1, t-j)\bar{z}_j + v(t-1)^{1/2}u_t.$$

by construction Z_n and $\mathcal{J}_n\bar{Z}_n$ are distributed $(0, \Sigma(n))$.

5 The Univariate Model

The covariance matrix $\Sigma(n)$ and $\bar{\Sigma}(n)$, for all k, are related by

$$\Sigma(n) = \mathcal{J}_n\bar{\Sigma}(n)\mathcal{J}_n.$$

This implies a major simplification for the univariate model. If $k = 1$, then $\Sigma(n) = \mathcal{J}_n\bar{\Sigma}(n)\mathcal{J}_n = \bar{\Sigma}(n)$, because in the univariate model $\gamma(n) = \gamma(n)'$. For the univariate model, the projection of x_t on lagged valued is equivalent to the projection on future values, this implies $A(n) = \bar{A}(n)$. This simplifies the recursive procedure to calculate the coefficients and variances to

$$A(n,n) = \frac{\gamma(n) - \Gamma(n-1)'\mathcal{J}_{n-1}A(n-1)}{v(n-1)}$$

$$\begin{bmatrix} A(n,1) \\ A(n,2) \\ \vdots \\ A(n,n) \end{bmatrix} = A(n-1) - A(n,n)\mathcal{J}_{n-1}A(n-1)$$

and

$$v(n) = \gamma(0) - A(n)'\Gamma(n)$$

The values of $A(n)$ and $v(n)$ can be used in the formulas above to calculate the inverse and determinant of the covariance matrix and to simulate data.

References

[1] Geweke, John F., "The Superneutrality of Money in the United STates: an Interpretation of the Evidence." Econometrica 54, 1986: 1-22.

[2] Sowell, Fallaw, "Fractionally Integrated Vector Time Series," Ph.D. Thesis, Department of Economics, Duke University, 1986.

[3] Whittle, Peter "On the Fitting of Multivariate Autoregression, and the Approximate Canonical Factorization of a Spectral Density Matrix." Biometrica 50, 1963: 129-134.