The $m(q,\Delta)$ estimator for fBM

Let $X_t = \sigma B_t^H$, and set $SS_n^{(q)} := \frac{1}{n}\sum_{i=1}^n |X_{i\Delta} - X_{(i-1)\Delta}|^q$. Then

$$\mathbb{E}(SS_n^{(q)}) \;=\; \frac{1}{n}\mathbb{E}(\sum_{i=1}^n |X_{i\Delta} - X_{(i-1)\Delta}|^q) \;=\; \Delta^{qH}\frac{1}{n}\sum_{i=1}^n \mathbb{E}(|X_i - X_{i-1}|^q) \;=\; \sigma^q\mathbb{E}(|Z|^q)\Delta^{qH} \;=\; \sigma^q K_q \Delta^{qH} \quad (1)$$

where $\Delta = \Delta_n = \frac{1}{n}$, $K_q = \mathbb{E}(|Z|^q) = \frac{2^{q/2}}{\sqrt{\pi}}\Gamma(\frac{q+1}{2})$, (for $q > -1$) and $Z \sim N(0,1)$.

This leads to a simple (and scale-independent) estimator for H given by

$$\hat{H}_n \;=\; -\frac{1}{2}\log_2 \frac{SS_n^{(2)}}{SS_{\frac{1}{2}n}^{(2)}} \quad (2)$$

and one can show using the Central Limit Theorem that

$$\sqrt{n}\,(\widehat{H}_n - H) \;\xrightarrow{d}\; \mathcal{N}(0, V_H), \quad (3)$$

where

$$V_H \;=\; \frac{1}{2(\ln 2)^2}(\sum_{k\in\mathbb{Z}}\rho_H(k)^2 \;+\; 2\sum_{m\in\mathbb{Z}} r_W(m)^2 \;-\; 4\sum_{m\in\mathbb{Z}} r_{ZW}(m)^2).$$

Now define, for $m \in \mathbb{Z}$,

$$r_W(m) : \;=\; 2^{-2H}(2\rho_H(2m) + \rho_H(2m-1) + \rho_H(2m+1)), \qquad r_{ZW}(m) := 2^{-H}(\rho_H(2m) + \rho_H(2m-1)).$$

which allows us to compute a confidence interval for H.

More generally, we can compute joint estimators $(\hat{H}_n, \hat{\sigma}_n)$ for (H,σ) defined by

$$SS_n^{(q)} \;=\; \hat{\sigma}_n^q K_q \Delta^{q\hat{H}_n}$$

if we have computed $SS_n^{(q)}$ for at least two Δ-values. Taking logs we see that

$$\log SS_n^{(q)} \;=\; q\log\hat{\sigma}_n + \log K_q \;+\; q\hat{H}_n \log\Delta$$

so we can perform **linear regression** on $\log SS_n^{(q)}$ vs $\log\Delta = \log\frac{1}{n}$ for a range of Δ-values (i.e. using a log-log plot, see plot overleaf). Then for the line of best fit, the **slope** will equal $q\hat{H}_n$ (q is chosen by you, e.g. $q = 1, 2, 2.5, 3$ etc), and the **intercept** at $\log\Delta = 0$ is $q\log\hat{\sigma}_n + \log K_q$, from which we can compute $\hat{\sigma}_n$ since K_q has an explicit formula above. This is the $m(q,\Delta)$ estimator discussed in [GJR18]. One can then also compute the R^2**-statistic** for the regression (which measures how close the data is to the line of best fit), and try to estimate the **sample variance** of $\hat{H}_n$ and $\hat{\sigma}_n$.

To approximate the effect of using **realized variance** with m subwindows to estimate V_t (as in Part 2 of the project), we can use the Central Limit Theorem approximation from FM02:

$$V_{i\Delta} \;=\; V_0 e^{\sigma B_{i\Delta}^H}(1 + \sqrt{\frac{2}{m}}\,\varepsilon_i)$$

where the ε_i's here are i.i.d. $N(0,1)$ (and independent of B^H), so (using that $\log(1+x) = x + O(x^2)$), we see that $\log V_{i\Delta} = \log V_0 + \sigma B_{i\Delta}^H + \sqrt{\frac{2}{m}}\varepsilon_i + O(\frac{1}{m}) = \log V_0 + X_{i\Delta} + \sqrt{\frac{2}{m}}\varepsilon_i + O(\frac{1}{m})$.

