8.6 Testing Cointegration

8.6.1 Engle-Granger Test

$y_t \sim I(1)$

$y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$

- $u_t \sim I(0) \implies$ Cointegration
- $u_t \sim I(1) \implies$ Spurious Regression

Estimate $y_{1,t} = \alpha + \gamma' y_{2,t} + u_t$ by OLS, and obtain $\hat{u}_t$.

Estimate $\hat{u}_t = \rho \hat{u}_{t-1} + \delta_1 \Delta \hat{u}_{t-1} + \delta_2 \Delta \hat{u}_{t-2} + \cdots + \delta_{p-1} \Delta \hat{u}_{t-p+1} + e_t$ by OLS.

ADF Test:

- $H_0: \ \rho = 1$ (Sprious Regression)
- $H_1: \ \rho < 1$ (Cointegration)

$\Longrightarrow$ **Engle-Granger Test**

For example, see Engle and Granger (1987), Phillips and Ouliaris (1990) and Hansen (1992).

Asymmptotic Distribution of Residual-Based ADF Test for Cointegration

# of Refressors,	(a) Regressors have no drift				(b) Some regressors have drift			
excluding constant	1%	2.5%	5%	10%	1%	2.5%	5%	10%
1	−3.96	−3.64	−3.37	−3.07	−3.96	−3.67	−3.41	−3.13
2	−4.31	−4.02	−3.77	−3.45	−4.36	−4.07	−3.80	−3.52
3	−4.73	−4.37	−4.11	−3.83	−4.65	−4.39	−4.16	−3.84
4	−5.07	−4.71	−4.45	−4.16	−5.04	−4.77	−4.49	−4.20
5	−5.28	−4.98	−4.71	−4.43	−5.36	−5.02	−4.74	−4.46

J.D. Hamilton (1994), *Time Series Analysis*, p.766.

8.6.2 Error Correction Representation

VAR(p) model:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where y_t, α and ϵ_t indicate $g \times 1$ vectors for $t = 1, 2, \cdots, T$, and ϕ_s is a $g \times g$ matrix for $s = 1, 2, \cdots, p$.

Rewrite:

$$y_t = \alpha + \rho y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \cdots + +\delta_{p-1} \Delta y_{t-p+1} + \epsilon_t,$$

where

$$\rho = \phi_1 + \phi_2 + \cdots + \phi_p,$$

$$\delta_s = -(\phi_{s+1} + \delta_{s+2} + \cdots + \phi_p), \qquad \text{for } s = 1, 2, \cdots, p-1.$$

Again, rewrite:

$$\Delta y_t = \alpha + \delta_0 y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \cdots + +\delta_{p-1} \Delta y_{t-p+1} + \epsilon_t,$$

where

$$\delta_0 = \rho - I_g = -\phi(1),$$

for $\phi(L) = I_g - \delta_1 L - \delta_2 L^2 - \cdots - \delta_p L^p$.

If y_t has h cointegrating relations, we have the following error correction representation:

$$\Delta y_t = \alpha - BA' y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \cdots + +\delta_{p-1} \Delta y_{t-p+1} + \epsilon_t,$$

where $A' y_{t-1}$ is a stationary $h \times 1$ vector (i.e., h I(0) processes), and B and A are $g \times h$ matrices.

Note that $\phi(1) = BA'$ for $\phi(L) = I_g - \delta_1 L - \delta_2 L^2 - \cdots - \delta_p L^p$.

Each row of A' denotes the cointegrating vector, i.e., A' consists of h cointegrating vectors.

Suppose that $\epsilon_t \sim N(0, \Sigma)$. The log-likelihood function is:

$$\log l(\alpha, \delta_1, \cdots, \delta_{p-1}, B|A)$$

$$
\begin{aligned}
= \; & -\frac{Tg}{2}\log(2\pi) - \frac{T}{2}\log|\Sigma| \\
& -\frac{1}{2}\sum_{t=1}^{T}(\Delta y_t - \alpha + BA' y_{t-1} - \delta_1 \Delta y_{t-1} - \cdots - \delta_{p-1}\Delta y_{t-p+1})'\Sigma^{-1} \\
& \qquad \times(\Delta y_t - \alpha + BA' y_{t-1} - \delta_1 \Delta y_{t-1} - \cdots - \delta_{p-1}\Delta y_{t-p+1})
\end{aligned}
$$

Given A and h, maximize $\log l$ with respect to $\alpha, \delta_1, \cdots, \delta_{p-1}, B$.

