Lectures on

The cointegrated vector autoregressive model

by Søren Johansen
Department of Economics
University of Copenhagen
and
CREATES, Aarhus University

Institute for the Economy in Transition

Moscow April 7-9, 2008

LECTURE 1

The vector autoregressive model, its solution, and interpretation

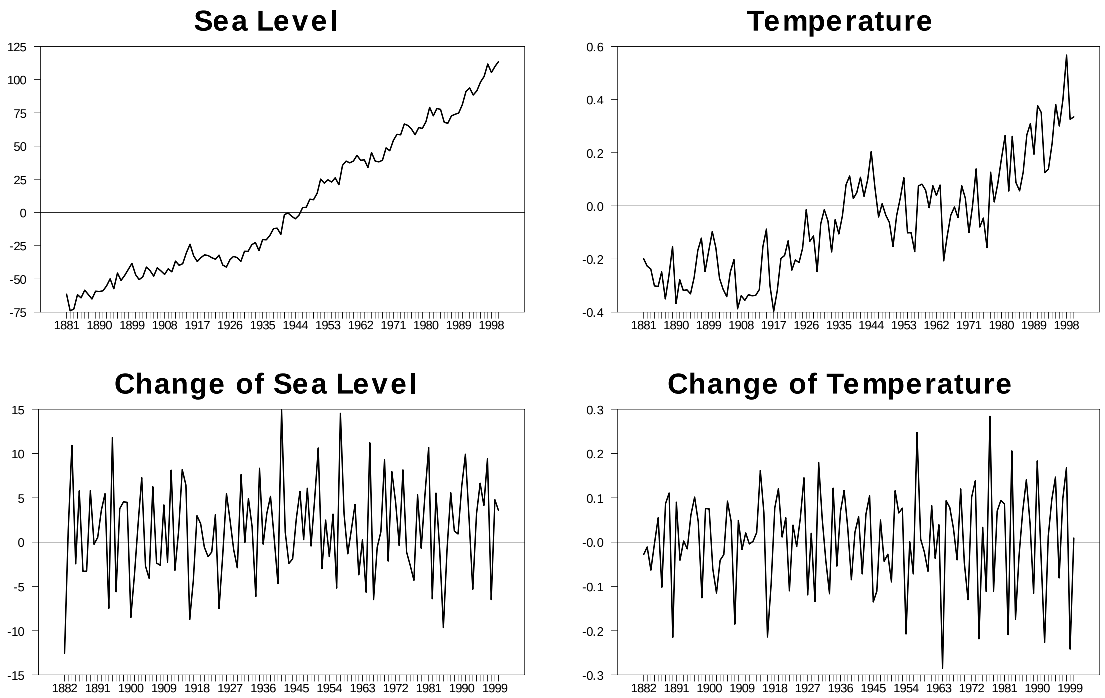
1. Some Examples of Integration and Cointegration
2. General Statistical Methodology
3. The Dynamic Properties of the VAR, Integration and Cointegration
4. Granger Representation Theorem
5. The Role of the Deterministic Terms
6. Interpretation of Cointegrating Coefficients
7. Shocks, Changes and Impulse Responses
8. Hypotheses of interest
9. Conclusions

1. Some Examples of Integration and Cointegration

1a: The global average temperature and sea level

Data : 1881:01 to 1995:01

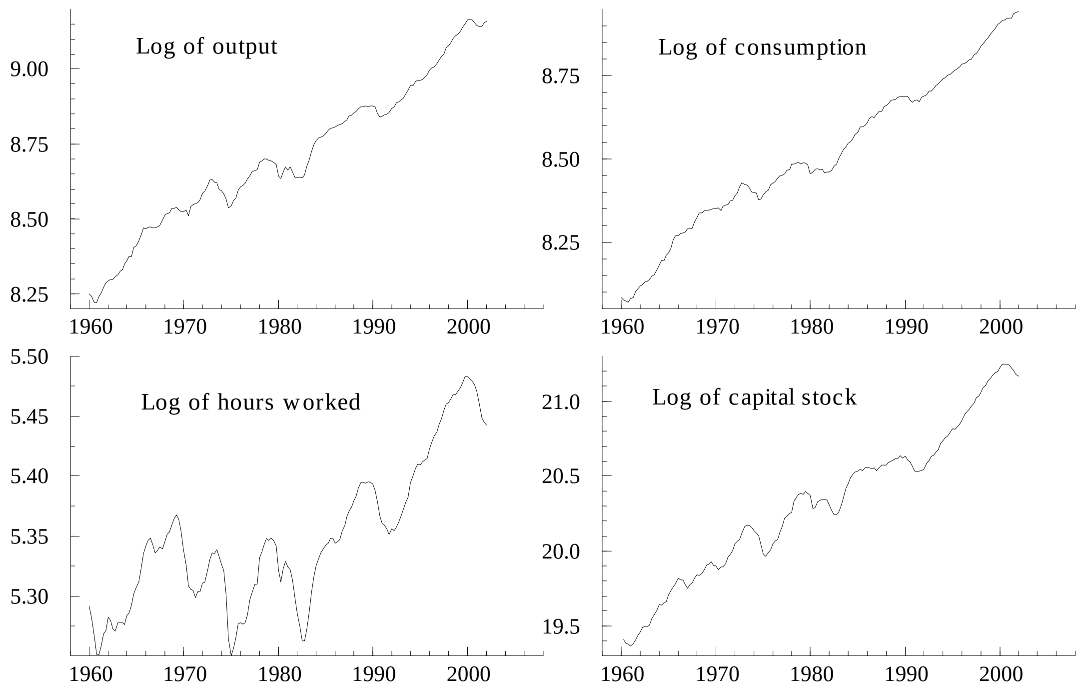
Hansen, J. et al *J. Geophys. Res. Atmos.* **106** 23947 (2001)

