Aug 2017 Final Practice

① Vector autoregression and cointegration

Consider $\underset{\sim}{x}_t \sim VAR(p)$, $k=2$;

$$\underset{\sim}{x}_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \sum_i^p \begin{bmatrix} \phi_{11}^{(i)} & \phi_{12}^{(i)} \\ \phi_{21}^{(i)} & \phi_{22}^{(i)} \end{bmatrix} \begin{bmatrix} x_{1,t-i} \\ x_{2,t-i} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}.$$

1. State how to check stationarity of $\underset{\sim}{x}_t$.

→ Define $A_i = \begin{bmatrix} \phi_{11}^{(i)} & \phi_{12}^{(i)} \\ \phi_{21}^{(i)} & \phi_{22}^{(i)} \end{bmatrix}$. WLOG, assume i starts w/ 1. That is,

$$\underset{\sim}{x}_t = \sum_{i=1}^p A_i B^i \underset{\sim}{x}_t + \underset{\sim}{a}_t, \quad \underset{\sim}{a}_t = \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}.$$

$$\Rightarrow \left(I - \sum_{i=1}^p A_i B^i\right) \underset{\sim}{x}_t = \underset{\sim}{a}_t, \quad \underset{\sim}{a}_t = \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}.$$

If the eigenvalues of $I - \sum_{i=1}^p A_i B^i$ are all greater than 1 in absolute value, then $\underset{\sim}{x}_t$ is stationary.

Or, write the following:

$$\underset{\sim}{\xi}_t = \begin{bmatrix} \underset{\sim}{x}_t \\ \underset{\sim}{x}_{t-1} \\ \vdots \\ \underset{\sim}{x}_{t-(p-1)} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} \underset{\sim}{x}_{t-1} \\ \underset{\sim}{x}_{t-2} \\ \vdots \\ \underset{\sim}{x}_{t-p} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{a}_t \\ \vec{0} \\ \vdots \\ \vec{0} \end{bmatrix}$$

$$= \underset{\sim}{A} \xi_{t-1} + \underset{\sim}{u}_t.$$

If all the absolute value of eigenvalues of A are less than 1, then $\underset{\sim}{x}_t$ is stationary.

2. Describe the methods to select the order for $\underset{\sim}{x}_t$.

→ There are a few ways to choose p. We can look at AIC or BIC; AIC(p) is defined as

$$AIC(p) = \ln \det(\hat{\Sigma}_a(p)) + \frac{2}{T} p k^2,$$

$$\text{with } \hat{\Sigma}_a(p) = \frac{1}{T} \sum_{t=1}^T \hat{a}_t \hat{a}_t'.$$

Choose p w/ lowest AIC.

We can also conduct a sequential likelihood ratio test such as VAR(p-1) vs. VAR(p).

3. State how to test Granger causality for the case that $X_{1,t}$ granger-causes $X_{2,t}$ but not the other way around. Based on the same condition, express $X_{2,t}$ as a TFNM of $X_{1,t}$.

$\rightarrow$ WLOG, assume $i = 1, \cdots, p$.

So $\underset{\sim}{x}_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \sum_{i=1}^{p} \begin{bmatrix} \phi_{11}^{(i)} & \phi_{12}^{(i)} \\ \phi_{21}^{(i)} & \phi_{22}^{(i)} \end{bmatrix} \begin{bmatrix} x_{1,t-i} \\ x_{2,t-i} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix}$.

Thus $\begin{cases} x_{2,t} = \sum_{i=1}^{p} [\phi_{21}^{(i)} x_{1,t-i} + \phi_{22}^{(i)} x_{2,t-i}] + a_{2,t}, \\ x_{1,t} = \sum_{i=1}^{p} [\phi_{11}^{(i)} x_{1,t-i} + \phi_{12}^{(i)} x_{2,t-i}] + a_{1,t}. \end{cases}$

~~We can test $H_0 : \phi_{12}^{(i)} = 0 \ \forall i$ vs. $H_1 : \phi_{12}^{(i)} \neq 0 \ \exists i$. If we reject H_0, conclude that~~

To test whether $x_{1,t}$ granger-causes $x_{2,t}$, test $H_0 : \phi_{21}^{(i)} = 0 \ \forall i$ vs. $H_1 : \phi_{21}^{(i)} \neq 0 \ \exists i$.
If we reject H_0, conclude that $x_{1,t}$ g-causes $x_{2,t}$.
To test whether $x_{2,t}$ doesn't granger-cause $x_{1,t}$, test $H_0 : \phi_{12}^{(i)} = 0 \ \forall i$ vs. $H_1 : \phi_{12}^{(i)} \neq 0 \ \exists i$.
If we cannot reject H_0, conclude that $x_{2,t}$ doesn't granger-cause $x_{1,t}$.

Write $x_{2,t} = \sum_{i=1}^{p} \phi_{21}^{(i)} x_{1,t-i} + \sum_{i=1}^{p} \phi_{22}^{(i)} B^i x_{2,t} + a_{2,t}$

$\Rightarrow (1 - \sum_{i=1}^{p} \phi_{22}^{(i)} B^i) x_{2,t} = \sum_{i=1}^{p} \phi_{21}^{(i)} x_{1,t-i} + a_{2,t}$

$\Rightarrow x_{2,t} = \dfrac{\sum_{i=1}^{p} \phi_{21}^{(i)} B^i}{1 - \sum_{i=1}^{p} \phi_{22}^{(i)} B^i} x_{1,t} + \dfrac{a_{2,t}}{1 - \sum_{i=1}^{p} \phi_{22}^{(i)} B^i}$

$= V(B) x_{1,t} + e_t$. ✓

$\left(\text{w/} \begin{cases} x_{1,t} = \sum_{i=1}^{p} \phi_{11}^{(i)} x_{1,t-i} + a_{1,t} \\ (1 - \sum_{i=1}^{p} \phi_{22}^{(i)} B) e_t = a_{2,t} \end{cases} \right)$

4. Suppose that $\phi_{kl}^{(i)} = 0$ for $i = 2, \cdots, p$, $k=1$, $l=2$.
Derive the implied model for $x_{2,t}$.

$\hookrightarrow$ Since $x_{1,t} = \sum_{i=1}^{p} \phi_{11}^{(i)} B^i x_{1,t} + \sum_{i=1}^{p} \phi_{12}^{(i)} B^i x_{2,t} + a_{1,t}$,
we have $x_{1,t} = \sum_{i=1}^{p} \phi_{11}^{(i)} B^i x_{1,t} + \phi_{12}^{(1)} B x_{2,t} + a_{2,t}$.
So $(1 - \sum_{i=1}^{p} \phi_{11}^{(i)} B^i) x_{1,t} - a_{2,t} = \phi_{12}^{(1)} B x_{2,t}$

$$\Rightarrow x_{2,t} = \frac{1 - \sum_{i=1}^{p} \phi_{11}^{(i)} B^i}{\phi_{12}^{(1)} B} x_{1,t} - \frac{a_{2,t}}{\phi_{12}^{(1)} B}$$

$$= v(B) x_{1,t} + e_t. \quad \checkmark$$

5. Suppose $\phi_1(B) x_{1,t} = \theta_1(B) u_{1,t}$ and $\phi_2(B) x_{2,t} = \theta_2(B) u_{2,t}$,
where $\phi_k(B) = 1 - \phi_1^{(k)} B - \cdots - \phi_{p_k}^{(k)} B^{p_k}$,
and $\theta_k(B) = 1 + \theta_1^{(k)} B + \cdots + \theta_{q_k}^{(k)} B^{q_k}$, $k = 1, 2$.
Describe how to test Granger causality using univariate ~~app~~ approach.

$\rightarrow$ WLOG, suppose we are testing whether $x_{1,t}$ g-causes $x_{2,t}$.

~~Define $\rho_{12}(k)$~~

Define $\rho_{12}(k) = \dfrac{E(u_{1,t} u_{2,t+k})}{\sqrt{E(u_{1,t}^2) E(u_{2,t}^2)}}$.

Set H_0: $x_{1,t}$ does not Granger-cause $x_{2,t}$.

$$Q_L = n^2 \sum_{k=0}^{L} (n-k)^{-1} r_{12}^2(k) \sim \chi^2_{L+1}.$$

$Q_L \sim \chi^2_{L+1}$ $\because$ we assume $x_{1,t}$ and $x_{2,t}$ are stationary and invertible ARMA(p_1, q_1) and ARMA(p_2, q_2) processes respectively (so no unit roots exist).
Reject H_0 for large Q_L, i.e. conclude that $x_{1,t}$ Granger-causes $x_{2,t}$.

6. Suppose that $x_{1,t}$ and $x_{2,t}$ are not weakly stationary. How do you model the joint dynamics of $\{x_{1,t}, x_{2,t}\}$? Discuss your decisions based on whether these two series are cointegrated.

→ Define $\underset{\sim}{x}_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix}$. I want to know γ such that $\gamma' \underset{\sim}{x}_t \sim I(0)$.

We can use Engle and Granger procedure:

① Test whether $x_{1,t}$ and $x_{2,t}$ are $I(1)$ using DF test, or ADF test, etc.

② If they are, regress one series against the other using least squares:

$$x_{2,t} = \beta_0 + \beta_1 x_{1,t} + e_t.$$

③ Run a URT on $\{\hat{e}_t\}$. If the residuals are stationary, these two series are cointegrated.

