Disjoint and sliding blocks estimators for heavy tailed time series

Rafał Kulik

Department of Mathematics and Statistics (University of Ottawa)

One World Extremes Seminar Series

14 December 2020

Plan

1. Introduction
   - Regularly Varying Time Series
2. Clusters of extremes
   - Definition. Existence. Representation
   - Examples
3. Estimation of cluster indices
   - Disjoint blocks estimators
   - Sliding blocks estimators
   - PoT vs. block maxima
   - Questions (with/without answers

Motivation

Let $\{X_j^\dagger, j \in \mathbb{Z}\}$ be a regularly varying sequence of i.i.d. nonnegative random variables with the tail distribution function $\overline{F}$. In particular:

- $\lim_{x\to\infty} \overline{F}(tx)/\overline{F}(x) = t^{-\alpha}$ for some $\alpha > 0$.
- There exists a sequence $a_n \to \infty$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(\max\{X_1^\dagger, \ldots, X_n^\dagger\} \le a_n x\right) = \exp(-x^{-\alpha}) \ .$$

Let now $\{X_j, j \in \mathbb{Z}\}$ be a stationary regularly varying sequence with the same marginal tail df $\overline{F}$. Then

$$\lim_{n\to\infty} \mathbb{P}\left(\max\{X_1, \ldots, X_n\} \le a_n x\right) = \exp(-\theta x^{-\alpha}) \ ,$$

where $\theta \in (0, 1]$ is called the *extremal index* (whenever exists). The extremal index can be represented as

$$\lim_{x\to\infty} \mathrm{E}[H(\{X_j/x, j \in \mathbb{Z}\})]$$

for some $H : \mathbb{R}_+^{\mathbb{Z}} \to \mathbb{R}$: $H(\boldsymbol{x}) = \mathbb{1}\{\max_{j\in\mathbb{Z}} x_j > 1\}$.

Questions:

- Can we consider different functionals $H : \mathbb{R}_+^{\mathbb{Z}} \to \mathbb{R}$?
- Yes, for specific choices of H we will define H-cluster indices.
- How to estimate H-cluster indices? Disjoint vs. sliding blocks estimators.

Let $\{X_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying nonnegative time series with marginal distribution function F and tail index $\alpha > 0$. This means that for all integers $s \leq t$, there exists a non zero Radon measure $\boldsymbol{\nu}_{s,t}$ such that

$$\lim_{x\to\infty} \frac{\mathbb{P}((X_s, \ldots, X_t) \in xA)}{\mathbb{P}(X_0 > x)} = \boldsymbol{\nu}_{s,t}(A) \,,$$

for all sets A separated from $\mathbf{0}$ satisfying $\boldsymbol{\nu}_{s,t}(\partial A) = 0$.

Tail measure

There exists a measure $\boldsymbol{\nu}$ defined on $\mathbb{R}_+^{\mathbb{Z}}$ such that

- $\boldsymbol{\nu}_{s,t} = \boldsymbol{\nu} \circ p_{s,t}^{-1}$;
- $\boldsymbol{\nu}(\{y \in \mathbb{R}_+^{\mathbb{Z}} : y_0 > 1\}) = 1$;
- $\boldsymbol{\nu}$ is homogeneous with index $-\alpha$.

Note that $\boldsymbol{\nu}$ is an infinite measure on $\mathbb{R}_+^{\mathbb{Z}}$. [1] Define

$$\eta = \boldsymbol{\nu}(\cdot \cap \{y \in \mathbb{R}_+^{\mathbb{Z}} : y_0 > 1\}) .$$

Let $\boldsymbol{Y} = \{Y_j, j \in \mathbb{Z}\}$ be a random element with distribution η. It is called the **tail process**. Different representation of the tail process:[2]

$$\mathbb{P}((Y_i, \dots, Y_j) \in \cdot) = \lim_{x \to \infty} \mathbb{P}(x^{-1}(X_i, \dots, X_j) \in \cdot \mid X_0 > x) .$$

The tail process $\boldsymbol{Y}$ is not stationary. Formulas exist for time series models.