For convenience we now define $\tilde{X}_{i\Delta} = X_{i\Delta} + \sqrt{\frac{2}{m}}\,\varepsilon_i$. Then $\tilde{X}_{i\Delta} - \tilde{X}_{(i-1)\Delta} = X_{i\Delta} - X_{(i-1)\Delta} + \sqrt{\frac{2}{m}}(\varepsilon_i - \varepsilon_{(i-1)}) \sim N(0, \sigma^2\Delta^{2H} + \frac{4}{m})$, and setting $\widetilde{SS}_n^{(2)} := \frac{1}{n}\sum_{i=1}^n |\tilde{X}_{i\Delta} - \tilde{X}_{(i-1)\Delta}|^2$, adding the effect of ε into the computation above we find that

$$\mathbb{E}(\widetilde{SS}_n^{(2)}) \;=\; \sigma^2\Delta^{2H} + \frac{4}{m}$$

(since $K_q = 1$ for $q = 2$) so we now regress $\log(\widetilde{SS}_n^{(2)} - \frac{4}{m})$ vs $\log\Delta$, which provides a smart adjustment to $\hat{H}_n$). (this adjustment is made in
`https://colab.research.google.com/drive/1jJGf4bVWETJqWRIMZ6STjsJ9jINgEaPd#scrollTo=e9fn6ABHNf52`

The $m(q,\Delta)$ estimator for the RL process

If $X_t = \sigma Z_t^H$ where $Z_t^H = \sqrt{2H}\int_0^t (t-s)^{H-\frac{1}{2}}dW_s$ is an RL process, then for $q=2$ and $H \in (0,\frac{1}{2})$ one can show that

$$\mathbb{E}(SS_n^{(2)}) \;=\; \frac{1}{n}\mathbb{E}(\sum_{i=1}^n |X_{i\Delta}-X_{(i-1)\Delta}|^2) \;=\; \sigma^q \Delta^{2H}\frac{1}{n}\sum_{i=1}^n \mathbb{E}(|Z_i^H - Z_{i-1}^H|^2) \;\sim\; \sigma^2 c_H \Delta^{2H}$$

as $n\to\infty$ where $c_H = -\frac{4H\Gamma(\frac{1}{2}+H)\Gamma(-2H)}{\Gamma(\frac{1}{2}-H)} > 0$. Hence for the regression we now consider

$$\log(SS_n^{(2)}) \;=\; 2\log\hat{\sigma}_n + \log c_{\hat{H}} + 2\hat{H}_n \log\Delta$$

so now the intercept is $2\log\hat{\sigma}_n + \log c_{\hat{H}}$, which leads to an adjusted estimate for $\hat{\sigma}_n$. This can also be used for driftless rough Heston model below since (modulo an unimportant constant) both processes have the same covariance.

Convergence of $\hat{H}_n$ to H for the first task in Part 2, and asymptotic normality

Above we considered sums of the form

$$SS_n^{(q)} \;:=\; \frac{1}{n}\sum_{j=1}^n |B_j^H - B_{j-1}^H|^q\,. \tag{4}$$

We cannot apply the SLLN to this since the RVs in the sum here are not i.i.d. However, since the increments process $X_j = B_j^H - B_{j-1}^H \sim N(0,1)$ is stationary (in particular $X_j \sim N(0,1)$ for all j) and moreover is known to be **ergodic**, which means that for any measurable function g with $\mathbb{E}(|g(X_1)|)<\infty$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^n g(X_j) \;=\; \mathbb{E}(g(X_1)) \qquad \text{a.s.} \tag{5}$$

In general, a stationary Gaussian process Y is ergodic if its covariance function $R(k)\to 0$ as $k\to\infty$, which we verified in the fBM.pdf document. Hence we can apply the general ergodic property in (5) to show that $SS_n^{(q)}$ in (4) tends to the **non-random** limit $K_q := \mathbb{E}(|Z|^q)$ a.s. (where $Z\sim N(0,1)$, see Homework 3); hence from the self-similarity of fBM

$$SS_n^{(q)} \;\overset{w}{\to}\; \frac{1}{n}n^{qH}\sum_{j=1}^n |B_{j/n}^H - B_{(j-1)/n}^H|^q$$

i.e. the right hand side tends weakly to the constant K_q as $n\to\infty$, which also implies convergence in probability.

