

Structural Vector Autoregressions and Asymptotic Theory for Time Series Econometrics

Part II. Asymptotic Theory

Helmut Lutkepohl

European University Institute, Florence

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

Multivariate Normal Distribution (Section B.1)

Notation

$$y \sim \mathcal{N}(\mu, \Sigma) \quad K\text{-dimensional}$$

Probability density function (p.d.f.)

$$f(y) = \frac{1}{(2\pi)^{K/2}} |\Sigma|^{-1/2} \exp \left[-\frac{1}{2} (y - \mu)' \Sigma^{-1} (y - \mu) \right]$$

Properties of Multivariate Normal Distribution

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

marginal distribution

$$y_1 \sim \mathcal{N}(\mu_1, \Sigma_{11})$$

conditional distribution

$$(y_1 | y_2 = c) \sim \mathcal{N}(\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (c - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$$

linear transformation

$$x = Ay + c \sim \mathcal{N}(A\mu + c, A\Sigma A')$$

Related Distributions (Section B.2)

χ^2 distribution with K degrees of freedom

$$y \sim \mathcal{N}(0, I_K) \quad \Rightarrow \quad z = y'y \sim \chi^2(K)$$

t distribution with m degrees of freedom

$z \sim \mathcal{N}(0, 1)$ and $u \sim \chi^2(m)$ stochastically independent

$$\Rightarrow \frac{z}{\sqrt{u/m}} \sim t(m)$$

F distribution with m and n degrees of freedom

$u \sim \chi^2(m)$ and $v \sim \chi^2(n)$ independent

$$\Rightarrow \frac{u/m}{v/n} \sim F(m, n)$$

Concepts of Stochastic Convergence (Section C.1)

convergence in probability

$$\text{plim } x_T = x \quad \text{or} \quad x_T \xrightarrow{P} x \quad \text{iff} \\ \forall \epsilon > 0, \quad \lim_{T \rightarrow \infty} \Pr(|x_T - x| < \epsilon) = 1$$

almost sure convergence, convergence with probability 1

$$x_T \xrightarrow{\text{a.s.}} x \quad \text{iff} \\ \forall \epsilon > 0, \quad \Pr(\lim_{T \rightarrow \infty} |x_T - x| < \epsilon) = 1$$

convergence in quadratic mean or mean square

$$x_T \xrightarrow{\text{q.m.}} x \quad \text{iff} \\ \lim_{T \rightarrow \infty} E(x_T - x)^2 = 0$$

convergence in distribution

$$x_T \xrightarrow{d} x \quad \text{iff} \\ \lim_{T \rightarrow \infty} F_T(c) = F(c) \quad (\text{at all continuity points})$$

where F_T and F distribution functions of x_T and x , respectively

Relations between Concepts of Stochastic Convergence

$$x_T \xrightarrow{a.s.} x \Rightarrow x_T \xrightarrow{p} x \Rightarrow x_T \xrightarrow{d} x$$

$$x_T \xrightarrow{q.m.} x \Rightarrow x_T \xrightarrow{p} x \Rightarrow x_T \xrightarrow{d} x$$

Nonstochastic limit

If x is a fixed, nonstochastic vector:

$$x_T \xrightarrow{q.m.} x$$

$$\Leftrightarrow [\lim E(x_T) = x \text{ and } \lim E\{(x_T - Ex_T)'(x_T - Ex_T)\} = 0]$$

$$x_T \xrightarrow{p} x \Leftrightarrow x_T \xrightarrow{d} x$$

Slutsky's Theorem

$g : \mathbb{R}^K \rightarrow \mathbb{R}^m$ continuous:

$$x_T \xrightarrow{p} x \Rightarrow g(x_T) \xrightarrow{p} g(x) \quad [\text{plim } g(x_T) = g(\text{plim } x_T)]$$

$$x_T \xrightarrow{d} x \Rightarrow g(x_T) \xrightarrow{d} g(x)$$

$$x_T \xrightarrow{a.s.} x \Rightarrow g(x_T) \xrightarrow{a.s.} g(x)$$

Order of Convergence (Section C.2)

a_T of smaller order than b_T

$$a_T = o(b_T) \text{ iff } \lim_{T \rightarrow \infty} a_T/b_T = 0$$

a_T at most of order b_T

$$a_T = O(b_T) \text{ iff } |a_T|/b_T \leq c$$

- ① $a_T = o(c_T), b_T = o(d_T) \Rightarrow a_T b_T = o(c_T d_T), a_T + b_T = o(\max[c_T, d_T])$ and $|a_T|^s = o(c_T^s)$ for $s > 0$.
- ② $a_T = O(c_T), b_T = O(d_T) \Rightarrow a_T b_T = O(c_T d_T), a_T + b_T = O(\max[c_T, d_T])$ and $|a_T|^s = O(c_T^s)$ for $s > 0$.
- ③ $a_T = o(c_T), b_T = O(d_T) \Rightarrow a_T b_T = o(c_T d_T)$.

Order in Probability (Section C.2)

A_T of smaller order in probability than b_T

$$A_T = o_p(b_T) \text{ iff } \text{plim}_{T \rightarrow \infty} A_T/b_T = 0$$

A_T is bounded in probability by b_T

$$A_T = O_p(b_T) \text{ iff } \forall \epsilon > 0, \exists c_\epsilon, \Pr\{|a_{ij,T}| \geq c_\epsilon b_T\} \leq \epsilon$$

Hence

$$a_T \xrightarrow{d} a \quad \Rightarrow \quad a_T = O_p(1)$$

- ① $A_T = o_p(c_T), B_T = o_p(d_T)$
 $\Rightarrow A_T B_T = o_p(c_T d_T)$ and $A_T + B_T = o_p(\max[c_T, d_T])$.
- ② $A_T = O_p(c_T), B_T = O_p(d_T)$
 $\Rightarrow A_T B_T = O_p(c_T d_T)$ and $A_T + B_T = O_p(\max[c_T, d_T])$.
- ③ $A_T = o_p(c_T), B_T = O_p(d_T)$
 $\Rightarrow A_T B_T = o_p(c_T d_T)$.

