Estimation of the Hurst parameter from continuous noisy data

P. Chigansky[1] M. Kleptsyna[2]

[1]University of Jerusalem, Israel

[2]University of Le Mans, France

Efficient inference for large and high-frequency data
An international online seminar
5 April 2022

Outline

Outline

Main objects

Definition (fractional Brownian motion — fBm)

Fractional Brownian motion is $B^H = (B_t^H)_{t \geq 0}$, a centered Gaussian process with the covariance function $K(s,t)$

$$K(s,t) = \mathbf{E} B_t^H B_s^H = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \quad s,t \in [0,T],$$

where $H \in]0,1]$ —the Hurst exponent.

and

Definition (Mixed fBm)

$X_t = B_t + B_t^H$, where

- B_t — the standard Brownian motion;
- B_t^H — an independent fractional Brownian motion.

What do we want

Estimate

A problem is to estimate

- the Hurst parameter $H \in (3/4, 1)$
- and the scaling parameter $\sigma \in \mathbb{R}_+$

from the data

Continuous time noisy data (mfBm)

$$X_t = \sigma B_t^H + \sqrt{\varepsilon} B_t,\ t \in [0, T]$$

where

- B^H is the fractional Brownian motion (fBm),
- B is an independent standard Brownian motion,
- $\varepsilon > 0$ is the noise intensity

Questions to answer and to discuss

to answer

- Is it possible to recover exactly the parameters for any finite T ? Why $H > 3/4$?
- How to write the likelihood function?
- Does $(\mathbb{P}^h_\theta)_{\theta\in\Theta}$satisfy the LAN property? In which regime?
- How to find the optimal minimax rates?
- How to construct the rate optimal estimators?

to discuss

- Fisher information: asymptotic or/and "one observation"...
- The rate optimal estimators: coupled/separated

Outline

LAN and asymptotic lower bound

Definition (LAN)

A family of probability measures $(\mathbb{P}_\theta^h)_{\theta\in\Theta}$ is LAN at a point θ_0 as $h \to 0$ if there exists a family of nonsingular $k \times k$ matrices $\phi(h) = \phi(h, \theta_0)$, such that for any $u \in \mathbb{R}^k$

$$\log \frac{d\mathbb{P}^h_{\theta_0+\phi(h)u}}{d\mathbb{P}^h_{\theta_0}}(X^h) = u^\top Z_{h,\theta_0} - \frac{1}{2}\|u\|^2 + r_h(u,\theta_0),$$

where

- Z_{h,θ_0} converges weakly under $\mathbb{P}^h_{\theta_0}$ to the standard normal law on $\mathbb{R}^k$
- $r_h(u,\theta_0)$ vanishes in $\mathbb{P}^h_{\theta_0}$-probability as $h \to 0$.

LAN and asymptotic lower bound

Theorem (Hájek)

Let $(\mathbb{P}_\theta^h)_{\theta\in\Theta}$ satisfy the LAN property at θ_0 with matrices $\phi(h,\theta_0) \to 0$ as $h \to 0$.
Then for any family of estimators $\widehat{\theta}_h$, a loss function ℓ and any $\delta > 0$

$$\varliminf_{h\to 0} \sup_{\|\theta-\theta_0\|<\delta} \mathbb{E}_\theta^h \ell\left(\phi(h,\theta_0)^{-1}(\widehat{\theta}_h - \theta)\right) \geq \int_{\mathbb{R}^k} \ell(x)\gamma(x)dx,$$

where γ is the standard normal density on $\mathbb{R}^k$.

Whittle approach, discrete time

Whittle approach, 1953

- $(X_j, j \in Z)$ — stationary Gaussian sequence with
  1. $\mathbb{E}X_j = 0$, $\mathbb{E}X_j X_0 = K(|j|)$,
  2. **continuous** spectral density $\widehat{K}_\theta(\lambda)$ (The Fourier transform of the covariance function);
- "Whittle's likelihood":

$$\widehat{\theta}_n = \underset{\theta}{\text{argmin}} \int_{-\pi}^{\pi} \Big(\log \widehat{K}_\theta(\lambda) + \frac{I_n(\lambda)}{\widehat{K}_\theta(\lambda)} \Big)\, d\lambda,$$

$$I_n(\lambda) = \left| \frac{1}{\sqrt{n}} \sum_{1}^{n} e^{-i\lambda j} X_j \right|^2$$ — the spectrogram