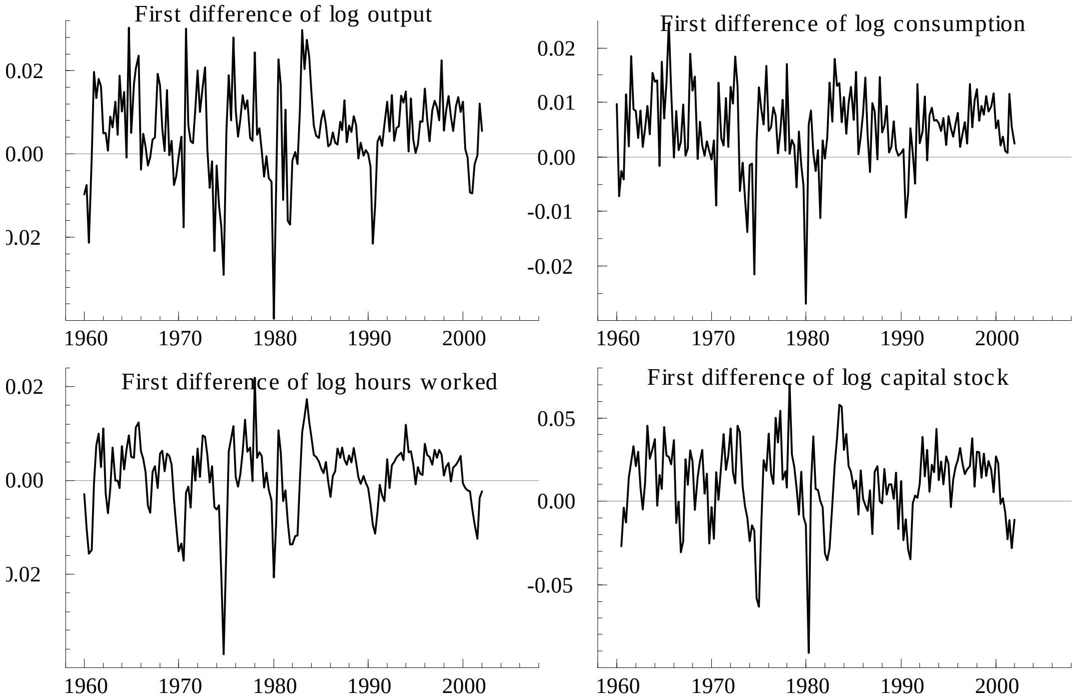

Sea Level
125
100
75
50
25
0
-25
-50
-75
1881 1890 1899 1908 1917 1926 1935 1944 1953 1962 1971 1980 1989 1998
Temperature
0.6
0.4
0.2
0.0
-0.2
-0.4
1881 1890 1899 1908 1917 1926 1935 1944 1953 1962 1971 1980 1989 1998
Change of Sea Level
15
10
5
0
-5
-10
-15
1882 1891 1900 1909 1918 1927 1936 1945 1954 1963 1972 1981 1990 1999
Change of Temperature
0.3
0.2
0.1
-0.0
-0.1
-0.2
-0.3
1882 1891 1900 1909 1918 1927 1936 1945 1954 1963 1972 1981 1990 1999

5.

1b: US data of consumption, income, investment, hours worked and capital

The data 1948:1 to 2002:2 from
Federal Reserve Bank of St. Louis' FRED database and
Bureau of Labor Statistics' Establishment Survey.
$N_t =$ Civilian, non-institutional population, age 16 and over.
$C_t =$ Real Personal Consumption Expenditures in chained 1996 dollars/N_t
$I_t =$ Real Gross Private Domestic Investment in chained 1996 dollars$/N_t$
$H_t =$ Hours of wage and salary workers on private, non-farm payrolls$/N_t$.
$Y_t = I_t + C_t$
$K_t =$ Capital Stock Formation$/N_t$
$x_t = (\log Y_t, \log C_t, \log K_t, \log H_t) = (y_t, c_t, k_t, h_t)$


Log of output
9.00
8.75
8.50
8.25
1960 1970 1980 1990 2000
Log of consumption
8.75
8.50
8.25
1960 1970 1980 1990 2000
Log of hours worked
5.50
5.45
5.40
5.35
5.30
1960 1970 1980 1990 2000
Log of capital stock
21.0
20.5
20.0
19.5
1960 1970 1980 1990 2000


First difference of log output
0.02
0.00
0.02
1960
1970
1980
1990
2000
First difference of log consumption
0.02
0.01
0.00
-0.01
-0.02
First difference of log hours worked
First difference of log capital stock
0.05
-0.05

2. General Statistical Methodology

Main question

WHICH STATISTICAL MODEL
DESCRIBES THE DATA?

General methodology is likelihood based inference

What is gained:

1. A coherent framework for formulating and testing the economic hypotheses of interest
2. The Gaussian likelihood is used for deriving test and estimators but their properties are derived under more general assumptions.

In particular:

3. Asymptotic distributions of maximum likelihood estimators
4. Consistent estimates of 'variances'
5. Asymptotic distribution of likelihood ratio statistics

The price paid:

One should carefully check the assumptions underlying the model, otherwise inference can be invalid

3. The Dynamic Properties of the VAR, Integration and Cointegration

The vector autoregressive model

$$x_t = \Pi_1 x_{t-1} + \ldots + \Pi_k x_{t-k} + \Phi D_t + \varepsilon_t,$$

where $t = 1, \ldots, T$ and ε_t i.i.d. $N_p(0, \Omega)$.

By recursive substitution the equations define x_t as function of

1. Initial values, $x_0, \ldots, x_{-k+1}$,
2. Errors $\varepsilon_1, \ldots, \varepsilon_t$,
3. Deterministic terms $D_1, \ldots, D_t$, (constant, linear term, seasonal and intervention dummies)
4. Parameters $(\Pi_1, \ldots, \Pi_k, \Phi)$

Example

Let $\varepsilon_t, t = 1, 2, \ldots$ be i.i.d. $(0, \sigma^2)$

$$\begin{aligned} x_t &= \rho x_{t-1} + \varepsilon_t \\ x_t &= \rho^t x_0 + \rho^{t-1}\varepsilon_1 + \ldots + \rho\varepsilon_{t-1} + \varepsilon_t \end{aligned}$$

If $|\rho| < 1$, then $x_t = \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i}$ stationary

If $\rho = 1$, then $x_t = x_0 + \sum_{i=1}^{t} \varepsilon_i$ random walk

$$x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \rho x_{1t-1} + \varepsilon_{1t} \\ x_{2t} + \varepsilon_{2t} \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{\infty} \rho^i \varepsilon_{1t-1} \\ x_{2t} + \sum_{i=1}^{t} \varepsilon_{2i} \end{pmatrix}$$

Example

$$x_t \;=\; \Pi_1 x_{t-1} + \Pi_2 x_{t-2} + \Phi D_t + \varepsilon_t$$ two-dimensional, ε_t i.i.d.(0,Ω)

