Asymptotic behaviour of estimators for (H,σ)

Let X be a real-valued zero-mean **stationary** Gaussian process $(X_t)_{t=0}^{\infty}$ (so in particular X_t has the same (Normal) distribution for all $t\in\mathbb{N}$) with a summable **autocovariance function** $r(k):=\mathbb{E}(X_tX_{t+k})$, i.e. $\sum_{k=1}^{\infty}|r(k)|<\infty$ (our interest will be **fractional Gaussian noise** (fGN) $X_n=B_n^H-B_{n-1}^H$, where B^H is fBM so $X_t\sim N(0,1)$ for all t). The **spectral density** of a stationary process X is the function $f_\theta(\omega)$ whose **Fourier series** coefficients are equal to $(r(k))_{k=0}^{\infty}$, i.e.

$$r(k) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega k} f_\theta(\omega)d\omega$$

so $f_\theta(\omega)=\sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$ when the infinite series here converges (which is the case when $r(k)$ is summable). For fGN there is an explicit formula for $f_\theta=f_H$ in Proposition 7.2.9 in [Taq02] (note the [Taq02] form for the spectral density is divided by 2π, and Eq 5.40 in [Ber94] has a spurious factor of $1/(2\pi)$).

The covariance matrix of a stationary process is clearly **Toeplitz** (which makes it fast to compute). The **Whittle approximation** for the determinant and the Inverse of the Covariance matrix Σ of X is

$$\log(\det\Sigma) \;\sim\; \frac{n}{2\pi}\int_{-\pi}^{\pi}\log f_\theta(\omega)\,d\omega \quad,\quad \Sigma^{-1}_{jk} \;\sim\; \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{f_\theta(\omega)}e^{i(j-k)\omega}d\omega \tag{1}$$

as $n\to\infty$ (this is the so-called **Szegö** limit (or Grenander–Szegö theorem) for Toeplitz matrices).

Define the normalized **discrete Fourier transform** (DFT) at frequency $\omega_j=\frac{2\pi j}{n}$ as:

$$Z_n(\omega_j) \;:=\; \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k e^{-i\omega_j k}$$

and we can also define $Z_n(\omega)$ for arbitrary $\omega\in(-\pi,\pi)$ as $\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_ke^{-i\omega k}$. Then can re-write the covariance part of the log likelihood (LL) of X in terms of $Z_n(\omega)$ as

$$\begin{aligned}\frac{1}{2\pi}\sum_{j=1}^{n}\sum_{k=1}^{n}X_j\Big(\int_{-\pi}^{\pi}\frac{1}{f_\theta(\omega)}e^{i(j-k)\omega}d\omega\Big)\,X_k &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{f_\theta(\omega)}\sum_{j,k}e^{i(j-k)\omega}X_jX_k\,d\omega\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{f_\theta(\omega)}\sum_{j,k}e^{ij\omega}X_je^{-ik\omega}X_k\,d\omega\\ &= \frac{n}{2\pi}\int_{-\pi}^{\pi}\frac{1}{f_\theta(\omega)}|Z_n(\omega)|^2\,d\omega \end{aligned}\tag{2}$$

where $|Z_n(\omega)|^2=I_n(\omega,y)=(\sum_{j=1}^{n}e^{ij\omega}X_j)(\sum_{k=1}^{n}e^{-ik\omega}X_k)=|\sum_{j=1}^{n}e^{ij\omega}X_j|^2$ is the **periodogram** of the random vector X (see plot below for fGN), and we can trivially verify that $I_n(\omega,y)$ is symmetric in ω.

Asymptotic independence of the DFT

For $\omega_j\neq\omega_k$, the coefficients $Z_n(\omega_j)$ and $Z_n(\omega_k)$ are asymptotically uncorrelated (we numerically test this in Python here:
`https://colab.research.google.com/drive/1ruxvZX8brSOKGTi5RFTznRNj0wpQlVUc?usp=sharing)`
Sketch proof:

$$\mathrm{Cov}(Z_n(\omega_j),\bar{Z}_n(\omega_k)) \;=\; \mathbb{E}[Z_n(\omega_j)\overline{Z_n(\omega_k)}] \;=\; \frac{1}{n}\sum_{s=1}^{n}\sum_{t=1}^{n}\mathbb{E}[X_sX_t]e^{-i\omega_j s}e^{i\omega_k t} \;=\; \frac{1}{n}\sum_{s=1}^{n}\sum_{t=1}^{n}r(t-s)e^{-i\omega_j s}e^{i\omega_k t}.$$

