

Multivariate Time Series - Sample Exam Solutions

TA: Jean Flemming

April 2, 2013

1. VAR(p) of $n \times 1$ time series vector y_t : $\Pi(L)y_t = u_t$ (1)

where $\Pi(L) = I_n - \Pi_1 L - \dots - \Pi_p L^p$, $u_t \sim iid N_n(0, \Omega)$

(a) Rewrite the VAR in reduced form:

$$y_t = \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + u_t \quad (2)$$

Write the system in companion form:

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \Pi_1 & \Pi_2 & \dots & \Pi_p \\ I & 0 & \dots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & I & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where the companion matrix is defined as

$$F = \begin{bmatrix} \Pi_1 & \Pi_2 & \dots & \Pi_p \\ I & 0 & \dots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & I & 0 \end{bmatrix}$$

The system is stable if the eigenvalues of F are less than one in modulus.

(b) Assume $u_t = B^{-1}\epsilon_t$, where $\epsilon_t \sim N_n(0, \Sigma)$ where Σ is a diagonal variance-covariance matrix. Then (1) can be written as:

$$\Pi(L)y_t = B^{-1}\epsilon_t$$

Multiplying both sides by B gives the structural VAR in lag operator notation:

$$B\Pi(L)y_t = \epsilon_t \quad (3)$$

Or equivalently:

$$By_t = B\Pi_1 y_{t-1} + \dots + B\Pi_p y_{t-p} + \epsilon_t$$

For identification of the structural VAR, again write the system in reduced form as in (2):

$$\begin{aligned} y_t &= \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + B^{-1} \epsilon_t \\ &= \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + u_t \end{aligned}$$

To obtain identification, check the order condition:

knowns: pn^2 elements in Φ_1, \dots, Φ_p , and $n(n+1)/2$ elements in the diagonal variance matrix Ω_e

unknowns: pn^2 elements in Π_1, \dots, Π_p , $n(n-1)$ parameters in the matrix B (since diagonal terms of B are 1), and n elements in the diagonal variance matrix Σ

$$\Rightarrow \#knowns - \#unknowns = n(n-1)/2$$

Therefore we need $n(n-1)/2$ restrictions for identification.

Imposing long run restrictions assumes that the j^{th} component of the structural error term ϵ_t has no cumulative long run impact on the i^{th} component of y_t . We can rewrite (3) as a VMA, renaming $(B\Pi(L))^{-1} = \Theta(L)$:

$$y_t = \Theta(L)\epsilon_t$$

In matrix form, we have:

$$\begin{aligned} \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{n,t} \end{bmatrix} &= \begin{bmatrix} \Theta_{11}^{(0)} & \dots & \Theta_{1n}^{(0)} \\ \vdots & \ddots & \vdots \\ \Theta_{n1}^{(0)} & \dots & \Theta_{nn}^{(0)} \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{n,t} \end{bmatrix} + \begin{bmatrix} \Theta_{11}^{(1)} & \dots & \Theta_{1n}^{(1)} \\ \vdots & \ddots & \vdots \\ \Theta_{n1}^{(1)} & \dots & \Theta_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} \epsilon_{1,t-1} \\ \epsilon_{2,t-1} \\ \vdots \\ \epsilon_{n,t-1} \end{bmatrix} + \dots \\ \Theta(1) &= \begin{bmatrix} \sum_{i=1}^{\infty} \Theta_{11}^{(i)} & \dots & \sum_{i=1}^{\infty} \Theta_{1n}^{(i)} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{\infty} \Theta_{n1}^{(i)} & \dots & \sum_{i=1}^{\infty} \Theta_{nn}^{(i)} \end{bmatrix} \end{aligned}$$

Since our restrictions concern the long run impact of error j on variable i , this is equivalent to saying that the ij^{th} component of matrix $\Theta(1)$ is equal to zero. The SVAR representation in this case would be:

$$\tilde{\Pi}(L)y_t = \epsilon_t$$

where $\tilde{\Pi}(L) = \Theta(L)^{-1}$ imposing the long run restrictions. This type of restriction makes the matrix $\Theta(1)$ lower triangular, giving us the necessary $n(n-1)/2$ restrictions for identification.

- (c) i. Writing the SVAR in the SMA representation:

$$y_t = \Theta(L)\epsilon_t = \Theta_0\epsilon_t + \dots + \Theta_k\epsilon_{t-k} + \dots$$

The impulse response function measures the effect of the j^{th} element of ϵ_t on the i^{th} element of y_{t+k} by plotting the ij^{th} element of matrix Θ_k as a function of k .

