SPURIOUS REGRESSION

Set-up:

$$\boxed{\begin{aligned} y_t &= y_{t-1} + u_t \\ x_t &= x_{t-1} + v_t \end{aligned}} \quad \begin{aligned} &; \quad u_t \sim IID(0, \sigma_u^2) \\ &; \quad v_t \sim IID(0, \sigma_v^2) \end{aligned}$$

$$E(u_t v_s) = 0 \quad \forall t, s$$
$$E(u_t u_{t-k}) = E(v_t v_{t-k}) = 0 \quad \forall k \neq 0$$

Regress

$$\boxed{y_t = \alpha + \beta x_t + \varepsilon_t}$$

- What do you expect from this regression?

$$\hat{\beta} \xrightarrow{p} 0$$
$$R^2 \xrightarrow{p} 0$$

$$t_{\hat{\beta}} \Rightarrow t \text{ DISTRIBUTION}$$

- What does it really happens?

$$\hat{\beta} \Rightarrow \text{Distribution}$$

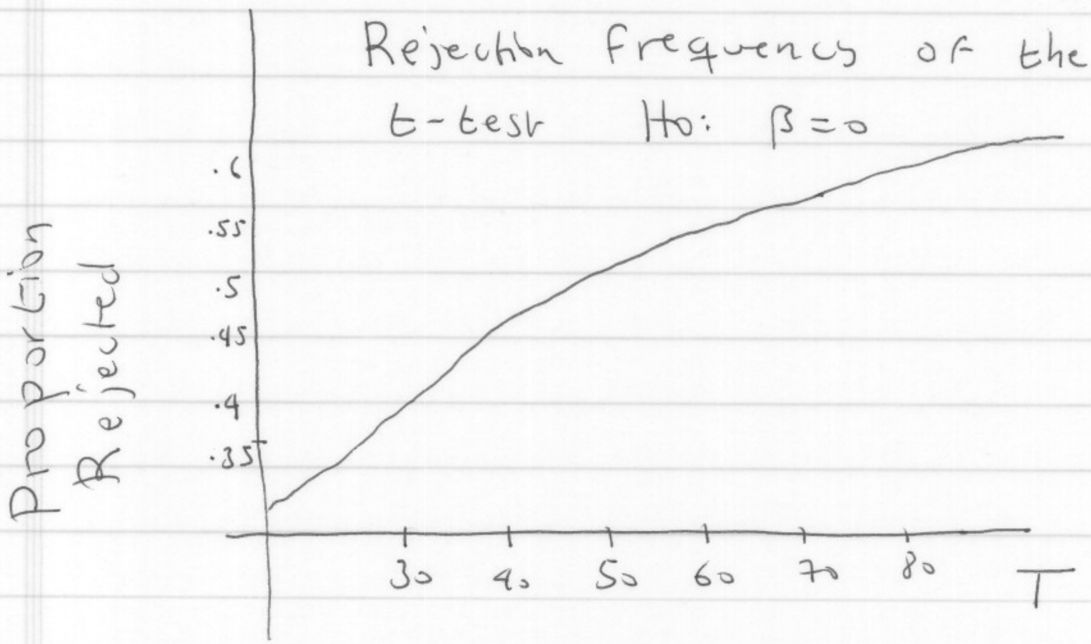
$$R^2 \Rightarrow \text{Distribution}$$

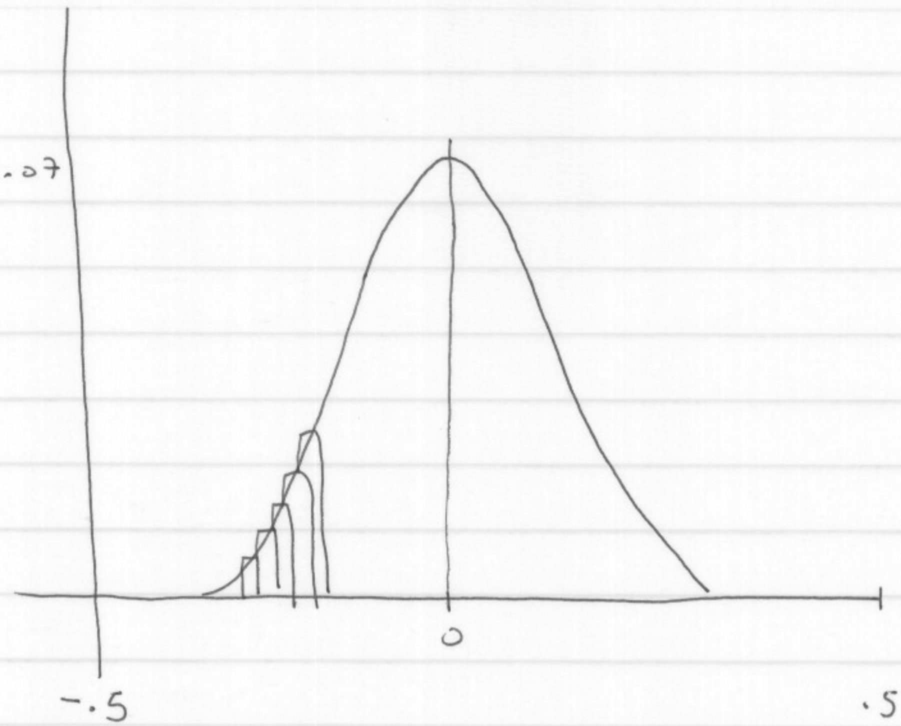
$$DW \xrightarrow{p} 0$$

$$T^{-\frac{1}{2}} t_{\hat{\beta}} \Rightarrow \text{Distribution}$$

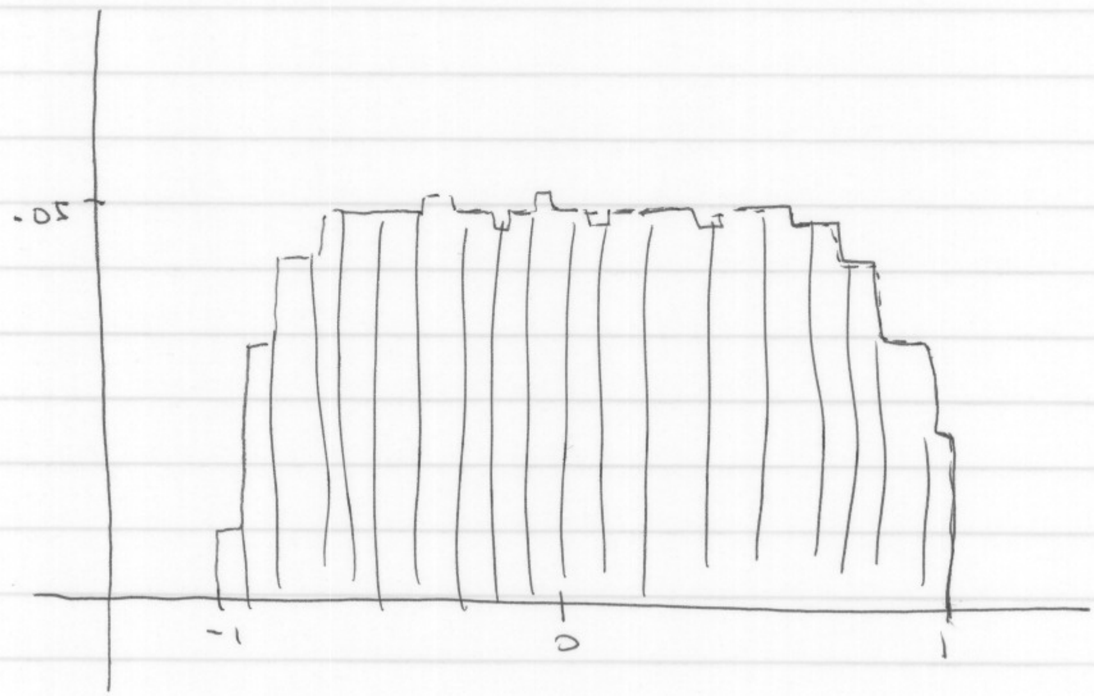
Spurious Regression problem:

Regression of an integrated series on another unrelated integrated series produces t-ratios on the slope parameter which indicate a relationship much more often that they should at the nominal test level. This problem does not dissapear as the sample size is increased





FREQUENCY DISTRIBUTION ~~FOR~~ FOR R^2 BETWEEN TWO IID processes



FREQUENCY DISTRIBUTION for R BETWEEN TWO I(1) PROCESSES WITH INDEPENDENT IID FIRST DIFFERENCES

Some Cures for Spurious Regression

- Spurious regression is a problem because we use the wrong set of critical values

I(1) case

$y_t = y_{t-1} + v_t \qquad v_t \sim iid(0, \sigma_v^2)$

$X_t = X_{t-1} + w_t \qquad w_t \sim iid(0, \sigma_w^2)$

$E(v_{t-s} w_{t-r}) = 0 \quad \forall\, r, s$

$y_t = \alpha + \beta X_t + u_t$

$$\hat{\beta} = \frac{\sum (y_t - \bar{y})(X_t - \bar{X})}{\sum (X_t - \bar{X})^2}$$

$* \; \boxed{T^{-\frac{1}{2}} t_{\hat{\beta}}} \Rightarrow$ Distribution (no nuisance parameters)

Why not to use (*)?

