Contributions to time series analysis: estimation, prediction and extremes

Habilitation

Olivier Wintenberger

CEREMADE, University Paris Dauphine & LFA-CREST

November 23, 2012
List of (pre)publications


List of (pre)publications


Outline

1. **Estimation (III, IV, IX, XI)**
   - Breaks of stationarity
   - Continuously invertible models

2. **Prediction (VIII, XIII, XIV)**
   - Weak transport inequalities
   - Exact oracle inequalities

3. **Extremes (VII, X, XV)**
   - The cluster index
   - Limit theorems for functions of Markov chains
Outline

1. **Estimation (III, IV, IX, XI)**
   - Breaks of stationarity
   - Continuously invertible models

2. **Prediction (VIII, XIII, XIV)**
   - Weak transport inequalities
   - Exact oracle inequalities

3. **Extremes (VII, X, XV)**
   - The cluster index
   - Limit theorems for functions of Markov chains
Motivation: model the dependence structure

Observations exhibit some temporal structure of dependence: iid case non realistic $\rightarrow$ need to model the dependence structure.

Example: log returns of a financial index

![Graph of Footse log-returns over time](image)
General affine models

Definition (Doukhan & W., 2008)

$T$ is an interval of $\mathbb{Z}$. The process $X = (X_t)_{t \in T}$ follows $\mathcal{M}_T(\sigma_\theta, f_\theta)$ iff

$$X_t = \sigma_\theta^* \left( (X_{t-i})_{i \geq 1} \right) Z_t + f_\theta^* \left( (X_{t-i})_{i \geq 1} \right)$$

for all $t \in T$,

with $(Z_t)$ iid such that $\mathbb{E} Z_0 = 0$ and $\mathbb{E} Z_0^2 = 1$, $X_t$ being independent of $(Z_{t+1}, Z_{t+2}, \ldots)$. The functions $\sigma_\theta, f_\theta$ are parametrized by $\theta \in \Theta \subset \mathbb{R}^d$.

Example

- Volatility models: $f_\theta = 0 \implies$ no correlations $\implies$ financial index analysis.
- Autoregressive models: $\sigma_\theta = 1 \implies$ prediction.
The GARCH(1,1) model with breaks

Let $T_j^*, 1 \leq j \leq K^*$ be successive intervals covering $\{1, \ldots, n\}$. Let

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0^*(j) + \alpha_1^*(j) X_{t-1}^2 + \beta_1^*(j) \sigma_{t-1}^2 \quad \forall t \in T_j^*.$$
Approximation of the volatility process

\[ \theta_j^* = (\alpha_0^*(j), \alpha_1^*(j), \beta_1^*(j)) \in \mathbb{R}^3, \ T_j^* = \{\tau_j^* n + 1, \ldots, \tau_{j+1}^* n\} \text{ and } K^* \text{ are unknown.} \]

**Definition**

Let \( \theta_j = (\alpha_0(j), \alpha_1(j), \beta_1(j)) \in \Theta \) and \( T_j = \{\tau_j n + 1, \ldots, \tau_{j+1} n\} \). Define recursively

\[ \hat{\sigma}_t^2(\theta_j) = \alpha_0(j) + \alpha_1(j) X_t^2 + \beta_1(j) \hat{\sigma}_t^2(\theta_j) \quad \forall t \in T_j, \]

from arbitrary initial \( \hat{\sigma}_{\tau_j n}^2(\theta_j) \).
The penalized Quasi Likelihood procedure

The QL over the period \( T_j \geq u_n \) is

\[
2\hat{\mathcal{L}}_n(T_j, \theta_j) := \sum_{t \in T_j} \hat{\ell}_t(\theta_j) \quad \text{with} \quad \hat{\ell}_t(\theta_j) := \frac{X^2_t}{\hat{\sigma}^2_t(\theta_j)} + \log(\hat{\sigma}^2_t(\theta_j)).
\]

We penalize the QL criterion \( \hat{J}_n(K, \pi, \underline{\theta}) := \sum_{j=1}^{K} \hat{\mathcal{L}}_n(T_j, \theta_k), \pi = (\tau_1, \ldots, \tau_K), \underline{\theta} = (\theta_1, \ldots, \theta_K) \), by the number \( K \) of periods:

\[
\tilde{J}_n(K, \pi, \underline{\theta}) := \hat{J}_n(K, \pi, \underline{\theta}) + \kappa_n K.
\]