Then, given h, how do we estimate A? $\Longrightarrow$ Johansen (1988, 1991)

(*) Canonical Correlatoion (正準相関)

$x' = (x_1, x_2, \cdots, x_n)$ and $y' = (y_1, y_2, \cdots, y_m)$, where $n \leq m$.

$$u = a'x = a_1x_1 + a_2x_2 + \cdots + a_nx_n,$$

$$v = b'y = b_1y_1 + b_2y_2 + \cdots + b_my_m,$$

where $\mathrm{V}(u) = \mathrm{V}(v) = 1$ and $\mathrm{E}(x) = \mathrm{E}(y) = 0$ for simplicity.

Define:

$$\mathrm{V}(x) = \Sigma_{xx}, \qquad \mathrm{E}(xy') = \Sigma_{xy}, \qquad \mathrm{V}(y) = \Sigma_{yy}, \qquad \mathrm{E}(yx') = \Sigma_{yx} = \Sigma'_{xy}.$$

The correlation coefficient between u and v, denoted by ρ, is:

$$\rho = \frac{\mathrm{Cov}(u, v)}{\sqrt{\mathrm{V}(u)}\sqrt{\mathrm{V}(v)}} = a'\Sigma_{xy}b,$$

where $\mathrm{V}(u) = a'\Sigma_{xx}a = 1$ and $\mathrm{V}(v) = b'\Sigma_{yy}b = 1$.

Maximize $\rho = a'\Sigma_{xy}b$ subject to $a'\Sigma_{xx}a = 1$ and $b'\Sigma_{yy}b = 1$.

The Lagrangian is:

$$L = a'\Sigma_{xy}b - \frac{1}{2}\lambda(a'\Sigma_{xx}a - 1) - \frac{1}{2}\mu(b'\Sigma_{yy}b - 1).$$

Take a derivative with respect to a and b.

$$\frac{\partial L}{\partial a} = \Sigma_{xy} b - \lambda \Sigma_{xx} a = 0, \qquad \frac{\partial L}{\partial b} = \Sigma'_{xy} a - \mu \Sigma_{yy} b = 0.$$

Using $a'\Sigma_{xx}a = 1$ and $b'\Sigma_{yy}b = 1$, we obtain:

$$\lambda = \mu = a'\Sigma_{xy} b.$$

From the first equation, we obtain:

$$a = \frac{1}{\lambda} \Sigma_{xx}^{-1} \Sigma_{xy} b,$$

which is substituted into the second equation as follows:

$$\frac{1}{\lambda} \Sigma'_{xy} \Sigma_{xx}^{-1} \Sigma_{xy} b - \lambda \Sigma_{yy} b = 0,$$

i.e.,

$$(\Sigma_{yy}^{-1} \Sigma'_{xy} \Sigma_{xx}^{-1} \Sigma_{xy} - \lambda^2 I_m) b = 0,$$

i.e.,

$$|\Sigma_{yy}^{-1} \Sigma'_{xy} \Sigma_{xx}^{-1} \Sigma_{xy} - \lambda^2 I_m| = 0.$$

The solution of λ^2 is given by the maximum eigen value of $\Sigma_{yy}^{-1} \Sigma'_{xy} \Sigma_{xx}^{-1} \Sigma_{xy}$, and b is the corresponding eigen vector.