We can use as a pseudo t-stat $\frac{t_{\hat{\beta}}}{\sqrt{T}}$

but then if $u_t \sim I(0)$,

$\frac{t_{\hat{\beta}}}{\sqrt{T}} \xrightarrow{p} 0$ and we will be unable

to reject $\beta = 0$ even if it is not true. (inconsistency)

I(0) case

$$X_t = \phi_X X_{t-1} + \varepsilon_{Xt} \qquad \varepsilon_{Xt}, \varepsilon_{Yt} \sim iid($$
$$Y_t = \phi_Y Y_{t-1} + \varepsilon_{Yt}$$
$$E[\varepsilon_{Xt-s}, \varepsilon_{Yt-r}] = 0 \quad \forall r, s$$

$$Y_t = \alpha + \beta X_t + u_t$$

$$\hat{\beta} = \frac{\sum_1^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_1^T (X_t - \bar{X})^2}$$

$$\tilde{t}_{\hat{\beta}} = \frac{\hat{\beta}}{\tilde{\sigma}_{\hat{\beta}}} = \frac{\hat{\beta}}{\hat{\sigma}_u \Big/ \sqrt{\sum_1^T (X_t - \bar{X})^2}}$$

$$\tilde{\sigma}_{\hat{\beta}}^2 = \frac{\hat{\sigma}_u^2}{\sqrt{\sum_1^T (X_t - \bar{X})^2}} \quad \overset{P}{\not\longrightarrow} V(\hat{\beta}) \quad \text{and}$$

therefore $\tilde{t}_{\hat{\beta}} \overset{d}{\not\longrightarrow} N(0,1)$

Serial correlation in the disturbances requires a different form of consistent estimator for the standard error $\hat{\beta}$ as follows

$$\sigma^2_{\hat\beta} = M^{-1} V M^{-1}$$

where $M \equiv E\left[(x_t - \bar{x})^2\right]$, $V = var\left[\frac{1}{\sqrt{T}} \sum_1^T (x_t - \bar{x}) u_t\right]$

Its consistent estimator is

$$\hat\sigma^2_{\hat\beta} = \hat{M}^{-1} \hat{V} \hat{M}^{-1} \xrightarrow{p} \sigma^2_{\beta} \text{ as } T \to \infty$$

where

$$\hat{V} = \frac{1}{T} \sum_1^T (x_t - \bar{x})^2 \hat{u}_t^2 + \frac{2}{T} \sum_{s=1}^{\ell} w(s,\ell) \sum_{t=s+1}^{T} (x_t - \bar{x}) \hat{u}_t \hat{u}_{t-s} (x_{t-s} - \bar{x})$$

$$\hat{V} \xrightarrow{p} V$$

$$\hat{M} = \frac{1}{T} \sum_1^T (x_t - \bar{x}) \xrightarrow{p} M \quad \text{as } T \to \infty$$

and $w(s,\ell)$ is an optimal weighting function already discussed when we studied the LRV estimation.

In this case

$$t_{\hat\beta} = \frac{\hat\beta}{\hat\sigma_{\hat\beta}} \overset{d}{\sim} N(0,1)$$

Notice that

$$t_{\hat\beta} = \frac{\tilde\sigma_{\hat\beta}}{\hat\sigma_{\hat\beta}} \tilde{t}_{\hat\beta}.$$

- Think if something like that can be done in the I(1) case

Other approaches

- $$y_t = \alpha + \phi y_{t-1} + \gamma x_t + \delta x_{t-1} + u_t$$

Notice that there are values for the coefficients ($\phi = 1$, $\gamma = \delta = 0$) such that
$$u_t \sim I(0).$$

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\phi} \\ \hat{\gamma} \\ \hat{\delta} \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \alpha \\ \phi \\ \gamma \\ \delta \end{pmatrix} ; \quad \sqrt{T}(\hat{\gamma} - \gamma) \xrightarrow{d} N(0, V_1)$$

$$\sqrt{T}(\hat{\delta} - \delta) \sim N(0, V_2)$$

$$\boxed{t_{\gamma=0} \xrightarrow{d} N(0,1)} \text{ as } T \to \infty$$

$$F_{\binom{\gamma}{\delta}=0} \xrightarrow{d} \text{non-standard distribution}$$

- $$\Delta y_t = \beta \Delta x_t + z_t$$

$$t_{\beta=0} \xrightarrow{d} N(0,1)$$

But this regression doesn't provide the information we want

We want a regression in levels

$$\downarrow$$

Cointegration

Cointegration

<u>Def</u> : $\boxed{\underset{n\times 1}{Y_t} \sim I(1)}$ {non-stationary with a unit root}

but $\boxed{a' Y_t \sim I(0)}$ {we may say stationary}

<u>Example</u>:
$$\boxed{\begin{aligned} y_{1t} &= \alpha y_{2t} + u_{1t} \\ y_{2t} &= y_{2t-1} + u_{2t} \end{aligned}} \qquad \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \sim WN$$

