High-dimensional autocovariance matrices: theory and application

Daning Bi, Xiao Han, Adam Nie, Yanrong Yang

Research School of Finance, Actuarial Studies and Statistics
The Australian National University

Seminar at Macquarie University

January 25, 2023

Outline

1. Motivation and Challenge: Inference on High-dimensional Time Series (HDTS)
2. Literature Review: PCA, Factor Modelling and Random Matrix Theory
3. Major Contribution:
    1. Asymptotic Theory for Spiked Eigenvalues
    2. Statistical Application: Equivalence Test of Two High-dimensional Time Series
    3. Simulation
    4. Empirical Application: Hierarchical Clustering for Multi-country Mortality Data
4. Conclusion and Future Works
5. References

Why to study High-dimensional Autocovariance Matrices

HDTS (1): Mortality Data

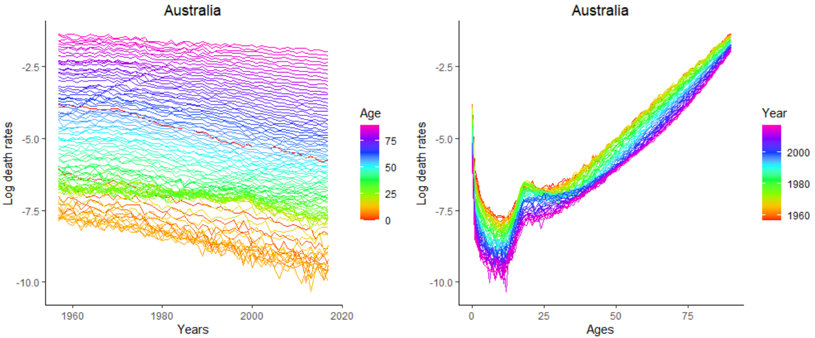


Figure 1: Log death rates for Australian

HDTS (2): Stock Returns

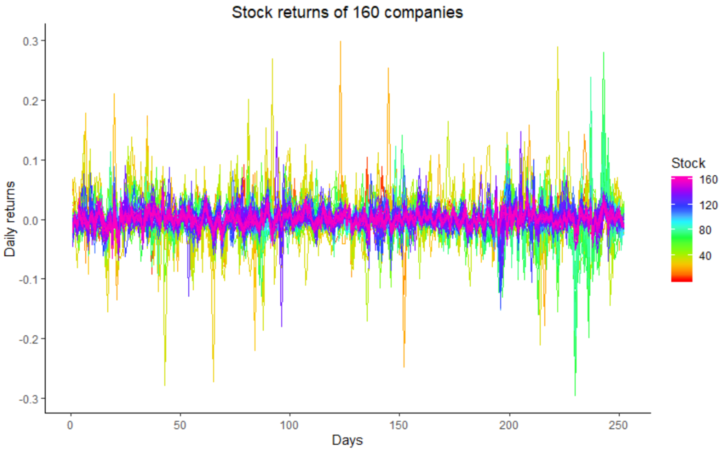


Figure 2: Daily returns of 160 US stocks in 2014

Challenges of HDTS Inference (1)

The major difficulty: curse of dimensionality.

Example:

For the population covariance matrix Σ (a $p \times p$ matrix), i.e.

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$$

the sample covariance matrix estimator $\widehat{\Sigma}$ is inaccurate in the sense of

$$\left\|\widehat{\Sigma} - \Sigma\right\|_F^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} (\widehat{\sigma}_{ij} - \sigma_{ij})^2 \asymp \frac{p^2}{T}.$$

Curse appears: when $T = O(p^2)$, $\left\|\widehat{\Sigma} - \Sigma\right\|_F^2$ does not converge to zero.

Challenges of HDTS Inference (2)

1. Common approaches to curse of dimensionality: (1) dimension reduction (2) variable selection.

   Example: dimension reduction projects a p-dimensional vector $\mathbf{y}_t$ into a K-dimensional subspace.

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{pt} \end{pmatrix} = \begin{pmatrix} \ell_{11} \\ \ell_{21} \\ \vdots \\ \ell_{pt} \end{pmatrix} \cdot f_{1t} + \begin{pmatrix} \ell_{12} \\ \ell_{22} \\ \vdots \\ \ell_{p2} \end{pmatrix} \cdot f_{2t} + \cdots + \begin{pmatrix} \ell_{1K} \\ \ell_{2K} \\ \vdots \\ \ell_{pK} \end{pmatrix} \cdot f_{Kt} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{pt} \end{pmatrix} \tag{0.1}$$

2. PCA: purse the subspace where the projected data holds the most variation of the original data.

3. Extra challenge on HDTS: the projected data from PCA may lose time-serial dependence.

Dimension Reduction based on Autocovariance Matrices

An ideal data structure on HDTS (for feasible dimension reduction):

$$\mathbf{y}_t = L\mathbf{f}_t + \epsilon_t, \quad t = 1, 2, \dots, T,$$

that satisfies (intuitively) that the low-dimensional projected data $\mathbf{f}_t$ holds the most time-serial dependence while the error component ϵ_t has almost independent observations.