Consider \( K_{\text{max}} > 0 \) fixed and define

\[
(\hat{K}_n, \hat{\pi}_n, \hat{\underline{\theta}}_n) \in \operatorname{Argmin}_{1 \leq K \leq K_{\text{max}}} \operatorname{Argmin}_{\pi, \underline{\theta}} \tilde{J}_n(K, \pi, \underline{\theta}).
\]
The penalized Quasi Likelihood procedure

**Theorem (Bardet & W., 2009, Bardet et al., 2012)**

*Under uniform moments and weak dependence assumptions, if \( \kappa_n \) is well-chosen then the procedure is consistent and for \( \theta \) asymptotically normal.*

In practice, calibrate \( \kappa_n \) from an adaptation of the slope heuristic:

![Graphs showing QL criterion and Plateau of models](image-url)
Breaks detection on the Footsee index

The statistical procedure detects three breaks:

- First break: announce of the subprime crisis,
- Second break: the Lehman Brothers bankrupt,
- Third break: announce of the G-20 measures to face the crisis.
The EGARCH(1,1) model

**Definition (Nelson, 1991)**

Volatility model with
\[ \log(\sigma_{t+1}^2) = \alpha^* + \beta^* \log(\sigma_t^2) + (\gamma^* Z_t + \delta^* |Z_t|), \forall t \in \mathbb{Z}. \]

The inverted model is not necessarily stable wrt the initial value \( \log(\hat{\sigma}_0^2) \)
\[ \log(\hat{\sigma}_{t+1}^2) = \alpha^* + \beta^* \log(\hat{\sigma}_t^2) + (\gamma^* X_t + \delta^* |X_t|) \exp\left(-\log(\hat{\sigma}_t^2)/2\right), \quad \forall t \geq 0. \]
Invertibility of the EGARCH(1,1) model

Definition (Straumann and Mikosch, 2006)

The model is called invertible when the inverted model is stable.

Example (EGARCH(1,1))

Invertibility conditions $\delta^* \geq |\gamma^*|$ and

$$
\mathbb{E}[\log(\max(\beta^*, 2^{-1}(\gamma^* X_t + \delta^*|X_t|) \exp(-2^{-1}\alpha^*/(1 - \beta^*)) - \beta^*))] < 0.
$$
Continuous invertibility of the EGARCH(1,1) model

QL: \( \sum_{t=1}^{n} \frac{X_t^2}{\hat{\sigma}_t^2(\theta)} + \log(\hat{\sigma}_t^2(\theta)) \) where, starting from \( \log(\hat{\sigma}_0^2(\theta)) \),
\[
\log(\hat{\sigma}_{t+1}^2(\theta)) = \alpha + \beta \log(\hat{\sigma}_t^2(\theta)) + (\gamma X_t + \delta |X_t|) \exp(-\log(\hat{\sigma}_t^2(\theta))/2), \quad \forall t \geq 0.
\]

**Definition (Wintenberger and Cai, 2011)**

The model is called **continuously invertible** on \( \Theta \) when the inverted model is stable for every \( \theta \in \Theta \) and when the asymptotic law is continuous on \( \Theta \).

**Theorem (Wintenberger and Cai, 2011)**

*Under continuous invertibility, the QMLE is strongly consistent and, for EGARCH(1,1), asymptotically normal under second moments of the score.*

**Example (EGARCH(1,1))**

Continuous invertibility on \( \Theta \) if all \( (\alpha, \beta, \gamma, \delta) \in \Theta \) satisfies \( \delta \geq |\gamma| \) and
\[
\mathbb{E}[\log(\max(\beta, 2^{-1}(\gamma X_t + \delta |X_t|) \exp(-2^{-1}\alpha/(1 - \beta)) - \beta))] < 0.
\]
Conclusion and perspectives on the estimation

- Conclusion
  1. Use the QMLE for general affine models or on a continuously invertible domain,
  2. Use the penalized QL for detecting breaks.