Whittle approach

It was shown for the "Whittle estimator":

- consistency
- asymptotic normality with rate $n^{-1/2}$ and the limiting variance :

$$J(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\nabla^\top \log \widehat{K}_\theta(\lambda)\right) \nabla \log \left(\widehat{K}_\theta(\lambda)\right) d\lambda$$

- "Whittle formula" for the Fisher information

$$J(\theta) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\, \nabla^\top L_n(X;\theta) \nabla L_n(X;\theta)$$

(for continuous $\widehat{K}_\theta$)

Outline

Singular spectral density, longe range dependence

When $\widehat{K}_\theta(\lambda) = |\lambda|^{-\alpha(\theta)} L(\theta,\lambda), 0 < \alpha(\theta) < 1$:

- Fox R. and Taqqu M. (1986): asymptotic normality, rate of convergence – $n^{-1/2}$
- Dahlhaus R.(1989) - (2006): — asymptotically efficiency

Continuous time

The problem is to estimate the Hurst parameter $H \in (0, 1)$ and the scaling parameter $\sigma \in \mathbb{R}_+$ given the data

$$X^T = (\sigma B_t^H,\ t \in [0, T]).$$

But both parameters can be recovered from X^T exactly for any $T > 0$.

Discrete observations

A meaningful statistical problem is to estimate the Hurst parameter $H \in (0, 1)$ and the scaling parameter $\sigma \in \mathbb{R}_+$ from the discrete data

Discrete data

$$\left(\sigma B^H_{\Delta(n)}, \ldots, \sigma B^H_{n\Delta(n)}\right)$$

where $\Delta(n) > 0$ is the discretization step.

Asymptotic regimes: $n \to \infty$

- the large time asymptotics: $\Delta(n) = \Delta > 0$ is fixed
- the high frequency asymptotics: $\Delta(n) \to 0$ (typically $\Delta(n) = 1/n$)

State-of-the-art

- Large time: Dahlhaus asymptotic normality, rate of convergence – $n^{-1/2}$
- High frequency:
  - J.F.Coeurjolly (2001); J.Istas and G.Lang (1997) for the separate estimation of H and σ^2, rate of convergence:

    $$n^{-1/2}\frac{1}{\log \Delta(n)^{-1}} \quad \text{and} \quad n^{-1/2}$$

  - A. Brouste, M.Fukasawa (2018) : efficient joint estimate, the rates of convergence

    $$n^{-1/2} \quad \text{and} \quad n^{-1/2}\log \Delta(n)^{-1}$$

    were shown to be **asymptotic minimax optimal**

Noisy discrete data

Noisy discrete data

$$\left(\sigma B^H_{\Delta(n)} + \xi_{1,n},\ \dots,\ \sigma B^H_{n\Delta(n)} + \xi_{n,n}\right),$$

where $\xi_{j,n}$ are i.i.d. independent of B^H.

Asymptotic regimes: $n \to \infty$

- the large time asymptotics: $\Delta(n) = \Delta > 0$ is fixed
- the high frequency asymptotics: $\Delta(n) \to 0$ (typically $\Delta(n) = 1/n$)

State-of-the-art

- Large time: Dahlhaus asymptotic normality, rate of convergence – $n^{-1/2}$
- High frequency: A.Gloter and M.Hoffman (2007) :
  - efficient joint estimate, rate of convergence

$$n^{-1/(4H+2)} \quad \text{and} \quad n^{-1/(4H+2)} \log n$$

  - "wavelet" type optimal rate estimator

Continuous time noisy data

Nothing was known about efficient estimate from continuous time noisy data (mfBm)

$$X_t = \sigma B_t^H + \sqrt{\varepsilon} B_t,\ t \in [0, T], H > 3/4.$$

Asymptotic regimes

- the large time asymptotics:
  $\varepsilon > 0$ is fixed and $T \to \infty$
- small noise asymptotics:
  T is fixed and $\varepsilon \to 0$

How to write the likelihood function?

Mixed fBm

$X_t = \sqrt{\varepsilon} B_t + \sigma B_t^H$, where

- B_t — the standard Brownian motion;
- B_t^H — an independent fractional Brownian motion.