Back to the Cointegration:

Estimate the following two regressions:

$$\Delta y_t = b_{1,0} + b_{1,1}\Delta y_{t-1} + b_{1,2}\Delta y_{t-2} + \cdots + b_{1,p-1}\Delta y_{t-p+1} + u_{1,t}$$

$$y_{t-1} = b_{2,0} + b_{2,1}\Delta y_{t-1} + b_{2,2}\Delta y_{t-2} + \cdots + b_{2,p-1}\Delta y_{t-p+1} + u_{2,t}$$

Obtain $\hat{u}_{i,t}$ for $i = 1, 2$ and $t = 1, 2, \cdots, T$, and compute as follow:

$$\hat{\Sigma}_{11} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{1,t}\hat{u}'_{1,t}, \qquad \hat{\Sigma}_{22} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{2,t}\hat{u}'_{2,t},$$

$$\hat{\Sigma}_{12} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{1,t}\hat{u}'_{2,t}, \qquad \hat{\Sigma}_{21} = \hat{\Sigma}'_{12}.$$

From $\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{\Sigma}_{12}$, compute h biggest eigenvalues, denoted by $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\cdots$, $\hat{\lambda}_h$, and the corresponding eigen vectors, denoted by $\hat{a}_1$, $\hat{a}_2$, $\cdots$, $\hat{a}_h$, where $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_h$,

The estimate of A, $\hat{A}$, is given by $\hat{A} = (\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_h)$.

How do we obtain h?

8.7 Testing the Number of Cointegrating Vectors

Trace Test (トレース検定): $H_0 : \lambda_{h+1} = 0$ and $H_1 : \lambda_h > 0$.

$$2(\log l_1 - \log l_0) = -T \sum_{i=h+1}^{g} \log(1 - \hat{\lambda}_i) \longrightarrow \mathrm{tr}(Q),$$

where

$$Q = \Big(\int_0^1 W(r)\mathrm{d}W(r)'\Big)' \Big(\int_0^1 W(r)W(r)'\mathrm{d}r\Big)^{-1} \Big(\int_0^1 W(r)\mathrm{d}W(r)'\Big).$$

Trace Test for # of Cointegrating Relations

# of Random	(a) Regressors have no drift				(b) Some regressors have drift			
Walks ($g-h$)	1%	2.5%	5%	10%	1%	2.5%	5%	10%
1	11.576	9.658	8.083	6.691	6.936	5.332	3.962	2.816
2	21.962	19.611	17.844	15.583	19.310	17.299	15.197	13.338
3	37.291	34.062	31.256	28.436	35.397	32.313	29.509	26.791
4	55.551	51.801	48.419	45.248	53.792	50.424	47.181	43.964
5	77.911	73.031	69.977	65.956	76.955	72.140	68.905	65.063

J.D. Hamilton (1994), *Time Series Analysis*, p.767.

Largest Eigenvalue Test (最大固有値検定):

$H_0: \lambda_{h+1} = 0$ and $H_1: \lambda_h > 0$.

$$2(\log l_1 - \log l_0) = -T \log(1 - \hat{\lambda}_{h+1}) \longrightarrow \text{maxmum eigen value of } Q,$$

Maximum Eigenvalue Test for # of Cointegrating Relations

# of Random	(a) Regressors have no drift				(b) Some regressors have drift			
Walks $(g-h)$	1%	2.5%	5%	10%	1%	2.5%	5%	10%
1	11.576	9.658	8.083	6.691	6.936	5.332	3.962	2.816
2	18.782	16.403	14.595	12.783	17.936	15.810	14.036	12.099
3	26.154	23.362	21.279	18.959	25.521	23.002	20.778	18.697
4	32.616	29.599	27.341	24.917	31.943	29.335	27.169	24.712
5	38.858	35.700	33.262	30.818	38.341	35.546	33.178	30.774

J.D. Hamilton (1994), *Time Series Analysis*, p.768.