Recall this also suggests the estimator of $\hat{H}=\hat{H}_n$ for H defined by the relation

$$K_q \;=\; \frac{1}{n}n^{q\hat{H}}\sum_{j=1}^n |B_{j/n}^H - B_{(j-1)/n}^H|^q \;=\; \frac{1}{n}n^{qH}\sum_{j=1}^n |B_{j/n}^H - B_{(j-1)/n}^H|^q \cdot n^{q(\hat{H}-H)}\,.$$

But from the discussion immediately before this, we know that $\frac{1}{n}n^{qH}\sum_{j=1}^n |B_{j/n}^H - B_{(j-1)/n}^H|^q$ tends to K_q in probability as $n\to\infty$, hence we must have that $n^{q(\hat{H}-H)}\to 1$ in probability as well, which implies that

$$q(\hat{H}_n - H)\log n \;\to\; 0$$

in probability (where we are now emphasizing the dependence of $\hat{H}$ on n), which can only be true if $\hat{H}_n - H \to 0$ in probability (recall that $\hat{H}$ depends on n).

Moreover, it can be shown that $\sqrt{n}\log n\,(\hat{H}_n - H)$ is asymptotically $N(0,2\gamma)$ as $n\to\infty$, where

$$\gamma \;=\; \sum_{r=-\infty}^{\infty}(|r+1|^{2H} - 2|r|^{2H} + |r-1|^{2H})$$

(see code and histogram numerically verifying this at `https://colab.research.google.com/drive/18ISGhNSN9ybbcpJlPO-Oje`

If a function X_t is $\frac{1}{q}$ Hölder continuous on $[0,T]$ i.e. $|X_t - X_s|\le c|t-s|^{1/q}$ for some constant c, then it has finite q-variation. To see this, we just note that

$$\sup_{\mathcal{P}}\sum_{t_k}|X_{t_k}-X_{t_{k-1}}|^q \;\le\; \sum_{t_k}|c(t_k - t_{k-1}|^{\frac{1}{q}})^q \;=\; c^q\sum_{t_k}|t_k - t_{k-1}| \;=\; c^q T \;<\; \infty \tag{6}$$

where $\mathcal{P}$ is the set of partitions of $[0,T]$.

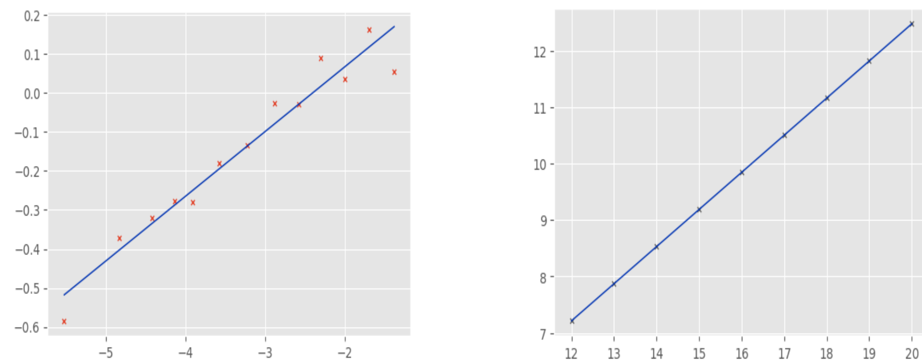

Figure 1: On the left, we see estimates of H for the SPX using the $m(q,\Delta)$ method for the SPX from 3rdJan22-15thJul24 for $q=2$ for which $\hat{H}=0.0830$ and $\hat{\sigma}=1.221$ (see similar plots in [GJR18]. On the right we see the linear regression in (9) for the Han-Schied method ($\log s_n$ vs n) for a true fBM path with 2^{20} time points, for which $\hat{H}=0.0508$, and $\hat{\sigma}=1.010$.

The Han-Schied [HS21] estimator for fBM

Let $X_t=\sigma B^H_t$ and let

$$\theta_{m,k} \;=\; 2^{\frac{m}{2}}(2X_{\frac{2k+1}{2^{m+1}}}-X_{\frac{k}{2^m}}-X_{\frac{k+1}{2^m}}) \;=\; -2^{\frac{m}{2}}(X_{\frac{2(k+1)}{2^{m+1}}}-2X_{\frac{2k+1}{2^{m+1}}}+X_{\frac{2k}{2^{m+1}}})$$