- Perspectives
  1. Asymptotic Normality of EGARCH(1,1) with no uniform moment condition
     ⇒ use similar arguments for proving AN in other models (AR-GARCH, BEKK-GARCH).
  2. Proof of the calibration of $\kappa_n$ based on the slope heuristic
     ⇒ non asymptotic criterion.
Outline

1 Estimation (III, IV, IX, XI)
   - Breaks of stationarity
   - Continuously invertible models

2 Prediction (VIII, XIII, XIV)
   - Weak transport inequalities
   - Exact oracle inequalities

3 Extremes (VII, X, XV)
   - The cluster index
   - Limit theorems for functions of Markov chains
Motivation: statistical learning for prediction

Observe \((X_1, \ldots, X_n)\) and predict \(X_{n+1} \in \mathbb{R}\) with \(\hat{X}_{n+1}^{\hat{\theta}}\) for \(\hat{\theta}\) well chosen. Let \(\ell\) be a loss function and \(R(\theta) = \mathbb{E}[\ell(\hat{X}_{n+1}^{\theta}, X_{n+1})]\) be the prediction risk.

**Definition (Oracle inequalities)**

The predictor \(\hat{\theta}\) satisfies an oracle inequality if

\[
R(\hat{\theta}) \leq R(\bar{\theta}) + \Delta_n(\varepsilon) \quad \text{w.p.} \quad 1 - \varepsilon
\]

where \(\bar{\theta}\) is the oracle \(R(\bar{\theta}) \leq \inf_\Theta R(\theta) + c n^{-1}\) and \((\Delta_n)\) is the rate of convergence \((\Delta_n \propto n^{-1} \text{ fast, } \Delta_n \propto n^{-1/2} \text{ slow})\).

Use \(r_n(\theta) = \frac{1}{n} \sum_{t=p+1}^{n} \ell(\hat{X}_t^{\theta}, X_t)\) and the concentration of \(r_n\) around \(R = \mathbb{E}[r_n]\).
Parametric and non parametric settings

Example (Models of predictors)

- Linear predictors \( \hat{X}_{n+1}^\theta = \sum_{j=0}^{p} \theta_j X_{n-j}, \quad \theta = (\theta_j) \in \Theta \subset \mathbb{R}^{p+1}, \)

- Functional predictors \( \hat{X}_{n+1}^\theta = \sum_{i=1}^{j} \theta_i \varphi_i(X_n, \ldots, X_{n-p}) \) for functions \( \varphi_i : \mathbb{R}^p \mapsto \mathbb{R}, \quad \theta \in \Theta = \bigcup_{j=1}^{m} \Theta_j \) with \( \Theta_j \subset \mathbb{R}^j. \)

No model assumption on the observations.
Transport inequalities ($n = 1$)


Let $1 \leq p \leq 2$. The law $\mathbb{P}$ on $\mathbb{R}$ satisfies $T_{p,d}(C)$ if for any law $Q$

$$W_{p,d}(\mathbb{P}, Q) = \inf_{\pi} \mathbb{E}_\pi [d^p(X, Y)]^{1/p} \leq \sqrt{2CK(Q|\mathbb{P})},$$

where $\pi$ is a coupling (the law of $(X, Y)$ with margins $(\mathbb{P}, Q)$) and $K(Q|\mathbb{P}) = \mathbb{E}_Q[\log(dQ/d\mathbb{P})]$.

Standard gaussian laws with different means: $W_{2,\cdot}(\mathbb{P}, Q) = \sqrt{2K(Q|\mathbb{P})}$
Weak transport inequalities

**Definition (Marton, 1996, W., 2012)**

Let $1 \leq p \leq 2$. The law $\mathbb{P} \in \mathbb{R}$ satisfies $\tilde{T}_{p,d}(C)$ if for any law $Q$

$$E_Q[E_\pi[d(X, Y)|Y]^p]^{1/p} \leq \sqrt{2CK(Q|\mathbb{P})}.$$ 

**Remark**

Any measure $\mathbb{P}$ on $\mathbb{R}$ satisfies $\tilde{T}_{p,1}(C)$ where 1 denotes the Hamming distance $d(x, y) = 1_{x \neq y}$ for any $1 \leq p \leq 2$. 
Extension to $\Gamma$-weakly dependent laws

Now $P$ is the law of $(X_t)_{1 \leq t \leq n}$ on $\mathbb{R}^n$. It satisfies $\tilde{T}_{p,d}(C)$ if for all $Q$

$$\mathbb{E}_Q \left[ \sum_{j=1}^{n} \mathbb{E}_\pi [d(X_j, Y_j)|Y]^p \right]^{1/p} \leq \sqrt{2CK(Q|P)}.$$ 

Let $X_0 = Y_0 = x_0 = y_0 \in \mathbb{R}$ and denote $P|_{x^{(i)}}$ the law of $(X_{i+1}, \ldots, X_n)$ conditionally on $(X_i, \ldots, X_0) = x^{(i)} := (x_i, \ldots, x_0)$.

