Extremes of Stationary Sequences: Clusters and Spectral Processes

JOHAN SEGERS

Université catholique de Louvain
Institut de statistique, biostatistique et sciences actuarielles

Indian Statistical Institute, Kolkata, January 14–15, 2012

Extremes of Stationary Sequences: Clusters and Spectral processes

I. Clusters of Extremes

II. Regular Variation and Tail Processes
— joint work with B. BASRAK and T. MEINGUET

III. Markov Processes
— joint work with A. JANSSEN

IV. Linear Processes
— joint work with T. MEINGUET

Part I

Clusters of Extremes

An informal view on clusters

For weakly dependent stationary sequences,
extremes arrive in clusters.

We are concerned with the asymptotic distribution of the 'block'

$$(X_1, \ldots, X_{r_n})$$

given that at least one 'extreme value' occurs

$$\sum_{i=1}^{r_n} I(X_i \text{ hits an exceptional set}) \geqslant 1 \qquad \text{(C)}$$

when the expected number of extremes is asymptotically negligible

$$r_n \, P(X_1 \text{ hits an exceptional set}) = o(1)$$

Formalizing the informal view requires some care

- The condition (C) is awkward to work with:
  when did the extreme value occur for the first time?
- If the expected number of extremes in a block remains finite, most variables X_i in the block $(X_1, \ldots, X_{r_n})$ will be irrelevant.

Formalizing the notion of a 'cluster' therefore requires some care.
Some possibilities:

- Cluster functionals
- Cluster distributions
- Cluster processes

Extremes of Stationary Sequences

Cluster functionals and the cluster map

Approximate cluster distributions

Limit cluster distributions

Beyond the cluster

Cluster statistics

Ingredients

- Stationary process $(X_n)_n$ on $\mathbb{R}$
- High threshold u_n
- Block size r_n

Interest is in cluster statistics of the form

$$c(X_1 - u_n, \ldots, X_{r_n} - u_n) \quad \text{conditionally on} \quad M_{r_n} > u_n$$

that only depend on the 'cluster':
the stretch between the first and the last exceedance over u_n.

We require that

$$r_n \to \infty, \qquad r_n \, P(X_1 > u_n) \to 0$$

Examples of cluster statistics

- Block maximum: maximal excess

$$c(y_1, \ldots, y_{r_n}) = \max(y_1, \ldots, y_{r_n})$$

- Aggregate excess: sum of excesses

$$c(y_1, \ldots, y_{r_n}) = \max(y_1, 0) + \cdots + \max(y_{r_n}, 0)$$

- Cluster size: number of excesses

$$c(y_1, \ldots, y_{r_n}) = I(y_1 > 0) + \cdots + I(y_{r_n} > 0)$$

- Cluster duration: time span between first and last excess

$$c(y_1, \ldots, y_{r_n}) = \max\{i : y_i > 0\} - \min\{i : y_i > 0\} + 1$$

- Number of threshold upcrossings

$$c(y_1, \ldots, y_{r_n}) = I(y_1 > 0) + I(y_1 \leqslant 0 < y_2) + \cdots + I(y_{r_n - 1} \leqslant 0 < y_{r_n})$$

Cluster functionals

Desirable properties of $c(\,\cdot\,)$:

- Its domain is a vector of arbitrary length
  with at least one non-zero component.
- It depends only on the 'extreme' part of the vector

Definition

A cluster functional is a map $c : \boldsymbol{A} \to \mathbb{R}$ with

$$\boldsymbol{A} = \boldsymbol{A}_1 \cup \boldsymbol{A}_2 \cup \dots$$
$$\boldsymbol{A}_r = \mathbb{R}^r \setminus (-\infty, 0]^r = \{(y_1, \dots, y_r) \in \mathbb{R}^r : \max(y_1, \dots, y_r) > 0\}$$

and neglecting everything that happened
before or after the first or last positive value:

$$c(y_1, \dots, y_r) = c(y_\alpha, \dots, y_\omega)$$
$$\alpha = \min\{i : y_i > 0\}$$
$$\omega = \max\{i : y_i > 0\}$$

[Yun 2000; Segers 2003; Drees & Rootzén 2010]

Cluster map

Definition

Recall $\boldsymbol{A} = \bigcup_{r \geqslant 1} \boldsymbol{A}_r$ and $\boldsymbol{A}_r = \mathbb{R}^r \setminus (-\infty, 0]^r$. Define the cluster map

$$C : \boldsymbol{A} \to \boldsymbol{A} : (y_1, \ldots, y_r) \mapsto (y_\alpha, \ldots, y_\omega)$$
$$\alpha = \min\{i : y_i > 0\}$$
$$\omega = \max\{i : y_i > 0\}$$

[Segers 2005]

Then $c : \boldsymbol{A} \to \mathbb{R}$ is a cluster functional if and only if

$$c = f \circ C \text{ for some } f : \boldsymbol{A} \to \mathbb{R}$$

Hence, to know the asymptotic distribution of cluster statistics,
it is sufficient to know the asymptotic distribution of the 'cluster' itself

$$C(X_1 - u_n, \ldots, X_{r_n} - u_n) \quad \text{conditionally on} \quad M_{r_n} > u_n$$

Extremes of Stationary Sequences

Cluster functionals and the cluster map

Approximate cluster distributions

Limit cluster distributions

Beyond the cluster

Aim: switch to a simpler conditioning event

We are interested in the cluster distribution

$$P[C(X_1 - u_n, \ldots, X_{r_n} - u_n) \in \cdot \mid M_{r_n} > u_n]$$

Recall $r_n \to \infty$ and $r_n P(X_1 > u_n) \to 0$.

The conditioning event $\{M_{r_n} > u_n\}$ is awkward to work with:
when exactly did the exceedances occur?

We'd rather prefer expressions in terms of the law of

$$(X_1, \ldots, X_k) \mid X_1 > u_n$$

This would be particularly convenient in the case of Markov chains.

Expected cluster size

Expected number of exceedances given that there is at least one:

$$E\left[\sum_{i=1}^{r_n} I(X_i > u_n) \,\middle|\, M_{r_n} > u_n\right] = \frac{r_n\, P(X_1 > u_n)}{P(M_{r_n} > u_n)} =: \frac{1}{\theta_n}$$

so

$$\theta_n = \frac{P(M_{r_n} > u_n)}{r_n\, P(X_1 > u_n)} \in (0, 1]$$

Example

In the iid case, since $r_n\, \overline{F}(u_n) \to 0$, we have

$$\theta_n = \frac{1 - (1 - \overline{F}(u_n))^{r_n}}{r_n \overline{F}(u_n)} \to 1$$

Finite-cluster condition

Suppose that the impact of a shock is somehow limited in time:

$$\underbrace{X_1}_{>u_n}, X_2, \ldots, X_m, \underbrace{X_{m+1}, \ldots, X_{r_n}}_{>u_n?}$$

$$\underbrace{X_1, \ldots, X_{r_n-m}}_{>u_n?}, X_{r_n-m+1}, \ldots, X_{r_n-1}, \underbrace{X_{r_n}}_{>u_n}$$

Formally, put $M_{i,j} = \max(X_i, \ldots, X_j)$ and suppose

$$\lim_{m\to\infty} \limsup_{n\to\infty} P(M_{m+1,r_n} > u_n \mid X_1 > u_n) = 0 \qquad \text{(FiCl1)}$$

$$\lim_{m\to\infty} \limsup_{n\to\infty} P(M_{1,r_n-m} > u_n \mid X_{r_n} > u_n) = 0 \qquad \text{(FiCl2)}$$

Sufficient condition:

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sum_{i=m+1}^{r_n} P(X_i > u_n \mid X_1 > u_n) = 0 \qquad \text{(FiCl)}$$

Bounded expected cluster sizes

If (FiCl), the expected cluster size remains bounded:

$$\limsup_{n\to\infty} \frac{r_n\, P(X_1 > u_n)}{P(M_{r_n} > u_n)} < \infty$$

$$\text{i.e. } \liminf_{n\to\infty} \theta_n > 0.$$

Proof: observe that $M_{r_n} \geqslant \max(X_1, X_{m+1}, X_{2m+1}, \ldots, X_{km+1})$ with $k \sim r_n/m$.

The approximant

Consider a bounded, measurable cluster functional $c : \boldsymbol{A} \to \mathbb{R}$.
Apply c to different stretches of the process:

$$c_n(i,j) = c(X_i - u_n, \ldots, X_j - u_n) \text{ on the event } M_{i,j} > u_n$$

Consider the approximation error

$$\left| \underbrace{E[c_n(1, r_n) \mid M_{r_n} > u_n]}_{\text{quantity of interest}} - \underbrace{\frac{\alpha_{n,m}(c)}{\theta_{n,m}}}_{\text{approximant}} \right|$$

where

$$\begin{aligned} \alpha_{n,m}(c) &= E[c_n(1,m) \mid X_1 > u_n] \\ &\qquad - E[c_n(2,m),\ M_{2,m} > u_n \mid X_1 > u_n] \\ \theta_{n,m} &= P[M_{2,m} \leqslant u_n \mid X_1 > u_n] \qquad \text{`runs'} \end{aligned}$$

The cluster approximation

Theorem

If (FiCl), then

$$\lim_{m\to\infty} \limsup_{n\to\infty} | \underbrace{\theta_{n,m}}_{\text{'runs'}} - \underbrace{\theta_n}_{\text{'blocks'}} | = 0$$

as well as

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sup_{c:|c|\leqslant 1} \left| E[c_n(1, r_n) \mid M_{r_n} > u_n] - \frac{\alpha_{n,m}(c)}{\theta_{m,n}} \right| = 0$$

[Segers (2005)]

Proof: elementary calculations, based on careful use of

- partitionings of the event $\{M_{r_n} > u_n\}$ and similar ones
- stationarity
- the cluster property
- (FiCl)

Main steps in the proof (1)

Consider the first time an exceedance occurs:

$$
\begin{aligned}
& E[c_n(1, r_n);\ M_{r_n} > u_n] \\
&= \sum_{j=1}^{r_n} E[c_n(j, r_n);\ M_{j-1} \leqslant u_n < X_j]
\end{aligned}
$$

By (FiCl), we can limit the (forward) horizon to m:

$$
\ldots \approx \sum_{j=1}^{r_n} E[c_n(j, j+m-1);\ M_{j-1} \leqslant u_n < X_j]
$$

Write each term as a difference by taking out the event $M_{j-1} \leqslant u_n$:

$$
\begin{aligned}
& E[c_n(j, j+m-1);\ \qquad\qquad\quad X_j > u_n] \\
-& E[c_n(j, j+m-1);\ M_{j-1} > u_n,\ X_j > u_n]
\end{aligned}
$$

By stationarity, the first term is already OK: $j = 1$.
What about the second term?

Main steps in the proof (2)

We need to consider

$$E[c_n(j, j+m-1);\ M_{1,j-1} > u_n,\ X_j > u_n]$$

By (FiCl), we can limit the (backward) horizon to m:

$$\ldots \approx E[c_n(j, j+m-1);\ M_{j-m,j-1} > u_n,\ X_j > u_n]$$

By stationarity (set $j = m+1$), this is

$$\ldots = E[c_n(m+1, 2m+1);\ M_{1,m} > u_n,\ X_{m+1} > u_n]$$

In $\{M_m > u_n\}$, consider the last time an exceedance occurs, apply stationarity, (FiCl), eventually yielding

$$\ldots \approx E[c_n(2, m);\ X_1 > u_n, M_{2,m} > u_n]$$

which is the second term in $\alpha_{n,m}(c)$.

Main steps in the proof (3)

Collect approximations to find that

$$E[c_n(1, r_n);\, M_{r_n} > u_n] \approx r_n\, \alpha_{n,m}(c)$$

Consider the special case $c \equiv 1$ to get

$$\theta_{n,m} \approx \theta_n$$

Combine the previous two displays to arrive at the desired approximation.

Without additional effort, the result is translated in a general framework

- Measurable state space $(S, \mathscr{S})$
- Measurable failure set $B \subset S$
- $\boldsymbol{A} = \bigcup_{k \geqslant 1} \boldsymbol{A}_k$ where $\boldsymbol{A}_k = S^k \setminus (S \setminus B)^k$
- Cluster map $C : \boldsymbol{A} \to \boldsymbol{A}$ is defined by

$$C(x_1, \ldots, x_k) = (x_\alpha, \ldots, x_\omega)$$

where

- $\alpha = \min\{i = 1, \ldots, k \,:\, x_i \in B\}$
- $\omega = \max\{i = 1, \ldots, k \,:\, x_i \in B\}$

The general framework encompasses multivariate extremes

Univariate extremes:

- state space $S = \mathbb{R}$
- failure set $B = (u, \infty)$

Multivariate extremes:

- state space $S = \mathbb{R}^d$
- failure sets $B = \mathbb{R}^d \setminus (-\boldsymbol{\infty}, \boldsymbol{u}]$ or $(\boldsymbol{u}, \boldsymbol{\infty})$ or $\{\boldsymbol{x} : \|\boldsymbol{x}\| > u\}$ or ...

What if the failure set is hit at least once?

- Stationary random vector $(X_1, \ldots, X_r)$ in S
- Assume $P[X_1 \in B] > 0$

Aim

To study the conditional distribution of

$$C(X_1, \ldots, X_r) \qquad \text{given} \qquad \bigcup_{i=1}^{r} \{X_i \in B\}$$

Cluster functionals and cluster map

A map $c : \boldsymbol{A} \to \mathbb{R}$ is a cluster functional
if it is measurable with respect to the cluster map, i.e.

$$c = f \circ C \qquad \text{for some} \qquad f : \boldsymbol{A} \to \mathbb{R}$$

that is, if

$$c(x_1, \ldots, x_r) = c(x_\alpha, \ldots, x_\omega)$$

in terms of the first and last hitting times, $1 \leqslant \alpha \leqslant \omega \leqslant r$ of B.

