EC402: MEI. Spring Term: The Analysis of Time Series*

Problem Set 3 – Solutions

1. Cointegration

Suppose that aggregate income, Y_t, follows a random walk with drift

$$Y_t = \mu_y + Y_{t-1} + \varepsilon_{yt}$$

and that the government always spends a fraction of the previous period output

$$G_t = \mu_g + gY_{t-1} + \varepsilon_{gt} \tag{1.1}$$

where $0 < g < 1$, ε_{yt} and ε_{gt} are mean zero serially independent *iid* variables.

Moreover, assume that the government runs a balanced budget ($T_t = G_t \ \forall t$) and that consumption follows

$$C_t = \alpha + \beta (Y_t - T_t) \tag{1.2}$$

with $\alpha, \beta > 0$

1. Are Y, C and G stationary? What are their orders of integration?

Answer *Y is a non stationary $I(1)$ variable, as well as G from equation (1.1). Substituting for T_t and Y_t in (1.2) we have that*

$$\begin{aligned} C_t &= \alpha + \beta \left(\mu_y + Y_{t-1} + \varepsilon_{yt} - \mu_g - gY_{t-1} - \varepsilon_{gt}\right) \\ &\propto \beta (1-g) Y_{t-1} + \beta (\varepsilon_{yt} - \varepsilon_{gt}) \end{aligned}$$

Therefore, since $0 < g < 1$, C_t is also a non stationary $I(1)$ variable.

2. Are C_t and G_t cointegrated? Are Y_t and G_t cointegrated? Are C_t and Y_t cointegrated? Are C_t, Y_t and G_t cointegrated?

Answer *If C_t and G_t are cointegrated there $\exists \lambda_1$ s.t. $C_t - \lambda_1 G_t \sim I(0)$.*

$$\begin{aligned} C_t - \lambda_1 G_t &= \alpha + \beta Y_t - (\beta + \lambda_1) G_t \\ &= \alpha + \beta \left(\mu_y + Y_{t-1} + \varepsilon_{yt}\right) - (\beta + \lambda_1) \left(\mu_g + g Y_{t-1} + \varepsilon_{gt}\right) \\ &\propto \beta Y_{t-1} - (\beta + \lambda_1) g Y_{t-1} + (\beta \varepsilon_{yt} - (\beta + \lambda_1) \varepsilon_{gt}) \end{aligned}$$

Since $Y_{t-1}$is $I(1)$, for cointegration among these variables we need $\beta - (\beta + \lambda) g = 0$ that is satisfied by $\lambda_1 = \frac{1}{g}(\beta - \beta g)$.

Y_t and G_t are cointegrated if there $\exists \lambda_2$ s.t. $Y_t - \lambda_2 G_t \sim I(0)$.

$$Y_t - \lambda_2 G_t \propto Y_{t-1} + \varepsilon_{yt} - \lambda_2 g Y_{t-1} - \lambda_2 \varepsilon_{gt}$$

that is we need $1 - \lambda_2 g = 0$ that is satisfied by $\lambda_2 = 1/g$

C_t and Y_t are cointegrated if there $\exists \lambda_3$ s.t. $C_t - \lambda_3 Y_t \sim I(0)$.

$$\begin{aligned} C_t - \lambda_3 Y_t &\propto (\beta - \lambda_3) Y_t - \beta G_t \\ &\propto (\beta - \lambda_3) Y_{t-1} - \beta g Y_{t-1} + (\beta - \lambda_3) \varepsilon_{yt} - \beta \varepsilon_{gt} \end{aligned}$$

that is we need $(\beta - \lambda_3) - \beta g = 0$ that is satisfied by $\lambda_3 = \beta - g\beta = g\lambda_1$

