Big n, Big p: Eigenvalues for Cov Matrices of Heavy-Tailed Multivariate Time Series

Richard A. Davis, Columbia University

Thomas Mikosch, University of Copenhagen
Oliver Pfaffel, Munich Re

June 5, 2014

Workshop on Quantitative Methods in Finance and Insurance
University of Zagreb
Croatia

The Setup

- Data matrix: A $p \times n$ matrix X consisting of n observations of a p-dimensional time series, i.e.,

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & \cdots & X_{pn} \end{bmatrix}.$$

- Sample covariance matrix: the $p \times p$ sample covariance matrix (normalized) is given by

$$XX^T = n\hat{\Gamma}(0) = \left[\sum_{t=1}^{n} X_{it} X_{jt} \right]_{i,j=1}^{p}.$$

- Objective: study the ordered eigenvalues

$$\lambda_{(1)} \geq \lambda_{(2)} \geq \ldots \geq \lambda_{(p)}$$

of the $p \times p$ sample covariance matrix XX^T.

The Setup-continued

Data matrix and sample covariance matrix:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & \cdots & X_{pn} \end{bmatrix} \text{ and } XX^T = n\hat{\Gamma}(0)$$

- Note that if the rows are independent and identically distributed ergodic time series (with mean 0 and variance 1), then for p fixed,

$$\hat{\Gamma}(0) \xrightarrow{P} I_p \,.$$

- Relation to PCA: $\lambda_{(1)}$ is the empirical variance of the first principal component, $\lambda_{(2)}$ of the second, and so on.

Known results for the largest eigenvalue

- Assume the entries of X are iid Gaussian (with mean zero and variance one)
- For $n \to \infty$ and fixed p, Anderson [1963] proved that

$$\sqrt{\frac{n}{2}}\left(\frac{\lambda_{(1)}}{n} - 1\right) \xrightarrow{d} \mathrm{N}(0,1)\,.$$

Known results for the largest eigenvalue

- Assume the entries of X are iid Gaussian (with mean zero and variance one)
- For $n \to \infty$ and fixed p, Anderson [1963] proved that

$$\sqrt{\frac{n}{2}}\left(\frac{\lambda_{(1)}}{n} - 1\right) \xrightarrow{d} \mathrm{N}(0,1)\,.$$

- Johnstone [2001] showed that for $p, n \to \infty$ s.t. $p/n \to \gamma \in (0, \infty)$

$$\frac{\sqrt{n} + \sqrt{p}}{\sqrt[3]{\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{p}}}}\left(\frac{\lambda_{(1)}}{\left(\sqrt{n} + \sqrt{p}\right)^2} - 1\right) \xrightarrow{d} \text{Tracy-Widom distribution}$$

Known results for the largest eigenvalue

- Assume the entries of X are iid Gaussian (with mean zero and variance one)
- For $n \to \infty$ and fixed p, Anderson [1963] proved that

$$\sqrt{\frac{n}{2}}\left(\frac{\lambda_{(1)}}{n} - 1\right) \xrightarrow{d} \mathrm{N}(0,1)\,.$$

- Johnstone [2001] showed that for $p, n \to \infty$ s.t. $p/n \to \gamma \in (0, \infty)$

$$\frac{\sqrt{n}+\sqrt{p}}{\sqrt[3]{\frac{1}{\sqrt{n}}+\frac{1}{\sqrt{p}}}}\left(\frac{\lambda_{(1)}}{\left(\sqrt{n}+\sqrt{p}\right)^2} - 1\right) \xrightarrow{d} \text{Tracy-Widom distribution}$$

- The assumption of Gaussianity in Johnstone's result can be relaxed to a *moment condition* (c.f. Four Moment Theorem by Tao and Vu [2011]; and work by Erdös, Johansson, Péché, Schlein, Soshnikov, Yau and others).

