Tail and memory: modeling and statistical inference

Rafał Kulik

University of Ulm, 12 December 2011

Plan of talk

1. $Typical$ financial data. Which models can capture such behaviour?

2. LMSV: Long Memory Stochastic Volatility

3. Partial sums for LMSV models

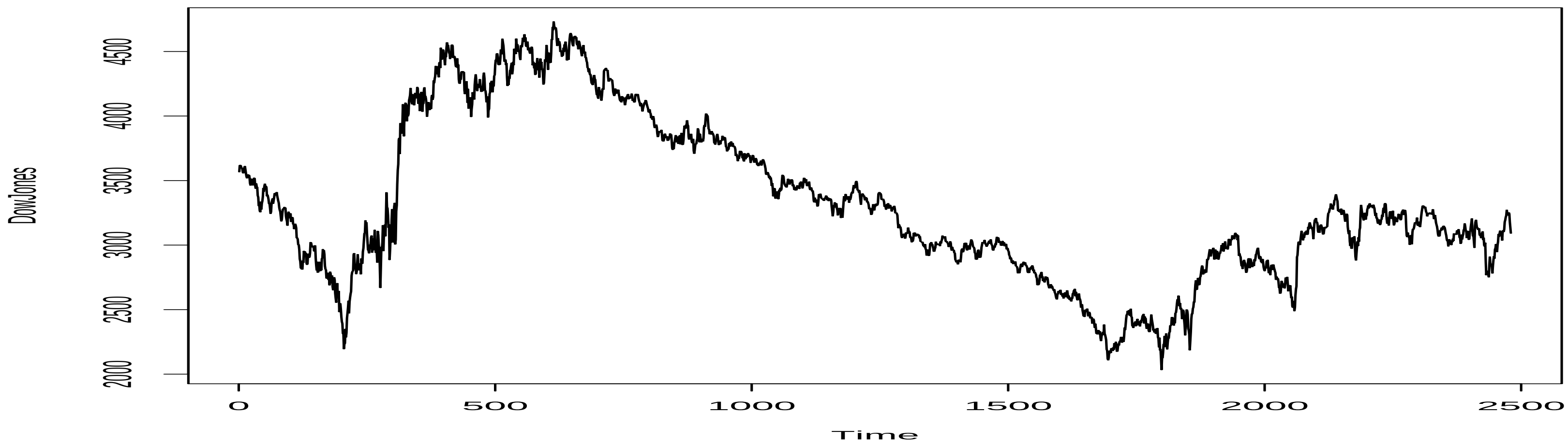
4. Estimation of memory

5. Estimation of tails (Kulik and Soulier 2011, 2012)

6. Some open problems

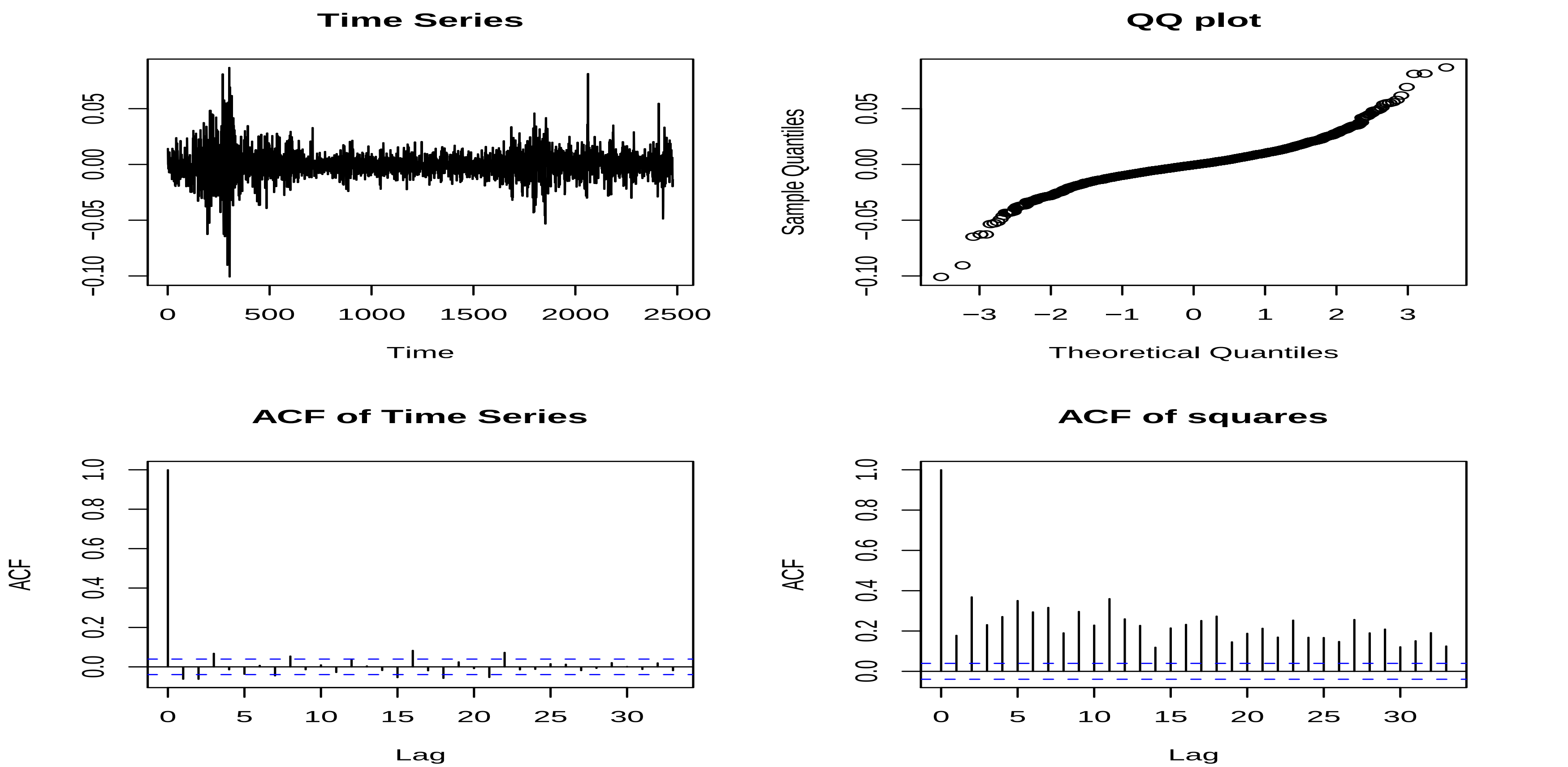
Dow Jones data

The following data set describes Dow Jones Composite Average from 1 Jan 2000 to 1 Jan 2010. (source Yahoo Finance).



Time series S_j, $j = 1, \ldots, n$, is not stationary. Apply transformation $Y_j = \log(S_j/S_{j-1})$, $j = 2, \ldots, n$.

Log returns


Time Series
0.05
0.00
−0.05
−0.10
0
500
1000
1500
2000
2500
Time
QQ plot
Sample Quantiles
−3
−2
−1
0
1
2
3
Theoretical Quantiles
ACF of Time Series
ACF
1.0
0.8
0.6
0.4
0.2
0.0
5
10
15
20
25
30
Lag
ACF of squares

Log returns - ctd.

- Y_j, $j = 2, \ldots, n$, are uncorrelated.
- Y_j^2, $j = 2, \ldots, n$, are correlated, with long memory.
- Non-normal behaviour. Possible heavy tails.

Which model can capture such behaviour?

MA(∞) models

Let ε_j, $j \in \mathbb{Z}$, be a sequence of iid centered random variables, with finite variance. Define