Cluster functionals and the cluster map are equivalent concepts:
for $E \subset \boldsymbol{A}$,

$$C(x_1, \ldots, x_r) \in E \iff \underbrace{I_E \circ C}_{=c}(x_1, \ldots, x_r) = 1$$

Extremal index variants

- Expected number of 'hits' of failure set B

$$E\left[\sum_{i=1}^{r} I\{X_i \in B\} \,\middle|\, \bigcup_{i=1}^{r}\{X_i \in B\}\right] = \frac{r\,P[X_1 \in B]}{P[\bigcup_{i=1}^{r}\{X_i \in B\}]} = \frac{1}{\theta}$$

- 'Hit' followed/preceded by a 'run' of 'non-hits' of failure set B

$$\begin{aligned} \theta_m &= P[\textstyle\bigcap_{i=2}^{m}\{X_i \notin B\} \mid X_1 \in B] \\ &= P[\textstyle\bigcap_{i=1}^{m-1}\{X_i \notin B\} \mid X_m \in B], \qquad m = 2, \ldots, r \end{aligned}$$

- Compare these with characterizations of extremal index
  - 'blocks' [Leadbetter 1983]
  - 'runs' [O'Brien 1987]
  - Multivariate extremal index [Nandagopalan 1994]

Approximate cluster distribution

- $\mathscr{C}$ is set of all cluster functionals $c : \boldsymbol{A} \to \mathbb{R}$ such that $|c| \leqslant 1$
- Cluster distribution: for $c \in \mathscr{C}$

$$\mu(c) = E[c(X_1, \ldots, X_r) \mid \bigcup_{i=1}^{r} \{X_i \in B\}]$$

- Approximant: for $c \in \mathscr{C}$

$$\begin{aligned} & \mu_m(c) \\ & = \quad \theta^{-1} \Big\{ E[c(X_1, \ldots, X_m) \mid X_1 \in B] \\ & \qquad - E[c(X_2, \ldots, X_m) I(\bigcup_{i=2}^{m} \{X_i \in B\}) \mid X_1 \in B] \Big\} \end{aligned}$$

Finite-sample cluster distribution approximation

Quantify (FiCl) via

$$\begin{aligned}\varepsilon = \max\{ & P[\textstyle\bigcup_{i=m+1}^{r}\{X_i \in B\} \mid X_1 \in B], \\ & P[\textstyle\bigcup_{i=1}^{r-m}\{X_i \in B\} \mid X_r \in B]\}\end{aligned}$$

Theorem

If $m \geqslant 2$ and $2m + 1 \leqslant r$,

$$\begin{aligned} \theta &\geqslant (2m)^{-1}(1-\varepsilon) \\ |\theta - \theta_m| &\leqslant \max(m/r, \varepsilon) \\ \sup_{c:|c|\leqslant 1} |\mu(c) - \mu_m(c)| &\leqslant \theta^{-1}(4m/r + 5\varepsilon) \end{aligned}$$

[Segers 20xx]

Interpretation: connection between distributions of

- $C(X_1, \ldots, X_r)$ given $\bigcup_{i=1}^{r}\{X_i \in B\}$
- $(X_1, \ldots, X_m)$ given $\{X_1 \in B\}$

Extremes of Stationary Sequences

Cluster functionals and the cluster map

Approximate cluster distributions

Limit cluster distributions

Beyond the cluster

Asymptotic cluster distribution

- State space: metric space (S, d)
- Failure set: non-empty open set $B \subset S$
- Random triangular array $\{X_{in} \,:\, n \geqslant 1,\ 1 \leqslant i \leqslant r_n\}$ in S
  - row length $r_n \to \infty$
  - every row $(X_{1n}, \ldots, X_{r_n n})$ is stationary
  - $p_n = P[X_{1n} \in B] > 0$
  - $r_n p_n = E[\sum_{i=1}^{r_n} I(X_{in} \in B)] \to 0$

Aim

To establish the limiting cluster distribution

$$C(X_{1n}, \ldots, X_{r_n n}) \quad \text{given} \quad \bigcup_{i=1}^{r_n} \{X_{in} \in B\}$$

with $C : \boldsymbol{A} \to \boldsymbol{A}$ the cluster map and $\boldsymbol{A} = \bigcup_{r \geqslant 1} \left(S^r \setminus (S \setminus B)^r\right)$

Example

- State space $S = \mathbb{R}$
- Failure set $B = \{x : |x| > 1\}$
- Random variables $X_{in} = X_i / a_n$, $1 \leqslant i \leqslant r_n$, with
  - $(X_i)_{i \geqslant 1}$ a stationary time series in $\mathbb{R}$
  - levels $0 < a_n \to \infty$ such that $nP[|X_1| > a_n] \to 1$
  - block sizes $r_n \to \infty$ and $r_n = o(n)$
- Rare events of interest:
  - $X_{in} \in B$ if and only if $|X_i| > a_n$
  - $\bigcup_{i=1}^{r_n} \{X_{in} \in B\}$ if and only if $M_{r_n} := \max(|X_1|, \ldots, |X_{r_n}|) > a_n$

Problem

To find the asymptotic cluster distribution

$$C(X_1/a_n, \ldots, X_{r_n}/a_n) \qquad \text{given} \qquad M_{r_n} > a_n?$$

Example (continued)

- Assume that the fidis of $(X_i)_i$ are multivariate regularly varying.
- Then there exists a process $(Y_k)_{k \geqslant 0}$ such that for every $k \geqslant 0$,

$$P[(X_1/a_n, \ldots, X_{k+1}/a_n) \in \cdot \mid |X_1| > a_n] \xrightarrow{d} P[(Y_0, \ldots, Y_k) \in \cdot]$$

  - Conceptually, given $|X_1| > a_n$,

$$\begin{array}{llll} X_1/a_n, & X_2/a_n, & \ldots, & X_{k+1}/a_n \\ Y_0, & Y_1, & \ldots, & Y_k \\ \text{`present'}, & \text{`future'} & & \end{array}$$

  - For Markov chains, the process $(Y_k)_{k \geqslant 0}$ can typically be written in terms of a random walk

    [Rootzén 1988; de Haan et al. 1989; Smith 1992; Perfekt 1994; S. 2007; Resnick and Zeber 2011]

- Can we express the asymptotic cluster distribution in terms of the tail process $(Y_k)_k$?

Assumptions

Tail process

Assume there exists a random sequence $(Y_k)_{k \geqslant 0}$ called tail process in S such that for every $k \geqslant 0$,

$$P[(X_{1n}, \ldots, X_{k+1,n}) \in \cdot \mid X_{1n} \in B] \xrightarrow{d} P[(Y_0, \ldots, Y_k) \in \cdot\,].$$

Also, assume $P[Y_k \in \partial B] = 0$ for all $k \geqslant 0$.

Finite cluster condition

The impact of a ‘hit’ does not last for too long:

$$\lim_{m\to\infty} \limsup_{n\to\infty} P[\textstyle\bigcup_{i=m+1}^{r_n} \{X_{in} \in B\} \mid X_{1n} \in B] = 0$$

$$\lim_{m\to\infty} \limsup_{n\to\infty} P[\textstyle\bigcup_{i=1}^{r_n - m} \{X_{in} \in B\} \mid X_{r_n n} \in B] = 0$$

Limiting cluster distributions

Theorem

[Segers 20xx] Under the above assumptions:

- *The tail process $(Y_k)_{k\geqslant 0}$ hits B only finitely often:*

$$Y_0 \in B \qquad \textit{and} \qquad \sharp\{k \geqslant 1 : Y_k \in B\} < \infty \qquad \textit{a.s.}$$

- *The expected number of hits converges to finite limit:*

$$\theta_n = 1/E[\textstyle\sum_{i=1}^{r_n} I(X_{in} \in B) \mid \bigcup_{i=1}^{r_n}\{X_{in} \in B\}] \\ \to P[\forall k \geqslant 1 : Y_k \notin B] =: \theta > 0$$

- *The cluster distribution converges:*

$$P[C((X_{in})_{i=1}^{r_n}) \in \cdot \mid \textstyle\bigcup_{i=1}^{r_n}\{X_i \in B\}] \\ \xrightarrow{d} \quad \theta^{-1}\Big\{P[C((Y_k)_{k\geqslant 0}) \in \cdot\,] \\ - P[\{C((Y_k)_{k\geqslant 1}) \in \cdot\,\} \cap \textstyle\bigcup_{k\geqslant 1}\{Y_k \in B\}]\Big\}$$

Extremes of Stationary Sequences

Cluster functionals and the cluster map

Approximate cluster distributions

Limit cluster distributions

Beyond the cluster

The prologue and epilogue of a cluster

- By definition, the 'cluster' starts and ends with the first and last extreme value in a block.
- What happened just before? What happens next? Maybe there are some 'less extreme' but still interesting values.

Aim

To find the (asymptotic) distribution of the *whole* block

$$X_1, \ldots, X_r \qquad \text{conditionally on} \qquad \exists i = 1, \ldots, r : X_i \in B$$

Challenges

- By (FiCl), however, most variables X_i will somehow 'vanish' asymptotically.
- The interesting observations will occur at some random time instant in the middle of the block.

Framework

- Metric space (S, d)
- Failure set $B = \{x \in S : d(x, q) > 1\}$ for some $q \in S$
- Random triangular array $\{X_{in} : n \geqslant 1,\ 1 \leqslant i \leqslant r_n\}$ in S
  - row length $r_n \to \infty$
  - stationary rows $(X_{1n}, \ldots, X_{r_n n})$
  - failure probability $p_n = P[d(X_{1n}, q) > 1] > 0$
  - $r_n p_n = E[\sum_{i=1}^{r_n} I\{d(X_{in}, q) > 1\}] \to 0$
- The point q acts as a 'black hole' for non-extreme values

Example

- $S = \mathbb{R}^d$
  $q = 0$
  $B = \{x \in \mathbb{R}^d : |x| > 1\}$
- $X_{in} = X_i / a_n$ for some sequence $0 < a_n \to \infty$

Problem statement

Aim

To find the limit distribution of quantities defined in terms of $(X_{1n}, \ldots, X_{r_n n})$ given $\bigcup_{i=1}^{r_n} I\{d(X_{in}, q) > 1\}$?

Example

Cluster point process N_n on $S \setminus \{q\}$:

$$N_n = \sum_{i=1}^{r_n} \delta_{X_{in}} \qquad \text{given} \qquad \bigcup_{i=1}^{r_n} \{d(X_{in}, q) > 1\}$$

Cluster process

On the event $\bigcup_{i=1}^{r_n} \{d(X_{in}, q) > 1\}$:

- First hitting time $\alpha_n = \min\{i = 1, \ldots, r_n \, : \, d(X_{in}, q) > 1\}$
- Cluster process $\xi_n = (\xi_{n,t})_{t \in \mathbb{Z}}$

$$\xi_{n,t} = \begin{cases} X_{\alpha_n + t, n} & \text{if } 1 \leqslant \alpha_n + t \leqslant r_n \\ q & \text{otherwise} \end{cases}$$

Intuitively, the vector $(X_{1n}, \ldots, X_{r_n n})$ is

- 'anchored' at the first hitting time α_n of the failure set;
- extended on the left and on the right by the constant sequence (q)

$$\begin{array}{ccccccccc} \ldots, & q, & X_{1n}, & \ldots, X_{\alpha_n - 1, n}, & X_{\alpha_n, n}, & X_{\alpha_n + 1, n}, \ldots, & X_{r_n, n}, & q, & \ldots \\ \ldots, \xi_{n, -\alpha_n}, & \xi_{n, -\alpha_n + 1}, \ldots, & \xi_{n,-1}, & & \xi_{n,0}, & \xi_{n,1}, & \ldots, \xi_{n, r_n - \alpha_n + 1}, & \xi_{n, r_n - \alpha_n}, \ldots & \end{array}$$

Mathematical problem statement

To establish weak convergence of
the cluster process ξ_n in the space $(\mathbb{E}, \rho)$, where

$$\begin{aligned} \mathbb{E} &= \{x \in S^{\mathbb{Z}} : d(x_0, q) > 1 \text{ and } x_t \to q \text{ as } t \to \pm\infty\} \\ \rho(x, y) &= \sup_{t \in \mathbb{Z}} d(x_t, y_t) \end{aligned}$$

- $\mathbb{E}$ is the space of S-valued sequences converging to q.
- The metric ρ induces the topology of uniform convergence.

Tentative application: point process convergence

Since the cluster point process N_n on $S \setminus \{q\}$ admits the representation

$$N_n = \sum_{i=1}^{r_n} \delta_{X_{in}} = T(\xi_n)$$

for a continuous map

$$\begin{array}{rccc} T: & (\mathbb{E}, e) & \to & M_p(S \setminus \{q\}) \\ & (x_t)_{t\in\mathbb{Z}} & \mapsto & \sum_{t\in\mathbb{Z}} \delta_{x_t} \end{array}$$

point process convergence would follow
from weak convergence of ξ_n in $\mathbb{E}$

Tentative application: cluster functionals

Recall $\boldsymbol{A} = \bigcup_{r \geqslant 1} \boldsymbol{A}_r$ and $\boldsymbol{A}_r = \{(x_1, \ldots, x_r) : \max_j d(x_j, q) > 1\}$

- disjoint union
- product topology

Consider the projection map

$$
\begin{array}{rccl}
\pi : & \mathbb{E} & \to & \boldsymbol{A} \\
 & (x_t)_t & \mapsto & (x_\alpha, \ldots, x_\omega) \\
 & & & \alpha(x) = \min\{t : d(x_t, q) > 1\} \\
 & & & \omega(x) = \max\{t : d(x_t, q) > 1\}
\end{array}
$$

Since π is continuous, weak convergence in $\mathbb{E}$ of $\xi_n = (\xi_{n,t})_t$ would imply weak convergence in $\boldsymbol{A}$ of the cluster

$$\pi(\xi_n) = (X_{\alpha,n}, \ldots, X_{\omega,n})$$

Assumption: tail process

Assume there exists a random sequence $(Y_t)_{t\in\mathbb{Z}}$ in S such that for every integer $k \geqslant 0$,

$$P[(X_{1n}, \ldots, X_{2k+1,n}) \in \cdot \mid d(X_{k+1,n}, q) > 1] \xrightarrow{d} P[(Y_{-k}, \ldots, Y_k) \in \cdot\,]$$

Schematically, we have

$$\begin{array}{cccc} & X_{1n}, \ldots, X_{k,n}, & X_{k+1,n}, & X_{k+2,n}, \ldots, X_{2k+1,n} \\ \xrightarrow{d} & Y_{-k}, \ldots, Y_{-1}, & Y_0, & Y_1, \ldots, Y_k \\ & \text{'past'} & \text{'present'} & \text{'future'} \end{array}$$

Also, assume $P[d(Y_t, q) = 1] = 0$ for all $t \in \mathbb{Z}$.

