Testing for a change in the tail parameter of regularly varying time series with long memory

Davide Giraudo
joint work with Annika Betken and Rafał Kulik

Ruhr-Universität Bochum
Research supported by the Collaborative Research Center 823 "Statistical modelling of nonlinear dynamic processes"

London, December 14th 2019
CFE-CMStatistics 2019

Hypothesis

Given observations $X_1, \ldots, X_n$, assume that $\mathbb{P}\left(X_j > x\right) = x^{-\alpha_j} L(x)$, $j = 1, \ldots, n$, where $\alpha_j > 0$ and L is a slowly varying function $\left(L\left(cx\right)/L\left(x\right) \to 1\right.$ for all positive c as $\left.x \to +\infty\right)$. The number α_j is called the tail index of X_j.

We consider the testing problem (H, A):

$$H : \alpha_1 = \cdots = \alpha_n$$

against

$$A : \alpha_1 = \cdots = \alpha_k \neq \alpha_{k+1} = \cdots = \alpha_n \quad \text{for some} \quad k \in \{1, \ldots, n-1\}.$$

Estimation of the tail index of a random variable

Let X be a random variable with distribution function F having regularly varying tail with index $-\alpha$, $\alpha > 0$, that is $\bar{F}(x) := \mathbb{P}(X > x) = x^{-\alpha} L(x)$, where L is slowly varying at infinity. It can be shown that

$$\lim_{u \to \infty} \mathbb{E}\left[\log\left(\frac{X}{u}\right) \mid X > u\right] = \lim_{u \to \infty} \frac{\mathbb{E}\left[\log\left(\frac{X}{u}\right) \mathbf{1}\{X > u\}\right]}{\mathbb{P}(X > u)} = \frac{1}{\alpha}.$$

Thus, the tail index of a stationary time series X_j, $j \in \mathbb{N}$, with marginal distribution F having regularly varying tail with index $-\alpha$, $\alpha > 0$, can be estimated using the estimator $\hat{\gamma}$ defined by

$$\hat{\gamma} = \frac{1}{\sum_{j=1}^{n} \mathbf{1}\{X_j > u_n\}} \sum_{j=1}^{n} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\{X_j > u_n\} ,$$

where u_n, $n \in \mathbb{N}$, is a sequence such that $u_n \to \infty$ and $n\bar{F}(u_n) \to \infty$.

Test statistic

Let

$$\hat{\gamma}_{n,k} = \frac{1}{\sum_{j=1}^{k} \mathbf{1}\left\{X_j > u_n\right\}} \sum_{j=1}^{k} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\left\{X_j > u_n\right\}$$

and

$$\Gamma_{k,n} = \frac{k}{n} \left| \frac{\hat{\gamma}_{n,k}}{\hat{\gamma}_{n,n}} - 1 \right| .$$

Since the location of the change of tail parameter k is unknown, we define the test-statistic

$$\Gamma_n := \max_{1 \leqslant k \leqslant n-1} \Gamma_{k,n}.$$

Empirical process

In order to find the limiting distribution of the test statistics, similar to Kulik and Soulier (2011) we consider the two-parameter tail empirical process (TEP)

$$e_n(s,t) = \left\{ \tilde{T}_n(s,t) - T(s,t) \right\}, \quad s \geqslant 1, t \in [0,1],$$

where

$$\tilde{T}_n(s,t) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \mathbf{1}\left\{X_j > u_n s\right\}$$

and

$$T(s,t) = ts^{-\alpha}.$$

Then

$$\int_1^{+\infty} \frac{1}{s} \tilde{T}_n(s,t) \mathrm{d}s = \sum_{j=1}^{n} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\left\{X_j > u_n\right\}.$$