The autocovariance matrix $\Sigma_\tau := \mathbb{E}[\mathbf{f}_t \mathbf{f}_{t+\tau}]$ is helpful.

Intuition: We see that $\Sigma_\tau = L \cdot \mathbb{E}[\mathbf{f}_t \mathbf{f}_{t+\tau}^\top] \cdot L^\top$. For the orthogonal complement matrix $B : p \times (p-K)$ (i.e. $B^\top L = \mathbf{0}$, $B^\top B = I_{p-K}$), we have $\Sigma_\tau \Sigma_\tau^\top B = 0$.

The $(p-K)$ columns of B are eigenvectors of the matrix $\Sigma_\tau \Sigma_\tau^\top$ corresponding to zero eigenvalues.

Subspace extracted from autocovariance matrices

In terms of analysis above, we conclude

1. The K columns of factor loading matrix L are eigenvectors of the matrix $\Sigma_\tau \Sigma_\tau^\top$ corresponding to non-zero eigenvalues.
2. The number K (the dimension of the subspace) is also the total number of non-zero eigenvalues of the matrix $\Sigma_\tau \Sigma_\tau^\top$.

A traditional estimator for $\Sigma_\tau \Sigma_\tau^\top$ is the sample version $\widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top$.

To study the dimension reduction on HDTS, it is equivalent to focus on empirical eigenvalues and eigenvectors from the symmetrized sample autocovariance matrix $\widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top$

Australian National University

Challenge

1. Similar to PCA under high-dimensional scenarios, the sample version $\widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top$ can be far from the population version $\Sigma_\tau \Sigma_\tau^\top$.

   Example: one sufficient condition for feasible PCA is

$$\frac{p}{T\lambda_1} \to 0, \quad as \;\; p, T \to \infty.$$

2. Few literature on empirical eigenvalues and corresponding eigenvectors (from the sample auto-covariance matrix).
   1. Lam, Yao and Bathia (2011, Biometrika)
   2. Lam and Yao (2012, Annals of Statistics)
   3. Li, Wang and Yao (2017, Annals of Statistics)
   4. Zhang, Pan, Yao and Zhou (2022, JASA to appear)

Some simulations before we go on...

- Consider the simple case where $L^\top = (I_2, 0)$, and

$$\mathbf{y}_t = \begin{pmatrix} f_{1t} \\ f_{2t} \\ \mathbf{0}_p \end{pmatrix} + \epsilon_t, \quad t = 1, \dots, T,$$

  where (ϵ_{it}) are i.i.d. standard Gaussians, and $(f_{1t})_t$ and $(f_{2t})_t$ are AR(1) processes.

- Parameters are chosen so that

$$\Sigma_1 \Sigma_1^\top = \mathbb{E}[\mathbf{y}_t \mathbf{y}_{t+1}^\top] \mathbb{E}[\mathbf{y}_t \mathbf{y}_{t+1}^\top]^\top = \mathrm{diag}(10, 3, \underbrace{0, \dots, 0}_{p}).$$

- $T = 1000, p \in \{100, 500, 800\}$.

What the fixed p asymptotic theory tells us

- Parameters are chosen so that

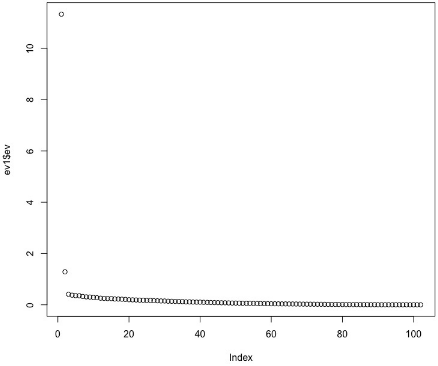
$$\Sigma_1 \Sigma_1^\top = \mathbb{E}[\mathbf{y}_t \mathbf{y}_{t+1}^\top] = \mathrm{diag}(10, 1, \underbrace{0, \dots, 0}_{p}).$$

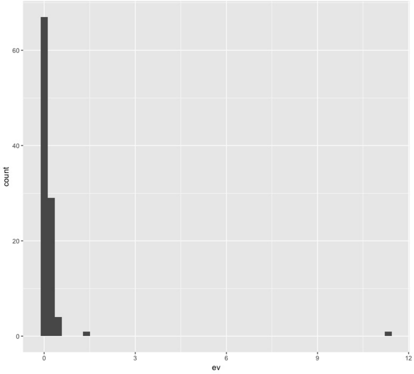
- Let λ_i be the (non-increasingly ordered) eigenvalues of $\widehat{\Sigma}_1 \widehat{\Sigma}_1^\top$. When p is fixed, we know that as $T \to \infty$, we have $\lambda_1 \to 10$, $\lambda_2 \to 1$ and $\lambda_i \to 0$ for all $i > 2$.