Infinite Sums of Random Variables (Section C.3)

$$\|A_i\| = [\text{tr}(A_i A_i')]^{1/2} = \left(\sum_m \sum_n a_{mn,i}^2 \right)^{1/2}$$

Existence

$$E(z_t' z_t) \leq c, \quad t = 0, \pm 1, \pm 2, \dots, \text{ and } \sum_{i=-\infty}^{\infty} \|A_i\| < \infty$$

$$\Rightarrow \sum_{i=-n}^n A_i z_{t-i} \xrightarrow{q.m.}_{n \rightarrow \infty} y_t$$

Moments

$$y_t = \sum_{i=-\infty}^{\infty} A_i z_{t-i} \quad \text{and} \quad x_t = \sum_{i=-\infty}^{\infty} B_i z_{t-i}$$

$$\Rightarrow E(y_t) = \lim_{n \rightarrow \infty} \sum_{i=-n}^n A_i E(z_{t-i})$$

$$\Rightarrow E(y_t x_t') = \lim_{n \rightarrow \infty} \sum_{i=-n}^n \sum_{j=-n}^n A_i E(z_{t-i} z_{t-j}') B_j'$$

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

Definitions and Tools (Section C.4)

martingale difference sequence with respect to Ω_t

$$\{x_t\} \text{ with } E(x_t) = 0 \forall t \text{ and } E(x_t | \Omega_{t-1}) = 0 \\ t = 2, 3, \dots$$

martingale difference sequence

$$\{x_t\} \text{ with } E(x_t) = 0 \text{ for } t = 1, 2, \dots, \text{ and} \\ E(x_t | x_{t-1}, \dots, x_1) = 0 \text{ for } t = 2, 3, \dots$$

martingale difference array

$$\{x_{T,t}\} (t = 1, 2, \dots, T; T = 1, 2, \dots) \text{ with } E(x_{T,t}) = 0 \\ \forall t, T \text{ and } E(x_{T,t} | x_{T,t-1}, \dots, x_{T,1}) = 0 \forall t \text{ and } T > 1$$

Chebychev's inequality

$$\forall c \in \mathbb{R}, \epsilon > 0, \quad \Pr\{|x - c| \geq \epsilon\} \leq \frac{E(|x - c|^r)}{\epsilon^r}$$

WLLNs I

Khinchine's Theorem

Let $\{x_t\}$ be a sequence of i.i.d. random variables with $E(x_t) = \mu < \infty$. Then

$$\bar{x}_T := \frac{1}{T} \sum_{t=1}^T x_t \xrightarrow{P} \mu.$$

LLN for non-i.i.d. random variables

Let $\{x_t\}$ be a sequence of independent random variables with $E(x_t) = \mu < \infty$ and $E|x_t|^{1+\epsilon} \leq c < \infty$ ($t = 1, 2, \dots$) for some $\epsilon > 0$ and a finite constant c . Then $\bar{x}_T \xrightarrow{P} \mu$.

Chebyshev's Theorem

Let $\{x_t\}$ be a sequence of uncorrelated random variables with $E(x_t) = \mu < \infty$ and $\lim_{T \rightarrow \infty} E(\bar{x}_T - \mu)^2 = 0$. Then $\bar{x}_T \xrightarrow{P} \mu$.

WLLNs II

Corollary to Chebyshev's Theorem

Let $\{x_t\}$ be a sequence of independent random variables with $E(x_t) = \mu < \infty$ and $\text{Var}(x_t) \leq c < \infty$ ($t = 1, 2, \dots$) for some finite constant c . Then $\bar{x}_T \xrightarrow{P} \mu$.

LLN for Martingale Differences

Let $\{x_t\}$ be a strictly stationary martingale difference sequence with $E|x_t| < \infty$ ($t = 1, 2, \dots$). Then $\bar{x}_T \xrightarrow{P} 0$.

LLN for Martingale Difference Arrays

Let $\{x_{T,t}\}$ be a martingale difference array with $E|x_{T,t}|^{1+\epsilon} \leq c < \infty$ for all t and T for some $\epsilon > 0$ and a finite constant c . Then $\bar{x}_T := T^{-1} \sum_{t=1}^T x_{T,t} \xrightarrow{P} 0$.

WLLNs III

Stationary Processes

Let $\{x_t\}$ be a stationary stochastic process with $E(x_t) = \mu < \infty$ and $E[(x_t - \mu)(x_{t-j} - \mu)] = \gamma_j$ ($t = 1, 2, \dots$) such that $\sum_{j=0}^{\infty} |\gamma_j| < \infty$.

Then $\bar{x}_T \xrightarrow{q.m.} \mu$ and, hence, $\bar{x}_T \xrightarrow{p} \mu$,
and $\lim_{T \rightarrow \infty} TE(\bar{x}_T - \mu)^2 = \sum_{j=-\infty}^{\infty} \gamma_j$.

