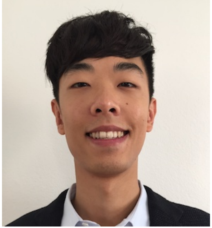
Statistical inference for rough volatility: Minimax theory

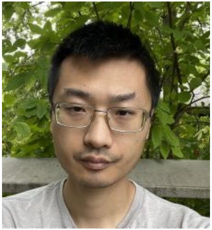
Gregoire Szymanski

École Polytechnique

May 2023

- This presentation is based on a joint work with Carsten Chong, Marc Hoffmann, Yanghui Liu, Mathieu Rosenbaum :






- Statistical inference for rough volatility : Minimax Theory, arXiv, 2022.
- Statistical inference for rough volatility : Central limit theorems, arXiv, 2022.

Table of contents

Setup

Rough volatility model

We consider an asset whose log price is modelled by a stochastic process S with dynamic

$$\begin{cases} dS_t = \sigma_t dB_t, \\ \sigma_t^2 = \exp(\eta W_t^H) \end{cases}$$

where $(H, \eta) \in \mathcal{D} = [H_-, H_+] \times [\eta_-, \eta_+]$, and where B is a Brownian motion and W^H is a Fractional Brownian Motion independent of B.

High frequency asymptotic

We aim at estimating the parameters H and η from high frequency observations of S. More precisely, we observe S at times $\frac{i}{n}$ with $i = 0, \ldots, n$ and we suppose that $n = 2^N$.

Fractional Brownian motions

Definition

A **fractional Brownian motion** W^H with Hurst index H with $0 < H < 1$ is the Gaussian process satisfying

$$\mathrm{Cov}(W_t^H, W_s^H) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}).$$

Proposition

- *W^H has stationary increments.*
- *$(a^{-H} W_{at}^H)_t$ is a fractional Brownian motion with Hurst index H.*

State of the art

- Estimation in stationary Gaussian sequences : Fox & Taqqu '86 and '87 ; Istas & Lang '97 ; Coeurjolly '01.
- Estimation of the Hurst parameter in additive noise models : Gloter and Hoffmann '07 **when $\mathbf{H > 1/2}$ only**.
- Estimation of the Hurst parameter in multiplicative noise models : Rosenbaum '08 **when $\mathbf{H > 1/2}$ only**.
- LAN property : Kawai '13 ; Brouste, Fukasawa '18.
- Whittle estimators : Fukasawa, Takabatake, Westphal '19

Table of contents

Minimax rates

Definition

The rate v_n is said to be a **lower rate of convergence** over $\mathcal{D}$ for estimating H if there exists $c > 0$ such that

$$\liminf_{n\to\infty} \inf_{\widehat{H}_n} \sup_{(H,\eta)\in\mathcal{D}} \mathbb{P}^n_{H,\eta}(v_n^{-1}|\widehat{H}_n - H| \geq c) > 0$$

where the infimum is taken over all admissible estimators.

Theorem

The rate $v_n(H) = n^{-1/(4H+2)}$ is a lower rate of convergence for estimating H over $\mathcal{D}$.

Table of contents

Construction of the estimator

- Our observations are given by

$$S_{i/n} = \int_0^{i/n} \sigma_s \, dB_s.$$

- Stochastic calculus gives

$$\left(S_{i/n} - S_{(i-1)/n}\right)^2 = \int_{(i-1)/n}^{i/n} \sigma_t^2 \, dt \; X_{i,n}^2$$

  where $X_{i,n}$ are independent standard Gaussian variables.

- Using the logarithm gives

$$\log\left(n\left(S_{i/n} - S_{(i-1)/n}\right)^2\right) = \log\left(n \int_{(i-1)/n}^{i/n} \sigma_t^2 \, dt\right) + \log\left(X_{i,n}^2\right).$$

Approximation of the spot volatility

- Here, we want to use the approximation

$$n\int_{(i-1)/n}^{i/n}\sigma_t^2 dt \approx \sigma_{i/n}^2 = \exp(\eta W_{i/n}^H)$$

so that we would be down to

$$Y_{i,n} = \eta W_{i/n}^H + \varepsilon_{i,n}$$

where the additive noise $\varepsilon_{i,n}$ is a well-behaved noise process.