④ Write in an ECM representation:

$$\begin{bmatrix} -\alpha\beta_1 & \rho_1 \\ -\alpha\beta_2 & \rho_2 \end{bmatrix}\begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix}$$

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix}[-\hat{\alpha} \;\; 1]$$

$$\nabla \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \hat{\phi}_1 \\ -\hat{\phi}_2 \end{bmatrix} \begin{bmatrix} \;\; \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} \beta_{x,i} & \beta_{y,i} \\ \gamma_{x,i} & \gamma_{y,i} \end{bmatrix} \begin{bmatrix} \nabla x_{1,t-i} \\ \nabla x_{2,t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

$$\nabla \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} [-\hat{\beta}_0 \;\; 1] \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} \beta_{x_1,i} & \beta_{x_2,i} \\ \gamma_{x_1,i} & \gamma_{x_2,i} \end{bmatrix} \begin{bmatrix} \nabla x_{1,t-i} \\ \nabla x_{2,t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

$$\nabla \underset{\sim}{x}_t = \vec{c} + \rho \beta' \underset{\sim}{x}_{t-1} + \sum_{i=1}^{m} \Phi_i \nabla \underset{\sim}{x}_{t-i} + \underset{\sim}{\varepsilon}_t.$$

We can also use Johansen procedure:

Consider a naive case: $\nabla \underset{\sim}{x}_t = \alpha\beta' \underset{\sim}{x}_{t-1} + \underset{\sim}{a}_t$.

From the assumption, $\underset{\sim}{x}_t \sim VAR(p)$.

That is, $\underset{\sim}{x}_t = \sum_{i=1}^{p} A_i B' \underset{\sim}{x}_{t-i} + \underset{\sim}{a}_t$.

Define $\Gamma_i = -(I - A_1 - A_2 - \cdots - A_i)$, $i = 1, \cdots, p-1$,

and $\Pi := \Gamma_p = -(I - A_1 - \cdots - A_p)$.

Then we have $\nabla \underset{\sim}{x}_t = \sum_{i=1}^{p-1} \Gamma_i \nabla \underset{\sim}{x}_{t-i} + \Pi \underset{\sim}{x}_{t-p} + \underset{\sim}{a}_t$.

If $\text{rank}(\Pi) = r \in (0, k)$ ($0 \Rightarrow \Pi = 0$; $k \Rightarrow \underset{\sim}{x}_t$ is stationary),
then it implies $\exists\, \alpha_{k \times r}$ and $\beta_{k \times r}$ such that
$\Pi = \alpha\beta'$, i.e. $\exists\, \beta'$ s.t. $\underset{\sim}{w}_t = \beta' \underset{\sim}{x}_t$ is stationary.

Estimate Π using two regressions:

$$\nabla \underset{\sim}{x}_t = \sum_{i=1}^{p-1} \Psi_i \nabla \underset{\sim}{x}_{t-i} + u_t,$$

$$\underset{\sim}{x}_{t-1} = \sum_{i=1}^{p-1} \Psi_i^* \nabla \underset{\sim}{x}_{t-i} + V_t.$$

Let $\hat{u}_t$ and $\hat{v}_t$ denote least square residuals.
Now we have $\hat{\underset{\sim}{u}}_t = \Pi \hat{\underset{\sim}{v}}_t + \underset{\sim}{E}_t$.

Define
$$\begin{cases} \hat{\Sigma}_{00} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t \hat{u}_t', & \hat{\Sigma}_{01} = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t \hat{v}_t', \\ \hat{\Sigma}_{10} = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{u}_t', & \hat{\Sigma}_{11} = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{v}_t'. \end{cases}$$

The sample matrix $\hat{\Pi}$ is $\hat{\Sigma}_{11}^{-1} \hat{\Sigma}_{10} \hat{\Sigma}_{00}^{-1} \hat{\Sigma}_{01}$.

In order to determine r, we first define
$\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_k \geq 0$ to be ordered eigenvalues of $\hat{\Pi}$.
The tests are:

① trace test : $H_0: r = r_0$ vs. $H_1: r > r_0$.
$\lambda_{trace}(r_0) = -T \sum_{i=r_0+1}^{k} \ln(1 - \hat{\lambda}_i)$. Reject H_0 if $\lambda_{tr}(r_0)$ large.

② maximum eigenvalue test : $H_0: r = r_0$ vs. $H_1: r = r_0 + 1$.
$\lambda_{max}(r_0) = -T \cdot \ln(1 - \hat{\lambda}_{r_0+1})$.

7. Discuss the reasons why we have to choose different models based on the condition of cointegration.

→ EG focuses on finding linear combination with minimum variance, whereas Johansen seeks linear combination which is more stationary. cut

Some implicit assumptions in the procedure makes the model only applicable in special cases. For example, EG procedure is only applicable to systems with more than two variables in a very special circumstances. Also, the focuses of each model are different, so it is natural to choose different models in different cases (that is, there is no single best procedure). insert & paste

From the risk management P.O.V., EG criterion may seem more important than J's criterion. Lastly, the presence of change points will affect the effectiveness of cointegration analysis.

8. Discuss the EG approach for modelling cointegrated $x_{1,t}$ and $x_{2,t}$.

→ ① Test whether $x_{1,t} \sim I(1)$ and $x_{2,t} \sim I(1)$ using unit root test.

② If both are $I(1)$, regress one to another using least squares: $x_{2,t} = \beta_0 + \beta_1 x_{1,t} + e_t$.

③ Run an unit root test on $\{\hat{e}_t\}$. If $\hat{e}_t \sim I(0)$, then $x_{1,t}$ and $x_{2,t}$ are cointegrated.

④ Write the following ECM:

$x_{2,t} - \hat{\beta}_1 x_{1,t} = \hat{e}_t$.

$$\nabla \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \begin{bmatrix} -\hat{\beta}_1 & 1 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} \beta_{x_1,i} & \beta_{x_2,i} \\ \gamma_{x_1,i} & \gamma_{x_2,i} \end{bmatrix} \nabla \begin{bmatrix} x_{1,t-i} \\ x_{2,t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

9. Discuss the implication of Granger representation theorem.
→ The theorem says that if x_t and y_t are cointegrated, then there exists an ECM representation.
VARs on differenced $I(1)$ processes will be a mis-specification if the component series are cointegrated.
Engle and Granger showed that an equilibrium specification is missing from a VAR representation.
However, when lagged disequilibrium terms are included as explanatory variables, the model becomes well specified.
The ECM is structured so that short-run deviation from the long-run equilibrium will be corrected.

Bootstrap time series

Consider an AR(2): $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t$, $a_t \overset{iid}{\sim} N(0,1)$.

1. Describe the steps of (unconditional) parametric bootstrap for the above AR(2).

→ Multiply both sides by y_{t-k}:

$$y_t y_{t-k} = \mu y_{t-k} + \phi_1 y_{t-1} y_{t-k} + \phi_2 y_{t-2} y_{t-k} + a_t y_{t-k}.$$

$k=0$ and $E(\cdot) \Rightarrow \gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + 1$

$k=1$ and $E(\cdot) \Rightarrow \gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 \Rightarrow \gamma_1 = \frac{\phi_1}{1-\phi_2}\gamma_0$.

$k=2$ and $E(\cdot) \Rightarrow \gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0$

So $\gamma_2 = \left(\phi_1 + \frac{\phi_2(1-\phi_2)}{\phi_1}\right)\gamma_1 = \left(\frac{\phi_1^2 - \phi_2^2 + \phi_2}{\phi_1}\right)\gamma_1$.

$$\Rightarrow \gamma_0 = \frac{\phi_1^2}{1-\phi_2}\gamma_0 + \phi_2\left(\frac{\phi_1^2 - \phi_2^2 + \phi_2}{\phi_1}\right)\gamma_1 + 1$$

$$= \frac{\phi_1^2}{1-\phi_2}\gamma_0 + \phi_2\left(\frac{\phi_1^2 - \phi_2^2 + \phi_2}{\cancel{\phi_1}}\right)\cdot\frac{\cancel{\phi_1}}{1-\phi_2}\gamma_0 + 1$$

$$= \gamma_0\left[\frac{\phi_1^2 + \phi_2\phi_1^2 - \phi_2^3 + \phi_2^2}{1-\phi_2}\right] + 1.$$

So we can find γ_0, which is $Var(y_t)$.

~~And $(1-\phi_1 B - \phi_2 B^2) y_t = \mu + a_t \sim N(\ldots)$~~

$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t$, $a_t \sim N(0,1)$.

$\Phi_2(B) y_t = e_t$, $e_t \sim N(\mu, 1)$.

Also, letting $m = E(y_t)$ gives

$m = \mu + \phi_1 m + \phi_2 m$

$\Rightarrow (1-\phi_1-\phi_2)m = \mu$

$\Rightarrow m = \frac{\mu}{1-\phi_1-\phi_2}$.

That is, (unconditional) $y_t \sim N\left(\frac{\mu}{1-\phi_1-\phi_2}, \gamma_0\right)$.

① Simulate y_0 by drawing a random number from above.

② Simulate y_1 by

③ Simulate $y_2 = \mu + \phi_1 y_1 + \phi_2 y_0 + a_2$.

④ Simulate $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t$, $t \geq 2$ recursively.

2. Describe the steps of carrying out the Sieve bootstrap for the above AR(2).

→ ~~(1) Assume $u_t \sim AR(p)$, and p wasn't specified, we could do this. Under $y_t = X_t \beta + u_t$.~~
~~(1) Estimate the model to obtain $\hat{u}_t$.~~
~~(2) Estimate~~

$y_t \sim AR(2)$, so

~~Suppose~~ The general procedure is the following:

$y_t = X_t \beta + u_t$. Assume $u_t \sim AR(p)$.

Choose p by AIC or sequential LRT.

① Estimate the model to obtain residuals, $\hat{u}_t$.