CLTs I

Lindeberg-Levy CLT

Let $\{x_t\}$ be a sequence of K -dimensional i.i.d. random vectors with mean μ and covariance matrix Σ_x . Then

$$\sqrt{T}(\bar{x}_T - \mu) \xrightarrow{d} \mathcal{N}(0, \Sigma_x).$$

CLT for Martingale Difference Arrays

Let $\{x_{T,t} = (x_{1T,t}, \dots, x_{KT,t})'\}$ be a K -dimensional martingale difference array with covariance matrices $E(x_{T,t}x_{T,t}') = \Sigma_{Tt}$ such that $T^{-1} \sum_{t=1}^T \Sigma_{Tt} \rightarrow \Sigma$, where Σ is positive definite. Moreover, suppose that $T^{-1} \sum_{t=1}^T x_{T,t}x_{T,t}' \xrightarrow{p} \Sigma$ and $E(x_{iT,t}x_{jT,t}x_{kT,t}x_{lT,t}) < \infty$ for all t and T and all $1 \leq i, j, k, l \leq K$. Then

$$\sqrt{T} \bar{x}_T \xrightarrow{d} \mathcal{N}(0, \Sigma).$$

CLTs II

CLT for Stationary Processes

Let $x_t = \mu + \sum_{j=0}^{\infty} \Phi_j u_{t-j}$ be a K -dimensional stationary stochastic process with $E(x_t) = \mu < \infty$, $\sum_{j=0}^{\infty} \|\Phi_j\| < \infty$ and $u_t \sim (0, \Sigma_u)$ i.i.d. white noise. Then

$$\sqrt{T}(\bar{x}_T - \mu) \xrightarrow{d} \mathcal{N} \left(0, \sum_{j=-\infty}^{\infty} \Gamma_x(j) \right),$$

where $\Gamma_x(j) := E[(x_t - \mu)(x_{t-j} - \mu)']$.

OLS estimator

Linear model: $y = X\beta + u$

Consistency

$$\begin{aligned} b &= (X'X)^{-1}X'y &= (X'X)^{-1}X'(X\beta + u) \\ & &= \beta + \left(\frac{X'X}{T}\right)^{-1} \left(\frac{X'u}{T}\right) \\ & &= \beta + O(1)o_p(1) \\ & &= \beta + o_p(1) \end{aligned}$$

Asymptotic normality

$$\sqrt{T}(b - \beta) = \left(\frac{X'X}{T}\right)^{-1} \left(\frac{X'u}{\sqrt{T}}\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2 Q^{-1})$$

Standard asymptotic properties of estimators (Section C.5)

Assumption $\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \Sigma)$

Linear transformation

$$\text{plim } \hat{A} = A \quad \Rightarrow \quad \sqrt{T}\hat{A}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, A\Sigma A')$$

Delta method

$$g(\beta) = (g_1(\beta), \dots, g_m(\beta))' \text{ with } \partial g / \partial \beta' \neq 0 \text{ at } \beta$$

$$\Rightarrow \sqrt{T}[g(\hat{\beta}) - g(\beta)] \xrightarrow{d} \mathcal{N}\left(0, \frac{\partial g(\beta)}{\partial \beta'} \Sigma \frac{\partial g(\beta)'}{\partial \beta}\right)$$

Quadratic form

$$\text{plim } \hat{\Sigma} = \Sigma \quad \Rightarrow \quad T(\hat{\beta} - \beta)' \hat{\Sigma}^{-1} (\hat{\beta} - \beta) \xrightarrow{d} \chi^2(K)$$

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 **Session 3. Estimation and Testing**
 - **Maximum Likelihood Estimation**
 - **Testing Principles**
 - **GMM Estimation**
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

Maximum Likelihood Estimation (Section C.7) I

Sample

$$y_1, \dots, y_T$$

Sample density function

$$f_T(y_1, \dots, y_T; \delta_0)$$

Likelihood function

$$l(\delta) = l(\delta|y_1, \dots, y_T) = f_T(y_1, \dots, y_T; \delta)$$

log-Likelihood function

$$\ln l(\delta|\cdot)$$

Maximum likelihood (ML) estimator

$$\begin{aligned} \tilde{\delta} \text{ such that } l(\tilde{\delta}) &= \sup_{\delta \in \mathbb{D}} l(\delta), \text{ i.e.,} \\ \tilde{\delta} &= \arg \max l(\delta) \end{aligned}$$

Maximum Likelihood Estimation (Section C.7) II

Score vector

$$s(\delta) = \partial \ln l(\delta) / \partial \delta$$

Information matrix

$$\mathcal{I}(\delta_0) = -E \left[\frac{\partial^2 \ln l}{\partial \delta \partial \delta'} \Big|_{\delta_0} \right]$$

Asymptotic information matrix

$$\mathcal{I}_a(\delta_0) = \lim_{T \rightarrow \infty} \mathcal{I}(\delta_0) / T$$

Asymptotic distribution of ML estimator

$$\sqrt{T}(\tilde{\delta} - \delta_0) \xrightarrow{d} \mathcal{N}(0, \mathcal{I}_a(\delta_0)^{-1})$$

under general regularity conditions

Asymptotic optimality of ML estimator

$$\sqrt{T}(\hat{\delta} - \delta_0) \xrightarrow{d} \mathcal{N}(0, \Sigma_{\hat{\delta}}) \Rightarrow \mathcal{I}_a(\delta_0)^{-1} \leq \Sigma_{\hat{\delta}}$$

LR, LM and Wald Tests

Hypotheses

$$H_0 : \varphi(\delta_0) = 0 \quad \text{against} \quad H_1 : \varphi(\delta_0) \neq 0$$

or

$$H_0 : \delta_0 = g(\gamma_0) \quad \text{against} \quad H_1 : \delta_0 \neq g(\gamma_0)$$

Likelihood ratio (LR) test

$$\lambda_{LR} = 2[\ln l(\tilde{\delta}) - \ln l(\tilde{\delta}_r)] \xrightarrow{d} \chi^2(N)$$

Lagrange multiplier (LM) test

$$\begin{aligned} \lambda_{LM} &= s(\tilde{\delta}_r)' \mathcal{I}(\tilde{\delta}_r)^{-1} s(\tilde{\delta}_r) \\ &= \tilde{\lambda}' \left[\left. \frac{\partial \varphi}{\partial \delta'} \right|_{\tilde{\delta}_r} \right] \mathcal{I}(\tilde{\delta}_r)^{-1} \left[\left. \frac{\partial \varphi'}{\partial \delta} \right|_{\tilde{\delta}_r} \right] \tilde{\lambda} \xrightarrow{d} \chi^2(N) \end{aligned}$$