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \Pi_{1.11} & \Pi_{1.12} \\ \Pi_{1.21} & \Pi_{1.22} \end{pmatrix} \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \end{pmatrix} + \begin{pmatrix} \Pi_{2.11} & \Pi_{2.12} \\ \Pi_{2.21} & \Pi_{2.22} \end{pmatrix} \begin{pmatrix} x_{1t-2} \\ x_{2t-2} \end{pmatrix} + \Phi D_t + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Characteristic polynomial

$$\Pi(z) = I_2 - \Pi_1 z - \Pi_2 z^2 = \begin{pmatrix} 1 - \Pi_{1.11} z - \Pi_{2.11} z^2 & -\Pi_{1.12} z - \Pi_{2.12} z^2 \\ -\Pi_{1.21} z - \Pi_{2.21} z^2 & 1 - \Pi_{1.22} z - \Pi_{2.22} z^2 \end{pmatrix}$$

with determinant $\det(\Pi(z)) = |\Pi(z)|$ of degree four, and inverse matrix

$$\Pi^{-1}(z) = \frac{adj(\Pi(z))}{\det(\Pi(z))}, z \neq \rho_i^{-1}$$

$$\det(\Pi(z)) = (1 - z\rho_1)(1 - z\rho_2)(1 - z\rho_3)(1 - z\rho_4)$$

The roots of $\det(\Pi(z)) = 0$ are $1/\rho_i$

Theorem: *Consider the model*

$$x_t = \Pi_1 x_{t-1} + \ldots + \Pi_k x_{t-k} + \Phi D_t + \varepsilon_t.$$

If $|\rho_i| < 1,$ the coefficients of $\Pi^{-1}(z) = \sum_{i=0}^{\infty} C_i z^i$ are exponentially decreasing and the process x_t is stationary with moving average representation

$$x_t = \sum_{i=0}^{\infty} C_i(\varepsilon_{t-i} + \Phi D_{t-i})$$

Example

$$\begin{aligned} x_t &= \rho x_{t-1} + \mu + \varepsilon_t \\ \pi(z) &= 1 - \rho z, \quad \pi(z)^{-1} = \sum_{i=0}^{\infty} \rho^i z^i \end{aligned}$$

If $|\rho| < 1,$ then $\rho^i \rightarrow 0$ and we use the coefficients to construct a stationary process

$$x_t = \sum_{i=0}^{\infty} \rho^i(\varepsilon_{t-i} + \mu) = \sum_{i=0}^{\infty} \rho^i \varepsilon_{t-i} + \frac{\mu}{1-\rho}$$

Integration and Cointegration

Definition 1. x_t integrated of order 1, $I(1)$, if $\Delta x_t = \sum_{i=0}^{\infty} C_i \varepsilon_{t-i}$, and $\sum_{i=0}^{\infty} C_i \neq 0$

Definition 2. If x_t is $I(1)$, and $\beta' x_t$ is stationary, then x_t is cointegrated with cointegration vector β.

Example

$$|\rho| < 1, x_{20} = 0,\ x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \rho x_{1t-1} + \varepsilon_{1t} \\ x_{2t} + \varepsilon_{2t} \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{\infty} \rho^i \varepsilon_{1t-i} \\ \sum_{i=1}^{t} \varepsilon_{2i} \end{pmatrix}$$

$$y_t = \begin{pmatrix} \frac{1}{2} x_{1t} + \frac{1}{2} x_{2t} \\ -\frac{1}{2} x_{1t} + \frac{1}{2} x_{2t} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \sum_{i=0}^{\infty} \rho^i \varepsilon_{1t-i} + \frac{1}{2} \sum_{i=1}^{t} \varepsilon_{2i} \\ -\frac{1}{2} \sum_{i=0}^{\infty} \rho^i \varepsilon_{1t-i} + \frac{1}{2} \sum_{i=1}^{t} \varepsilon_{2i} \end{pmatrix} \text{ both } I(1)$$

But $y_{1t} - y_{2t} (= x_{1t})$ is stationary and $y_{1t} + y_{2t} (= x_{2t})$ is random walk.

Thus y_t is cointegrated with $\beta' = (1, -1)$ and common trend $\sum_{i=1}^{t} \varepsilon_{2i}$

Another example

$$
\begin{aligned}
x_{1t} &= a\sum_{i=1}^{t}\varepsilon_{1i} + \varepsilon_{2t} \sim I(1) \\
x_{2t} &= b\sum_{i=1}^{t}\varepsilon_{1i} + \varepsilon_{3t} \sim I(1) \\
bx_{1t} - ax_{2t} &= b\varepsilon_{2t} - a\varepsilon_{3t} \sim I(0)
\end{aligned}
$$

$x(t)$ is $I(1)$ and cointegrated with $\beta = (b, -a)'$ and common trend $\sum_{i=1}^{t}\varepsilon_{1i}$.

Granger's result: If $\Delta x_t = \sum_{i=0}^{\infty} C_i\varepsilon_{t-i}$ then $C(z) = C(1) + (1-z)C^*(z)$ and

$$
x_t = C(1)\sum_{i=1}^{t}\varepsilon_i + \sum_{i=0}^{\infty} C_i^*\varepsilon_{t-i}
$$

Reduced rank of $C(1) = \gamma\eta'$ implies the common trends $\eta'\sum_{i=1}^{t}\varepsilon_i$ and hence cointegration $\gamma_\perp' x_t$. Inverting $C(z)^{-1}$ we find an infinite order autoregressive model with unit roots.

Some history.

The Error Correction Model (ECM)

$$\begin{aligned} AR &: \; x_t = \Pi_1 x_{t-1} + \Pi_2 x_{t-2} + \varepsilon_t \\ x_t - x_{t-1} &= (\Pi_1 + \Pi_2 - I_p) x_{t-1} + \Pi_2 (x_{t-2} - x_{t-1}) + \varepsilon_t \\ ECM &: \; \Delta x_t = \Pi x_{t-1} + \Gamma_1 \Delta x_{t-1} + \varepsilon_t \end{aligned}$$

Note that

$$\Pi(z) = I_p - z\Pi_1 - z^2 \Pi_2 = (1-z) I_p - \Pi z - \Gamma_1 z (1-z)$$

If $\Pi(z)$ has unit roots, then

$$\Pi(1) = -\Pi = -\alpha\beta',$$

for some α and β of dimension $p \times r$ and rank $r < p$

Error Correction Model:

$$\Delta x_t = \alpha\beta' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \varepsilon_t$$

4. Granger Representation Theorem

Question: If the VAR has unit roots and the other roots are larger than one, what is the moving average representation?