Setting

- Suppose $X = (X_{it})_{i,t}$, $i = 1, \ldots, p$, $t = 1, \ldots, n$, with

$$X_{it} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} h(k,j) Z_{i-k,t-j} \,.$$

- The noise $(Z_{i,t})$ is iid with regularly varying tails of index $\alpha \in (0,4)$ (infinite fourth moment), i.e.,

$$n\,P(|Z_{11}| > a_n x) \to x^{-\alpha} \text{ as } n \to \infty, \text{ for } x > 0,$$

($a_n = L(n) n^{1/\alpha}$) and

$$\lim_{x\to\infty} \frac{P(Z_{11} > x)}{P(|Z_{11}| > x)} = p_+ \quad \text{and} \quad \lim_{x\to\infty} \frac{P(Z_{11} \le -x)}{P(|Z_{11}| > x)} = 1 - p_+$$

Conditions on h

Summability assumptions on $h(k,l)$:

$$\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}|h(k,j)|^{\delta}<\infty \quad \text{for some } \delta<\min\{1,\alpha\}$$

and

$$\sum_{t=0}^{\infty}\left(\sum_{j=t}^{\infty}|h(k,j)|\right)^{\alpha/2-\epsilon}<\infty, \quad \text{for } k=0,1,2\ldots,.$$

Note: latter condition is implied by

$$\sum_{j=0}^{\infty} j^{2/\alpha+\epsilon'}\,|h(k,j)|<\infty\,, \quad k=0,1,\ldots,,$$

for $\epsilon' > 0$ arbitrarily close to zero.

Setting (cont)

- Let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of

$$\begin{cases} XX^T, & \text{if } \alpha \in (0,2), \\ XX^T - E(XX^T), & \text{if } \alpha \in (2,4). \end{cases}$$

- Let (D_s) be the iid sequence given by

$$D_s = D_s^{(n)} = \sum_{t=1}^{n} Z_{s,t}^2 \, .$$

Note:

1. The D_s play a key role in determining the asymptotic properties of the ordered eigenvalues $\lambda_{(1)} \geq \cdots \geq \lambda_{(p)}$.
2. Large deviations result implies $pP(D_1 \geq a_{np}^2 x) \to x^{-\alpha/2}$ for $\alpha \in (0,2)$. (Mean correct D_1 for $\alpha \in (2,4)$.)

One more thing!

Set $\mathbf{h}_i = (h_{i0}, h_{i1}, \ldots)^T$ and define the matrix $H = (\mathbf{h}_0, \mathbf{h}_1, \ldots,)$. Let

$$M = H^T H.$$

i.e., the (i, j)th entry of M is

$$M_{ij} = \mathbf{h}_i^T \mathbf{h}_j = \sum_{l=0}^{\infty} h_{il} h_{jl}, \quad i, j = 0, 1, \ldots, .$$

By construction, M is symmetric and non-negative definite and has ordered eigenvalues

$$v_1 \geq v_2 \geq v_3 \geq \cdots$$

Let $r \leq \infty$ be the rank of M so that $v_r > 0$ while $v_{r+1} = 0$ if $r < \infty$.

One more thing!

Set $\mathbf{h}_i = (h_{i0}, h_{i1}, \ldots)^T$ and define the matrix $H = (\mathbf{h}_0, \mathbf{h}_1, \ldots,)$. Let

$$M = H^T H.$$

i.e., the (i, j)th entry of M is

$$M_{ij} = \mathbf{h}_i^T \mathbf{h}_j = \sum_{l=0}^{\infty} h_{il} h_{jl}, \quad i, j = 0, 1, \ldots, .$$

By construction, M is symmetric and non-negative definite and has ordered eigenvalues

$$v_1 \geq v_2 \geq v_3 \geq \cdots$$

Let $r \leq \infty$ be the rank of M so that $v_r > 0$ while $v_{r+1} = 0$ if $r < \infty$.
Remark: M is the covariance matrix of the vector $\mathbf{X}^* = (X_0^*, X_1^*, \ldots)^T$,

$$X_i^* = \sum_{l=0}^{\infty} h(i, l) Z_l, \quad \{Z_l\} \sim \text{IID}(0, 1)$$

Example

$$X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$$

$$H^T = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots \\ -2 & 2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ldots \end{pmatrix} \quad M = H^T H = \begin{pmatrix} 2 & 0 & 0 & \cdots \\ 0 & 8 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ldots \end{pmatrix}$$

which has non-negative eigenvalues $v_1 = 8$ and $v_2 = 2$ ($r = 2$).

