Prob. Sem. / Math Finance Colloq.

Parameter estimation for Gaussian sequences: long memory motivations in finance, sharp asymptotic normality and non-normality.

Frederi G. Viens, Dept. Statistics, Purdue University;

Division of Mathematical Sciences, National Science Foundation

Los Angeles, Sep 18, 2015

Plan of the talk

0. [Blackboard] Motivation from long memory stochastic volatility
1. Partially observed fractional Ornstein-Uhlenbeck processes: from continuous to discrete observations
2. Scale parameter estimators for discretely observed general stationary Gaussian processes
3. Optimal fourth and third moment theorems (Malliavin)
4. Speed of normal convergence in total variation for stationary Gaussian power variations
5. Non-normal convergence

Joint work with

- Léo Neufcourt, Columbia University,
- Khalifa Es-Sebaiy, B. El Onsy, Cadi Ayyad University (Morocco)
- Luis Barboza, Universidad de Costa Rica (Costa Rica)
- Work partially funded by
  - CIMFAV Universidad de Valparaíso (Chile)
  - Ecole Polytechnique, Palaiseau (France)
  - NSF
  - Fulbright Commission

1 Partially observed Gaussian processes

- **Classical case: OU process (Lipster-Shiryaev 78)**

  * $dX(t) = -\theta X(t)\,dt + dV(t)\,,\ \ t \in [0,T]$, increas, horiz. $T \to \infty$

  * explicit non-stationary or stationary solution

  $\int_0^t e^{-\theta(t-s)} dV(s)$ or $\int_{-\infty}^t e^{-\theta(t-s)} dV(s)$ no matter what V is.

- **QUESTION: estimate** $\theta > 0$.

- **ANSWERS:** when V is Brownian motion, since $d\langle V\rangle = dt$

  * using Girsanov, get MAXIMUM LIKELIHOOD ESTIMATOR

$$\hat{\theta}_T = -\frac{\int_0^T X(s)\,dX(s)}{\int_0^T |X(s)|^2\,ds}.$$

  * another interpretation: (ORDINARY) LEAST SQUARES ESTIM

  indeed, interpret Itô integral $\int_0^T X(s)\,dX(s) = \int_0^T X(s)\,\dot{X}(s)\,ds$

  then minimize $L^2(ds)$-norm of the regression error

$$\theta \mapsto \int_0^T \left|\dot{X}(s) + \theta X(s)\right|^2 ds.$$

- **OU process driven by an OU process :**
  - * Replace V with another OU process:
  - * $dV(t) = -\rho X(t)\,dt + dB(t)$
  - * DIFFICULTY: ONLY OBSERVE X BUT ESTIMATE (θ, ρ).
  - * Bravely and naïvely assume $\hat{\theta}_T$ still works
  - * Observed version of V : $\hat{V}(t) = X(t) + \hat{\theta}_T \int_0^t X(s)\,ds$
  - * Bravely and naïvely try

$$\hat{\rho}_T = \frac{\int_0^T \hat{V}(s)\,d\hat{V}(s)}{\int_0^T \left|\hat{V}(s)\right|^2 ds}.$$

  - * Bercu Proia and Savy SPL 2014 proved this intuition is incorrect:

$$\lim_{T\to\infty} \left(\hat{\theta}_T, \hat{\rho}_T\right) = \left(\theta + \rho, \frac{\theta\rho(\theta+\rho)}{(\theta+\rho)^2 + \rho\theta}\right).$$

- **OU process driven by a fractional OU process ($H > 1/2$):**

  * Same situation as above, but B is fractional Brownian motion.