《Why y_{1t} and y_{2t} are cointegrated?》

<u>Different Representation,</u>

- <u>Triangular System Representation</u> (TS)

$$y_{1t} = \alpha y_{2t} + u_{1t}$$
$$\Delta y_{2t} = u_{2t}$$

- <u>Moving average representation</u>

$$\Delta y_{1t} = \alpha u_{2t} + \Delta u_{1t}$$
$$\Delta y_{2t} = u_{2t}$$

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{bmatrix} (1-L) & \alpha \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

$$\Delta Y_t = C(L) u_t = C(1) u_t + (1-L) \tilde{C}(L) u_t$$

Note:

(i) $C(1) = \begin{bmatrix} 1-1 & \alpha \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \alpha \\ 0 & 1 \end{bmatrix} = \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix}$

has <u>not</u> full rank

(ii) $a' C(1) = \begin{pmatrix} 1 & -\alpha \end{pmatrix} \begin{pmatrix} 0 & \alpha \\ 0 & 1 \end{pmatrix} = 0$

- error correction term (ECM)

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} u_{1t-1} + \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

ERROR TERM
FROM THE COINTEGRATING RELATIONSHIP

$$= \begin{pmatrix} -1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & -\alpha \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$\boxed{\underset{2\times1}{\Delta Y_t} = \underset{2\times1}{\gamma} \underset{1\times2}{\alpha'} Y_{t-1} + \varepsilon_t}$$ VAR ECM

Cointegration and economic theory

Example 1: THEORY OF PURCHASING POWER PARITY

"Apart from transportation costs, goods should sell for the same effective price in two countries"

$$\boxed{P_t = S_t \; P_t^*}$$

An index of the price level in USA

dolars per peseta

Price Index for Spain

$p_t = s_t + p_t^*$ in logs

A weaker version of PPP:

$$\boxed{p_t = s_t + p_t^* + z_t}$$

Example 2: PRESENT VALUE MODELS

$$\boxed{Y_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t y_{t+i} + c} \quad (*)$$

- Long-term yields — Short-term yields
- Stock Prices — Dividends
- Consumption — Labor Income

If y_t has a unit root and $(*)$ holds

then Y_t and y_t will be Cointegrated

$$\boxed{S_t = Y_t - \theta y_t \sim I(0)}$$ Campbell-Shiller

It can be proved that

$\left\langle\left\langle Y_t \text{ and } \frac{S_t}{1-L} \right\rangle\right\rangle$ are cointegrated too
(J. Gonzalo)
???

So Y_t and y_t are multicointegrated

Implications of Cointegration

- **For the MA**

$$\underset{p\times 1}{\Delta Y_t} = C(L)\,\varepsilon_t = C(1)\,\varepsilon_t + (1-L)\,\tilde{C}(L)\,\varepsilon_t$$

(*) $C(1)$ reduced rank $\quad \underset{r\times p}{\alpha'}\ \underset{p\times p}{C(1)} = 0$

(*) $C(L)$ is not invertible

(*) $C(1) = \underset{p\times m}{A_1}\ \underset{m\times p}{B_1} \qquad m = p-r$

Common Trend Representation (Stock-Watson)

$$\Delta Y_t = A_1 B_1 \varepsilon_t + (1-L)\,\tilde{C}(L)\,\varepsilon_t$$

$$Y_t = A_1 \frac{\eta_t}{(1-L)} + \frac{\cancel{(1-L)}}{\cancel{(1-L)}}\,\tilde{C}(L)\,\varepsilon_t$$

$$= \underset{p\times m}{A_1}\ \underset{m\times 1}{f_t} + \tilde{C}(L)\,\varepsilon_t$$

Trends ??? $\qquad (1-L) f_t = \eta_t$

Common ??? $\qquad m < p$

QUESTION:

Can α' have full rank ??