Assumption: finite-cluster condition

For all $\delta > 0$, as $m \to \infty$,

$$\left.\begin{array}{r} \limsup_{n\to\infty} P[\bigcup_{i=m+1}^{r_n}\{d(X_{in},q) > \delta\} \mid d(X_{1n},q) > 1] \\ \limsup_{n\to\infty} P[\bigcup_{i=1}^{r_n-m}\{d(X_{in},q) > \delta\} \mid d(X_{r_n n},q) > 1] \end{array}\right\} \to 0$$

This will ensure, among others, that $\lim_{|t|\to\infty} Y_t = q$ a.s.

Weak convergence of the cluster process

Theorem

When the tail process exists and the finite-cluster condition holds,

- *the tail sequence $(Y_t)_{t\in\mathbb{Z}}$ hits the failure set finitely often:*

$$P[d(Y_0, q) > 1,\ Y_t \to q \text{ as } t \to \pm\infty] = 1$$

- *with positive probability, the tail process hits the failure set for the first time at $t = 0$:*

$$\theta = P[\forall t \leqslant -1 : d(Y_t, q) \leqslant 1] > 0$$

- *the cluster process converges weakly in $\mathbb{E}$:*

$$\begin{aligned} &P[\xi_n \in \cdot \mid \textstyle\bigcup_{i=1}^{r_n} \{d(X_{in}, q) > 1\}] \\ &\quad \xrightarrow{d} \quad P[(Y_t)_{t\in\mathbb{Z}} \in \cdot \mid \forall t \leqslant -1 : d(Y_t, q) \leqslant 1] \end{aligned}$$

[Segers 20xx]

Corollary: Point process convergence

Under the conditions of the theorem,

$$N_n \xrightarrow{d} N$$

in $M_p(S \setminus \{q\})$, where

$$\begin{aligned} N_n &\stackrel{d}{=} \textstyle\sum_{i=1}^{r_n} \delta_{X_{in}} \quad \text{given} \quad \bigcup_{i=1}^{r_n} \{d(X_{in}, q) > 1\} \\ N &\stackrel{d}{=} \textstyle\sum_{t \in \mathbb{Z}} \delta_{Y_t} \quad \text{given} \quad \bigcap_{t \leqslant -1} \{d(Y_t, q) \leqslant 1\} \end{aligned}$$

Corollary: Convergence of cluster stretches

Recall the cluster map $C : \boldsymbol{A} \to \boldsymbol{A}$,
with $\boldsymbol{A} = \bigcup_{r \geqslant 1} \boldsymbol{A}_r$ and $\boldsymbol{A}_r = \{(x_1, \ldots, x_r) \in S^r : \max_j d(x_j, q) > 1\}$.

Under the conditions of the theorem, we have

$$
\begin{aligned}
C(X_{1n}, \ldots, X_{r_n n}) &= (X_{\alpha_n, n}, \ldots, X_{\omega_n, n}) \\
&\xrightarrow{d} [(Y_0, \ldots, Y_\tau) \quad \text{given} \quad \forall t \leqslant -1 : d(Y_t, q) \leqslant 1] \\
&\qquad \text{with} \quad \tau = \max\{t \in \mathbb{Z} : d(Y_t, q) > 1\}
\end{aligned}
$$

How does this relate to previous results on cluster functionals?

Linking up with cluster functional theory

For a bounded, continuous cluster functional $c : \boldsymbol{A} \to \mathbb{R}$,

$$
\begin{aligned}
&E[c(X_{1n},\ldots,X_{r_n n}) \mid \exists i = 1,\ldots,r_n : d(X_{in},q) > 1] \\
\to &E[c(Y_0,\ldots,Y_\tau) \mid \forall t \leqslant -1 : d(Y_t,q) \leqslant 1] \\
= &E[c((Y_t)_{t\geqslant 0}) \mid \forall t \leqslant -1 : d(Y_t,q) \leqslant 1] \\
= &\frac{E[c((Y_t)_{t\geqslant 0}); \forall t \leqslant -1 : d(Y_t,q) \leqslant 1]}{P[\forall t \leqslant -1 : d(Y_t,q) \leqslant 1]} \\
= &\frac{1}{\theta}\{E[c((Y_t)_{t\geqslant 0})] - E[c((Y_t)_{t\geqslant 0}); \exists t \leqslant -1 : d(Y_t,q) > 1]\}
\end{aligned}
$$

However, by the earlier limiting-cluster-distribution theorem,

$$
\begin{aligned}
&E[c(X_{1n},\ldots,X_{r_n n}) \mid \exists i = 1,\ldots,r_n : d(X_{in},q) > 1] \\
&\qquad \to \frac{1}{\theta}\{E[c((Y_t)_{t\geqslant 0})] - E[c((Y_t)_{t\geqslant 1}); \exists t \geqslant 1 : d(Y_t,q) > 1]\}
\end{aligned}
$$

Equality follows from a ‘time-change formula’.

Summary: Cluster of extremes

- Description via cluster functionals or the cluster map
- General state space
- Change of conditioning event:

  From: Conditional distribution of an excited block

  To: Conditional distribution of a stretch given an excited initial value
- Approximate cluster distributions
- Limiting cluster distributions if the tail process exists
- Looking beyond the cluster: convergence in sequence space
  - First hitting time serves as time origin

Part II

Regular Variation and Tail Processes
— with B. Basrak and T. Meinguet

Tail processes and spectral processes: Concise descriptions of extremal dependence

- Point processes of extremes [Davis & Hsing 1995; Davis & Mikosch 1998; Basrak & S. 2009]
- Cluster functionals [Yun 2000; S. 2003]
- Extremograms [Davis & Mikosch 2009]
- Empirical tail processes [Drees & Rootzén 2010]
- Joint survival functions, tail dependence coefficients [S. 2007; Meinguet & S. 2010]
- Large deviations [Mikosch & Wintenberger 2012a,b]
- Central limit theorems with non-Gaussian stable limits [Barkiewicz et al. 2011; Basrak, Krizmanić & S. 2012]
- ...

Time series of random functions: Dependence over space in time

Physical quantity observed in space and over time

$$X_t(x) = \text{value at time } t \text{ at location } x$$

Space coordinate x varies over a grid – *high-dimensional!*

Think of x as varying continously over space
$\rightsquigarrow$ For fixed t, view $X_t(\,\cdot\,)$ as a *random function*
$\rightsquigarrow$ Time series $\big(X_t(\,\cdot\,)\big)_{t\in\mathbb{Z}}$ of random functions

Goal: to model

Space – cross-sectional tail dependence

Time – clusters

The proper function space depends on the context

- Maximal temperature over $S \subset [0,1]^2$:

$$\sup_{x \in S} X_t(x)$$

$\rightsquigarrow$ Space of $C([0,1]^2)$ of *continuous functions*

- Aggregated rainfall over $S \subset [0,1]^2$:

$$\int_S X_t(x)\, dx$$

$\rightsquigarrow$ Space $L^1([0,1]^2)$ of *integrable functions*

Extremes of Stationary Sequences

Regular variation

Heavy tails: power-law behaviour

Mathematical description: regular variation

space	*tail*
$\mathbb{R}$	$x \to \infty$
$\mathbb{R}$	$\lvert x \rvert \to \infty$
$\mathbb{R}^d$	$\max_j x_j \to \infty$
$\mathbb{R}^d$	$\lVert x \rVert \to \infty$
$\mathbb{B}$	$\lVert x \rVert \to \infty$

Defining regular variation

Regular variation can be defined/characterized in multiple ways:

- limits of functions
- vague/M_0 convergence of measures on punctured spaces
- weak convergence of finite measures on the unit sphere
- weak convergence of conditional probability distributions

To study regular variation of time series and clustering extremes, the latter view is quite convenient:

1. On $\mathbb{R}$, at ∞
2. On $\mathbb{R}$, at $\pm\infty$
3. On $\mathbb{R}^d$
4. On a Banach space $\mathbb{B}$

Regular variation at infinity is equivalent to weak convergence of relative excesses

A rv X is regularly varying (RV) at infinity with index $\alpha > 0$ if

$$\lim_{u \to \infty} \frac{P(X > uy)}{P(X > u)} = y^{-\alpha}, \qquad y > 0$$

For $y \geqslant 1$, this is can be written as

$$\lim_{u \to \infty} P(X/u > y \mid X > u) = y^{-\alpha} = P(Y > u)$$

RV(α) $\Leftrightarrow$ *weak convergence of relative excesses*:

$$\mathscr{L}(X/u \mid X > u) \xrightarrow{d} \mathscr{L}(Y) = \text{Pareto}(\alpha), \qquad u \to \infty$$

Regular variation on the real line is equivalent to weak convergence of certain conditional distributions (1)

A rv X is regularly varying with index $\alpha > 0$ if, as $u \to \infty$,

$$\frac{P(|X| > uy)}{P(|X| > u)} \to y^{-\alpha} \qquad (y > 0)$$

$$\frac{P(X > u)}{P(|X| > u)} \to p$$

Equivalent to weak convergence of conditional distributions:

$$\mathscr{L}\left(|X|/u \mid |X| > u\right) \xrightarrow{d} \mathscr{L}(Y) \sim \text{Pareto}(\alpha) \qquad \text{radius}$$

$$\mathscr{L}\Big(\underbrace{X/|X|}_{\text{sign}(X)} \mid |X| > u\Big) \xrightarrow{d} \mathscr{L}(\Theta) \qquad \text{angle}$$

as $u \to \infty$, where

$$\begin{aligned} P(\Theta = +1) &= p \\ P(\Theta = -1) &= 1 - p \end{aligned}$$

Regular variation on the real line is equivalent to weak convergence of certain conditional distributions (2)

Also jointly: X is RV with index $\alpha > 0$ if, as $u \to \infty$,

$$\mathscr{L}\left(\frac{|X|}{u}, \frac{X}{|X|} \;\middle|\; |X| > u \right) \xrightarrow{d} \mathscr{L}(Y, \Theta)$$

where

- $Y \sim \text{Pareto}(\alpha)$
- $$\begin{aligned} P(\Theta = +1) &= p \\ P(\Theta = -1) &= 1 - p \end{aligned}$$
- Y and Θ are independent

Regular variation also equivalent to

$$\mathscr{L}(X/u \mid |X| > u) \xrightarrow{d} \mathscr{L}(Y\Theta)$$

Regular variation in Euclidean space is equivalent to weak convergence of certain conditional distributions (1)

A random vector X in $\mathbb{R}^d$ is regularly varying with index $\alpha > 0$ if for all $y > 0$,

$$\frac{P(\|X\| > uy,\, X/\|X\| \in \cdot\,)}{P(\|X\| > u)} \xrightarrow{w} y^{-\alpha} H(\,\cdot\,), \qquad u \to \infty$$

for some probability measure H on $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d \mid \|x\| = 1\}$.

Equivalent to weak convergence of conditional distributions:

$$\mathscr{L}(\|X\|/u \mid \|X\| > u) \xrightarrow{d} \mathscr{L}(Y) = \text{Pareto}(\alpha) \qquad \text{radius}$$

$$\mathscr{L}(X/\|X\| \mid \|X\| > u) \xrightarrow{d} \mathscr{L}(\Theta) = H \qquad \text{angle}$$

as $u \to \infty$

Weak convergence of the radius and the angle separately implies their weak convergence jointly

For bounded, continuous $f : \mathbb{S}^{d-1} \to \mathbb{R}$ and for $y \geqslant 1$, as $u \to \infty$,

$$E\left[f\left(\frac{X}{\|X\|}\right); \frac{\|X\|}{u} > y \,\middle|\, \|X\| > u\right]$$
$$= \underbrace{E\left[f\left(\frac{X}{\|X\|}\right) \,\middle|\, \|X\| > uy\right]}_{\to E[f(\Theta)]} \underbrace{\frac{P(\|X\| > uy)}{P(\|X\| > u)}}_{\to y^{-\alpha}=P(Y>y)}$$
$$\to E[f(\Theta); Y > y]$$

for $Y \sim \text{Pareto}(\alpha)$, independent of Θ

Regular variation in Euclidean space is equivalent to weak convergence of certain conditional distributions (2)

A random vector X is RV with index $\alpha > 0$ and angular measure H if

$$\mathscr{L}\left(\frac{\|X\|}{u}, \frac{X}{\|X\|} \;\middle|\; \|X\| > u \right) \xrightarrow{d} \mathscr{L}(Y, \Theta)$$

where

- $Y \sim \text{Pareto}(\alpha)$
- $\Theta \sim H$
- Y and Θ are independent

Finally, regular variation is also equivalent to

$$\mathscr{L}(X/u \mid \|X\| > u) \xrightarrow{d} \mathscr{L}(Y\Theta), \qquad u \to \infty$$

Regular variation in a Banach space: weak convergence of conditional distributions

Multivariate regular variation in normed spaces: similarly.

[Hult & Lindskog 2005]

A random element X of a Banach space $\mathbb{B}$ is regularly varying if

$$\mathscr{L}(X/u \mid \|X\| > u) \xrightarrow{d} \mathscr{L}(Y), \qquad u \to \infty$$

and Y is such that $\|Y\| \geqslant 1$ is non-degenerate.

Necessarily

- $\|Y\| \sim \text{Pareto}(\alpha)$ for some $\alpha > 0$
- $\|Y\|$ and $\Theta = Y/\|Y\|$ are independent

and therefore

$$\mathscr{L}\left(\frac{\|X\|}{u}, \frac{X}{\|X\|} \,\middle|\, \|X\| > u \right) \xrightarrow{d} \mathscr{L}(\|Y\|, \Theta), \qquad u \to \infty$$

For the vague-convergence aficionados: yes you can, but…

Regular variation on Euclidean spaces often defined via vague convergence of measures:

- Convergence of integrals of continuous functions with compact support
- Multivariate regular variation on $\mathbb{R}^d$: for some $V \in RV_{-\alpha}$,

$$\frac{1}{V(u)} P\left[\frac{X}{u} \in \cdot\,\right] \xrightarrow{v} \mu(\,\cdot\,), \qquad u \to \infty.$$

Vague convergence on $[-\infty, +\infty]^d \setminus \{\mathbf{0}\}$

For infinite-dimensional $\mathbb{B}$, vague convergence collapses:

- $\mathbb{B}$ not locally compact
- $f : \mathbb{B} \to \mathbb{R}$ continuous and compactly supported implies $f \equiv 0$

Replace vague convergence by M_0-convergence

M_0-convergence:

"Weak convergence of finite measures on sets bounded away from the origin."