For example, to estimate impulse responses for a VAR(1), we first estimate Π_1 in the reduced form VAR, then, given $\hat{\Pi}_1$, estimate the matrix $\Psi(L) = \Pi(L)^{-1}$ through the equation: $\hat{\Psi}_k = \hat{\Pi}_1^k$. Using the long run restrictions on the SVAR from part b), we can recover an estimate of the matrix B , and finally can compute estimates of Θ_k with $\hat{\Theta}_k = \hat{\Psi}_k \hat{B}^{-1}$ for $k = 0, 1, \dots$

- ii. The forecast error variance decomposition is calculated by computing the h-step-ahead forecast errors of y_t . Let

$$y_{t|t-h} = E(y_t | \{u_s\}_{s=-\infty}^{t-h})$$

denote the h-step-ahead forecast of y_t at time $t - h$, and the resulting forecasting error is then

$$u_{t|t-h} = y_t - y_{t|t-h} = \sum_{k=1}^{h-1} C(L)u_{t-k}$$

where again $C(L) = \Pi(L)^{-1}$ from (1). Since Ω is assumed to be diagonal, assume its diagonal elements are σ_j^2 . Then the variance of the i^{th} element of $u_{t|t-h}$ is

$$\sum_{j=1}^n [\sigma_j^2 \sum_{k=0}^{h-1} c_{ij,k}^2]$$

where $c_{ij,k}$ is the ij^{th} element of C_k . The variance decomposition of $y_{i,t}$ at horizon h is given by:

$$\rho_{ij,h}^2 = \frac{\sigma_j^2 \sum_{k=0}^{h-1} c_{ij,k}^2}{\sum_{j=1}^n [\sigma_j^2 \sum_{k=0}^{h-1} c_{ij,k}^2]}$$

$$2. \text{ VAR}(p) \text{ of } nx1 \text{ time series vector } Y_t: \quad \Pi(L)Y_t = \Phi D_t + \epsilon_t \quad (3)$$

where $\Pi(L) = \Pi_0 - \Pi_1 L - \dots - \Pi_p L^p$,

$\epsilon_t \sim iid N_n(0, \Omega)$, and D_t is a deterministic n -dimensional vector.

- (a) The process is CI(1,1) if there exists an α such that $y_t - \alpha y_{t-1}$ is I(0). Normalizing $\Pi_0 = I_n$ and rewriting the system in first differences:

$$\Delta Y_t = \Phi D_t + \Pi Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_p \Delta Y_{t-p} + \epsilon_t$$

where $\Pi = \Pi_1 + \Pi_2 + \dots + \Pi_p - I_n$

To find unit roots, compute: $\text{rank}(\Pi)$. If the VAR in equation (3) has unit roots, then Π must be of reduced rank: $\text{rank}(\Pi) < n$. There are two cases to consider:

- i. $\text{rank}(\Pi) = 0 \Rightarrow \Pi = 0$ and Y_t is not cointegrated.
- ii. $0 < \text{rank}(\Pi) = r < n \Rightarrow Y_t$ is I(1) with r linearly independent cointegrating vectors. We can write

$$\Pi = \alpha \beta'$$

with β an rxn matrix whose rows form a basis for the r cointegrating vectors. This is not unique because the rows of β only form a basis for the cointegrating vectors, therefore they generate the entire space spanned by the cointegrating relationships.

- (b) Under the above conditions, and following Johansen (1995), the deterministic terms of the series are restricted to the form

$$\Phi D_t = \mu_0 + \mu_1 t$$

Assuming this term is a restricted trend, we have

$$\mu_t = \mu_0 + \alpha \rho_1 t$$

Rewriting the first differenced version of (3), we get:

$$\Delta Y_t = \mu_0 + \alpha(\rho_1 t + \beta' Y_{t-1}) + \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_p \Delta Y_{t-p} + \epsilon_t$$

Where μ_0 is the drift and $\rho_1 t$ is the linear trend.

By the Granger Representation Theorem, we have

$$\Psi(1) = \beta_{\perp} (\alpha'_{\perp} \Gamma_1 \beta_{\perp})^{-1} \alpha'_{\perp}$$

where α_{\perp} and β_{\perp} are $n \times (n + r)$ full rank matrices.

β_{\perp} is the common trends loading matrix, and the common trend is given by

$$TS_t = \Psi(1) \sum_{k=1}^t u_k$$

- (c) Johansen's trace statistic tests for cointegration with a null hypothesis of r_0 cointegrating vectors against the alternative of $r > r_0$ cointegrating vectors, while the eigenvalue statistic tests whether there are r_0 versus $r_0 + 1$ cointegrating relationships. Clearly, the alternative hypotheses are not the same for both tests. Both tests are likelihood ratio tests.

The trace test considers the following hypotheses:

$$\begin{aligned} H_0 : & \quad r = r_0 \\ H_1 : & \quad r > r_0 \end{aligned}$$

And the trace statistic is calculated as:

$$J_{trace}(r_0) = -T \sum_{i=r_0+1}^n \log(1 - \hat{\lambda}_i)$$

where $\hat{\lambda}_i$ $i = 1, \dots, n$ are the estimated eigenvalues of Π and T is the length of the time series.

The maximum eigenvalue test considers the following hypotheses:

$$\begin{aligned} H_0 : & \quad r = r_0 \\ H_1 : & \quad r = r_0 + 1 \end{aligned}$$

And the maximum eigenvalue statistic is calculated as:

$$J_{max}(r_0) = -T \log(1 - \hat{\lambda}_{r_0+1})$$