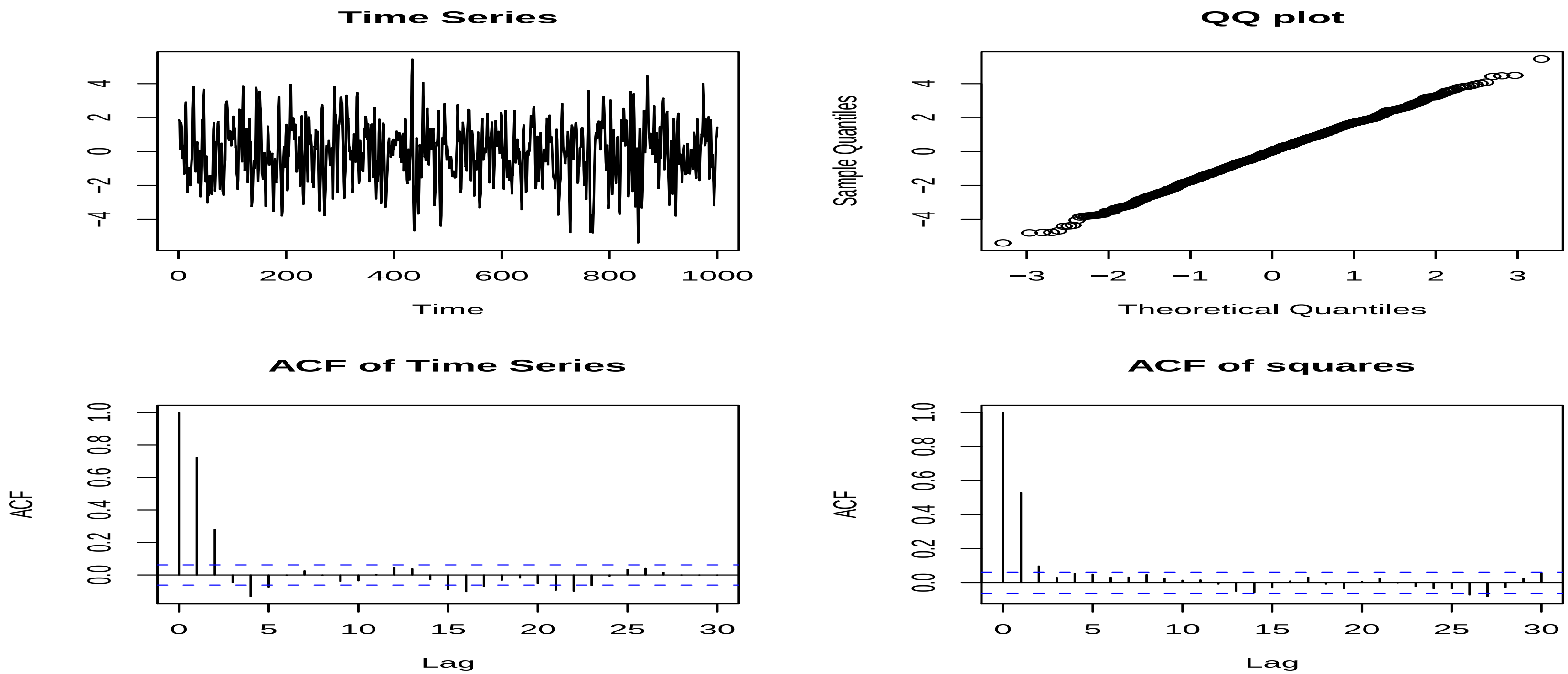
$$Y_j = \sum_{k=1}^{\infty} c_k \varepsilon_{j-k} \ ,$$

where $\sum_{k=1}^{\infty} c_k^2 < \infty$. This sequence is correlated since $\mathrm{Cov}(Y_0, Y_j) = \sum_{k=1}^{\infty} c_k c_{k+j}$. If $\sum_{k=1}^{\infty} |c_k| < \infty$, then covariances are summable and sequence is *short range dependent*.

In particular, ARMA(p, q) models,

$$Y_j = \beta_0 + \sum_{r=1}^{p} \beta_r Y_{j-r} + \sum_{s=1}^{q} \beta'_s \varepsilon_{j-s} \ .$$

can be represented in terms of MA(∞). ARMA$(0, q)$ sequence is q-dependent, whereas covariances in ARMA$(p, 0)$ decay exponentially fast.



- Y_j, $j = 2, \ldots, n$, are uncorrelated. NO
- Y_j^2, $j = 2, \ldots, n$, are correlated. YES
- Non-normal behaviour. Possible heavy tails. YES

GARCH(p, q) models

Consider a stationary solution of

$$Y_j = \sigma_j \varepsilon_j \; , \sigma_j^2 = \beta_0 = \sum_{r=1}^{p} \beta_r Y_{j-r}^2 + \sum_{s=1}^{q} \beta_s' \sigma_{j-s}^2 \; ,$$

where $\beta_0 > 0$. Let $\mathcal{F}_j$ - sigma field generated by $\varepsilon_j, \varepsilon_{j-1}, \ldots$.

