Large deviations for (pseudo-)regenerative Markov chains

In collaboration with T. Mikosch

Olivier Wintenberger
wintenberger@ceremade.dauphine.fr

CEREMADE, University Paris Dauphine & LFA-CREST.

CPH, May 30, 2013
Motivation: characterization of the limit of partial sums

Let \((X_t)_{t \geq 1}\) be a process with dependent extreme values.

Motivation

Characterization of the limit of \(S_n = \sum_{t=1}^{n} X_t\) under tractable hypothesis?

Example (Errors of empirical statistics)

1. Empirical mean \(\overline{X}_n = \frac{1}{n} \sum_{t=1}^{n} X_t\) when \(\mathbb{E}|X| < \infty\) but \(\mathbb{E}X^2 = \infty\), limit distribution of the error \((\overline{X}_n - \mathbb{E}(X))\) correctly normalized?

2. Empirical autocovariances: for any lag \(h \geq 1\) we have

\[
\hat{\gamma}_n(h) = \frac{1}{n-h} \sum_{j=1}^{n-h} (X_j - \overline{X}_n)(X_{j+h} - \overline{X}_n).
\]
A r.v. \( Y \) is strictly \( \alpha \)-stable distributed iff \( \exists \ a > 0, \ Y_1 \) and \( Y_2 \) independent, distributed as \( Y \) such that \( Y_1 + Y_2 = aY \) in distribution. Then \( Y \) is strictly \( \alpha \)-stable with \( 0 < \alpha \leq 2 \) and c.f. \( \exp(-|x|^\alpha \chi_\alpha(x, b_+, b_-)), \)

\[
\chi_\alpha(x, b_+, b_-) = \frac{\Gamma(2 - \alpha)}{1 - \alpha} ((b_+ + b_-) \cos(\pi \alpha/2) - i \pm_x (b_+ - b_-) \sin(\pi \alpha/2)).
\]
**Theorem (Feller, 1977)**

If $\exists (a_n), a_n > 0$ and $Y$ strict. stable such that

$$a_n^{-1}S_n \rightarrow Y$$  \hspace{1cm} (SSL)

then $X_t$ are iid $RV(\alpha)$ centered r.v. if $\alpha > 1$.

For $\alpha < 2$ and $a_n = L(n)n^{1/\alpha}$ s.t. $\lim_n n\mathbb{P}(\vert X \vert > a_n) = 1$ then $b_+ + b_- = 1$.

Remark that if $0 < \alpha < 1$ then $\mathbb{E}|X| = \infty$. 
Regularly varying sequences

Stationary RV(\(\alpha\)) processes, Basrak & Segers (2009)

\( (X_t) \) is RV(\(\alpha\)) iff \( \exists \) its spectral tail process (\(\Theta_t\)) defined for \( k \geq 0 \), \( u \geq 1 \) when \( x \to \infty \)

\[ \mathbb{P}(X_0 > ux, |X_0|^{-1}(X_0, \ldots, X_k) \in \cdot | |X_0| > x) \xrightarrow{w} u^{-\alpha} \mathbb{P}((\Theta_0, \ldots, \Theta_k) \in \cdot). \]

Example

If \((X_t)\) is iid, \(\Theta_t = 0\) for \( t \geq 1 \) and \( b_\pm = \mathbb{E}[\Theta_0^\alpha] \) for \( \alpha \in (1, 2) \).

Remark that \( b_+ + b_- = \mathbb{E}[\Theta_0^\alpha] + \mathbb{E}[\Theta_0^\alpha] = \mathbb{E}|\Theta_0|^{\alpha} = 1 \) because \( |\Theta_0| = 1 \).
A necessary condition

Theorem (Jakubowski, 1993)

If (SSL) with $a_n = L(n)n^{1/\alpha}$ then it exists a sequence $k_n, n/k_n \to \infty$ such that

$$|\mathbb{E}(e^{i\alpha a_n^{-1}S_n}) - \mathbb{E}(e^{i\alpha a_n^{-1}S_n/k_n})^{k_n}| \to 0.$$  \hspace{1cm} (MX)

Example

(MX) is satisfied for

1. $(X_t)$ iid,
2. $X_t = Y$ strictly stable for all $t \geq 1$!!!
Toward coupling conditions

Remark that $X_t = Y \in RV(\alpha)$ is a stationary sequence satisfying

1. $RV(\alpha)$,
2. $MX$.

