Optimal tests for the absence of random individual effects in large n and small T dynamic panels

Yuichi Goto Kyushu University

Nezar Bennala Université libre de Bruxelles

Marc Hallin Université libre de Bruxelles

Time Series, Random Fields and beyond

Fall School

September 23 -- 27, 2024

Ulm University

Introduction

Panel Data (or longitudinal data):

Panel data are multi- or high- dimensional time series in econometrics.

Dynamic panel data models with serial correlation (Lillard & Willis (1978)):

$$y_{it} = x_{it}\beta + \mu + \alpha_i + e_{it} \qquad i = 1, \ldots, n, t = 1, \ldots, T,$$

where x_{it} : an explanatory variable,

α_i : a fixed or random effect,

e_{it} : a time series.

Introduction

Many studies for panel data analysis focus on
estimation or testing for $\boldsymbol{\beta}$.

local asymptotic normality (LAN) base-optimal tests for the existence of
- random autoregressive coefficients of AR(p): Akharif & Hallin 2003
- random coefficients of regression: Fihri et al. 2020, Akharif et al. 2020
- random individual effects of panel data: Bennala et al (2012)

A test for the existence of random effects in time series one- & two-way models: Goto et al. 2020 (a,b)

$$\boldsymbol{y}_{it} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{e}_{it}, \qquad i = 1, \ldots, n \text{ (fixed)}; \; t = 1, \ldots, T_i \; (\to \infty),$$

For i.i.d. disturbances, local asymptotic normality does not hold under the asymptotic regime $T_i \to \infty, n : \text{fixed}$: Goto et al. 2022

→ We would like develop an optimal test for the existence of individual effects in dynamic panel data

Introduction

Local asymptotic normality is introduced by LeCam (1960)
And this concept plays a vital role in optimal inference and testing.

Roussas (1972) showed LAN for Markov process.
Roussas (1979) extended Roussas (1972) to non-Markov process.

Kreiss (1987): LAN for ARMA model
Linton (1993): LAN for ARCH model
Garel and Hallin (1995): LAN for multiple-output linear model with VARMA error term
Benghabrit and Hallin (1996): LAN for bilinear models
Hallin, Taniguchi, Serroukh, & Choy (1999): LAN for regression model with long memory error
Kato, Taniguchi, & Honda (2006): LAN for CHARN model
Cutting, Paindaveine, & Verdebout (2017): LAN for rotationally symmetric densities (directional statistics)

Books: Roussas (1972), Ibragimov & Khasminskii (1981), Taniguchi & Kakizawa (2000), Ley and Verdebout (2017)

Introduction

For a sequence of statistical experiments $\mathcal{E}^{(n)} = \left\{\mathcal{X}^{(n)}, \mathcal{F}^{(n)}, \left\{P_{\boldsymbol{\theta}}^{(n)} : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p\right\}\right\}$

the statistical experiments is local asymptotic normal (LAN) if

there exists, $\boldsymbol{\Delta}^{(n)}$ such that $\boldsymbol{\Delta}^{(n)} \Rightarrow N(\mathbf{0}, \boldsymbol{I}(\boldsymbol{\theta}))$ under $P_{n,\boldsymbol{\theta}}$

and for every sequence $\mathbf{h}_n \to \mathbf{h}$,

$$\Lambda(\boldsymbol{\theta}, \boldsymbol{\theta}_n) = \log \frac{\mathrm{d}P_{n,\boldsymbol{\theta}_n}}{\mathrm{d}P_{n,\boldsymbol{\theta}}} = \mathbf{h}^T \boldsymbol{\Delta}^{(n)} - \frac{1}{2}\mathbf{h}^T \boldsymbol{I}(\boldsymbol{\theta})\mathbf{h} + o_p(1) \quad \text{as } n \to \infty \text{ under } P_{n,\boldsymbol{\theta}},$$

where $\boldsymbol{I}(\boldsymbol{\theta})$ is a Fisher information and $\boldsymbol{\theta}_n = \boldsymbol{\theta} + \mathbf{h}/\sqrt{n}$.

