Estimation of scaling parameter for continuous processes

Mark Podolskij, University of Heidelberg

joint work with O.E. Barndorff-Nielsen and J.M. Corcuera

Le Mans, 21-24 March

Gaussian-like processes

- Our basic model is giving by the integral

$$X_t = X_0 + \int_0^t \sigma_s dG_s, \qquad t \geq 0,$$

where σ is a *volatility* process and $(G_s)_{s\geq 0}$ is a Gaussian process with *centered* and *stationary* increments.

Gaussian-like processes

- Our basic model is giving by the integral

$$X_t = X_0 + \int_0^t \sigma_s dG_s, \qquad t \geq 0,$$

where σ is a *volatility* process and $(G_s)_{s\geq 0}$ is a Gaussian process with *centered* and *stationary* increments.

- Under Hölder continuity conditions on σ and G the above integral is well-defined in the Riemann-Stieltjes sense.

Gaussian-like processes

- Our basic model is giving by the integral

$$X_t = X_0 + \int_0^t \sigma_s dG_s, \qquad t \geq 0,$$

where σ is a *volatility* process and $(G_s)_{s\geq 0}$ is a Gaussian process with *centered* and *stationary* increments.

- Under Hölder continuity conditions on σ and G the above integral is well-defined in the Riemann-Stieltjes sense.

- The stochastic process X is assumed to be observed at time points $t_i = i\Delta_n$, $i = 0, \ldots, [t/\Delta_n]$ with $\Delta_n \to 0$.

The scaling parameter

- Let $(R_t)_{t\geq}$ denote the variogram of the Gaussian driver G, i.e.

$$R_t = \mathbb{E}(G_{s+t} - G_s)^2.$$

The scaling parameter

- Let $(R_t)_{t\geq}$ denote the variogram of the Gaussian driver G, i.e.

$$R_t = \mathbb{E}(G_{s+t} - G_s)^2.$$

- We assume that R has a polynomial behaviour in the neighborhood of 0:

$$R_t = \text{const} \cdot t^\alpha + O(t^{\alpha+\alpha'}) \qquad \text{as } t \to 0$$

for some $\alpha \in (0, 2)$ and $\alpha' > 0$.

The scaling parameter

- Let $(R_t)_{t\geq}$ denote the variogram of the Gaussian driver G, i.e.

$$R_t = \mathbb{E}(G_{s+t} - G_s)^2.$$

- We assume that R has a polynomial behaviour in the neighborhood of 0:

$$R_t = \text{const} \cdot t^{\alpha} + O(t^{\alpha+\alpha'}) \qquad \text{as } t \to 0$$

  for some $\alpha \in (0,2)$ and $\alpha' > 0$.

- Our aim is to estimate the scaling parameter

$$\alpha \in (0,2)$$

  from high frequency data $X_{i\Delta_n}$.

Estimation

- Our estimation procedure relies on the power variation statistic

$$V(X,p)_t^n = \Delta_n \tau_{\Delta_n}^{-p} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p, \qquad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n},$$

with $\tau_{\Delta_n}^2 = R_{\Delta_n}$.

Estimation

- Our estimation procedure relies on the power variation statistic

$$V(X,p)_t^n = \Delta_n \tau_{\Delta_n}^{-p} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p, \qquad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n},$$

  with $\tau_{\Delta_n}^2 = R_{\Delta_n}$.

- In contrast to the semimartingale framework the statistic $V(X,p)_t^n$ is not feasible as the normalizing constant τ_{Δ_n} is unknown.

Estimation

- Our estimation procedure relies on the power variation statistic

$$V(X,p)^n_t = \Delta_n \tau_{\Delta_n}^{-p} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p, \qquad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n},$$

with $\tau^2_{\Delta_n} = R_{\Delta_n}$.

- In contrast to the semimartingale framework the statistic $V(X,p)^n_t$ is not feasible as the normalizing constant τ_{Δ_n} is unknown.

- Recall the identity

$$\tau^2_{\Delta_n} = \text{const} \cdot \Delta_n^{\alpha} + O(\Delta_n^{\alpha+\alpha'}),$$

where $\alpha \in (0,2)$ is the parameter of our interest.

Law of large numbers

- **Theorem:** Under certain regularity conditions on the variogram R we obtain the convergence

$$V(X,p)_t^n \xrightarrow{ucp} m_p \int_0^t |\sigma_s|^p ds$$

with $m_p = \mathbb{E}(|N(0,1)|^p)$.