[1] Owada and Samorodnitsky (2010)

[2] Basrak and Segers (2009)

Clusters of extremes, cluster functionals

- Let H be a functional defined on $\mathbb{R}_+^{\mathbb{Z}}$ and such that its values do not depend on coordinates that are equal to zero. That is, for $\mathbf{x} = \{x_j, j \in \mathbb{Z}\} \in \mathbb{R}_+^{\mathbb{Z}}$ we denote $\mathbf{x}_{i,j} = (x_i, \dots, x_j) \in \mathbb{R}_+^{(j-i+1)}$. Then, we identify $H(\mathbf{x}_{i,j})$ with $H((\mathbf{0}, \mathbf{x}_{i,j}, \mathbf{0}))$, where $\mathbf{0} \in \mathbb{R}_+^{\mathbb{Z}}$ is a zero vector.
- Such functionals H will be called cluster functionals.
- Given a cluster functional H on $\mathbb{R}_+^{\mathbb{Z}}$, we want to define the limiting quantity (cluster index):

$$\boldsymbol{\nu}^*(H) = \lim_{n\to\infty} \boldsymbol{\nu}_n^*(H) = \lim_{n\to\infty} \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n \mathbb{P}(X_0 > u_n)} .$$

with $r_n, u_n \to \infty$.

Cluster indices - existence and representation

When does the limit exist? We need assumptions on r_n, u_n; time series; and functionals H.

- Let $r_n\mathbb{P}(X_0 > u_n) \to 0$ and $n\mathbb{P}(X_0 > u_n) \to \infty$;
- Anticlustering condition (extremes cannot persists for infinite horizon time): We say that Condition $\mathcal{AC}(r_n, u_n)$ holds if for every $x, y \in (0, \infty)$, [3]

$$\lim_{\ell\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{\ell\le|j|\le r_n} X_j > u_n x \mid X_0 > u_n y\right) = 0 \,.$$

  The condition is valid for e.g. geometrically ergodic Markov chains. [4]
- H cannot be arbitrary: take $H \equiv 1$. Then $\boldsymbol{\nu}^*(H) = \infty$. On the other hand, $\Upsilon(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\}$, then $\boldsymbol{\nu}^*(\Upsilon) \equiv 1$.

[3] Davis and Hsing (1995)

[4] Kulik, Soulier, Wintenberger (2019)

Cluster indices - existence and representation

Recall:

$$\boldsymbol{\nu}_n^*(H) = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n \mathbb{P}(X_0 > u_n)} \; .$$

Theorem 1

Assume that $\mathcal{AC}(r_n, u_n)$ holds. Then the sequence of measures $\boldsymbol{\nu}_n^*$ converges to $\boldsymbol{\nu}^*$. The mode of convergence is vague# convergence on $\tilde{\ell}_0 \setminus \{\mathbf{0}\}$, where ℓ_0 is the set of elements of $\mathbb{R}^{\mathbb{Z}}$ that vanish at infinity and $\tilde{\ell}_0$ is the set of equivalence classes.

In other words, $\boldsymbol{\nu}_n^*(H) \to \boldsymbol{\nu}^*(H)$ for all bounded continuous shift invariant functions H with support separated from $\mathbf{0}$.

Representation:

$$\boldsymbol{\nu}^*(H) = \mathbb{E}\left[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}\right] = \mathbb{E}\left[H(\boldsymbol{Y})\mathbb{1}\left\{\sup_{j\le -1} Y_j \le 1\right\}\right] \; .$$

Cluster indices - examples

- the "extremal index" obtained with $H(\boldsymbol{x}) = \mathbb{1}\left\{\sup_{j\in\mathbb{Z}} x_j > 1\right\}$:
- the cluster size distribution obtained with

$$H(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j\in\mathbb{Z}} \mathbb{1}\{x_j > 1\} = m\right\}, \quad m \in \mathbb{N} ;$$