However, (SSL) holds iff $Y$ is strictly $\alpha$-stable.

Mixing type conditions sufficient for (MX) excluding the case $X_t = Y$. 
Assume that $X_t = f(\Phi_t)$ where $(\Phi_t)$ is a Markov chain:

$\Phi_t = F(\Phi_{t-1}, \xi_t)$, where $(\xi_t)$ is iid.

**Definition (Coupling scheme, Thorisson (2000))**

Consider $X_t^* = f(\Phi_t^*)$ with $\Phi_t^* = F(\Phi_{t-1}^*, \xi_t)$ for $t \geq 1$ and $(\Phi_0^*, \Phi_0)$ iid.
Coupling conditions

Proposition
If $\sum_t \mathbb{E}|X_t - X^*_t| < \infty$ then (MX) is satisfied

Example (AR(1): $X_t = \rho^t X_0 + \sum_{j=1}^{t} \rho^{t-j} \xi_j$)
When small jumps matter.

The point process approach deals with \( \sum_{t=1}^{n} \delta_{X_t/a_n} \) on some set vanishing around 0.

**Example (Coupled regularly varying Markov chain)**

For \( (T_t) \) iid positive \( RV(\alpha') \), \( (B_t) \) iid Rademacher, \( (\xi_t) \) iid centered \( RV(\alpha) \) with \( \alpha > \alpha' > 1 \) consider \( X_t = B_{N_T(t)} + \xi_t, N_T(t) = \inf\{k \geq 1, T_1 + \cdots + T_k \geq t\} \).

Then

\[
\begin{align*}
\sum_{t=1}^{n} \delta_{X_t/a_n} &\sim \sum_{t=1}^{n} \delta_{\xi_t/a_n} \\
S_n &\sim \sum_{j=1}^{N_T(n)} \pm T_j, N_T(n)E(T) \sim n
\end{align*}
\]

\( \Rightarrow \alpha\)-stable limit,

\( L(n)n^{-\alpha'}S_n\) \( \alpha'\)-stable limit.

**Remark**

- \( \mathbb{E}|X_t - X_t^*| = \mathbb{E}|B_{N_T(t)} - B_{N^*_T(t)}^*| \leq 2\mathbb{P}(T_1 \geq t) = 2L(t)t^{-\alpha'} \).
- Does not work for \( 0 < \alpha' < 1 \).
Vanishing small values condition

Additional hypothesis

Davis and Hsing (1995)

\[
\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{P}\left( \left| \sum_{t=1}^{n} X_t I_{\{|X_t| \leq \epsilon a_n\}} - \mathbb{E}(X_t I_{\{|X_t| \leq \epsilon a_n\}}) \right| > x a_n \right) = 0, \quad x > 0.
\]

(VSV)

Example

Iid \((X_t)\) satisfies (VSV).

Condition (VSV) has to be verified for dependent \((X_t)\).
Identification of the clusters

**SRE:** \( X_t = A_t X_{t-1} + B_t, \ t \geq 1 \) with \( (A_t, B_t) \) iid, \( A_t > 0, \ E A_0^\alpha = 1 \) and \( E |B_0|^{\alpha + \epsilon} < \infty, \ \epsilon > 0 \). The unique stationary solution \( (X_t) \) is RV(\( \alpha \)).

How to identify the clusters?
Approximation by local dependance (Rootzen, 1978)

When is it a good approximation when $m \to \infty$?

Davis & Hsing (1995), Basrak & Segers (2009)

$$\lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{P}\left( \max_{m \leq |i| \leq n/k_n} |X_i| > x a_n \mid |X_0| > x a_n \right) = 0, \quad x > 0.$$  \hspace{1cm} \text{(ALD)}

Under (ALD) $\theta > 0$, i.e. average size of clusters are finite.
Drift condition (DCp)