Law of large numbers

- **Theorem:** Under certain regularity conditions on the variogram R we obtain the convergence

$$V(X,p)_t^n \xrightarrow{ucp} m_p \int_0^t |\sigma_s|^p ds$$

with $m_p = \mathbb{E}(|N(0,1)|^p)$.

- For the corresponding CLT we require some stronger conditions.

Central limit theorem

Theorem: Let the volatility process σ be smooth enough and further assume that

$$R_t = \text{const} \cdot t^{\alpha} + O(t^{\alpha+\alpha'}) \qquad \text{as } t \to 0$$

for some $\alpha \in (0, 3/2) \setminus \{1\}$ and $\alpha' > 0$. Then we deduce the stable convergence

$$\Delta_n^{-1/2}\Big(V(X,p)_t^n - m_p \int_0^t |\sigma_s|^p ds\Big) \xrightarrow{\mathcal{D}_{st}} \rho \int_0^t |\sigma_s|^p dW_s',$$

where W' is a new Brownian motion (independent of everything) and

$$\rho^2 = \lim_{n\to\infty} \Delta_n^{-1} \text{Var}(V(B^H, p)_1^n)$$

with B^H being a fBm with Hurst parameter $H = \alpha/2$.

Some remarks

- The central limit theorem is proved via a combination of a blocking technique, Malliavin calculus and the properties of stable convergence (and many many approximations).

Some remarks

- The central limit theorem is proved via a combination of a blocking technique, Malliavin calculus and the properties of stable convergence (and many many approximations).
- Joint convergence for a family of powers $(p_1, \ldots, p_k)$ and frequencies $(d_1\Delta_n, \ldots, d_l\Delta_n)$ is straightforward.

Some remarks

- The central limit theorem is proved via a combination of a blocking technique, Malliavin calculus and the properties of stable convergence (and many many approximations).
- Joint convergence for a family of powers $(p_1, \ldots, p_k)$ and frequencies $(d_1\Delta_n, \ldots, d_l\Delta_n)$ is straightforward.
- The restriction
$$\alpha \in (0, 3/2) \setminus \{1\}$$
is explained by the fact that for $\alpha \in (3/2, 2)$ we obtain a non-central limit theorem with a slower rate of convergence.

Ratio statistics

- Even though all asymptotic results are infeasible we can use the relationship

$$\tau^2_{\Delta_n} = \text{const} \cdot \Delta_n^{\alpha} + O(\Delta_n^{\alpha+\alpha'}),$$

to estimate $\alpha \in (0,2)$. This implies that $\tau^2_{2\Delta_n}/\tau^2_{\Delta_n} \to 2^{\alpha}$.

Ratio statistics

- Even though all asymptotic results are infeasible we can use the relationship

$$\tau^2_{\Delta_n} = \text{const} \cdot \Delta_n^\alpha + O(\Delta_n^{\alpha+\alpha'}),$$

to estimate $\alpha \in (0,2)$. This implies that $\tau^2_{2\Delta_n}/\tau^2_{\Delta_n} \to 2^\alpha$.

- Our estimation method is based on the change of frequency:

$$R_t^n = \frac{\sum_{i=2}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-2)\Delta_n}|^2}{\sum_{i=1}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-1)\Delta_n}|^2} \xrightarrow{P} 2^\alpha.$$

Ratio statistics

- Even though all asymptotic results are infeasible we can use the relationship

$$\tau_{\Delta_n}^2 = \text{const} \cdot \Delta_n^{\alpha} + O(\Delta_n^{\alpha+\alpha'}),$$

  to estimate $\alpha \in (0,2)$. This implies that $\tau_{2\Delta_n}^2/\tau_{\Delta_n}^2 \to 2^{\alpha}$.

- Our estimation method is based on the change of frequency:

$$R_t^n = \frac{\sum_{i=2}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-2)\Delta_n}|^2}{\sum_{i=1}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-1)\Delta_n}|^2} \xrightarrow{P} 2^{\alpha}.$$

- A central limit theorem for the normalized statistic

$$\Delta_n^{-1/2}\Big(\frac{\log R_t^n}{\log 2} - \alpha\Big)$$

  holds for $\alpha \in (0,3/2) \setminus \{1\}$ and $\alpha' > 1/2$. The latter condition is required to ensure that $\Delta_n^{-1/2}(\tau_{2\Delta_n}^2/\tau_{\Delta_n}^2 - 2^{\alpha}) \to 0$.