- a stop-loss index of a univariate time series obtained with

$$H(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j\in\mathbb{Z}} (x_j - 1)_+ > \eta\right\}, \quad \eta > 0 ;$$

- a large deviation index of a univariate time series obtained with [5]

$$H(\boldsymbol{x}) = \mathbb{1}\{K(\boldsymbol{x}) > 1\}, \quad K(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} x_j ;$$

[5] Mikosch and Wintenberger (2013, 2014)

"Proof" of Theorem 1: [6] Assume $H(\boldsymbol{x}) = 0$ whenever $\max_{j\in\mathbb{Z}} x_j < 1$. Then

$$\frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n\mathbb{P}(X_0 > u_n)} = \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{\max_{i=1,\dots,r_n} X_i > u_n\}]}{r_n\mathbb{P}(X_0 > u_n)}$$

$$= \sum_{j=1}^{r_n} \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{\max_{i=1,\dots,j-1} X_i \le u_n\}\mathbb{1}\{X_j > u_n\}]}{r_n\mathbb{P}(X_0 > u_n)}$$

$$= \frac{1}{r_n}\sum_{j=1}^{r_n} \frac{\mathbb{E}[H(\boldsymbol{X}_{1-j,r_n-j}/u_n)\mathbb{1}\{\max_{i=1-j,\dots,-1} X_i \le u_n\}\mathbb{1}\{X_0 > u_n\}]}{\mathbb{P}(X_0 > u_n)}$$

$$\approx \int_0^1 g_n(s)\mathrm{d}s$$

with $g_n(s)$ defined by

$$g_n(s) = \mathbb{E}\left[H(\boldsymbol{X}_{1-[r_ns],r_n-[r_ns]}/u_n)\mathbb{1}\left\{\max_{i=1-[r_ns],\dots,-1} X_i \le u_n\right\} \mid X_0 > u_n\right] .$$

[6] Planinic and Soulier (2018); Chapter VI of Kulik and Soulier (2020)

Disjoint blocks estimators

Define $m_n = [n/r_n]$ and consider the statistic

$$\widetilde{\mathrm{DB}}_n(H) = \frac{1}{n\mathbb{P}(X_0 > u_n)} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/u_n) \ .$$

Note that

$$\boldsymbol{\nu}^*(H) = \lim_{n\to\infty} \mathbb{E}[\widetilde{\mathrm{DB}}_n(H)] \ .$$

In order to replace u_n, take a sequence of integers $k \to \infty$ such that $k/n \to 0$ and define $u_n = F^{\leftarrow}(1 - k/n)$. Define the blocks estimator

$$\widehat{\mathrm{DB}}_n(H) = \frac{1}{k} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/X_{(n:n-k)}) \ ,$$

where $X_{(n:n)} \geq \cdots \geq X_{(n:1)}$.

Blocks estimators - conditions

Let $H_s = H(\cdot/s)$, $s > 0$ and $\mathcal{H} = \{H_s, s \in [s_0, t_0]\}$, $(0 < s_0 < 1 < t_0)$ be a linear subspace of $L^2(\boldsymbol{\nu}^*)$ such that

(BCLT1) $\lim_{n\to\infty} \boldsymbol{\nu}_n^*(H_s) = \boldsymbol{\nu}^*(H_s)$, $\lim_{n\to\infty} \boldsymbol{\nu}_n^*(H_s H_t) = \boldsymbol{\nu}^*(H_s H_t)$.

(BCLT2) $\lim_{n\to\infty} \boldsymbol{\nu}_n^* \left(H_s^2 \mathbb{1}\left\{ |H_s| > \eta\sqrt{n\mathbb{P}(X_0 > u_n)} \right\} \right) = 0$.