9　周波数領域

周波数領域（Frequency Domain）：

1. スペクトラム (パワー・スペクトラム) の定義：

$$\begin{aligned} f(\lambda) &= (2\pi)^{-1} \sum_{\tau=-\infty}^{\infty} \gamma(\tau)\cos(\lambda\tau) \\ &= (2\pi)^{-1} \sum_{\tau=-\infty}^{\infty} \gamma(\tau)\exp(-i\lambda\tau) \end{aligned}$$

2. 三角関数と指数関数： $\exp(i\theta) = \cos(\theta) + i\sin(\theta)$

   $\Longrightarrow$

$$\cos(\theta) = \frac{1}{2}\Big(\exp(i\theta) + \exp(-i\theta)\Big), \qquad \sin(\theta) = \frac{1}{2i}\Big(\exp(i\theta) - \exp(-i\theta)\Big)$$

加法定理： $\exp\Big(i(\theta_1+\theta_2)\Big) = \exp(i\theta_1)\exp(i\theta_2)$

$\Longrightarrow$

$$\exp\Big(i(\theta_1+\theta_2)\Big) = \cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)$$

$$\begin{aligned}\exp(i\theta_1)\exp(i\theta_2) &= \Big(\cos(\theta_1)+i\sin(\theta_1)\Big)\Big(\cos(\theta_2)+i\sin(\theta_2)\Big)\\ &= \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2) + i\Big(\cos(\theta_1)\sin(\theta_2)+\sin(\theta_1)\cos(\theta_2)\Big)\end{aligned}$$

よって，

$$\cos(\theta_1+\theta_2) = \cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2)$$

$$\sin(\theta_1+\theta_2) = \cos(\theta_1)\sin(\theta_2) + \sin(\theta_1)\cos(\theta_2)$$

が得られる。

3. y_t がホワイト・ノイズであれば， $f(\lambda) = (2\pi)^{-1}\sigma_\epsilon^2$

4. スペクトラムと自己相関関数との関係

$$\gamma(\tau) = \int_{-\pi}^{\pi} f(\lambda)\cos(\lambda\tau)\mathrm{d}\lambda$$

従って，スペクトラムは自己相関関数のすべての情報を持っている。

5. $\sum w_j^2 < \infty$ とする。

$$y_t = \sum_{j=-r}^{s} w_j x_{t-j}$$

$f_x(\lambda)$ を x_t のスペクとラムとする。 $W(\lambda)$ を次のように定義する。

$$W(\lambda) = \sum_{j=-r}^{s} w_j e^{-i\lambda j}$$

このとき， y_t のスペクとラムは以下のようになる。

$$f_y(\lambda) = |W(\lambda)|^2 f_x(\lambda)$$

$|W(\lambda)|^2$ は伝達関数 (transfer function) と呼ばれ,

$$\begin{aligned} |W(\lambda)|^2 &= W(\lambda)\overline{W(\lambda)} \\ &= \sum_{j=-r}^{s} w_j e^{-i\lambda j} \sum_{j=-r}^{s} w_j e^{i\lambda j} \end{aligned}$$

$\overline{W(\lambda)}$ は $W(\lambda)$ の共役複素数とする。

6. MA(q) モデルの場合：

$$\begin{aligned} y_t &= \epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_q\epsilon_{t-q} \\ &= (1 + \theta_1 L + \cdots + \theta_q L^q)\epsilon_t \\ &= \theta(L)\epsilon_t \end{aligned}$$

$y_t = \theta(L)\epsilon_t$ のとき，ϵ_t のパワー・スペクトラム $f_\epsilon(\lambda)$ から y_t のパワー・スペクトラム $f_y(\lambda)$ への変換：

$$\begin{aligned} f_y(\lambda) &= \theta(e^{-i\lambda})\theta(e^{i\lambda})f_\epsilon(\lambda) \\ &= \theta(e^{-i\lambda})\theta(e^{i\lambda})\frac{\sigma_\epsilon^2}{2\pi} \end{aligned}$$

→ 基本パターン

7. AR(p) モデルの場合:

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

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

$\phi(L)y_t = \epsilon_t$ のとき，ϵ_t のパワー・スペクトラム $f_\epsilon(\lambda)$ から y_t のパワー・スペクトラム $f_y(\lambda)$ への変換：

$$
\begin{aligned}
f_y(\lambda) &= \frac{1}{\phi(e^{-i\lambda})\phi(e^{i\lambda})} f_\epsilon(\lambda) \\
&= \frac{1}{\phi(e^{-i\lambda})\phi(e^{i\lambda})} \frac{\sigma_\epsilon^2}{2\pi}
\end{aligned}
$$