Unit roots imply reduced rank of $\Pi(1)$:

$$\begin{aligned}\Delta x_t &= \alpha\beta' x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \varepsilon_t \\ \Pi(z) &= (1-z)I_p - \alpha\beta' z - \sum_{i=1}^{k-1} (1-z) z^i \Gamma_i\end{aligned}$$

$I(1)$ condition avoids $I(2)$ processes: The roots satisfy $z = 1$ or $|z| > 1$ and

$$\det(\alpha'_\perp (I_p - \sum_{i=1}^{k-1} \Gamma_i)\beta_\perp) \neq 0$$

Theorem: *If the roots of* $\det \Pi(z) = 0$ *are either* $|z| > 1$ *or* $z = 1$, *and* $\det(\alpha'_\perp(I_p - \sum_{i=1}^{k-1} \Gamma_i)\beta_\perp) \neq 0$ *is satisfied then*

$$\Pi(z)^{-1} = C\frac{1}{1-z} + \sum_{i=0}^{\infty} C_i^* z^i, \; |C_i^*| \leq a^i, \text{ for } a < 1$$

$$C = \beta_\perp(\alpha'_\perp(I_p - \sum_{i=1}^{k-1} \Gamma_i)\beta_\perp)^{-1}\alpha'_\perp$$

$$x_t = C\sum_{i=1}^{t} \varepsilon_i + \sum_{i=0}^{\infty} C_i^* \varepsilon_{t-i} + A, \quad \beta' A = 0$$

1. Δx_t *is stationary,* x_t *is* $I(1)$
2. $\beta' x_t$ *is stationary,* x_t *is cointegrating (*r *cointegrating or long-run relations)*
3. *The common trends are* $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ *(*$p - r$ *common trends)*

Example

$$\begin{aligned}\Delta x_{1t} &= \alpha_1(x_{1t-1} - x_{2t-1}) + \varepsilon_{1t} \\ \Delta x_{2t} &= \alpha_2(x_{1t-1} - x_{2t-1}) + \varepsilon_{2t}\end{aligned}$$

Subtracting we find an AR(1) process

$$\Delta(x_{1t} - x_{2t}) = (\alpha_1 - \alpha_2)(x_{1t-1} - x_{2t-1}) + \varepsilon_{1t} - \varepsilon_{2t}$$

if $|1 + \alpha_1 - \alpha_2| < 1$ then

$$x_{1t} - x_{2t} = \sum_{i=0}^{\infty}(1 + \alpha_1 - \alpha_2)^i(\varepsilon_{1t-i} - \varepsilon_{2t-i})(= y_t)$$

is stationary.

Example continued

$$\Delta x_{1t} = \alpha_1(x_{1t-1} - x_{2t-1}) + \varepsilon_{1t}$$
$$\Delta x_{2t} = \alpha_2(x_{1t-1} - x_{2t-1}) + \varepsilon_{2t}$$

$y_t = x_{1t} - x_{2t}$ is stationary. Another linear combination is a random walk:

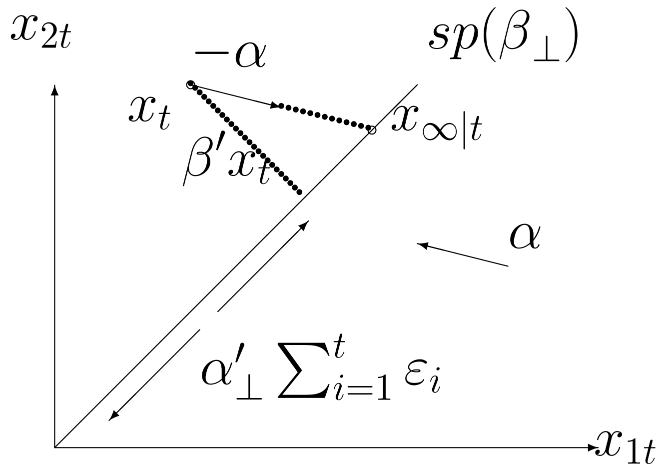
$$\alpha_2 \Delta x_{1t} - \alpha_1 \Delta x_{2t} = \alpha_2 \varepsilon_{1t} - \alpha_1 \varepsilon_{2t}$$
$$\alpha_2 x_{1t} - \alpha_1 x_{2t} = \alpha_2 x_{10} - \alpha_1 x_{20} + \sum_{i=1}^{t} (\alpha_2 \varepsilon_{1i} - \alpha_1 \varepsilon_{2i}) (= S_t)$$

$$x_{1t} = \frac{1}{\alpha_2 - \alpha_1}(S_t - \alpha_1 y_t) \text{ and } x_{2t} = \frac{1}{\alpha_2 - \alpha_1}(S_t - \alpha_2 y_t)$$

Thus if $|1 + \alpha_1 - \alpha_2| < 1$ then

1. $x_{1t} - x_{2t}$ is stationary
2. $\alpha_2 x_{1t} - \alpha_1 x_{2t}$ is random walk
3. x_t is $I(1)$
4. x_t cointegrated, cointegration vector $\beta' = (1, -1)$,

and common trend $\sum_{i=1}^{t} (\alpha_2 \varepsilon_{1i} - \alpha_1 \varepsilon_{2i})$.



1.In the model $\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t$, the point $x_t = (x_{1t}, x_{2t})$ is moved towards the long-run value $x_{\infty|t}$ on the attractor set $\{x|\beta' x = 0\} = sp(\beta_\perp\}$ by the force $-\alpha$ or $+\alpha$, and pushed along the attractor set by the common trends $\alpha'_\perp \sum_{i=1}^t \varepsilon_i$

Examples

$$\begin{aligned}\Delta x_{1t} &= -\frac{1}{4}(x_{1t-1} - x_{2t-1}) + \varepsilon_{1t} \\ \Delta x_{2t} &= \frac{1}{4}(x_{1t-1} - x_{2t-1}) + \varepsilon_{2t}\end{aligned}$$

gives $I(1)$ integration and cointegration because $1 + \alpha_1 - \alpha_2 = 1/2 < 1$.

Another example

$$\begin{aligned}\Delta x_{1t} &= \frac{1}{4}(x_{1t-1} - x_{2t-1}) + \varepsilon_{1t} \\ \Delta x_{2t} &= -\frac{1}{4}(x_{1t-1} - x_{2t-1}) + \varepsilon_{2t}\end{aligned}$$

is explosive and not cointegrated because $1 + \alpha_1 - \alpha_2 = 3/2 > 1$.

A strange example

$$\begin{aligned}\Delta x_{1t} &= \frac{1}{4}(x_{1t-1} - x_{2t-1}) + \frac{9}{4}\Delta x_{2t-1} + \varepsilon_{1t} \\ \Delta x_{2t} &= -\frac{1}{4}(x_{1t-1} - x_{2t-1}) + \varepsilon_{2t}\end{aligned}$$

is $I(1)$ and cointegrated despite $1 + \alpha_1 - \alpha_2 = 3/2 > 1$.
The sign of the adjustment is not intuitive
The processes do not adjust properly, yet are $I(1)$ because the term $\frac{9}{4}\Delta x_{2t-1}$. Roots are $4/3$ and 1.

$$\alpha = \frac{1}{4}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \beta = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, I_2 - \Gamma_1 = \begin{pmatrix} 1 & -\frac{9}{4} \\ 0 & 1 \end{pmatrix}, \alpha'_\perp(I_2 - \Gamma_1)\beta_\perp = -\frac{1}{4} \neq 0$$

(and the roots satisfy $|z| > 1$ or $z = 1$)