$\boldsymbol{\Delta}^{(n)}$ is called the central sequence

$\hat{\boldsymbol{\theta}}_n$ is said to be an efficient estimator (at $\boldsymbol{\theta}$)

if $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) - \boldsymbol{I}(\boldsymbol{\theta})^{-1}\boldsymbol{\Delta}^{(n)} = o_{P_{n,\boldsymbol{\theta}}}(1)$

Introduction

Once the LAN property is shown,

we can construct,

$$I(\theta_0)^{-1/2}\Delta^{(n)} > z_{1-\alpha}$$

(i, univariate) for the null $\theta \leq \theta_0$ and the alternative $\theta > \theta_0$,

a locally and asymptotically uniformly most powerful test

and

(ii) for the null $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and the alternative $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$,

a locally and asymptotically maximin test

by using the central sequence

$$\boldsymbol{\Delta}^{(n)}\boldsymbol{I}(\boldsymbol{\theta}_0)^{-1}\boldsymbol{\Delta}^{(n)} > \chi^2_p[1-\alpha]$$

Regarding to estimation, it holds that

an asymptotically centering estimator is efficient among the regular estimators.

Main

Dynamic panel data model with AR(p) disturbances

$$y_{it} = \mu + \boldsymbol{\beta}^\top \mathbf{x}_{it} + \sigma_u u_i + \nu_{it} \qquad i = 1, \ldots, n, \quad t = 1, \ldots, T$$

with $\nu_{it} = \sum_{j=1}^{p} \rho_j \nu_{i(t-j)} + \sigma_\epsilon \varepsilon_{it}$

where

$\mathbf{x_{it}}$: explanatory variables

u_i: individual effect with pdf h

ε_{it}: disturbance with pdf f_1

Suppose $\mathrm{E}(u_i) = \mathrm{E}(\epsilon_{it}) = 0$, $\mathrm{Var}(u_i) = \mathrm{Var}(\epsilon_{it}) = 1$ (for identifiability)

f_1 is possibly non-differentiable but $f_1^{1/2}$ is quadratic mean differentiable

Main

Test for the absence of individual effect

Null:

$$\mathcal{H}_{f_1}^{(n)} := \left\{ \mathrm{P}_{\mu,\boldsymbol{\beta},\boldsymbol{\rho},\sigma_\epsilon^2,0;f_1}{}^{(n)}; (\mu, \boldsymbol{\beta}, \boldsymbol{\rho}, \sigma_\epsilon^2) \in \mathbb{R}^{1+K+p} \times \mathbb{R}_+ \right\}$$

Alternative:

$$\mathcal{K}_{f_1}^{(n)} := \left\{ \mathrm{P}_{\mu,\boldsymbol{\beta},\boldsymbol{\rho},\sigma_\epsilon^2,\sigma_u^2;f_1}{}^{(n)}; (\mu, \boldsymbol{\beta}, \boldsymbol{\rho}, \sigma_\epsilon^2, \sigma_u^2) \in \mathbb{R}^{1+K+p} \times \mathbb{R}_+^2 \right\}$$

Consider

$$\boldsymbol{\vartheta}_0 := (\mu, \boldsymbol{\beta}^\top, \sigma_\epsilon, \boldsymbol{\rho}^\top, 0)^\top$$

and a local parameter sequence

$$\boldsymbol{\vartheta}_0 + n^{-1/2}\boldsymbol{\iota}^{(n)}\boldsymbol{\tau}^{(n)} = \boldsymbol{\vartheta}_0 + \frac{1}{\sqrt{n}} \begin{pmatrix} \tau_1^{(n)} \\ \mathbf{K}^{(n)}\boldsymbol{\tau}_2^{(n)} \\ \tau_3^{(n)} \\ \boldsymbol{\tau}_4^{(n)} \\ \tau_5^{(n)} \end{pmatrix}.$$

Main

Likelihood ratio (conditionally on $y_{i0}, \ldots, y_{i(1-p)}$ for all $i = 1, \ldots, p$)

$$\Lambda_{f_1}^{(n)} = \log \prod_{i=1}^{n} \prod_{t=1}^{T} \frac{\left(\gamma_{n,2} \int_{\mathbb{R}} f_1\left(\gamma_{n,1} u + \gamma_{n,2} Z_{it} + \gamma_{i,t,n,3}\right) h(u) \mathrm{d}u\right)}{f_1\left(Z_{it}\right)}$$

where $$Z_{it} := Z_{it}(\boldsymbol{\vartheta}_0) := \frac{y_{it} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{it} - \sum_{j=1}^{p} \rho_j (y_{it-j} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{i(t-j)})}{\sigma_\epsilon},$$