Two issues

1. Condition (MX) or sufficient coupling is not sufficient for (VSV),
2. Condition (ALD) is not very tractable.

One solution

Let $X_t = f(\Phi_t)$ where $\Phi_t$ is a nice Markov chain. It satisfies Condition (DCp) for $p > 0$ if there exist $\beta \in (0, 1), \ b > 0$ such that for any $y$,

$$
\mathbb{E}(|f(\Phi_1)|^p \mid \Phi_0 = y) \leq \beta |f(y)|^p + b. \quad (\text{DCp})
$$

Remark that (DCp) implies (DCp’) for $p > p’$ (Jensen’s inequality).
Examples for (DCp)

Examples

1. $(X_t)$ iid $RV(\alpha)$ then $\mathbb{E}(|X_1|^p \mid X_0 = y) = \mathbb{E}|X_1|^p =: b$, $0 < p < \alpha$,

2. AR(1): $X_t = \rho X_{t+1} + \xi_t$ with $(\xi_t)$ iid $RV(\alpha)$ then

$$\mathbb{E}(|\rho y + \xi_1|^p \mid X_0 = y) \leq (|\rho|y + (\mathbb{E} |\xi_1|^p)^{1/p})^p \leq \beta y^p + b$$

for $|\rho|^p < \beta < 1$ and all $1 \leq p < \alpha$,

3. $X_t = Y$ then $\mathbb{E}(|X_1|^p \mid X_0 = y) = |y|^p$ does not satisfy (DCp).
Examples for (DCp)

Example

SRE: $X_t = A_t X_{t-1} + B_t$ with $\mathbb{E} A_0^\alpha = 1$ and $\mathbb{E} B_0^{\alpha+\varepsilon} < \infty$ then

$$\mathbb{E}(|A_1 y + B_1|^p \mid X_0 = y) \leq ((\mathbb{E} A_0^p)^{1/p} y + (\mathbb{E} |\xi_1|^p)^{1/p})^p \leq \beta y^p + b$$

for $\mathbb{E} A_0^p < \beta < 1$ as $(\mathbb{E} A_0^p)^{1/p} < (\mathbb{E} A_0^\alpha)^{1/\alpha} = 1$ for $1 \leq p < \alpha$,

Conjecture

If the Markov chain $(\Phi_t) \in \text{RV}(\alpha)$ then it satisfies (DCp).
Regeneration of Markov chains with an accessible atom (Doeblin, 1939)

Definition

$(\Phi_t)$ is a Markov chain of kernel $P$ on $\mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$.

- $A$ is an atom if $\exists$ a measure $\nu$ on $\mathcal{B}(\mathbb{R}^d)$ st $P(x, B) = \nu(B)$ for all $x \in A$.
- $A$ is accessible, i.e. $\sum_k P^k(x, A) > 0$ for all $x \in \mathbb{R}^d$.

Let $(\tau_A(j))_{j \geq 1}$ visiting times to the set $A$, i.e.

$\tau_A(1) = \tau_A = \min\{k > 0 : X_k \in A\}$ and $\tau_A(j + 1) = \min\{k > \tau_A(j) : X_k \in A\}$.

Regeneration cycles

1. $N_A(t) = \#\{j \geq 1 : \tau_A(j) \leq t\}$, $t \geq 0$, is a renewal process,
2. The cycles $(\Phi_{\tau_A(t)+1}, \ldots, \Phi_{\tau_A(t+1)})$ are iid.
Irreducible Markov chain and Nummelnin scheme

**Definition (Minorization condition, Meyn and Tweedie, 1993)**

\[ \exists \, \delta > 0, \text{ a small set } C \in \mathcal{B}(\mathbb{R}^d) \text{ and a distribution } \nu \text{ on } C \text{ such that} \]

\[ P^k(x, B) \geq \delta \nu(B), \quad x \in C, \quad B \in \mathcal{B}(\mathbb{R}^d). \quad (\text{MC}_k) \]

(\text{MC}_1) is called the strongly aperiodic case.

Any irreducible aperiodic Markov chain \((\Phi_t)\) satisfies (\text{MC}_k) for some \(k \geq 1\).

**Nummelnin splitting scheme for pseudo-regenerative Markov chain**

Under (\text{MC}_1) an enlargement of \((\Phi_t)\) on \(\mathbb{R}^d \times \{0, 1\} \subset \mathbb{R}^{d+1}\) possesses an accessible atom \(A = C \times \{1\} \implies \text{the enlarged Markov chain regenerates.}\)
Inference on real data, Bertail and Clementecon (2009)

Squared of log-ratios $X_t = \log(P_t/P_{t-1})^2$ where $(P_t)$ are CAC 40 prices.