5. The Role of the Deterministic Terms

The linear 'innovation term'

$$\Delta x_t = \alpha\beta' x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + (\mu_0 + \mu_1 t + \varepsilon_t)$$

Granger Representation Theorem

$$\begin{aligned} x_t &= C\sum_{i=1}^{t}(\varepsilon_i + \mu_0 + \mu_1 i) + \sum_{i=0}^{\infty} C_i^*(\varepsilon_{t-i} + \mu_0 + \mu_1(t-i)) + A \\ &= C\sum_{i=1}^{t}\varepsilon_i + C\mu_0 t + \frac{1}{2}C\mu_1 t(t+1) + \sum_{i=0}^{\infty} C_i^*(\varepsilon_{t-i} + \mu_0 + \mu_1(t-i)) + A \end{aligned}$$

Thus

1. Quadratic trend in general, $\frac{1}{2}C\mu_1 t^2 = \frac{1}{2}\beta_\perp(.)^{-1}\alpha_\perp'\mu_1 t^2$
2. If $\alpha_\perp'\mu_1 = 0$, only linear trend (from both terms) because $C\mu_1 = 0$
3. If $\mu_1 = 0$, still linear trend, $C\mu_0 t$, but $\beta' x_t$ no trend, because $\beta' C = 0$
4. If $\mu_1 = 0,\ \alpha_\perp'\mu_0 = 0$ no linear trend because $C\mu_0 = 0$
5. If $\mu_1 = \mu_0 = 0$ no deterministics

The linear 'additive term'

$$\begin{aligned} x_t &= \tau_0 + \tau_1 t + y_t \\ \Delta y_t &= \alpha\beta' y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \\ \Delta x_t - \tau_1 &= \alpha\beta'(x_{t-1} - \tau_0 - \tau_1(t-1)) + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} - \sum_{i=1}^{k-1} \Gamma_i \tau_1 + \varepsilon_t \end{aligned}$$

The 'innovation' form is

$$\begin{aligned} \Delta x_t &= \alpha\beta' x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \mu_0 + \mu_1 t + \varepsilon_t \\ \mu_0 &= \alpha\beta'(\tau_1 - \tau_0) + (I_p - \sum_{i=1}^{k-1} \Gamma_i)\tau_1 \\ \mu_1 &= -\alpha\beta'\tau_1 \text{ so that } \alpha_\perp' \mu_1 = 0 \end{aligned}$$

Note that τ_1 and $\beta'\tau_0$ are identified

Other deterministics

The ‘innovation’ (blip-)dummy

$$d_t = 1_{\{t=t_0\}} = \begin{cases} 1, & t = t_0 \\ 0, & t \neq t_0 \end{cases}$$

Model

$$\Delta x_t = \alpha\beta' x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi d_t + \varepsilon_t$$

Granger Representation Theorem

$$x_t = C \sum_{i=1}^{t} (\varepsilon_i + \Phi d_i) + \sum_{i=0}^{\infty} C_i^* (\varepsilon_{t-i} + \Phi d_{t-i}) + A$$

The deterministic part of x_t is

$$C\Phi \sum_{i=1}^{t} d_i + \sum_{i=0}^{\infty} C_i^* \Phi d_{t-i} = C\Phi 1_{\{t \geq t_0\}} + C_{t-t_0}^* \Phi 1_{\{t \geq t_0\}}$$

The ‘additive’ (step-)dummy

$$
\begin{aligned}
x_t &= \phi 1_{\{t \geq t_0\}} + y_t \\
\Delta y_t &= \alpha\beta' y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \\
\Delta x_t - \phi 1_{\{t=t_0\}} &= \alpha\beta'(x_{t-1} - \phi 1_{\{t-1 \geq t_0\}}) + \sum_{i=1}^{k-1} (\Gamma_i \Delta x_{t-i} - \Gamma_i \phi 1_{\{t-i=t_0\}}) + \varepsilon_t \\
\Delta x_t &= \alpha\beta' x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} - \alpha\beta'\phi 1_{\{t-1 \geq t_0\}} + \phi 1_{\{t=t_0\}} - \sum_{i=1}^{k-1} \Gamma_i \phi 1_{\{t-i=t_0\}} + \varepsilon_t
\end{aligned}
$$

Note the step dummy $1_{\{t-1 \geq t_0\}}$ and the k blip dummies $1_{\{t-i=t_0\}}, i = 0, \ldots, k$.

6. Interpretation of Cointegrating Coefficients

Regression coefficients (elasticities)

$$x_{1t} = \beta_2 x_{2t} + \beta_3 x_{3t} + \varepsilon_t$$

are interpreted via a counter factual experiment

β_2 is the effect on x_1 of a change in $x_2,$ keeping x_3 constant

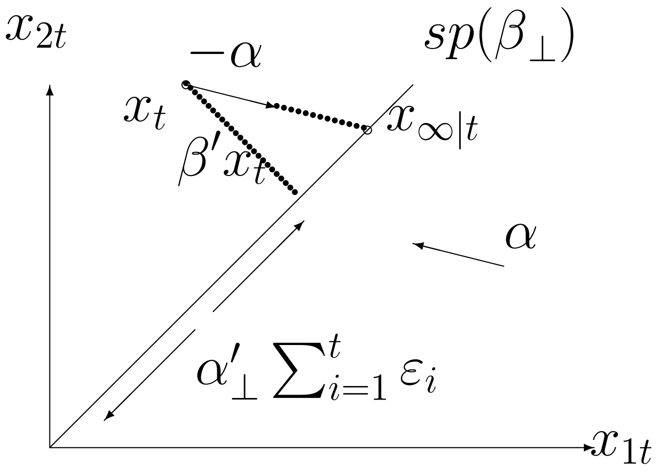
The estimate of β_2 is the regression of x_{1t} on x_{2t} correcting for x_{3t} :
$\hat{\beta}_2 = S_{12.3}S_{22.3}^{-1}, \; (S_{12.3} = S_{12} - S_{13}S_{33}^{-1}S_{32})$

$$\Delta x_t = \alpha(\beta' x_{t-1} - \beta_0) + \varepsilon_t$$

The cointegrating relations are long-run relations in the sense that they have been there all the time.
They influence the movement of the process by pulling towards the attractor set

$$\{x \in R^p | Cx = \alpha(\beta'\alpha)^{-1}\beta_0\} = \{x | \beta' x = \beta_0\}.$$

The set of steady states or long-run values.



1.In the model $\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t$, the point $x_t = (x_{1t}, x_{2t})$ is moved towards the long-run value $x_{\infty|t}$ on the attractor set $\{x|\beta' x = 0\} = sp(\beta_\perp\}$ by the force $-\alpha$ or $+\alpha$, and pushed along the attractor set by the common trends $\alpha'_\perp \sum_{i=1}^t \varepsilon_i$

Definition of long-run value

Model $\Delta x_t = \alpha(\beta' x_{t-1} - \beta_0) + \varepsilon_t$.