**Definition ($\Gamma$-weak dependence)**

Let $d_1 \leq Md_2$ for $M > 0$. $P$ is $\Gamma_{d_1,d_2}(p)$-weakly dependent if for all $1 \leq i \leq n$, $(x^{(i)}, y_i) \in E^{i+2}$, a trajectory coupling $\pi|_i$ of $(P|_{x^{(i)}}, P|_{y_i,x^{(i-1)}})$ satisfying

$$\mathbb{E}_\pi|_i (d_1^p(X_k, Y_k))^{1/p} \leq \gamma_{k,i}(p) d_2(x_i, y_i), \quad \text{for all} \quad i + 1 \leq k \leq n.$$
Weak transport inequalities

Weak transport of $\Gamma$-weakly dependent laws

$$\Gamma(p) := \begin{pmatrix}
M & 0 & 0 & \cdots & 0 \\
\gamma_{2,1}(p) & M & 0 & \cdots & 0 \\
\gamma_{3,1}(p) & \gamma_{3,2}(p) & M & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma_{n,1}(p) & \gamma_{n,2}(p) & \cdots & \gamma_{n,n-1}(p) & M \\
\end{pmatrix}.$$ 

Theorem (W., 2012)

If $\mathbb{P}$ is $\Gamma_{d_1,d_2}(p)$-weakly dependent and if $\mathbb{P}_{x_j|x^{(j-1)}}$ satisfies $\tilde{T}_{p,d_2}(C)$ for all $1 \leq j \leq n$ then $\mathbb{P}$ satisfies $\tilde{T}_{p,d_1}(C\|\Gamma(p)\|_p^2 n^{2/p-1}).$

Example

$$\begin{array}{|c|c|c|}
\hline
\Gamma_{d_1,d_2}(p) & d_1(x, y) = |x - y| & d_1(x, y) = 1_{x \neq y} \\
\hline
\hline
\Gamma_{d_1,d_2}(p) & d_2(x, y) = 1_{x \neq y} & d_2(x, y) = |x - y| & d_2(x, y) = 1_{x \neq y} \\
\hline
\hline
p=1 & \text{Rio, 2000} & \text{Djellout et al., 2004} & \text{Rio, 2000} \\
\hline
p=2 & \text{W., 2012} & \text{W., 2012} & \text{Samson, 2000} \\
\hline
\end{array}$$
Conditional weak transport inequalities for $\mathbb{P} \in \tilde{T}_{2,d_1}(C)$

Let $f_\theta = r_n(\theta) - r_n(\bar{\theta})$ and $L_j : \mathbb{R}^n \times \Theta \mapsto \mathbb{R}_+$, $1 \leq j \leq n$ such that

$$f_\theta(y) - f_\theta(x) \leq \sum_{j=1}^{n} L_j(y, \theta)d_1(x_j, y_j) \quad \forall x, y \in \mathbb{R}^n.$$ 

Let $\mu$ be a measure on $\Theta$ depending on $\hat{\theta}$ (then on $(X_t)_{1 \leq t \leq n}$). Choose $\nu$ independent of $(X_t)_{1 \leq t \leq n}$ and $Q_\theta$ such that $\nu Q_\theta = Q \mu$. Then for all $\lambda > 0$

$$\mathbb{E}_Q \left[ \mathbb{E}_\mathbb{P}[f_\theta] - f_\theta \right] \leq \frac{C \sum_{j=1}^{n} \mathbb{E}_Q[L_j(\theta)^2]}{\lambda} + \lambda \quad [\mathcal{K}(Q | \mathbb{P})].$$
Conditional weak transport inequalities for $\mathbb{P} \in \tilde{T}_{2,d_1}(C)$