The principal object for analisis:

Fundamental Martingale

$$M_t = \mathbb{E}\big(B_t | \mathbb{F}_t^X\big), \quad t \in [0, T].$$

Fundamental Martingale Representation via X

Fundamental Martingale Representation

$$M_t = \int_0^t q(t,s)dX_s, \quad \langle M\rangle_t = \int_0^t q(t,s)ds, \quad t \geq 0.$$

The kernel $q(t,s)$ solves integral equation:

Equation for the Kernel

$$\varepsilon q(t,s) + H(2H-1)\sigma^2 \int_0^t q(t,r)|s-r|^{2H-2}dr = 1.$$

X est **diffusion** type process

X est **diffusion** type process

Let $H \in (\frac{3}{4}, 1]$.Then

- X is a **diffusion** type process:

$$X_t = \sqrt{\varepsilon}\overline{B}_t + \int_0^t \rho_s(X)ds,$$

- $\overline{B}_t = \int_0^t \frac{dM_s}{q(s,s)}$ is an F^X-Brownian motion,
- $\rho_t(X) = \int_0^t \frac{\dot{q}(s,t)}{q(t,t)} dX_s = \int_0^t g(s,t) dX_s$
- where g solves the equation:

$$\varepsilon g(s,t) + c_H \sigma^2 \int_0^t g(r,t)|s-r|^{2H-2} dr = c_H \sigma^2 |t-s|^{2H-2}.$$

The density w.r.t $\mu^{\sqrt{\varepsilon}W}$

The density w.r.t $\mu^{\sqrt{\varepsilon}W}$

Moreover, the measures μ^X and $\mu^{\sqrt{\varepsilon}W}$ are equivalent and

$$\frac{d\mu^X}{d\mu^{\sqrt{\varepsilon}W}}(X) = \exp\left\{\frac{1}{\varepsilon}\int_0^T \rho_t(X)dX_t - \frac{1}{2\varepsilon}\int_0^T \rho_t^2(X)dt\right\}.$$

Main difficulties

- Continuous time analog "Whittle likelihood"
- Continuous time "Whittle formula" for Fisher information
- "Ergodic" properties of $\rho_t(X)$?
- The kernel in the integral equation $K \notin L_1(R)$, $K \notin L_2(R)$. What happened with $g(t,.)$ when $t \to \infty$? (The classical theory of I.Gohberg, I.Feldman does not answer).
- No limit for the solution of the integral equation when $\varepsilon \to 0$.

Outline

Large time asymptotics: optimal rate

Theorem (LAN, large time)

The family $(\mathbb{P}_\theta^T)_{\theta\in\Theta}$ is LAN at any $\theta_0 \in \Theta$ as $T \to \infty$ with

$$\phi(\theta_0, T) = T^{-1/2} I(\theta_0, \varepsilon)^{-1/2},$$

where

- $I(\theta, \varepsilon) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \nabla^\top \log\left(\varepsilon + \widehat{K}_\theta(\lambda)\right) \nabla \log\left(\varepsilon + \widehat{K}_\theta(\lambda)\right) d\lambda,$
- $\widehat{K}_\theta(\lambda) = \sigma^2 \Gamma(2H+1) \sin(\pi H) |\lambda|^{1-2H}, \quad \lambda \in \mathbb{R} \setminus \{0\},$

or, $\phi(\theta_0, T) = T^{-1/2}\mathrm{Id}$
with "Fisher information" $I(\theta_0, \varepsilon)$
($I(\theta_0, \varepsilon)$ looks like the "Whittle formula".)

Large time asymptotics: optimal rate estimator

- Base the estimation on the discrete data $X_{k\Delta}$, $k = 1, ..., [T/\Delta]$, $\Delta > 0$
- Use the simpler Whittle's spectral estimator

Such estimators achieve the optimal $T^{-\frac{1}{2}}$ rate and the limit risks can be made arbitrarily close to the bound provided by Theorem by choosing Δ small enough.

Outline

Small noise asymptotics, optimal rate

Define the matrix

$$M(\varepsilon,\theta) = \begin{pmatrix} 1 & -2\sigma^2 \log \varepsilon^{-1/(2H-1)} \\ 0 & 1 \end{pmatrix}.$$

Theorem (LAN, small noise)

Assume that $\phi(\varepsilon,\theta_0)$ satisfies the scaling condition

$$\varepsilon^{-1/(2H_0-1)}\phi(\varepsilon,\theta_0)^\top M(\varepsilon,\theta_0) \mathcal{I}(\theta_0;1) M(\varepsilon,\theta_0)^\top \phi(\varepsilon,\theta_0) \xrightarrow[\varepsilon\to 0]{} \mathrm{Id}.$$

Then the family $(\mathbb{P}^\varepsilon_\theta)_{\theta\in\Theta}$ is LAN at any $\theta_0\in\Theta$ as $\varepsilon\to 0$.