- <u>For the VAR</u>

We do not have a finite VAR for ΔY_t

Suppose

$$\boxed{\Phi(L)\,Y_t = \varepsilon_t}\ ;\ Y_t = \phi_1 Y_{t-1} - \cdots \phi_k Y_{t-k} + \varepsilon_t$$

From the Wold's representation

$$(1-L)\,Y_t = C(L)\,\varepsilon_t$$

So

$$(1-L)\,\phi(L)\,Y_t = \phi(L)\,C(L)\,\varepsilon_t$$

$$(1-L)\,\varepsilon_t = \phi(L)\,C(L)\,\varepsilon_t$$

Therefore

$$\boxed{\phi(1)\,C(1) = 0}$$

rank $\phi(1) = \underline{r}$ rank$(C(1)) = m = p-r$

$$\boxed{C(1)\,\phi(1) = 0}$$

We can decompose $\phi(1) = AB$ such that

$$\boxed{B\,C(1) = 0} \Rightarrow ??$$
$$\boxed{C(1)\,A = 0}$$

From the VAR $\boxed{\Phi(L) Y_t = \varepsilon_t}$ how do we get the ECM ??

$$\Phi(L) = \Phi(1) + (1-L)\tilde{\Phi}(L)$$

$$= \Phi(1) + \Phi(1)L - \Phi(1)L + (1-L)\tilde{\Phi}(L)$$

$$= \Phi(1)L + (1-L)\Phi(1) + (1-L)\tilde{\Phi}(L)$$

$$= \Phi(1)L + (1-L)\underbrace{\left(\tilde{\Phi}(L) + \Phi(1)\right)}_{\Gamma(L)}$$

Note that

$$\Phi(0) = \Phi(1) + \tilde{\Phi}(0) \Rightarrow \tilde{\Phi}(0) = \underset{\overset{\shortparallel}{I}}{\Phi(0)} - \Phi(1)$$

So $\Gamma(0) = I$

Then

$$\Phi(L) Y_t = \varepsilon_t$$

can be written as

$$(1-L) Y_t = -\Phi(1) Y_{t-1} + (1-L)\left(\Gamma(L) - I\right) Y_t + \varepsilon_t$$

$$= -AB\, Y_{t-1} + \text{LAGS OF } \Delta Y_t + \varepsilon_t$$

Testing for cointegration

For simplicity assume at most one cointegrating vector

- a' is known

$a' Y_t = Z_t$

D-F test on Z_t $\left\{ \begin{array}{l} \Delta Z_t = \rho Z_{t-1} + \text{Lags}\, \Delta Z_t + e_t \\ t_{\hat{\rho}} \sim \text{D-F Distribution} \end{array} \right.$

Ho: no-cointegration ($\rho = 0$)
HA: cointegration

- a' is unknown so it has to be estimated

$\hat{a}' Y_t = \hat{Z}_t$

D-F test on $\hat{Z}_t$ $\left\{ \begin{array}{l} \Delta \hat{Z}_t = \rho \hat{Z}_{t-1} + \text{lags}\, \Delta \hat{Z}_t + e_t \\ t_{\hat{\rho}} \nsim \text{D-F distribution} \end{array} \right.$

We will see other tests that are valid even for $r \geq 1$.

Estimation of the Cointegrating Vector

Proposition 19.2 (Hamilton)

Suppose (bivariate case for simplicity)

$$y_{1t} = \alpha y_{2t} + z_t^*$$
$$\Delta y_{2t} = u_{2t}$$

and $\underset{2\times 1}{\begin{pmatrix} z_t^* \\ u_{2t} \end{pmatrix}} = \underset{2\times 2}{\psi^*(L)} \underset{2\times 1}{\varepsilon_t}$

where ε_t is iid, mean zero, finite fourth moments and $E(\varepsilon_t \varepsilon_t') = PP'$. Further $\sum_{s=0}^{\infty} |s\,\psi_s^*| < \infty$.

Under general conditions

$$\underset{2\times 1}{v_t} = \begin{bmatrix} z_t^* \\ u_t \end{bmatrix} ; \quad X_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} v_t \Rightarrow B(r)$$

where $B = \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix} \equiv BM(\Omega)$

and
$$\Omega = \psi^*(1) PP' \psi^*(1)'.$$

Remember that in the univariate case the LRV of $\underset{1\times 1}{v_t}$ is

$$\lim_{T\to\infty} Var\left(\frac{1}{\sqrt{T}} \sum_1^T v_t\right) = \sigma^2 + 2\lambda$$

where $\lambda = \sum_{k=1}^{T} E(v_t v_{t-k})$

Now in the multivariate case

$$\Omega = \lim_{T\to\infty} Var\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T v_t\right)$$
$$= \lim_{T\to\infty} \frac{1}{T} E\left(\sum_1^T v_t \sum_1^T v_t'\right)$$
$$= \lim_{T\to\infty} \frac{1}{T} E\left(S_T \, S_T'\right) \quad \text{with } S_T = \sum_{t=1}^{T} v_j$$
$$= \Sigma + \Lambda + \Lambda'$$

with

$$\Sigma_1 = \lim_{T \to \infty} \frac{1}{T} \sum_{1}^{T} E(v_t v_t')$$

$$\Lambda = \lim_{T \to \infty} \frac{1}{T} \sum_{1}^{T} E(S_{t-1} v_t')$$

$$= \lim_{T \to \infty} \frac{1}{T} \sum_{t=2}^{T} \sum_{j=1}^{t-1} E(v_j v_t')$$