Let $f_\theta = r_n(\theta) - r_n(\bar{\theta})$ and $L_j : \mathbb{R}^n \times \Theta \mapsto \mathbb{R}_+$, $1 \leq j \leq n$ such that

$$f_\theta(y) - f_\theta(x) \leq \sum_{j=1}^n L_j(y, \theta)d_1(x_j, y_j) \quad \forall x, y \in \mathbb{R}^n.$$ 

Let $\mu$ be a measure on $\Theta$ depending on $\hat{\theta}$ (then on $(X_t)_{1 \leq t \leq n}$). Choose $\nu$ independent of $(X_t)_{1 \leq t \leq n}$ and $Q_\theta$ such that $\nu Q_\theta = Q \mu$. Then for all $\lambda > 0$

$$\mathbb{E}_{Q_\theta} [\mathbb{E}_{\mathbb{P}}[f_\theta] - f_\theta] \leq \frac{C \sum_{j=1}^n \mathbb{E}_{Q_\theta}[L_j(\theta)^2]}{\lambda} + \lambda \quad [\mathcal{K}(Q_\theta|\mathbb{P})].$$
Conditional weak transport inequalities for $\mathbb{P} \in \tilde{T}_{2, d_1}(C)$

Let $f_\theta = r_n(\theta) - r_n(\bar{\theta})$ and $L_j : \mathbb{R}^n \times \Theta \mapsto \mathbb{R}_+$, $1 \leq j \leq n$ such that

$$f_\theta(y) - f_\theta(x) \leq \sum_{j=1}^n L_j(y, \theta) d_1(x_j, y_j) \quad \forall x, y \in \mathbb{R}^n.$$ 

Let $\mu$ be a measure on $\Theta$ depending on $\hat{\theta}$ (then on $(X_t)_{1 \leq t \leq n}$). Choose $\nu$ independent of $(X_t)_{1 \leq t \leq n}$ and $Q_\theta$ such that $\nu Q_\theta = Q \mu$. Then for all $\lambda > 0$

$$\mathbb{E}_\nu \mathbb{E}_{Q_\theta} \left[ \mathbb{E}_{\mathbb{P}}[f_\theta] - f_\theta \right] \leq \frac{C \sum_{j=1}^n \mathbb{E}_\nu \mathbb{E}_{Q_\theta}[L_j(\theta)^2]}{\lambda} + \lambda \mathbb{E}_\nu [\mathcal{K}(Q_\theta | \mathbb{P})].$$

From $\mathbb{E}_\nu \mathbb{E}_{Q_\theta} = \mathbb{E}_Q \mathbb{E}_\mu$, $\mathbb{E}_\nu [\mathcal{K}(Q_\theta | \mathbb{P})] = E_Q[\mathcal{K}(\mu | \nu)] + \mathcal{K}(Q | \mathbb{P})$ and $Q = \mathbb{P}_A$

w.p. $1 - \varepsilon$  

$$\mathbb{E}_\mu[R(\theta)] \leq R(\bar{\theta}) + \lambda \log(\varepsilon^{-1})$$

$$+ \mathbb{E}_{\mathbb{P}_A}[\mathbb{E}_\mu[f_\theta] + \lambda \mathcal{K}(\mu | \nu)] + \frac{C \sum_{j=1}^n \mathbb{E}_{\mathbb{P}_A} \mathbb{E}_\mu[L_j(\theta)^2]}{\lambda}.$$
Oracle inequalities for the ERM predictor

The model of predictors $\Theta$ of dimension $d(\Theta)$ is fixed.

Let $\ell(x, y) = (x - y)^2$, $\hat{\theta}_{ERM} = \min_{\theta} r_n(\theta)$ such that $f_{\hat{\theta}_{ERM}} \leq 0$ and
\[ \| \hat{X}_{\theta}^{\theta} \|_{\infty} \leq K \mathbb{E}[(\hat{X}_{\theta}^{\theta})^2] \text{ a.s. for all } \theta \in \Theta \text{ and } K > 0, \]

**Theorem (W., 2012)**

*If $\mathbb{P}$ is $\Gamma_{1,1}(2)$-weakly dependent, w.p. $1 - \varepsilon$ for all $T > 0$*

\[ R(\hat{\theta}_{ERM}) \leq R(\overline{\theta}) + C \frac{d(\Theta) + \log(\varepsilon^{-1}) - \log \mathbb{P}(r_n(\overline{\theta}) > T)}{n}. \]
Oracle inequalities for the Gibbs predictor

Now \( \Theta = \bigcup_{j=1}^{m} \Theta_j \) for models of predictors \( \Theta_j \) of dimension \( d(\Theta_j) \) and \( \bar{\theta} \in \Theta_j \).