[Hult & Lindskog 2006]

X is *regularly varying* of index α if for some $V \in RV_{-\alpha}$,

$$\frac{1}{V(u)} P\left[\frac{X}{u} \in \cdot\,\right] \xrightarrow{M_0} \mu(\,\cdot\,), \qquad u \to \infty$$

the limit measure μ being non-null.

Extension to regular variation on star-shaped metric spaces.

Extremes of Stationary Sequences

Joint regular variation of a time series: What does it mean?

Let $\mathbb{B}$ be a separable Banach space

- E.g. $\mathbb{R}^d$, $C([0,1])$, L^p, ℓ^p
- Separability assumed out of convenience.
  Probably not needed everywhere.
  Excludes for instance $D([0,1])$ and spaces of usc functions

Let $(X_t)_{t\in\mathbb{Z}}$ be a strictly stationary time series in $\mathbb{B}$.

- Law of $(X_{s+h},\ldots,X_{t+h})$ does not depend on h.

Joint regular variation of the *whole* series $(X_t)_{t\in\mathbb{Z}}$?

The raw definition involves a cascade of angular measures

$(X_t)_{t \in \mathbb{Z}}$ is (jointly) regularly varying with index $\alpha > 0$
if for all $s \leqslant t \in \mathbb{Z}$, the vector $(X_s, \ldots, X_t)$ in $\mathbb{B}^{t-s+1}$
is regularly varying with the same index.

Wlog $s = 1 \leqslant t$. Let H_t be the spectral measure of $(X_1, \ldots, X_t)$:

$$\mathscr{L}\left(\frac{(X_1, \ldots, X_t)}{\|(X_1, \ldots, X_t)\|} \,\middle|\, \|(X_1, \ldots, X_t)\| > u \right) \xrightarrow{d} H_t, \qquad u \to \infty$$

- H_t is a probability measure on the unit sphere in $\mathbb{B}^t$.
- The measures $H_1, H_2, H_3, \ldots$ are linked somehow.
- Idem for M_0-convergence to limit measures μ_t.

Changing the conditioning event yields a unique limit object

Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary time series in $\mathbb{B}$ and let $\alpha > 0$.

Theorem

The following statements are equivalent:

(i) $(X_t)_{t\in\mathbb{Z}}$ *is regularly varying with index* α.

(ii) *The function* $u \mapsto P(\|X_0\| > u)$ *belongs to* $RV_{-\alpha}$ *and*

$$\mathscr{L}\left((X_t/\|X_0\|)_{t\in\mathbb{Z}} \,\middle|\, \|X_0\| > u\right) \xrightarrow{d} (\Theta_t)_{t\in\mathbb{Z}} \qquad (u \to \infty)$$

(iii) *For* $Y \sim \text{Pareto}(\alpha)$ *independent from some* $(\Theta_t)_{t\in\mathbb{Z}}$,

$$\mathscr{L}\left(\|X_0\|/u, (X_t/\|X_0\|)_{t\in\mathbb{Z}} \,\middle|\, \|X_0\| > u\right) \xrightarrow{d} \left(Y, (\Theta_t)_{t\in\mathbb{Z}}\right) \qquad (u \to \infty)$$

(iv) *For* $Y \sim \text{Pareto}(\alpha)$ *independent from some* $(\Theta_t)_{t\in\mathbb{Z}}$,

$$\mathscr{L}\left((X_t/u)_{t\in\mathbb{Z}} \mid \|X_0\| > u\right) \xrightarrow{d} (Y\Theta_t)_{t\in\mathbb{Z}} \qquad (u \to \infty)$$

Reconstructing the M_0-limit measures from the spectral process or tail process

- Spectral process: the unique limit process $(\Theta_t)_{t \in \mathbb{Z}}$ in (ii)–(iv).
- Tail process: the process $Y_t = Y\Theta_t$ in (iii)

The M_0-limit in $\mathbb{B}^t$ punctured at the origin

$$\frac{1}{P(\|X_0\| > u)} P[(X_1/u, \ldots, X_t/u) \in \cdot\,] \xrightarrow{M_0} \mu_t \qquad (u \to \infty)$$

is given by

$$\int_{\mathbb{B}^t} f \, d\mu_t = \sum_{j=1}^{t} \int_0^\infty E\Big[f(0, \ldots, 0, r\Theta_0, \ldots, r\Theta_{t-j}) \\ I\Big(\max_{-j+1 \leqslant i \leqslant -1} \|\Theta_i\| = 0 \Big) \Big] \, d(-r^{-\alpha})$$

The spectral process versus the spectral measure

- Special case $t = 0$:

$$\mathscr{L}(X_0/\|X_0\| \mid \|X_0\| > u) \xrightarrow{d} \mathscr{L}(\Theta_0), \qquad u \to \infty$$

  so $\mathscr{L}(\Theta_0)$ is the spectral measure H_0 of X_0.
  Clearly, $\|\Theta_0\| = 1$.

- For general $t \in \mathbb{Z}$,

$$\mathscr{L}(X_t/\|X_0\| \mid \|X_0\| > u) \xrightarrow{d} \mathscr{L}(\Theta_t), \qquad u \to \infty$$

  so $\|\Theta_t\| \neq 1$ in general if $t \neq 0$.
  By stationarity, the spectral measure of X_t is H_0 too.

The tail and spectral processes of a stationary process are in general non-stationary

Example (Independence)

If $(X_t)_{t\in\mathbb{Z}}$ is iid and X_0 is regularly varying,

$$\mathscr{L}\big((u^{-1}X_t)_{t\in\mathbb{Z}} \mid \|X_0\| > u\big) \xrightarrow{\text{fidi}} \mathscr{L}(\ldots, 0, 0, Y_0, 0, 0, \ldots)$$

Example (Full dependence)

If $X_t = X_0$ for all $t \in \mathbb{Z}$ and X_0 is regularly varying,

$$\mathscr{L}\big((u^{-1}X_t)_{t\in\mathbb{Z}} \mid \|X_0\| > u\big) \xrightarrow{\text{fidi}} \mathscr{L}(\ldots, Y_0, Y_0, Y_0, \ldots)$$

Extremes of Stationary Sequences

Stationarity of $(X_t)_{t\in\mathbb{Z}}$ induces a subtle structure on the tail/spectral process

Claim. $P(\Theta_{-t} \neq 0) = E[\|\Theta_t\|^\alpha]$

Proof – step 1:
Since $Y_{-t} = \|Y_0\|\Theta_{-t}$,

$$\begin{aligned} P(\Theta_{-t} \neq 0) &= P(Y_{-t} \neq 0) \\ &= \lim_{r\to 0} P(\|Y_{-t}\| > r) \\ &= \lim_{r\to 0} \lim_{u\to\infty} P(\|X_{-t}\|/u > r \mid \|X_0\| > u) \end{aligned}$$

Calculate the two limits.

Stationarity of $(X_t)_{t\in\mathbb{Z}}$ induces a subtle structure on the tail/spectral process

Claim. $P(\Theta_{-t} \neq 0) = E[\|\Theta_t\|^\alpha]$

Proof – step 2:
Limit as $u \to \infty$: By stationarity and regular variation

$$
\begin{aligned}
& P(\|X_{-t}\|/u > r \mid \|X_0\| > u) \\
& = P(\|X_0\|/u > r \mid \|X_t\| > u) \\
& = \frac{P(\|X_0\| > ur,\ \|X_t\| > u)}{P(\|X_t\| > u)} \\
& = \underbrace{\frac{P(\|X_0\| > ru)}{P(\|X_t\| > u)}}_{\to r^{-\alpha}} \underbrace{P(r\|X_t\| > ru \mid \|X_0\| > ru)}_{\to P(r\|Y_t\|>1)} \\
& \to r^{-\alpha} P(r\|Y_t\| > 1)
\end{aligned}
$$

as $u \to \infty$.

Stationarity of $(X_t)_{t\in\mathbb{Z}}$ induces a subtle structure on the tail/spectral process

Claim. $P(\Theta_{-t} \neq 0) = E[\|\Theta_t\|^\alpha]$

Proof – step 3:
Limit as $r \to 0$: Since $Y_t = \|Y_0\|\Theta_t$,

$$r^{-\alpha} P(r\|Y_t\| > 1) = r^{-\alpha} \int_1^\infty P(ry\|\Theta_t\| > 1)\, d(-y^{-\alpha})$$
$$= \int_0^{r^{-\alpha}} P(\|\Theta_t\|^\alpha > x)\, dx$$
$$\xrightarrow{r\to 0} \int_0^\infty P(\|\Theta_t\|^\alpha > x)\, dx = E[\|\Theta_t\|^\alpha]$$

QED

Forward and backward process: Restricting the spectral process to the future or the past

A stationary process $(X_t)_{t \in \mathbb{Z}}$ in $\mathbb{B}$ has a forward tail process $(Y_t)_{t \geqslant 0}$ if

$$\mathscr{L}\left((X_t/u)_{t \geqslant 0} \mid \|X_0\| > u \right) \xrightarrow{\text{fidi}} \mathscr{L}\left((Y_t)_{t \geqslant 0} \right)$$

Idem: backward tail process, forward/backward spectral process.

The property $P(\Theta_{-t} \neq 0) = E[\|\Theta_t\|^\alpha]$ suggests that we can infer the distribution of the backward process from the forward one.

Time-change formula:
How a time-shift affects the spectral process

Theorem

Statements (ii)–(iv) in the previous theorem are equivalent to the same statements with $\mathbb{Z}$ replaced by $\mathbb{Z}_+$ or $\mathbb{Z}_-$.

In that case,

$$E[f(\Theta_{-s},\dots,\Theta_t)] = E\left[f\left(\frac{\Theta_0}{\|\Theta_s\|},\dots,\frac{\Theta_{t+s}}{\|\Theta_s\|}\right)\|\Theta_s\|^\alpha I(\|\Theta_s\|>0)\right]$$

for all nonnegative integer s and t and for all integrable functions $f:\mathbb{B}^{t+s+1}\to\mathbb{R}$ such that $f(\theta_{-s},\dots,\theta_t)=0$ whenever $\theta_{-s}=0$.

Considering the time-reversed process $\tilde{X}_t = X_{-t}$
yields a similar reduction to the backward spectral process.

Understanding the time-change formula (1)

Assume $\mathbb{B} = \mathbb{R}$, $\alpha = 1$, and $X_t > 0$ a.s., so $\Theta_0 = 1$.

The time-change formula at $s = 1$ and $t = 0$ implies that for integrable $f : [0, \infty) \to \mathbb{R}$ such that $f(0) = 0$,

$$E[f(\Theta_{-1})] = E[f(1/\Theta_{+1})\,\Theta_{+1}]$$
$$E[f(\Theta_{+1})] = E[f(1/\Theta_{-1})\,\Theta_{-1}]$$

Let μ be the limit measure of (X_{t-1}, X_t) on $[0, \infty]^2 \setminus \{(0,0)\}$:

$$\frac{1}{P(X_0 > u)} P[u^{-1}(X_{t-1}, X_t) \in \cdot\,] \xrightarrow{v} \mu(\,\cdot\,) \qquad (u \to \infty)$$

To be applied to both (X_0, X_1) and to (X_{-1}, X_0): duality relation between Θ_1 and Θ_{-1}.

Understanding the time-change formula (2)

By definition of μ, Θ_1 and Θ_{-1} (Picture!):

$$
\begin{aligned}
P(\Theta_1 \leqslant z) &= \lim_{u\to\infty} P\left[\frac{X_1}{X_0} \leqslant z \,\middle|\, X_0 > u\right] \\
&= \lim_{u\to\infty} \frac{1}{P(X_0 > u)} P\left[\frac{X_1/u}{X_0/u} \leqslant z,\, X_0/u > 1\right] \\
&= \mu\{(x,y) : y/x \leqslant z,\, x > 1\} \\
P(\Theta_{-1} \leqslant z) &= \lim_{u\to\infty} P\left[\frac{X_0}{X_1} \leqslant z \,\middle|\, X_1 > u\right] \\
&= \dots \\
&= \mu\{(x,y) : x/y \leqslant z,\, y > 1\}
\end{aligned}
$$

Link between Θ_1 and Θ_{-1} follows if we can solve for μ.

Solving for the limit measure

If $f : [0,\infty)^2 \to \mathbb{R}$ (bounded, continuous) vanishes on $[0,\delta] \times [0,\infty)$,

$$\begin{aligned}
\int f \, d\mu &= \lim_{u\to\infty} \frac{1}{P(X_0 > u)} E[f(X_0/u, X_1/u)] \\
&= \lim_{u\to\infty} \frac{P(X_0 > \delta u)}{P(X_0 > u)} E[f(X_0/u, X_1/u) \mid X_0 > \delta u] \\
&= \delta^{-1} E[f(\delta Y_0, \delta Y_1)] \\
&= \delta^{-1} \int_1^\infty E[f(\delta y, \delta y \Theta_1)] \, d(-y^{-1}) \\
&= \int_\delta^\infty E[f(r, r\Theta_1)] \, d(-r^{-1}) \\
&= \int_0^\infty E[f(r, r\Theta_1)] \, d(-r^{-1})
\end{aligned}$$

- Formula extends to f such that $f(0,y) = 0$.
- For more general f, decompose

$$f(x,y) = \{f(x,y) - f(0,y)\} + f(0,y)$$

Symmetry

μ is symmetric if and only if $\Theta_{-1} \stackrel{d}{=} \Theta_1$.

Example

If μ corresponds to the Hüsler–Reiss max-stable distribution, we have $\Theta_{-1} \stackrel{d}{=} \Theta_1$ Lognormal with unit expectation.