If v_t is weakly stationary then

$$\Omega = \underbrace{E[v_1 v_1']}_{=\Sigma_1} + \underbrace{\sum_{k=2}^{\infty} E(v_1 v_k')}_{=\Lambda} + \underbrace{\sum_{k=2}^{\infty} E(v_k v_1')}_{=\Lambda'}$$

$$\Sigma_1 = \begin{bmatrix} E(z_t^*)^2 & E(z_t^* u_{2t}) \\ E(u_{2t} z_t^*) & E(u_{2t}^2) \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \sum_{k=1}^{\infty} E(z_{t-k} z_t) & \sum_{k=1}^{\infty} E(z_{t-k} u_t) \\ \sum_{k=1}^{\infty} E(z_t u_{t-k}) & \sum_{k=1}^{\infty} E(u_t z_{t-k}) \end{bmatrix}$$

To get the asymptotic distribution of $\hat{\alpha}$ notice that

$$\frac{1}{T^2}\sum_{1}^{T} y_{2t}^2 \Rightarrow \int B_2^2(r)\,dr$$

$$\frac{1}{T}\sum_{1}^{T} y_{2t} z_t^* \Rightarrow \int B_2\, dB_1 + M$$

where $M = \sum_{j=0}^{\infty} E(u_{2t} z^*_{t+j}) = \Lambda_{[2,1]} + E(u_{2t} z_t^*)$

Remember that in the univariate case $y_t = \alpha y_{t-1} + u_t$ we had

$$\frac{1}{T}\sum_{t=2} y_{t-1} u_t \Rightarrow \int B\, dB + \frac{1}{2}(w^2 - w_0^2)$$

with $w^2 = w_0^2 + 2\sum_{j=1}^{\infty} E(u_t u_{t+j})$

Regress $B_1(r)$ on $B_2(r)$: independent BM

$$B_1(r) = \Omega_{22}^{-1}\Omega_{12} B_2(r) + \Omega_{1.2} L(r)$$

with $\Omega_{1.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$

substituting above we have

$$\int B_2\, dB_1 + M = \Omega_{22}^{-1}\Omega_{12}\int B_2\, dB_2 + \Omega_{1.2}\int B_2\, dL(r) + M$$

$$= \Omega_{1.2}\int B_2\, dL(r) + \Omega_{22}^{-1}\Omega_{12}\int B_2\, dB_2 + M$$

So

$$T(\hat{\alpha}-\alpha) \Rightarrow \left(\int B_2^2(r)\,dr\right)^{-1}\left(\int B_2\,dB_1 + M\right)$$

$$= \left(\int B_2^2(r)\,dr\right)^{-1}\left[\Omega_{1.2}\int B_2\,dL(r) + \Omega_{22}^{-1}\Omega_{12}\int B_2\,dB_2 + M\right]$$

UNIT ROOT PAR

simultaneous Equation part

Note that if Z_t^* and U_{2t} are not correlated (leads and lags) ($\Omega_{12} = 0 = M$) then the A.D simplifies to

$$T(\hat{\alpha}-\alpha) \Rightarrow \left(\int B_2^2\,dr\right)^{-1}\Omega_{11}\int B_2\,dL(r) \equiv N\left(0, \left(\int B_2^2\,dr\right)^{-1}\Omega_{11}^2\right)$$

In this case we can run the usual tests and they will have the standard distributions.

Let F_2 denote the sub σ-field of $\mathbf{F}$ generated by $\{B_2(r) : 0 \le r \le 1\}$. Since $B(r)$ is Gaussian and independent of $B_2(r)$

$$\int B_2(r)\, dL(r) \Big|_{F_2} \equiv N\left(0, \int B_2^2\, dr\right)$$

and

$$\left[\int_0^1 B_2^2\, dr\right]^{-\frac{1}{2}} \left[\int_0^1 B_2\, dL\right]\Big|_{F_2} \equiv N(0, I)$$

Since the latter distribution does not depend on realizations of $B_2(r)$, it is also the unconditional distribution.

Summarizing

$$T(\hat{\alpha}-\alpha) \Longrightarrow \left[\int B_2^2(r)\,dr\right]^{-1}\left[\underbrace{\square\int B_2(r)\,dL(r)}_{A1} + \underbrace{\square\int B_2\,dB_2}_{A2} + \underbrace{M}_{A3}\right]$$

UNIT ROOT

SIMULTANEOUS EQUATION BIAS

These two terms create a <u>bias</u> in finite sample.

Sometimes more important than the "<u>bias</u>" problem is the "<u>inference</u>" problem.