② Estimate AR(p): $\hat{u}_t = \sum_{i=1}^{p} \phi_i \hat{u}_{t-i} + \varepsilon_t$. (1)

③ Generate bootstrap error terms:

$u_t^* = \sum_{i=1}^{p} \hat{\phi}_i u_{t-i}^* + \varepsilon_t^*$, where ε_t^* is resampled from the (rescaled) residuals from (1).

$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$

④ Generate the bootstrap data according to

$y_t^* = X_t \hat{\beta} + u_t^*$.

Write $\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \end{bmatrix} = \begin{bmatrix} 1 & y_{t-1} & y_{t-2} \\ 1 & y_{t-2} & y_{t-3} \\ 1 & y_{t-3} & y_{t-4} \end{bmatrix} \begin{bmatrix} \mu \\ \phi_1 \\ \phi_2 \end{bmatrix} + \begin{bmatrix} a_t \\ a_{t-1} \\ a_{t-2} \end{bmatrix}$.

Assume $a_t \sim AR(p)$. ✓

$(1-\phi_1 B-\phi_2 B^2)y_t=\mu+a_t=e_t \Longrightarrow y_t=\sum_{j=0}^{\infty}\psi_j e_{t-j} \sim AR(2).$

$\cancel{y_t=\mu\sum_{j=0}^{\infty}\psi_j+a_t\sum_{j=0}^{\infty}\psi_j.}$

$\cancel{y_t=\mu+\phi_1 y_{t-1}+\phi_2 y_{t-2}+a_t.}$ 103

3. Describe the steps of carrying out the block bootstrap method for the above AR(2).

→ Suppose $t=1,\cdots,T$. Then we have:

$y_t=\{\,\boxed{y_1}\boxed{y_2}\boxed{y_3}\cdots\boxed{y_{T-1}}\boxed{y_T}\,\}.$

Define $\underset{\sim}{Z}_t=[y_t, y_{t-1}, y_{t-2}]$.

Construct length l overlapping blocks:

Block 1 : $(\underset{\sim}{Z}_1,\cdots,\underset{\sim}{Z}_l)$

Block 2 : $(\underset{\sim}{Z}_2,\cdots,\underset{\sim}{Z}_{l+1})$

$\vdots \quad \vdots$

Block S: $(\underset{\sim}{Z}_S,\cdots,\underset{\sim}{Z}_{l+S-1})$

$\vdots \quad \vdots$

Block b: $(\underset{\sim}{Z}_b,\cdots,\underset{\sim}{Z}_T)$.

So. $T=l+b-1 \Rightarrow b=T-l+1$.

Resample the blocks with replacement.

Or, with length l, define:

Block 1 : $(y_1, y_2,\cdots, y_l)$

Block 2 : $(y_2,\cdots, y_{l+1})$

$\vdots$

Block b : $(y_b,\cdots, y_T)$ $(T=l+b-1 \Rightarrow b=T-l+1)$.

4. Describe the pros and cons of the above models.

→ For the parametric bootstrap:

Pros : · Straightforward

Cons : · possible risk of misspecification of a_t's distribution.

· assuming independent errors

· computational difficulty if p large

→ For the sieve bootstrap:

Pros : · Straightforward

Cons : · assuming iid innovations of errors, thereby ruling out other forms of heteroscedasticity

*

→ For the block bootstrap: more reliable s.e.

Pros: · dependency well preserved; works w/ hetero. as well as serial corr.

Cons: · pseudo time series generated is not stationary although the original series is.

· the mean $\bar{x}_n^*$ is biased

· the estimator of the variance of the $\sqrt{n}\cdot\bar{x}_n$ is also biased.

· higher-order ~~accuracy~~ accuracy than asymptotic methods only by a modest extent.

SSM

Express a given TSM as a SSM.

1. $y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$, $a_t \overset{iid}{\sim} N(0,1)$

$y_t = F' \underset{\sim}{x}_t + \epsilon_t$ 2. $y_t = \alpha + \sum_{i=1}^{p} \beta f_{i,t} + a_t$, $a_t \overset{iid}{\sim} N(0,1)$.

$\underset{\sim}{x}_t = G \underset{\sim}{x}_{t-1} + \underset{\sim}{w}_t$. 1. $\Phi_2(B) y_t = \Theta_2(B) a_t$

$$\rightarrow y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + a_t$$

$$= \begin{bmatrix} \phi_1 & \phi_2 & \theta_1 & \theta_2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ a_{t-1} \\ a_{t-2} \end{bmatrix} + a_t .$$

$$\begin{bmatrix} y_{t-1} \\ y_{t-2} \\ a_{t-1} \\ a_{t-2} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \theta_1 & \theta_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-2} \\ y_{t-3} \\ a_{t-2} \\ a_{t-3} \end{bmatrix} + \begin{bmatrix} a_{t-1} \\ 0 \\ a_{t-1} \\ 0 \end{bmatrix}$$

$$\text{Or, } y_t = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \\ \theta_1 a_t \\ \theta_2 a_{t-1} \end{bmatrix} + 0 ,$$

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \theta_1 a_t \\ \theta_2 a_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \theta_1 a_{t-1} \\ \theta_2 a_{t-2} \end{bmatrix} + \begin{bmatrix} a_t \\ 0 \\ \theta_1 a_t \\ \theta_2 a_{t-1} \end{bmatrix}$$

2. $y_t = \alpha + \sum_{i=1}^{p} \beta f_{i,t} + a_t$

$$= \alpha + \beta f_{1,t} + \beta f_{2,t} + \cdots + \beta f_{p,t} + a_t.$$

$$= \begin{bmatrix} 1 & f_{1,t} & \cdots & f_{p,t} \end{bmatrix}_{1\times(p+1)} \begin{bmatrix} \alpha \\ \beta \\ \vdots \\ \beta \end{bmatrix}_{(p+1)\times 1} + a_t$$

$$\begin{bmatrix} \alpha \\ \beta \\ \vdots \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \vdots \\ \beta \end{bmatrix} + \vec{0}.$$

APR 2018 Final Practice

2. Define Granger causality in terms of VAR process.

$\rightarrow$ Consider $\underset{\sim}{z}_t = \begin{bmatrix} \underset{\sim}{x}_t \\ \underset{\sim}{y}_t \end{bmatrix}_{k\times 1}$, with $\underset{\sim}{x}_t$ a $k_1 \times 1$ vector

and $\underset{\sim}{y}_t$ a $k_2 \times 1$ vector, $k = k_1 + k_2$.

Suppose $\underset{\sim}{z}_t \sim VAR(p)$. Write $\underset{\sim}{z}_t = \sum_{i=1}^{p} A_i \underset{\sim}{z}_{t-i} + \underset{\sim}{E}_i$.

Now write $\begin{bmatrix} \underset{\sim}{x}_t \\ \underset{\sim}{y}_t \end{bmatrix} = \sum_{i=1}^{p} \begin{bmatrix} \phi_{11}^{(i)} & \phi_{12}^{(i)} \\ \phi_{21}^{(i)} & \phi_{22}^{(i)} \end{bmatrix} \begin{bmatrix} \underset{\sim}{x}_{t-i} \\ \underset{\sim}{y}_{t-i} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{e}_{x,t} \\ \underset{\sim}{e}_{y,t} \end{bmatrix}$,

where $\phi_{11}^{(i)}$ is a $k_1 \times k_1$ matrix, $\phi_{12}^{(i)}$ a $k_1 \times k_2$, $\phi_{21}^{(i)}$ a $k_2 \times k_1$, and $\phi_{22}^{(i)}$ a $k_2 \times k_2$.

If $\phi_{12}^{(i)} = \underset{\sim}{0}\ \forall i$, say $\underset{\sim}{y}_t$ does not Granger-cause $\underset{\sim}{x}_t$;
if $\phi_{21}^{(i)} = \underset{\sim}{0}\ \forall i$, say $\underset{\sim}{x}_t$ does not Granger-cause $\underset{\sim}{y}_t$.

3. Test Granger causality using VAR or univariate approach.

→ ① Univariate approach

Suppose $x_t \sim ARMA(p_x, q_x)$ and $y_t \sim ARMA(p_y, q_y)$, i.e. $\Phi_{p_x}(B)x_t = \Theta_{q_x}(B)a_t$ and $\Phi_{p_y}(B)y_t = \Theta_{q_y}(B)e_t$.

~~Define~~ $\cancel{\rho_{xy}(k) = \frac{E(x_t y_{t+k})}{\sqrt{E(x_t^2)E(y_t^2)}}}$

Define $\rho_{ae}(k) = \dfrac{E(a_t e_{t+k})}{\sqrt{E(a_t^2)E(e_t^2)}}$.

H_0: x_t does not Granger-cause y_t.

Then let $Q_L = n^2 \sum_{k=0}^{L} (n-k)^{-1} \gamma_{ae}^2(k) \sim \chi^2_{L+1}$.

Reject H_0 for large Q_L.

② VAR

As in the previous question;

H_0: x_t does not Granger-cause y_t

$\phi_{21}^{(i)} = 0 \ \forall i$

H_0: y_t does not Granger-cause x_t

$\phi_{12}^{(i)} = 0 \ \forall i$

4. State the approaches for cointegration modelling.

→ ① Engle - Granger and ECM.

1) Test whether x_t and y_t are $I(1)$ using unit root test.

2) If they are $I(1)$, regress one series against the other : $y_t = \beta_0 + \beta_1 x_t + e_t$.

3) Run an unit root test on $\{\hat{e}_t\}$. If $\hat{e}_t \sim I(0)$, conclude that x_t and y_t are cointegrated.