  Or in other words, we have

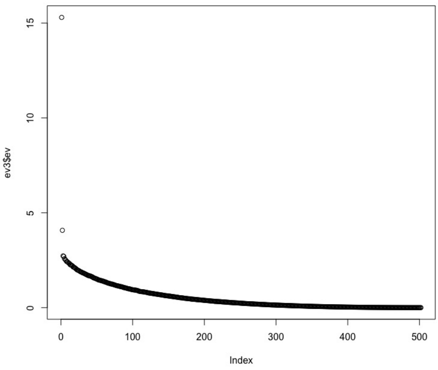
$$\frac{1}{p+2} \sum_{i=1}^{p+2} \delta_{\lambda_i}(dx) \Rightarrow \frac{1}{p+2}(\delta_{10} + \delta_1) + \frac{p}{p+2}\delta_0.$$

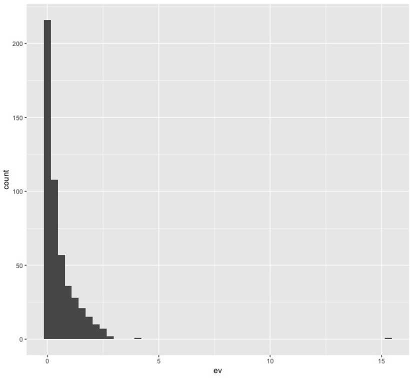
Sample eigenvalues, p=100



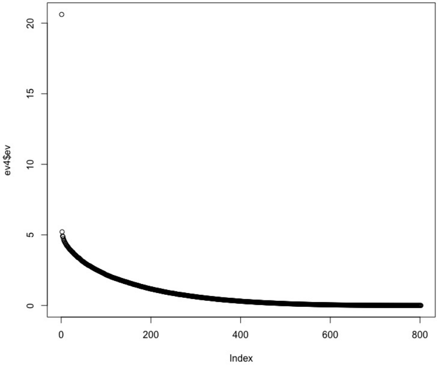


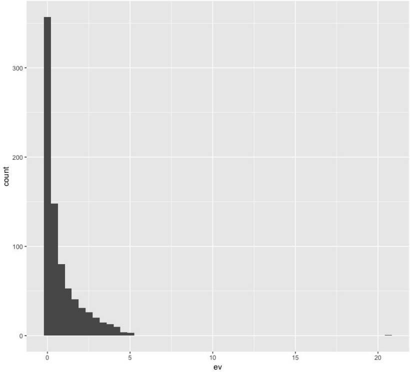
Sample eigenvalues, p=500





Sample eigenvalues, p=800





Review on Available Results for Autocovariance

The setting

- Consider a stationary time series $(\mathbf{y}_t)_{t=1,\ldots,T} \subseteq \mathbb{R}^{K+p}$ arising from the factor model

$$\mathbf{y}_t = L\mathbf{f}_t + \epsilon_t, \quad t = 1, \ldots, T,$$

  where the matrix $(\mathbf{f}_t)_{t=1,\ldots,T}$ contains K independent factors and $L^\top L = I_K$.
- High dimensional setting: $p = p_T \to \infty$ as $T \to \infty$ and $p/T \to c > 0$.
- Each factor $(f_{it})_{t=1,\ldots,T}$ is itself a stationary time series of the form

$$f_{it} = \sigma_i \sum_{l=0}^{\infty} \phi_{il} z_{i,t-l}, \quad i = 1, \ldots, K, \quad t = 1, \ldots, T,$$

  where (z_{it}) are i.i.d. with zero mean and unit variance.
- Normalization: take $\|\phi_i\|_{\ell_2} = 1$ so that $\mathrm{Var}(f_{it}) = \sigma_i^2$ for all $i \leq K$ and $t > 0$.