Small sets $C = \{X_t^2 \leq a_n\}$ for any $a_n > 0$ (T-chains).
Coupling under (DCp)

Under (DCp) then \( \mathbb{E}e^{c\tau_A(1)} < \infty, \mathbb{P}(\tau_A(1) \geq t) \leq \mathbb{E}e^{c\tau_A(1)}e^{-ct} \) and

\[
\mathbb{E}|X_t - X^*_t| \leq 2\mathbb{E}|X_t|\mathbb{P}(\tau_A(1) \geq t) \leq 2\mathbb{E}|X_t|Ce^{-ct}.
\]
(SSL) for sums of $m$-dependent r.v.

Assume $(X_t, t \leq 0)$ is independent of $(X_t, t \geq m)$ then $\Theta_t = 0$ for $|t| \geq m$.

**Theorem**

If $(X_t)$ is centered RV($\alpha$) with $\alpha > 1$ then it satisfies (SSL) $a_n^{-1} S_n \rightarrow Y$ where $Y$ has c.f. $\exp(-|x|^\alpha \chi_\alpha(x, b_+, b_-))$ with cluster indices

$$b_\pm = \mathbb{E} \left[ \left( \sum_{t=0}^{m-1} \Theta_t \right)_\alpha \pm \left( \sum_{t=1}^{m-1} \Theta_t \right)_\alpha \right].$$
Assume \((X_t = f(\Phi_t))\) where \((\Phi_t)\) (possibly enlarged) possesses an accessible atom \(A\) and an invariant measure \(\pi\) s.t. \(\Phi_0 \sim \pi\).

**Theorem**

*If \((X_t)\) is centered \(RV(\alpha)\) with \(\alpha > 1\) and satisfies \((DCp)\) for \(p < \alpha\) then it satisfies \((SSL)\) with cluster indices*

\[
b_{\pm} = \mathbb{E}\left[\left(\sum_{t=0}^{\infty} \Theta_t\right)^{\alpha}_{\pm} - \left(\sum_{t=1}^{\infty} \Theta_t\right)^{\alpha}_{\pm}\right].
\]
Sketch of the proof

Under (DCp) we have $E|\Theta_k|^p \leq C \rho^k$ for some $C > 0$, $0 < \rho < 1$. In particular $(\Theta_t)$ is a convergent series in $L^{\alpha-1}$.

By the mean value theorem we have there exists $C > 0$

$$E \left[ \left( \sum_{t=0}^{m-1} \Theta_t \right)^\alpha \pm \left( \sum_{t=1}^{m-1} \Theta_t \right)^\alpha \right] \leq C E \left| \sum_{t=0}^{m-1} \Theta_t \right|^{\alpha-1}.$$

By the dominated convergence theorem the cluster index exists.
SRE: $X_t = A_t X_{t-1} + B_t$, then $\Theta_t = \prod_{j=1}^{t} A_j \Theta_0$ satisfies

$$\mathbb{E}|\Theta_t|^\alpha = 1 \implies \mathbb{E}\left(\sum_{t=1}^{\infty} |\Theta_t|^\alpha\right) = \infty.$$
Application to autocorrelograms of squared log-ratios

Assume that $X_t = \log\left(\frac{P_t}{P_{t-1}}\right)^2$ is RV($\alpha$) satisfying (DCp).

Hill’s estimator: $\hat{\alpha} \approx 2$. 
Autocorrelogram in presence of extremes

\[ \hat{\gamma}_n(h) \approx \gamma(h) + Y_1(h) \] asymptotically \( \alpha \approx 1 \)-stable asymmetric distributed.