Theorem (Main result to the point process convergence)

Let $p = p_n \to \infty$ be a sequence satisfying certain growth conditions (to be specified later) and suppose $k = k_p \to \infty$ is any sequence such that $k^2 = o(p)$.

a) *If $\alpha \in (0, 2)$, then*

$$a_{np}^{-2} \max_{i=1,\ldots,p} \left|\lambda_{(i)} - \delta_{(i)}\right| \xrightarrow{P} 0\,, \quad n \to \infty\,,$$

where

- $\lambda_{(1)} \geq \cdots \geq \lambda_{(p)}$ *are the ordered eigenvalues of XX^T.*
- $\delta_{(1)} \geq \cdots \geq \delta_{(p)}$ *are the ordered values from the set $\{D_{(i)} v_j,\ i = 1, \ldots, k,\ j = 1, 2, \ldots,\}$.*

Note: $\delta_{(1)} = v_1 D_{(1)},\ \delta_{(2)} = v_2 D_{(1)} \vee v_1 D_{(2)}$, etc.

Theorem (Main result cont)

b) *If* $\alpha \in (2,4)$, *then*

$$a_{np}^{-2} \max_{i=1,\dots,p} \left| \widetilde{\lambda}_{(i)} - \widetilde{\delta}_i \right| \xrightarrow{P} 0\,, \quad n \to \infty\,,$$

where

- $\tilde{\lambda}_{(1)}, \dots, \tilde{\lambda}_{(p)}$ *are the ordered eigenvalues* (λ_i) *according to their absolute values.*
- $\tilde{\delta}_{(1)} \geq \dots \geq \tilde{\delta}_{(p)}$ *are the ordered values from the set* $\{(D_{l_i} - ED)v_j,\ i = 1,\dots,k,\ j = 1,2,\dots,\}$.

Theorem (Point process convergence)

Let $p = p_n \to \infty$ be a sequence satisfying certain growth conditions (to be specified later). Then we have the point process convergence,

$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \xrightarrow{d} N = \sum_{j=1}^{r} \sum_{i=1}^{\infty} \epsilon_{v_j \Gamma_i^{-2/\alpha}},$$

where $\Gamma_i = E_1 + \ldots + E_i$ is the cumulative sum of iid standard (i.e., mean one) exponentially distributed rv's,

Note: The point process $N^* = \sum_{i=1}^{\infty} \epsilon_{\Gamma_i^{-2/\alpha}}$ is a Poisson process with $E(N^*(dx)) = \alpha/2 x^{-\alpha/2-1} dx$.

The largest eigenvalues

Let $d_{(1)} \geq d_{(2)} \geq \cdots$ be the ordered values of the set

$$\{v_j \Gamma_i^{-2/\alpha},\ i = 1, \ldots,\ j = 1, 2, \ldots,\}$$

The largest eigenvalues

Let $d_{(1)} \geq d_{(2)} \geq \cdots$ be the ordered values of the set

$$\{v_j \Gamma_i^{-2/\alpha},\ i = 1, \ldots,\ j = 1, 2, \ldots,\}$$

$$d_{(1)} = v_1 \Gamma_1^{-2/\alpha}, \qquad d_{(2)} = \max(v_2 \Gamma_1^{-2/\alpha}, v_1 \Gamma_2^{-2/\alpha})$$

The largest eigenvalues

Let $d_{(1)} \geq d_{(2)} \geq \cdots$ be the ordered values of the set

$$\{v_j \Gamma_i^{-2/\alpha}, \, i = 1, \ldots, \, j = 1, 2, \ldots, \}$$

$$d_{(1)} = v_1 \Gamma_1^{-2/\alpha}, \qquad d_{(2)} = \max(v_2 \Gamma_1^{-2/\alpha}, v_1 \Gamma_2^{-2/\alpha})$$

- The theorem implies the joint convergence of the m-largest eigenvalues

$$a_{np}^{-2} \left(\lambda_{(1)}, \ldots, \lambda_{(m)} \right) \stackrel{d}{\to} \left(d_{(1)}, \ldots, d_{(m)} \right) .$$