Now let $h=t-s$, so $t=s+h$. Then we see that

$$\mathrm{Cov}(Z_n(\omega_j),\bar{Z}_n(\omega_k)) \;=\; \frac{1}{n}\sum_{s=1}^{n}\sum_{h=-(s-1)}^{n-s}r(h)e^{-i\omega_j s}e^{i\omega_k(s+h)} \;=\; \frac{1}{n}\sum_{s=1}^{n}e^{i(\omega_k-\omega_j)s}\sum_{h=-(s-1)}^{n-s}r(h)e^{i\omega_k h}.$$

As $n\to\infty$, and since $r(h)$ is absolutely summable, we see that

$$\sum_{h=-(s-1)}^{n-s}r(h)e^{i\omega_k h} \;\to\; \sum_{h=-\infty}^{\infty}r(h)e^{i\omega_k h} \;=\; f_\theta(\omega_k)\,.$$

Thus $\mathrm{Cov}(Z_n(\omega_j),\bar{Z}_n(\omega_k))\approx\frac{1}{n}\sum_{s=1}^{n}e^{i(\omega_k-\omega_j)s}f_\theta(\omega_k)$. This is a geometric series with partial sum $\frac{1}{n}\sum_{s=1}^{n}e^{i(\omega_k-\omega_j)s}=\frac{1}{n}\cdot\frac{e^{i(\omega_k-\omega_j)}(1-e^{in(\omega_k-\omega_j)})}{1-e^{i(\omega_k-\omega_j)}}$ which is clearly 1 if $j=k$, or tends to zero if $j\neq k$, as claimed.

Remark 0.1 With a bit more analysis, one can also show that $\omega_j\in(0,\pi)$, as $n\to\infty$ $(\Re(Z_n(\omega_j)),\Im(Z_n(\omega_j))\xrightarrow{w}N(\begin{bmatrix}0\\0\end{bmatrix},\frac{1}{2}f(\omega_j)I_2)$ where I_2 is the 2x2 identity matrix.

Using the Whittle approximation to derive the Local Asymptotic Normality (LAN) property

Using (1) and (2), the Whittle approximation for the LL of X is

$$\ell_n(\theta) \;=\; -\frac{1}{2}\log(\det\Sigma) - \frac{1}{2}\mathbf{x}^\top\Sigma^{-1}\mathbf{x} + const. \;=\; -\frac{n}{4\pi}\int_{-\pi}^{\pi}(\log f_\theta(\omega) + \frac{|Z_n(\omega)|^2}{f_\theta(\omega)})d\omega + const.$$

where the log part approximates the determinant part, the Z_n part approximates the covariance part, and the constant is unimportant as it doesn't depend on θ.

The **Whittle estimator** is then defined as argmax of the Whittle approximation for the LL, which we can re-write (using similar notation to [FTW21] and [Szy23] as $\hat{\theta}_n = \arg\min_\theta U_n(\theta)$, where

$$U_n(\theta) \;=\; \frac{1}{4\pi}\int_{-\pi}^{\pi}(\log f_\theta(\omega) + \frac{|Z_n(\omega)|^2}{f_\theta(\omega)})d\omega \tag{3}$$

since the $\frac{n}{4\pi}$ prefactor doesn't affect the maximizer(s), and removing the minus sign here just transforms this to a minimization problem. For fGN we can re-write this in the form

$$U_n(H,\sigma) \;=\; \frac{1}{4\pi}\int_{-\pi}^{\pi}(\log(\sigma^2 f_H(\omega)) + \frac{|Z_n(\omega)|^2}{\sigma^2 f_H(\omega)})d\omega$$

and solving $\frac{\partial}{\partial\sigma}U_n(\hat{H}_n,\sigma) = 0$, we see that

$$\hat{\sigma}_n^2 \;=\; \frac{1}{\pi}\int_0^{\pi}\frac{|Z_n(\omega)|^2}{f_{\hat{H}_n}(\omega)}d\omega$$

so we can reduce to problem to a one-dimensional minimization over H.