Wald test

$$\lambda_W = T \varphi(\tilde{\delta})' \left(\left[\left. \frac{\partial \varphi}{\partial \delta'} \right|_{\tilde{\delta}} \right] \tilde{\Sigma}_{\tilde{\delta}} \left[\left. \frac{\partial \varphi'}{\partial \delta} \right|_{\tilde{\delta}} \right] \right)^{-1} \varphi(\tilde{\delta}) \xrightarrow{d} \chi^2(N)$$

Generalized Method of Moments (GMM) (Hamilton 1994)

Suppose $E[h(\theta_0, w_t)] = 0$

Objective function

$$Q_T(\theta, W) = \left(\frac{1}{T} \sum_{t=1}^T h(\theta, w_t) \right)' \Omega_T^{-1} \left(\frac{1}{T} \sum_{t=1}^T h(\theta, w_t) \right)$$

where Ω_T is a positive definite weighting matrix

GMM estimator

$$\hat{\theta} = \arg \min Q_T(\theta, W)$$

Properties of GMM estimator

$$\frac{1}{\sqrt{T}} \sum_t h(\theta_0, w_t) \xrightarrow{d} \mathcal{N}(0, \Omega) \text{ by CLT}$$

choose $\Omega_T = \Omega$

$$\frac{1}{T} \sum_t \frac{\partial h}{\partial \theta'} \xrightarrow{p} H'$$

$$\Rightarrow \sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, (H'\Omega^{-1}H)^{-1})$$

In practice choose $\Omega_T = \frac{1}{T} \sum_t h(\tilde{\theta}, w_t)h(\tilde{\theta}, w_t)'$
where $\tilde{\theta}$ is a consistent estimator of θ .

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics**
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

Standard Brownian Motion/Standard Wiener Process (Section C.8)

$\mathbf{W}(t)$ defined for $t \in [0, 1]$:

- ① $\mathbf{W}(0) = 0$ with probability one.
- ② $\mathbf{W}(t)$ is continuous in t with probability one.
- ③ For any partitioning of the unit interval, $0 \leq t_1 < t_2 < \dots < t_k \leq 1$, the vector

$$\begin{bmatrix} \mathbf{W}(t_2) - \mathbf{W}(t_1) \\ \vdots \\ \mathbf{W}(t_k) - \mathbf{W}(t_{k-1}) \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} t_2 - t_1 & & 0 \\ & \ddots & \vdots \\ 0 & \dots & t_k - t_{k-1} \end{bmatrix} \right),$$

that is, the differences have a multivariate normal distribution with independent components, means of zero, and variances $t_i - t_{i-1}$.

Nonstandard version: $Z(t) = \sigma \mathbf{W}(t)$

Functional Central Limit Theorem (FCLT)

Quantities of interests

$$X_T(r) = \frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} u_t, \quad r \in [0, 1]$$

where u_t is i.i.d. $(0, \sigma_u^2)$

and $\lfloor Tr \rfloor$ largest integer less than or equal to Tr

A limit result

$$\sqrt{T}X_T(r) = \frac{\sqrt{\lfloor Tr \rfloor}}{\sqrt{T}} \frac{1}{\sqrt{\lfloor Tr \rfloor}} \sum_{t=1}^{\lfloor Tr \rfloor} u_t \xrightarrow{d} \mathcal{N}(0, r\sigma_u^2)$$

FCLT

$$\sqrt{T}X_T(\cdot)/\sigma_u \xrightarrow{d} \mathbf{W}(\cdot)$$

Results on Random Functions

Convergence of random function

$G_T(\cdot)$, $G(\cdot)$ random functions.

$$G_T \xrightarrow{p} G \quad \text{iff} \quad \sup_{t \in [0,1]} |G_T(t) - G(t)| \xrightarrow{p} 0$$

Continuous Mapping Theorem

$\{G_T(\cdot)\}$, $G(\cdot)$ random functions and $g(\cdot)$ a continuous function defined on a space of functions:

$$G_T \xrightarrow{d} G \quad \Rightarrow \quad g(G_T) \xrightarrow{d} g(G)$$

Example

$$\int_0^1 \sqrt{T} X_T(r) dr \xrightarrow{d} \sigma_u \int_0^1 \mathbf{W}(r) dr$$

Properties of Random Walks

$$x_t = x_{t-1} + u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma_u^2), \quad x_0 = 0$$

- ① $T^{-1/2} \sum_{t=1}^T u_t \xrightarrow{d} \sigma_u \mathbf{W}(1) = \mathcal{N}(0, \sigma_u^2).$
- ② $T^{-1} \sum_{t=1}^T x_{t-1} u_t \xrightarrow{d} \frac{1}{2} \sigma_u^2 [\mathbf{W}(1)^2 - 1] = \frac{1}{2} \sigma_u^2 [\chi^2(1) - 1].$
- ③ $T^{-3/2} \sum_{t=1}^T t u_t \xrightarrow{d} \sigma_u \mathbf{W}(1) - \sigma_u \int_0^1 \mathbf{W}(r) dr = \mathcal{N}(0, \sigma_u^2/3).$
- ④ $T^{-3/2} \sum_{t=1}^T x_{t-1} \xrightarrow{d} \sigma_u \int_0^1 \mathbf{W}(r) dr = \mathcal{N}(0, \sigma_u^2/3).$
- ⑤ $T^{-2} \sum_{t=1}^T x_{t-1}^2 \xrightarrow{d} \sigma_u^2 \int_0^1 \mathbf{W}(r)^2 dr.$
- ⑥ $T^{-5/2} \sum_{t=1}^T t x_{t-1} \xrightarrow{d} \sigma_u \int_0^1 r \mathbf{W}(r) dr.$
- ⑦ $T^{-3} \sum_{t=1}^T t x_{t-1}^2 \xrightarrow{d} \sigma_u^2 \int_0^1 r \mathbf{W}(r)^2 dr.$
- ⑧ $T^{-(n+1)} \sum_{t=1}^T t^n \rightarrow 1/(n+1)$ for $n = 0, 1, \dots$