Small noise asymptotics, comments

Remarks

- $\phi(\varepsilon, \theta_0)$ cannot be diagonal (analogous to singularity of "Fisher information": A. Brouste, M.Fukasawa (2018))
- The choice of $\phi(\varepsilon, \theta_0)$ is not unique: for example, we can take

$$\phi(\varepsilon, \theta_0) = \frac{\varepsilon^{1/(4H_0-2)}}{\sqrt{T}} \begin{pmatrix} 1 & 0 \\ 2\sigma^2 \log \varepsilon^{-1/(2H-1)} & 1 \end{pmatrix} I(\theta_0, 1)^{-1/2}$$

1. Different choices of $\phi(\varepsilon, \theta_0)$ give lower bounds for different linear combinations of individual errors (risks)
2. Upper/lower Cholesky decompositions give bounds for individual components

Optimal rates of H and σ^2, small noises

- The optimal minimax estimation rates of H and σ^2 follows from the upper and lower triangular Cholesky decompositions of the matrix $M(\varepsilon,\theta_0)I(\theta_0;1)M(\varepsilon,\theta_0)^\top$:

Optimal rates of H and σ^2

$$\frac{\varepsilon^{1/(4H_0-2)}}{J_H(\theta)} \quad \text{and} \quad \frac{\varepsilon^{1/(4H_0-2)}}{J_{\sigma^2}(\theta)} \log(\varepsilon^{-1}),$$

where

$$J_H(\theta) = \sqrt{T\left(I_{11}(\theta;1) - \frac{I_{12}(\theta;1)^2}{I_{22}(\theta;1)}\right)}.$$

and

$$J_{\sigma^2}(\theta) = \frac{H-\frac{1}{2}}{\sigma^2} J_H(\theta)$$

Small noise, one parameter is unknown

Theorem (one parameter is to be estimated)

1) For any fixed $\sigma_0^2 \in \mathbb{R}_+$, *the family* $\left(\mathbb{P}^{\varepsilon}_{(H,\sigma_0^2)}\right)_{H\in(\frac{3}{4},1)}$ *is LAN at any* $H_0 \in (\frac{3}{4}, 1)$ *as* $\varepsilon \to 0$ *with*

$$\phi(\varepsilon, H_0) := \varepsilon^{1/(4H_0-2)} \frac{1}{\log \varepsilon^{-1}} \frac{H_0 - \frac{1}{2}}{\sigma_0^2} \frac{1}{\sqrt{TI_{22}(\theta_0; 1)}}.$$

2) For any fixed $H_0 \in (\frac{3}{4}, 1)$, *the family* $\left(\mathbb{P}^{\varepsilon}_{(H_0,\sigma^2)}\right)_{\sigma^2\in(0,\infty)}$ *is LAN at any* $\sigma_0^2 \in \mathbb{R}_+$ *as* $\varepsilon \to 0$ *with*

$$\phi(\varepsilon, \sigma_0) := \varepsilon^{1/(4H_0-2)} \frac{1}{\sqrt{TI_{22}(\theta_0; 1)}}.$$

The "Fisher information" in this model — ???
May be, if we take :

$$\phi(\varepsilon, \theta_0) = \frac{\varepsilon^{1/(4H_0-2)}}{\sqrt{T}} \begin{pmatrix} 1 & 0 \\ 2\sigma^2 \log \varepsilon^{-1/(2H-1)} & 1 \end{pmatrix}$$

with "Fisher information" $I(\theta_0, 1)$???

Small noise asymptotics: optimal rate joint estimator

We follow the ideas of A.Gloter and M.Hoffman (2007).

Wavelet method objects

- $J_\varepsilon^* = \max\left\{1 \le j \le [2\log_2 \varepsilon^{-1}] : \widehat{Q}_j \ge 2^j \varepsilon\right\}$
- $\widehat{H}_j = 1 - \frac{1}{2}\log_2 \frac{\widehat{Q}_{j+1}}{\widehat{Q}_j}$
- $\widehat{Q}_j = \sum_{k=0}^{2^{j-1}-1} \widetilde{d_{j,k}}^2 - \varepsilon\|\psi\|^2$—the energy of the j-th resolution level
- $\widetilde{d_{j,k}} = \int_{\mathbb{R}} \psi_{j,k}(t)dX_t$—estimations for the wavelet coefficients of σB^H
- $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \quad j \in \mathbb{N},\ k \in \mathbb{Z}$ — translates and dilations of any mother wavelet function ψ with compact support and two vanishing moments.