Analysis of Autocovariance

- Autocovariance of each factor is given by $\mathrm{Cov}(f_{it}, f_{i,t+\tau}) = \sigma_i^2\gamma_i(\tau)$, $\tau > 0$. Under this setup, for $\tau > 0$ we have

$$\Sigma_\tau := \mathbb{E}[\mathbf{y}_t\mathbf{y}_{t+\tau}^\top] = L\mathbb{E}[\mathbf{f}_t\mathbf{f}_{t+\tau}^\top]L^\top = L\begin{pmatrix}\sigma_1^2\gamma_1(\tau) & & \\ & \ddots & \\ & & \sigma_K^2\gamma_K(\tau)\end{pmatrix}L^\top.$$

- The spectrum of the ($(K+p) \times (K+p)$ dimensional) matrix $M := \Sigma_\tau\Sigma_\tau^\top$:

$$\sigma(M) = \left\{\sigma_1^4\gamma_1(\tau)^2, \dots, \sigma_K^4\gamma_K(\tau)^2, \underbrace{0, \dots, 0}_{p}\right\}$$

Australian National University

Asymptotics of sample eigenvalues

- In practice, we often estimate the eigenvalues

$$\sigma(M) = \left\{\sigma_1^4\gamma_1(\tau)^2, \dots, \sigma_K^4\gamma_K(\tau)^2, 0, \dots, 0\right\}$$

  using eigenvalues $\lambda_{1,\tau}, \dots, \lambda_{K+p,\tau}$ of the matrix $\widehat{M} := \widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top$.
- The asymptotic properties of $\{\lambda_{i,\tau}\}_{i=1,\dots,K+p}$ are the focus of several recent papers including Lam, Yao & Bathia (2011), Lam & Yao (2012), Li, Wang & Yao (2017).
- Main goal of our work is to establish the asymptotic normality of $\{\lambda_{1,\tau}, \dots, \lambda_{K,\tau}\}$.

The "low dimensional" regime where p is fixed

- When the dimension p is fixed, as the sample size $T \to \infty$,

$$\widehat{M} := \widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top \xrightarrow{\mathbb{P}} \Sigma_\tau \Sigma_\tau^\top =: M$$

  in the operator (and hence in any) norm.

- By continuity (w.r.t. the operator norm), for any $k \leq K + p$ and fixed $\tau > 0$,

$$\lambda_{k,\tau} \xrightarrow{\mathbb{P}} \sigma_k^4 \gamma_k(\tau)^2$$

  and the asymptotic fluctuation of $\lambda_{k,\tau}$ is Gaussian.

- However, when $p \to \infty$, this is no longer true.

The "high dimensional" regime where p diverges

- Suppose now that $p = p_T \to \infty$ as $T \to \infty$ and $p/T \to c > 0$.
- $\widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top$ still "consistently" estimates $\Sigma_\tau \Sigma_\tau$, but only entry-wise, so in general

$$\liminf_{p,T\to\infty} \|\widehat{\Sigma}_\tau \widehat{\Sigma}_\tau^\top - \Sigma_\tau \Sigma_\tau^\top\|_{op} > 0$$

and as a result, we have $\lambda_{k,\tau} \not\to \sigma_k^4 \gamma_k(\tau)^2$. The asymptotic fluctuations of $\lambda_{k,\tau}$ (around its limiting mean) may not be Gaussian either.

Recent works in the “high dimensional” regime

- Assume $p/T \to c > 0$.
- When $K = 0$ (the so-called null case), Li, Pan & Yao (2015) derives the limiting spectral distribution of $\widehat{\Sigma}\widehat{\Sigma}^\top$, i.e. as $p, T \to \infty$,

$$\sum_{i=1}^{K+p} \delta_{\lambda_{i,\tau}}(dx) \Rightarrow \text{some non-degenerate distribution } \nu.$$

- The phase transition of $\{\lambda_{k,\tau}\}$ is shown in Li, Wang & Yao (2017): there exists a critical threshold $\eta > 0$ such that the following dichotomy exists:
  - if $\sigma_i^4 \gamma_i(\tau)^2 > \eta$ then $\lambda_{i,\tau} \to \mu_i > \sigma_i^4 \gamma_i(\tau)^2$ in probability, i.e. $\lambda_{i,\tau}$ is detectable,
  - if $\sigma_i^4 \gamma_i(\tau)^2 < \eta$ then $\lambda_{i,\tau} \to \{\max x : \nu[x, \infty) > 0\}$, i.e. $\lambda_{i,\tau}$ "blends in" with all the other small eigenvalues which are estimators of zero.