C_t, Y_t and G_t are cointegrated if there $\exists \lambda_4, \lambda_5$ and s.t. $C_t - \lambda_4 Y_t - \lambda_5 G_t \sim I(0)$

$$\begin{aligned} C_t - \lambda_4 Y_t - \lambda_5 G_t &\propto (\beta - \lambda_4) Y_t - (\beta + \lambda_5) G_t \\ &\propto (\beta - \lambda_4) Y_{t-1} - (\beta + \lambda_5) g Y_{t-1} + (\beta - \lambda_4) \varepsilon_{yt} - (\beta + \lambda_5) \varepsilon_{gt} \end{aligned}$$

So we need $(\beta - \lambda_4) - (\beta + \lambda_5) g = 0$. This has infinite solutions given by $\lambda_4 = \beta - g(\beta + \lambda_5)$. One of the solutions is $\lambda_5 = -\beta$ and $\lambda_4 = \beta$.

3. How many linearly independent cointegration vectors are there? Why?

Answer *Ordering the variables as C_t, Y_t and G_t, we can write a matrix that has in each row the cointegration vectors we have found before, that is*

$$\begin{bmatrix} 1 & 0 & -\frac{1}{g}(\beta-\beta g) \\ 0 & 1 & -1/g \\ 1 & -(\beta-g\beta) & 0 \\ 1 & -\beta & \beta \end{bmatrix}$$

Note that the third row can be obtained by multiplying the second row by $(\beta - g\beta)$ and subtracting the result from the first row. That is, the third row is not linearly independent from the first two rows. Moreover, the fourth row can be obtained multiplying the second row by β and subtracting the result from the first row. That is, there are only two linearly independent cointegrating vectors among these three variables. This is due to the fact that N variables can at most share $N-1$ common trends.

4. What are the long-run trends of Y_t, G_t and C_t?

Answer

$$\begin{aligned} E\left[Y_{t+T}-Y_t\right] &= E\left[(Y_{t+T}-Y_{t+T-1})+...+(Y_{t+1}-Y_t)\right]=\mu_y T \\ E\left[G_{t+T}-G_t\right] &= gE\left[Y_{t+T-1}-Y_{t-1}\right]=g\mu_y T \\ E\left[C_{t+T}-C_t\right] &= \beta\left\{E\left[Y_{t+T}-Y_t\right]-E\left[G_{t+T}-G_t\right]\right\}=\beta(1-g)\mu_y T \end{aligned}$$

That is, the long run trend of C_t is a linear combination of the long run trends of Y_t and G_t.

2. Dynamic Simultaneous Equations

Consider the following model

$$\begin{aligned} y_{1t} &= \gamma y_{2t}+\beta x_t+\varepsilon_{1t} \\ y_{2t} &= \alpha y_{1t-1}+\varepsilon_{2t} \end{aligned}$$

where ε_{1t} and ε_{2t} are serially uncorrelated disturbances which may be contemporaneously correlated with each other.

1. Are the parameters identifiable?

Answer *Two endogenous variables* y_{1t}, y_{2t}*, two exogenous variables* $y_{1,t-1}, x_t$*. (Note* $\varepsilon_{1t}, \varepsilon_{2t}$ *must be serially uncorrelated for* y_{1t-1} *to be treated as exogenous). As* $N = 2$*,* $N - 1 = 1$*. One exclusion restriction per equation. Therefore, the first equation is just identified by order condition as well as the second equation. Both pass rank condition.*

2. Write down the final form and the autoregressive final form of the model.

Answer *By substitution, we have*

$$y_{1t} = \gamma\alpha y_{1,t-1} + \gamma\varepsilon_{2t} + \beta x_t + \varepsilon_{1t}$$

$$\therefore \quad (1 - \gamma\alpha L)y_{1t} = \beta x_t + (\varepsilon_{1t} + \gamma\varepsilon_{2t})$$