Dickey-Fuller Tests I

$\rho = 1$, $\hat{\rho}$ OLS estimator, $u_t \sim \text{i.i.d.}(0, \sigma_u^2)$

$$y_t = \rho y_{t-1} + u_t$$

$$T(\hat{\rho} - 1) \xrightarrow{d} \frac{\frac{1}{2}[\mathbf{W}(1)^2 - 1]}{\int_0^1 \mathbf{W}(r)^2 dr}$$

$$t_{\hat{\rho}-1} = \frac{\hat{\rho}-1}{\widehat{\sigma}_{\hat{\rho}}} \xrightarrow{d} \frac{\frac{1}{2}[\mathbf{W}(1)^2 - 1]}{\left[\int_0^1 \mathbf{W}(r)^2 dr\right]^{1/2}}$$

Dickey-Fuller Tests II

$$y_t = \mu + x_t, \quad x_t = \rho x_{t-1} + u_t$$

$$T(\hat{\rho} - 1) \xrightarrow{d} \frac{\frac{1}{2}[\mathbf{W}(1)^2 - 1] - \mathbf{W}(1) \int_0^1 \mathbf{W}(r) dr}{\int_0^1 \mathbf{W}(r)^2 dr - \left[\int_0^1 \mathbf{W}(r) dr \right]^2}$$

$$t_{\hat{\rho}-1} = \frac{\hat{\rho}-1}{\hat{\sigma}_{\hat{\rho}}} \xrightarrow{d} \frac{\frac{1}{2}[\mathbf{W}(1)^2 - 1] - \mathbf{W}(1) \int_0^1 \mathbf{W}(r) dr}{\left\{ \int_0^1 \mathbf{W}(r)^2 dr - \left[\int_0^1 \mathbf{W}(r) dr \right]^2 \right\}^{1/2}}$$

$$y_t = \nu + y_{t-1} + u_t, \quad \nu \neq 0$$

$$T^{3/2}(\hat{\rho} - 1) \xrightarrow{d} \mathcal{N}(0, 12\sigma_u^2/\nu^2)$$

$$t_{\hat{\rho}-1} = \frac{\hat{\rho}-1}{\hat{\sigma}_{\hat{\rho}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Dickey-Fuller Tests III

$$y_t = \mu_0 + \mu_1 t + x_t, \quad x_t = \rho x_{t-1} + u_t$$

$$T(\hat{\rho} - 1) \xrightarrow{d} a/b$$

$$t_{\hat{\rho}-1} = \frac{\hat{\rho}-1}{\hat{\sigma}_{\hat{\rho}}} \xrightarrow{d} a/\sqrt{b}$$

Extension to AR(p)

Rewrite $y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t$ as

$$y_t = \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t$$

or

$$\Delta y_t = (\rho - 1)y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t$$

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics**
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

Standard Brownian Motion/Standard Wiener Process (Section C.8.2)

$\mathbf{W}(t)$ ($K \times 1$) defined for $t \in [0, 1]$:

- ① $\mathbf{W}(0) = 0$ with probability one.
- ② A realization $\mathbf{W}(t)$ is a continuous function in t on the unit interval with probability one.
- ③ For any partitioning of the unit interval, $0 \leq t_1 < t_2 < \dots < t_k \leq 1$, the vector

$$\begin{bmatrix} \mathbf{W}(t_2) - \mathbf{W}(t_1) \\ \vdots \\ \mathbf{W}(t_k) - \mathbf{W}(t_{k-1}) \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} (t_2 - t_1)I_K & & 0 \\ & \ddots & \vdots \\ 0 & \dots & (t_k - t_{k-1})I_K \end{bmatrix} \right)$$

that is, the differences have multivariate normal distributions with independent components, means of zero, and variances of the form $t_i - t_{i-1}$, depending on their difference in time.

(Non-standard) Brownian Motion/Wiener Process

Definition

$$Z(t) := P\mathbf{W}(t) \text{ with } P (K \times K)$$

Properties

$$Z(t) - Z(s) \sim \mathcal{N}(0, (t-s)PP') \text{ for } s < t$$

$$Z(t) \sim \mathcal{N}(0, tPP')$$

Results on Multivariate Random Functions

Convergence of random function

$G_T(\cdot)$, $G(\cdot)$ multivariate random functions

$$G_T \xrightarrow{P} G \text{ iff } \sup_{t \in [0,1]} \|G_T(t) - G(t)\| \xrightarrow{P}_{T \rightarrow \infty} 0$$

FCLT

$$X_T(r) = \frac{1}{T} \sum_{t=1}^{\lfloor Tr \rfloor} u_t \quad (K \times 1), \quad u_t \sim \text{i.i.d.}(0, \Sigma_u)$$

$$\Rightarrow \sqrt{T} \Sigma_u^{-1/2} X_T(\cdot) \xrightarrow{d} \mathbf{W}(\cdot)$$

Beveridge-Nelson decomposition

$x_t = x_{t-1} + w_t$, where

$w_t = \Xi(L)u_t = \sum_{j=0}^{\infty} \Xi_j u_{t-j}$, with $\sum_{j=0}^{\infty} j \|\Xi_j\| < \infty$
and $u_t \sim (0, \Sigma_u = (\sigma_{ij}))$ white noise

$\Rightarrow x_t = x_0 + w_1 + \dots + w_t = x_0 + \Xi(1) \sum_{s=1}^t u_s + \sum_{j=0}^{\infty} \Xi_j^* u_{t-j} - w_0^*$

where

- $\Xi(1) = \sum_{j=0}^{\infty} \Xi_j$
- $\Xi_j^* = -\sum_{i=j+1}^{\infty} \Xi_i$, $j = 0, 1, \dots$
- $w_0^* = \sum_{j=0}^{\infty} \Xi_j^* u_{-j}$

Define $\Lambda = \Xi(1)P$ with P such that $PP' = \Sigma_u$.