8. ARMA(p, q) モデルの場合:

$$
\phi(L)y_t = \theta(L)\epsilon_t
$$

$$
y_t = \phi(L)^{-1}\theta(L)\epsilon_t
$$

$\phi(L)y_t = \theta(L)\epsilon_t$ のとき，ϵ_t のパワー・スペクトラム $f_\epsilon(\lambda)$ から y_t のパワー・スペクトラム $f_y(\lambda)$ への変換：

$$
\begin{aligned}
f_y(\lambda) &= \frac{\theta(e^{-i\lambda})\theta(e^{i\lambda})}{\phi(e^{-i\lambda})\phi(e^{i\lambda})} f_\epsilon(\lambda) \\
&= \frac{\theta(e^{-i\lambda})\theta(e^{i\lambda})}{\phi(e^{-i\lambda})\phi(e^{i\lambda})} \frac{\sigma_\epsilon^2}{2\pi}
\end{aligned}
$$

10 Generalized Method of Moments (GMM, 一般化積率法)

10.1 Method of Moments (MM, 積率法)

As $n \longrightarrow \infty$, we have the result: $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mathrm{E}(X) = \mu$.

$\Longrightarrow$ **Law of Large Number (大数の法則)**

$X_1, X_2, \cdots, X_n$ are n realizations of X.

[Review] Chebyshev's inequality (チェビシェフの不等式) is given by:

$$P(|X-\mu| > \epsilon) \le \frac{\sigma^2}{\epsilon^2} \quad \text{or} \quad P(|X-\mu| \le \epsilon) \ge 1 - \frac{\sigma^2}{\epsilon^2},$$

where $\mu = \mathrm{E}(X)$, $\sigma^2 = \mathrm{V}(X)$ and any $\epsilon > 0$.

Note that $P(|X-\mu| > \epsilon) + P(|X-\mu| \le \epsilon) = 1$.

Replace X, $\mathrm{E}(X)$ and $\mathrm{V}(X)$ by $\overline{X}$, $\mathrm{E}(\overline{X}) = \mu$ and $\mathrm{V}(\overline{X}) = \dfrac{\sigma^2}{n}$.

As $n \longrightarrow \infty$,

$$P(|\overline{X}-\mu| \le \epsilon) \ge 1 - \frac{\sigma^2}{n\epsilon^2} \longrightarrow 1.$$

That is, $\overline{X} \longrightarrow \mu$ as $n \longrightarrow \infty$.

[End of Review]

$\overline{X}$ is an approximation of $E(X) = \mu$.

Therefore, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is taken as an estimator of μ.

$\Longrightarrow$ $\overline{X}$ is MM estimator of $E(X) = \mu$.

MM is applied to the regression model as follows:

Regression model: $y_i = x_i\beta + u_i$, where x_i and u_i are assumed to be stochastic.

Familiar Assumption: $\mathrm{E}(x'u) = 0$, called the **orthogonality condition (直交条件)**, where x is a $1 \times k$ vector and u is a scalar.

We consider that $(x_1, x_2, \cdots, x_n)$ and $(u_1, u_2, \cdots, u_n)$ are realizations generated from random variables x and u, respectively.

From the law of large number, we have the following:

$$\frac{1}{n}\sum_{i=1}^{n} x_i' u_i = \frac{1}{n}\sum_{i=1}^{n} x_i'(y_i - x_i\beta) \longrightarrow \mathrm{E}(x'u) = 0.$$

Thus, the MM estimator of β, denoted by β_{MM}, satisfies:

$$\frac{1}{n}\sum_{i=1}^{n} x_i'(y_i - x_i\beta_{MM}) = 0.$$

Therefore, β_{MM} is given by:

$$\beta_{MM} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i' x_i\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i' y_i\right) = (X'X)^{-1}X'y,$$

which is equivalent to OLS and MLE.

Note that X and y are:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$