4) Write in an ECM representation:

$$\nabla \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \begin{bmatrix} -\beta_1 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} \beta_{x,i} & \beta_{y,i} \\ \gamma_{x,i} & \gamma_{y,i} \end{bmatrix} \begin{bmatrix} \nabla x_{t-i} \\ \nabla y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

② Johansen's procedure:

~~1) Write two regressions:~~ $\cancel{\nabla \underset{\sim}{x}_t = \sum_{i=1}^{p-1} \Psi_i \nabla \underset{\sim}{x}_{t-i} + u_t}$, $\cancel{\underset{\sim}{x}_{t-1} = \sum_{i=1}^{p-1} \Psi_i^* \nabla \underset{\sim}{x}_{t-i} + v_t}$.

~~2) Let u_t and v_t denote regression residuals.~~

~~Let $\hat{u}_t = \Pi \hat{v}_t + E_t$.~~

1) Consider $\nabla \underset{\sim}{x}_t = \alpha\beta' \underset{\sim}{x}_{t-1} + \underset{\sim}{a}_t$, $\underset{\sim}{x}_t \sim VAR(p)$.

$\underset{\sim}{x}_t = \sum_{i=1}^{p} A_i \underset{\sim}{x}_{t-i} + \underset{\sim}{a}_t$

2) Define
$$\begin{cases} \Gamma_i = -(I - A_1 - \cdots - A_i), & i = 1, \cdots, p-1; \\ \Gamma_p = \Pi = -(I - A_1 - \cdots - A_p) \end{cases}$$

3) Rewrite the following:
$$\nabla \underset{\sim}{x}_t = \sum_{i=1}^{p-1} \Gamma_i \nabla \underset{\sim}{x}_{t-i} + \Pi \underset{\sim}{x}_{t-p} + \underset{\sim}{a}_t .$$

If $\operatorname{rank}(\Pi) = 0$, then $\Pi = 0$.

If $\operatorname{rank}(\Pi) = k$, then $\underset{\sim}{x}_t$ is stationary.

If $\operatorname{rank}(\Pi) = r \in (0, k)$, then $\exists\ \alpha_{k\times r}$ and $\beta_{k\times r}$ such that $\Pi = \alpha\beta'$, i.e. $\underset{\sim}{w}_t = \beta' \underset{\sim}{x}_t$ is stationary.

4) If $\operatorname{rank}(\Pi) = r$, write two regressions:
$$\begin{cases} \nabla \underset{\sim}{x}_t = \sum_{i=1}^{p-1} \Psi_i \nabla \underset{\sim}{x}_{t-i} + u_t, \\ \underset{\sim}{x}_{t-1} = \sum_{i=1}^{p-1} \Psi_i^* \nabla \underset{\sim}{x}_{t-i} + v_t \end{cases}$$
and get regression residuals $\hat{u}_t$ and $\hat{v}_t$.

5) Write $\hat{u}_t = \Pi \hat{v}_t + E_t$.

Define
$$\begin{cases} \hat{\Sigma}_{00} = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t \hat{u}_t', & \hat{\Sigma}_{01} = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t \hat{v}_t', \\ \hat{\Sigma}_{10} = \frac{1}{T} \sum_{t=1}^{T} \hat{v}_t \hat{u}_t', & \hat{\Sigma}_{11} = \frac{1}{T} \sum_{t=1}^{T} \hat{v}_t \hat{v}_t' \end{cases}$$

6) Let $\hat{\Pi} = \hat{\Sigma}_{11}^{-1} \hat{\Sigma}_{10} \hat{\Sigma}_{00}^{-1} \hat{\Sigma}_{01}$.

7) To determine r, conduct the following tests:

① Trace test: $H_0: r_0 = r$ vs $H_1: r > r_0$

$\lambda_{trace}(r_0) = -T \sum_{i=r_0+1}^{k} \ln(1 - \hat{\lambda}_i)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ are ordered, and $\hat{\lambda}_i$'s are eigenvalues of $\hat{\Pi}$.

Reject H_0 for large $\lambda_{tr}(r_0)$.

② Maximum eigenvalue test: $H_0: r_0 = r$ vs. $H_1: r = r_0 + 1$.

$\lambda_{max}(r_0) = -T \ln(1 - \hat{\lambda}_{r_0+1})$.

Reject H_0 for large $\lambda_{max}(r_0)$.

5. State EG's representation theorem and its implication for modelling multivariate time series.
→ The theorem says if x_t and y_t are cointegrated, then there exists an ECM representation.
VAR on differenced I(1) process will be a misspecification if the components series are cointegrated.
Engle and Granger showed that an equilibrium specification is missing from a VAR representation.
However, when lagged disequilibrium terms are included as explanatory variables, the model becomes well specified.
The ECM is structured so that short-run deviation from the long-run equilibrium will be corrected.

6. State time series bootstrapping methods. In particular, ~~for depe~~ state for dependent time series and dynamic regression models taught in class.
→ Bootstrapping is a procedure of sampling with replacement from data and then computing estimates of parameters from these samples.
For dependent data, there are mainly three ways of bootstrapping procedures. One is a parametric bootstrapping. Consider $y_t = \phi_1 y_{t-1} + a_t + \alpha$, $a_t \sim N(0, \sigma_a^2)$.
Then we can compute the unconditional distribution of y_t, which is $N\left(\frac{\alpha}{1-\phi_1}, \frac{\sigma_a^2}{1-\phi_1^2}\right)$. And we have $y_{t-1} \sim N\left(\frac{\alpha}{1-\phi_1}, \frac{\sigma_a^2}{1-\phi_1^2}\right)$, with ρ_1 being a correlation.
Also, $y_t | y_{t-1} \sim N(\phi_1 y_{t-1}, \sigma_a^2)$.

(If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, then
$X_1 | X_2 = x_2 \sim N\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(x_2 - \mu_2), \sigma_1^2(1-\rho^2)\right)$.)

Simulate $y_0 \sim N\left(\frac{\alpha}{1-\phi_1}, \frac{\sigma_a^2}{1-\phi_1^2}\right)$, and $y_1 = \phi_1 y_0 + \alpha + a_t$.
$y_t = \phi_1 y_{t-1} + a_t + \alpha$, $t \geq 1$, and repeat.

The next is the Sieve bootstrapping. Consider $y_t = \underset{\sim}{X_t}\underset{\sim}{\beta} + u_t$.
Assume u_t follows an unknown, stationary process with homoscedastic innovations. Approximate u_t with AR(p) where p is selected by AIC or by sequential testing.
Now estimate the model to obtain $\hat{u}_t$.
Estimate AR(p): $\hat{u}_t = \sum_{i=1}^{p} \phi_i \hat{u}_{t-i} + \varepsilon_t$. Choose p by testing or criterion, and ensure its stationarity.
Generate bootstrap error terms: $u_t^* = \sum_{i=1}^{p} \hat{\phi}_i u_{t-i}^* + \varepsilon_t^*$, where ε_t^*'s are resampled from the (rescaled) residuals from $\hat{u}_t = \sum_{i=1}^{p} \phi_i \hat{u}_{t-i} + \varepsilon_t$.
Finally, generate bootstrap data according to
$y_t^* = \underset{\sim}{X_t}\hat{\underset{\sim}{\beta}} + u_t^*$.

The last is the block bootstrapping. The idea is to divide the data of n observations into blocks of length l and select b of these blocks by resampling with replacement. Overlapping blocks and fixed l is preferred. Often $l \propto n^{1/3}$. Block-of-blocks bootstrap is the analogue of the pairs bootstrap for dynamic models.
Consider $\underset{\sim}{z_t} = [y_t \ y_{t-1} \ \underset{\sim}{X_t}]$, $y_t = \underset{\sim}{X_t}\underset{\sim}{\beta} + \gamma y_{t-1} + u_t$.

$$
\begin{aligned}
\text{Define: } \text{Block}_1 &= (\underset{\sim}{z_1}, \cdots, \underset{\sim}{z_l}) \\
\text{Block}_2 &= (\underset{\sim}{z_2}, \cdots, \underset{\sim}{z_{l+1}}) \\
&\ \vdots \\
\text{Block}_k &= (\underset{\sim}{z_k}, \cdots, \underset{\sim}{z_{l+k-1}}) \\
&\ \vdots \\
\text{Block}_b &= (\underset{\sim}{z_b}, \cdots, \underset{\sim}{z_n}).
\end{aligned}
$$

So $n = l + b - 1 \Rightarrow b = n - l + 1$ blocks are defined in total.
This works w/ non-constant variance as well as serial correlation. Generally, blocks bootstrap is more accurate than asymptotic methods but only by a modest extent.
This bootstrap yields more reliable s.e.s.

7. In finance, ~~the~~ the appraisal returns y_t on private assets may be modeled as $y_t = \sum_{i=0}^{m} w_i r_{t-i}$, $w_i \ge 0$, $\sum w_i = 1$, where r_t denotes ~~the~~ unobservable economic returns on private assets.

1) Geltner suggests to estimate w_0 using $y_t = (1-w_0)y_{t-1} + w_0 r_t$. Express w_i in terms of w_0.

$\cancel{y_t = w_0 r_t + \sum_{i=1}^{m} w_i r_{t-i}}$

$\cancel{= w_0 r_t + (1-w_0) y_{t-1}}$

$\cancel{= w_0 r_t + (1-w_0)\sum_{i=0}^{m} w_i r_{t-1-i}}$

$\cancel{= w_0 r_t + (1-w_0)[w_0 r_{t-1} + \sum_{i=1}^{m} w_i r_{t-1-i}]}$

$\cancel{= w_0 r_t + w_0(1-w_0) r_{t-1} + (1-w_0)\sum_{i=1}^{m} w_i r_{t-1-i}}$

~~Since~~ $\cancel{y_t = w_0 r_t + \sum_{i=1}^{m} w_i r_{t-i}}$, ~~we have~~

$\cancel{y_{t-1} - w_0 r_{t-1} = \sum_{i=1}^{m} w_i r_{t-1-i}}$.