- $\mathrm{E}(Y_0 Y_j) = \mathrm{E}\left[\mathrm{E}(Y_0 Y_j | \mathcal{F}_{j-1})\right] = \mathrm{E}\left[Y_0 \sigma_j \mathrm{E}(\varepsilon_j | \mathcal{F}_{j-1})\right] = 0$.
- $(Y_j, \mathcal{F}_j)$ is a martingale.

- Y_j, $j = 2, \ldots, n$, are uncorrelated. YES
- Y_j^2, $j = 2, \ldots, n$, are correlated. YES, but long memory not possible.
- Non-normal behaviour. Possible heavy tails. YES

Long memory Gaussian sequence

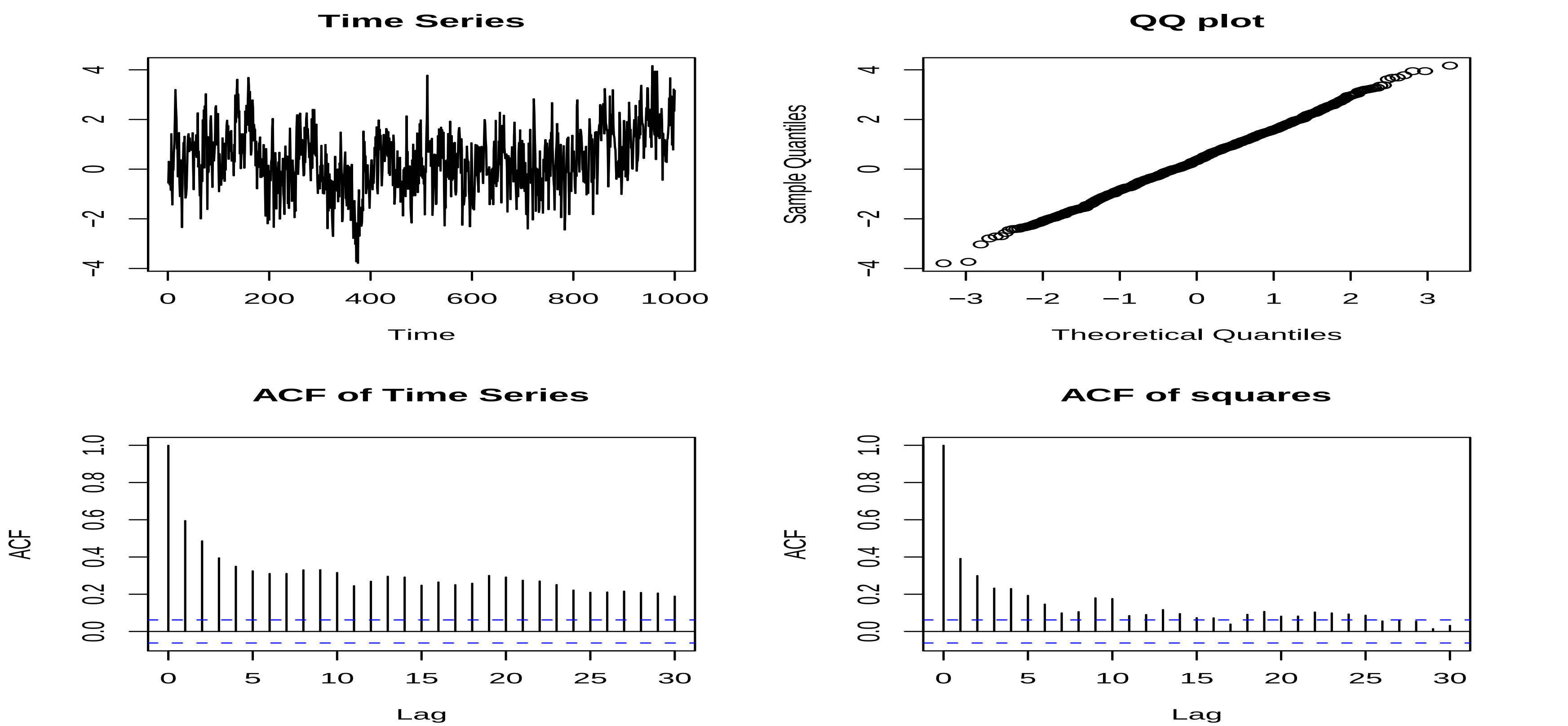
Assume that X_j, $j \geq 1$, is a stationary Gaussian process with covariance