CLT of Spiked Empirical Eigenvalues

Conditions on Dimension, Factors and Idiosyncratic Error

Assumptions 1

1. $p, T \to \infty$ *and* $p/T \to c > 0$.
2. $\sigma_i \to \infty$ *and there exists* $C > 0$ *such that* $\sigma_i/\sigma_j < C$ *for all* $i, j = 1, \ldots, K$.
3. $(z_{it})_{1 \leq i \leq K, 1-L \leq t \leq T+1}$ *is independent, identically distributed with* $\mathbb{E}[z_{it}] = 0$, $\mathbb{E}[z_{it}^2] = 1$ *and uniformly bounded* $4 + \epsilon$ *moment for some* $\epsilon > 0$.
4. $(\epsilon_{it})_{1 \leq i \leq p+K, 1 \leq t \leq T+1}$ *is i.i.d. standard Gaussian.*
5. $\sup_i \|\phi_i\|_{\ell_1} < \infty$.

Conditions on Number of Factors, Auto Time-lag and Factor Strength

Assumptions 2

1. τ *is a fixed, non-negative integer*
2. $K = o(T^{1/16})$ *and* $K = o(\sigma_1^2)$ *as* $T \to \infty$.
3. *the sequence* $(\mu_{1,\tau}, \ldots, \mu_{K,\tau})$ *is arranged in decreasing order and there exists* $\epsilon > 0$ *such that* $\mu_{i,\tau}/\mu_{i+1,\tau} > 1 + \epsilon$ *for all* $i = 1, \ldots, K - 1$.

Assumptions 3

1. $\tau \in \mathbb{N}$ *and* $\tau \to \infty$ *as* $T \to \infty$.
2. $K = o(T^{1/16}\gamma_1(\tau)^{1/2})$ *and* $K = o(\sigma_1^2\gamma_1(\tau)^3)$ *as* $T \to \infty$.
3. *there exists* $C_1 > 0$ *such that* $\mu_{i,\tau}/\mu_{j,\tau} \leq C_1$ *for all* $i, j = 1, \ldots, K$ *and* $\tau \geq 0$.
4. *there exists* T_0 *large enough and some* $\epsilon > 0$ *such that* $\mu_{i,\tau}/\mu_{i+1,\tau} > 1 + \epsilon$ *for all* $i = 1, \ldots, K - 1$ *and* $T > T_0$.

Location of Spiked Empirical Eigenvalues

Theorem 1

Under Assumption 1 and either Assumption 2 or 3, we have

$$\frac{\lambda_{n,\tau}}{\mu_{n,\tau}} - 1 = O_p\left(\frac{1}{\gamma_n(\tau)\sqrt{T}}\right) + O_p\left(\frac{K}{\sigma_n^2\gamma_n(\tau)^2}\right), \quad n = 1, \ldots, K. \tag{0.2}$$

where $\mu_{n,\tau}$ *is*

$$\mu_{i,\tau} := \mathbb{E}[y_{i,t}y_{i,t+\tau}]^2 = \sigma_i^4\gamma_i(\tau)^2, \quad i = 1, \ldots, K, \quad \tau \geq 0. \tag{0.3}$$

Australian National University

- The asymptotic distribution of $\lambda_{i,\tau}$ remains unknown.
- Our work is a first step in answering this question - we show that:

Theorem 2

Under Assumption 1 and either Assumption 2 or 3, we have

$$\sqrt{T}\frac{\gamma_i(\tau)}{2\nu_{i,\tau}}\left(\frac{\lambda_{i,\tau}}{\theta_{i,\tau}} - 1\right) \Rightarrow N(0,1),$$

where $\theta_{i,\tau}$ is defined implicitly as the solution to some (non-random) equation.

- For generality we allow $K \to \infty$ and even $\tau \to \infty$.

Statistical Application: Equivalance Test for two HDTS's

Statistical applications: auto-covariance test

- Hypothesis testing for comparing two populations is a traditional statistical problem
  - T-test and/or Z-test for equality of two population mean
  - F-test for equality of two population variance
- Comparing two populations of high-dimensional time series
  - Provide better inference if they share similar information (both temporal and cross-sectional)
  - Aggregated analysis for multiple populations of high-dimensional time series
    - Human mortality data from different countries
  - Interest: spiked eigenvalues of high-dimensional auto-covariance matrices for two populations

Hypothesis testing for two populations

- Testing for the equivalence of spiked eigenvalues for auto-covariance matrices of two high-dimensional time series
- For two high-dimensional time series $\left\{\mathbf{y}_t^{(1)}\right\}$ and $\left\{\mathbf{y}_t^{(2)}\right\}$ following the factor models in canonical form under assumptions of Theorem 2,
  - H_0: $\mu_{i,\tau}^{(1)} = \mu_{i,\tau}^{(2)}$ for all $i = 1, 2, ..., K$;
  - H_1: $\mu_{i,\tau}^{(1)} \neq \mu_{i,\tau}^{(2)}$ for at least one i, $i = 1, 2, ..., K$.