Small noise asymptotics: optimal rate joint estimator

Let

$$\widehat{H}(\varepsilon) = \widehat{H}_{J_\varepsilon^*}$$

and

$$\widehat{\sigma}^2(\varepsilon) = \frac{2}{c_{\widehat{H}(\varepsilon)}(\psi)} \frac{\widehat{Q}_{J_\varepsilon^*}}{2^{J_\varepsilon^*(2-2\widehat{H}(\varepsilon))}},$$

Proposition

The estimations errors $\varepsilon^{-1/(4H-2)}(\widehat{H}(\varepsilon) - H)$ *and*

$$\varepsilon^{-1/(4H-2)} \frac{1}{\log \varepsilon^{-1}} \left(\widehat{\sigma}^2(\varepsilon) - \sigma^2\right)$$

are bounded in $\mathbb{P}_\theta$*-probability, uniformly over compacts in* Θ*, as* $\varepsilon \to 0$*.*

Outline

Recall

Continuous time noisy data

$$X_t = \sigma B_t^H + \sqrt{\varepsilon} B_t,\ t \in [0, T],\ H \in (\frac{3}{4}, 1].$$

X is a **diffusion** type process:

$$X_t = \sqrt{\varepsilon}\overline{B}_t + \int_0^t \rho_s^{\varepsilon}(X, \theta)ds,$$

where

- $\rho_t^{\varepsilon}(X, \theta) = \int_0^t g_{\varepsilon}(t, t-s)dX_s$
- g_{ε} solves the equation:

$$\varepsilon g_{\varepsilon}(t, s) + \sigma^2 c_H \int_0^t g_{\varepsilon}(t, r)|s - r|^{2H-2} dr = \sigma^2 c_H |s|^{2H-2}.$$

Recall

The probability measure $\mathbb{P}_\theta^T$ and $\mathbb{P}_0^T$ induced by $\sqrt{\varepsilon}\overline{B}$ are mutually absolutely continuous and

$$\frac{d\mathbb{P}_\theta^T}{d\mathbb{P}_0^T}(X^T) = \exp\left(\frac{1}{\varepsilon}\int_0^T \rho_t^\varepsilon(X,\theta)dX_t - \frac{1}{2}\frac{1}{\varepsilon}\int_0^T \rho_t^\varepsilon(X,\theta)^2 dt\right).$$

The likelihood ratio takes the form

$$\log\frac{d\mathbb{P}^T_{\theta_0+\phi(T)^{-1}u}}{d\mathbb{P}^T_{\theta_0}}(X^T) =$$

$$\frac{1}{\sqrt{\varepsilon}}\int_0^T \left(\rho_t^\varepsilon(X,\theta_0+\phi(T)^{-1}u) - \rho_t^\varepsilon(X,\theta_0)\right)d\overline{B}_t$$

$$-\frac{1}{2}\frac{1}{\varepsilon}\int_0^T \left(\rho_t^\varepsilon(X,\theta_0+\phi(T)^{-1}u) - \rho_t^\varepsilon(X,\theta_0)\right)^2 dt,$$

What should we check?

"Ergodic" type conditions to check

1.
$$\frac{1}{\varepsilon}\frac{1}{T}\int_0^T \nabla^\top \rho_t(X,\theta_0)\nabla\rho_t(X,\theta_0)dt \xrightarrow[T\to\infty]{L^2(\Omega)} I(\theta_0;\varepsilon)$$

2. for all sufficiently small $\delta > 0$,
$$\frac{1}{T^2}\int_0^T \sup_{\theta:\|\theta-\theta_0\|\leq\delta} \mathbb{E}\left\|\nabla^2\rho_t(X,\theta)\right\|^2 dt \xrightarrow[T\to\infty]{} 0.$$

How to check?