Remark

- In practice it is more informative to consider a *power plot* to infer the parameter α. Consider the power variation ratio

$$R(q)_t^n = \frac{\sum_{i=2}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-2)\Delta_n}|^q}{\sum_{i=1}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-1)\Delta_n}|^q} \xrightarrow{P} 2^{q\alpha/2}.$$

Remark

- In practice it is more informative to consider a *power plot* to infer the parameter α. Consider the power variation ratio

$$R(q)_t^n = \frac{\sum_{i=2}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-2)\Delta_n}|^q}{\sum_{i=1}^{[t/\Delta_n]} |X_{i\Delta_n} - X_{(i-1)\Delta_n}|^q} \xrightarrow{P} 2^{q\alpha/2}.$$

- Example: For $q \in (\underline{a}, \overline{a})$ the scaling parameter α can be estimated via

$$\hat{\alpha} = \frac{1}{\overline{a} - \underline{a}} \int_{\underline{a}}^{\overline{a}} \frac{2 \log R(q)_t^n}{q \log 2} dq \xrightarrow{P} \alpha.$$

Higher order differences

- Now we provide an estimation method for the values

$$\alpha \in (3/2, 2)$$

Higher order differences

- Now we provide an estimation method for the values

$$\alpha \in (3/2, 2)$$

- It turns out that considering higher order differences solves the problem. Let $\Delta_i^{(q)n} X$ denote the qth order difference of X, e.g.

$$\Delta_i^{(2)n} X = X_{i\Delta_n} - 2X_{(i-1)\Delta_n} + X_{(i-2)\Delta_n}.$$

  Define the power variation via

$$V^{(q)}(X, p)_t^n = \Delta_n (\tau_{\Delta_n}^{(q)})^{-p} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^{(q)n} X|^p$$

  with $(\tau_{\Delta_n}^{(q)})^2 = \mathbb{E}(\Delta_i^{(q)n} G)^2$.

Asymptotic theory

Theorem: For all $\alpha \in (0,2)$ and $q \geq 1$ it holds that

$$V^{(q)}(X,p)_t^n \xrightarrow{ucp} m_p \int_0^t |\sigma_s|^p ds.$$

Under further assumptions on R and σ we obtain

$$\Delta_n^{-1/2}\Big(V^{(q)}(X,p)_t^n - m_p \int_0^t |\sigma_s|^p ds\Big) \xrightarrow{\mathcal{D}_{st}} \rho^{(q)} \int_0^t |\sigma_s|^p dW_s'$$

for all $q \geq 2$. Here W' is a new Brownian motion (independent of everything) and

$$(\rho^{(q)})^2 = \lim_{n\to\infty} \Delta_n^{-1} \mathrm{Var}(V^{(q)}(B^H,p)_1^n)$$

with B^H being a fBm with Hurst parameter $H = \alpha/2$.

A model with smooth drift

- Let us now consider the model

$$Z = X + Y,$$

where $X_t = X_0 + \int_0^t \sigma_s dG_s$ is our basic process and Y is a drift process with

$$Y \in C^r(\mathbb{R}_{\geq 0}) \qquad \text{a.s.}$$

A model with smooth drift

- Let us now consider the model

$$Z = X + Y,$$

where $X_t = X_0 + \int_0^t \sigma_s dG_s$ is our basic process and Y is a drift process with

$$Y \in C^r(\mathbb{R}_{\geq 0}) \qquad \text{a.s.}$$

- Our main object of interest is the scaling parameter

$$\alpha \in (0,2)$$

of X.

A model with smooth drift

- Let us now consider the model

$$Z = X + Y,$$

where $X_t = X_0 + \int_0^t \sigma_s dG_s$ is our basic process and Y is a drift process with

$$Y \in C^r(\mathbb{R}_{\geq 0}) \qquad \text{a.s.}$$

- Our main object of interest is the scaling parameter

$$\alpha \in (0, 2)$$

of X.

- It turns out that the higher order differences have a second useful property: they make the power variation robust to certain smooth drift processes.

Robust asymptotic results

Theorem: Let $Z = X + Y$ and $Y \in C^r(\mathbb{R}_{\geq 0})$ a.s.

(i) If $r > \alpha/2$ then it holds that

$$V^{(q)}(Z,p)_t^n - V^{(q)}(X,p)_t^n \xrightarrow{ucp} 0$$

for all $\alpha \in (0,2)$, $p \geq 0$ and $q \geq 1$.