Australian National University

Test statistic

- For two HD time series, a test statistic can be considered as,

$$Z_{i,\tau} = \sqrt{T}\frac{\gamma_{i,\tau}}{2\sqrt{2}v_{i,\tau}}\frac{\lambda_{i,\tau}^{(1)} - \lambda_{i,\tau}^{(2)}}{\theta_{i,\tau}}, \quad (0.4)$$

where

$$\theta_{i,\tau} = \frac{\theta_{i,\tau}^{(1)} + \theta_{i,\tau}^{(2)}}{2}, \; v_{i,\tau} = \frac{v_{i,\tau}^{(1)} + v_{i,\tau}^{(2)}}{2}, \text{ and } \gamma_{i,\tau} = \frac{\gamma_{i,\tau}^{(1)} + \gamma_{i,\tau}^{(2)}}{2}, \quad (0.5)$$

and $\theta_{i,\tau}^{(m)}$ is the asymptotic centering of $\lambda_{i,\tau}^{(m)}$.

Theorem 3

Under the assumptions of Theorem 2, for two independent high-dimensional time series $\left\{\mathbf{y}_t^{(1)}\right\}$ and $\left\{\mathbf{y}_t^{(2)}\right\}$ following the same factors in canonical form , we have

$$Z_{i,\tau} = \sqrt{T}\frac{\gamma_{i,\tau}}{2\sqrt{2}v_{i,\tau}}\frac{\lambda_{i,\tau}^{(1)} - \lambda_{i,\tau}^{(2)}}{\theta_{i,\tau}} \Rightarrow \mathcal{N}(0,1), \tag{0.6}$$

as $T, p \to \infty$, where $\theta_{i,\tau}$, $v_{i,\tau}$ and $\gamma_{i,\tau}$ are defined in (0.5).

Theorem 3 is a direct result of Theorem 2, since an asymptotic distribution of $\frac{\lambda_{i,\tau}^{(1)} - \lambda_{i,\tau}^{(2)}}{\theta_{i,\tau}}$ can be derived using the independence between $\lambda_{i,\tau}^{(1)}$ and $\lambda_{i,\tau}^{(2)}$.

Theorem 4

Under the assumptions of Theorem 2, if we additionally assume two independent high-dimensional time series $\left\{\mathbf{y}_t^{(1)}\right\}$ *and* $\left\{\mathbf{y}_t^{(2)}\right\}$ *follow two different canonical factor models with*

$$K_1 = K_2 = K,\ \gamma_{i,\tau}^{(1)} = \gamma_{i,\tau}^{(2)} = \gamma_{i,\tau},\ v_{i,\tau}^{(1)} = v_{i,\tau}^{(2)} = v_{i,\tau},\ \text{and } \theta_{i,\tau}^{(1)} = (1+c)\theta_{i,\tau}^{(2)}.$$

Then, for any c such that $\sqrt{T}\frac{2c}{2+c} \to \infty$ *as* $T, p \to \infty$ *and* $\lambda_{i,\tau}^{(1)} \neq \lambda_{i,\tau}^{(2)}$*, it holds that*

$$Pr\left(|Z_{i,\tau}| > z_\alpha | H_1\right) \to 1, \tag{0.7}$$

for $T, p \to \infty$*, where* z_α *is the* α*-th quantile of the standard normal distribution.*

Implementation

- Step 1: Estimations of the factor model
  - Use symmetrized lag-τ sample auto-covariance matrix to estimate the number of factors $\widehat{r}^{(\cdot)}$ and factor loading matrices $\widehat{L}$ from the two samples and then estimate the factors.
- Step 2: Standardizing the estimated factor models to the canonical form
  - This can be achieved by normalizing the estimated loading matrix to a diagonal matrix and the variance of each factors to be 1.

Implementation

- Step 3: Estimates of unknown parameters for the test statistic
  - Bootstrap methods for time series such as the sieve bootstrap needs to be conducted on the estimated factors for estimating $v_{i,\tau}^{(\cdot)}$ and $\theta_{i,\tau}^{(\cdot)}$.
- Step 4: Computing the test statistic and *p*-values
  - The test statistic can be computed as

$$\widetilde{Z}_{i,\tau} := \left(\lambda_{i,\tau}^{(1)} - \lambda_{i,\tau}^{(2)}\right) \sqrt{\frac{T_1 T_2}{T_1 + T_2} \frac{\widetilde{\gamma}_{i,\tau}^{*}}{2\widetilde{v}_{i,\tau}^{*} \widetilde{\theta}_{i,\tau}^{*}}},$$

where $\widetilde{\theta}_{i,\tau}^{*}$, $\widetilde{v}_{i,\tau}^{*}$ and $\widetilde{\gamma}_{i,\tau}^{*}$ are bootstrap estimates.