(BCLT3) For all $H_s \in \mathcal{H}$ there exist functions K_n such that

$$|H_s\left(\frac{X_1,\dots,X_{r_n}}{u_n}\right) - H_s\left(\frac{X_1,\dots,X_{r_n-\ell_n}}{u_n}\right)| \le K_n(X_{r_n-\ell_n+1},\dots,X_{r_n}) ,$$

and

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(X_0 > u_n)} \mathbb{E}\left(K_n^2(X_1,\dots,X_{\ell_n})\right) = 0 .$$

Disjoint blocks estimator - CLT

Theorem 2

Let $\{X_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying univariate and β-mixing time series (with "good" rates) and $s > 0$. Under the above conditions

$$\left(\sqrt{n\mathbb{P}(X_0 > u_n)}\left\{\widetilde{\mathrm{DB}}_n(H_s) - \mathbb{E}[\widetilde{\mathrm{DB}}_n(H_s)]\right\}, H_s \in \mathcal{H}\right) \xrightarrow{fi.di.} (\mathbb{G}(H_s), H_s \in \mathcal{H}) ,$$

where $\mathbb{G}$ is a centered Gaussian process with the covariance $\boldsymbol{\nu}^*(H_s H_t)$. Under additional conditions (a version of the anticlustering condition $\mathcal{AC}(r_n, u_n)$, bias conditions and assumptions on $H_s, s \in [s_0, t_0]$ with $0 < s_0 < 1 < t_0 < \infty$) we have

$$\sqrt{k}\left\{\widehat{\mathrm{DB}}_n(H) - \boldsymbol{\nu}^*(H)\right\} \xrightarrow{\mathrm{d}} \mathbb{G}^*(H) ,$$

where $\mathbb{G}^*(H) = \mathbb{G}(H - \boldsymbol{\nu}^*(H)\Upsilon)$, $\Upsilon(\boldsymbol{x}) = \sum_j \mathbb{1}\{x_j > 1\}$.

Disjoint blocks estimator - CLT

"Proof" of Theorem 2:

- Thanks to mixing, the blocks can be considered as independent.
- The variance of the disjoint blocks estimator is then

$$\frac{m_n}{n\mathbb{P}(X_0 > u_n)}\mathrm{Var}(H(\boldsymbol{X}_{1,r_n}/u_n))$$
$$\sim \frac{1}{r_n\mathbb{P}(X_0 > u_n)}\mathbb{E}[H^2(\boldsymbol{X}_{1,r_n}/u_n)] \sim \boldsymbol{\nu}^*(H^2)\ .$$

[7]

[7] Drees and Rootzen (2010), Chapter X of Kulik and Soulier (2020)

Example 1: Disjoint Blocks Estimator of the Extremal Index

The data-based disjoint blocks estimator of the extremal index θ is

$$\widehat{\theta}_{n,k} = \frac{1}{k} \sum_{i=1}^{[n/r_n]} \mathbb{1}\left\{\max\{X_{(i-1)r_n+1}, \ldots, X_{ir_n}\} > X_{(n:n-k)}\right\} .$$

The limiting distribution is normal with mean zero and variance

$$\sigma^2(\theta) := \theta + \theta^2 \boldsymbol{\nu}^*(\Upsilon^2) - 2\theta^2 \,, \quad \boldsymbol{\nu}^*(\Upsilon^2) = \sum_{j\in\mathbb{Z}} \mathbb{P}(Y_j > 1) \,.$$

Example 2: Blocks Estimator of the Large Deviation Index

The data-based disjoint blocks estimator of the large deviation index is

$$\widehat{\theta}_{\mathrm{largedev},n,k} := \frac{1}{k}\sum_{i=1}^{m_n} \mathbb{1}\left\{\left(\sum_{j=(i-1)r_n+1}^{ir_n} X_j\right) > X_{(n:n-k)}\right\} .$$

The limiting variance is

$$\sigma^2 := \theta_{\mathrm{largedev}} + \theta_{\mathrm{largedev}}^2 \boldsymbol{\nu}^*(\Upsilon^2) - 2\theta_{\mathrm{largedev}}\mathbb{P}\left(\left(\sum_{j\in\mathbb{Z}} Y_j\right) > 1\right) .$$