  * First look at fully observed case: MLE and LSE *do not* coincide

    – MLE (Kleptsyna and Le Breton SISP 2002)

    – LSE (Nualart and Hu SPL 2010)

    they propose $\hat{\theta}$ but stoch. integ. is of Skorohod type

    they propose an unrelated method of moments

  * We adapt method of Bercu Proia and Savy:

    – need the Malliavin calculus to recompute all needed asymptotics;

    – evaluate Skorohod integs by converting to Young integs;

    – invoke the Nualart-Peccati criterion to prove asymp. normality

$$\lim_{T\to\infty}\left(\hat{\theta}_T,\hat{\rho}_T\right)=\left(\theta+\rho,\frac{\theta\rho\left(\theta+\rho\right)}{\left(\theta+\rho\right)^2+\frac{\rho^{2-2H}-\theta^{2-2H}}{\theta^{-2H}-\rho^{-2H}}}\right)=:\left(\theta^*,\rho^*\right)\ \text{a.s.;}$$

$$\lim_{T\to\infty}\sqrt{T}\left(\hat{\theta}_T-\theta^*,\hat{\rho}_T-\rho^*\right)=N\left(0,\Xi\right)\ \text{with } \Xi \text{ explicit, } H<3/4.$$

- **Same problem, now assume data is discrete:** timestep Δ_n,

  * Generic quadratic variation $Q_n(Z) := \frac{1}{n}\sum_{k=1}^{n} Z(k\Delta_n)$.

  * Replace $\int_0^T |X(s)|^2\, ds$ by $Q_n(X)$.

  * Also need to discretize $\Sigma_t := \int_0^t X(s)\, ds$ with $\hat{\Sigma}_{k\Delta_n} = \Delta_n \sum_{i=1}^{k} X(i\Delta_n)$.

  * Replace $\int_0^T \left|\hat{V}(s)\right|^2 ds$ by $Q_n(X) + \left|\check{\theta}_T\right|^2 Q_n\left(\hat{\Sigma}\right)$

  * to define $\left(\check{\theta}, \check{\rho}\right)$ as solution to $F\left(\check{\theta}, \check{\rho}\right) = \left(Q_n(X), Q_n\left(\hat{\Sigma}\right)\right)$ where

$$F(x,y) = H\Gamma(2H) \times \begin{cases} \frac{1}{y^2-x^2}\left(y^{2-2H} - x^{2-2H}, x^{-2H} - y^{-2H}\right) & \text{if } x \neq y \\ \left((1-H)x^{-2H}, Hx^{-2H-2}\right) & \text{if } x = y. \end{cases}$$

  * We prove a.s. cvce via control of errors from continuous case;

  need intermediate estimator based on $Q_n(\Sigma)$;

  need sampling rate $\Delta_n \leq n^{-\frac{1}{1+H}-\varepsilon}$.

  * We prove CLT based on a.s. control of errors, and our continuous

  CLT

  but need a stronger sampling rate $\Delta_n = o\left(n^{-\frac{1}{1+2H/3}}\right)$.

2 Discretey observed Gaussian processes (in progr.)

- **Work directly with the discrete observations from stationary continuous-time process, and exploit 4th or 3rd moment theorem**

  * For quadratic estimators with stationary processes, setup is equivalent to

$$Q_n(Z) := \frac{1}{n}\sum_{k=1}^{n} |Z(k)|^2$$

  where Z is stationary Gaussian sequence with auto-covariance ρ

  * Let $\gamma = \rho(0) = Var(Z(k))$; let $V_n(Z) := \sqrt{n}(Q_n(Z) - \gamma)$

  * By 3rd moment theorem

$$d_{TV}\left(\frac{V_n(Z)}{\sqrt{Var(V_n(Z))}}, N(0,1)\right) \asymp \frac{\left(\sum_{k=1}^{n}\rho(k)^{3/2}\right)^2}{\left(\sum_{k=1}^{n}\rho(k)^2\right)^{3/2}}$$

  and in fact $Var(V_n(Z)) \asymp n^{-1}\sum_{k=1}^{n}\rho(k)^2$ and need not converge.

  * More powerful strategy: no sampling restriction: $\Delta_n = 1$ is OK.