$$\mathrm{cov}(X_0, X_j) = \rho_j = j^{2H-2} L_0(j) = j^{2d-1} L_0(j),$$

where $H \in (1/2, 1)$ is the Hurst exponent, $d \in (0, 1/2)$ is the fractional difference parameter and $L_0(\cdot)$ is a slowly varying function. Note that covariances are not summable.
1. X_j, $j = 2, \ldots, n$, are uncorrelated. NO
2. X_j^2, $j = 2, \ldots, n$, are correlated. YES, long memory possible.
3. Non-normal behaviour. Possible heavy tails. NO
Note: such processes have the following representation

$$X_j = \sum_{k=1}^{\infty} c_k \varepsilon_{j-k},$$

where $c_k \sim L_1(j) j^{d-1}$ and ε_j are i.i.d. Gaussian. If Gaussian is replaced with another distribution $\Rightarrow$ linear long memory models .



References: Beran (1994); Doukhan, Oppenheim, Taqqu ed. (2003)

Long Memory Stochastic Volatility

Assume that $Z, Z_j,\ j \ge 1$, is a sequence of i.i.d. random variables and that $X_j,\ j \ge 1$, is a stationary LRD Gaussian process written as $\sum_{k=1}^{\infty} c_k \varepsilon_{j-k}$. We assume that Z_j, $j \ge 1$, and X_j, $j \ge 1$, are independent. The stochastic volatility process is defined as

$$Y_j = \sigma(X_j) Z_j,$$

where $\sigma(\cdot)$ is a nonnegative function. We note, in particular, that if $\mathrm{E}[Z_1^2] < \infty$ and $\mathrm{E}[Z_1] = 0$, then Y_j, $j \ge 1$, are uncorrelated.

- Y_j, $j = 2, \ldots, n$, are uncorrelated. YES. In particular, Y_j is a martingale.

- Y_j^2, $j = 2, \ldots, n$, are correlated. YES.

$$\mathrm{Cov}(Y_0, Y_j) = \mathrm{E}[Z_1^2]\mathrm{Cov}(\sigma(X_0), \sigma(X_j)) \; .$$

- Non-normal behaviour. Possible heavy tails. YES

  We will assume that for some $\alpha \in (0, \infty)$,

$$\bar{F}_Z(z) = \mathrm{P}(Z > x) = x^{-\alpha}\ell(x) \; , \tag{1}$$

  where ℓ is a slowly varying function. Having (1) and $\mathrm{E}[\sigma^{\alpha+\epsilon}(X_1)] < \infty$,

$$\bar{F}(x) = \mathrm{P}(Y_1 > x) = \mathrm{P}(\sigma(X_1)Z_1 > x) \sim \mathrm{E}[\sigma^{\alpha}(X_1)]\mathrm{P}(Z_1 > x) \; , \quad \text{as } x \to \infty.$$

Partial Sums for LMSV models

References: Davis and Mikosch (2001), Kulik and Soulier (2012).

Assume that $\mathrm{E}[Z_1] = 0$. Let $\mathcal{H}_j = \mathcal{F}_j \vee \mathcal{Z}_j$, where $\mathcal{F}_j$ and $\mathcal{Z}_j$ are sigma fields generated by ε_j and Z_j, respectively.

- If $\mathrm{E}[Z_1^2] < \infty$, then $n^{-1/2} \sum_{j=1}^n Y_j \xrightarrow{\mathrm{d}} \mathcal{N}(0, \omega^2)$ by of martingale CLT.
- If $\mathrm{E}[Z_1^2] = \infty$, then $n^{-1/\alpha} \sum_{t=1}^n Y_t$ converges to a stable random variable.

Furthermore, assume that $\mathrm{E}[Z_j^2] < \infty$. Let $\sigma_j = \sigma(X_j)$.