Extremes of Stationary Sequences

Joint survival function when applying linear functionals

- Let $\{0, t\} \subset I \subset \{0, \ldots, t\}$.
- For $i \in I$, let $0 \neq b_i^* \in \mathbb{B}^*$, the dual of $\mathbb{B}$
  - $b_i^* : \mathbb{B} \to \mathbb{R}$ linear and bounded

By conditioning on the events $\|X_0\| > u/\|b_0^*\|$ or $\|X_t\| > u/\|b_t^*\|$,

$$\lim_{u \to \infty} \frac{P(\forall i \in I : b_i^* X_i > u)}{P(\|X_0\| > u)} = E[\min\{(b_i^* \Theta_i)_+^\alpha : i \in I\}]$$
$$= E[\min\{(b_i^* \Theta_{i-t})_+^\alpha : i \in I\}]$$

Equality of the expectations follows from the time-change formula.

Proof via conditioning and the spectral representation

Proof of $\dfrac{P(\forall i \in I : b_i^* X_i > u)}{P(\|X_0\| > u)} \to E[\min\{(b_i^* \Theta_i)_+^\alpha : i \in I\}$:

Step 1: calculate the limit as $u \to \infty$.
Since $b_0^* X_0 > u$ implies $\|X_0\| > u / \|b_0^*\|$,

$$\begin{aligned}
&\frac{P(\forall i \in I : b_i^* X_i > u)}{P(\|X_0\| > u)} \\
&= \frac{P(\|X_0\| > u/\|b_0^*\|)}{P(\|X_0\| > u)} P(\forall i \in I : b_i^* X_i > u \mid \|X_0\| > u/\|b_0^*\|) \\
&\to \|b_0^*\|^\alpha P(\forall i \in I : b_i^* Y_i > \|b_0^*\|)
\end{aligned}$$

Proof via conditioning and the spectral representation

Proof of $\dfrac{P(\forall i \in I : b_i^* X_i > u)}{P(\|X_0\| > u)} \to E[\min\{(b_i^*\Theta_i)_+^\alpha : i \in I\}$:

Step 2: Reduce the tail process to the spectral process.
Recall $Y_i = Y\Theta_i$ with $Y \sim \text{Pareto}(\alpha)$ independent of $(\Theta_i)_i$.

$$
\begin{aligned}
&\|b_0^*\|^\alpha \, P(\forall i \in I : b_i^* Y_i > \|b_0^*\|) \\
&= \|b_0^*\|^\alpha \int_1^\infty P(\forall i \in I : b_i^*(y\Theta_i) > \|b_0^*\|) \, d(-y^{-\alpha}) \\
&= \int_0^{\|b_0^*\|^\alpha} P\{\forall i \in I : (b_i^*\Theta_i)_+^\alpha > u\} \, du \\
&= \int_0^\infty P\{\forall i \in I : (b_i^*\Theta_i)_+^\alpha > u\} \, du \\
&= E[\min\{(b_i^*\Theta_i)_+^\alpha : i \in I\}]
\end{aligned}
$$

using $|b_0^*\Theta_0| \leqslant \|b_0^*\| \, \|\Theta_0\| = \|b_0^*\|$. QED

Joint survival of the sequence of norms

Similarly, for $b_i \in (0, \infty)$,

$$\lim_{u \to \infty} \frac{P(\forall i \in I : b_i \|X_i\| > u)}{P(\|X_0\| > u)} = E[\min\{b_i^\alpha \|\Theta_i\|^\alpha : i \in I\}]$$
$$= E[\min\{b_i^\alpha \|\Theta_{i-t}\|^\alpha : i \in I\}]$$

Equality of the expectations follows from the time-change formula.

Tail dependence coefficients

The coefficient of upper tail dependence between $b^* X_0$ and $b^* X_h$, for $b^* \in \mathbb{B}^*$ such that $P(b^* \Theta_0 > 0) > 0$:

$$\lim_{u \to \infty} P(b^* X_h > u \mid b^* X_0 > u) = \frac{E[\min\{(b^* \Theta_0)_+^\alpha,\ (b^* \Theta_h)_+^\alpha\}]}{E[(b^* \Theta_0)_+^\alpha]}$$
$$= \frac{E[\min\{(b^* \Theta_0)_+^\alpha,\ (b^* \Theta_{-h})_+^\alpha\}]}{E[(b^* \Theta_0)_+^\alpha]}$$

The coefficient of tail dependence between $\|X_0\|$ and $\|X_h\|$:

$$\lim_{u \to \infty} P(\|X_h\| > u \mid \|X_0\| > u) = E[\min(\|\Theta_h\|^\alpha, 1)]$$
$$= E[\min(\|\Theta_{-h}\|^\alpha, 1)]$$

Extremogram

Extremogram: Extreme-value analogue of the correllogram:

$$\rho_{A,B}(h) = \lim_{n\to\infty} n\,P(X_0/a_n \in A,\ X_h/a_n \in B),$$

- Regions A, B at least one of which stays away from the origin
- $a_n > 0$ satisfies $nP(\|X_0\| > a_n) \to 1$ as $n \to \infty$

[Davis & Mikosch 2009]

If A and B are continuity sets of the distributions of Y_0 and Y_h respectively and if $A \subset \{x \in \mathbb{B} : \|x\| > 1\}$, then

$$\begin{aligned}\rho_{A,B}(h) &= \lim_{n\to\infty} P(X_0/a_n \in A,\ X_h/a_n \in B \mid \|X_0\| > a_n) \\ &= P(Y_0 \in A,\ Y_h \in B).\end{aligned}$$

Extremogram of the image under linear functionals

If

$$A = \{x \in \mathbb{B} : a^* x > 1\},$$
$$B = \{x \in \mathbb{B} : b^* x > 1\}$$

for some $a^*, b^* \in \mathbb{B}^*$, then

$$\begin{aligned}\rho_{A,B}(h) &= \lim_{n\to\infty} n\,P(a^* X_0 > a_n,\, b^* X_h > a_n) \\ &= E[\min\{(a^*\Theta_0)_+^\alpha,\, (b^*\Theta_h)_+^\alpha\}]\end{aligned}$$

Extremal index of the sequence of norms

The (candidate) extremal index [Leadbetter 1983] of $(\|X_t\|)_{t\in\mathbb{Z}}$:

$$
\begin{aligned}
\theta &= \lim_{m\to\infty} \lim_{u\to\infty} P\left(\max_{t=1,\ldots,m} \|X_t\| \leqslant u \,\middle|\, \|X_0\| > u \right) \\
&= P\left(\sup_{t\geqslant 1} \|Y_t\| \leqslant 1 \right) \\
&= E\left[\sup_{t\geqslant 0} \|\Theta_t\|^\alpha - \sup_{t\geqslant 1} \|\Theta_t\|^\alpha \right]
\end{aligned}
$$

Passing from the tail process to the spectral process

Proof of $P(\sup_{t\geqslant 1}\|Y_t\| \leqslant 1) = E[\sup_{t\geqslant 0}\|\Theta_t\|^\alpha - \sup_{t\geqslant 1}\|\Theta_t\|^\alpha]$:
Writing $Y = \|Y_0\|$, since $Y^{-\alpha} \sim \text{Uniform}(0,1)$ and since $\|\Theta_0\| = 1$,

$$\begin{aligned}
P\left(\sup_{t\geqslant 1}\|Y_t\| \leqslant 1\right) &= P\left(Y\sup_{t\geqslant 1}\|\Theta_t\| \leqslant 1\right) \\
&= P\left(\sup_{t\geqslant 1}\|\Theta_t\|^\alpha \leqslant Y^{-\alpha}\right) \\
&= \int_0^1 P\left(\sup_{t\geqslant 1}\|\Theta_t\|^\alpha \leqslant u\right) du \\
&= 1 - E\left[\min\left(1, \sup_{t\geqslant 1}\|\Theta_t\|^\alpha\right)\right] \\
&= E\left[\sup_{t\geqslant 0}\|\Theta_t\|^\alpha - \sup_{t\geqslant 1}\|\Theta_t\|^\alpha\right]
\end{aligned}$$

using the identity $\int_0^1 P(\xi \leqslant u)\,du = 1 - \int_0^\infty P\{\min(1,\xi) > u\}\,du$

Extremal index of the image under a linear functional

Let $b^* \in \mathbb{B}^*$ be such that $P(b^*\Theta_0 > 0) > 0$.
The (candidate) extremal index of $(b^*X_t)_{t\in\mathbb{Z}}$:

$$\begin{aligned}
\theta(b^*) &= \lim_{m\to\infty} \lim_{u\to\infty} P\left(\max_{t=1,\dots,m} b^*X_t \leqslant u \,\middle|\, b^*X_0 > u \right) \\
&= 1 - \frac{E[\min\{(b^*\Theta_0)_+^\alpha, \sup_{t\geqslant 1}(b^*\Theta_t)_+^\alpha\}]}{E[(b^*\Theta_0)_+^\alpha]} \\
&= \frac{E[\sup_{t\geqslant 0}(b^*\Theta_t)_+^\alpha - \sup_{t\geqslant 1}(b^*\Theta_t)_+^\alpha]}{E[(b^*\Theta_0)_+^\alpha]}
\end{aligned}$$

Large deviations and the cluster index

$\mathbb{B} = \mathbb{R}$. Partial sums $S_k = X_1 + \cdots + X_k$.
For $a_n > 0$ such that $n\,P(|X_0| > a_n) \to 1$, put

$$b_+(k) = \lim_{n\to\infty} n\,P(S_k > a_n)$$

For certain Markov chains, the cluster index b_+ exists:

$$b_+ = \lim_{k\to\infty} \{b_+(k+1) - b_+(k)\} = E\left[\left(\sum_{t\geqslant 0}\Theta_t\right)_+^\alpha - \left(\sum_{t\geqslant 1}\Theta_t\right)_+^\alpha\right]$$

Large deviations principle: for appropriate $u_n, v_n \to \infty$,

$$\lim_{n\to\infty} \sup_{x\in(u_n,v_n)} \left| \frac{P(S_n > x)}{n\,P(|X_0| > x)} - b_+ \right| = 0$$

[Mikosch & Wintenberger 2012a,b; Wintenberger 2012]

Central limit theorems with stable, non-Gaussian limits

$\mathbb{B} = \mathbb{R}$ and $0 < \alpha < 2$. Partial sums $S_n = X_1 + \cdots + X_n$

- Stable limits of the partial sums

  [Bartkiewicz, Jakubowski, Mikosch, and Wintenberger 2011]

- Functional limit theorem in $D[0, 1]$ with Skorohod's M_1 topology (weaker than J_1)

  [Basrak, Krizmanić & S. 2012]

Limiting characteristic functions (Lévy measures) expressed in terms of spectral process.

Extremes of Stationary Sequences

Describing heavy tails: Regular variation

Tail and spectral processes

Time-change formula

Using the spectral process

Conclusion

To take home...

1. Regular variation and existence of the spectral process:

$$\mathscr{L}\left(\|X_0\|/u,\ (X_t/\|X_0\|)_{t\in\mathbb{Z}} \mid \|X_0\| > u\right) \xrightarrow{\text{fidi}} \mathscr{L}\left(Y,\ (\Theta_t)_{t\in\mathbb{Z}}\right)$$

   with $Y \sim \text{Pareto}(\alpha)$ independent of $(\Theta_t)_{t\in\mathbb{Z}}$
2. Time-change formula:
   backward ($t \leqslant 0$) versus forward ($t \geqslant 0$) spectral process
3. Using the spectral process for describing extremal dependence

Part III

Markov Processes
— with A. Janßen

Extremes of Stationary Sequences

Set-up and main finding

Forward spectral processes

Time-change formula

Adjoint distributions

Back-and-forth spectral processes and the spectral process

Set-up: multivariate Markov chain with regularly varying initial distribution

Discrete-time, $\mathbb{R}^d$-valued random process $(X_t)_{t \geqslant 0}$ defined by

$$X_t = \Psi(X_{t-1}, \varepsilon_t), \qquad t = 1, 2, \ldots,$$

where

- $\varepsilon_1, \varepsilon_2, \ldots$ are iid in a measurable space $(\mathbb{E}, \mathscr{E})$, independent of X_0
- $\Psi : \mathbb{R}^d \times \mathbb{E} \to \mathbb{R}^d$ is measurable
- the law of X_0 is multivariate regularly varying

If $(X_t)_t$ is stationary, it will be assumed to be defined for all $t \in \mathbb{Z}$.

Commenting the framework: Representation of the Markov chain

Rather than transition kernels, use the representation

$$X_t = \Psi(X_{t-1}, \varepsilon_t)$$

- Non-unique
- General, e.g. inverse (conditional) Rosenblatt (1952) transform
  - ε_t iid uniform $[0, 1]^d$
  - $\Psi(x, \cdot)$ vector of (conditional)^2 quantile functions
- Arises naturally in examples, e.g. stochastic recurrence equation

$$X_t = A_t X_{t-1} + B_t, \qquad \varepsilon_t = (A_t, B_t)$$

Aim: to find the spectral process of a multivariate regularly varying Markov chain

We are looking for the weak limit $(M_t)_t$, called spectral process, in

$$\mathscr{L}\big(\|X_0\|/u,\ (X_t/\|X_0\|)_t \mid \|X_0\| > u\big) \xrightarrow{d} \mathscr{L}\big(Y,\ (M_t)_t\big), \qquad u \to \infty$$

- $\alpha > 0$ is the index of regular variation of X_0
- Y is Pareto(α), i.e. $P[Y > y] = y^{-\alpha}$ for $y \geqslant 1$
- Y is independent of $(M_t)_t$

Continuous mapping theorem:

$$\mathscr{L}\big((X_t/u)_t \mid \|X_0\| > u\big) \xrightarrow{d} \mathscr{L}\big((YM_t)_t\big), \qquad u \to \infty$$

The spectral process and the extremogram: two sides of the same coin

Linking the spectral process and the extremogram [Davis & Mikosch 2009]:

- For nice sets $A, B \subset \mathbb{R}^d$ such that $A \subset \{x : \|x\| \geqslant 1\}$,

$$\rho_{AB}(h) = \lim_{u \to \infty} P[u^{-1} X_h \in B \mid u^{-1} X_0 \in A]$$
$$= P[Y M_h \in B \mid Y M_0 \in A], \qquad h = 0, 1, 2, \ldots$$

- Conversely, from the extremogram of the lagged-h process

$$Y_{t,h} = \operatorname{vec}(X_{t-h+1}, \ldots, X_t),$$

one deduces the $2hd$-dimensional distributions of the spectral process.