Remark 0.2 $U_n(\theta) \to U_\infty(\theta) := \frac{1}{4\pi}\int_{-\pi}^{\pi}(\log(f_\theta(\omega)) + \frac{f_{\theta_0}(\omega)}{f_\theta(\omega)})d\omega = \frac{1}{4\pi}\int_{-\pi}^{\pi}(\log(\sigma^2 f_H(\omega)) + \frac{\sigma^2 f_{H_0}(\omega)}{\sigma^2 f_H(\omega)})d\omega$ under $\mathbb{P}_{\theta_0}$, and this expression is minimized at $\theta = \theta_0$ and if $\arg\min U_n(\theta) \to \arg\min U_\infty(\theta) = \theta_0$, the Whittle estimator $\hat{\theta}_n$ is **consistent**, i.e. $\hat{\theta}_n \to \theta_0$ in probability.

Then the **score** is

$$\ell_n'(\theta) \;=\; -\frac{n}{4\pi}\int_{-\pi}^{\pi}\frac{f_\theta(\omega) - |Z_n(\omega)|^2}{f_\theta(\omega)^2}\nabla_\theta f_\theta(\omega)d\omega\,.$$

$|Z_n(\omega)|^2$ is symmetric in ω (see above), and approximating this integral as a sum over each $\omega_j = \frac{2\pi j}{n}$ we get

$$\ell_n'(\theta) \;\approx\; -\frac{2n\pi}{4n\pi}\sum_{j=1}^{n}\frac{f_\theta(\omega_j) - |Z_n(\omega_j)|^2}{f_\theta(\omega_j)^2}\nabla_\theta f_\theta(\omega_j) \;=\; -\frac{1}{2}\sum_{j=1}^{n}\frac{f_\theta(\omega_j) - |Z_n(\omega_j)|^2}{f_\theta(\omega_j)^2}\nabla_\theta f_\theta(\omega_j)\,. \tag{4}$$

From the result in previous section, $Z_n(\omega)$ is a sequence of asymptotically independent Normal random variables and hence $f_\theta(\omega_j) - |Z_n(\omega_j)|^2$ is a sequence of approximately independent shifted $\chi^2(df = 2)$ random variables.

Hence using a Lyapunov-type CLT, this sum has expectation zero, and variance equal to

$$(\frac{2}{4})^2\sum_{j=1}^{n}\frac{2f_\theta(\omega_j)^2}{f_\theta(\omega_j)^4}\nabla f_\theta(\omega_j)^\top\nabla f_\theta(\omega_j) \;\approx\; \frac{1}{2}\cdot\frac{n}{\pi}\int_0^{\pi}(...)d\omega \;=\; \frac{n}{4\pi}\int_{-\pi}^{\pi}\frac{\nabla_\theta f_\theta(\omega)^\top\nabla_\theta f_\theta(\omega)}{f_\theta(\omega)^2}d\omega \;=\; nI(\theta)$$

where $I(.)$ is the **Fisher information matrix**. n term here then cancels with $(\frac{u}{\sqrt{n}})^2$ when we make perturbation to get the LAN property, as claimed in Cohen et. al.[CGL13].

Asymptotic normality of MLEs

Using the Taylor remainder theorem and setting the answer to zero, we see that

$$\nabla\ell_n(\hat{\theta}_n) \;=\; \nabla\ell_n(\theta_0) + \nabla^2\ell_n(\tilde{\theta}_n)(\hat{\theta}_n - \theta_0) \;=\; 0$$

for some $\tilde{\theta}_n \in [\theta, \theta_n]$, which we can re-arrange as

$$\sqrt{n}\,(\hat{\theta}_n - \theta_0) \;=\; -(\frac{1}{n}\nabla^2\ell_n(\tilde{\theta}_n))^{-1}\frac{1}{\sqrt{n}}\nabla\ell_n(\theta_0)\,.$$

Asymptotic normality of the score

The (approximate) score function under the Whittle likelihood is:

$$\nabla \ell_n(\theta_0) \quad = \quad \int_{-\pi}^{\pi} [\frac{\partial \log f_{\theta_0}(\omega)}{\partial \theta} (\frac{I_n(\omega)}{f_{\theta_0}(\omega)} - 1)] d\omega$$

where $I_n = |Z_n(\omega)|^2$ as before. Recall from above that $\frac{1}{\sqrt{n}} \nabla \ell_n(\theta_0) \xrightarrow{d} N(0, I(\theta_0))$, where the Whittle likelihood approximation for $I(\theta_0)$ is $I(\theta_0) = \frac{1}{4\pi} \int_{-\pi}^{\pi} (\frac{\partial \log f_{\theta_0}(\omega)}{\partial \theta})(\frac{\partial \log f_{\theta_0}(\omega)}{\partial \theta})^\top d\omega$