Analysis on basis of autocorrelogram are not adapted to heavy tailed cases.
Regular variation of cycles

Denoting the independent cycles \( S_A(t) = \sum_{i=1}^{\tau_A(t+1)} f(\Phi_{\tau_A(t)+i}) \),

\[
S_n = \sum_{i=1}^{\tau_A} X_i + \sum_{t=1}^{N_A(n)-1} S_A(t) + \sum_{\tau_A(N_A(n)+1)}^{n} X_i.
\]

**Theorem**

If \((X_t)\) RV\((\alpha)\) with \(\alpha > 0\), \(\alpha \notin \mathbb{N}\) and (DCp) with \(p < \alpha\) and \(b\pm \neq 0\) then

\[
P_A\left( S_A(1) > x \right) \sim_{x\to\infty} b_\pm \mathbb{E}_A(\tau_A) P(|X| > x).
\]

**Remarks**

1. The full cycles \( S_A(t) = \sum_{i=1}^{\tau_A(t+1)} f(\Phi_{\tau_A(t)+i}) \) are regularly varying with the same index \(\alpha > 0\) than \(X_t\),
2. If \(\tau_A\) is independent of \((X_t)\) then \(P_A(S_A(1) > x) \sim_{x\to\infty} \mathbb{E}_A(\tau_A) P(X > x)\),
3. Under (DCp) and \(\mathbb{E}|X|^p\) then \(\mathbb{E}_A|S_A(1)|^p < \infty\).
Precise large deviations for sums

**Corollary (Under the hypothesis of the Theorem)**

If $0 < \alpha < 1$ then 

$$\lim_{n \to \infty} \sup_{x \geq b_n} \left| \frac{\mathbb{P}(\pm S_n > x)}{n \mathbb{P}(|X| > x)} - b_{\pm} \right| = 0,$$

where

$$b_n = n^{1/\alpha} + 1/2 + \epsilon$$

else, if $\mathbb{P}(\tau_A > n) = o(n \mathbb{P}(|X| > c_n))$,

$$\lim_{n \to \infty} \sup_{b_n \leq x \leq c_n} \left| \frac{\mathbb{P}(\pm S_n > x)}{n \mathbb{P}(|X| > x)} - b_{\pm} \right| = 0.$$

Determination of the constant in LD of Davis and Hsing (1995) valid for $\alpha < 2$.

**Sketch of the proof:**

Under $\mathbb{P}(\tau_A > n) = o(n \mathbb{P}(|X| > c_n))$,

$$S_n \approx \sum_{t=1}^{N_A(n)-1} S_A(t).$$

Use Nagaev's precise LD result on the iid regularly varying cycles $S_A(t)$. 
Link between extremal and cluster index, $\Theta_0 = 1$

Under RV($\alpha$) and (DCp), extremal index $\theta_+ = \mathbb{E}[(\sup_{t \geq 0} \Theta_t)^\alpha_+] - (\sup_{t \geq 1} \Theta_t)^\alpha_+]$.

**Example (Asymptotic independence)**

$\Theta_t = 0$ for all $t > 0$ then $b_+ = \theta_+ = 1$.

**Example (AR(1): $X_t = \rho X_{t-1} + \xi_t$, $\forall t \in \mathbb{Z}$ with $\rho > 0$)**

$\Theta_t = \rho^t$ for all $t \geq 0$ then $\theta_+ = 1 - \rho^\alpha$ and $b_+ = \theta_+/(1 - \rho)^\alpha$.

**Example (GARCH(1,1)$^2$: $X_t^2 = \sigma_t^2 Z_t^2$, $\sigma_t^2 = \alpha_0^* + \alpha_1^* X_{t-1}^2 + \beta_1^* \sigma_{t-1}^2$)**

$\Theta_t = (Z_t/Z_0)^2 \prod_{i=1}^t (\alpha_1^* Z_{i-1}^2 + \beta_1^*)$ for all $t \geq 0$ then $b_+$ and $\theta_+$ are explicit.
Peaks over thresholds

Process of exceedances of the squared log-ratios

Exceedances

Time
Description of the clusters

Renormalization by the first exceedance in the cluster
Representation of the average clusters

As. ind., observations, AR(1)  GARCH(1,1)
Conclusions and perspectives on the extremes

**Conclusions**

1. Cluster indices $b_\pm$ determine the asymptotic distribution of the sums of dependent and regularly varying variables,
2. The extremal and cluster indices describe the clusters of extreme values.

**Perspectives**

1. We use Markovian processes and their regenerative structures $\implies$ use also regenerative structures to identify the clusters.
2. Model the extremal dependence in view of the observed clusters $\implies$ introduce new models with extremal behaviors similar than the observed ones.