Main findings

Markov spectral processes $(M_t)_t$ verify the following properties:

- The forward ($t \geqslant 0$) and backward ($t \leqslant 0$) chains are adjoint
- The forward and backward spectral processes are Markov chains
- They enjoy a certain scaling property

Univariate case: (to be thought of as) multiplicative random walks

[Smith 1992; Perfekt 1994; Yun 2000; Bortot & Coles 2000/2003; S. 2007; Resnick & Zeber 2011]

General: back-and-forth tail chain

Extremes of Stationary Sequences

Set-up and main finding

Forward spectral processes

Time-change formula

Adjoint distributions

Back-and-forth spectral processes and the spectral process

Condition: regularly varying initial distribution

The distribution of X_0 is regularly varying with

- index $\alpha > 0$
- *spectral/angular measure* H on the unit sphere $\mathbb{S}^{d-1}$

$$\mathscr{L}(\|X_0\|/u,\ X_0/\|X_0\| \mid \|X_0\| > u) \xrightarrow{d} \mathscr{L}(Y, M_0), \qquad u \to \infty$$

where

- $M_0 \sim H$
- Y is Pareto(α), i.e. $P[Y > y] = y^{-\alpha}$ for $y \geqslant 1$
- Y and M_0 are independent

Condition: asymptotic scaling of the update function

Recall

$$X_t = \Psi(X_{t-1}, \varepsilon_t)$$

1. With probability one and for all H-almost every $s \in \mathbb{S}^{d-1}$,

$$\lim_{u \to \infty} \frac{\Psi\big(u\, s(u), \varepsilon_t\big)}{u} = \phi(s, \varepsilon_t)$$

   whenever $s(u) \to s$ as $u \to \infty$.

2. If $P[\phi(s, \varepsilon_t) = 0] > 0$ for some s in the support of H, then with probability one,

$$\sup_{\|x\| \leqslant u} |\Psi(x, \varepsilon_t)| = O(u), \qquad u \to \infty$$

Conditions easily verified in examples such as $X_t = A_t X_{t-1} + B_t$.

Unfolding the recursion

Aim: to find the weak limit M_t, of $X_t/\|X_0\|$, given $\|X_0\| > u \to \infty$.
If $\|X_0\|$ is 'large':

$$M_0 \stackrel{d}{\approx} \frac{X_0}{\|X_0\|} \sim H$$

$$M_1 \stackrel{d}{\approx} \frac{X_1}{\|X_0\|} = \frac{\Psi(X_0,\varepsilon_1)}{\|X_0\|} \approx \phi\left(\frac{X_0}{\|X_0\|},\varepsilon_1\right) \stackrel{d}{\approx} \phi(M_0,\varepsilon_1),$$

$$\begin{aligned} M_2 &\stackrel{d}{\approx} \frac{X_2}{\|X_0\|} = \frac{\|X_1\|}{\|X_0\|}\,\frac{\Psi(X_1,\varepsilon_2)}{\|X_1\|} \\ &\approx \frac{\|X_1\|}{\|X_0\|}\,\phi\left(\frac{X_1}{\|X_1\|},\varepsilon_2\right) \\ &= \frac{\|X_1\|}{\|X_0\|}\,\phi\left(\frac{X_1/\|X_0\|}{\|(X_1/\|X_0\|)\|},\varepsilon_2\right) \stackrel{d}{\approx} \|M_1\|\,\phi\left(\frac{M_1}{\|M_1\|},\varepsilon_2\right) \end{aligned}$$

Existence and description of the forward spectral process

Theorem

For a time-homogeneous Markov chain $(X_t)_{t\geqslant 0}$, under the previous conditions,

$$\mathscr{L}\left(\frac{\|X_0\|}{u}; \frac{X_0}{\|X_0\|}, \frac{X_1}{\|X_0\|}, \dots \,\middle|\, \|X_0\| > u\right) \xrightarrow{d} \mathscr{L}(Y; M_0, M_1, \dots)$$

with, for $t \geqslant 1$,

$$M_t = \|M_{t-1}\| \, \phi\left(\frac{M_{t-1}}{\|M_{t-1}\|}, \varepsilon_t\right) I_{\{\|M_{t-1}\|>0\}}$$

and

- $Y, M_0, \varepsilon_1, \varepsilon_2, \dots$ *are independent*
- $Y \sim \text{Pareto}(\alpha)$
- $M_0 \sim H$
- $\varepsilon_1, \varepsilon_2, \dots$ *iid (copies) as in the definition of* (X_t)

Example: vector AR(1) – angular measure

$$X_t = AX_{t-1} + \varepsilon_t, \qquad t \geqslant 0$$

- deterministic $A \in \mathbb{R}^{d \times d}$ such that $\|A^m\| < 1$ for some $m \geqslant 1$
- ε_t iid regularly varying $\alpha > 0$, angular measure λ
- $X_0, \varepsilon_1, \varepsilon_2, \ldots$ independent

Then X_0 is regularly varying with index α too and spectral measure

$$H = \sum_{k \geqslant 0} p_k \lambda_k$$

- λ_k the angular measure of $A^k \varepsilon_t$
- $(p_k)_{k \geqslant 0}$ a discrete probability distribution given by A, λ and α

See Part IV *Linear processes*.

Example: vector AR(1) – forward tail process

The update function has the asymptotic scaling property:

$$\begin{aligned}\phi(s, \varepsilon_t) &= \lim_{u\to\infty} \frac{\Psi(u\,s(u), \varepsilon_t)}{u} \\ &= \lim_{u\to\infty} \frac{A\,u\,s(u) + \varepsilon_t}{u} \\ &= As, \qquad s \in \mathbb{S}^{d-1}\end{aligned}$$

The forward spectral process $(M_t)_{t\geqslant 0}$ is then simply

$$\begin{aligned}M_t &= \|M_{t-1}\|\,\phi\left(\frac{M_{t-1}}{\|M_{t-1}\|}, \varepsilon_t\right) \\ &= AM_{t-1} \\ &= \cdots \\ &= A^t M_0\end{aligned}$$

Extremes of Stationary Sequences

Set-up and main finding

Forward spectral processes

Time-change formula

Adjoint distributions

Back-and-forth spectral processes and the spectral process

Stationarity: existence of the full spectral process

Suppose in addition that $(X_t)_t$ is *strictly stationary*.
Without loss of generality, assume that X_t is defined for all $t \in \mathbb{Z}$.

Corollary

Under the previous conditions, there exists a process $(M_t)_{t\in\mathbb{Z}}$ *s.t.*

$$\mathscr{L}\left(\frac{\|X_0\|}{u}; \dots, \frac{X_{-1}}{\|X_0\|}, \frac{X_0}{\|X_0\|}, \frac{X_1}{\|X_0\|}, \dots \,\middle|\, \|X_0\| > u \right) \xrightarrow{d} \mathscr{L}(Y; \dots, M_{-1}, M_0, M_1, \dots), \qquad u \to \infty$$

[S. 2007; Basrak & S. 2009; Meinguet & S. 2010]

Existence of the spectral process and regular variation

- The fidis of the Markov chain $(X_t)_{t\in\mathbb{Z}}$ are regularly varying
- Existence of the forward spectral process M_t, $t \geqslant 0$, implies existence of the full spectral process M_t, $t \in \mathbb{Z}$
- Reconstruct the full spectral process from the forward spectral process via time-change formulas

Time-change formula:
Reconstructing the full tail process from the forward part

Corollary

For all integer h, s, t *with* $s, t \geqslant 0$ *and for every measurable function* $f : (\mathbb{R}^d)^{s+1+t} \to \mathbb{R}$ *satisfying* $f(x_{-s}, \ldots, x_t) = 0$ *whenever* $x_0 = 0$,

$$E[f(M_{-s-h}, \ldots, M_{t-h})] = E\left[f\left(\frac{M_{-s}}{\|M_h\|}, \ldots, \frac{M_t}{\|M_h\|}\right) \|M_h\|^\alpha \, I_{\{\|M_h\|>0\}}\right]$$

[Basrak & S. 2009, Theorem 3.1(iii)]

- Change in distribution due to a time-shift of lag $h \in \mathbb{Z}$
- Choosing $s = 0 \leqslant h$ yields at the right-hand side an expression that depends on the forward tail process only

Extremes of Stationary Sequences

Set-up and main finding

Forward spectral processes

Time-change formula

Adjoint distributions

Back-and-forth spectral processes and the spectral process

Recapitulation

$(X_t)_t$	regularly varying Markov chain in $\mathbb{R}^d$
Ψ	update function: $X_t = \Psi(X_{t-1}, \varepsilon_t)$
ϕ	scaling limit: $\Psi(x, \varepsilon_t) \approx \|x\| \, \phi(\frac{x}{\|x\|}, \varepsilon_t)$ if $\|x\|$ is large
Y	Pareto(α) random variable weak limit of $\|X_0\|/u$ given $\|X_0\| > u$ as $u \to \infty$ $\alpha > 0$ is the index of regular variation of $\|X_0\|$
M_t	spectral process weak limit of $X_t/\|X_0\|$ given $\|X_0\| > u$ as $u \to \infty$
H	spectral/angular measure of X_0 law of M_0, taking values in $\mathbb{S}^{d-1} = \{x : \|x\| = 1\}$

How to reconstruct the backward spectral process?

For Markov spectral processes $(M_t)_t$:

- The forward spectral process admitted an explicit representation:

$$M_t = \|M_{t-1}\| \, \phi\left(\frac{M_{t-1}}{\|M_{t-1}\|}, \varepsilon_t\right) I_{\{\|M_{t-1}\|>0\}}, \qquad t \geqslant 1$$

- By the time-change formula,
  the law of the backward spectral process ($t \leqslant 0$)
  is determined by the forward spectral process ($t \geqslant 0$)

How does the backward spectral process look like?

A first step: let us study the law of (M_{-1}, M_0). Recall:

$$\mathscr{L}(X_{-1}/\|X_0\| \mid \|X_0\| > u) \xrightarrow{d} \mathscr{L}(M_{-1}), \qquad u \to \infty$$

A special case of the time-change formula motivates an adjoint relation between probability measures

The distributions of (M_0, M_1) and (M_0, M_{-1}) are "adjoint".

- In the time-change formula, set $s = 0$ and $h = t = 1$:

$$E[f(M_{-1}, M_0)] = E\left[f\left(\frac{M_0}{\|M_1\|}, \frac{M_1}{\|M_1\|}\right) \|M_1\|^{\alpha} I_{\{\|M_1\|>0\}}\right]$$

  for all $f : (\mathbb{R}^d)^2 \to \mathbb{R}$ satisfying $f(y_0, y_1) = 0$ whenever $y_0 = 0$

- Similarly, set $s = 1$, $h = -1$ and $t = 0$:

$$E[f(M_0, M_1)] = E\left[f\left(\frac{M_{-1}}{\|M_{-1}\|}, \frac{M_0}{\|M_{-1}\|}\right) \|M_{-1}\|^{\alpha} I_{\{\|M_{-1}\|>0\}}\right]$$

  for all $f : (\mathbb{R}^d)^2 \to \mathbb{R}$ such that $f(y_{-1}, y_0) = 0$ whenever $y_0 = 0$

Admissible distributions for the definition of the adjoint

The adjoint relation will be defined on a certain set $\mathscr{M}_\alpha$ of probability measures P on $\mathbb{S}^{d-1} \times \mathbb{R}^d$.

- Think of P as the law of (M_0, M_1) or (M_0, M_{-1}).

By definition, P belongs to $\mathscr{M}_\alpha$ if for every Borel set $S \subset \mathbb{S}^{d-1}$

$$\int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} I\left(\frac{m}{\|m\|} \in S \right) \|m\|^\alpha \, P(\mathrm{d}s, \mathrm{d}m) \leqslant P(S \times \mathbb{R}^d)$$

We call $\mathscr{M}_\alpha$ the set of admissible distributions.

In particular, setting $S = \mathbb{S}^{d-1}$ yields

$$\int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} \|m\|^\alpha \, P(\mathrm{d}s, \mathrm{d}m) \leqslant 1$$

Tail chain distributions are admissible

Let $(M_t)_{t\in\mathbb{Z}}$ be the spectral process of a regularly varying stationary Markov chain $(X_t)_{t\in\mathbb{Z}}$ as before.

Lemma

The law of (M_0, M_1) *belongs to* $\mathscr{M}_\alpha$*, i.e.*

$$E\left[I\left(\frac{M_1}{\|M_1\|} \in S\right) \|M_1\|^\alpha\right] \leqslant P(M_0 \in S)$$

for every Borel set $S \subset \mathbb{S}^{d-1}$.

In particular, setting $S = \mathbb{S}^{d-1}$ gives

$$E[\|M_1\|^\alpha] \leqslant 1$$

An adjoint relation between probability measures

For $P \in \mathscr{M}_\alpha$, define a signed Borel measure P^* on $\mathbb{S}^{d-1} \times \mathbb{R}^d$ by:

- Restriction to $\mathbb{S}^{d-1} \times \{0\}$: for $S \subset \mathbb{S}^{d-1}$,

$$P^*(S \times \{0\})$$
$$= P(S \times \mathbb{R}^d) - \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} I\left(\frac{m}{\|m\|} \in S\right) \|m\|^\alpha \, P(\mathrm{d}s, \mathrm{d}m)$$

- Restriction to $\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$: for $E \subset \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$,

$$P^*(E) = \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} I\left(\left(\frac{m}{\|m\|}, \frac{s}{\|m\|}\right) \in E\right) \|m\|^\alpha \, P(\mathrm{d}s, \mathrm{d}m)$$

We call P^* the adjoint measure of P in $\mathscr{M}_\alpha$.

The adjoint is a true 'adjoint'

Lemma

Let $P \in \mathscr{M}_\alpha$ and let P^* be its adjoint measure.

(i) P^* *is a probability measure.*

(ii) *The marginal distributions of P and P^* on $\mathbb{S}^{d-1}$ are the same.*

(iii) $P^* \in \mathscr{M}_\alpha$.

(iv) $(P^*)^* = P$.

(v) *For every measurable function $f : \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\}) \to \mathbb{R}$,*

$$\int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(s^*, m^*) \, P^*(\mathrm{d}s^*, \mathrm{d}m^*) = \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f\left(\frac{m}{\|m\|}, \frac{s}{\|m\|} \right) \|m\|^\alpha \, P(\mathrm{d}s, \mathrm{d}m)$$

The forward and backward increments of the spectral process satisfy the adjoint relation

Let $(M_t)_{t\in\mathbb{Z}}$ be the spectral process of a regularly varying stationary Markov chain $(X_t)_{t\in\mathbb{Z}}$ as before.