Then

$$\begin{aligned}\sum_{j=1}^{n}\left(Y_j^2-\mathrm{E}[Y_j^2]\right) &= \sum_{j=1}^{n}\left(Y_j^2-\mathrm{E}[Y_j^2|\mathcal{H}_{j-1}]\right)+\sum_{j=1}^{n}\left(\mathrm{E}[Y_j^2|\mathcal{H}_{j-1}]-\mathrm{E}[Y_j^2]\right)\\ &= \mathrm{martingale}+\mathrm{E}[Z_1^2]\sum_{j=1}^{n}\left(\sigma_j^2-\mathrm{E}[\sigma_j^2]\right)\ .\end{aligned}$$

- If $\mathrm{E}[Z_1^4]<\infty$, then the first part converges to a normal law with rate $n^{-1/2}$, on account of martingale CLT.

- If $\mathrm{E}[Z_1^4]=\infty$, the first part will converge to a stable law, and the limit will be the same as if the summands were iid.

- For the second part, possible long memory.

Estimation of dependence parameter

- Parametric methods: time domain MLE, spectral domain MLE (so called Whittle estimator) - Dahlhaus (1989), Fox and Taqqu (1989) for long memory Gaussian processes, Kokoszka and Taqqu (1997) for long memory infinite variance linear processes. Gaussian and stable limits respectively. Not feasible for stochastic volatility models.

- Semiparametric methods: logperiodogram regression and local Whittle estimation: Robinson (1995) for long memory Gaussian processes, Moulines, Hurvich, Soulier (2003) for stochastic volatility models. Nothing known for infinite variance linear processes

- Wavelet methods: Moulines, Roueff, Taqqu (2008-2011). Unknown for stochastic volatility models.

Tail empirical process

Define (see Rootzén (2009), Drees (2000)),

$$\tilde{T}_n(s) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{n} 1_{\{Y_j > u_n + u_n s\}} \ ,$$

and

$$e_n(s) = \tilde{T}_n(s) - T_n(s) \ , \ s \in [0, \infty) \ , \tag{2}$$

where

$$T_n(x) := \frac{\bar{F}(u_n + u_n x)}{\bar{F}(u_n)} \to T(x) = (1+x)^{-\alpha} \, , \qquad x \geq 0, \ n \geq 1. \tag{3}$$

Tail empirical process - limiting behaviour

Theorem 1. *Let q be the Hermite rank of $G(x) = \sigma^{\alpha}(x)$ with $q(1-H) \neq 1/2$ + some technical assumptions.*

(i) *If $n\bar{F}(u_n)\rho_n^q \to 0$ as $n \to \infty$, then $\sqrt{n\bar{F}(u_n)}\, e_n$ converges weakly in $D([0,\infty))$. to the Gaussian process $W \circ T$, where W is the standard Brownian motion.*

(ii) *If $n\bar{F}(u_n)\rho_n^q \to \infty$ as $n \to \infty$ then $\rho_n^{-q/2} e_n(s)$ converges weakly in $D([0,\infty))$ to the process $\mathrm{const}.T(s)$.*

Comments

- The meaning of the above result is that for u_n big, long memory does not play any role. However, if u_n is small, long memory comes into play and the limit is degenerate. Furthermore, *small* and *big* depends on the relative behaviour of the tail of Y_1 and the memory parameter. Note that the condition $n\bar{F}(u_n)\rho_n^q \to \infty$ implies that $1 - 2q(1-H) > 0$, in which case the partial sums of the subordinate process $\{G(X_i)\}$ weakly converge to the Hermite process of order q. The two cases will be referred to as the limits *in the i.i.d. zone* and *in the LRD zone*.

- One can replace T_n with T in the definition of the tail empirical process, provided a second order condition is fulfilled.

Tail empirical process with random levels

Let $U(t) = F^{\leftarrow}(1 - 1/t)$, where $F^{\leftarrow}$ is the left-continuous inverse of F. Define $u_n = U(n/k)$. If F is continuous, then $n\bar{F}(u_n) = k$.