Hessian convergence to a constant

The Hessian of the Whittle log-likelihood is

$$\nabla^2 \ell_n(\theta_0) \quad = \quad -\int_{-\pi}^{\pi} [\frac{\partial^2 \log f_{\theta_0}(\omega)}{\partial \theta \partial \theta^\top} (1 - \frac{I_n(\omega)}{f_\theta(\omega)}) + (\frac{\partial}{\partial \theta} \log f_{\theta_0}(\omega))^\top (\frac{\partial}{\partial \theta} \log f_{\theta_0}(\omega)) \frac{I_n(\omega)}{f_{\theta_0}(\omega)}] \, d\omega \, .$$

Using that $\mathbb{E}_{\theta_0}(I_n(\omega)) = f_{\theta_0}(\omega)$, the expectation of this expression is just $-I(\theta_0)$, and (under ergodicity and mixing condition) we have the convergence $\frac{1}{n} \nabla^2 \ell_n(\tilde{\theta}_n) \xrightarrow{p} -I(\theta_0)$ where $\tilde{\theta}_n$ lies between $\hat{\theta}_n$ and θ_0, and the Fisher information $I(\theta_0)$ is as defined above. By Slutsky's theorem, we conclude that $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I^{-1}(\theta_0))$.

Estimating (H, σ) under high-frequency observations

The high-frequency (HF) regime corresponds to observations of the process $X = (\sigma B^H_{T/n}, \sigma B^H_{2T/n}, ..., \sigma B^H_T)$ (with $\theta = (H, \sigma)$ unknown (as in Part 2, where ν plays the role of σ), and W.L.O.G. we set $T = 1$. Note that $Y_j = \sigma n^H (B^H_{j/n} - B^H_{(j-1)/n})$ is a standard fGN, but the issue now is that the true H is unknown, so the Y process here is unobserved.

Without any add-on noise, we can easily verify that the true MLE is **scale-independent** (since there is an explicit expression for the MLE for $\hat{\sigma}$, see FM14 2023 chapter2, and the **Whittle estimator** for H (defined above, which we will henceforth denote by $\hat{H}_n$) is also scale-independent (assuming σ is unknown)[1].

To see this, for fGN we can re-write $U_n(\theta)$ in (3) in the form

$$U_n(\theta) \quad = \quad \frac{1}{4\pi} \int_{-\pi}^{\pi} (\log(\sigma^2 f_H(\omega)) + \frac{|Z_n(\omega)|^2}{\sigma^2 f_H(\omega)}) d\omega \, .$$

Minimizing the integrand in σ we find that $\hat{\sigma}^2_n = \int_{-\pi}^{\pi} \frac{|Z_n(\omega_1)|^2}{f_H(\omega_1)} d\omega_1$, and evaluating the integrand at $\hat{\sigma}_n$, we see that

$$U_n(H, \hat{\sigma}^2_n) \quad = \quad \frac{1}{4\pi} \int_{-\pi}^{\pi} d\omega \cdot \log(\int_{-\pi}^{\pi} \frac{|Z_n(\omega_1)|^2}{f_H(\omega_1)} d\omega_1) \; + \; \frac{1}{4\pi} \int_{-\pi}^{\pi} (\log f_H(\omega) + \frac{|Z_n(\omega)|^2}{\int_{-\pi}^{\pi} \frac{|Z_n(\omega_1)|^2}{f_H(\omega_1)} d\omega_1 \cdot f_H(\omega)}) d\omega$$

which only changes by constant (which doesn't depend on H) if we multiply Z_n by a constant; thus we still obtain the same $\hat{H}_n$ when we minimize this expression over H, so $\hat{H}_n$ is scale-independent as claimed.