Solution of these equations for x_{t+h} for given x_t

$$x_{t+h} = (I_p + \alpha\beta')^h x_t + \sum_{i=0}^{h-1} (I_p + \alpha\beta')^i (\varepsilon_{t+h-i} - \alpha\beta_0),$$

implies

$$E(x_{t+h}|x_t) = (I_p + \alpha\beta')^h x_t - \sum_{i=0}^{h-1} (I_p + \alpha\beta')^i \alpha\beta_0 \to ?$$

$$E(\alpha'_\perp x_{t+h}|x_t) = \alpha'_\perp x_t$$

$$E(\beta' x_{t+h}|x_t) = (I_r + \beta'\alpha)^h \beta' x_t - \sum_{i=0}^{h-1} (I_r + \beta'\alpha)^i \beta'\alpha\beta_0 \to \beta_0$$

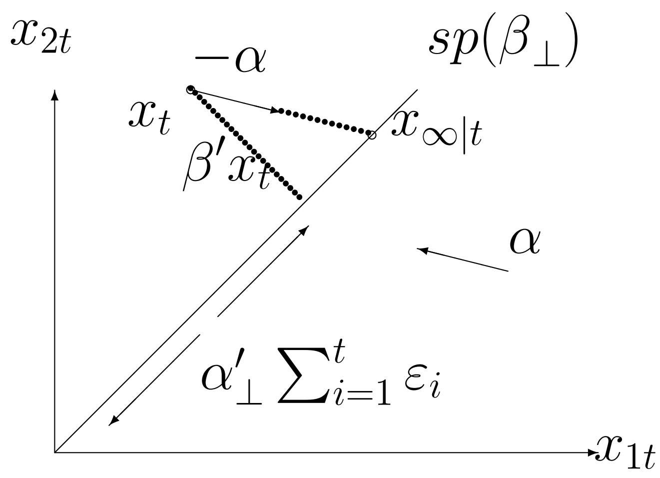
$$x_{\infty|t} = \lim_{h\to\infty} E(x_{t+h}|x_t) = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp x_t + \alpha(\beta'\alpha)^{-1}\beta_0$$

The long-run value $x_{\infty|t}$

$$x_{\infty|t} = \lim_{h\to\infty} E(x_{t+h}|x_t) = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp x_t + \alpha(\beta'\alpha)^{-1}\beta_0 = Cx_t + \alpha(\beta'\alpha)^{-1}\beta_0,$$

is a point in the attractor set, because $\beta' x_{\infty|t} = \beta_0$

A cointegrating relation is a relation between long-run values



1.In the model $\Delta x_t = \alpha\beta' x_{t-1} + \varepsilon_t$, the point $x_t = (x_{1t}, x_{2t})$ is moved towards the long-run value $x_{\infty|t}$ on the attractor set $\{x|\beta' x = 0\} = sp(\beta_\perp\}$ by the force $-\alpha$ or $+\alpha$, and pushed along the attractor set by the common trends $\alpha'_\perp \sum_{i=1}^t \varepsilon_i$

Counter factual experiment (lag one model):
If the current value is changed from x_t to $x_t + c$, then the long-run value is changed from $x_{\infty|t}$ to $x_{\infty|t} + Cc = x_{\infty|t} + \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp c$, which is a vector in $sp(\beta_\perp)$, because $\beta'(x_{\infty|t} + Cc) = 0$.
If $c = C\xi$ (or $\beta' c = 0$), we *can produce a change from* $x_{\infty|t}$ *to* $x_{\infty|t} + c$ *by adding* c *to* x_t :

$$C(x_t + c) + \alpha(\beta'\alpha)^{-1}\beta_0 = x_{\infty|t} + C^2\xi = x_{\infty|t} + C\xi = x_{\infty|t} + c.$$

Consider the cointegrating relation $\beta' x = \beta_0$ solved for x_1

$$x_1 = \gamma_2 x_2 + \gamma_3 x_3 + \beta_0.$$

Then

γ_2 is the effect of x_2 on x_1 keeping x_3 fixed

in the sense that the long-run change $c = (\gamma_2, 1, 0)'$ is orthogonal to $\beta : (1, -\gamma_2, -\gamma_3)c = 0$, so we can achieve the long-run change c by moving the current value to $x_t + c$.

The Structural Error Correction Model

$$\Gamma_0^* \Delta x_t = \alpha^*(\beta' x_{t-1} - \beta_0) + \sum_{i=1}^{k-1} \Gamma_i^* \Delta x_{t-i} + \varepsilon_t^*,$$

where $\varepsilon_t^* = \Gamma_0^* \varepsilon_t$ is i.i.d. $(0, \Omega^*)$, $\Omega^* = \Gamma_0^* \Omega \Gamma_0^{*\prime}$, and $\alpha^* = \Gamma_0^* \alpha, \Gamma_i^* = \Gamma_0^* \Gamma_i$.
The parameters β and β_0 are unchanged, but all the other coefficients have changed.
Two identification problems
1. Identify β by restrictions
2. Identify the remaining coefficients by restrictions (possibly on Ω^*)
Long-run **relations between** variables
But **equations for** the variables in the system
$m - y - c(i^b - i^d) = \beta_0$ is **money relation** if solved for money
$\Delta m_t = \alpha_1 (m - y - c(i^b - i^d))_{t-1} + ...$is **money equation**

7. Shocks, Changes and Impulse Responses

The model

$$\begin{aligned}\Delta x_t &= \alpha\beta' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \varepsilon_t \\ x_t &= (I_p + \alpha\beta')x_{t-1} + \Gamma_1 \Delta x_{t-1} + \varepsilon_t\end{aligned}$$

shows that a change in ε_t $(\varepsilon_t \mapsto \varepsilon_t + c)$ is equivalent to a change in x_t $(x_t \mapsto x_t + c)$. We call ε_t a **shock** and c a **change**. Granger Representation Theorem

$$x_{t+h} = C\sum_{i=1}^{t+h}\varepsilon_i + \sum_{i=0}^{\infty} C_i\varepsilon_{t+h-i} + A = C(\varepsilon_1 + \ldots + \varepsilon_t + \ldots + \varepsilon_{t+h}) + C_0\varepsilon_{t+h} + \ldots + C_h\varepsilon_t + \ldots + A$$

implies that the effect at time $t+h$ of a change c to ε_t (or x_t) is

$$\frac{\partial x_{t+h}}{\partial \varepsilon_t} c = (C + C_h)c \to Cc, h \to \infty$$

The function $h \to (C + C_h)c$ is the reduced form impulse response function. A change c to the system at time t propagates through the system and becomes Cc in the long run. The permanent effect of a change is Cc.