- Other powers (Hermite variations): 3rd moment theorem not valid;
  * optimal 4th moment theorem still holds, repeat detailed analysis;
  * or just use non-optimal result of Nourdin & Peccati (PTRF, 2008):

$$d_{TV}\left(\frac{V_n(Z)}{\sqrt{Var(V_n(Z))}}, N(0,1)\right) \leq C\sqrt{\kappa_4\left(\frac{V_n(Z)}{\sqrt{Var(V_n(Z))}}\right)}.$$

- The actual speeds depend on how quicky $Var(V_n(Z))$ stabilizes.
  * Ease of extension to non-stationary data.
  * Have strategy to avoid $H < 3/4$ condition for fGn
- Application to fractional OU and OUfOU:
  * fOU: $\lim Q_n(X) = H\Gamma(2H)\theta^{2H}$ is a method of moments;
  * OUfOU: *conjec* : a.s. & CLT for our previous $\left(\check{\theta}, \check{\rho}\right)$ with $\Delta_n = 1$.
- When $\rho(k) = k^{2H-2}L(k)$ with L slowly varying and $H > 3/4$,
  * fluctuations have H-Rosenblatt-distributed limit ;
  * open problem to find optimal TV or Wasserstein speed.

3 Malliavin calculus tools

References

[NuP] Nualart, D.; Peccati, G. (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (1), 177-193.

[BBL] Biermé, H.; Bonami, A.; León, J.R. (2011). Central limit theorems and quadratic variations in terms of spectral density. *Electron. J. Probab.* **16** (13), 362-395.

[BBNP] Biermé, H.; Bonami, A.; Nourdin, I.; Peccati, G. (2013), Optimal Berry-Esseen rates on the Wiener space: the barrier of third and fourth cumulants. *ALEA* **9** (2), 473-500.

[NP] Nourdin, I.; Peccati, G. (2013). The optimal fourth moment theorem. In *Proceedings of the Amer. Math. Soc*, 2014.

Theorem 3.1 [Sharp 4th moment thm in total variation: [NuP], [NP]].

- *Assumptions:*

  (1) $(F_n)_{n\geq 0}$ *sequence in a fixed Wiener chaos;*

  (2) $Var\,[F_n] = 1$

- *Conclusions:*

  (1) [NuP] The following two statements are equivalent:

  (a) $(F_n)_{n\geq 0}$ *converges in law towards* $\mathcal{N}(0,1)$ *;*

  (b) $\mathbf{E}[F_n^4] \to 3 = \mathbf{E}[\mathcal{N}(0,1)^4]$.

  (2) [NP] Let $M_n := \max(\mathbf{E}[F_n^4] - 3, \left|\mathbf{E}[F_n^3]\right|)$. *If (1a) holds, then*

$$\exists c, C > 0 : cM_n \leq d_{TV}(F_n, N) \leq CM_n.$$

Where does such a result come from?

Nourdin-Peccati Analysis on Malliavin calculus with Stein's lemma.

Stein's lemma $\Longrightarrow$ bound on d_{TV} (uniform dist. of probab. measures):
Let $\mathcal{F}_{TV} = \left\{ f : \|f\|_\infty \leq \sqrt{\pi/2}\ ;\ \|f'\|_\infty \leq 2 \right\}$. Then for X w/ density,

$$d_{TV}\left(X, \mathcal{N}(0,1)\right) \leq \sup_{f \in \mathcal{F}_{TV}} \left| \mathbf{E}\left[f'(X)\right] - \mathbf{E}\left[X f(X)\right] \right|.$$

* Assume $L^2(\Omega)$ is w.r.t a fixed BM on $[0,1]$, say.
* For $X \in \mathbf{D}^{1,2}$ with $\mathbf{E}[X] = 0$, $Var[X] = 1$, let $G_X := \int_0^1 D_s X\, D_s\left(-L^{-1}\right) X\, ds$.
* Then

$$\mathbf{E}\left[X f(X)\right] = \mathbf{E}\left[G_X f'(X)\right].$$

* Hence

$$d_{TV}\left(X, \mathcal{N}(0,1)\right) \leq 2\mathbf{E}\left[\left|1 - G_X\right|\right].$$