Thus for the high-frequency regime, $\hat{H}_n$ has same behaviour as for original regime, and in particular $\hat{H}_n$ tends asymptotically to a Normal RV with variance equal to the $(1, 1)$ component of the inverse of the Fisher information matrix:

$$I(H, \sigma) \quad = \quad \frac{1}{4\pi} \begin{bmatrix} \int_{-\pi}^{\pi} (\frac{\partial}{\partial H} \log f_\theta(\omega))^2 d\omega & \int_{-\pi}^{\pi} \frac{\partial}{\partial H} \log f_H(\omega) \frac{\partial}{\partial \sigma} \log f_\theta(\omega) d\omega \\ \int_{-\pi}^{\pi} \frac{\partial}{\partial H} \log f_H(\omega) \frac{\partial}{\partial \sigma} \log f_\theta(\omega) d\omega & \int_{-\pi}^{\pi} (\frac{\partial}{\partial \sigma} \log f_\theta(\omega))^2 d\omega \end{bmatrix} .$$

Then using that $f_\theta(\omega) = f_\theta((H, \sigma)) = \sigma^2 f_H(\omega)$, this simplifies to

$$I(H, \sigma) \quad = \quad \frac{1}{4\pi} \begin{bmatrix} \int_{-\pi}^{\pi} (\frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)})^2 d\omega & \frac{2}{\sigma} \int_{-\pi}^{\pi} \frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)} d\omega \\ \frac{2}{\sigma} \int_{-\pi}^{\pi} \frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)} d\omega & \int_{-\pi}^{\pi} \frac{4}{\sigma^2} d\omega \end{bmatrix} \quad = \quad \begin{bmatrix} \frac{1}{4\pi} \int_{-\pi}^{\pi} (\frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)})^2 d\omega & \frac{1}{2\pi\sigma} \int_{-\pi}^{\pi} \frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)} d\omega \\ \frac{1}{2\pi\sigma} \int_{-\pi}^{\pi} \frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)} d\omega & \frac{2}{\sigma^2} \end{bmatrix}$$

(we can compute these integrals in Mathematica with the `NIntegrate` command). Using that the inverse of a symmetric 2×2 matrix $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, is $A^{-1} = \frac{1}{ad - b^2} \begin{bmatrix} d & -b \\ -b & a \end{bmatrix}$. we see in particular that

$$(I(H, \sigma)^{-1})_{1,1} \quad = \quad \frac{1}{a - \frac{b^2}{d}} \quad = \quad (\frac{1}{4\pi} \int_{-\pi}^{\pi} (\frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)})^2 d\omega - \frac{1}{8\pi^2} (\int_{-\pi}^{\pi} \frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)} d\omega)^2)^{-1} \tag{5}$$

which agrees with Theorem 1 in [Szy23]. Note $(I(H, \sigma)^{-1})_{1,1} > \frac{1}{I(H, \sigma)_{1,1}} = \frac{1}{\frac{1}{4\pi} \int_{-\pi}^{\pi} (\frac{\frac{\partial}{\partial H} f_H(\omega)}{f_H(\omega)})^2 d\omega}$, which is the asymptotic variance for $\hat{H}$ if σ were known, so σ being unknown adds to the variance of $\hat{H}_n$ as we intuitively expect.

[1] As an aside we remark the Han-Schied [HS21] estimator for H is not scale-independent but has the advantage of being model agnostic

The Szymanski asymptotic normality result for the Whittle estimator under high frequency observations

We can define a re-scaled estimator $\hat{\nu}_n$ as

$$\hat{\sigma}_n \quad = \quad n^{\hat{H}_n}\hat{\nu}_n$$

and set $\tilde{\sigma}_n = n^H\hat{\nu}_n$, where $\hat{H}_n$ is the (scale-independent) Whittle estimator for H for X and $\hat{\nu}_n$ is the Whittle estimator of the multiplicative constant for the observed "time-stretched" process $\tilde{Y}_j = X_{j/n} = \sigma(B^H_{j/n} - B^H_{(j-1)/n})$ (note $\tilde{Y} = \nu\tilde{X}$ where $\tilde{X}$ is an fGN with $\nu \ll 1$ for n large), and we note that $\hat{\sigma}_n$ *is* computable from observations of X. Moreover, we can re-write $\hat{\sigma}_n$ as $\hat{\sigma}_n = n^{\hat{H}_n - H}\tilde{\sigma}_n$ and we know that

$$\sqrt{n}(\hat{H}_n - H) \quad \to \quad N(0, (I^{-1})_{1,1})$$

from the standard theory for the non high-frequency regime above since (as discussed in previous section) $\hat{H}$ is unaffected by switching between the HF and the original regime. Hence (following the top of page 13 in [Szy23] (after correcting some typos there) we see that