Properties of Multivariate Unit Root Processes I

- 1 $T^{-1/2} \sum_{t=1}^T w_t \xrightarrow{d} \Lambda \mathbf{W}(1).$
- 2 $T^{-1/2} \sum_{t=1}^T W_t u_{it} \xrightarrow{d} \mathcal{N}(0, \sigma_{ii} \Sigma_W)$ for $i = 1, \dots, K.$
- 3 $T^{-1} \sum_{t=1}^T w_t w'_{t-h} \xrightarrow{P} \Gamma_w(h)$ for $h = 0, 1, 2, \dots.$
- 4 $T^{-1} \sum_{t=1}^T (x_{t-1} w'_{t-h} + w_{t-h} x'_{t-1})$
 $\xrightarrow{d} \begin{cases} \Lambda \mathbf{W}(1) \mathbf{W}(1)' \Lambda' - \Gamma_w(0) & \text{for } h = 0, \\ \Lambda \mathbf{W}(1) \mathbf{W}(1)' \Lambda' - \Gamma_w(0) + \sum_{j=-h+1}^{h-1} \Gamma_w(j) & \text{for } h = 1, 2, \dots \end{cases}$
- 5 $T^{-1} \sum_{t=1}^T x_{t-1} w'_t \xrightarrow{d} \Lambda \left\{ \int_0^1 \mathbf{W}(r) d\mathbf{W}(r)' \right\} \Lambda' + \sum_{j=1}^{\infty} \Gamma_w(j).$
- 6 $T^{-1} \sum_{t=1}^T x_{t-1} u'_t \xrightarrow{d} \Lambda \left\{ \int_0^1 \mathbf{W}(r) d\mathbf{W}(r)' \right\} P'.$
- 7 $T^{-3/2} \sum_{t=1}^T x_{t-1} \xrightarrow{d} \Lambda \int_0^1 \mathbf{W}(r) dr.$
- 8 $T^{-3/2} \sum_{t=1}^T t w_{t-h} \xrightarrow{d} \Lambda \left\{ \mathbf{W}(1) - \int_0^1 \mathbf{W}(r) dr \right\}$ for $h = 0, 1, 2, \dots.$

Properties of Multivariate Unit Root Processes II

- 9 $T^{-2} \sum_{t=1}^T x_{t-1} x'_{t-1} \xrightarrow{d} \Lambda \left\{ \int_0^1 \mathbf{W}(r) \mathbf{W}(r)' dr \right\} \Lambda'.$
- 10 $T^{-5/2} \sum_{t=1}^T t x_{t-1} \xrightarrow{d} \Lambda \int_0^1 r \mathbf{W}(r) dr.$
- 11 $T^{-3} \sum_{t=1}^T t x_{t-1} x'_{t-1} \xrightarrow{d} \Lambda \left\{ \int_0^1 r \mathbf{W}(r) \mathbf{W}(r)' dr \right\} \Lambda'.$

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model**
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

OLS Estimation (Section 7.1) I

The model

$$\Delta y_t = \mathbf{\Pi} y_{t-1} + u_t = \alpha \beta' y_{t-1} + u_t$$

$$\text{rk}(\alpha) = \text{rk}(\beta) = r \neq 0$$

OLS estimator

$$\hat{\mathbf{\Pi}} = \left(\sum_{t=1}^T \Delta y_t y_{t-1}' \right) \left(\sum_{t=1}^T y_{t-1} y_{t-1}' \right)^{-1}$$

Property 1

$$\hat{\mathbf{\Pi}} - \mathbf{\Pi} = \left(\sum_{t=1}^T u_t y_{t-1}' \right) \left(\sum_{t=1}^T y_{t-1} y_{t-1}' \right)^{-1}$$

OLS Estimation (Section 7.1) II

Transformation

$$Q := \begin{bmatrix} \beta' \\ \alpha'_{\perp} \end{bmatrix}, \quad Q^{-1} = [\alpha(\beta'\alpha)^{-1} : \beta_{\perp}(\alpha'_{\perp}\beta_{\perp})^{-1}]$$

$$v_t := Qu_t, \quad z_t := Qy_t$$

$$\Rightarrow \Delta z_t = Q\Pi Q^{-1}z_{t-1} + v_t = \begin{bmatrix} \beta'\alpha & 0 \\ 0 & 0 \end{bmatrix} z_{t-1} + v_t$$

Transformed estimator

$$Q(\hat{\Pi} - \Pi)Q^{-1} =$$

$$\begin{bmatrix} \sum_{t=1}^T v_t z_{t-1}^{(1)'} & \sum_{t=1}^T v_t z_{t-1}^{(2)'} \end{bmatrix} \begin{bmatrix} \sum_t z_{t-1}^{(1)} z_{t-1}^{(1)'} & \sum_t z_{t-1}^{(1)} z_{t-1}^{(2)'} \\ \sum_t z_{t-1}^{(2)} z_{t-1}^{(1)'} & \sum_t z_{t-1}^{(2)} z_{t-1}^{(2)'} \end{bmatrix}^{-1}$$

Lemma 7.1 I

$$\textcircled{1} \quad T^{-1} \sum_{t=1}^T z_{t-1}^{(1)} z_{t-1}^{(1)'} = T^{-1} \sum_{t=1}^T \beta' y_{t-1} y_{t-1}' \beta \xrightarrow{p} \Gamma_z^{(1)}.$$

$$\textcircled{2} \quad T^{-1/2} \text{vec} \left(\sum_{t=1}^T v_t z_{t-1}^{(1)'} \right) \xrightarrow{d} \mathcal{N}(0, \Gamma_z^{(1)} \otimes \Sigma_v),$$

where $\Sigma_v := Q \Sigma_u Q'$ is the covariance matrix of v_t .

$$\textcircled{3} \quad T^{-1} \sum_{t=1}^T v_t z_{t-1}^{(2)'} \xrightarrow{d} \Sigma_v^{1/2} \left(\int_0^1 \mathbf{W}_K d\mathbf{W}'_K \right)' \Sigma_v^{1/2} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix},$$

where \mathbf{W}_K abbreviates a standard Wiener process $\mathbf{W}_K(s)$ of dimension K .