$$\gamma_{n,1} := \frac{\left(-1 + \sum_{j=1}^{p} \left(\rho_j + \frac{\tau_{4j}^{(n)}}{\sqrt{n}}\right)\right) \sqrt{\frac{\tau_5^{(n)}}{\sqrt{n}}}}{\sigma_\epsilon + \frac{\tau_3^{(n)}}{\sqrt{n}}}, \quad \gamma_{n,2} := \frac{\sigma_\epsilon}{\sigma_\epsilon + \frac{\tau_3^{(n)}}{\sqrt{n}}},$$

$$\gamma_{i,t,n,3} := \frac{-\tau_1^{(n)} - \boldsymbol{\tau}_2^{(n)\top} \mathbf{K}^{(n)\top} \mathbf{x}_{it} + \sum_{j=1}^{p} \left(\rho_j + \frac{\tau_{4j}^{(n)}}{\sqrt{n}}\right) \left(\tau_1^{(n)} + \boldsymbol{\tau}_2^{(n)\top} \mathbf{K}^{(n)\top} \mathbf{x}_{i(t-j)}\right)}{\sqrt{n}(\sigma_\epsilon + \frac{\tau_3^{(n)}}{\sqrt{n}})}$$
$$- \frac{\sum_{j=1}^{p} \tau_{4j}^{(n)} \left(y_{i(t-j)} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{i(t-j)}\right)}{\sqrt{n}(\sigma_\epsilon + \frac{\tau_3^{(n)}}{\sqrt{n}})}.$$

Main

Quadratic mean differentiability (qmd)

$$\frac{1}{h^2}\int_{\mathbb{R}}\left\{f_1^{1/2}(v+h)-f_1^{1/2}(v)-\frac{h}{2}\frac{\dot{f}_1(v)}{f_1^{1/2}(v)}\right\}^2 \mathrm{d}v = o_p(1),$$

where $\dot{f}_1(v)$: a.e. derivative of $f_1(v)$.

Multivariate qmd $\Leftrightarrow$ partial qmd (Lind & Roussas 1972)

Our case:

$v \to Z_{it}$

$h \to$ highly non-linear , even involving integral

Main

Theorem (Local asymptotic normality)

Under $P^{(n)}_{\boldsymbol{\theta}_0, f_1}$,

the following asymptotic expansion holds true

$$\begin{aligned}\Lambda^{(n)}_{f_1} =& \prod_{i=1}^{n}\prod_{t=1}^{T} \frac{\left(\gamma_{n,2}\int_{\mathbb{R}} f_1\left(\gamma_{n,1}u + \gamma_{n,2}Z_{it} + \gamma_{i,t,n,3}\right) h(u)\mathrm{d}u\right)}{f_1\left(Z_{it}\right)} \\ =& \boldsymbol{\tau}^{(n)^\top}\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta}_0) - \frac{1}{2}\boldsymbol{\tau}^{(n)^\top}\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta}_0)\boldsymbol{\tau}^{(n)} + o_{\mathrm{p}}(1) \quad \text{as } n \to \infty,\end{aligned}$$

and

$$\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta}_0) \Rightarrow N\left(\mathbf{0}, \boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta}_0)\right),$$

where

Main

$$\boldsymbol{\vartheta} := (\mu, \boldsymbol{\beta}^\top, \sigma_\epsilon, \boldsymbol{\rho}^\top, \sigma_u^2)^\top \text{ and } \boldsymbol{\vartheta}_0 := (\mu, \boldsymbol{\beta}^\top, \sigma_\epsilon, \boldsymbol{\rho}^\top, 0)^\top.$$

the central sequence

$$\boldsymbol{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta}_0) := \left(\Delta_{f_1;1}^{(n)}(\boldsymbol{\vartheta}_0), \boldsymbol{\Delta}_{f_1;2}^{(n)}(\boldsymbol{\vartheta}_0), \Delta_{f_1;3}^{(n)}(\boldsymbol{\vartheta}_0), \boldsymbol{\Delta}_{f_1;4}^{(n)}(\boldsymbol{\vartheta}_0), \Delta_{f_1;5}^{(n)}(\boldsymbol{\vartheta}_0)\right)^\top$$