The largest eigenvalues

Let $d_{(1)} \geq d_{(2)} \geq \cdots$ be the ordered values of the set

$$\{v_j \Gamma_i^{-2/\alpha},\, i = 1, \ldots,\, j = 1, 2, \ldots,\}$$

$$d_{(1)} = v_1 \Gamma_1^{-2/\alpha}, \qquad d_{(2)} = \max(v_2 \Gamma_1^{-2/\alpha}, v_1 \Gamma_2^{-2/\alpha})$$

- The theorem implies the joint convergence of the m-largest eigenvalues

$$a_{np}^{-2} \left(\lambda_{(1)}, \ldots, \lambda_{(m)} \right) \stackrel{d}{\rightarrow} \left(d_{(1)}, \ldots, d_{(m)} \right).$$

- $(\lambda_{(1)}/a_{np}^2)^{-\alpha/2} \stackrel{d}{\rightarrow} v_1^{-\alpha/2} \Gamma_1$

The largest eigenvalues

Let $d_{(1)} \geq d_{(2)} \geq \cdots$ be the ordered values of the set

$$\{v_j \Gamma_i^{-2/\alpha},\ i = 1, \ldots,\ j = 1, 2, \ldots,\}$$

$$d_{(1)} = v_1 \Gamma_1^{-2/\alpha}, \qquad d_{(2)} = \max(v_2 \Gamma_1^{-2/\alpha}, v_1 \Gamma_2^{-2/\alpha})$$

- The theorem implies the joint convergence of the m-largest eigenvalues

$$a_{np}^{-2} \left(\lambda_{(1)}, \ldots, \lambda_{(m)} \right) \stackrel{d}{\to} \left(d_{(1)}, \ldots, d_{(m)} \right) .$$

- $(\lambda_{(1)}/a_{np}^2)^{-\alpha/2} \stackrel{d}{\to} v_1^{-\alpha/2} \Gamma_1$
-

$$\frac{\lambda_{(1)}}{\lambda_{(1)} + \cdots + \lambda_{(m)}} \stackrel{d}{\to} \frac{v_1 \Gamma_1^{-2/\alpha}}{d_{(1)} + \cdots + d_{(m)}}, \quad n \to \infty .$$

The largest eigenvalues

Let $d_{(1)} \geq d_{(2)} \geq \cdots$ be the ordered values of the set

$$\{v_j \Gamma_i^{-2/\alpha},\ i = 1, \ldots,\ j = 1, 2, \ldots,\}$$

$$d_{(1)} = v_1 \Gamma_1^{-2/\alpha}, \qquad d_{(2)} = \max(v_2 \Gamma_1^{-2/\alpha}, v_1 \Gamma_2^{-2/\alpha})$$

- The theorem implies the joint convergence of the m-largest eigenvalues

$$a_{np}^{-2}\left(\lambda_{(1)}, \ldots, \lambda_{(m)}\right) \xrightarrow{d} \left(d_{(1)}, \ldots, d_{(m)}\right).$$

- $(\lambda_{(1)}/a_{np}^2)^{-\alpha/2} \xrightarrow{d} v_1^{-\alpha/2}\Gamma_1$
-

$$\frac{\lambda_{(1)}}{\lambda_{(1)} + \cdots + \lambda_{(m)}} \xrightarrow{d} \frac{v_1 \Gamma_1^{-2/\alpha}}{d_{(1)} + \cdots + d_{(m)}}, \quad n \to \infty.$$

In fact, we have more!!

Self-normalization

Under the conditions of the theorem, the following limit results hold.

1. If $\alpha \in (0,2)$, then

$$a_{np}^{-2}\Big(\lambda_{(1)}, \sum_{i=1}^{p} \lambda_i\Big) \stackrel{d}{\to} \Big(\Gamma_1^{-2/\alpha}, \sum_{j=1}^{r}\sum_{i=1}^{\infty} v_j \Gamma_i^{-2/\alpha}\Big),$$

and in particular,

$$\frac{\lambda_{(1)}}{\lambda_1 + \cdots + \lambda_p} \stackrel{d}{\to} \frac{v_1}{v_1 + \ldots + v_r} \frac{\Gamma_1^{-2/\alpha}}{\sum_{i=1}^{\infty} \Gamma_i^{-2/\alpha}}, \quad n \to \infty .$$