Energy levels

Definition (Energy Level)

For $j \leq N$, we define the energy levels of the Fractional Brownian trajectory by

$$Q_j = \eta^2 2^{-j} \sum_{k=0}^{2^{j-1}} \left(W^H_{(k+1)2^{-j}} - W^H_{k2^{-j}}\right)^2.$$

Proposition

$$\mathbb{E}[(Q_j - \tfrac{1}{2}\eta^2 2^{-2jH})^2] \leq C 2^{-j(1+4H)}.$$

Energy levels

- Therefore

$$2^{-2H} \approx \frac{Q_{j+1}}{Q_j}$$

  and we shall estimate H through this ratio.
- Therefore, we need to estimate correctly the energy levels, and we start by estimating the increments

$$d_{j,k} = \eta 2^{-j/2}\left(W^H_{(k+1)2^{-j}} - W^H_{k2^{-j}}\right).$$

- This is done by averaging the observations of the fractional Brownian motion to dampen the effects of the noise, which gives

$$\widetilde{d}_{j,k} = 2^{-j/2} \sum_{l=0}^{2^{N-j}-1} \left(Y_{(k+1)2^{N-j}+l,n} - Y_{k2^{N-j}+l}\right).$$

Energy levels

- d can be decomposed as

$$\widetilde{d}_{j,k} = 2^{-j/2} \sum_{l=0}^{2^{N-j}-1} \left(Y_{(k+1)2^{N-j}+l,n} - Y_{k2^{N-j}+l} \right)$$
$$= \underbrace{2^{-j/2}\eta \sum_{l=0}^{2^{N-j}-1} \left(W^H_{(k+1)2^{-j}+(l+1)/n} - W^H_{(k+1)2^{-j}+l/n} \right)}_{\approx d_{j,k}}$$
$$+ \underbrace{2^{-j/2} \sum_{l=0}^{2^{N-j}-1} \left(\varepsilon_{(k+1)2^{N-j}+l,n} - \varepsilon_{k2^{N-j}+l} \right)}_{=: e_{j,k}}.$$

Energy levels

- **Provided ε is well behaved**, $\widetilde{d}_{j,k}$ is a good estimator for $d_{j,k,n}$.
- We need to correct the effect of the second moment of the noise in

$$\widetilde{d}_{j,k}^2 \approx (d_{j,k,n})^2 + 2d_{j,k,n}e_{j,k} + e_{j,k}^2,$$

and therefore we write

$$\widehat{Q}_{j,n} = \sum_{k=0}^{2^{j-1}} \widetilde{d}_{j,k}^2 - \mathbb{E}[e_{j,k}^2].$$

- Again, **provided ε is well behaved**, we have

$$\mathbb{E}[(\widehat{Q}_{j,n} - Q_j)^2] \leq C2^j/n^2.$$

Construction of the estimator

- Since $2^{-2H} \approx Q_{j+1}/Q_j$, our final estimator of H is

$$-\frac{1}{2}\log\left(\frac{\widehat{Q}_{j+1,n}}{\widehat{Q}_{j,n}}\right)$$

  but we still need to find an appropriate level j.
- We do a Biais-variance decomposition, the biais is of order $2^{-j/2}$ while the variance is of order $n^{-1}2^{j(4H+1)/2}$.
- Balancing both term yields $2^j \approx n^{1/(2H+1)}$

Construction of the estimator

Adaptive estimation

We can cook up $\widehat{J}_n^* \approx n^{1/(2H+1)}$ by choosing $\widehat{J}_n^*$ so that

$$Q_{J_n^*,n} \approx n^{-1} 2^{J_n^*}$$

The estimator of H is eventually given by

$$\widehat{H}_n = -\frac{1}{2} \log \frac{\widehat{Q}_{\widehat{J}_n^*+1}}{\widehat{Q}_{\widehat{J}_n^*}}.$$

Theorem

$n^{1/(4H+2)}(\widehat{H}_n - H)$ is bounded in probability uniformly over $\mathcal{D}$.