Let \( \ell(x, y) = (x - y)^2 \), the a priori measure \( \nu \propto \sum_{j=1}^{m} 2^{-d(\Theta_j)} \pi_j \) \((\pi_j \text{ uniform})\) and \( \hat{\theta}_\lambda \) the Gibbs predictor \( \hat{\theta}_\lambda = \mathbb{E}_{\mu, \lambda}[\theta] \) where

\[
d_{\mu, \lambda} = \arg \min \mathbb{E}_{\mu}[f_\theta] + \lambda \mathcal{K}(\mu | \nu) = \frac{\exp(-\lambda^{-1} r_n(\theta))}{\mathbb{E}_\nu[\exp(-\lambda^{-1} r_n(\theta))]} d\nu.
\]

Theorem (Alquier et al., 2012)

Let \( \mathbb{P} \) be \( \Gamma_{1,1}(2) \)-weakly dependent and \( f_\theta \) be \( L \)-Lipschitz continuous. Under boundedness and w.p. \( 1 - \varepsilon \), we have for \( \eta > 0 \) well chosen

\[
R(\hat{\theta}_{\eta/n}) \leq R(\bar{\theta}) + C \frac{d(\Theta_j) \log(n/d(\Theta_j)) + \log(\varepsilon^{-1})}{n}.
\]
Prediction of the $\tau-$quantile, $\tau \in (0, 1)$

**Definition (Koenker, 2005)**

The quantile loss functions are given by

$$
\ell_{\tau}(x, y) = \begin{cases} 
\tau(x - y), & \text{if } x - y > 0, \\
-(1 - \tau)(x - y), & \text{otherwise.}
\end{cases}
$$

**Theorem (Alquier et al., 2012)**

Let $\mathbb{P}$ be $\Gamma_{1}\Gamma_{1}$-weakly dependent and $f_{\theta}$ be $L$-Lipschitz continuous. Under boundedness and w.p. $1 - \varepsilon$, we have for $\lambda^{\ast} > 0$ well chosen

$$
R(\hat{\theta}_{\text{ERM}}) \leq R(\bar{\theta}) + C \left( \sqrt{d(\Theta)} \log^{5/2}(n) + \log(\varepsilon^{-1}) \right) / \sqrt{n},
$$

$$
R(\hat{\theta}_{\lambda^{\ast}}) \leq R(\bar{\theta}) + C \left( \sqrt{d(\Theta_{j})} \log^{5/2}(n) + \log(\varepsilon^{-1}) \right) / \sqrt{n}.
$$
Application: confident intervals of prediction

Example (Forecast of the French GDP, Cornec, 2010)
Application: VaR forecast (Bale III and Solvancy II)

Example (Forecast of the VaR of a financial index, Giot and Laurent, 2004)
Conclusion and perspectives on the prediction

Conclusion

1. Use the ERM predictor when the model of predictors is satisfactory,
2. Use the Gibbs predictor for aggregating different models of predictors, penalizing a priori by the dimensions.

Perspectives

1. Extension of dimension free inequalities (Tsirelson’s ones) for convex functions of $\Gamma_{||,1}(2)$-weakly dependent processes
   $\implies$ new exponential inequalities for bounded dynamical systems.
2. Nonexact oracle inequalities for the ERM predictor in $\Gamma_{||,1}(2)$-weakly dependent processes
   $\implies$ investigate subgradient descent predictors to obtain exact oracle inequalities.
Outline

1 Estimation (III, IV, IX, XI)
   - Breaks of stationarity
   - Continuously invertible models

2 Prediction (VIII, XIII, XIV)
   - Weak transport inequalities
   - Exact oracle inequalities

3 Extremes (VII, X, XV)
   - The cluster index
   - Limit theorems for functions of Markov chains
Motivation: description of the clusters of extremes

An extreme event can be directly followed by another extreme event due to the dependence of the process and then form a cluster.