2. If $\alpha \in (2,4)$ then

$$\frac{\lambda_{(1)}}{\lambda_1 + \cdots + \lambda_p} \stackrel{d}{\to} \frac{v_1}{v_1 + \ldots + v_r} \frac{\Gamma_1^{-2/\alpha}}{\xi_{\alpha/2}}, \quad n \to \infty ,$$

where

$$\xi_{\alpha/2} = \lim_{\gamma \downarrow 0} \sum_{i=1}^{\infty} \Big(\Gamma_i^{-2/\alpha} I_{\{\Gamma_i^{-2/\alpha} > \gamma\}} - E\Gamma_i^{-2/\alpha} I_{\{\Gamma_i^{-2/\alpha} > \gamma\}}\Big)$$

Example (cont)

Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$

Example (cont)

Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$

Then,

$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \xrightarrow{d} N = \sum_{i=1}^{\infty} \left(\epsilon_{8\Gamma_i^{-2/\alpha}} + \epsilon_{2\Gamma_i^{-2/\alpha}} \right).$$

Example (cont)

Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$

Then,

$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \xrightarrow{d} N = \sum_{i=1}^{\infty} \left(\epsilon_{8\Gamma_i^{-2/\alpha}} + \epsilon_{2\Gamma_i^{-2/\alpha}} \right).$$

Results:

- $a_{np}^{-2}\lambda_{(1)} \xrightarrow{d} 8\Gamma_1^{-2/\alpha}$

Example (cont)

Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$

Then,

$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \xrightarrow{d} N = \sum_{i=1}^{\infty} \left(\epsilon_{8\Gamma_i^{-2/\alpha}} + \epsilon_{2\Gamma_i^{-2/\alpha}} \right).$$

Results:

- $a_{np}^{-2}\lambda_{(1)} \xrightarrow{d} 8\Gamma_1^{-2/\alpha}$
- $a_{np}^{-2}(\lambda_{(1)}, \lambda_{(2)}) \xrightarrow{d} (8\Gamma_1^{-2/\alpha}, 2\Gamma_1^{-2/\alpha} \vee 8\Gamma_2^{-2/\alpha})$

Example (cont)

Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$

Then,

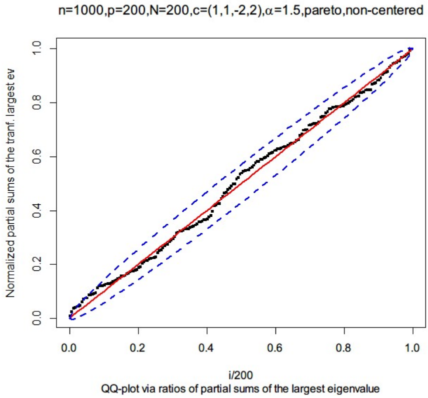
$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_i} \xrightarrow{d} N = \sum_{i=1}^{\infty} \left(\epsilon_{8\Gamma_i^{-2/\alpha}} + \epsilon_{2\Gamma_i^{-2/\alpha}} \right).$$

Results:

- $a_{np}^{-2}\lambda_{(1)} \xrightarrow{d} 8\Gamma_1^{-2/\alpha}$
- $a_{np}^{-2}(\lambda_{(1)}, \lambda_{(2)}) \xrightarrow{d} (8\Gamma_1^{-2/\alpha}, 2\Gamma_1^{-2/\alpha} \vee 8\Gamma_2^{-2/\alpha})$
-
$$\frac{\lambda_{(1)}}{\lambda_1 + \cdots + \lambda_p} \xrightarrow{d} \frac{8}{10} \frac{\Gamma_1^{-2/\alpha}}{\sum_{i=1}^{\infty} \Gamma_i^{-2/\alpha}}, \quad n \to \infty.$$

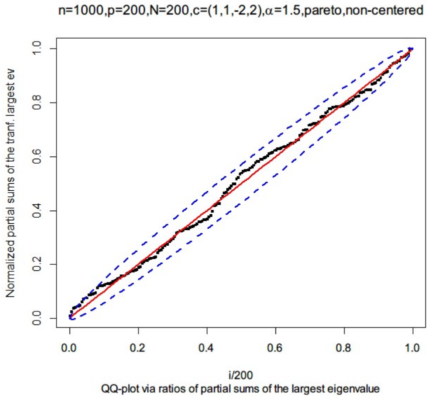
QQ-Plot via ratio of partial sums to $\lambda_{(1)}$