* Moreover if $X \in \mathbf{D}^{1,4}$, then

$$\mathbf{E}\left[\left|1 - G_X\right|\right] \leq 2\sqrt{\mathbf{E}\left[\left|1 - G_X\right|^2\right]} = 2\sqrt{Var}\left[G_X\right].$$

Special case: $X \in q^{\mathrm{th}}$ **chaos.**

* Then by def, $\left(-L^{-1}\right) X = q^{-1} X$ and thus $G_X = q^{-1} \|DX\|^2$.

* Then work a little hard to estimate $\sqrt{Var}\,[G_X] \asymp \mathbf{E}\left[X^4\right] - 3$. Thus

$$d_{TV}(F_n, N) \leq C_q \sqrt{\mathbf{E}\left[X^4\right] - 3}.$$

How did [NP] get better estimate?

* NOTATION: cumulants $\kappa_4\,(X) := \mathbf{E}\left[X^4\right] - 3$ and $\kappa_3\,(X) := \mathbf{E}\left[X^3\right]$

* Iterate a polarization and iteration of G_X up to order 4, to get

$$d_{TV}(F_n, N) \leq C_q \left(|\kappa_3| + \kappa_4 + \sqrt{\kappa_4 \left(\kappa_3^2 + \kappa_4^{3/2} \right)} \right) \leq c_q \max\left(|\kappa_3|\,, \kappa_4\right)$$

* use $f = \sin$ and $f = \cos$ to find lower bounds with $|\kappa_3|$ and κ_4 respectively.

- **Stationary centered Gaussian sequence:** $(X_n)_{n\in\mathbb{Z}}$

  X is jointly Gaussian and $\mathbf{E}[X_n] = 0$;

  $\exists\rho$ on $\mathbb{Z} : \forall k, n \in \mathbb{Z}$, $\mathbf{E}[X_n X_{n+k}] = \rho(k)$;

  $\rho(0) = Var[X_n] = 1$ without loss of generality;

- **Only technical assumptions:**

  ρ is of constant sign;

  $|\rho|$ decreases near $+\infty$.

- **Normalized centered quadratic variation**

$$V_n := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \left(X_k^2 - 1\right),$$

$$v_n := \mathbf{E}\left[V_n^2\right] = Var[V_n],$$

$$F_n := \frac{V_n}{\sqrt{v_n}}.$$

Therefore F_n is a sequence in second Wiener chaos with $\mathbf{E}\left[F_n^2\right] = 0$.

- **Breuer-Major theorem (special case):**
  if $\sigma^2 := \sum_{k\in\mathbb{Z}} \rho(k)^2 = \lim_{n\to\infty} v_n < \infty$ then $V_n \underset{law}{\to} \mathcal{N}\left(0,\sigma^2\right)$.

- Beyond this theorem, there are cases where $F_n \to \mathcal{N}\left(0,\tilde{\sigma}^2\right)$ even if $\sigma^2 = \infty$ and law(V_n) diverges. In that case, we say that normal convergence holds without a Breuer-Major theorem.

- **Speed of Breuer-Major convergence in total variation:**
  * **Classical Berry-Esséen (iid case):** $d_{TV}\left(V_n, \mathcal{N}(0,1)\right) \le c/\sqrt{n}$
  * **the case of fGn (increments of fBm):** $\rho(k) \sim H(2H-1)\,k^{2H-2}$
  – Before [BBNP], it was believed that classical *Berry-Esséen rate holds* for fGn only if $H < 3/5$;
  – thanks to [BBNP], and [NP], we know *Berry-Esséen rate holds* if and only if $H < 2/3$, and other rates can be computed, for example:
    * if $H = 2/3$, then $d_{TV}\left(F_n, \mathcal{N}(0,1)\right) \asymp n^{-1/2}\log^2 n$;
    * if $H \in (2/3, 3/4)$ then $d_{TV}\left(F_n, \mathcal{N}(0,1)\right) \asymp n^{6H-9/2}$.