Reduced form impulse responses

$$(C+C_h)\varepsilon_t = \sum_{i=1}^{p}(C+C_h)e_i e_i'\varepsilon_t = \sum_{i=1}^{p}(C+C_h)e_i\varepsilon_{it}$$

e_i is i'th unit vector

The effect at time horizon h of variable i on variable j is given by

$$e_j'(C+C_h)e_i = (C+C_h)_{ji}, h = 1, 2, 3, \ldots$$

The effect of the change c on the system is

$$(C+C_h)c = \sum_{i=1}^{p}(C+C_h)e_i e_i'c = \sum_{i=1}^{p}(C+C_h)e_i c_i$$

a combination of the changes c_i to the i'th variable.

Structural form impulse responses

$$(C + C_h)\varepsilon_t = (C + C_h)B^{-1}B\varepsilon_t = \sum_{i=1}^{p}(C + C_h)w_i v_i'\varepsilon_t$$

where $v_i'\varepsilon_t$ is a structural shock and w_i is its effect

$$B^{-1} = (w_1, \ldots, w_p),\ B' = (v_1, \ldots, v_p),\ \sum_{i=1}^{p} w_i v_i' = I_p$$

Construct v_i so that $v_i'\varepsilon_t$ ‘makes economic sense’.
(Cholesky decomposition: $B\Omega B' = I_p$, B triangular, requires ordering of variables)
A change of c_i of the i'th structural shock has the effect $w_i c_i$ on the system, that is,

$$(C + C_h)w_i c_i, h = 1, 2, 3, \ldots$$

The structural impulse response function shows the effect on individual variables of changing many variables at the same time.

Permanent and transitory effects and changes

The response to a change c is

$$(C + C_h)c \rightarrow Cc = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp c$$

Definition of permanent and transitory effects

1. Cc is the *permanent effect* of a change c
2. $C_h c, h = 1, 2, \dots.$ is the *transitory effect* of a change c

Permanent and transitory component of a shock
The decomposition

$$\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} + \Omega\alpha_\perp(\alpha'_\perp\Omega\alpha_\perp)^{-1}\alpha'_\perp = I_p$$

gives rise to the decomposition

$$\varepsilon_t = \underbrace{\alpha(\alpha'\Omega^{-1}\alpha)^{-1}}_{effect}\underbrace{\alpha'\Omega^{-1}\varepsilon_t}_{trans\ shock} + \underbrace{\Omega\alpha_\perp(\alpha'_\perp\Omega\alpha_\perp)^{-1}}_{effect}\underbrace{\alpha'_\perp\varepsilon_t}_{perm\ shock}$$

of the shock ε_t into the **permanent shock** $\alpha'_\perp\varepsilon_t$ and the (independent) **transitory shock** $\alpha'\Omega^{-1}\varepsilon_t$.

$$Cov(\alpha'\Omega^{-1}\varepsilon_t, \alpha'_\perp\varepsilon_t) = \alpha'_\perp\Omega^{-1}\Omega\alpha = \alpha'_\perp\alpha = 0$$

The effect $\alpha(\alpha'\Omega^{-1}\alpha)^{-1}$ of the transitory shock gives transitory change:

$$(C + C_h)\alpha(\alpha'\Omega^{-1}\alpha)^{-1} = C_h\alpha(\alpha'\Omega^{-1}\alpha)^{-1} \underset{h\to\infty}{\longrightarrow} 0,$$

The effect $\Omega\alpha_\perp(\alpha'_\perp\Omega\alpha_\perp)^{-1}$ of the permanent shock gives permanent change

$$(C + C_h)\Omega\alpha_\perp(\alpha'_\perp\Omega\alpha_\perp)^{-1} \underset{h\to\infty}{\longrightarrow} C\Omega\alpha_\perp(\alpha'_\perp\Omega\alpha_\perp)^{-1} = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}$$

Transitory and permanent structural shocks

$$\varepsilon_t = \underbrace{\alpha(\alpha'\Omega^{-1}\alpha)^{-1}B_{tr}^{-1}}_{effect} \underbrace{B_{tr}\alpha'\Omega^{-1}\varepsilon_t}_{struct.trans.shock} + \underbrace{\Omega\alpha_{\perp}(\alpha'_{\perp}\Omega\alpha_{\perp})^{-1}B_{pm}^{-1}}_{effect} \underbrace{B_{pm}\alpha'_{\perp}\varepsilon_t}_{struct.perm\ shock}$$

Choose B_{tr} and B_{pm} to make ‘structural’ transitory and permanent shocks

8. HYPOTHESES OF INTEREST

Hypotheses on the rank

The model

$$\mathcal{H}_r : \Delta x_t = \alpha\beta' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \varepsilon_t, \quad \alpha_{p\times r}, \beta_{p\times r}$$

$$\mathcal{H}_0 \subset \cdots \subset \mathcal{H}_r \subset \mathcal{H}_{r+1} \subset \cdots \subset \mathcal{H}_p$$

The rank is the number of ‘stable’ economic long-run relations between the ‘unstable’ economic data

1. Test for a given rank consistent with economic theory or
2. Estimate the rank from the data

Hypotheses on β

The model and the data $x_t = (y_t, c_t, k_t, h_t)$

$$\Delta x_t = \alpha(\beta' x_{t-1} - \beta_0) + \Gamma_1 \Delta x_{t-1} + \varepsilon_t$$

Assume first that $r = 1$:
Hypothesis: income-consumption ratio stationary

$$\beta' = (1, -1, 0, 0) : \beta' x_t = y_t - c_t$$

Hypothesis: Cobb-Douglas relation stationary

$$\beta' = (1, 0, -(1-\theta), -\theta) : \beta' x_t = y_t - \theta h_t - (1-\theta) k_t$$

Hypothesis: hours worked is stationary

$$\beta' = (0, 0, 0, 1) : \beta' x_t = h_t$$

A general formulation of the same linear restriction on all vectors

Data $x_t = (y_t, c_t, k_t, h_t)$
Hypothesis: income-consumption ratio is a function of hours worked

$$\beta' = (1, -1, 0, \phi) : \beta' x_t = y_t - c_t + \phi h_t$$

$$\mathcal{H} : \beta = H\phi, \text{ or } R'\beta = 0 \text{ where } R = H_\perp.$$

$$\beta = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \phi_0 \\ \phi_1 \end{pmatrix} = \begin{pmatrix} \phi_0 \\ -\phi_0 \\ 0 \\ \phi_1 \end{pmatrix} : \beta' x_t = \phi_0(y_t - c_t) + \phi h_t$$

(normalize on $\phi_0 = 1$)