$$\begin{aligned}
\frac{\sqrt{n}}{\log n}(\hat{\sigma}_n - \sigma) &= \frac{\sqrt{n}}{\log n}(\tilde{\sigma}_n - \sigma) + \frac{\sqrt{n}}{\log n}(\hat{\sigma}_n - \tilde{\sigma}) \\
&= \frac{\sqrt{n}}{\log n}(\tilde{\sigma}_n - \sigma) + \frac{\sqrt{n}}{\log n}\tilde{\sigma}_n(n^{\hat{H}_n - H} - 1) \\
&= \frac{\sqrt{n}}{\log n}(\tilde{\sigma}_n - \sigma) + \frac{\sqrt{n}}{\log n}\tilde{\sigma}_n \log n\,(\hat{H}_n - H) + O((\hat{H}_n - H)^2) \\
&= \frac{\sqrt{n}}{\log n}(\tilde{\sigma}_n - \sigma) + \tilde{\sigma}_n\sqrt{n}\,(\hat{H}_n - H) + O((\hat{H}_n - H)^2)
\end{aligned}$$

where we have used the Taylor expansion $n^x - 1 = x\log n + O(x^2)$ in the penultimate line, with $x = \hat{H}_n - H$ here. Then we see that the first term here is a higher order (i.e. smaller) term than the second term because $\sqrt{n}(\tilde{\sigma}_n - \sigma)$ is asymptotically Normal from the standard theory above for the non high-frequency regime because $\tilde{\sigma}_n$ behaves the same as the usual Whittle estimator $\hat{\sigma}_n$ in the non-HF regime. Moreover, since $\tilde{\sigma}_n$ is a **consistent estimator**, $\tilde{\sigma}_n = \sigma(1 + o(1))$, so (at leading order) we find that

$$\frac{\sqrt{n}}{\log n}(\hat{\sigma}_n - \sigma) \quad = \quad \sigma\sqrt{n}\,(\hat{H}_n - H) \quad \to \quad N(0, \sigma^2(I^{-1})_{1,1})$$

where $(I^{-1})_{1,1} = (I(H,\sigma)^{-1})_{1,1}$ was computed in (5), which also agrees with Theorem 1 in [Szy23] (note for us the $\gamma(H)$ function in [Szy23] is just $\gamma(H) = H$, see Remark just above Section 1.2 in [Szy23]).

Remark 0.3 Note that $\hat{\sigma}_n - \sigma$ now has a larger standard deviation i.e. $O(\frac{\log n}{\sqrt{n}})$ as opposed to just $O(\frac{1}{\sqrt{n}})$ due to the HF regime.

Singular Fisher information for the high-frequency regime

If we naively try to perform a LAN analysis using the Whittle approximation in the HF regime, we get

$$U_n(H,\sigma) \quad = \quad \frac{1}{4\pi}\int_{-\pi}^{\pi}(\log(\frac{\sigma^2 f_H(\omega)}{n^{2H}}) + \frac{|Z_n(\omega)|^2}{\sigma^2 f_\theta(\omega)/n^{2H}})d\omega\,.$$

Then

$$\begin{aligned}
\frac{\partial}{\partial H}U_n(H,\sigma) &= \frac{1}{2\pi}\int_0^{\pi}(\frac{n^{2H}|Z_n(\omega)|^2}{\sigma^2 f_H(\omega)} - 1)(2\log n - \frac{\frac{\partial}{\partial H}f_H(\omega)}{f_H(\omega)})d\omega \\
&= \frac{1}{2\pi}\int_0^{\pi}(\frac{n^{2H}|Z_n(\omega)|^2}{\sigma^2 f_H(\omega)} - 1)d\omega \cdot 2\log n\,(1 + o(1))
\end{aligned}$$

as $n \to \infty$, which (using the same argument as in (4)) has asymptotic variance

$$\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{2\sigma^4 f_H(\omega)^2}{\sigma^4 f_H(\omega)^2}d\omega \cdot 2(\log n)^2\,(1 + o(1)) \quad \sim \quad 2(\log n)^2\,.$$

Similarly

$$\frac{\partial}{\partial\sigma}U_n(H,\sigma) \quad = \quad \frac{1}{\pi\sigma^3}\int_0^{\pi}(\sigma^2 - \frac{n^{2H}|Z_n(\omega)|^2}{f_H(\omega)})d\omega\,.$$

Then $\mathbb{E}(\frac{\partial}{\partial H}U_n(H,\sigma)\frac{\partial}{\partial\sigma}U_n(H,\sigma))$ tends to $\frac{2}{\sigma}\log n$ and $\mathbb{E}((\frac{\partial}{\partial\sigma}U_n(H,\sigma))^2)$ tends to $\frac{2}{\sigma^2}$ (this agree with the matrix entries at the end of page 2 in [BF18] when we scale out the $\log n$; see also Theorem 2.1 and the lower part of the matrix above it in [Kaw13] where the H and σ entries are swapped around). Hence the associated Fisher information here is **singular** (i.e. zero determinant) so we cannot apply a **minmax theorem**, which is why we need the alternative Szymanski approach with $\hat{\nu}_n$ above.