4 Third moment theorem

IDEA: use [NP]'s Theorem 3.1;

NOTATION: $\kappa_4(F_n) := \mathbf{E}\left[F_n^4\right] - 3$ and $\kappa_3(F_n) := \mathbf{E}\left[F_n^3\right]$

REQUIREMENT: compute $\kappa_4(F_n)$ and $\kappa_3(F_n)$ as sharply as possible.

LUCKY BREAK: don't need $\kappa_4(F_n)$ as sharply as $\kappa_3(F_n)$:)

Proposition 4.1 [Sharp cumulants estimates]

$$\frac{2}{v_n^{3/2}\sqrt{n}}\left(\sum_{|k|<n}|\rho(k)|^{3/2}\right)^2 \leq |\kappa_3(F_n)| \leq \frac{8}{v_n^{3/2}\sqrt{n}}\left(\sum_{|k|<n}|\rho(k)|^{3/2}\right)^2, \tag{1}$$

$$\kappa_4(F_n) = \frac{const}{v_n^2\ n}\left(\sum_{|k|<n}|\rho(k)|^{4/3}\right)^3, \tag{2}$$

$$\sum_{k=0}^{n-1}\rho^2(k) \leq v_n \leq 2\sum_{k=0}^{n-1}\rho^2(k). \tag{3}$$

By Holder's inequality for $\sum_{|k|<n}|\rho(k)|^{2/3+2/3}$ with $p = \frac{9}{4}, q = \frac{9}{5}$, inequalities (1) and (2) imply:

Corollary 4.2 [Lucky shortcut] $n^{1/4}\kappa_4(F_n)^{3/4} = \mathcal{O}\left(|\kappa_3(F_n)|\right)$.

Consequently, by 4th moment theorem of [NuP],

$\kappa_3(F_n) \to 0 \implies \kappa_4(F_n) - 3 \to 0 \implies \lim_{n\to\infty} law(F_n) = \mathcal{N}(0,1)$.

And by sharp 4th moment theorem of [NP], the convergence is in total variation, and all the converses also holds. More precisely, $\kappa_4(F_n) = o\left(|\kappa_3(F_n)|\right)$ and we have the following.

Theorem 4.3 [Sharp third moment theorem in TV]

(i) $\lim_{n\to\infty} law(F_n) = \mathcal{N}(0,1)$.

$\iff$ *(ii)* $\lim_{n\to\infty} \kappa_3\left(F_n\right) = 0$;

$\iff$ *(iii)* $\lim_{n\to\infty} \kappa_4\left(F_n\right) = 0$;

$\iff$ *(iv)* $\varepsilon_n := \left(\sum_{|k|<n} |\rho(k)|^{3/2}\right)^2 v_n^{-3/2} n^{-1/2} = o\left(1\right)$.

In this case, $\kappa_3\left(F_n\right) \asymp \varepsilon_n$ *, and* ε_n *is the exact TV convergence rate:*

$$d_{TV}(F_n, \mathcal{N}(0,1)) \asymp \kappa_3\left(F_n\right) \asymp \varepsilon_n.$$

5 Non-normal convergence

References

[BN] Breton, J.C.; Nourdin, I. (2008). Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion. *Electron. Comm. Probab.* **13**, 482-493.

[DM] Davydov, Y.A.; Martynova, G. V. (1987). Limit behavior of multiple stochastic integral. *Statistics and control of random process. Preila, Nauka, Moscow*, 55-57 (in Russian).

[DM79] Dobrushin, R.L.; Major, P (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z.Wahrsch. Verw. Gebiete* **50** (9), 27-52.

[T79] Taqqu, M. S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50** (1), 53-83.

Theorem 5.1 [Dobrushin-Major / Taqqu 79 (special case)]

Assumption: $\rho(k) = k^{2H-2} L(k)$ *where* L *is slowly varying at* ∞ *and* $H \in (3/4, 1)$.