Next assume that $r = 2$

Hypothesis: Income-consumption ratio and hours worked are stationary

$$\beta' = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, R'\beta = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \beta = 0$$

Hypothesis: Homogeneity between income and consumption

$$R'\beta = (1, 1, 0, 0)'\beta = 0$$

or

$$\beta_{4\times 2} = H_{4\times 3}\phi_{3\times 2} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \phi_{3\times 2} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ -\phi_{11} & -\phi_{12} \\ \phi_{21} & \phi_{22} \\ \phi_{31} & \phi_{32} \end{pmatrix}$$

β not identified because linear combinations of the columns have the same restrictions

Hypothesis: Hours worked stationary

$$\beta = \begin{pmatrix} 0 & \phi_1 \\ 0 & \phi_2 \\ 0 & \phi_3 \\ 1 & \phi_4 \end{pmatrix} = (b, \phi_{4\times 1})$$

($\beta = (b, H\phi)$ not identified)

Identification problem for β

The vector β_1 is identified by restrictions $R'\beta_1 = 0$ if there is no linear combination $\sum_j a_j\beta_j$ satisfying the restrictions $R'\sum_j a_j\beta_j = 0$, (other than $a_1 = 1, a_i = 0, i \neq 1)$

General linear restrictions

$$\beta_{4\times 2} = (H_1\phi_1, H_2\phi_2) \text{ or } R_i'\beta_i = 0, \quad R_i = H_{i\perp}$$

is identified if

$$rank(R_i'\beta) \geq 1, i = 1, 2 \qquad \text{The Wald criterion}$$

Another criterion (independent of the unknown parameter) is

$$rank(R_i'H_j) \geq 1, \text{ for } i \neq j = 1, 2.$$

Example

$$
\begin{aligned}
y_t + a_2 c_t + a_3 k_t + a_4 h_t &\sim I(0) \\
y_t - c_t &\sim I(0)
\end{aligned}
$$

The first is not identified and the second is. If instead

$$
\begin{aligned}
k_t + a_2 h_t &\sim I(0) \\
y_t - c_t &\sim I(0)
\end{aligned}
$$

then both identified.

Weak exogeneity (Gaussian errors)

$$\begin{aligned}\Delta x_{1t} &= \alpha_1\beta' x_{t-1} + \Gamma_{11}\Delta x_{t-1} + \varepsilon_{1t}\\ \Delta x_{2t} &= \alpha_2\beta' x_{t-1} + \Gamma_{21}\Delta x_{t-1} + \varepsilon_{2t}\end{aligned}$$

Model for Δx_{1t} given Δx_{2t} $(\omega = \Omega_{12}\Omega_{22}^{-1})$

$$\Delta x_{1t} = \omega\Delta x_{2t} + (\alpha_1 - \omega\alpha_2)\beta' x_{t-1} + (\Gamma_{11} - \omega\Gamma_{21})\Delta x_{t-1} + (\varepsilon_{1t} - \omega\varepsilon_{2t})$$

Definition. *x_{2t} is weakly exogenous for β, if marginal model for Δx_{2t} has no information on β, that is, if* $\alpha_2 = 0$:

$$\begin{aligned}\text{conditional model} &: \Delta x_{1t} = \omega\Delta x_{2t} + \alpha_1\beta' x_{t-1} + (\Gamma_{11} - \omega\Gamma_{21})\Delta x_{t-1} + (\varepsilon_{1t} - \omega\varepsilon_{2t})\\ \text{marginal model} &: \Delta x_{2t} = \Gamma_{21}\Delta x_{t-1} + \varepsilon_{2t}\end{aligned}$$

Parameters

$$\begin{aligned}\theta_{cond} &= (\omega, \alpha_1, \beta, \Gamma_{11} - \omega\Gamma_{21}, \Omega_{11.2})\\ \theta_{marg} &= (\Gamma_{21}, \Omega_{22})\end{aligned}$$

are variation free: reparametrization in terms of $\theta_{cond}, \theta_{marg}$ is one to one.
Strong and super exogeneity.

Hypotheses on α
The model and the data

$$x_t = (y_t, c_t, k_t, h_t)'$$

$$\Delta x_t = \alpha\beta' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \varepsilon_t$$

If the shocks to k_t (technology shocks or supply shocks) drive the system, then k_t is weakly exogenous: Third row in α equals to zero

$$(0,0,1,0)'\alpha = 0 \text{ or } \alpha_{4\times 2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \phi_{3\times 2}$$

Then $\alpha'_\perp = (0,0,1,0)'$ and $\alpha'_\perp \sum_{i=1}^t \varepsilon_i = \sum_{i=1}^t \varepsilon_{3i}$. Thus the shocks to the weakly exogenous variable cumulate to stochastic trends (pushing forces)
If the shocks to c_t (demand shocks) drive the system, then c_t is weakly exogenous: Second row in α equals to zero.

If the supply shocks (to k_t) do not contribute to the pushing forces, then a column of α is proportional to unit vector

$$\alpha_{4\times 2} = \begin{pmatrix} 0 & \phi_1 \\ 0 & \phi_2 \\ 1 & \phi_3 \\ 0 & \phi_4 \end{pmatrix} = (a, \phi) \text{ or } (0,0,1,0)'\alpha_\perp = 0$$

Thus $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ does not contain $\sum_{i=1}^{t} \varepsilon_{3i}$, or ε_{3t} is not a pushing force.

Identification problem for $\alpha_\perp$

From Granger Representation Theorem the common trends are $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$, which 'create' the non-stationarity, but this does not identify the trends.
For any $\eta_{(p-r)\times(p-r)}$: $\eta\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ can also be used:

$$C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp = \beta_\perp(\eta\alpha'_\perp \Gamma \beta_\perp)^{-1}\eta\alpha'_\perp$$

Linear restrictions on $\alpha_\perp$ can identify the trends uniquely.

9. Conclusions

The cointegrated vector autoregressive model
$\Delta x_t = \alpha(\beta' x_{t-1} - \beta_0) + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \varepsilon_t$
is a dynamic stochastic model for all the variables, that allows the simultaneous modelling of the long-run relations $\beta' x = \beta_0$, and the adjustment towards the disequilibrium errors.

1. The long-run relations $\beta' x = \beta_0$ define the attractor set

$$\{x \in R^p | Cx = \alpha(\beta'\alpha)^{-1}\beta_0\} = \{x | \beta' x = \beta_0\}$$

the set of equilibria or steady states. The coefficients are long-run elasticities.
2. The adjustment coefficients α define the direction of adjustment, the ‘pulling forces’
3. The common trends, given by $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$, define the ‘pushing forces’

The Granger Representation Theorem gives the solution of the autoregressive equations and is useful for

1. The role of deterministics
2. Asymptotics

The equations are **equations for** the variables and the cointegrating relations are **relations between** the variables