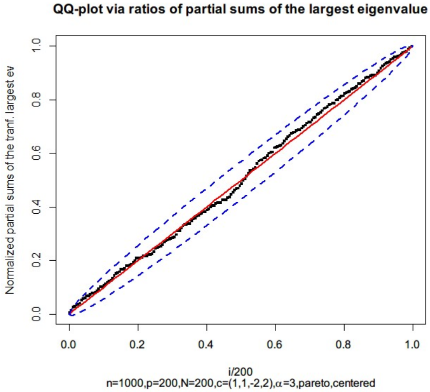
Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$, Pareto noise wih $\alpha = 1.5$ and $\alpha = 3.0$, replications = 200



QQ-Plot via ratio of partial sums to $\lambda_{(1)}$

Model: $X_{i,t} = Z_{i,t} + Z_{i,t-1} - (2Z_{i-1,t} - 2Z_{i-1,t-1})$, Pareto noise wih $\alpha = 1.5$ and $\alpha = 3.0$, replications = 200





Example: Ratio of largest to second largest, $\lambda_{(1)}/\lambda_{(2)}$:

Recall:

$$\frac{\lambda_{(1)}}{\lambda_{(2)}} \xrightarrow{d} \begin{cases} 4, & \text{if } 8\Gamma_2^{-2/\alpha} < 2\Gamma_1^{-2/\alpha}, \\ \dfrac{\Gamma_1^{-2/\alpha}}{\Gamma_2^{-2/\alpha}}, & \text{otherwise} \end{cases}$$

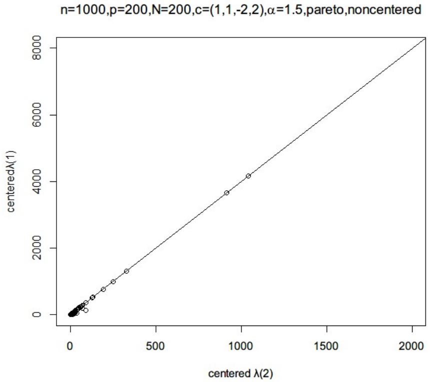
It follows that

$$\begin{aligned} \lim_{n\to\infty} P(\lambda_{(1)} = 4\lambda_{(2)}) &= P(2\Gamma_1^{-2/\alpha} > 8\Gamma_2^{-2/\alpha}) \\ &= P(\frac{E_1}{E_1+E_2} < 2^{-\alpha}) = 2^{-\alpha} = .354(\alpha = 1.5) \end{aligned}$$

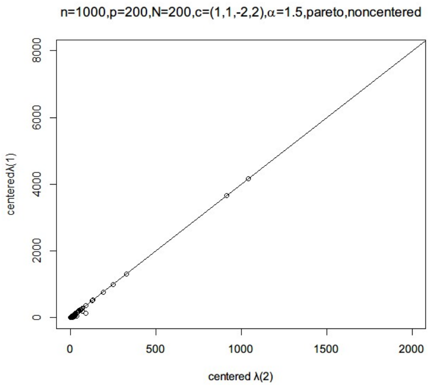
and

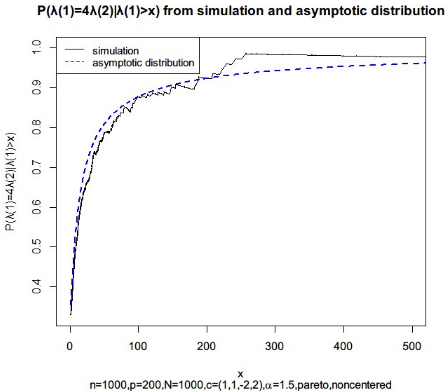
$$\lim_{n\to\infty} P(\lambda_{(1)} = 4\lambda_{(2)} | \lambda_{(1)} > a_{np}^2 x) = P(\frac{E_1}{E_1+E_2} < 2^{-\alpha} | E_1 < 8x^{-\alpha/2}).$$