(ii) If $r - \alpha/2 > 1/2$ then it holds that

$$\Delta_n^{-1/2}\Big(V^{(q)}(Z,p)_t^n - V^{(q)}(X,p)_t^n\Big) \xrightarrow{ucp} 0$$

for all $\alpha \in (0,2)$, $p \geq 0$ and $q \geq \min([r],1) + 1$.

Estimation with gaps

- Let us go back to the original model

$$X_t = X_0 + \int_0^t \sigma_s dG_s, \qquad t \geq 0.$$

Estimation with gaps

- Let us go back to the original model

$$X_t = X_0 + \int_0^t \sigma_s dG_s, \qquad t \geq 0.$$

- Assume now that the variogram R of G satisfies the relation

$$R_t = \text{const} \cdot t^\alpha + O(t^{\alpha+\alpha'}), \qquad \text{as } t \to 0,$$

with $\alpha \in (0,2)$ and $\alpha' \in (0,1/2)$.

Estimation with gaps

- Let us go back to the original model

$$X_t = X_0 + \int_0^t \sigma_s dG_s, \qquad t \geq 0.$$

- Assume now that the variogram R of G satisfies the relation

$$R_t = \text{const} \cdot t^{\alpha} + O(t^{\alpha+\alpha'}), \qquad \text{as } t \to 0,$$

  with $\alpha \in (0,2)$ and $\alpha' \in (0,1/2)$.

- As $\tau_{2\Delta_n}^2/\tau_{\Delta_n}^2 - 2^{\alpha} = O(\Delta_n^{\alpha})$ the CLT for the ratio statistic R_t^n does not hold anymore, because the quantity

$$\Delta_n^{-1/2}(\tau_{2\Delta_n}^2/\tau_{\Delta_n}^2 - 2^{\alpha})$$

  explodes as $\Delta_n \to 0$.

Estimation with gaps

- The idea is now to introduce gaps in the definition of the ratio statistic to achieve a slower convergence rate. This prevents the afore-mentioned bias from the explosion.

Estimation with gaps

- The idea is now to introduce gaps in the definition of the ratio statistic to achieve a slower convergence rate. This prevents the afore-mentioned bias from the explosion.

- Let $M_n \to \infty$ with $M_n\Delta_n \to 0$. Define a new ratio statistic via

$$\overline{R}_t^n = \frac{\sum_{i=2}^{[t/M_n\Delta_n]} |X_{iM_n\Delta_n} - X_{iM_n\Delta_n-2}|^2}{\sum_{i=1}^{[t/M_N\Delta_n]} |X_{iM_n\Delta_n} - X_{iM_n\Delta_n-1}|^2}.$$

Estimation with gaps

- The idea is now to introduce gaps in the definition of the ratio statistic to achieve a slower convergence rate. This prevents the afore-mentioned bias from the explosion.

- Let $M_n \to \infty$ with $M_n\Delta_n \to 0$. Define a new ratio statistic via

$$\overline{R}_t^n = \frac{\sum_{i=2}^{[t/M_n\Delta_n]} |X_{iM_n\Delta_n} - X_{iM_n\Delta_n - 2}|^2}{\sum_{i=1}^{[t/M_N\Delta_n]} |X_{iM_n\Delta_n} - X_{iM_n\Delta_n - 1}|^2}.$$

- Choose M_n such that

$$(M_n\Delta_n)^{-1/2}(\tau_{2\Delta_n}^2/\tau_{\Delta_n}^2 - 2^\alpha) \to 0.$$

Robust asymptotic results

Theorem: Assume that

$$R_t = \text{const} \cdot t^{\alpha} + O(t^{\alpha+\alpha'}) \qquad \text{as } t \to 0$$

for some $\alpha \in (0, 3/2) \setminus \{1\}$ and $\alpha' \in (0, 1/2)$.

(i) We deduce that $\overline{R}_t^n \xrightarrow{P} 2^{\alpha}$ and

$$(M_n \Delta_n)^{-1/2} \left(\overline{R}_t^n - 2^{\alpha} \right) \xrightarrow{\mathcal{D}_{st}} \int_0^t f_s dW_s'$$

for a known process $(f_s)_{s \geq 0}$.

(ii) In the critical case $M_n \sim \Delta_n^{2\alpha'-1}$ we obtain the convergence rate

$$\Delta_n^{-\alpha'}.$$

This is indeed the optimal convergence rate.

Thank you!