Conclusion [DM79]: F_n *converges in law to the law of a Rosenblatt r.v.*

$$F_\infty = \iint_{\mathbf{R}^2} \frac{|xy|^{H-1/2}}{\sqrt{K_H}} \frac{e^{i(x+y)} - 1}{i(x+y)} W(dx) W(dy).$$

Here W is a $\mathbb{C}$-valued white noise: on $\mathbb{R}_+$, $W(dx) = B_1(dx) + iB_2(dx)$ for B_1, B_2 iid BMs; $W(-dx) = \overline{W(dx)}$, $W(dx)^2 = 0$, $|W(dx)|^2 = dx$.

[BN] find TV speed of convergence by using a classical result [DM]:

* if F_n and F_∞ are on in the same Wiener chaos, and $Var[F_n]$ bounded
* then for large n, $d_{TV}(F, F_\infty) \leq c_{F_\infty} \left(\hat{\mathbf{E}} \left[(F - F_\infty)^2 \right] \right)^{1/4}$.

Case $L = 0$ **:** using self-similarity of fBm [BN] prove

$$d_{TV}(F, F_\infty) \leq c_{F_\infty} n^{3/4-H}.$$

Unfortunately, we find this good speed only works for fGn...

- **Log-modulated power spectral density:**
  - let $H \in (3/4, 1)$ and $\beta \geq 0$, let $L(y) = \log^{2\beta}(|y|)$ (or asymp). Let

$$q(x) := C_{H,\beta} |x|^{1-2H} L\left(\frac{e\pi}{|x|}\right).$$

  - This $q \in L^1\left(S^1, dx\right)$, q is C^∞ except at 0; thus $q \equiv$ Fourier series of its Fourier inverse ρ.
  - I.e. let X with covariance function

$$\rho(k) := \frac{1}{2\pi} \int_{-\pi}^{\pi} q(x) \cos(kx)\, dx$$

    then X has spectral density q and def of ρ is Fourier inversion.
  - Recompute ρ by changing variables:

$$\rho(k) = \frac{C_{H,\beta}}{2\pi} k^{2H-2} \int_{-k\pi}^{k\pi} |x|^{1-2H} \cos(x)\, L\left(k\frac{e\pi}{|x|}\right) dx.$$

Theorem 5.2 *Using above X, with a c depending only on H and β,*

$$d_{TV}\left(F_n, F_\infty\right) \leq \frac{c}{\log^{1/2} n}$$

where, with $K' = \frac{(2\Gamma(2-2H)\cos(\pi(1-H)))^2}{(4H-2)(4H-3)}$*, F_∞ is Rosenblatt law*

$$F_\infty = \iint_{\mathbf{R}^2} \frac{|xy|^{H-1/2}}{\sqrt{K'_H}} \frac{e^{i(x+y)}-1}{i\left(x+y\right)} W\left(dx\right) W\left(dy\right).$$

- Technical proof:

  * Typically: ρ has long memory, thus $\rho \notin \ell^1(\mathbb{Z})$, cannot invoke classical Fourier inversion; prove it or assume it by working with q directly.

  * To apply a meta-theorem: use a trade off between a speed of cvce to limiting kernel and speed of integrability of cutoff kernel at ∞.

  * need precise estimates of ρ and nv_n :

with $K_H := \frac{1}{\pi}\int_0^\infty |x|^{1-2H}\cos(x)\,dx = 2\Gamma(2-2H)\cos(\pi(1-H))$,

$$\rho(k) = C_{H,\beta}K_H L(k)\,k^{2H-2}\left(1+\mathcal{O}\left(\frac{1}{L(k)}\right)\right);$$

$$nv_n = \left(C_{H,\beta}\right)^2 K'_H\ n^{4H-2}\ L^2(n)\left(1+\mathcal{O}\left(\frac{1}{\log n}\right)\right).$$