$$:= \begin{pmatrix} \frac{(1-\sum_{j=1}^p \rho_j)}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \phi_{f_1}(Z_{it}) \\ \frac{1}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \mathbf{K}^{(n)\top} \left(\mathbf{x}_{it} - \sum_{j=1}^p \rho_j \mathbf{x}_{i(t-j)}\right) \phi_{f_1}(Z_{it}) \\ \frac{1}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (Z_{it}\phi_{f_1}(Z_{it}) - 1) \\ \frac{1}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (y_{i(t-1)} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{i(t-1)}, \ldots, y_{i(t-p)} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{i(t-p)})^\top \phi_{f_1}(Z_{it}) \\ \frac{(1-\sum_{j=1}^p \rho_j)^2}{2\sigma_\epsilon^2 \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \psi_{f_1}(Z_{it}) \end{pmatrix},$$

the Fisher information matrix

$$\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta}) := \begin{pmatrix} \Gamma_{f_1;11}(\boldsymbol{\vartheta}) & \mathbf{0}_K^\top & \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \mathbf{0}_p^\top & \Gamma_{f_1;15}(\boldsymbol{\vartheta}) \\ \mathbf{0}_K & \boldsymbol{\Gamma}_{f_1;22}(\boldsymbol{\vartheta}) & \mathbf{0}_K & \boldsymbol{O}_{K\times p} & \mathbf{0}_K \\ \Gamma_{f_1;13}(\boldsymbol{\vartheta}) & \mathbf{0}_K^\top & \Gamma_{f_1;33}(\boldsymbol{\vartheta}) & \mathbf{0}_p^\top & \Gamma_{f_1;35}(\boldsymbol{\vartheta}) \\ \mathbf{0}_p & \boldsymbol{O}_{p\times K} & \mathbf{0}_p & \boldsymbol{\Gamma}_{f_1;44}(\boldsymbol{\vartheta}) & \mathbf{0}_p \\ \Gamma_{f_1;15}(\boldsymbol{\vartheta}) & \mathbf{0}_K^\top & \Gamma_{f_1;35}(\boldsymbol{\vartheta}) & \mathbf{0}_p^\top & \Gamma_{f_1;55}(\boldsymbol{\vartheta}) \end{pmatrix}.$$

h does not appear in $\boldsymbol{\Delta}_{f_1}^{(n)}(\boldsymbol{\vartheta}_0)$ nor in $\boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta})$

Main

Unspecified parameter case:

σ_u^2-efficient central sequence

$$\Delta^{*(n)}_{f_1;5}(\boldsymbol{\vartheta}) := \Delta^{(n)}_{f_1;5}(\boldsymbol{\vartheta}) - \begin{pmatrix}\Gamma_{f_1,15}(\boldsymbol{\vartheta}) & \Gamma_{f_1,35}(\boldsymbol{\vartheta})\end{pmatrix} \begin{pmatrix}\Gamma_{f_1,11}(\boldsymbol{\vartheta}) & \Gamma_{f_1,13}(\boldsymbol{\vartheta}) \\ \Gamma_{f_1,13}(\boldsymbol{\vartheta}) & \Gamma_{f_1,33}(\boldsymbol{\vartheta})\end{pmatrix}^{-1} \begin{pmatrix}\Delta^{(n)}_{f_1;1}(\boldsymbol{\vartheta}) \\ \Delta^{(n)}_{f_1;3}(\boldsymbol{\vartheta})\end{pmatrix}$$

$$= \Delta^{(n)}_{f_1;5}(\boldsymbol{\vartheta}) - \Gamma^*_{f_1,15}(\boldsymbol{\vartheta})\Delta^{(n)}_{f_1;1}(\boldsymbol{\vartheta}) - \Gamma^*_{f_1,35}(\boldsymbol{\vartheta})\Delta^{(n)}_{f_1;3}(\boldsymbol{\vartheta}),$$

efficient information matrix

$$\Gamma^*_{f_1,55} := \Gamma_{f_1,55}(\boldsymbol{\vartheta}) - \Gamma^*_{f_1,15}(\boldsymbol{\vartheta})\Gamma_{f_1,15}(\boldsymbol{\vartheta}) - \Gamma^*_{f_1,35}(\boldsymbol{\vartheta})\Gamma_{f_1,35}(\boldsymbol{\vartheta})$$

Suppose that

$n^{1/2}\boldsymbol{\iota}^{(n)^{-1}}\left(\hat{\boldsymbol{\vartheta}}^{(n)} - \boldsymbol{\vartheta}_0\right) = O_p(1)$, where $\hat{\sigma}_u^2 = 0$, under $P^{(n)}_{\boldsymbol{\theta}_0,f_1}$.