Table of contents

About the approximations

Unfortunately, when $H < 1/2$, The procedure detailed previously does not work because ε is not nice...

$$\eta W^H_{i/n} + \varepsilon_{i,n} = \log\left(n \int_{(i-1)/n}^{i/n} \sigma_t^2 \, dt\right) + \log\left(X_{i,n}^2\right)$$
$$\approx \log\left(\sigma^2_{(i-1)/n}\right) + \log\left(X_{i,n}^2\right).$$

Taylor development

Proposition

There exists random variable Z_0 bounded in $L^2(\mathbb{P}_{H,\eta})$ uniformly on $\mathcal{D}$ such that

$$\log\left(n\int_{i/n}^{(i+1)/n}\sigma_u^2 du\right) = \cdots$$

$$= \sum_{b=2}^{2S}\sum_{s=1}^{2S}\frac{(-1)^{s-1}}{s}\sum_{\substack{\mathbf{r}\in\{1,\dots,S\}^s\\ \sum_j \mathbf{r}_j=b}}\prod_{j=1}^{s}\frac{\eta^{\mathbf{r}_j}}{\mathbf{r}_j!}\frac{1}{/n}\int_{i/n}^{(i+1)/n}(W_u^H - W_{i/n}^H)^{\mathbf{r}_j}du$$

$$+ n\int_{i/n}^{(i+1)/n}\eta W_u^H du + Z(i,n)\cdot n^{-H^*(S+1)}$$

where the random variables $Z(i,n)$ satisfy $|Z(i,n)| \le Z_0$.

About the approximations

- We can still define the energy levels associated with these observations, but the scaling is hugely influenced by the presence of the additional terms.
- Indeed, the we cannot hope having an expression like $\mathbb{E}[(Q_{j,n} - \eta^2\kappa(H)2^{-2jH})^2] \leq C2^{-j(1+4H)}$ because the terms $(W_u^H - W_{i/n}^H)^{\mathbf{r}_j}$ create additional scaling terms of order 2^{2Hj}.

Energy levels

Proposition

There exist explicit functions of H *denoted* κ_a *such that if* $S \geq 1/(4H_-) + 1/2$ *and* $S > H_+/(2H_-) - 1/2$*, we have*

$$\mathbb{E}_{H,\eta}\Big[\big(Q_j - \sum_{a=1}^{S} \eta^{2a} 2^{-2aHj} \kappa_a(H)\big)^2\Big] \leq C 2^{-j(1+4H)}$$

for some constant C *depending only on* S*.*

Energy levels

- Therefore the scaling $Q_{j+1}/Q_j = 2^{-2H}$ is no longer exact.
- Instead we have

$$\frac{Q_{j+1}}{Q_j} \approx \frac{\sum_{a=1}^{S} \eta^{2a} 2^{-2aH(j+1)} \kappa_a(H)}{\sum_{a=1}^{S} \eta^{2a} 2^{-2aHj} \kappa_a(H)} \approx 2^{-2H} + O(2^{-2Hj})$$

- We can build an estimator using the same procedure as in the simplified model, but it will exhibit a slower rate of convergence because of the additional contribution of order 2^{-2Hj} in the biais of this estimator.

Theorem

$(n^{1/(4H+2)} \wedge n^{2H})(\widehat{H}_n - H)$ is bounded in probability uniformly over $\mathcal{D}$.

Biais correction

- We need an additional biais correction procedure to improve the convergence rate of this estimator.
- First we need an estimator for η.

Proposition

If $v_n(\widehat{H}_n - H)$ is bounded in probability uniformly over $\mathcal{D}$, we can build an estimator $\widehat{\eta}_n$ of η such that $v_n \log(n)^{-1}(\widehat{\eta}_n - \eta)$ is bounded in probability uniformly over $\mathcal{D}$.