Lemma 7.1 II

$$\bullet 4 \quad T^{-3/2} \sum_{t=1}^T z_{t-1}^{(1)} z_{t-1}^{(2)'} \xrightarrow{p} 0.$$

$$\bullet 5 \quad T^{-2} \sum_{t=1}^T z_{t-1}^{(2)} z_{t-1}^{(2)'} \xrightarrow{d} [0 : I_{K-r}] \Sigma_v^{1/2} \left(\int_0^1 \mathbf{w}_K \mathbf{w}_K' ds \right) \Sigma_v^{1/2} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix}.$$

The quantities in (2), (3), and (5) converge jointly.

Result 1

$$\text{Let } D = \begin{bmatrix} T^{1/2}I_r & 0 \\ 0 & TI_{K-r} \end{bmatrix}.$$

Then

$$\text{vec}[Q(\hat{\boldsymbol{\pi}} - \boldsymbol{\pi})Q^{-1}D]$$

$$\xrightarrow{d} \begin{bmatrix} \mathcal{N}(0, (\Gamma_z^{(1)})^{-1} \otimes \Sigma_v) \\ \text{vec} \left\{ \Sigma_v^{1/2} \left(\int_0^1 \mathbf{w}_K d\mathbf{w}'_K \right)' \Sigma_v^{1/2} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix} \right. \\ \left. \times \left([0 : I_{K-r}] \Sigma_v^{1/2} \left(\int_0^1 \mathbf{w}_K \mathbf{w}'_K ds \right) \Sigma_v^{1/2} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix} \right)^{-1} \right\} \end{bmatrix}.$$

Result 2

The estimator $\hat{\boldsymbol{\Pi}}$ is asymptotically normal,

$$\sqrt{T} \text{vec}(\hat{\boldsymbol{\Pi}} - \boldsymbol{\Pi}) \xrightarrow{d} \mathcal{N}\left(0, \beta(\Gamma_z^{(1)})^{-1} \beta' \otimes \Sigma_u\right),$$

and $\beta(\Gamma_z^{(1)})^{-1} \beta'$ can be estimated consistently by

$$\left(T^{-1} \sum_{t=1}^T y_{t-1} y_{t-1}' \right)^{-1}.$$

A testing problem

Hypotheses

$$H_0 : \boldsymbol{\pi} = 0 \quad \text{versus} \quad H_1 : \boldsymbol{\pi} \neq 0$$

Test statistic

$$\lambda_W = T \text{vec}(\hat{\boldsymbol{\pi}})' \left(\left(T^{-1} \sum_{t=1}^T y_{t-1} y_{t-1}' \right) \otimes \hat{\boldsymbol{\Sigma}}_u^{-1} \right) \text{vec}(\hat{\boldsymbol{\pi}})$$

Problem

$$\lambda_W \rightsquigarrow \chi^2(K^2)$$

Overview

- 1 Session 1. Review of Basic Definitions and Results
 - Multivariate Normal and Related Distributions
 - Concepts of Stochastic Convergence
 - Order in Probability
 - Infinite Sums of Random Variables
- 2 Session 2. Laws of Large Numbers and Central Limit Theorems
 - Weak Laws of Large Numbers
 - Central Limit Theorems
 - Properties of OLS
- 3 Session 3. Estimation and Testing
 - Maximum Likelihood Estimation
 - Testing Principles
 - GMM Estimation
- 4 Session 4. Unit Root Asymptotics
- 5 Session 5. Multivariate Unit Root Asymptotics
- 6 Session 6. OLS Estimation of the Cointegrated VAR(1) Model
- 7 Session 7. Other Estimators of the Cointegrated VAR(1) Model

OLS estimation of α given β

The estimator

$$\hat{\alpha} = \left(\sum_{t=1}^T \Delta y_t y'_{t-1} \beta \right) \left(\sum_{t=1}^T \beta' y_{t-1} y'_{t-1} \beta \right)^{-1}$$

Asymptotic properties

$$\sqrt{T} \text{vec}(\hat{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}(0, (\Gamma_z^{(1)})^{-1} \otimes \Sigma_u)$$

and, thus,

$$\sqrt{T} \text{vec}(\hat{\alpha} \beta' - \boldsymbol{\Pi}) \xrightarrow{d} \mathcal{N}(0, \beta (\Gamma_z^{(1)})^{-1} \beta' \otimes \Sigma_u)$$

GLS estimation of β given α I

Normalization of β

$$\beta = \begin{bmatrix} I_r \\ \beta_{(K-r)} \end{bmatrix}$$

The model

$$\Delta y_t - \alpha y_{t-1}^{(1)} = \alpha \beta'_{(K-r)} y_{t-1}^{(2)} + u_t = (y_{t-1}^{(2)'} \otimes \alpha) \text{vec}(\beta'_{(K-r)}) + u_t$$

GLS estimator

$$\begin{aligned} & \widehat{\beta}'_{(K-r)} \\ &= (\alpha' \Sigma_u^{-1} \alpha)^{-1} \alpha' \Sigma_u^{-1} \\ & \quad \times \left(\sum_{t=1}^T (\Delta y_t - \alpha y_{t-1}^{(1)}) y_{t-1}^{(2)'} \right) \left(\sum_{t=1}^T y_{t-1}^{(2)} y_{t-1}^{(2)'} \right)^{-1} \end{aligned}$$