$\hat{\boldsymbol{\vartheta}}^{(n)}$ is locally discrete (Kreiss 1987)

asymptoic linearity: for $\boldsymbol{\vartheta}_n$ such that $\left\| n^{1/2}\boldsymbol{\iota}^{(n)^{-1}}(\boldsymbol{\vartheta}_n - \boldsymbol{\vartheta}_0)\right\| \leq C$,

$\boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta}_n) - \boldsymbol{\Delta}^{(n)}_{f_1}(\boldsymbol{\vartheta}_0) + \boldsymbol{\Gamma}_{f_1}(\boldsymbol{\vartheta}) n^{1/2}\boldsymbol{\iota}^{(n)^{-1}}(\boldsymbol{\vartheta}_n - \boldsymbol{\vartheta}_0) = o_p(1)$ under $P^{(n)}_{\boldsymbol{\theta}_0,f_1}$

Main

A test statistic:

$$T^{*(n)}_{f_1}\left(\hat{\boldsymbol{\vartheta}}^{(n)}\right) := \left(\Gamma^{*}_{f_1,5}\left(\hat{\boldsymbol{\vartheta}}^{(n)}\right)\right)^{-1/2} \Delta^{*(n)}_{f_1;5}\left(\hat{\boldsymbol{\vartheta}}^{(n)}\right).$$

A test which rejects $\mathcal{H}^{(n)}_{f_1}$ whenever $T^{*(n)}_{f_1}\left(\hat{\boldsymbol{\vartheta}}^{(n)}\right) > \Phi^{-1}(1-\alpha)$is locally asymptotically most stringent, against $\mathcal{K}^{(n)}_{f_1}$.

Main

Pseudo-Gaussian test: f_1 is generally unknown

$\rightarrow$ use the central sequence for Gaussian density

Gaussian versions of the central sequence

$$\boldsymbol{\Delta}_{\mathcal{N}}^{(n)}(\boldsymbol{\vartheta}_0) := \left(\Delta_{\mathcal{N};1}^{(n)}(\boldsymbol{\vartheta}_0), \boldsymbol{\Delta}_{\mathcal{N};2}^{(n)}(\boldsymbol{\vartheta}_0), \Delta_{\mathcal{N};3}^{(n)}(\boldsymbol{\vartheta}_0), \boldsymbol{\Delta}_{\mathcal{N};4}^{(n)}(\boldsymbol{\vartheta}_0), \Delta_{\mathcal{N};5}^{(n)}(\boldsymbol{\vartheta}_0)\right)^\top$$

$$:= \begin{pmatrix} \frac{(1-\sum_{j=1}^p \rho_j)}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T Z_{it} \\ \frac{1}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T \mathbf{K}^{(n)\top} \left(\mathbf{x}_{it} - \sum_{j=1}^p \rho_j \mathbf{x}_{i(t-j)}\right) Z_{it} \\ \frac{1}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (Z_{it}^2 - 1) \\ \frac{1}{\sigma_\epsilon \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (y_{i(t-1)} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{i(t-1)}, \ldots, y_{i(t-p)} - \mu - \boldsymbol{\beta}^\top \mathbf{x}_{i(t-p)})^\top Z_{it} \\ \frac{\left(1-\sum_{j=1}^p \rho_j\right)^2}{2\sigma_\epsilon^2 \sqrt{n}} \sum_{i=1}^n \sum_{t=1}^T (Z_{it}^2 - 1) \end{pmatrix}$$