$$\lim_{n\to\infty} P(\lambda_{(1)} = 4\lambda_{(2)}) \quad = \quad P\left(\frac{E_1}{E_1 + E_2} < 2^{-\alpha}\right) = 2^{-\alpha} = .354(\alpha = 1.5)$$



$$\lim_{n\to\infty} P(\lambda_{(1)} = 4\lambda_{(2)}) \quad = \quad P(\frac{E_1}{E_1 + E_2} < 2^{-\alpha}) = 2^{-\alpha} = .354(\alpha = 1.5)$$





$$\lim_{n\to\infty} P(\lambda_{(1)} = 4\lambda_{(2)} | \lambda_{(1)} > a_{np}^2 x) = P(\tfrac{E_1}{E_1+E_2} < 2^{-\alpha} | E_1 < 8x^{-\alpha/2}).$$

Growth conditions on p_n

Case $P(Z_1 > x) \sim cx^{-\alpha}$: Here $L_2(x) = C\log(x)$.

- For $\alpha \in (0, 2)$,

$$p_n = O(n^\beta), \quad \text{for any } \beta > 0.$$

  Can allow for a touch faster growth rate ($p_n = O(\exp\{c_n\})$, where $c_n^2/n \to 0$ in the $\alpha \in (0, 1)$ case.

- For $\alpha \in (2, 4)$,

$$p_n = O(n^\beta), \quad \beta \in (0, (4-\alpha)/[2(\alpha-1)]).$$

  This excludes the case $p_n \sim cn$.

Elements of the proof I:

Special case: $X_{i,t} = \theta_0 Z_{i,t} + \theta_1 Z_{i-1,t}$

$$\begin{aligned}\sum_{t=1}^{n} X_{it}^2 &= \sum_{t=1}^{n} \underbrace{\theta_0^2 Z_{i,t}^2 + \theta_1^2 Z_{i-1,t}^2}_{\text{tail index } \alpha/2} + 2\theta_0\theta_1 \sum_{t=1}^{n} \underbrace{Z_{i,t} Z_{i-1,t}}_{\text{tail index } \alpha} \\ &= \theta_0^2 D_i + \theta_1^2 D_{i-1} + o_p(a_{np}^2)\end{aligned}$$

and

$$\begin{aligned}\sum_{t=1}^{n} X_{it} X_{i+1,t} &= \theta_0\theta_1 \sum_{t=1}^{n} Z_{i,t}^2 + o_p(a_{np}^2) \\ &= \theta_0\theta_1 D_i + o_p(a_{np}^2)\end{aligned}$$

Elements of the proof I:

Special case: $X_{i,t} = \theta_0 Z_{i,t} + \theta_1 Z_{i-1,t}$

$$\begin{aligned}\sum_{t=1}^{n} X_{it}^2 &= \sum_{t=1}^{n} \underbrace{\theta_0^2 Z_{i,t}^2 + \theta_1^2 Z_{i-1,t}^2}_{\text{tail index } \alpha/2} + 2\theta_0\theta_1 \sum_{t=1}^{n} \underbrace{Z_{i,t} Z_{i-1,t}}_{\text{tail index } \alpha} \\ &= \theta_0^2 D_i + \theta_1^2 D_{i-1} + o_p(a_{np}^2)\end{aligned}$$

and

$$\begin{aligned}\sum_{t=1}^{n} X_{it} X_{i+1,t} &= \theta_0\theta_1 \sum_{t=1}^{n} Z_{i,t}^2 + o_p(a_{np}^2) \\ &= \theta_0\theta_1 D_i + o_p(a_{np}^2)\end{aligned}$$

$$\begin{pmatrix} \mathbf{X}_i^T \mathbf{X}_i & \mathbf{X}_{i+1}^T \mathbf{X}_i \\ \mathbf{X}_{i+1}^T \mathbf{X}_i & \mathbf{X}_{i+1}^T \mathbf{X}_{i+1} \end{pmatrix} \approx \begin{pmatrix} \theta_0^2 & \theta_0\theta_1 \\ \theta_0\theta_1 & \theta_1^2 \end{pmatrix} D_i + \begin{pmatrix} \theta_1^2 & 0 \\ 0 & 0 \end{pmatrix} D_{i-1} + \begin{pmatrix} 0 & 0 \\ 0 & \theta_0^2 \end{pmatrix} D_{i+1}$$