Lag the first equation of the system and substitute into the second,

$$y_{2t} = \alpha\gamma y_{2,t-1} + \alpha\beta x_{t-1} + \alpha\varepsilon_{1t-1} + \varepsilon_{2t}.$$

$$\therefore \quad (1 - \alpha\gamma L)y_{2t} = \alpha\beta x_{t-1} + (\varepsilon_{2t} + \alpha\varepsilon_{1t-1})$$

$$\therefore \quad (1 - \alpha\gamma L)\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha\beta L \end{bmatrix} x_t + \begin{bmatrix} 1 & \gamma \\ \alpha L & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \qquad \begin{array}{c} \textit{Autoregressive} \\ \textit{Final Form} \end{array}$$

$$\therefore \quad \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \frac{1}{(1-\gamma\alpha L)}\begin{bmatrix} \beta \\ \alpha\beta L \end{bmatrix} x_t + \frac{1}{(1-\alpha\gamma L)}\begin{bmatrix} 1 & \gamma \\ \alpha L & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \qquad \textit{Final Form}$$

1. What is the necessary condition for the model to be stable?

Answer *Stable iff* $|\alpha\gamma| < 1$.

3. VAR Estimation

Consider the VAR

$$y_t = c + \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + \varepsilon_t$$

where y_t is a $n \times 1$ vector of time series, c is $n \times 1$ vector of constants, the Φ's are $n \times n$ matrixes of coefficients and ε_t is a vector of disturbances s.t. $\varepsilon_t \sim iidN(0, \Omega)$.

1. Define the $(np+1)\times 1$ vector

$$x_t = \begin{bmatrix} 1 \\ y_{t-1} \\ \cdots \\ y_{t-p} \end{bmatrix}$$

and the $n \times (np+1)$ matrix

$$\Pi' = [c,\ \Phi_1, ..., \Phi_p].$$

Write down the sample log likelihood of this model.

Answer *The VAR can be rewritten as*

$$y_t = \Pi' x_t + \varepsilon_t$$

therefore

$$y_t | x_t \sim N\left(\Pi' x_t, \Omega\right).$$

implying that

$$f\left(y_t | x_t; \Pi, \Omega\right) = (2\pi)^{-n/2} \left|\Omega^{-1}\right|^{1/2} \exp\left\{-\frac{1}{2}\left(y_t - \Pi' x_t\right)' \Omega^{-1}\left(y_t - \Pi' x_t\right)\right\}.$$

Therefore, conditioning on the first p observations $(y_0, ..., y_{-p+1})$, the sample log likelihood is given by

$$\begin{aligned} \log L\left(\Pi, \Omega\right) &= -\frac{Tn}{2}\log 2\pi + \frac{T}{2}\log\left|\Omega^{-1}\right| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}\left[\left(y_t - \Pi' x_t\right)' \Omega^{-1}\left(y_t - \Pi' x_t\right)\right] \end{aligned} \tag{3.1}$$

2. Show that the term $\sum_{t=1}^{T}\left(y_t - \Pi' x_t\right)' \Omega^{-1}\left(y_t - \Pi' x_t\right)$ in the log likelihood you derived above can be rewritten as

$$\sum_{t=1}^{T}\left\{\left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right)' x_t\right]' \Omega^{-1}\left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right) x_t\right]\right\} \tag{3.2}$$

where $\hat{\Pi}$ is the OLS equation-by-equation estimate of Π and $\hat{\varepsilon}_t$ is the vector of OLS residuals.

Answer *Note that*

$$
\begin{aligned}
& \left(y_t - \Pi' x_t\right)' \Omega^{-1} \left(y_t - \Pi' x_t\right) \\
= & \left(y_t - \hat{\Pi}' x_t + \hat{\Pi}' x_t - \Pi' x_t\right)' \Omega^{-1} \left(y_t - \hat{\Pi}' x_t + \hat{\Pi}' x_t - \Pi' x_t\right) \\
= & \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right)' x_t\right]' \Omega^{-1} \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right)' x_t\right]
\end{aligned}
$$

since $\hat{\varepsilon}_t \equiv y_t - \hat{\Pi}' x_t$, implying that

$$
\begin{aligned}
& \sum_{t=1}^{T} \left(y_t - \Pi' x_t\right)' \Omega^{-1} \left(y_t - \Pi' x_t\right) \\
\equiv & \sum_{t=1}^{T} \left\{ \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right)' x_t\right]' \Omega^{-1} \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right) x_t\right] \right\}
\end{aligned}
$$