Corollary

The distributions of (M_0, M_1) *and* (M_0, M_{-1}) *are adjoint.*

Proof: Time-change formula.

Special case:

$$P[M_{-1} \neq 0] = E[\|M_1\|^\alpha],$$
$$P[M_1 \neq 0] = E[\|M_{-1}\|^\alpha]$$

Special case: univariate and positive

- $d = 1, \mathbb{S}^{d-1} = \{-1, 1\}$
- If $P \in \mathscr{M}_\alpha$ has $P(\{-1\} \times \mathbb{R}) = 0$,
  then P must be concentrated on $\{+1\} \times [0, \infty)$
- Then so is P^* and for $B \subset (0, \infty)$

$$P^*(\{+1\} \times B) = \int_{s=+1,\, m>0} I\left(\frac{1}{m} \in B\right) m^\alpha \, P(\mathrm{d}s, \mathrm{d}m)$$

- Examples if $\alpha = 1$:
  - If P is lognormal with unit expectation, then $P = P^*$
  - If P is Bernoulli, then $P = P^*$
  - If P is unit exponential,
    then P^* is the law of $1/(E_1 + E_2)$, with E_1, E_2 iid unit exponential

Extremes of Stationary Sequences

Set-up and main finding

Forward spectral processes

Time-change formula

Adjoint distributions

Back-and-forth spectral processes and the spectral process

Taking stock

- Initial state: $M_0 \sim H$ angular measure of X_0
- Forward spectral process: $M_0, \varepsilon_1, \varepsilon_2, \ldots$ are independent and

$$M_j = \|M_{t-1}\| \, \phi\left(\frac{M_{t-1}}{\|M_{t-1}\|}\right) I_{\{\|M_{t-1}\|>0\}}, \qquad t = 1, 2, \ldots$$

- Laws of (M_0, M_1) and (M_0, M_{-1}) are adjoint
- Time-change formula

How does the backward spectral process M_t, $t \leqslant 0$, look like?

Back-and-forth spectral process

A process $(M_t)_{t\in\mathbb{Z}}$ in $\mathbb{R}^d$ is called a *back-and-forth tail chain* with index $\alpha \in (0,\infty)$, notation $\textsc{bftc}(\alpha)$, if:

(i) $\mathscr{L}(M_0, M_1)$ and $\mathscr{L}(M_0, M_{-1})$ belong to $\mathscr{M}_\alpha$ and are adjoint;

(ii) the forward chain $(M_t)_{t\geqslant 0}$ is a Markov chain with respect to the filtration $\sigma(M_s, s \leqslant t)$, $t \geqslant 0$, and the Markov kernel satisfies

$$P[M_t \in \cdot \mid M_{t-1} = x_{t-1}] = \begin{cases} \delta_0(\,\cdot\,) & \text{if } x_{t-1} = 0, \\ P[\|x_{t-1}\| M_1 \in \cdot \mid M_0 = x_{t-1}/\|x_{t-1}\|] & \text{if } x_{t-1} \neq 0; \end{cases}$$

(iii) the backward chain $(M_{-t})_{t\geqslant 0}$ is a Markov chain with respect to the filtration $\sigma(M_{-s}, s \leqslant t)$, $t \geqslant 0$, and satisfies the same relation as in (ii) with $t-1$ and t replaced by $-t+1$ and $-t$ respectively

Time-change formula for a BFTC

Let $(M_t)_{t\in\mathbb{Z}}$ be a $\text{BFTC}(\alpha)$.

Theorem

For all integer $s, t \geqslant 0$ *and for all measurable functions* $f : \mathbb{R}^{(s+1+t)d} \to \mathbb{R}$ *vanishing on* $\{0\} \times \mathbb{R}^{(s+t)d}$*, the* $s+1$ *numbers*

$$E\left[f\left(\frac{M_{-s+h}}{\|M_h\|}, \dots, \frac{M_{t+h}}{\|M_h\|}\right) \|M_h\|^\alpha \, I_{\{M_h \neq 0\}}\right], \qquad h = 0, \dots, s,$$

are all the same, in the sense that if one integral exists, then they all exist and they are equal.

The case $s = 1$ and $t = 0$ is just the adjoint relation between the distributions of (M_0, M_1) and (M_0, M_{-1}).

Identifying a back-and-forth tail chain from its forward part

Theorem
Let $(Y_t)_{t\in\mathbb{Z}}$ be a process in $\mathbb{R}^d$ and let $(M_t)_{t\in\mathbb{Z}}$ be a $\mathrm{BFTC}(\alpha)$ *in $\mathbb{R}^d$.*

If

1. $\mathscr{L}(Y_0, \ldots, Y_t) = \mathscr{L}(M_0, \ldots, M_t)$ *for all* $t \geqslant 0$
2. *for all $h, s, t \in \mathbb{Z}$ with $s, t \geqslant 0$ and for all bounded, measurable $f : (\mathbb{R}^d)^{s+1+t} \to \mathbb{R}$ satisfying $f(y_{-s}, \ldots, y_t) = 0$ whenever $y_0 = 0$,*

$$E[f(Y_{-s-h}, \ldots, Y_{t-h})] = E\left[f\left(\frac{Y_{-s}}{\|Y_h\|}, \ldots, \frac{Y_t}{\|Y_h\|}\right)\|Y_h\|^\alpha I_{\{\|Y_h\|>0\}}\right]$$

then

$$\mathscr{L}(Y_{-s}, \ldots, Y_t) = \mathscr{L}(M_{-s}, \ldots, M_t), \qquad s, t \geqslant 0.$$

Markov spectral processes are back-and-forth tail chains

Every $(M_t)_{t\in\mathbb{Z}}$ whose forward part $(t \geqslant 0)$ has a $\text{BFTC}(\alpha)$ structure, must be a full $(t \in \mathbb{Z})$ $\text{BFTC}(\alpha)$. In particular:

Corollary

The spectral process $(M_t)_{t\in\mathbb{Z}}$ of a regularly varying, stationary Markov chain $(X_t)_{t\in\mathbb{Z}}$ satisfying the earlier conditions is a $\text{BFTC}(\alpha)$.

Univariate back-and-forth tail chains are sign-sensitive multiplicative random walks

- P a law on $\{-1,+1\} \times \mathbb{R}$ in $\mathscr{M}_\alpha$; adjoint P^*
- $(M_t)_{t \in \mathbb{Z}}$ a BFTC(α) with $(M_0, M_1) \sim P$ and $(M_0, M_{-1}) \sim P^*$
- Then for $t \geqslant 1$,

$$M_t = \begin{cases} |M_{t-1}| A_t & \text{if } M_{t-1} > 0, \\ 0 & \text{if } M_{t-1} = 0, \\ |M_{t-1}| B_t & \text{if } M_{t-1} < 0; \end{cases}$$

$$M_{-t} = \begin{cases} |M_{-t+1}| A_{-t} & \text{if } M_{-t+1} > 0, \\ 0 & \text{if } M_{-t+1} = 0, \\ |M_{-t+1}| B_{-t} & \text{if } M_{-t+1} < 0; \end{cases}$$

where the increments $A_{\pm t}$ and $B_{\pm t}$ are independent, with laws determined by P and P^*, and independent of $M_0 \in \{-1, 1\}$

'Tail switching potential' [Bortot & Coles 2003; S. 2007]

Example: vector AR(1) – back-and-forth tail process

Recall: deterministic $A \in \mathbb{R}^{d \times d}$, iid regularly varying $(\varepsilon_t)_{t \in \mathbb{Z}}$,

$$X_t = AX_{t-1} + \varepsilon_t = \sum_{k \geqslant 0} A^k \varepsilon_{t-k} \qquad t \in \mathbb{Z}$$

$$M_t = A^t M_0 \qquad t \geqslant 0$$

Full $\textsc{bftc}(\alpha)$:

$$M_{-N+h} = \begin{cases} A^h M_{-N} & \text{if } h \geqslant 0 \\ 0 & \text{otherwise} \end{cases}$$

where

- N is a certain random nonnegative integer
- conditionally on N, the distribution of M_{-N} is determined by A, the angular measure of ε_t, and $\alpha > 0$.

Conclusion: structure of Markov spectral processes

- Tail chains give information on the extremes of multivariate regularly varying Markov chains
- Markov spectral processes are back-and-forth tail chains:
  - The forward and backward spectral processes are Markov chains too
  - The forward ($t \geqslant 0$) and backward ($t \leqslant 0$) chains are adjoint
  - They enjoy a certain scaling property

Part IV

Linear Processes
— with T. Meinguet

Extremes of Stationary Sequences

Time series of random functions

Physical quantity observed in space and over time

$$X_t(x) = \text{value at time } t \text{ at location } x$$

Space coordinate x varies over a grid – *high-dimensional!*

Think of x as varying continously over space

$\rightsquigarrow$ For fixed t, view $X_t(\,\cdot\,)$ as a *random function*

$\rightsquigarrow$ Time series $\big(X_t(\,\cdot\,)\big)_{t\in\mathbb{Z}}$ of random functions

Goal: to model extremal dependence in

Space – cross-sectional tail dependence

Time – clusters

Example: Autoregressive process

Define $X_t(\cdot)$ recursively by

$$X_t(x) = \int K(x,y)\, X_{t-1}(y)\, dy + Z_t(x)$$

Model ingredients:

- Kernel $K(x,y)$: from location y now to location x tomorrow
- Z_t iid random functions: innovations – heavy tails!

More general: *linear time series*

Regular variation in a Banach space is weak convergence of conditional distributions

A random element X of a Banach space $\mathbb{B}$ is regularly varying if

$$\mathscr{L}(X/u \mid \|X\| > u) \xrightarrow{d} \mathscr{L}(Y), \qquad u \to \infty$$

for Y such that $\|Y\| \geqslant 1$ is non-degenerate.

Necessarily

- $\|Y\| \sim \text{Pareto}(\alpha)$ for some $\alpha > 0$
- $\|Y\|$ and $\Theta = Y/\|Y\|$ are independent

and therefore

$$\mathscr{L}\left(\frac{\|X\|}{u}, \frac{X}{\|X\|} \;\middle|\; \|X\| > u \right) \xrightarrow{d} \mathscr{L}(Y, \Theta), \qquad u \to \infty$$

Linear processes taking values in a Banach space

Two Banach spaces $\mathbb{B}_1, \mathbb{B}_2$.
A linear process $(X_t)_{t\in\mathbb{Z}}$ is of the form

$$X_t = \sum_{i\in\mathbb{Z}} T_i(Z_{t-i})$$

where

- Z_t are iid in $\mathbb{B}_1$
- Bounded linear operators $T_i : \mathbb{B}_1 \to \mathbb{B}_2$

E.g.: AR(1) process ($\mathbb{B}_1 = \mathbb{B}_2$)

$$X_t = T(X_{t-1}) + Z_t = \sum_{i\geqslant 0} T^i Z_{t-i}$$

Extremes of Stationary Sequences

Linear operators preserve regular variation

Let X be a regularly varying random element in $\mathbb{B}_1$ with index $\alpha > 0$ and spectral measure H and let $A : \mathbb{B}_1 \to \mathbb{B}_2$ be a bounded linear operator. We have

$$\frac{P(\|AX\| > u)}{P(\|X\| > u)} \to \int_{\mathbb{S}_1} \|A\theta\|^\alpha \, H(d\theta) \qquad (u \to \infty).$$

If $H(\{\theta \in \mathbb{S}_1 : A\theta \neq 0\}) > 0$, this limit is positive and AX is regularly varying in $\mathbb{B}_2$ with index α and spectral measure H_A

$$\int_{\mathbb{S}_2} g(\theta) \, H_A(d\theta) = \frac{1}{\int_{\mathbb{S}_1} \|A\theta\|^\alpha \, H(d\theta)} \int_{\mathbb{S}_1} g\left(\frac{A\theta}{\|A\theta\|}\right) \|A\theta\|^\alpha \, H(d\theta).$$

for H_A-integrable $g : \mathbb{S}_2 \to \mathbb{R}$.

The transformed spectral measure can be simulated from by a rejection algorithm

The expression for H_A has the following probabilistic meaning:

$$H_A = \mathscr{L}\left(\frac{A\Theta}{\|A\Theta\|} \,\middle|\, U \leqslant \frac{\|A\Theta\|^\alpha}{\|A\|^\alpha} \right).$$

- Θ is a random element in $\mathbb{S}_1$ with distribution H
- $U \sim \text{Uniform}(0, 1)$ independent of Θ

Rejection algorithm

Generating a random draw Θ_A from H_A:

1. Draw $\Theta \sim H$ and $U \sim \text{Uniform}(0, 1)$ independently.
2. If $U \leqslant \|A\Theta\|^\alpha / \|A\|^\alpha$, then return $\Theta_A = A\Theta / \|A\Theta\|$ and stop.
3. Otherwise, go back to step 1.

Extremes of Stationary Sequences

Infinite random sums

Let $\mathbb{B}_1$ and $\mathbb{B}_2$ be real, separable Banach spaces.

Tail behavior of the $\mathbb{B}_2$-valued infinite random sum

$$X = \sum_n T_n Z_n$$

- $(Z_n)_{n\in\mathbb{Z}}$ iid random elements in $\mathbb{B}_1$
- $T_n : \mathbb{B}_1 \to \mathbb{B}_2$ bounded linear operators.