Example: squares of the log returns in the last period
Assume that $X_t = f(\Phi_t)$ for a Markov chain $(\Phi_t)$.

**Definition (Doeblin 1939)**

$(\Phi_t)$ 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 $x \in \mathbb{R}^d$.

**Regeneration cycles**

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

The excursions $(\Phi_{\tau_A(t)+1}, \ldots, \Phi_{\tau_A(t+1)})$ and thus $(X_{\tau_A(t)+1}, \ldots, X_{\tau_A(t+1)})$ are iid.
The cluster index

Splitting Markov chains

Excursions of the squares of the log returns

(DC_p) \( \mathbb{E}(|f(\Phi_1)|^p \mid \Phi_0 = y) \leq \beta |f(y)|^p + b 1_A(y), \quad \beta \in (0, 1). \)
Regularly varying processes

Definition (Basrak & Segers, 2009)

The stationary process \((X_t)\) is regularly varying of index \(\alpha > 0\) if the spectral tail process \((\Theta_t)\) satisfies, for \(k \geq 0\),

\[
\underset{w}{\mathbb{P}(\frac{|X_0|^{-1}(X_0, \ldots, X_k) \in \cdot)}{\mathbb{P}(|X_0| > x)} \rightarrow u^{-\alpha} \mathbb{P}((\Theta_0, \ldots, \Theta_k) \in \cdot)}.
\]
The regular variation index $\alpha$

Fat-tailed queues $\sim x^{-\alpha} L(x)$ with $\alpha \approx 2$
Peaks over thresholds
Realizations of the spectral tail process

By definition $\Theta_0 = 1$
Theorem (Mikosch and Wintenberger, 2012)

Under \((RV_\alpha)\) and \((DC_p)\) with \(p > (\alpha - 1)_+\) then the cluster index exists

\[
b_+ = \mathbb{E}[(\sum_{t=0}^{\infty} \Theta_t)^{\alpha}_+ - (\sum_{t=1}^{\infty} \Theta_t)^{\alpha}_+].
\]

In this case, the extremal index exists

\[
\theta_+ = \mathbb{E}[(\sup_{t \geq 0} \Theta_t)^{\alpha}_+ - (\sup_{t \geq 1} \Theta_t)^{\alpha}_+].
\]
Examples with $\Theta_0 = 1$

<table>
<thead>
<tr>
<th>Example (Asymptotic independence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta_t = 0$ for all $t &gt; 0$ then $b_+ = \theta_+ = 1$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example (AR(1): $X_t = \rho X_{t-1} + \xi_t, \forall t \in \mathbb{Z}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta_t = \rho^t$ for all $t \geq 0$ then $\theta_+ = 1 - \rho^\alpha$ and $b_+ = \theta_+/(1 - \rho)^\alpha$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example (GARCH(1,1)$^2$: $X_t^2 = \sigma_t^2 Z_t^2$, $\sigma_t^2 = \alpha_0^* + \alpha_1^* X_{t-1}^2 + \beta_1^* \sigma_{t-1}^2, \forall t \in \mathbb{Z}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta_t = (Z_t/Z_0)^2 \prod_{i=1}^t (\alpha_1^* Z_{i-1}^2 + \beta_1^*)$ for all $t \geq 0$ then $b_+$ and $\theta_+$ are explicit.</td>
</tr>
</tbody>
</table>
Representation of the clusters using the extremal and cluster indices

As. ind., observations, AR(1) GARCH(1,1)

Habilitation, November 23, 2012
Olivier Wintenberger, CEREMADE and CREST-LFA
Classical limit theorems

Let \((X_t')\) be an iid process with the same margins than \((X_t)\).

**Theorem (Meyn and Tweedie, 1993)**

Assume that \((DC_p)\) holds for \(p = 1\), \(\mathbb{E}|X_0|^2 < \infty\) and \(\mathbb{E}X_0 = 0\). Then

1. The central limit theorem holds \(n^{-1/2}S_n \xrightarrow{d} \mathcal{N}(0, \sigma^2)\)
2. The full iid cycles \(S(t) = \sum_{i=1}^{\tau_A(t+1)} f(\Phi_{\tau_A(t)+i})\) satisfies
   \[
   \mathbb{E}_A(S(1)) = 0 \quad \text{and} \quad \mathbb{E}_A[S(1)^2] = \sigma^2.
   \]

Let \((c_n)\) satisfying \(\lim_{n \to \infty} n\mathbb{P}(|X_0| > c_n) = 0\).