GLS estimation of β given α II

Properties I

$$T(\widehat{\beta}'_{(K-r)} - \beta'_{(K-r)})$$

$$\xrightarrow{d} \left(\int_0^1 \mathbf{W}_{K-r}^* d\mathbf{W}_r^{*'} \right)' \left(\int_0^1 \mathbf{W}_{K-r}^* \mathbf{W}_{K-r}^{*'} ds \right)^{-1}$$

- $\mathbf{W}_{K-r}^* := Q^{22}[0 : I_{K-r}] \Sigma_v^{1/2} \mathbf{W}_K$
- Q^{22} denotes the lower right-hand $((K-r) \times (K-r))$ block of Q^{-1}
- $\mathbf{W}_r^* := (\alpha' \Sigma_u^{-1} \alpha)^{-1} \alpha' \Sigma_u^{-1} Q^{-1} \Sigma_v^{1/2} \mathbf{W}_K$

Properties II

$$\text{vec} \left[\left(\widehat{\beta}'_{(K-r)} - \beta'_{(K-r)} \right) \left(\sum_{t=1}^T y_{t-1}^{(2)} y_{t-1}^{(2)'} \right)^{1/2} \right]$$

$$\xrightarrow{d} \mathcal{N} \left(0, I_{K-r} \otimes (\alpha' \Sigma_u^{-1} \alpha)^{-1} \right)$$

EGLS estimation of β

EGLS estimator

$$\begin{aligned} \widehat{\beta}'_{(K-r)} = & \\ & (\widehat{\alpha}' \widehat{\Sigma}_u^{-1} \widehat{\alpha})^{-1} \widehat{\alpha}' \widehat{\Sigma}_u^{-1} \left(\sum_{t=1}^T (\Delta y_t - \widehat{\alpha} y_{t-1}^{(1)}) y_{t-1}^{(2)'} \right) \\ & \times \left(\sum_{t=1}^T y_{t-1}^{(2)} y_{t-1}^{(2)'} \right)^{-1} \end{aligned}$$

Properties

$$T(\widehat{\beta}'_{(K-r)} - \beta'_{(K-r)}) = o_p(1)$$

ML estimation I

log likelihood function

$$\ln l = -\frac{KT}{2} \ln 2\pi - \frac{T}{2} \ln |\Sigma_u|$$

$$-\frac{1}{2} \sum_{t=1}^T (\Delta y_t - \mathbf{\Pi} y_{t-1})' \Sigma_u^{-1} (\Delta y_t - \mathbf{\Pi} y_{t-1})$$

Equivalent optimization problem

$$\text{minimize } \left| T^{-1} \sum_{t=1}^T (\Delta y_t - \alpha \beta' y_{t-1}) (\Delta y_t - \alpha \beta' y_{t-1})' \right|$$

ML estimation II

ML estimator

$\lambda_1 \geq \dots \geq \lambda_K$ eigenvalues with associated orthonormal eigenvectors ν_1, \dots, ν_K of the matrix

$$\left(\sum_{t=1}^T y_{t-1} y_{t-1}' \right)^{-1/2} \left(\sum_{t=1}^T y_{t-1} \Delta y_t' \right) \left(\sum_{t=1}^T \Delta y_t \Delta y_t' \right) \\ \times \left(\sum_{t=1}^T \Delta y_t y_{t-1}' \right) \left(\sum_{t=1}^T y_{t-1} y_{t-1}' \right)^{-1/2}$$

$$\tilde{\beta} = [\nu_1, \dots, \nu_r]' \left(\sum_{t=1}^T y_{t-1} y_{t-1}' \right)^{-1/2}$$

$$\tilde{\alpha} = \left(\sum_{t=1}^T \Delta y_t y_{t-1}' \tilde{\beta} \right) \left(\sum_{t=1}^T \tilde{\beta}' y_{t-1} y_{t-1}' \tilde{\beta} \right)^{-1}$$

Properties

$$\sqrt{T} \text{vec}(\tilde{\alpha} \tilde{\beta}' - \Pi) \xrightarrow{d} \mathcal{N}(0, \beta(\Gamma_z^{(1)})^{-1} \beta' \otimes \Sigma_u)$$

ML estimation III

If $\beta' = [I_r : \beta'_{(K-r)}]$ and $\check{\beta}$ corresponding ML estimator

- $$\check{\alpha} = \left(\sum_{t=1}^T \Delta y_t y_{t-1}' \check{\beta} \right) \left(\sum_{t=1}^T \check{\beta}' y_{t-1} y_{t-1}' \check{\beta} \right)^{-1}$$

$$\sqrt{T} \text{vec}(\check{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}(0, (\Gamma_z^{(1)})^{-1} \otimes \Sigma_u)$$
- $$\check{\beta}'_{(K-r)}$$

$$= (\check{\alpha}' \tilde{\Sigma}_u^{-1} \check{\alpha})^{-1} \check{\alpha}' \tilde{\Sigma}_u^{-1}$$

$$\times \left(\sum_{t=1}^T (\Delta y_t - \check{\alpha} y_{t-1}^{(1)}) y_{t-1}^{(2)'} \right) \left(\sum_{t=1}^T y_{t-1}^{(2)} y_{t-1}^{(2)'} \right)^{-1}$$

$$T(\check{\beta}'_{(K-r)} - \beta'_{(K-r)})$$

$$\xrightarrow{d} \left(\int_0^1 \mathbf{W}_{K-r}^* d\mathbf{W}_r^{*'} \right)' \left(\int_0^1 \mathbf{W}_{K-r}^* \mathbf{W}_{K-r}^{*'} ds \right)^{-1}$$