Biais correction

- The scaling of the energy levels is given by $Q_j \approx \sum_{a=1}^{S} \eta^{2a} 2^{-2aHj} \kappa_a(H)$ and we want to cut this sum to $a = 1$.
- Thus we need to replace Q_j by

$$Q_j - \sum_{a=2}^{S} \eta^{2a} 2^{-2aHj} \kappa_a(H) \approx \eta^2 2^{-2Hj} \kappa_1(H).$$

- So we replace $\widehat{Q}_j$ by $\widehat{Q}_j^c(\widehat{H}_n, \widehat{\eta}_n)$

$$\widehat{Q}_j^c(\widetilde{H}, \widetilde{\eta}) = \widehat{Q}_j - \sum_{a=2}^{S} \widetilde{\eta}^{2a} 2^{-2a\widetilde{H}j} \kappa_a(\widetilde{H})$$

Biais correction

- Notice that

$$\frac{\widehat{Q}_j^c(H,\eta)}{\widehat{Q}_j^c(H,\eta)} = 2^{-2H}$$

  as in the simplified setup.
- Thus we define

$$\widehat{H}_n^{(1)} = -\frac{1}{2}\log\frac{\widehat{Q}_{J_n^*+1}(\widehat{H}_n,\widehat{\eta}_n)}{\widehat{Q}_{J_n^*}(\widehat{H}_n,\widehat{\eta}_n)}$$

Biais correction

- However, we cannot immediately retrieve the convergence rate of the simplified setup because $\widehat{Q}_j^c(\widehat{H}_n, \widehat{\eta}_n) \neq \widehat{Q}_j^c(H, \eta)$. The difference $\widehat{Q}_j^c(\widehat{H}_n, \widehat{\eta}_n) - \widehat{Q}_j^c(H, \eta)$ also creates a biais that we can still control.

Theorem

$(n^{1/(4H+2)} \wedge n^{4H(H+1)/(2H+1)})(\widehat{H}_n^{(1)} - H)$ *is bounded in probability uniformly over* $\mathcal{D}$.

Biais correction

- We repeat the biais correction procedure by defining a sequence of estimators

$$\widehat{H}_n^{(m+1)} = -\frac{1}{2}\log\frac{\widehat{Q}_{J_n^*+1}(\widehat{H}_n^{(m)},\widehat{\eta}_n^{(m)})}{\widehat{Q}_{J_n^*}(\widehat{H}_n^{(m)},\widehat{\eta}_n^{(m)})}$$

Theorem

$(n^{1/(4H+2)} \wedge n^{2H(2H+m+1)/(2H+1)})(\widehat{H}_n^{(m)} - H)$ is bounded in probability uniformly over $\mathcal{D}$.

- We take $m_{opt} > m > 1/(4H) - 2H - 1$ for any $H_- < H < H_+$ and we get the same convergence rate $n^{1/(4H+2)}$ as in the simplified model.

Table of contents

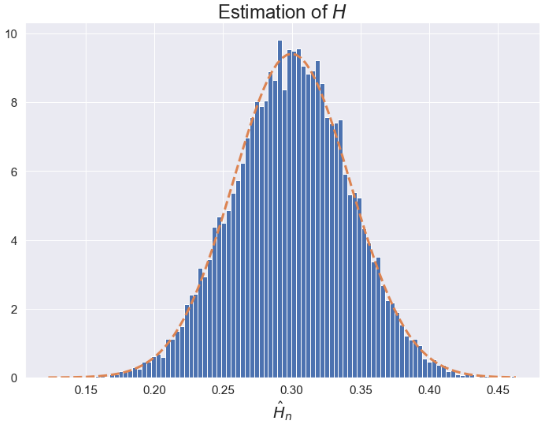
Summary

- The lower bound for estimating the parameter H is $n^{-1/(4H+2)}$, and we have developed an estimator achieving this rate.
- This rate is unusual and could seem counter-intuitive at first glance since it the optimal rate for estimating β-Holder continuous function in most models is usually $n^{-\beta/(2\beta+1)}$.
- But we do not seek to reconstruct the spot volatility.
- The rougher the volatility is, the easier it is to see it.
- Beyond this simplified framework, we can extend these idea to (rough) stochastic Volterra differential equations.

Thank you !

Asymptotic behaviour

- These Estimators are asymptotically Gaussian...



Non parametric estimation

- In practice, the models considered in finance do not use Fractional Brownian motion but non-parametric approximations of the fractional Brownian motion :

$$X_t = X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s \, dB_s$$
$$\sigma_t^2 = \sigma_0^2 + \int_0^t a_s \, ds + \int_0^t g(t-s)\eta_s \, d\widetilde{B}_s$$

  where $g(t) \approx t^{H-1/2}$.

- We can extend our estimators to embrace this framework.