Note that $(\hat{H}_n,\hat{\sigma}_n)$ satisfy

$$\begin{aligned}\int_0^\pi (\frac{n^{2\hat{H}_n}|Z_n(\omega)|^2}{\sigma^2 f_{\hat{H}_n}(\omega)}-1)(2\log n-\frac{\frac{\partial}{\partial H}f_{\hat{H}_n}(\omega)}{f_{\hat{H}_n}(\omega)})d\omega &= 0\\ \hat{\sigma}_n^2 &= \frac{1}{\pi}\int_0^\pi \frac{n^{2\hat{H}_n}|Z_n(\omega_1)|^2}{f_{\hat{H}_n}(\omega_1)}d\omega_1\end{aligned}$$

but substituting $\hat{\sigma}$ into the first equation, two π terms cancel and we are left with

$$\int_0^\pi (\frac{n^{2\hat{H}_n}|Z_n(\omega)|^2}{\sigma^2 f_{\hat{H}_n}(\omega)}-1)\frac{\frac{\partial}{\partial H}f_{\hat{H}_n}(\omega)}{f_{\hat{H}_n}(\omega)}d\omega = 0$$

which is the equation for the usual Whittle MLE in the non-HF case (as we expect since $\hat{H}_n$ is scale-dependent). Check if $\hat{\sigma}_n^2$ is the same estimator proposed in Theorem 1 in [Szy23].

TO DO: Consider stationary fBM X (**s-fBM**) as discussed in [BW22] and Hwk1 q3 in FM14 2024 with additional ξ param since V_0 is random now for rBergomi model driven by S-fBM with $V_t=\xi e^{\sigma X_t}$ (spectral density $f_H(\omega)$ appears to be wobbly using Fourier series..). Or add drift parameter as in [Kaw13], or a second fBM with $H>\frac{1}{2}$ to incorporate long-range dependence.

Adding additive noise

To incorporate the **microstructure noise** which arises from using realized variance to estimate B^H under high-frequency observations, we add $\sqrt{\frac{2}{n}}$ times a standard Gaussian to the original observed νB^H (see CLT part of Brownian motion chapter in FM02 to see where the $\frac{2}{n}$ comes from; this has nothing to do with fBM), or equivalently (if we work with increments instead as we do above) we add $\sqrt{\frac{2}{n}}Y$ to fGN observed on $[0,1]$, where $Y_t=\varepsilon_t-\varepsilon_{t-1}$ and the $\varepsilon_t' s$ are i.i.d. standard Normals (note Y is an MA(1) process of the form $Y_t=\mu+\varepsilon_t+\theta\varepsilon_{t-1}$ with $\theta=-1$ here). Then

$$\mathbb{E}(Y_sY_t) = \mathbb{E}((\varepsilon_t-\varepsilon_{t-1})(\varepsilon_s-\varepsilon_{s-1}))$$

so Y is also a stationary Gaussian process.

Cramer-Rao bound for biased estimators

For an estimator with **bias** $b(\theta)$, Cramer-Rao bound for MSE is

$$\mathbb{E}((\hat{\theta}-\theta)^2) \geq \frac{(1+b'(\theta))^2}{I(\theta)}+b(\theta)^2$$

Correction to Whittle approximation

Let C_n^{-1} denote the Whittle approx to Σ_n^{-1}, and $\Delta_n=\Sigma_n-C_n$. For small Δ_n, we have the standard matrix expansions:

$$\begin{aligned}\Sigma_n^{-1} &= (C_n+\Delta_n)^{-1} = C_n^{-1}-C_n^{-1}\Delta_nC_n^{-1}+C_n^{-1}\Delta_nC_n^{-1}\Delta_nC_n^{-1}+o(\|\Delta_n\|^2),\\ \log\det(C_n+\Delta_n) &= \log\det C_n+\mathrm{trace}(C_n^{-1}\Delta_n)-\frac{1}{2}\mathrm{trace}((C_n^{-1}\Delta_n)^2)+o(\|\Delta_n\|^2).\end{aligned}$$