Define

$$\hat{T}_n(s) = \frac{1}{k}\sum_{j=1}^{n} 1_{\{Y_j > Y_{n-k:n}(1+s)\}} \ .$$

Here we consider the *practical* process

$$\hat{e}_n^*(s) = \hat{T}_n(s) - T(s), \qquad s \in [0, \infty) \ .$$

Tail empirical process with random levels - result

Theorem 2. *Under the conditions of Theorem 1, together with a second order assumptions, we have: $\sqrt{k}\hat{e}_n^*$ converges weakly in $D([0,\infty))$ to $B \circ T$, where B is the Brownian bridge (regardless of the behaviour of $k\rho_n^q$).*

The behaviour described in Theorem 2 is quite unexpected, since the process with $estimated$ levels $Y_{n-k:n}$ has a faster rate of convergence than the one with the deterministic levels u_n. A similar phenomenon was observed in the context of LRD based empirical processes with estimated parameters.

Applications to Hill estimator

A natural application of the asymptotic results for tail empirical process $\hat{e}_n^*$ is the asymptotic normality of the Hill estimator of the extreme value index γ defined by

$$\hat{\gamma}_n = \frac{1}{k}\sum_{i=1}^{k} \log\left(\frac{Y_{n-i+1:n}}{Y_{n-k:n}}\right) = \int_0^\infty \frac{\hat{T}_n(s)}{1+s}\,\mathrm{d}s \; .$$

Since $\gamma = \int_0^\infty (1+s)^{-1} T(s)\,\mathrm{d}s$, we have

$$\hat{\gamma}_n - \gamma = \int_0^\infty \frac{\hat{e}_n^*(s)}{1+s}\,\mathrm{d}s \; .$$

Corollary 3. *Under the assumptions of Theorem 2, $\sqrt{k}(\hat{\gamma}_n - \gamma)$ converges weakly to the centered Gaussian distribution with variance γ^2.*

Nice Hill plot

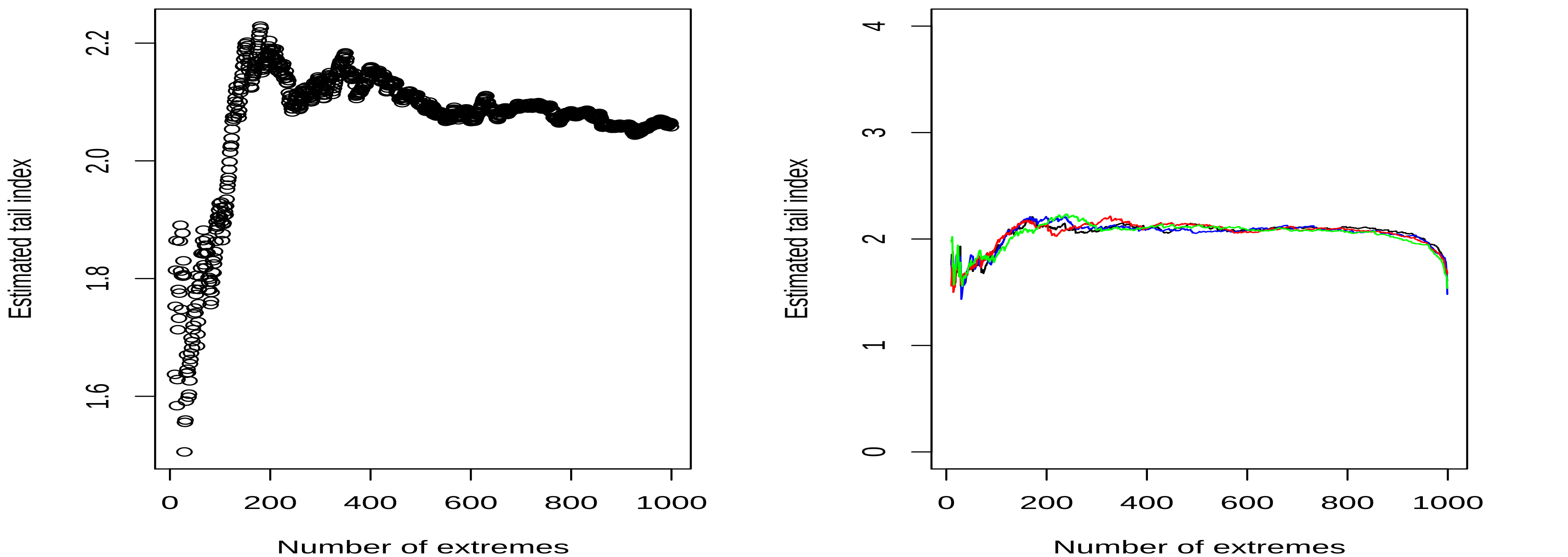


Figure 1: Hill estimator: $\alpha = 2$ and Pareto iid (left panel), $\sigma = 0.05$ (right panel)