Stochastic volatility model

We will consider the case of the stochastic volatility model:

$$X_j = \sigma\left(Y_j\right)\varepsilon_j,$$

where

- $\left(\varepsilon_j\right)_{j\in\mathbb{Z}}$ is an i.i.d. sequence of random variables with $\mathbb{E}\left[\varepsilon_1\right] = 0$;
- σ is a non-negative measurable function;
- Y_j, $j \geqslant 1$, is a stationary, long-range dependent Gaussian process, that is,

$$Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}\ , \quad \sum_{k=1}^{\infty} c_k^2 = 1\ ,$$

for i.i.d. Gaussian random variables η_j, $j \in \mathbb{Z}$, with $\mathbb{E}\left[\eta_1\right] = 0$, $\operatorname{var}\eta_1 = 1$, $\left(\left(\varepsilon_j, \eta_j\right)\right)_{j\in\mathbb{Z}}$ is independent and

$$\gamma_Y(k) := \operatorname{cov}\left(Y_j, Y_{j+k}\right) = \sum_{\ell\geqslant 0} c_\ell c_{\ell+k} = k^{-D} L_\gamma\left(k\right),$$

where $D \in (0,1)$ and L_γ slowly varying at ∞.

Covariance of the process $(\sigma(Y_j))_{j\geqslant 1}$

Let φ be the density of the standard normal distribution. Every $G \in L^2(\mathbb{R}, \varphi(x)dx)$ has an expansion in Hermite polynomials, i.e. for $G \in L^2(\mathbb{R}, \varphi(x)dx)$ and X standard normally distributed, we have

$$G(X) = \sum_{r=0}^{\infty} \frac{J_r(G)}{r!} H_r(X),$$

where the so-called *Hermite coefficient* $J_r(G)$ is given by

$$J_r(G) := \langle G, H_r \rangle_{L^2} = \mathbb{E}\left[G(X)H_r(X)\right],$$

and H_r is the r-th Hermite polynomial.

Let $m := \min\left\{k \geqslant 1 : J_k(G) \neq 0\right\}$. be the Hermite rank of G. Then

$$d_n^2 := \operatorname{var}\left(\sum_{j=1}^{n} H_m(Y_j)\right) \sim c_m n^{2-mD} L^m(n), \ c_m = \frac{2m!}{(1-Dm)(2-Dm)}.$$

Type of assumptions

(TA.1)

$$\mathbb{P}\left(X_1 > x\right) = cx^{-\alpha} \exp\left(\int_1^x \frac{\eta^*\left(u\right)}{u} du\right),$$

where $\eta^*\left(u\right) = u^{-\rho} L_{\eta^*}(u)$, $\rho > 0$, L_{η^*} slowly varying;

(TA.2) Assumptions on the moments of $\sigma\left(Y_1\right)$ and $1/\sigma\left(Y_1\right)$.

(TA.3) $\eta^*(u_n) = o\left(\frac{d_n}{n} + \frac{1}{\sqrt{n\bar{F}(u_n)}}\right)$.

Example

Assume that $\eta^*(x) = x^{-\alpha\beta}$ for some $\beta > 0$; then, for $x \to \infty$, $\mathbb{P}\left(X_1 > x\right) = C\left(x^{-\alpha} + \mathcal{O}(x^{-\alpha(\beta+1)})\right)$. Taking σ such that $0 < c \leqslant \sigma\left(x\right) \leqslant C$, $x \in \mathbb{R}$, the assumptions (TA.1-2) are satisfied.

Limit of the empirical process

Theorem (Betken, G., Kulik (2019+))

Assume that the technical assumptions hold. Let

$$e_n(s,t) := \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{[nt]} \mathbf{1}\{X_j > u_n s\} - ts^{-\alpha}.$$

- *If* $\frac{n}{d_n} = o\left(\sqrt{n\bar{F}(u_n)}\right)$,

$$\frac{n}{d_n} e_n(s,t) \Rightarrow \frac{s^{-\alpha}}{\mathbb{E}\left[\sigma^{\alpha}(Y_1)\right]} \frac{J_q(\Psi)}{q!} Z_q(t),$$

  where $\Rightarrow$ *denotes weak convergence in* $D([1,\infty] \times [0,1])$, $\Psi(y) = \sigma^{\alpha}(y)$, q *is the Hermite rank of* Ψ.