Simulation Studies

Data Generating Process

- DGP:
  - consider a one-factor model for both populations, where the factor is generated by

$$f_{1,t}^{(m)} = \phi_1^{(m)} f_{1,t-1}^{(m)} + z_{1,t}^{(m)}, \; m = 1, 2, \tag{0.8}$$

    where $\phi_1^{(m)} = 0.5$ and $\left\{z_{1,t}^{(m)}\right\}$ are i.i.d. $\mathcal{N}\left(0, \left(\sigma_z^{(m)}\right)^2\right)$ with $\left(\sigma_z^{(m)}\right)^2 = 3/4$, so that $Var\left(f_{1,t}^{(m)}\right) = 1$
  - And the data is generated by

$$\mathbf{y}_t^{(m)} = \begin{pmatrix} \sigma_1^{(m)} \\ \mathbf{0}_{N-1} \end{pmatrix} f_{1,t}^{(m)} + \boldsymbol{\epsilon}_t^{(m)}, \tag{0.9}$$

    where $\sigma_1^{(m)} = N^{1-\delta}$, $\{\epsilon_{j,t}\}$ are i.i.d. $\mathcal{N}(0,1)$, and $\left\{f_{1,t}^{(m)}\right\}$ are generated by (0.8).
  - Note that δ represents the factor strength and $\delta = 0$ is the strongest case.

Empirical Sizes

- Empirical sizes

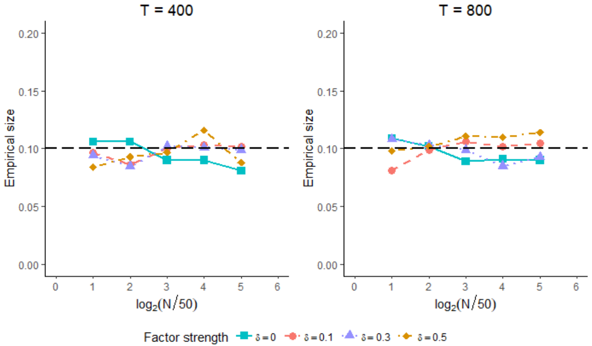


Figure 3: Empirical sizes of the auto-covariance test with $T = 400, 800$, $N = 100, 200, 400, 800, 1600$, and $\delta = 0, 0.1, 0.3, 0.5$.

Australian National University

Empirical Powers: factor strength

- Empirical powers: scenario 1 - study the effect of different variances
  - Consider different groups of data generated with $\sigma^{2(2)}$ set as $1.1\left(\sigma_1^{(1)}\right)^2, 1.3\left(\sigma_1^{(1)}\right)^2, 1.5\left(\sigma_1^{(1)}\right)^2, 1.7\left(\sigma_1^{(1)}\right)^2$, and $1.9\left(\sigma_1^{(1)}\right)^2$, respectively.

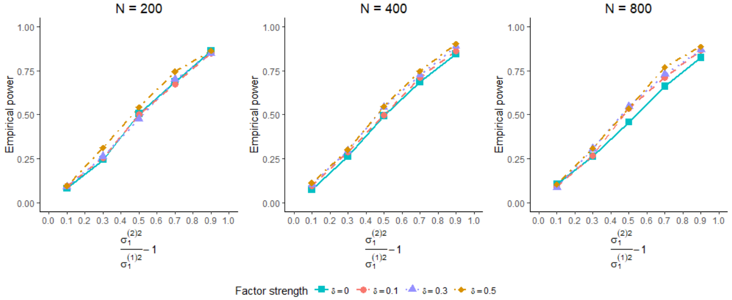


Figure 4: Empirical powers of the auto-covariance test in the first scenario with $T = 400$, $N = 200, 400, 800$, and $\delta = 0, 0.1, 0.3, 0.5$.

Empirical Powers: Spikeness

- Empirical powers: scenario 2 - study the effect of different auto-covariances (auto-correlations) of $f_{i,t}$
  - Consider different groups of data generated with $\phi_{i,1}^{(2)}$ set as $0.9\phi_1^{(1)}, 0.8\phi_1^{(1)}, 0.7\phi_1^{(1)}, 0.6\phi_1^{(1)}$, and $0.5\phi_1^{(1)}$, respectively.