(note the similarity of the second expression to a 2nd order finite difference estimate). Then (with some tedious algebra) using the formula for $R(s,t)=\mathbb{E}(B^H_s B^H_t)$, one can check that

$$\mathbb{E}(\theta^2_{m,k}) \;=\; \sigma^2 2^{m-2H(1+m)}(4-4^H)\,. \tag{7}$$

Then setting $s_n^2=\sum_{m=0}^{n-1}\sum_{k=0}^{2^m-1}\theta^2_{m,k}$, we see that $\mathbb{E}(s_n^2)=\sum_{m=0}^{n-1}2^m\mathbb{E}(\theta^2_{m,k})$ (since (7) does not depend on k) which simplifies to

$$\mathbb{E}(s_n^2) \;=\; \sigma^2(4^{n(1-H)}-1) \;\sim\; \sigma^2 4^{n(1-H)} \;=\; \sigma^2 2^{2n(1-H)} \tag{8}$$

as $n\to\infty$, which suggests an estimator $\hat{H}_n$ defined by $s_n=\hat{\sigma}_n\,2^{n(1-\hat{H}_n)}$ which (assuming $\hat{\sigma}_n=O(1)$ as $n\to\infty$) we can re-arrange as

$$\hat{H}_n \;=\; 1-\frac{1}{n}\log_2(\frac{s_n}{\hat{\sigma}_n}) \;=\; 1-\frac{1}{n}\log_2 s_n \,+\, O(\frac{1}{n})$$

where $\log_2$ denotes the base-2 logarithm, so (ignoring the $O(\frac{1}{n})$ remainder term), we recover the Han-Schied[HS21] estimator $\hat{H}_n=1-\frac{1}{n}\log_2 s_n$. Then

$$\begin{aligned}\mathbb{E}(\hat{H}_n) \;=\; 1-\frac{1}{n}\mathbb{E}(\log_2(s_n)) \;=\; 1-\frac{1}{2n}\mathbb{E}(\log_2(s_n^2)) \;\geq\; 1-\frac{1}{2n}\log_2\mathbb{E}(s_n^2) \;&=\; 1-\frac{1}{2n}\log_2(\sigma^2(4^{n(1-H)}-1))\\ &\geq\; 1-\frac{1}{2n}\log_2(\sigma^2(4^{n(1-H)}))\\ &=\; 1-\frac{1}{2n}\log_2(\sigma^2)\;-\;\frac{1}{2n}\log_2(4^{n(1-H)})\\ &=\; H-\frac{1}{n}\log_2\sigma\end{aligned}$$

and the final line is $>H$ if $\sigma<1$, so $\mathbb{E}(\hat{H}_n)>H$ if $\sigma<1$. Note this argument does not show that $\mathbb{E}(\hat{H}_n)<H$ if $\sigma>1$.

Jointly estimating H and σ using [HS21]

If $X_t=\sigma B^H_t$, we can jointly estimate H and σ by performing linear regression since

$$\log s_n \;=\; \log\hat{\sigma}_n \,+\, n(1-\hat{H}_n)\log 2 \tag{9}$$

but we now have to compute $\log s_n$ for a range of different n-values to get a line of best fit, for which the slope is $(1-\hat{H}_n)\log 2$ and the intercept is $\log\hat{\sigma}_n$.

Proof of scale-invariance of the [HS21] SSE

Recall Definition 8.1 in [HS21]:

$$\lambda_n^s(x) \quad := \quad \mathrm{argmin}_{\lambda>0} \sum_{k=n-m}^{n} \alpha_{n-k} \, (\hat{R}_k(\lambda x) - \hat{R}_{k-1}(\lambda x))^2 \tag{10}$$

[1] where $\hat{R}_n(x)$ is the basic [HS21] estimator at order n. Then we see that $\lambda_n^s(cx) = \frac{\lambda_n^s(x)}{c}$, so

$$\hat{R}_n(\lambda_n^s(cx)\, cx) \quad = \quad \hat{R}_n(\lambda_n^s(x)\, x)\,.$$

But the right hand side here is the SSE $R_n^s(x) := \hat{R}_n(\lambda_n^s(x)x)$ applied to the path x, and the left hand side is the SSE applied to the scaled path cx. Hence the SSE is **scale-invariant**.

Corollary 8.4 in [HS21] shows if $X = \sigma B^H$, then all of their three SSEs (R_n^s, R_n^t and R_n^r) have error

$$|R_n - H| \quad = \quad O(2^{-n/2}\sqrt{\log n}) \quad = \quad O(\frac{1}{\sqrt{N}}\sqrt{\log_2\log_2 N}) \tag{11}$$

since $N = 2^n$ so $2^{-n/2} = 2^{-\log_2 N/2} = \frac{1}{\sqrt{N}}$. Note this is slightly slower convergence than $O(\frac{1}{\sqrt{N}})$ because $\sqrt{\log_2\log_2 N}$ tends to infinity (very slowly). The intuition behind the SSE is that we are trying to find the λ-value that minimizes the difference between consecutive iterates: $|\hat{R}_k(\lambda x) - \hat{R}_{k-1}(\lambda x)|$, since (loosely speaking) this difference being smaller implies that the sequence $R_k(\lambda x)$ is converging quicker. Note (11) does not say anything about the asymptotic variance of $|R_n - H|$.