3. Show that (3.2) can be further simplified as

$$
\sum_{t=1}^{T} \hat{\varepsilon}_t' \Omega^{-1} \hat{\varepsilon}_t + \sum_{t=1}^{T} \left[x_t' \left(\hat{\Pi} - \Pi\right) \Omega^{-1} \left(\hat{\Pi} - \Pi\right)' x_t\right] \tag{3.3}
$$

[Hint: use the fact that this last expression is a scalar and recall that the OLS residuals are, by construction, orthogonal to the regressors]

Answer *Note that*

$$
\begin{aligned}
& \sum_{t=1}^{T} \left\{ \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right)' x_t\right]' \Omega^{-1} \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi\right) x_t\right] \right\} \\
= & \sum_{t=1}^{T} \hat{\varepsilon}_t' \Omega^{-1} \hat{\varepsilon}_t + 2 \sum_{t=1}^{T} \hat{\varepsilon}_t' \Omega^{-1} \left(\hat{\Pi} - \Pi\right)' x_t \\
& + \sum_{t=1}^{T} \left[x_t' \left(\hat{\Pi} - \Pi\right) \Omega^{-1} \left(\hat{\Pi} - \Pi\right)' x_t\right]
\end{aligned}
$$

Note that each of these elements is a scalar and that

$$
\begin{aligned}
\sum_{t=1}^{T} \hat{\varepsilon}_t' \Omega^{-1} \left(\hat{\Pi} - \Pi\right)' x_t &= trace \left[\sum_{t=1}^{T} \hat{\varepsilon}_t' \Omega^{-1} \left(\hat{\Pi} - \Pi\right)' x_t\right] \\
&= trace \left[\sum_{t=1}^{T} \Omega^{-1} \left(\hat{\Pi} - \Pi\right)' x_t \hat{\varepsilon}_t'\right] \\
&= trace \left[\Omega^{-1} \left(\hat{\Pi} - \Pi\right)' \sum_{t=1}^{T} x_t \hat{\varepsilon}_t'\right]
\end{aligned}
$$

But the OLS residuals are, by construction, orthogonal to the regressors, therefore $\sum_{t=0}^{T} x_t \hat{\varepsilon}_t'$ *is a matrix of zeros and the last expression is therefore identically zero implying that*

$$\begin{aligned}
& \sum_{t=1}^{T} \left\{ \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi \right)' x_t \right]' \Omega^{-1} \left[\hat{\varepsilon}_t + \left(\hat{\Pi} - \Pi \right) x_t \right] \right\} \\
= & \sum_{t=1}^{T} \hat{\varepsilon}_t' \Omega^{-1} \hat{\varepsilon}_t + \sum_{t=0}^{T} \left[x_t' \left(\hat{\Pi} - \Pi \right) \Omega^{-1} \left(\hat{\Pi} - \Pi \right)' x_t \right]
\end{aligned}$$

4. Use this last result to show that the OLS estimator of Π ($\hat{\Pi}$) is the MLE.

Answer *Define* $x_t^* = \left(\hat{\Pi} - \Pi \right)' x_t$. *The last term in (3.3) is therefore*

$$\sum_{t=1}^{T} \left[x_t' \left(\hat{\Pi} - \Pi \right) \Omega^{-1} \left(\hat{\Pi} - \Pi \right)' x_t \right] = \sum_{t=1}^{T} x_t^{*\prime} \Omega^{-1} x_t^*.$$

Note that since Ω *is a positive definite matrix,* Ω^{-1} *is as well. Thus, the last expression is positive for any sequence* $\{x_t^*\}_{t=0}^{T}$ *other than* $x_t^* = 0 \; \forall t$. *Thus the smallest value that (3.3) can take is achieved when* $\Pi = \hat{\Pi}$. *It than follows that the log likelihood (3.1) is maximized by setting* $\Pi = \hat{\Pi}$, *proving that OLS equation by equation is the MLE estimator of the unrestricted VAR.*