

Alternate estimators for H

The $m(q, \Delta)$ estimator from [GJR18]

For the first task, let $SS_n^{(q)} := \frac{1}{n} \sum_{i=1}^n |B_{i\Delta}^H - B_{(i-1)\Delta}^H|^q$. Then

$$\mathbb{E}(SS_n^{(q)}) = \frac{1}{n} \mathbb{E}\left(\sum_{i=1}^n |B_{i\Delta}^H - B_{(i-1)\Delta}^H|^q\right) = \Delta^{qH} \frac{1}{n} \sum_{i=1}^n \mathbb{E}(|B_i^H - B_{i-1}^H|^q) = \mathbb{E}(|Z|^q) \Delta^{qH} = K_q \Delta^{qH} \quad (1)$$

where $K_q = \mathbb{E}(|Z|^q) = \frac{2^{q/2}}{\sqrt{\pi}} \Gamma(\frac{q+1}{2})$, with $q > -1$ and $Z \sim N(0, 1)$. From this we can then derive the estimator for the first task. Alternatively, if let $X = \sigma B_t^H$ and now define $SS_n^{(q)} = \frac{1}{n} \mathbb{E}\left(\sum_{i=1}^n |X_{i\Delta}^H - X_{(i-1)\Delta}^H|^q\right)$, and assume H and σ are unknown, then (1) changes to

$$\mathbb{E}(SS_n^{(q)}) = \sigma^q K_q \Delta^{qH}$$

which leads to the estimate

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

for H and σ . Taking logs we see that

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

so we can perform **linear regression** on $\log SS_n^{(q)}$ vs $\log \Delta$ for a range of Δ -values (i.e. using a log-log plot). Then for the line of best fit, the **slope** will equal $q\hat{H}$ (q is chosen by you, e.g. $q = 1, 2, 2.5, 3$ etc), and the **intercept** at $\log \Delta = 0$ is $q \log \hat{\sigma} + \log K_q$, from which we can compute $\hat{\sigma}$ 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} and $\hat{\sigma}$.

The Han-Schied [HS21] estimator

Let $X_t = \sigma B_t^H$ and let

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

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

$$\mathbb{E}(\theta_{m,k}^2) = \sigma^2 2^{m-2H(1+m)} (4 - 4^H). \quad (2)$$

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

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

as $n \rightarrow \infty$, which suggests an estimator \hat{H} defined by

$$s_n = \hat{\sigma} 2^{n(1-\hat{H})}$$

which (assuming $\hat{\sigma} = O(1)$ as $n \rightarrow \infty$) we can re-arrange as

$$\hat{H} = 1 - \frac{1}{n} \log_2 \left(\frac{s_n}{\hat{\sigma}} \right) = 1 - \frac{1}{n} \log_2 s_n + O\left(\frac{1}{n}\right)$$

where \log_2 denotes the base-2 logarithm, so (ignoring the $O(\frac{1}{n})$ remainder term), we recover the Han-Schied[HS21] estimator $\hat{H} = 1 - \frac{1}{n} \log_2 s_n$ (which in principle also works for general processes which aren't fBM if n is sufficiently large).

Or we can jointly estimate H and σ by performing linear regression since

$$\log s_n = \log \hat{\sigma} + n(1 - \hat{H}) \log 2$$

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 $(1 - \hat{H}) \log 2$ and the intercept is $\log \hat{\sigma}$.

You can then draw histograms of \hat{H} if you simulate M fBM paths and compute the sample variance for \hat{H} (or a confidence interval), or compute \hat{H} for real data, e.g. using the SPX data file or data from yahoo finance.

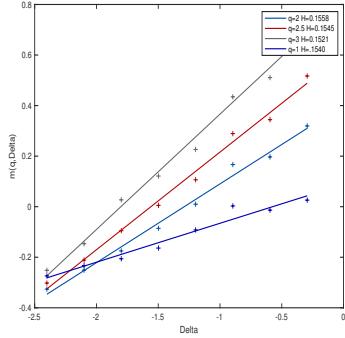


Figure 1: Estimates of H using the $m(q, \Delta)$ method using Oxford Mann data from 2000-2018 (see similar plots in [GJR18]).

Discussion on the rough Heston model, and the Riemann-Liouville process

The driftless rough Heston model satisfies

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

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

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

for $0 \leq s \leq t$, 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.

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

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

in terms of the RL process Z above (and note that A_t is known at time zero for all $t \geq 0$), and $c_H = (\frac{2H \Gamma(\frac{3}{2}-H)}{\Gamma(H+\frac{1}{2}) \Gamma(2-2H)})^{\frac{1}{2}}$.

References

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