It is sufficient to

1. obtain the decomposition:

$$\mathbb{E}\,\nabla^\top \rho_s(X;\theta_0)\nabla\rho_t(X;\theta_0) = Q(t-s) + R(s,t)$$

   where

   - $Q(0) = \varepsilon I(\theta_0;\varepsilon)$
   - $\big\|R(s,t)\big\| \le C\Big(t^{-1/2} + s^{-1/2} + (st)^{-b}\Big), \quad \forall s,t \in [T_{\min},\infty)$
   - $\big\|Q(t-s)\big\| \le C \wedge |t-s|^{-1}\log^2|t-s|, \quad \forall s,t \in \mathbb{R}_+$

   with constants $b \in (0,\frac{1}{2})$, $C > 0$ and $T_{\min} > 0$.

2. check that

$$\sup_{\theta:\|\theta-\theta_0\|\le\delta} \mathbb{E}\big\|\nabla^2\rho_t(X,\theta)\big\|^2 \le C, \quad \forall t \ge T_{\min}.$$

Try to "solve" the integral equation

- Roughly speaking, we will find the "asymptotically stationary ergodic" version of ρ^ε:

$$\rho_t^\varepsilon(X,\theta) = \int_0^t g_\varepsilon(t, t-s)dX_s$$

$$\widetilde{\rho}_t^\varepsilon(X,\theta) = \int_{-\infty}^t g_\varepsilon(t-s)dX_s$$

- Solution of the integral equation can be decomposed into the sum of the principle term independent of t: g_ε , and the residual term which vanishes with $t \to \infty$.

$$\varepsilon g_\varepsilon(t,s) + \sigma^2 c_H \int_0^t g_\varepsilon(t,r)|s-r|^{2H-2}dr = \sigma^2 c_H |s|^{2H-2}.$$

and

$$\varepsilon g_\varepsilon(s) + \sigma^2 c_H \int_0^\infty g_\varepsilon(r)|s-r|^{2H-2}dr = \sigma^2 c_H |s|^{2H-2}.$$

Try to "solve" the integral equation

- Roughly speaking, we will find the "asymptotically stationary ergodic" version of ρ^ε:

$$\rho_t^\varepsilon(X,\theta) = \int_0^t g_\varepsilon(t,t-s)dX_s$$

$$\widetilde{\rho}_t^\varepsilon(X,\theta) = \int_{-\infty}^t g_\varepsilon(t-s)dX_s$$

- Solution of the integral equation can be decomposed into the sum of the principle term independent of t: g_ε , and the residual term which vanishes with $t \to \infty$.

$$\varepsilon g_\varepsilon(t,s) + \sigma^2 c_H \int_0^t g_\varepsilon(t,r)|s-r|^{2H-2}dr = \sigma^2 c_H |s|^{2H-2}.$$

and

$$\varepsilon g_\varepsilon(s) + \sigma^2 c_H \int_0^\infty g_\varepsilon(r)|s-r|^{2H-2}dr = \sigma^2 c_H |s|^{2H-2}.$$

- Why g_ε exist? In which classe?
- What about convergence?

Take care: $K \notin L_1(R)$, $K \notin L_2(R)$!

Laplace Transforme

The analysis is based on the Laplace transform

$$\widehat{g}_t(z) = \int_0^t g(t,s)e^{-zs}ds, \quad z \in \mathbb{C}.$$

Then $g(t,\cdot)$ can be recovered by the inverse Laplace transform

$$g(t,s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{g}_t(i\lambda)e^{is\lambda}d\lambda.$$

Asymptotic analysis

Small perturbation?

The Laplace transform satisfies

$$\widehat{g}_t(z) = \widehat{g}(z) + \widehat{R}_t(z), \quad z \in \mathbb{C},$$

where $\widehat{g}(z)$ can be written explicitly and

Importent estimate

For all sufficiently small $\delta > 0$, there exist positive constants C, $T_{\min}$ and c so that

$$\sup_{\|\theta-\theta_0\|\le\delta} \int_{-\infty}^{\infty} \left|\partial_i \partial_j \widehat{R}_t(i\lambda;\theta)\right|^2 \Lambda(i\lambda;\theta_0) d\lambda \le Ct^{-c}, \quad \forall t \ge T_{\min}.$$

Principal representation

The Laplace transform satisfies

$$\widehat{g}_t(z) - 1 = \frac{\Phi_0(z) + e^{-tz}\Phi_1(-z)}{\Lambda(z)}$$

where

$$\Lambda(z) = \varepsilon + \frac{\sigma^2}{2}\Gamma(2H+1)\big(z^{1-2H} + (-z)^{1-2H}\big) = \varepsilon + \widehat{K}(z)$$

the "spectral density" of the "derivative" of X in $\mathbb{C} \setminus \mathbb{R}$
and
$\Phi_0(z)$ and $\Phi_1(z)$ are sectionally holomorphic on $\mathbb{C} \setminus \mathbb{R}_+$.
But $\widehat{g}_t$ should be analytic in $\mathbb{C}$.
A discontinuity along the real axis should be removable.