These two problems disappear when Z_t^* and u_{2t} are uncorrelated at all leads and lags.

ii What to do if Z_t^* and u_t are <u>not</u> uncorrelated??

Find a way of eliminating $\boxed{A2}$ and $\boxed{A3}$

One possible solution:

(a) Correct for correlation between Z_t^* and u_{2t} by adding leads and lags of $\Delta y_{2t} = u_{2t}$

[Saikkonen, Stock-Watson, Phillips-Loretan ...]

Define $\tilde{Z}_t$ as the residuals of

$$Z_t^* = \sum_{s=-p}^{p} \beta'_s u_{2t-s} + \tilde{Z}_t$$

[Projection of Z_t^* on $(u_{2t-p}, \ldots u_{2t-1}, u_{2t}, u_{2t+1} \ldots u_{2t+p})$]

It is clear that $\tilde{Z}_t$ is uncorrelated with u_{2t-s} for $s = -p, -p+1, \ldots, p$.

Therefore instead of estimating α in

$$\boxed{y_{1t} = \alpha' y_{2t} + Z_t^*}$$

we can do it in

$$\boxed{y_{1t} = \alpha' y_{2t} + \sum_{s=-p}^{p} \beta'_s \Delta y_{2t-s} + \tilde{Z}_t}$$

Note that $\tilde{Z}_t$ can be correlated. This problem is easily solved.

Example consumption - Income (1948:2 - 1988:3)

$$C_t = \text{constant} + \underset{(.00306)}{.99216}\, y_t + \square\, \Delta y_{t+4} + \square\, \Delta y_{t+3} \ldots + \square\, \Delta y_t + \square\, \Delta y_{t-1} + \cdots + \square\, \Delta y_{t-4} + \hat{u}_t$$

$$s_T^2 = (T-11)^{-1} \sum_{t=1}^{T} \hat{u}_T^2 = (1.516)^2$$

We want to test that $\underline{a = (1, -1)}$

$$t = \frac{(.99216 - 1)}{.00306} = \boxed{-2.562} \quad \text{then what??}$$

This t-statistic has been calculated using the short-run variance (s_T^2) of $\hat{u}_t$. But $\hat{u}_t$ may be correlated. If this is the case the statistic that follows a $\underline{t}$ distribution is

$$\frac{s_T}{\tilde{\lambda}_{11}} \underbrace{\boxed{\frac{(\hat{\alpha} - 1)}{s_T (X'X)^{-1}}}}_{= \; t} = \frac{s_T}{\tilde{\lambda}_{11}}\, t \sim \underline{t\text{-distributi}}$$

Long-run variance of $\hat{u}_t$

• This argument is only applicable to the coefficient of the I(1) regressor

To find the long-run variance of $\hat{u}_t$:

$$\hat{u}_t = .718\,\hat{u}_{t-1} + .2057\,\hat{u}_{t-2} + \hat{e}_t$$

where

$$\hat{\sigma}_1^2 = (T-2)^{-1} \sum_{t=3}^{T} \hat{e}_t^2 = .38$$

Thus the estimate of $\tilde{\lambda}_{11}$ is

$$\tilde{\lambda}_{11} = \frac{(.38)^{\frac{1}{2}}}{(1-.718-.2057)} = \underline{\underline{8.084}}$$

Hence a test of the null hypothesis that $a = (1, -1)$ can be based on

$$t\left(S_T / \hat{\lambda}_{11}\right) = (-2.562)\,\frac{1.516}{8.084} = \boxed{-.48}$$

then what??

another possible solution is

(b) Maximum Likelihood in an ECM
(Johansen, Ahn-Reinsel)

$$\underset{p\times 1}{\Delta X_t} = \Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_k \Delta X_{t-k} + e_t$$

$$\Pi = \boxed{\underset{p\times r}{\gamma}\ \underset{r\times p}{\alpha'}} \qquad e_t \sim N(0, \Omega)$$

$$L\left(\Gamma_1\, \Gamma_2 \ldots \Gamma_k, \Pi, \Omega \,\middle|\, (X_1 \ldots X_T)\right) = \frac{Tp}{2}\log(2\pi) - \frac{T}{2}|\Omega| - \frac{1}{2}\sum_{t=1}^{T} e_t' \Omega^{-1} e_t$$

We want $\boxed{\begin{array}{l} \text{MAX } L(\ \mid\ ) \\ \text{s.t. } \quad \Pi = \underset{p\times r}{\gamma}\ \underset{r\times p}{\alpha'} \end{array}}$