Sliding blocks estimators

We consider

$$\widetilde{\mathrm{SB}}_n(H) = \frac{1}{r_n n \mathbb{P}(X_0 > u_n)} \sum_{i=1}^{n-r_n} H(\boldsymbol{X}_{(i-1)r_n+1, ir_n}/u_n) \;,$$

and

$$\widehat{\mathrm{SB}}_n(H) = \frac{1}{kr_n} \sum_{i=1}^{n-r_n} H(\boldsymbol{X}_{(i-1)r_n+1, ir_n}/X_{(n:n-k)}) \;.$$

Sliding blocks estimator - CLT

Theorem 3

Let $\{X_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying univariate and β-mixing time series (with "good" rates) and $s > 0$. Under the above conditions

$$\left(\sqrt{n\mathbb{P}(X_0 > u_n)}\left\{\widetilde{\mathrm{SB}}_n(H_s) - \mathbb{E}[\widetilde{\mathrm{SB}}_n(H_s)]\right\}, H_s \in \mathcal{H}\right) \stackrel{\text{fi.di.}}{\longrightarrow} (\mathbb{G}(H_s), H_s \in \mathcal{H}) ,$$

where $\mathbb{G}$ is a centered Gaussian process with the covariance $\boldsymbol{\nu}^*(H_s H_t)$. Under additional conditions (a version of the anticlustering condition $\mathcal{AC}(r_n, u_n)$, bias conditions and assumptions on $H_s, s \in [s_0, t_0]$) we have

$$\sqrt{k}\left\{\widehat{\mathrm{SB}}_n(H) - \boldsymbol{\nu}^*(H)\right\} \stackrel{\mathrm{d}}{\longrightarrow} \mathbb{G}^*(H) ,$$

where $\mathbb{G}^*(H) = \mathbb{G}(H - \boldsymbol{\nu}^*(H)\Upsilon)$, $\Upsilon(\boldsymbol{x}) = \sum_j \mathbb{1}\{x_j > 1\}$.

8

[8] Cissokho and Kulik (2020)

Sliding blocks estimator - CLT

"Proof" of Theorem 3:

- Contribution to the variance from two adjacent blocks.
- For $\xi \in (0,1)$

$$\lim_{n\to\infty} \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n) H(\boldsymbol{X}_{1+[\xi r_n],r_n+[\xi r_n]}/u_n)\right]$$
$$= (1-\xi)\boldsymbol{\nu}^*(H^2) = (1-\xi)\mathbb{E}[H^2(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}] .$$

- Integrate over $\xi \in [0,1]$. Multiply by 2.

PoT vs. block maxima

PoT approach:

- Hsing (1991) considers estimation of the extremal index. Drees and Rootzen (2010) is a seminal paper on estimation of cluster functionals using disjoint blocks.
- In Drees and Neblung (2020) the authors study asymptotic normality of the sliding blocks estimators in a general setting. They show that the limiting variance of such estimators does not exceed the one for the disjoint blocks estimators.
- For the extremal index they found the variances to be equal.

PoT vs. block maxima

Block maxima:

- Robert, Segers, Ferro (2009) and Bücher and Segers (2018a, 2018b): Sliding blocks estimators have smaller variance than disjoint blocks estimators.
- Note that the threshold used in Robert et al. (2009) is $r_n\overline{F}(c_{r_n}) \sim 1$. (Here $n\overline{F}(u_n) \to \infty$). The threshold c_n is related to asymptotics for block maxima.
- In the context of Bücher and Segers (2018a, 2018b), contribution to the variance from two adjacent blocks; but in a non-linear way.

Questions

- Runs estimators (Cissokho and Kulik (2021));
- Linear combinations of disjoint blocks, sliding blocks, runs estimators;
- CLT beyond mixing (consistency of disjoint blocks under m-dependent approximations; Kulik and Soulier (2020, Chapter X));
- Resampling (Drees (2015), Jentsch and Kulik (2020), Kulik and Soulier (2020, Chapter XII));
- Bias???
- Long memory???

Thank you!!!!