~~Substituting this into equation, and we get~~

$\cancel{y_t = w_0 r_t + w_0(1-w_0) r_{t-1} + (1-w_0)(y_{t-1} - w_0 r_{t-1})}$

$\cancel{= w_0 r_t + w_0(1-w_0) r_{t-1} + (1-w_0) y_{t-1}}$

~~Estimate~~ $\cancel{w_1 = w_0(1-w_0)}$. ~~Now,~~

$\cancel{y_t = w_0 r_t + w_0(1-w_0) r_{t-1} + (1-w_0)[w_1 r_{t-2} + \cdots + w_m r_{t-1-m}]}$

$\cancel{= w_0 r_t + w_0(1-w_0) r_{t-1} + w_0(1-w_0)^2 r_{t-2} + (1-w_0)[w_2 r_{t-3} + \cdots + w_m r_{t-1-m}]}$.

~~Inductively,~~ $\cancel{w_i = w_0(1-w_0)^i, \; i = 1, \dots, m}$ ~~[illegible]~~.

$\rightarrow \; y_t = w_0 r_t + \sum_{i=1}^{m} w_i r_{t-i}$

$= w_0 r_t + (1-w_0) y_{t-1}$

$= w_0 r_t + (1-w_0)[w_0 r_{t-1} + \sum_{i=1}^{m} w_i r_{t-1-i}]$

$= w_0 r_t + (1-w_0) w_0 r_{t-1} + (1-w_0)(1-w_0) y_{t-2}$.

Inductively, $\underline{w_i = w_0(1-w_0)^i, \; i = 0, \dots, \infty}$. //

$\sum_{i=0}^{\infty} w_i = \sum_{i=0}^{\infty} w_0(1-w_0)^i = w_0 \cdot \frac{1}{1-(1-w_0)} = w_0 \cdot \frac{1}{w_0} = 1$ ✓

2) GLM suggests estimating w_i, $i=1,\cdots,m$ by fitting an MA(m) : $y_t=\sum_{i=0}^{m}\theta_i a_{t-i}$, $\theta_0=1$, $\theta_i\geq 0$. Express w_i, $i=0,\cdots,m$ and r_t in terms of $\{\theta_i\}_{i=0}^{m}$ and $\{a_t\}$ in equation.

$\rightarrow$ $y_t=\theta_0 a_t+\theta_1 a_{t-1}+\cdots+\theta_m a_{t-m}$

$$=\frac{\theta_0}{\Sigma\theta_i}\Sigma\theta_i a_t+\frac{\theta_1}{\Sigma\theta_i}\Sigma\theta_i a_{t-1}+\cdots+\frac{\theta_m}{\Sigma\theta_i}\Sigma\theta_i a_{t-m}.$$

That is, $w_i=\dfrac{\theta_i}{\sum_{j=0}^{m}\theta_j}$ $\forall i=0,\cdots,m$,

and $r_t=\sum_{j=0}^{m}\theta_j a_t$ $\forall t$.

3) Suppose that $r_t=\alpha+\beta f_t+e_t$. Substitute this in above. Express y_t using a DLM with single input f_t.

$\rightarrow$ $y_t=\sum_{i=0}^{m}w_i r_{t-i}$

$$=\sum_{i=0}^{m}w_i(\alpha+\beta f_{t-i}+e_{t-i})$$

$$=\alpha+\sum_{i=0}^{m}w_i\beta f_{t-i}+\sum_{i=0}^{m}w_i e_{t-i}$$

$$=\alpha+\sum_{i=0}^{m}\beta w_i f_{t-i}+\sum_{i=0}^{m}w_i e_{t-i}$$

$$=\alpha+\sum_{i=0}^{m}\beta w_i f_{t-i}+\xi_t,\quad \xi_t=\sum_{i=0}^{m}w_i e_{t-i}$$

Obs. equation

$$=\begin{bmatrix}\alpha & \beta w_0 & \beta w_1 & \cdots & \beta w_m\end{bmatrix}\begin{bmatrix}1\\ f_t\\ f_{t-1}\\ \vdots\\ f_{t-m}\end{bmatrix}+\begin{bmatrix}w_0 & \cdots & w_m\end{bmatrix}\begin{bmatrix}e_t\\ \vdots\\ e_{t-m}\end{bmatrix}.$$

State eqn.

$$\begin{bmatrix}e_t\\ e_{t-1}\\ \vdots\\ e_{t-m}\end{bmatrix}=\begin{bmatrix}0 & 0 & \cdots & 0 & 0\\ 1 & 0 & \cdots & 0 & 0\\ 0 & 1 & \cdots & 0 & 0\\ \vdots & & \ddots & & \vdots\\ 0 & 0 & \cdots & 1 & 0\end{bmatrix}\begin{bmatrix}e_{t-1}\\ e_{t-2}\\ \vdots\\ e_{t-(m+1)}\end{bmatrix}+\begin{bmatrix}e_t\\ 0\\ \vdots\\ 0\end{bmatrix}$$

$$y_t = \alpha + \sum_{i=0}^{m} \beta w_i f_{t-i} + \xi_t, \quad \xi_t = \sum_{i=0}^{m} w_i e_{t-i}$$

$$= \alpha + \beta \sum_{i=0}^{m} w_i B^i f_t + \sum_{i=0}^{m} w_i B^i e_t$$

$$= V(B) f_t + U_t, \quad V(B) = \beta \sum_{i=0}^{m} w_i B^i, \quad U_t = \alpha + \xi_t.$$

1. Exercises in MTS-R.

pg.10 1) Discuss how to test the stationarity of a VAR of order n using the idea of the companion matrix.

→ Suppose $\underset{\sim}{x}_t \sim VAR(n)$. That is, $\underset{\sim}{x}_t = \sum_{i=1}^{n} A_i \underset{\sim}{x}_{t-i} + \underset{\sim}{e}_t$.

Define $\underset{\sim}{Z}_t = \begin{bmatrix} \underset{\sim}{x}_t \\ \vdots \\ \underset{\sim}{x}_{t-(n-1)} \end{bmatrix}$ and $\underset{\sim}{E}_t = \begin{bmatrix} \underset{\sim}{e}_t \\ \vec{0} \\ \vdots \\ \vec{0} \end{bmatrix}$.

So $$\underset{\sim}{Z}_t = \begin{bmatrix} \underset{\sim}{x}_t \\ \vdots \\ \underset{\sim}{x}_{t-(n-1)} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{n-1} & A_n \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \begin{bmatrix} \underset{\sim}{x}_{t-1} \\ \vdots \\ \underset{\sim}{x}_{t-n} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{e}_t \\ \vec{0} \\ \vdots \\ \vec{0} \end{bmatrix}.$$

So $\underset{\sim}{Z}_t = \underset{\sim}{A} \underset{\sim}{Z}_{t-1} + \underset{\sim}{E}_t$, i.e. $\underset{\sim}{Z}_t \sim VAR(1)$.
If all mods of eigenvalues of $\underset{\sim}{A}$ are less than 1, then we say $\underset{\sim}{x}_t$ is stationary.

pg.12 2) Describe and implement the univariate approach to test whether e doesn't G-cause prod, rw, and U.

→ Fit an ARMA($p_\square, q_\square$), $\square$ = e, prod, rw, U, to each of time series. Denote the white noise of each time series as $\varepsilon_{e,t}$, $\varepsilon_{prod,t}$, $\varepsilon_{rw,t}$, and $\varepsilon_{U,t}$. Define CCR function at lag k of these residuals:

$$\rho_{ep}(k) = \frac{E(\varepsilon_{e,t}\varepsilon_{prod,t+k})}{\sqrt{E(\varepsilon_{e,t}^2)E(\varepsilon_{prod,t}^2)}}$$

$$\rho_{er}(k) = \frac{E(\varepsilon_{e,t}\varepsilon_{rw,t+k})}{\sqrt{E(\varepsilon_{e,t}^2)E(\varepsilon_{rw,t}^2)}}$$

$$\rho_{eu}(k) = \frac{E(\varepsilon_{e,t}\varepsilon_{u,t+k})}{\sqrt{E(\varepsilon_{e,t}^2)E(\varepsilon_{u,t+k}^2)}}$$

H_0: e does not Granger-cause prod

$$Q_{L,1} = n^2\sum_{k=0}^{L}(n-k)^{-1}\gamma_{ep}^2(k) \sim \chi^2_{L+1}$$

Reject H_0 for large $Q_{L,1}$.

H_0: e does not Granger-cause rw

$$Q_{L,2} = n^2\sum_{k=0}^{L}(n-k)^{-1}\gamma_{er}^2(k) \sim \chi^2_{L+1}.$$

H_0: e does not G-cause U.

$$Q_{L,3} = n^2\sum_{k=0}^{L}(n-k)^{-1}\gamma_{eu}^2(k) \sim \chi^2_{L+1}.$$

pg.15 3) Fit the CO_2 and gas model using Box and Tiao transformation.

$$\rightarrow \ co2_t = a + \sum_{i=3}^{7} v_i gas_{t-i} + e_t, \quad e_t \sim MA(6).$$

$$= a + V(B)gas_t + e_t, \quad e_t = \Theta_6(B)a_t.$$

Multiply each side by $\frac{1}{\Theta_6(B)}$ and obtain

$$\frac{1}{\Theta_6(B)}co2_t = \frac{a}{\Theta_6(B)} + V(B)\frac{gas_t}{\Theta_6(B)} + a_t$$

$$= a + V(B)\frac{gas_t}{\Theta_6(B)} + a_t$$

$$\Rightarrow \tilde{co2}_t = a + V(B)\tilde{gas}_t + a_t, \quad a_t \overset{iid}{\sim} N(0, \sigma_a^2).$$

WN.

pg. 16 4) What are two portmanteau tests in TFNM?