- *If* $\sqrt{n\bar{F}(u_n)} = o\left(\frac{n}{d_n}\right)$,

$$\sqrt{n\bar{F}(u_n)}\, e_n(s,t) \Rightarrow B_{s^{-\alpha},t}$$

  in $D([1,\infty] \times [0,1])$, *where* B *denotes a standard Brownian sheet.*

Explanation of the two cases

We consider the following decomposition:

$$\begin{aligned} e_n(s,t) &= \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathbf{1}\left\{X_j > u_n s\right\} - \mathbb{E}\left[\mathbf{1}\left\{X_j > u_n s\right\} \mid \mathcal{F}_{j-1}\right]\right) \\ &\quad + \frac{1}{n\bar{F}_Z(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathbb{E}\left[\mathbf{1}\left\{X_j > u_n s\right\} \mid \mathcal{F}_{j-1}\right) - \bar{F}(u_n s)\right] \\ &= M_n(s,t) + R_n(s,t), \end{aligned}$$

where

$$\mathcal{F}_j := \sigma\left(\varepsilon_k, \eta_k, k \in \mathbb{Z}, k \leqslant j\right).$$

We call M_n the *martingale part*, while we refer to R_n as the *long memory part*.

If $\frac{n}{d_n} = o\left(\sqrt{n\bar{F}(u_n)}\right)$, the martingale part is negligible.

Convergence of the tail estimator

$$\hat{\gamma}_n(t) = \frac{\sum_{j=1}^{[nt]} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\{X_j > u_n\}}{\sum_{j=1}^{\lfloor nt \rfloor} \mathbf{1}\{X_j > u_n\}}.$$

Corollary (Betken, G., Kulik (2019+))

Under the technical assumptions,

- *if* $\frac{n}{d_n} = o\left(\sqrt{n\bar{F}(u_n)}\right)$, *then*

$$\frac{n}{d_n} t \left(\hat{\gamma}_n(t) - \gamma\right) \to 0 \textit{ in probability in } D[0,1];$$

- *if* $\sqrt{n\bar{F}(u_n)} = o\left(\frac{n}{d_n}\right)$, *then*

$$\sqrt{n\bar{F}(u_n)}t\left(\hat{\gamma}_n(t) - \alpha^{-1}\right) \Rightarrow \frac{1}{\alpha} B \textit{ in distribution in } D[0,1],$$

where B *is a standard Brownian motion.*

Convergence of the test statistic

Corollary (Betken, G., Kulik (2019+))

Under the technical assumptions,

- *if* $\frac{n}{d_n} = o\left(\sqrt{n\bar{F}(u_n)}\right)$, *then*

$$\frac{n}{d_n} \sup_{t\in[0,1]} t \left| \frac{\hat{\gamma}_n(t)}{\hat{\gamma}_n} - 1 \right| \to 0 \text{ in probability};$$

- *if* $\sqrt{n\bar{F}(u_n)} = o\left(\frac{n}{d_n}\right)$, *then*

$$\sqrt{n\bar{F}(u_n)} \sup_{t\in[0,1]} t \left| \frac{\hat{\gamma}_n(t)}{\hat{\gamma}_n} - 1 \right| \overset{\mathcal{D}}{\Longrightarrow} \frac{1}{\alpha} \sup_{t\in[0,1]} |B(t) - tB(1)|$$

  where B *is a standard Brownian motion.*

Conclusion

What we have done:

- Convergence of the empirical process.
- Convergence of the tail estimator and test statistic when the martingale part dominates.

Remaining questions:

- Find the good normalisation for the tail estimator when the long memory part dominates.
- Treat the case where n/d_n and $\sqrt{n\bar{F}\left(u_n\right)}$ are equivalent.

References

A. Betken, D. Giraudo, R. Kulik (2019+). Testing for change in the tail parameter for regularly varying time series with long memory via Hill statistics. In preparation

A. Betken, R. Kulik (2019).Testing for change in stochastic volatility with long range dependence. *Journal of Time Series Analysis*, DOI: 10.1111/jtsa.12449.

V. Pipiras, M. S. Taqqu (2017). Long-Range Dependence and Self-Similarity. Vol. 45. *Cambridge university press*.

R. Kulik, P. Soulier (2011). The tail empirical process for long memory stochastic volatility sequences. *Stochastic Processes and their Applications*, 121:109 - 134.

A. C. Harvey (2002). Long memory in stochastic volatility. *In: Forecasting Volatility in the Financial Markets*, 307-320.

F. Breidt, N. Crato, and P. De Lima (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83:325 - 348.

M. S. Taqqu (1979). Convergence of integrated processes of arbitrary Hermite rank. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50:55 - 83.