STEPS

(i) Concentrate L with respect $\Gamma_1 \ldots \Gamma_k$

ΔX_t on $\Delta X_{t-1} \ldots \Delta X_{t-k}$ + $\boxed{R_{0t}}$ RESIDUALS

X_{t-1} on ΔX_{t-1} ΔX_{t-k} + $\boxed{R_{1t}}$ Residuals

$$L(\ ) = -\frac{Tp}{2}\log(2\pi) - \frac{T}{2}\log|\Omega| - \frac{1}{2}\sum_{t=1}^{T}(R_{0t} - \Pi R_{1t})'\Omega^{-1}(\quad)$$

(ii) Concentrate $L(\ )$ with respect Ω

$$L(\gamma\alpha) = \underbrace{\boxed{-\frac{Tp}{2}\log(2\pi) - \frac{Tp}{2}}}_{\text{constant}} - \frac{T}{2}\log\left|\left(\frac{1}{T}\right)\sum_{1}^{T}(R_{0t} - \gamma\alpha' R_{1t})'(R_{0t} - \gamma\alpha' R_{1t})\right|$$

$$\text{MAX } L(\ ) \iff \text{Minimize } \left|\frac{1}{T}\sum_{1}^{T}(R_{0t} - \gamma\alpha' R_{1t})'(R_{0t} - \gamma\alpha' R_{1t})\right|$$

Define $S_{ij} = T^{-1}\sum_{t=1}^{T} R_{it}R_{jt}'$, $i, j = 0, 1$.

$$L(\gamma,\alpha) = K_0 - \frac{T}{2}\log\left|S_{00} - \Pi S_{10} - S_{01}\Pi' + \Pi S_{11}\Pi'\right|$$

(iii) Now concentrate $L(\ )$ with respect γ

$$\frac{\partial L(\gamma,\alpha)}{\partial \gamma} = 0 \implies \boxed{\hat{\gamma} = S_{01}\alpha(\alpha' S_{11}\alpha)^{-1}}$$

Substituting

$$L(\alpha) = K_1 - \frac{T}{2}\log\left|S_{00} - S_{01}\alpha(\alpha' S_{11}\alpha)^{-1}\alpha' S_{10}\right|$$

Consider the following matrices with eigenvalues and eigenvectors

$$(d_i S_{11} - S_{10} S_{00}^{-1} S_{01}) v_i = 0 \qquad \text{s.t. } V' S_{11} V = I$$
$$i = 1, 2, \ldots p$$

$$(d_i S_{00} - S_{01} S_{11}^{-1} S_{10}) m_i = 0 \qquad \text{s.t. } m' S_{00} m = I$$
$$i = 1, 2, \ldots p$$

where $d_1 > d_2 > \ldots > d_r > d_{r+1} > \ldots d_p$.

It can be proved that

$$\boxed{\alpha' = \begin{pmatrix} v_1' \\ \vdots \\ v_r' \end{pmatrix} \quad \text{The first } \underline{r} \text{ eigenvectors of } \underline{S_{11}^{-1} S_{10} S_{00}^{-1} S_{01}}}$$

and

$$\boxed{L(\;) = K_2 - T/2 \sum_{i=1}^{r} \log(1 - d_i)}$$

How do we find r?

What about Π_i?

So far we have assumed we know $\underline{r}$.

- **How do we estimate $\underline{r}$?**

H_0: rank of cointegration $= \underline{r}$

$$L_0 = K_2 - \frac{T}{2} \sum_{i=1}^{r} \log(1-\hat{\lambda}_i)$$

H_A: rank of cointegration $= p$ (unconstrained)

$$L_A = K_2 - \frac{T}{2} \sum_{i=1}^{p} \log(1-\hat{\lambda}_i)$$

$$L_A - L_0 = \left(-T/2\right) \sum_{i=r+1}^{p} \log(1-\hat{\lambda}_i)$$

$$2(\log L_A - \log L_0) = \boxed{-T \sum_{i=r+1}^{p} \log(1-\hat{\lambda}_i)}$$

$$\Downarrow$$

$$\text{Trace}\left[\int_0^1 W(s)\,dW(s)'\right]'\left[\int_0^1 W(s)\,W(s)'\,ds\right]^{-1}\left[\int_0^1 W(s)\,dW(s)'\right]$$

Why we should care about cointegration?

- Testing Economic Theory

  Once it is assumed that $\underset{p\times 1}{X_t} \sim I(1)$, any economic theory that relates the variables in $\underset{p\times 1}{X_t}$ implies cointegration

- NO VARs in first differences

- Forecasting

  The model $\Delta X_t = \gamma \alpha' X_{t-1} + \text{lags}\ \Delta X_t + \varepsilon_t$

  should forecast better than

  $$\Delta X_t = \Pi X_{t-1} + \text{lags}\ \Delta X_t + \varepsilon_t$$

- Common factor Representation

- Shocks' identification

- - - - -