Time Series Econometrics
Supplementary Lecture 6

Yukai Yang

Department of Statistics
Uppsala University

1 Exercise 20.2

It was claimed in the text that the maximized log likelihood function under the null hypothesis of h cointegrating relations was given by [20.3.2]. What is the nature of the restriction on the VAR in [20.3.1] when $h = 0$? Show that the value of [20.3.2] for this case is the same as the log likelihood for a VAR($p-1$) process fitted to the differenced data $\Delta \mathbf{y}_t$.

We are now using the error-correction form of the VAR:

$$\Delta \mathbf{y}_t = \boldsymbol{\zeta}_1 \Delta \mathbf{y}_{t-1} + \cdots + \boldsymbol{\zeta}_{p-1} \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\alpha} + \boldsymbol{\zeta}_0 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t. \tag{1}$$

Any VAR of order p can be written in this way. What is special here is that under H_0 of h cointegrating relations, we can write

$$\boldsymbol{\zeta}_0 = -\mathbf{B}\mathbf{A}',$$

for $\mathbf{B}$ $n \times h$ and $\mathbf{A}$ $n \times h$ full rank. The maximized log likelihood is given by:

$$\mathcal{L}^* = -(Tn/2)\log(2\pi) - (Tn/2) - (T/2)\log|\hat{\boldsymbol{\Sigma}}_{\mathbf{UU}}| - (T/2)\sum_{i=1}^{h}\log(1-\hat{\lambda}_i) \tag{2}$$

where $\hat{\boldsymbol{\Sigma}}_{\mathbf{UU}}$ is the covariance matrix of the residuals in the auxiliary regression

$$\Delta \mathbf{y}_t = \hat{\boldsymbol{\pi}}_0 + \hat{\boldsymbol{\Pi}}_1 \Delta \mathbf{y}_{t-1} + \cdots + \hat{\boldsymbol{\Pi}} \Delta \mathbf{y}_{t-p+1} + \hat{\mathbf{u}}_t. \tag{3}$$

The $\hat{\lambda}_1 > \cdots > \hat{\lambda}_p$ are eigenvalues found from

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{VV}}^{-1} \hat{\boldsymbol{\Sigma}}_{\mathbf{VU}} \hat{\boldsymbol{\Sigma}}_{\mathbf{UU}}^{-1} \hat{\boldsymbol{\Sigma}}_{\mathbf{UV}} \tag{4}$$

where $\hat{\mathbf{\Sigma}}_{\mathbf{VV}}$ is the covariance matrix of the residuals in the regression

$$\mathbf{y}_{t-1} = \hat{\boldsymbol{\theta}}_0 + \hat{\mathbf{\Theta}}_1 \Delta \mathbf{y}_{t-1} + \cdots + \hat{\mathbf{\Theta}}_{p-1} \Delta \mathbf{y}_{t-p+1} + \hat{\mathbf{v}}_t \tag{5}$$

and similarly, $\hat{\mathbf{\Sigma}}_{\mathbf{UV}} = T^{-1} \sum_{t=1}^{T} \hat{\mathbf{u}}_t \hat{\mathbf{v}}_t'$. Here, you can note that the idea here is based on the Frisch-Waugh-Lovell theorem. We are correcting both $\Delta \mathbf{y}_t$ and $\mathbf{y}_{t-1}$ for the lagged differences. By doing so, we can reduce the problem to a simpler regression:

$$\begin{aligned} \hat{\mathbf{u}}_t &= \boldsymbol{\zeta}_0 \hat{\mathbf{v}}_t + \boldsymbol{\varepsilon}_t \\ &= -\mathbf{B}\mathbf{A}' \hat{\mathbf{v}}_t + \boldsymbol{\varepsilon}_t. \end{aligned}$$

Maximization of the log likelihood is then done in steps. The log likelihood of the above model is maximized with respect to $\mathbf{B}$ and $\mathbf{A}$. Given the estimate $\hat{\mathbf{A}}$, everything else can be found by usual regression techniques (or by exploiting algebraic identities, as in the book).

Since h is the number of cointegrating relations, if $h = 0$, that means that there are no cointegrating relations. No linear combination of the variables is stationary. This means that $\boldsymbol{\zeta}_0 = \mathbf{0}$. For the log likelihood in (2), the last term drops out so

$$\mathcal{L}^* = -(Tn/2)\log(2\pi) - (Tn/2) - (T/2)\log|\hat{\mathbf{\Sigma}}_{\mathbf{UU}}|. \tag{6}$$

Compare this to the log likelihood for a VAR(p) in chapter 11 [11.1.32]:

$$\mathcal{L} = -(Tn/2)\log(2\pi) + (T/2)\log|\hat{\mathbf{\Omega}}^{-1}| - (Tn/2)$$

where $\hat{\mathbf{\Omega}}$ is the residual covariance matrix. As you can see, these are the same (since $\log|\mathbf{A}^{-1}| = -\log|\mathbf{A}|$), so the log likelihood of the model in (1) when $h = 0$ is the same as the log likelihood for a VAR($p-1$) in $\Delta\mathbf{y}_t$. If $h > 0$, then these would not be the same. What this tells us is that if we have some data and neglect cointegration, we are implicitly assuming $h = 0$. If, in reality, $h = 1$, for example, we would be better off by taking that into account.