Terrible Hill plot

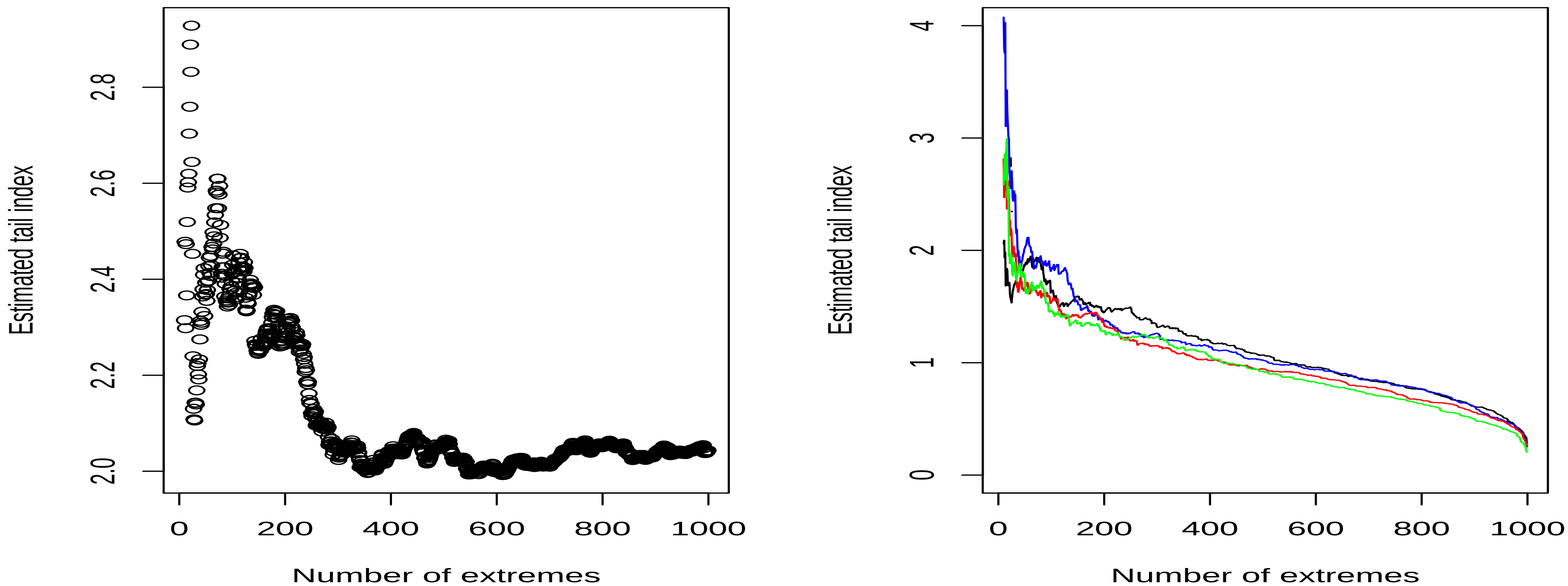


Figure 2: Hill estimator: $\alpha = 2$ and Pareto iid (left panel), $\sigma = 1$ (right panel)

Comments

The model considered is $\sigma(X_i) = \exp(\sigma X_i)$, where X_i, $i \geq 1$, are LRD standard normal. We may observe that for small volatility parameter σ there is not too much difference between i.i.d. Pareto-based Hill estimator and those for stochastic volatility models. However, if σ becomes bigger, estimation is completely inappropriate and is as bad for very strong memory as for i.i.d. case.

Open problems

- Estimation of dependence parameter: long memory linear processes with infinite variance, wavelets methods for infinite variance and stochastic volatility.

- Tail index estimation for long memory linear processes with infinite variance. (Beran 2012 considers M-estimation and gets stable limits).

- Traffic models.