And its Fisher information under the unknown density g

$$\boldsymbol{\Gamma}_{N;g}(\boldsymbol{\vartheta}) := \begin{pmatrix} \Gamma_{\mathcal{N};g;11}(\boldsymbol{\vartheta}) & \mathbf{0}_K^\top & \Gamma_{\mathcal{N};g;13}(\boldsymbol{\vartheta}) & \mathbf{0}_p^\top & \Gamma_{\mathcal{N};g;15}(\boldsymbol{\vartheta}) \\ \mathbf{0}_K & \boldsymbol{\Gamma}_{\mathcal{N};g;22}(\boldsymbol{\vartheta}) & \mathbf{0}_K & \boldsymbol{O}_{K\times p} & \mathbf{0}_K \\ \Gamma_{\mathcal{N};g;13}(\boldsymbol{\vartheta}) & \mathbf{0}_K^\top & \Gamma_{\mathcal{N};g;33}(\boldsymbol{\vartheta}) & \mathbf{0}_p^\top & \Gamma_{\mathcal{N};g;35}(\boldsymbol{\vartheta}) \\ \mathbf{0}_p & \boldsymbol{O}_{p\times K} & \mathbf{0}_p & \boldsymbol{\Gamma}_{\mathcal{N};g;44}(\boldsymbol{\vartheta}) & \mathbf{0}_p \\ \Gamma_{\mathcal{N};g;15}(\boldsymbol{\vartheta}) & \mathbf{0}_K^\top & \Gamma_{\mathcal{N};g;35}(\boldsymbol{\vartheta}) & \mathbf{0}_p^\top & \Gamma_{\mathcal{N};g;55}(\boldsymbol{\vartheta}) \end{pmatrix}.$$

However, $\Delta_{\mathcal{N};3}^{(n)}(\boldsymbol{\vartheta}_0)$ and $\Delta_{\mathcal{N};5}^{(n)}(\boldsymbol{\vartheta}_0)$ are collinear...

Main

Possible reasons:

Conditional likelihood could be a problem:

$$\prod_{i=1}^{a} p_{y_{i1},\ldots,y_{iT}|y_{i0},\ldots,y_{i(1-p)}}\left(z_{i1},\ldots,z_{iT}|z_{i0},\ldots,z_{i(1-p)}\right)$$

$$=\prod_{i=1}^{a}\prod_{t=1}^{T} p_{y_{it}|y_{i(t-1)},\ldots,y_{i(1-p)}}\left(z_{it}|z_{i(t-1)},\ldots,z_{i(1-p)}\right)$$

$$=\frac{1}{\sigma_\epsilon^T}\prod_{i=1}^{a}\prod_{t=1}^{T}\int_{\mathbb{R}} h(u) f_1\left(\frac{z_{it}-\mu-\boldsymbol{\beta}^\top\mathbf{x}_{it}-\sum_{j=1}^{p}\rho_j(y_{it-j}-\mu-\boldsymbol{\beta}^\top\mathbf{x}_{i(t-j)})-\sigma_u(1-\sum_{j=1}^{p}\rho_j)u_i}{\sigma_\epsilon}\right)\mathrm{d}u$$

it seems that y_{it} is $y_{it'}$ independent even if u_i is non-degenerate.

Relation with (full) likelihood:

$$\prod_{i=1}^{a} p_{y_{i1},\ldots,y_{iT}}\left(z_{i1},\ldots,z_{iT}\right)$$

$$=\prod_{i=1}^{a} p_{y_{i0},\ldots,y_{i(1-p)}}\left(z_{i0},\ldots,z_{i(1-p)}\right)\prod_{t=1}^{T} p_{y_{i1},\ldots,y_{iT}|y_{i0},\ldots,y_{i(1-p)}}\left(z_{i1},\ldots,z_{iT}|z_{i0},\ldots,z_{i(1-p)}\right)$$

Main

Possible reasons:

Full likelihood, provided $y_{i0}, \cdots, y_{i(1-p)}$ as constants, works:

$$\prod_{i=1}^{a} p_{y_{i1},\ldots,y_{iT}}(z_{i1},\ldots,z_{iT})$$
$$= \frac{1}{\sigma_\epsilon^T}\prod_{i=1}^{a}\int_{\mathbb{R}} h(u)\prod_{t=1}^{T} f_1\left(\frac{z_{it}-\mu-\boldsymbol{\beta}^\top\mathbf{x}_{it}-\sum_{j=1}^{p}\rho_j(y_{it-j}-\mu-\boldsymbol{\beta}^\top\mathbf{x}_{i(t-j)})-\sigma_u(1-\sum_{j=1}^{p}\rho_j)u_i}{\sigma_\epsilon}\right)\mathrm{d}u$$

Then, the central sequence $\Delta_{\mathcal{N};5}^{(n)}(\boldsymbol{\vartheta}_0)$ is not collinear to $\Delta_{\mathcal{N};3}^{(n)}(\boldsymbol{\vartheta}_0)$.

The difference is how to deal with initial values (constants or random variables)

Future works:

LAN for full likelihood, prove asymptotic linearity, a rank-based test

→ Acquire LAN-based inference and test

Thank you so much
for your time and kind attention.