Note that C_n can also be computed as

$$C_n(f) = F^*\Lambda_fF$$

where

$$\Lambda_f = \mathrm{diag}\big(f(\lambda_j;\theta)\big),\quad \lambda_j=\frac{2\pi j}{n},\quad j=0,\dots,n-1$$

(where diag means the diagonal matrix with the given entries along the diagonal), and the DFT matrix F is $F=\frac{1}{\sqrt{n}}[e^{-i\lambda_jk}]_{j,k=0}^{n-1}$ i.e. $F_{j,k}=\frac{1}{\sqrt{n}}e^{-i\frac{2\pi}{n}jk}$. However Σ_n and hence Δ_n also expensive to re-compute for n large for different H-values when minimizing.

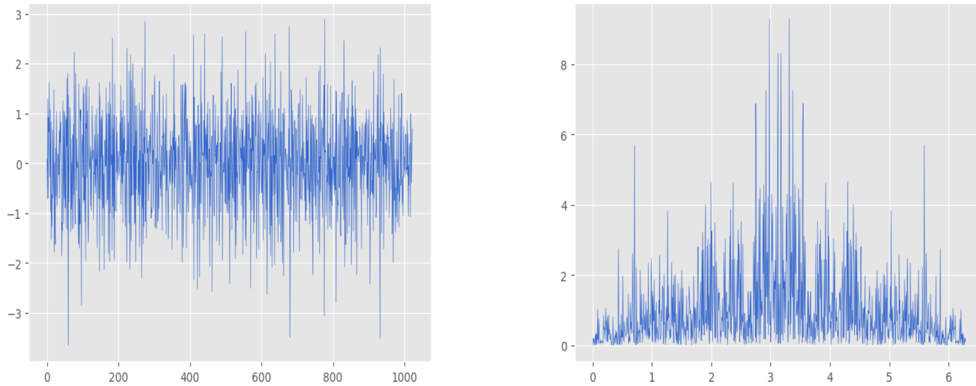

Figure 1: Simulation of fGN (left) and its periodogram (right) for $H = .25$

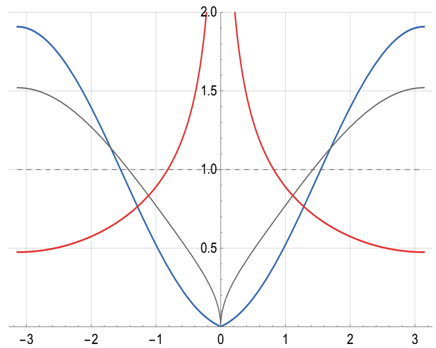

Figure 2: Spectral density $f_H(\omega)$ of fGN for $H = .1$ (blue), $H = .25$ (grey), $H = .5$ (grey dashed) and $H = .75$ (red)

.

References

[Ber94] Beran, J., "Statistics for Long-Memory Processes," *Chapman and Hall/CRC*, New York, 1994.

[BW22] Bacry, E. and Wu, D., "S-fractional Brownian motion: Definition and properties," *Stochastic Processes and their Applications*, vol. 152, pp. 585–614, 2022.

[CGL13] Cohen, S., F. Gamboa, C. Lacaux, and J.-M. Loubes, "LAN property for some fractional type Brownian motion," *ALEA, Lat. Am. J. Probab. Math. Stat.*, 10 (2013), 91–106.

[BF18] Brouste, A. and Fukasawa, M., "Local asymptotic normality property for fractional Gaussian noise under high-frequency observations", *Statistical Inference for Stochastic Processes*, 2018.

[FTW21] M.Fukasawa, T.Takabatake, and R.Westphal, "Is volatility rough?", *Quantitative Finance*, vol.21, no.11, pp. 1769–1783, 2021.

[HS21] Han, X. and A.Schied, "The roughness exponent and its model-free estimation," arXiv preprint

[Kaw13] Kawai, R., "Fisher Information for Fractional Brownian Motion Under High-Frequency Discrete Sampling", *Communications in Statistics – Theory and Methods*, vol. 42, no. 9, pp. 1628–1636, 2013.

[Szy23] Szymanski, G., "Optimal Estimation of the Rough Hurst parameter in additive noise," *Electronic Journal of Statistics*, 2023.

[Taq02] Taqqu, M.S., "Theory and Applications of Long-Range Dependence", *Birkhäuser*, 2002.