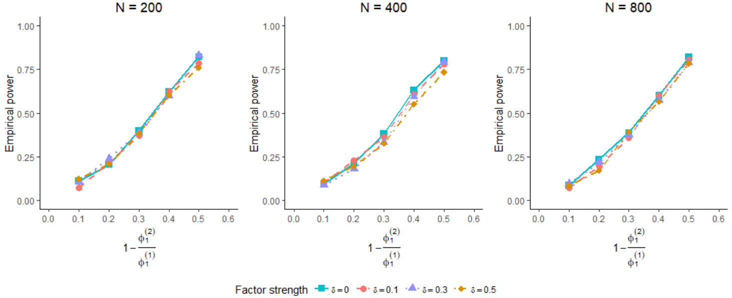


Figure 5: Empirical powers of the auto-covariance test in the second scenario with $T = 400$, $N = 200, 400, 800$, and $\delta = 0, 0.1, 0.3, 0.5$.

Empirical Application on Clustering Mortality Data

Human mortality data across countries

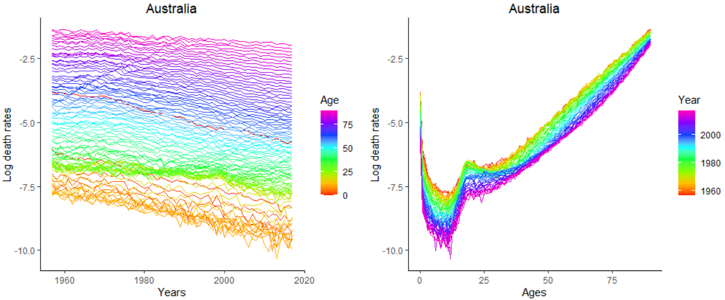


Figure 6: Log death rates for Australian

Human mortality data across countries

- We study total death rates from selected countries where the data is available from 1957 to 2017
- The data is prepared by taking first order difference on the log death rates as the original data is not stationary

Table 1: Estimated number of factors in the factor model for each country

Estimated number of factors	Countries
1	Australia, Belgium, Bulgaria, Czechia, Finland, Greece, Hungary, Japan, Netherlands, Sweden, Switzerland, U.K., U.S.A.
2	Denmark
3	Canada, France, Italy, Portugal
5	Poland

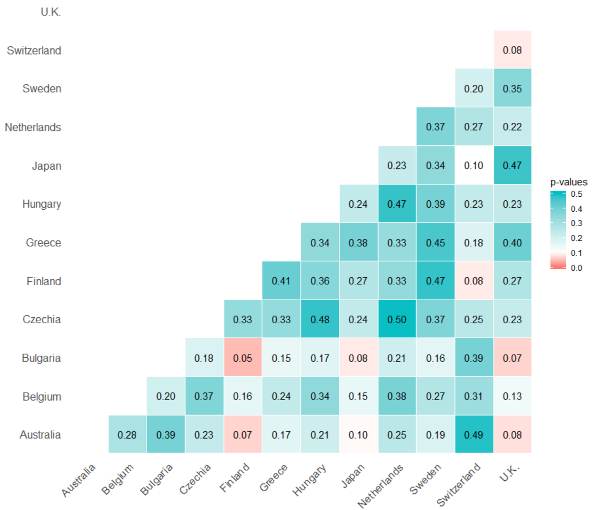


Figure 7: *p*-values of the auto-covariance test for each pair of countries that have one factor in the estimated factor model

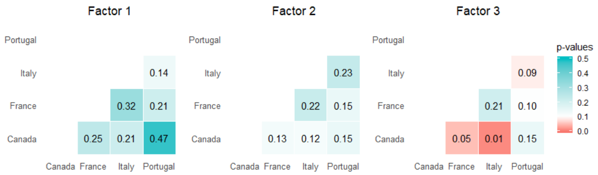


Figure 8: *p*-values of the auto-covariance test for each pair of countries that have three factors in the estimated factor model

Figure 9: *p*-values of the auto-covariance test of the first factor for all countries except U.S.A.

References

Lam, C. & Yao, Q. (2012), 'Factor modeling for high-dimensional time series: Inference for the number of factors', *The Annals of Statistics* **40**(2), 694–726.

Lam, C., Yao, Q. & Bathia, N. (2011), 'Estimation of latent factors for high-dimensional time series', *Biometrika* **98**(4), 901–918.

Li, Z., Pan, G. & Yao, J. (2015), 'On singular value distribution of large-dimensional autocovariance matrices', *Journal of Multivariate Analysis* **137**, 119–140.

Li, Z., Wang, Q. & Yao, J. (2017), 'Identifying the number of factors from singular values of a large sample auto-covariance matrix', *The Annals of Statistics* **45**(1), 257–288.

Pan, J. & Yao, Q. (2008), 'Modelling multiple time series via common factors', *Biometrika* **95**(2), 365–379.

Thank you !