→ First, suppose $y_t = V(B)x_t + \varepsilon_t$, $x_t \sim ARMA$.

So ~~$\Phi(B)x_t$~~ $\Phi(B)x_t = \Theta(B)\alpha_t$.

We first need to go through prewhitening process.

Multiply each side by $\frac{\Phi(B)}{\Theta(B)}$:

$$\frac{\Phi(B)}{\Theta(B)} y_t = V(B)\alpha_t + \frac{\Phi(B)}{\Theta(B)} \varepsilon_t$$

$$= \tau_t = V(B)\alpha_t + n_t. \quad \text{Assume } \alpha_t \perp n_t.$$

Multiply each by α_{t-j}, $j \geq 0$:

$$\tau_t \alpha_{t-j} = V(B)\alpha_t \alpha_{t-j} + n_t \alpha_{t-j}$$

Take $E(\cdot)$:

$$Cov(\tau_t, \alpha_{t-j}) = V_j Var(\alpha_{t-j})$$

$$\Rightarrow V_j = \frac{Cov(\tau_t, \alpha_{t-j})}{Var(\alpha_{t-j})} = Corr(\tau_t, \alpha_{t-j}) \cdot \frac{se(\tau_t)}{se(\alpha_t)}.$$

① To see if $\alpha_t \perp n_t$, we conduct the following portmanteau test:

$$Q_0 = m(m+2)\sum_{j=0}^{k} (m-j)^{-1} \hat{\rho}^2_{\hat{\alpha}\hat{n}}(j) \sim \chi^2_{k+1-M}$$

with $\hat{\alpha}_t = \hat{\Phi}(B)x_t - \hat{\theta}_1\alpha_{t-1} - \cdots - \hat{\theta}_q \alpha_{t-q}$,

m = # of residuals ($\hat{n}_t$) calculated

M = # of V_j's estimated.

② To see if $n_t \sim WN$,

$$Q_1 = m(m+2)\sum_{j=0}^{k} (m-j)^{-1} \hat{\rho}_n(j) \sim \chi^2_{k-(p+q)},$$

where m is the same as above,

p, q are from $n_t \sim ARMA(p,q)$.

pg. 17 5) Consider the following: $\begin{cases} \nabla z_{1,t} = \alpha_1 (z_{1,t-1} - z_{2,t-1}) + u_{1,t} \\ \nabla z_{2,t} = \alpha_2 (z_{1,t-1} - z_{2,t-1}) + u_{2,t} \end{cases}$

Show that the process $y_t = z_{1,t} - z_{2,t}$ is AR and stationary if $|1+\alpha_1-\alpha_2| < 1$.

→ Rewrite the system as follows:

$$\nabla \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 & -\alpha_1 \\ \alpha_2 & -\alpha_2 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-2} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-2} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}.$$

So $\nabla \underset{\sim}{z}_t = \alpha\beta' \underset{\sim}{z}_{t-1} + \underset{\sim}{u}_t$

$= \underset{\sim}{z}_t - \underset{\sim}{z}_{t-1} = \alpha\beta' \underset{\sim}{z}_{t-1} + \underset{\sim}{u}_t$.

Multiply both by β':

$\beta' \underset{\sim}{z}_t - \beta' \underset{\sim}{z}_{t-1} = \beta'\alpha\beta' \underset{\sim}{z}_{t-1} + \beta' \underset{\sim}{u}_t$.

$\Rightarrow \beta' \underset{\sim}{z}_t = \beta' \underset{\sim}{z}_{t-1} + \beta'\alpha\beta' \underset{\sim}{z}_{t-1} + \beta' \underset{\sim}{u}_t$.

$= (I + \beta'\alpha)\beta' \underset{\sim}{z}_{t-1} + \beta' \underset{\sim}{u}_t$.

$\Rightarrow \beta' \underset{\sim}{z}_t = (I + \beta'\alpha)\beta' \underset{\sim}{z}_{t-1} + \underset{\sim}{e}_t, \quad \underset{\sim}{e}_t = \beta' \underset{\sim}{u}_t$.

So $\beta' \underset{\sim}{z}_t = z_{1,t} - z_{2,t}$ is stationary if all the absolute values of eigenvalues of $I + \beta'\alpha$ are less than 1.

And $I + \beta'\alpha = I + \begin{bmatrix} 1 & -1 \end{bmatrix}_{1\times 2} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}_{2\times 1}$

$= 1 + \alpha_1 - \alpha_2$

~~Since $I+\beta'\alpha$ is 1×1 matrix, the eigenvalue of this is itself.~~

Thus $(1 - (1+\alpha_1-\alpha_2)B)\beta' \underset{\sim}{z}_t = \underset{\sim}{e}_t$. All the roots of $1 - (1+\alpha_1-\alpha_2)B = 0$ must lie outside the unit circle. i.e. $|B| = \left|\frac{1}{1+\alpha_1-\alpha_2}\right| = \frac{1}{|1+\alpha_1-\alpha_2|} > 1$,

i.e. $\underline{|1+\alpha_1-\alpha_2| < 1}$. //

pg.20 6) Suppose $\nabla \underset{\sim}{Z}_t = \alpha\beta' \underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t$. Show that we need the absolute values of the eigenvalues of $I_r + \beta'\alpha$ all less than 1 for $\beta' \underset{\sim}{Z}_t$ to be stationary.

→ So $\nabla \underset{\sim}{Z}_t = \underset{\sim}{Z}_t - \underset{\sim}{Z}_{t-1} = \alpha\beta' \underset{\sim}{Z}_{t-1} + \underset{\sim}{U}_t$.

Multiply both by β' and obtain:

$$\beta' \underset{\sim}{Z}_t - \beta' \underset{\sim}{Z}_{t-1} = \beta'\alpha\beta' \underset{\sim}{Z}_{t-1} + \beta' \underset{\sim}{U}_t.$$

So $\beta' \underset{\sim}{Z}_t = (I + \beta'\alpha)\beta' \underset{\sim}{Z}_{t-1} + \beta' \underset{\sim}{U}_t$.

Thus we require $|\lambda_i| < 1 \ \forall i$, λ_i's are eigenvalues of $I + \beta'\alpha$. ✓

7) Suppose conflicting results/conclusions are found in trace test and maximum eigenvalue test. How would you resolve the conflicting inference results? Explain your reasons.

→ Go with results of trace test. ~~Trace test~~

Trace statistic, $-T \sum_{i=r_0+1}^{k} \ln(1-\lambda_i)$, considers all of the smallest eigenvalues. So it holds more power than the maximum eigenvalue statistic, $-T\ln(1-\lambda_{r_0+1})$.

pg.21 8) Do we see conflicting results in the example below? Explain your reasons.

	trace	10pct	5pct	1pct
r<=1	7.78	7.52	9.24	12.97
r=0	47.77	17.85	19.96	24.60

	Eigenmax	10pct	5pct	1pct
$r \le 1$	7.78	7.52	9.24	12.97
r=0	40.00	13.75	15.67	20.20

→ From "trace", we choose $r=1$ because $7.78 < 5$pct. Also, "eigenmax" suggests $r=1$ ∵ $7.78 < 5$pct. They are not conflicting.

pg. 28 9) Write down ECM for BHP and VALE using the results below.

6.1	BHP.l1	VALE.l1	constant
Eigenvalues	0.04	0.01	0
BHP.l1	1	1	1
VALE.l1	-0.72	-0.73	2.05
constant	-1.83	-1.54	-5.71

6.2.1	BHP.l1	VALE.l1	constant
BHP.d	-0.067	0.005	0
VALE.d	0.025	0.008	0

6.2.2	BHP.dl1	VALE.dl1
BHP.d	-0.115	0.069
VALE.d	0.053	0.045

$$[1 \;\; -0.72]\begin{bmatrix} BHP \\ VALE \end{bmatrix} = 1.83$$

$\rightarrow$ ~~So~~ ~~BHP~~ So 6.1 suggests $BHP = 0.72 VALE + 1.83$

6.2.1 suggests $\hat{\alpha} = (-0.067, 0.025)'$.

ECM is:

$$\nabla \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} [-\hat{\beta}, \; 1] \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \sum_{i=1}^{1} \begin{bmatrix} \gamma_{x,i} & \gamma_{y,i} \\ \gamma_{x,i} & \gamma_{y,i} \end{bmatrix} \nabla \begin{bmatrix} x_{t-i} \\ y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

$\nabla \underset{\sim}{z_t} = \alpha\beta' \underset{\sim}{z_{t-1}} + \underset{\sim}{u_t}$.

$$\therefore \nabla \begin{bmatrix} VALE_t \\ BHP_t \end{bmatrix} = \begin{bmatrix} 0.008 & 0.025 \\ 0.005 & -0.067 \end{bmatrix} \begin{bmatrix} VALE_{t-1} \\ BHP_{t-1} \end{bmatrix} + \begin{bmatrix} 0.045 & 0.053 \\ 0.069 & -0.115 \end{bmatrix} \nabla \begin{bmatrix} VALE_{t-1} \\ BHP_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

~~10) Redo the cointegration analysis using EG approach.~~

~~$\rightarrow \nabla BHP_t = \begin{bmatrix} -0.067 \\ 0.025 \end{bmatrix} E$~~

$\hat{\rho}_X(k) \sim N(0, \frac{1}{N})$, $\hat{\rho}_{XY}(k) \sim N(0, \frac{1}{n}(1+2\sum_j \rho_X(j)\rho_Y(j)))$

$$Q_{BP} = n\sum_{k=1}^{m} \hat{\rho}^2(k) \sim \chi^2_{m-(p+q)}$$

$$Q_{LB} = n(n+2)\sum_{k=1}^{m}(n-k)^{-1}\hat{\rho}^2(k) \sim \chi^2_{m-(p+q)}$$

$\hat{\rho}^2(k)$ " $\hat{\rho}_k^2$

\# of params in the model

$$AIC = -2\log ML + 2k$$

$$BIC = -2\log ML + k\log(n)$$

BDF:

$\nabla X_t = \pi X_{t-1} + a_t$, $\pi = \phi_1 - 1$.