Principal representation

The Laplace transform satisfies

$$\widehat{g}_t(z) - 1 = \frac{\Phi_0(z) + e^{-tz}\Phi_1(-z)}{\Lambda(z)}$$

where

$$\Lambda(z) = \varepsilon + \frac{\sigma^2}{2}\Gamma(2H+1)\big(z^{1-2H} + (-z)^{1-2H}\big) = \varepsilon + \widehat{K}(z)$$

the "spectral density" of the "derivative" of X in $\mathbb{C} \setminus \mathbb{R}$
and
$\Phi_0(z)$ and $\Phi_1(z)$ are sectionally holomorphic on $\mathbb{C} \setminus \mathbb{R}_+$.
But $\widehat{g}_t$ should be analytic in $\mathbb{C}$.
A discontinuity along the real axis should be removable.

Removing the singularities

A discontinuity along the real axis — Hilbert type problem

$$\Phi_0^+(t) - e^{2i\alpha(t)}\Phi_0^-(t) = -2ie^{i\alpha(t)}\sin\alpha(t)e^{-t}\Phi_1(-t),$$

$$\Phi_1^+(t) - e^{2i\alpha(t)}\Phi_1^-(t) = 2ie^{i\alpha(t)}\sin\alpha(t)e^{-t}\Phi_0(-t), \quad t > 0$$

under conditions: $\Phi \sim \varepsilon$, $z \sim \infty$ and $\Phi \lesssim z^{1-2H}$ $z \sim 0$.

where $\alpha(t) = \arg(\Lambda^+(t))$.

by the Sokhotski-Plemelj theorem

To find Φ and/or to analyse the behaviour of $\widehat{g}_t$ is equivalent to analyse an integral system of equations:

$$\begin{aligned} p_t(s) &= \quad \frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+s}p_t(\tau)d\tau - \frac{1}{2}, \\ q_t(s) &= -\frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+s}q_t(\tau)d\tau - \frac{1}{2}, \end{aligned} \qquad s \in \mathbb{R}_+.$$

How to obtain the principal representation?
We use that

$$|x-y|^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t|x-y|}\, dt.$$

Outline

From small noise to large time

Main equality

$$\nabla \rho_t^{\varepsilon}(X^{\varepsilon}, \theta_0) = \varepsilon^{(1-\gamma_0)/2} \nabla \rho^1_{t\varepsilon^{-\gamma_0}}(X^1, \theta_0) M(\varepsilon, \theta_0)^{\top}$$

where $\gamma = 1/(2H-1)$ and

$$M(\varepsilon, \theta) = \begin{pmatrix} 1 & -2\sigma^2 \log \varepsilon^{-\gamma} \\ 0 & 1 \end{pmatrix}$$

From small noise to large time

The solution g_ε satisfies the following scaling property with respect to ε: for any $\varepsilon > 0$ and $t > s > 0$,

$$g_\varepsilon(t,s) = \varepsilon^{-\gamma} g_1\left(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\right),$$

$$\nabla g_\varepsilon(t,s) = \varepsilon^{-\gamma} \nabla g_1\left(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\right) M(\varepsilon,\theta)^\top,$$

$$\nabla^2 g_\varepsilon(t,s) = \varepsilon^{-\gamma}\nu(\varepsilon,\theta)\frac{\partial}{\partial\sigma^2} g_1\left(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\right)$$
$$+\varepsilon^{-\gamma} M(\varepsilon,\theta)\nabla^2 g_1\left(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\right) M(\varepsilon,\theta)^\top.$$

Try to differentiate the integral equation

The solution g_1 satisfies

$$t\frac{\partial}{\partial t}g_1(t,s)+s\frac{\partial}{\partial s}g_1(t,s)+g_1(t,s)=(2H-1)\sigma^2\frac{\partial}{\partial \sigma^2}g_1(t,s).$$

But g is not derivable in 0!

Work in progress

- Continuous time "Whittle formula" for a large class of processes
- Continuous time "Whittle likelihood"
- MLE and Bayes estimator