Note if $m = 0$ and $\alpha_0 = 1$, then (10) simplifies to

$$\lambda_n^s(x) \quad := \quad \mathrm{argmin}_{\lambda>0}(\hat{R}_k(\lambda x) - \hat{R}_{k-1}(\lambda x))^2$$

and the SSE simplifies to:

$$R_n \quad = \quad 1 - \frac{1}{2}\log_2(\frac{s_n^2}{s_{n-1}^2})\,. \tag{12}$$

(this is conceptually similar to (2), since in both cases we are just using a ratio of the same statistic at the finest and next finest resolution). Note H-Hölder continuity alone is not enough to guarantee convergence of R_n here, since if e.g. x is linear then $s_n = 0$ for all n, so the right hand side is undefined.

The driftless rough Heston model as a non-Gaussian process with the same covariance as the RL process

The driftless rough Heston model satisfies

$$V_t \quad = \quad V_0 + \nu\int_0^t (t-u)^{H-\frac{1}{2}}\sqrt{V_u}dW_u\,.$$

Then $\mathbb{E}(V_t) = V_0$, and V has covariance function:

$$\begin{aligned}
\mathbb{E}((V_s - V_0)(V_t - V_0)) \quad &= \quad \nu^2\mathbb{E}(\int_0^s (s-u)^{H-\frac{1}{2}}\sqrt{V_u}dW_u \cdot \int_0^t (t-r)^{H-\frac{1}{2}}\sqrt{V_r}dW_r) \\
&= \quad \nu^2\int_0^s (s-u)^{H-\frac{1}{2}}(t-u)^{H-\frac{1}{2}}\mathbb{E}(V_u)du \\
&= \quad V_0\nu^2\int_0^s (s-u)^{H-\frac{1}{2}}(t-u)^{H-\frac{1}{2}}du \quad = \quad V_0\nu^2\bar{R}(s,t)
\end{aligned}$$

for $0 \le s \le t$, where $\bar{R}(s,t)$ is the covariance function for the Riemann-Liouville (RL) process $Z_t = \int_0^t (t-u)^{H-\frac{1}{2}}dW_u$ used for the rough Bergomi model (note Z is a Gaussian process but V is not), but the explicit formula for $\bar{R}(s,t)$ is more complicated than the $R(s,t)$ formula for fBM.

The Mandelbrot-van Ness representation for fBM and relation to the RL process

We also have the **Mandelbrot-van Ness** representation for fBM:

$$W_t^H \quad = \quad c_H(\int_{-\infty}^0 ((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}})dW_s + \int_0^t (t-s)^{H-\frac{1}{2}}dW_s) \quad = \quad c_H(A_t + Z_t)$$

for $t \ge 0$, in terms of an RL process Z (and note that A_t is known at time zero for all $t \ge 0$), and $c_H = (\frac{2H\,\Gamma(\frac{3}{2}-H)}{\Gamma(H+\frac{1}{2})\,\Gamma(2-2H)})^{\frac{1}{2}}$. Note also that A_t and Z_t are independent.

[1] one can check the minimizer is unique here since the objective function is strictly convex

References

[BFG16] Bayer, C., P.K.Friz, and J.Gatheral, “Pricing under rough volatility”, *Quantitative Finance*,16(6), pp. 887–904, 2016.

[Bib20] Bibinger, M., “Cusum tests for changes in the Hurst exponent and volatility of fractional Brownian motion”, 2020

[Cha14] Chang, Y-C., “Efficiently Implementing the Maximum Likelihood Estimator for Hurst Exponent”, *Mathematical Problems in Engineering*, vol. 2014

[CD22] Cont, R. and P.Das, “Rough Volatility: Fact or Artefact?”, preprint, 2022

[F23] Forde, M., “Statistical issues and calibration problems under rough and Markov volatility”, `https://martinforde.github.io/Talk.pdf`

[GJR18] Gatheral, J., T.Jaisson and M.Rosenbaum, “Volatility is rough”, *Quantitative Finance*, 18(6), 2018

[HS21] Han, X. and A.Schied, “The Hurst roughness exponent and its model-free estimation”, preprint, 2021

[HS24] Han, X. and A.Schied, “Estimating the roughness exponent of stochastic volatility from discrete observations of the integrated variance”, preprint, 2024