$H_0: \pi = 0$ vs. $H_1: \pi > 0$ ($I(1)$ vs. $I(0)$).

BDF general:

$\nabla X_t = \pi X_{t-1} + a_t + \alpha + \tau^T DR_t$,

$\tau^T = (\alpha_1, \alpha_2, \dots)$

$DR_t = (t, t^2, \dots)$

→ Problems:

- → correct model specification assumed (correct trend and intercept)
- → DGP may have contained both AR and MA term
- → structural breaks in the data may exist.

ADF:

Use estimation on time trend, and ∇X_{t-j} terms to correct the effect of autocorrelated error terms.

$$\nabla X_t = \pi X_{t-1} + a_t + \tau^T DR_t + \sum_{j=1}^{k} \delta_j \nabla X_{t-j}, \quad k = p-1.$$

$H_0: \pi = 0$ vs. $H_1: \pi \neq 0$.

Selecting p:

① AR approx: $ARIMA(p,1,q) \approx ARIMA(k,1,0)$,

$k \leq T^{1/3}$.

(2) General to specific

Start with large p^*. Conduct the usual t- or F-test. If the statistic is insignificant (at lag p^*) at some specified critical value, reestimate the reg. using p^*-1. Continue until the last is sig. diff. from 0. Once k^*, tentative lag length, is determined, conduct the diagnostic check by plotting residual ACF plot and portmanteau tests on reg. residuals to ensure our k^* is legitimate.

✓ I(2) test.

$\nabla^2 x_t = \pi_1 \nabla x_{t-1} + a_t$. $H_0: \pi_1 = 0$ vs. $H_1: \pi_1 \neq 0$.

If H_0 not rejected, conclude $x_t \sim I(2)$.

Then write $\nabla^2 x_t = \pi_1 \nabla x_{t-1} + \pi_2 x_{t-1} + a_t$.

$H_0: \pi_1 < 0$ and $\pi_2 = 0$ (presence of single unit root).

If H_0 rejected, conclude $x_t \sim I(0)$.

✓ TFNM

(1) $y_t = \sum_{i=0}^{k} v_i x_{t-i} + a_t$, $a_t \overset{iid}{\sim} N(0,1)$.

(2) $(1-\phi B)x_t = e_t$, $e_t \overset{iid}{\sim} N(0,1)$.

Since $y_t = V(B)x_t + a_t$, we multiply both by $(1-\phi B)$:

$(1-\phi B)y_t = \tau_t = V(B)e_t + \varepsilon_t$, $\varepsilon_t = (1-\phi B)a_t$.

Assume $e_t \perp\!\!\!\perp \varepsilon_t$. Now multiply both by e_{t-j}:

$\tau_t e_{t-j} = V(B) e_t e_{t-j} + e_{t-j}\varepsilon_t$.

Take $E(\cdot)$:

$Cov(\tau_t, e_{t-j}) = v_j Var(e_{t-j})$

$$\Rightarrow v_j = \frac{Cov(\tau_t, e_{t-j})}{Var(e_{t-j})} = Corr(\tau_t, e_{t-j}) \cdot \frac{se(\tau_t)}{se(e_t)} = \rho_{\tau e}(j) \cdot \frac{se\,\tau_t}{se\,e_t}$$

✓ TFNM portmanteau test

① Is really $e_t \perp \varepsilon_t$?

$$Q_0 = m(m+2)\sum_{j=0}^{k^*}(m-j)^{-1}\hat{\rho}^2_{e\varepsilon}(j) \sim \chi^2_{k^*+1-M}$$

m = # of res. ($\hat{\varepsilon}_t$) calculated,
M = # of parameters estimated in TFNM

② Is $\varepsilon_t \sim WN$?

$$Q_1 = m(m+2)\sum_{j=1}^{k^*}(m-j)^{-1}\hat{\rho}^2_{\varepsilon}(j) \sim \chi^2_{k^*-(p+q)}.$$

$\varepsilon_t \sim ARMA(p,q)$.

✓ Box and Tiao.

$y_t = V(B)x_t + e_t$, $e_t \sim ARMA(p,q)$,
So $\Phi_p(B)e_t = \Theta_q(B)a_t$.

$$\Rightarrow \frac{\Phi_p(B)}{\Theta_q(B)}y_t = V(B)\cdot\frac{\Phi_p}{\Theta_q}x_t + a_t$$

$$\Rightarrow \tilde{y}_t = V(B)\tilde{x}_t + a_t.$$

① Run OLS on $y_t = V(B)x_t + e_t$. Collect $\{\hat{e}_t\}$.
② Fit $ARMA(p,q)$ on $\hat{e}_t$.
③ Apply the procedure above.
④ Run OLS on $\tilde{y}_t = V(B)\tilde{x}_t + a_t$.
⑤ Check whether $\{\hat{a}_t\}$ are serially correlated.
If it is, repeat ② to ④.

✓ VAR(p)

$$\underset{\sim}{y}_t = \sum_{i=1}^{p} A_i \underset{\sim}{y}_{t-i} + \underset{\sim}{u}_t, \quad \underset{\sim}{u}_t \overset{iid}{\sim} N_k(\vec{0}, \Sigma_u).$$

$$A_p(B)\underset{\sim}{y}_t = \underset{\sim}{u}_t.$$

$\underset{\sim}{y}_t$ stationary if $A_p(B)$ has all the eigenvalues greater than 1 in its abs. value.

Or, $\underset{\sim}{\xi}_t = \underset{\sim}{A}\,\underset{\sim}{\xi}_{t-1} + \underset{\sim}{V}_t$, $\underset{\sim}{A} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}$,

$\underset{\sim}{\xi}_t = \begin{bmatrix} \underset{\sim}{y}_t \\ \vdots \\ \underset{\sim}{y}_{t-(p-1)} \end{bmatrix}$, $\underset{\sim}{V}_t = \begin{bmatrix} \underset{\sim}{u}_t \\ \vec{0} \\ \vdots \\ \vec{0} \end{bmatrix}$,

and all $|\lambda_i|$'s < 1 of $\underset{\sim}{A}$.

✓ p selection

→ $AIC = \ln\det(\hat{\Sigma}_u(p)) + \frac{2}{T}pk^2$.

→ Sequential LRT : p-1 vs. p

$$\hat{\Sigma}_u = \frac{1}{T}\sum_{t=1}^{T} \hat{\underset{\sim}{u}}_t \hat{\underset{\sim}{u}}_t'.$$

✓ Granger causality : VAR approach

$$\begin{bmatrix} \underset{\sim}{y}_{1,t} \\ \underset{\sim}{y}_{2,t} \end{bmatrix} = \sum_{j=1}^{p} \begin{bmatrix} \phi_{11}^{(j)} & \phi_{12}^{(j)} \\ \phi_{21}^{(j)} & \phi_{22}^{(j)} \end{bmatrix} \begin{bmatrix} \underset{\sim}{y}_{1,t-i} \\ \underset{\sim}{y}_{2,t-i} \end{bmatrix} + \underset{\sim}{u}_t.$$

H_0: $y_{2,t}$ does not G-cause $y_{1,t}$.

$\phi_{12}^{(j)} = 0\ \forall j$ vs. H_1: $\phi_{12}^{(j)} \neq 0\ \exists j$.

H_0: $y_{1,t}$ does not G-cause $y_{2,t}$.

$\phi_{21}^{(j)} = 0\ \forall j$ vs. H_1: $\phi_{21}^{(j)} \neq 0\ \exists j$.

✓ Granger-causality: univariate approach

$$\Phi_{p_x}(B) x_t = \Theta_{q_x}(B) u_t.$$
$$\Phi_{p_y}(B) y_t = \Theta_{q_y}(B) v_t$$

Define $\rho_{uv}(k) = \dfrac{E(u_t v_{t+k})}{\sqrt{E(u_t^2)E(v_t^2)}}$.

H_0: x_t does not G-cause y_t.

$$Q_L = n^2 \sum_{k=0}^{L} (n-k)^{-1} \hat{\gamma}_{uv}^2(k) \sim \chi^2_{L+1}$$

✓ Cointegration

$\underset{\sim}{z}_t \sim I(1)$ becomes $\beta' \underset{\sim}{z}_t \sim I(0)$.

Or $\underset{\sim}{x}_t, \underset{\sim}{y}_t \sim I(1)$ becomes

$a\underset{\sim}{x}_t + b\underset{\sim}{y}_t \sim I(0)$.

✓ Granger representation thm

If x_t and y_t are cointegrated, then there is a ECM representation for these two.

Idea: If component series are cointegrated, then VAR on I(1) process will be a misspecification. E and G showed that in VAR, an equilibrium specification is missing from a VAR representation. If, however, we include lagged disequilibrium terms, then the model becomes well specified. Such a model is called ECM ∵ the model is structured so that short-run deviation from the long-run equilibrium will be corrected.

✓ EG procedure

① Check if $x_t \sim I(1)$ and $y_t \sim I(1)$ using URT.