Possible extension: *random* linear operators (e.g. random matrices)

[Hult & Samorodnitsky 2008]

Convergence of the series

Put $V(x) = P(\|Z_n\| > x)$. Assume $V \in RV_{-\alpha}$.
Suppose there exists δ with $0 < \delta < \min(\alpha, 1)$ such that

$$\sum_n \|T_n\|^\delta < \infty.$$

As $E[\|Z_n\|^\delta] = \int_0^\infty V(x^{1/\delta})\, dx < \infty$, we have

$$E[(\sum_n \|T_n Z_n\|)^\delta] \leqslant \sum_n \|T_n\|^\delta \, E[\|Z_n\|^\delta] < \infty,$$

so that the series $X = \sum_n T_n Z_n$ converges absolutely almost surely.
Moreover, the tail of $\|X\|$ is of the same order as the one of $\|Z_n\|$:

$$\frac{P(\|X\| > x)}{V(x)} \leqslant \frac{P(\sum_n \|T_n\| \|Z_n\| > x)}{V(x)} \to \sum_n \|T_n\|^\alpha < \infty$$

[Resnick 1987, Lemma 4.24; A.s. convergence under weaker conditions in Mikosch & Samorodnitsky (2000), Lemma A.3]

Regular variation of the summands

Now assume that the common distribution of the random elements Z_n is regularly varying with index α and spectral measure H. We have

$$\lim_{x\to\infty} \frac{P(\|T_n Z_n\| > x)}{V(x)} = \int_{\mathbb{S}_1} \|T_n\theta\|^\alpha \, H(d\theta) =: c_n.$$

Moreover, if $c_n > 0$, then $T_n Z_n$ is regularly varying in $\mathbb{B}_2$ with index α and with spectral measure H_n given by

$$\int_{\mathbb{S}_2} f(\theta)\, H_n(d\theta) = \frac{1}{c_n} \int_{\mathbb{S}_1} f(T_n\theta/\|T_n\theta\|)\, \|T_n\theta\|^\alpha \, H(d\theta)$$

for H_n-integrable functions $f : \mathbb{S}_2 \to \mathbb{R}$.

The single-shock heuristic (1)

- Let $(Z_i)_{i\in\mathbb{Z}}$ be an iid sequence in $\mathbb{B}_1$.
- Let $T_i : \mathbb{B}_1 \to \mathbb{B}_2$, $i \in \mathbb{Z}$, be bounded linear operators.

Proposition

If

(i) $x \mapsto V(x) = P(\|Z_i\| > x)$ is $RV_{-\alpha}$ for some $\alpha > 0$,

(ii) $\lim_{x\to\infty} P(\|T_i Z_i\| > x)/V(x) = c_i \in [0,\infty)$ for all $i \in \mathbb{Z}$,

(iii) $\sum_i \|T_i\|^\delta < \infty$ for some $0 < \delta < \min(\alpha, 1)$,

then the series $\sum_i T_i Z_i$ is almost surely absolutely convergent and

$$\begin{aligned} &\lim_{x\to\infty} \frac{1}{V(x)} E\Big|I(\|\textstyle\sum_i T_i Z_i\| > x) - \sum_i I(\|T_i Z_i\| > x)\Big| \\ =&\lim_{x\to\infty} \frac{1}{V(x)} E\Big|I(\textstyle\sum_i \|T_i Z_i\| > x) - \sum_i I(\|T_i Z_i\| > x)\Big| \\ =&0 \end{aligned}$$

The single-shock heuristic (2)

Corollary

$$
\begin{aligned}
\lim_{x\to\infty} \frac{P(\|\sum_i T_i Z_i\| > x)}{V(x)} &= \lim_{x\to\infty} \frac{P(\sum_i \|T_i Z_i\| > x)}{V(x)} \\
&= \lim_{x\to\infty} \frac{\sum_i P(\|T_i Z_i\| > x)}{V(x)} = \sum_i c_i < \infty.
\end{aligned}
$$

Extension of Lemma 4.24 in Resnick (1987).

The spectral measure of the series is a mixture of those of the summands

Proposition

If the common distribution of the independent random elements Z_n ($n \in \mathbb{Z}$) is regularly varying with index α and spectral measure H and if $\sum_n \|T_n\|^\delta < \infty$, then

$$\lim_{x \to \infty} \frac{P(\|\sum_n T_n Z_n\| > x)}{V(x)} = \lim_{x \to \infty} \frac{P(\sum_n \|T_n Z_n\| > x)}{V(x)} = \sum_n c_n < \infty.$$

If $\sum_n c_n > 0$, then the random series $X = \sum_n T_n Z_n$ is regularly varying with index α too, its spectral measure H_X being given by

$$H_X = \sum_n p_n H_n$$

$$p_n = \frac{c_n}{\sum_k c_k} = \lim_{x \to \infty} P(\|T_n Z_n\| > x \mid \|\sum_k T_k Z_k\| > x).$$

The spectral measure reflects the biggest-shock heuristic

The spectral measure H_X can be written as

$$\int f \, dH_X = \frac{\displaystyle\sum_{n\in\mathbb{Z}} E\left[f\left(\frac{T_n(\Theta_Z)}{\|T_n(\Theta_Z)\|}\right)\|T_n(\Theta_Z)\|^\alpha\right]}{\displaystyle\sum_{n\in\mathbb{Z}} E[\|T_n(\Theta_Z)\|^\alpha]},$$

with Θ_Z distributed according to the spectral measure of Z.

Special case: Linear combinations with random coefficients

In case $\mathbb{B}_1 = \mathbb{R}$ we can write $\mathbb{B}_2 = \mathbb{B}$ and the series X is an infinite linear combination of the elements $\psi_i = T_i(1) \in \mathbb{B}$ with random coefficients Z_i:

$$X = \sum_i Z_i \, \psi_i.$$

The spectral measure of X is equal to

$$H_X = \mathscr{L}(\Theta_Z \psi_N / \|\psi_N\|)$$

with

- Θ_Z a random variable in $\{-1, +1\}$
- N an integer-valued random variable independent of Θ_Z and s.t.

$$P(N = n) = p_n = \frac{\|\psi_n\|^\alpha}{\sum_k \|\psi_k\|^\alpha} \qquad (n \in \mathbb{Z})$$

[Davis & Resnick 1985; Embrechts, Klüppelberg, Mikosch, 1997; Davis & Mikosch 2006]

Extremes of Stationary Sequences

Linear processes with regularly varying innovations

Rather than a single random series, we now study the linear process

$$X_t = \sum_i T_i Z_{t-i}, \qquad t \in \mathbb{Z}.$$

with

- $(Z_n)_{n\in\mathbb{Z}}$ is an iid sequence of $RV(\alpha)$ random elements in $\mathbb{B}_1$
- $T_n : \mathbb{B}_1 \to \mathbb{B}_2$ are bounded linear operators such that $\sum_n \|T_n\|^\delta < \infty$ for some $0 < \delta < \min(\alpha, 1)$

The random series defining X_t converges absolutely and $(X_t)_{t\in\mathbb{Z}}$ is a stationary time series in $\mathbb{B}_2$.

The signature of the series given a shock at a certain moment

If $c_n > 0$, where

$$c_n = \int_{\mathbb{S}_1} \|T_n\theta\|^\alpha \, H(d\theta)$$

we can define a probability measure κ_n on the space $\mathbb{B}_2^{\mathbb{Z}}$ of $\mathbb{B}_2$-valued sequences endowed with the product topology by

$$\int_{\mathbb{B}_2^{\mathbb{Z}}} f(\theta_{-s}, \ldots, \theta_t) \, \kappa_n\big(d(\theta_n)_{n\in\mathbb{Z}}\big)$$
$$= \frac{1}{c_n} \int_{\mathbb{S}_1} f\left(\frac{T_{-s+n}\theta}{\|T_n\theta\|}, \ldots, \frac{T_{t+n}\theta}{\|T_n\theta\|}\right) \|T_n\theta\|^\alpha \, H(d\theta), \quad (1)$$

for nonnegative integer s, t and for bounded and continuous $f : \mathbb{B}_1^{t+s+1} \to \mathbb{R}$.

The spectral process is a mixture over the signature patterns

Proposition

If $\sum_n c_n > 0$, then $(X_t)_{t\in\mathbb{Z}}$ is a regularly varying stationary time series in $\mathbb{B}_2$ with index α, its spectral process $(\Theta_t)_{t\in\mathbb{Z}}$ having law κ equal to

$$\kappa = \textstyle\sum_n p_n \kappa_n$$

$$\text{where } p_n = \frac{c_n}{\sum_k c_k}$$

i.e.

$$E\big[f\big((\Theta_t)_{t\in\mathbb{Z}}\big)\big] = \frac{\displaystyle\sum_{n\in\mathbb{Z}} E\left[f\left(\frac{T_{n+t}(\Theta_Z)}{\|T_n(\Theta_Z)\|}\right)\|T_n(\Theta_Z)\|^\alpha\right]}{\displaystyle\sum_{n\in\mathbb{Z}} E[\|T_n(\Theta_Z)\|^\alpha]},$$

Simulating the spectral process

1. Draw a random integer N from $(p_n)_{n\in\mathbb{Z}}$.
2. Independently from N and from each other, draw $\Theta_Z \sim H$ and $U \sim \mathrm{Uniform}(0, 1)$.
3. If $U \leqslant \|T_N\Theta_Z\|^\alpha / \|T_N\|^\alpha$, then return $\Theta_t = T_{N+t}\Theta_Z / \|T_N\Theta_Z\|$ for all $t \in \mathbb{Z}$ and stop.
4. Otherwise, go back to step 2.

Extremes of Stationary Sequences

Autoregressive equation

AR(1) process in $\mathbb{B} = \mathbb{B}_1 = \mathbb{B}_2$:

$$X_t = TX_{t-1} + Z_t, \qquad t \in \mathbb{Z}.$$

- iid innovations Z_t in Banach, $RV(\alpha, H)$
- $T : \mathbb{B} \to \mathbb{B}$ bounded linear operator such that $\|T^m\| < 1$ for some integer $m \geqslant 1$

Note: faily general, since by considering sequence spaces,
an arbitrary linear process can be represented as
the image of a linear operator applied to an AR(1) process

The AR(1) equation has a regularly varying solution

The AR(1) equation has a stationary solution given by

$$X_t = \sum_{n\geqslant 0} T^n Z_{t-n}, \qquad t \in \mathbb{Z},$$

The tail of $\|X_t\|$ satisfies

$$\lim_{x\to\infty} \frac{P(\|X_t\| > x)}{P(\|Z_0\| > x)} = \sum_{n\geqslant 0} \int_{\mathbb{S}} \|T^n\theta\|^\alpha H(d\theta)$$

$(X_t)_{t\in\mathbb{Z}}$ is regularly varying with index $\alpha > 0$ and with spectral process as described above.

- $p_n = 0$ for all $n < 0$
- If $p_{n_0} = 0$ for some integer $n_0 \geqslant 1$, then $p_n = 0$ for all $n \geqslant n_0$

Simulating the spectral process of an AR(1) process

1. Draw a random nonnegative integer N from $(p_n)_{n\geqslant 0}$.
2. Independently from N and from each other, draw $\Theta_Z \sim H$ and $U \sim \text{Uniform}(0,1)$.
3. If $U \leqslant \|T^N \Theta_Z\|^\alpha / \|T^N\|^\alpha$, then return

$$\Theta_{-N} = \frac{\Theta_Z}{\|T^N \Theta_Z\|}, \qquad \Theta_{-N+h} = \begin{cases} T^h \Theta_{-N} & \text{if } h > 0, \\ 0 & \text{if } h < 0. \end{cases}$$

4. Otherwise, go back to step 2.

Extremes of Stationary Sequences

Main findings

- Regular variation is preserved by bounded linear operators
- Tails of random series with independent, regularly varying components governed by the single-shock heuristic
- AR(1) processes: simple structure of the spectral process, readily simulated

Thank you!

References (1)

- Basrak, B., Krizmanić, D., Segers, J. (2011), *Ann. Probab.* 40, 2008–2033.
- Bartkiewicz, K., Jakubowski, A., Mikosch, T., Wintenberger, O. (2011) *Probab. Th. Rel. Fields* 150, 337–372.
- Basrak, B., Segers, J. (2009), *Stoch. Process. Applic.* 119, 1055–1080
- Bortot, P., Coles, S.G. (2000) *Bernoulli* 6, 183–190.
- Bortot, P., Coles, S.G. (2003) *J. R. Statist. Soc. B* 65, 851–867.
- Davis, R.A., Hsing, T. (1995), *Ann. Probab.* 23, 879–917.
- Davis, R.A., Mikosch, T. (1998), *Ann. Statist.* 26, 2049–2080.
- Davis, R.A., Mikosch, T. (2006), *Stoch. Process. Applic.* 118, 560–584.
- Davis, R.A., Mikosch, T. (2009), *Bernoulli* 15, 977–1009.
- Davis, R.A., Resnick, S.I. (1985), *Ann. Probab.* 13, 179–195.

References (2)

- De Haan, L., Resnick, S.I., Rootzén, H., & de Vries, C. (1989) *Stoch. Process. Applic.* 32, 213–224.
- Drees, H., Rootzén, H. (2010), *Ann. Statist.* 38, 2145–2186.
- Hult, H., Lindskog, F. (2005) *Stoch. Process. Applic.* 115, 249–274.
- Hult, H., Lindskog, F. (2006) *Publ. Inst. Math. (Beograd) (N.S.)* 80, 121–140.
- Hult, H., Samorodnitsky, G. (2008) *Bernoulli* 14, 838–864.
- Meinguet, T., Segers, J. (2010) arXiv:1001.3262.
- Mikosch, T., Samorodnitsky, G. (2000) *Ann. Appl. Probab.* 10, 1025–1064.
- Leadbetter, R. (1983) *Z. Wahrscheinlichkeitsth.* 65, 291–306.
- Mikosch, T., Wintenberger, O. (2012a) *Probab. Th. Rel. Fields* DOI 10.1007/s00440-012-0445-0.
- Mikosch, T., Wintenberger, O. (2012b) Mimeo.
- Nandagopalan, S. (1994) *J. Research of the National Institute of Standards and Technology* 99, 543–550.

References (3)

- O'Brien, G.L. (1987) *Ann. Probab.* 15, 281–291.
- Perfekt, R. (1994) *Ann. Appl. Probab.* 4, 529–548.
- Resnick, S. I. (1987) *Extreme Values, Regular Variation and Point Processes*, Springer.
- Resnick, S. I. (2007) *Heavy-Tailed Phenomena*, Springer.
- Resnick, S. I., Zeber, D. (2011) arXiv:...
- Rootzén, H. (1988) *Adv. Appl. Probab.* 20, 371–390.
- Segers, J. (2003) *Adv. Appl. Probab.* 35, 1028–1045.
- Segers, J. (2005) *Statist. Probab. Letters* 74, 330–336.
- Segers, J. (2007) arXiv:0701411.
- Smith, R. (1992) *J. Appl. Probab.* 29, 37–45.
- Wintenberger, O. (2012) Mémoire d'Habilitation, Université Paris–Dauphine.
- Yun, S. (2000) *J. Appl. Probab.* 37, 29–44.