**Theorem (Leadbetter, 1983)**

Under \((RV_\alpha)\) and \((DC_p)\), \(\alpha, p > 0\), the extremal index \(\theta_+\) satisfies for all \(u > 0\)

\[
\mathbb{P}\left(\max_{1 \leq i \leq n} X_i > c_n\right) \sim \theta_+ \mathbb{P}\left(\max_{1 \leq i \leq n} X_i' > c_n\right), \quad n \to \infty.
\]
Let \((a_n)\) satisfying \(\lim_{n \to \infty} np(\mid X_0 \mid > a_n) = 1\). Let \(\mathbb{E} X_0 = 0\) if \(X_0\) is integrable and \(b_- = \mathbb{E}[(\sum_{t=0}^{\infty} \Theta_t)^{\alpha} - (\sum_{t=1}^{\infty} \Theta_t)^{\alpha}]\).

**Theorem (Bartkiewicz et al., 2011)**

Assume \((RV_\alpha)\) with \(0 < \alpha < 2\), \(\alpha \neq 1\) and \((DC_p)\) with \(p > (\alpha - 1)_+\) then

\[ a_n^{-1} S_n \xrightarrow{d} \xi_\alpha \]

is satisfied for a centered \(\alpha\)-stable r.v. \(\xi_\alpha\) with cf \(\psi_\alpha(x) = \exp(-|x|^{\alpha} \chi_\alpha)\), where

\[ \chi_\alpha = \frac{\Gamma(2 - \alpha)}{1 - \alpha} ((b_+ + b_-) \cos(\pi \alpha/2) - i \text{sign}(x)(b_+ - b_-) \sin(\pi \alpha/2)). \]
Theorem (Mikosch and W., 2012)

Under the assumptions of the previous Theorem, if $b_+ \neq 0$ then

$$P_A(S(1) > x) \sim b_+ E_A(\tau_A) P(|X_0| > x) \sim b_+ P\left( \sum_{t=1}^{\tau_A} |X'_t| > x \right), \quad x \to \infty.$$ 

Extended to $\alpha > 2$ under $b_+ \neq 0$ and $(\text{DC}_p)$ for all $p < \alpha$. 
Let $b_n = n^{\delta + 1/\alpha} \wedge 1/2$ for any $\delta > 0$ and $S'_n = \sum_{i=1}^n X'_i$.

**Theorem (Mikosch and W., 2012)**

Assume $(\text{RV}_\alpha)$ with $\Theta_0 = 1$.

- If $0 < \alpha < 1$ and $(\text{DC}_p)$ with $p > 0$ then
  $$
  \lim_{n \to \infty} \sup_{x \geq b_n} \left| \frac{\mathbb{P}(S_n > x)}{\mathbb{P}(S'_n > x)} - b_+ \right| = 0,
  $$

- If $1 < \alpha < 2$ and $(\text{DC}_p)$ with $p > \alpha - 1$ or $\alpha > 2$ and $(\text{DC}_p)$ for all $p < \alpha$
  $$
  \lim_{n \to \infty} \sup_{b_n \leq x \leq c_n} \left| \frac{\mathbb{P}(S_n > x)}{\mathbb{P}(S'_n > x)} - b_+ \right| = 0
  $$

with $\mathbb{P}(\tau_A > n) = o(\mathbb{P}(S'_n > c_n))$. 

Habilitation, November 23, 2012
Olivier Wintenberger, CEREMADE and CREST-LFA
Conclusion and perspectives on the extremes

Conclusion

1. The extremal and cluster indices describe the clusters of extreme values,
2. $\alpha$-stable central limit theorems and large deviations provides theoretical justifications of the introduction of a new index.

Perspectives

1. We need to consider Markovian processes and their regenerative structures $\implies$ use also regenerative structures to validate the estimation procedure in theory.
2. Model the extremal dependence in view of the observed clusters represented with the extremal and cluster indices $\implies$ introduce new models with extremal behaviors similar than the observed ones.
Thank you for your attention!