② If they are, regress series one another:
$y_t = \beta_0 + \beta_1 x_t + e_t$.

③ Collect $\{\hat{e}_t\}$ (OLS residuals) and run URT on $\{\hat{e}_t\}$.

④ If $\hat{e}_t \sim I(0)$, conclude x_t and y_t are cointegrated.

⑤ Write the ECM:

$$\nabla \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \begin{bmatrix} -\hat{\beta}_1 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} \beta_{x,i} & \beta_{y,i} \\ \delta_{x,i} & \delta_{y,i} \end{bmatrix} \nabla \begin{bmatrix} x_{t-i} \\ y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

✓ Pairs trading

① Define spread $= S_1 - bS_2$.

② Determine the trading signal.

③ Determine your position and construct a portfolio.

④ Size your position (risk/capital mgmt).

⑤ Calculate transaction cost.

✓ VAR(p) portmanteau tests

$\underset{\sim}{z}_t \sim VAR(p)$,

$$Q_h = T \sum_{j=1}^{h} tr(\hat{C}_j^T \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}) \sim \chi^2_{k^2(h-p)}$$

$$\hat{C}_i = \frac{1}{T} \sum_{t=1}^{T} \hat{\underset{\sim}{u}}_t \hat{\underset{\sim}{u}}_{t-i}$$

$$Q_h^* = T \sum_{j=1}^{h} (T-j)^{-1} tr(\hat{C}_j^T \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}) \sim \chi^2_{k^2(h-p)}.$$

✓ Model for cointegration:

$$\nabla \underset{\sim}{Z}_t = \alpha\beta' \underset{\sim}{Z}_{t-1} + \underset{\sim}{u}_t .$$

If $\underset{\sim}{Z}_t \sim VAR(p)$, i.e. $\underset{\sim}{Z}_t = \sum_{i=1}^{p} \Phi_i \underset{\sim}{Z}_{t-i} + \underset{\sim}{u}_t$,

then define $\begin{cases} \Gamma_i = -(I - \Phi_1 - \cdots - \Phi_i), \ i=1,\cdots,p-1 \\ \Gamma_p = -(I - \Phi_1 - \cdots - \Phi_p) = \Pi . \end{cases}$

$\underset{\sim}{Z} \in k \times 1$ Then $\nabla \underset{\sim}{Z}_t = \sum_{i=1}^{p-1} \Gamma_i \nabla \underset{\sim}{Z}_{t-i} + \Pi \underset{\sim}{Z}_{t-p} + \underset{\sim}{u}_t .$

rank$(\Pi) = 0 \rightarrow \Pi = 0$.
rank$(\Pi) = k \rightarrow \underset{\sim}{Z}_t$ is stationary.
rank$(\Pi) = r \in (0,k)$
$\Rightarrow \exists\, \alpha_{k\times r}$ and $\beta_{k\times r}$ s.t. α is the impact on cointegrating series on $\nabla \underset{\sim}{Z}_t$, and $\beta' \underset{\sim}{Z}_t \sim I(0)$;
$\Pi = \alpha\beta'$.

✓ Johansen procedure

Estimate Π by the following:

$\nabla \underset{\sim}{Z}_t = \sum_{i=1}^{p-1} \Phi_i \nabla \underset{\sim}{Z}_{t-i} + \underset{\sim}{u}_t$ $\quad \Pi$ is related to

$\underset{\sim}{Z}_{t-1} = \sum_{i=1}^{p-1} \Phi_i^* \nabla \underset{\sim}{Z}_{t-i} + \underset{\sim}{V}_t$. Var matrix btwn $\nabla \underset{\sim}{Z}_t$ and $\underset{\sim}{Z}_{t-1}$.

Collect OLS residuals $\{\hat{\underset{\sim}{u}}_t\}$ and $\{\hat{\underset{\sim}{V}}_t\}$.

Let $\hat{\underset{\sim}{u}}_t = \Pi \hat{\underset{\sim}{V}}_t + \underset{\sim}{\varepsilon}_t$.

$$\hat{\Pi} = \hat{\Sigma}_{11}^{-1} \hat{\Sigma}_{10} \hat{\Sigma}_{00}^{-1} \hat{\Sigma}_{01} ,$$

$$\hat{\Sigma}_{00} = \frac{1}{T}\sum_{t=1}^{T} \hat{\underset{\sim}{u}}_t \hat{\underset{\sim}{u}}_t' , \quad \hat{\Sigma}_{01} = \hat{u}_t \hat{V}_t' ,$$

$$\hat{\Sigma}_{10} = \frac{1}{T}\sum_{t=1}^{T} \hat{\underset{\sim}{V}}_t \hat{\underset{\sim}{u}}_t' , \quad \hat{\Sigma}_{11} = \hat{V}_t \hat{V}_t' .$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ be ordered eigenvalues of $\hat{\Pi}$.

① Trace test: $H_0: r = r_0$ vs. $H_1: r > r_0$.

$$\lambda_{tr}(r_0) = -T \sum_{i=r_0+1}^{k} \ln(1-\lambda_i).$$

Reject H_0 for large $\lambda_{tr}(r_0)$.

② Max eigenvalue test: $H_0: r = r_0$ vs. $H_1: r = r_0+1$.

$$\lambda_{max}(r_0) = -T \ln(1-\lambda_{r_0+1}).$$

Reject H_0 for large $\lambda_{max}(r_0)$.

✓ State-space model (SSM)

Two eqn's:

(1) obs. eqn: relationship btwn observed variables and unobserved state variables

$$y_t = F' x_t + \varepsilon_t.$$

(2) State eqn: dynamics of state variables x_t.

$$x_t = G x_{t-1} + w_t.$$

✓ Bootstrapping: regression

$y_t = X_t \beta + u_t$, $E(u_t|X_t) = 0$, $u_t \overset{iid}{\sim} (0, \sigma^2)$.

n obs, k regressors.

① Residual bootstrap: (minimal dist. assumption require, indep. of regressors and iid

1) Run OLS to obtain $\hat{\beta}$ and $\hat{u}_t$.

2) Generate a sample as

$$y_t^* = X_t \hat{\beta} + u_t^*, \quad u_t^* \sim edf(\ddot{u}_t), \ \ddot{u}_t = \left(\frac{n}{n-k}\right)^{\frac{1}{2}} \hat{u}_t.$$

② Parametric : assume $u_t \sim$ specific dist.
1) Run OLS to obtain $\hat{\beta}$ and $\hat{u}_t$.
2) Generate data using $y_t^* = \underset{\sim}{X}_t \hat{\beta} + u_t^*$,
$u_t^* \overset{iid}{\sim} N(0, s^2)$, s^2 = sample var of $\hat{u}_t$.

③ Wild : designed to handle heteroscedasticity.
e.g. $y_t^* = \underset{\sim}{X}_t \hat{\beta} + f(\hat{u}_t) v_t^*$,
$$f(\hat{u}_t) = \frac{\hat{u}_t}{\sqrt{1-h_{ii}}}.$$

④ Pairs : generate $[y_t \; \underset{\sim}{X}_t]$.
Pros:
① valid even when errors have non-constant var.
② works even for dynamic models
③ applicable to enormous amt of models
④ in case of multivariate, we can combine the pairs and residual bootstrap.
Cons:
① DGP doesn't impose H_0's restrictions on β.
② Doesn't yield very accurate results.

Bootstrapping : dependent data
① Parametric.
1) Find unconditional $x_t \sim dist(\mu_1, \sigma_1^2)$.
2) Simulate $x_0 \sim dist(\mu_1, \sigma_1^2)$.
3) Find $X_t | X_{t-1} \sim dist(\mu_2, \sigma_2^2)$

Note: if $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, then $X_1 | X_2 = x_2 \sim N(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(x_2 - \mu_2), \sigma_1^2(1-\rho^2))$

4) Simulate X_t's. Iterate.

② Sieve : $y_t = X_t\beta + u_t$.
Assumption: $u_t \sim ARMA$ of stationary process, constant var.
1) Estimate the model, and obtain $\hat{u}_t$.
2) Estimate $\hat{u}_t \sim AR(p)$: $\hat{u}_t = \sum_{i=1}^{p} \phi_i \hat{u}_{t-i} + \varepsilon_t$. (1).

3) Generate bootstrap error terms:
$u_t^* = \sum \hat{\phi}_i u_{t-i}^* + \varepsilon_t^*$ - (2) where $\{\varepsilon_t^*\}$ are (rescaled) resampled residuals from (1).
4) Generate the bootstrap data according to (2):
$y_t^* = \underset{\sim}{X_t}\hat{\beta} + u_t^*$.

③ Block
· Divide the data of n obs into b blocks of length l. Resample with replacement all the possible blocks.
· Cons:
1) Even the original is stationary, blocks aren't.
2) Mean of MBB $\bar{x}^*$ is biased:
$E(\bar{x}^* | x_1, \cdots, x_n) - \bar{x} \neq 0$.
3) Var est. of MBB also biased.
· Choice of l is critical: it must ↑ as n↑.
Usually, $l \propto n^{1/3}$.
· Block-of-blocks:
Consider $y_t = \underset{\sim}{X_t}\beta + \gamma y_{t-1} + u_t$.
Define $z_t = [y_t, y_{t-1}, \underset{\sim}{X_t}]$.
$Block_1 = (\underset{\sim}{z_1}, \cdots, \underset{\sim}{z_l})$
⋮
$Block_b = (\underset{\sim}{z_{b1}}, \cdots, \underset{\sim}{z_n})$.
→ works w/ non-constant var as well as serial correlation; offer more accuracy than asymptotic methods but only by a modest extent; yields more reliable s.e.'s.