The covariance matrix can be *approximated* by

$$XX^T = \sum_{i=1}^{p} D_i M_i + o_p(a_{np}^2),$$

where M_i is the $p \times p$ matrix consisting of all zeros except for a 2×2 matrix,

$$M = \begin{pmatrix} \theta_0^2 & \theta_0\theta_1 \\ \theta_0\theta_1 & \theta_1^2 \end{pmatrix},$$

whose NW corner is pinned to the i^{th} position on the diagonal. For example,

$$M_1 = \begin{pmatrix} \theta_0^2 & \theta_0\theta_1 & 0 & \cdots & 0 \\ \theta_0\theta_1 & \theta_1^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & \theta_0^2 & \theta_0\theta_1 & \cdots & 0 \\ 0 & \theta_0\theta_1 & \theta_1^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Denote the order statistics of the D_i's by $D_{(1)} \geq D_{(2)} \geq \cdots \geq D_{(p)}$ and write $D_{L_i} = D_{(i)}$.

Then,

- $XX^T = \sum_{i=1}^{p} D_{L_i} M_{L_i} + o_p(a_{np}^2)$ in the sense that

$$a_{np}^{-2} \| XX^T - \sum_{i=1}^{p} D_{L_i} M_{L_i} \|_2 \xrightarrow{P} 0 \, ,$$

  where

$$\|A\|_2 = \sqrt{\text{largest eigenvalue of } AA^T} \quad \text{(operator 2-norm).}$$

Stochastic volatility models—special case

Suppose the rows are independent copies of the SV process given by

$$X_t = \sigma_t Z_t$$

where (Z_t) is iid RV(α) and $(\ln \sigma_t^2)$ is a purely nondeterministic stationary Gaussian process (this can be weakened), independent of (Z_t).

Theorem *Suppose $p_n, n \to \infty$ such that*

$$\limsup_{n\to\infty} \frac{p_n}{n^\beta} < \infty \,, \quad \textit{for some } \beta > 0 \textit{ satisfying}$$

1. $\beta < \infty$ *if* $\alpha \in (0,1)$*, and*
2. $\beta < \frac{2-\alpha}{\alpha-1}$ *if* $\alpha \in (1,2)$.

Then, we have the point process convergence,

$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_{(i)}} \xrightarrow{d} N = \sum_{i=1}^{\infty} \epsilon_{\Gamma_i^{-2/\alpha}} \,.$$

Stochastic volatility models—special case

Point process convergence:

$$N_p := \sum_{i=1}^{p} \epsilon_{a_{np}^{-2}\lambda_{(i)}} \xrightarrow{d} N = \sum_{i=1}^{\infty} \epsilon_{\Gamma_i^{-2/\alpha}} .$$

Remarks:

- Proof uses a large deviation result of Davis and Hsing (1995); see also Mikosch and Wintenberger (2012).
- Likely that we can weaken the restriction on β
- Similar results hold for GARCH processes if X_t is RV(α) with $\alpha \in (0, 2)$.

References

Richard A. Davis, Oliver Pfaffel and Robert Stelzer
Limit Theory for the largest eigenvalues of sample covariance matrices with heavy-tails. Stoch. Proc. Appl. 24, 2014, 18–50.

Richard A. Davis, Thomas Mikosch, and Oliver Pfaffel
Asymptotic Theory for the Sample Covariance Matrix of a Heavy-Tailed Multivariate Time Series. Preprint 2013.

Ian M. Johnstone
On the Distribution of the Largest Eigenvalue in Principal Components Analysis. Ann. Statist., 2001, 29, 295-327.

Alexander Soshnikov
Poisson Statistics for the Largest Eigenvalues in Random Matrix Ensembles. Lect. Notes in Phys., 2006, 690, 351-364.

Antonio Auffinger, Gérard Ben Arous, and Sandrine Péché
Poisson convergence for the largest eigenvalues of heavy tailed random matrices. Ann. Inst. H. Poincaré Probab. Statist., 2009, 45, 589-610.