Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 347 (2009) 663–666
COMPTES RENDUS MATHEMATIQUE

Probability Theory/Statistics

Application of Malliavin calculus to long-memory parameter estimation for non-Gaussian processes

Alexandra Chronopoulou [a], Ciprian A. Tudor [b], Frederi G. Viens [a]

[a] *Department of Statistics, Purdue University, 150 N. University St., West Lafayette, IN 47907-2067, USA*
[b] *SAMOS-MATISSE, centre d'économie de La Sorbonne, Université de Paris 1, 90, rue de Tolbiac, 75634 Paris, France*

Received 22 March 2009; accepted 24 March 2009

Available online 22 April 2009

Presented by Paul Malliavin

Abstract

Using multiple Wiener–Itô stochastic integrals and Malliavin calculus we study the rescaled quadratic variations of a general Hermite process of order q with long-memory (Hurst) parameter $H \in (\frac{1}{2}, 1)$. We apply our results to the construction of a strongly consistent estimator for H. It is shown that the estimator is asymptotically non-normal, and converges in the mean-square, after normalization, to a standard Rosenblatt random variable. ***To cite this article: A. Chronopoulou et al., C. R. Acad. Sci. Paris, Ser. I 347 (2009).***

Résumé

Application du calcul de Malliavin à l'estimation du paramètre de mémoire longue pour des processus non-gaussiens. Nous servant des intégrales multiples de Wiener–Itô et du calcul de Malliavin, nous étudions la variation quadratique renormalisée d'un processus de Hermite général d'ordre q avec paramètre de mémoire longue $H \in (\frac{1}{2}, 1)$. Nous appliquons nos résultats à la construction d'un estimateur fortement consistent pour H. Il est démontré que l'estimateur est asymptotiquement non-normal, et converge en moyenne de carrés, après normalisation, vers une variable aléatoire de Rosenblatt standard. ***Pour citer cet article : A. Chronopoulou et al., C. R. Acad. Sci. Paris, Ser. I 347 (2009).***

1. Introduction

A stochastic process $\{X_t\colon t \in [0, 1]\}$ is called self-similar with self-similarity parameter $H \in (0, 1)$ when typical sample paths look qualitatively the same irrespective of the distance from which we look at them, i.e. for any fixed time-scaling constant for $c > 0$, the processes $c^{-H} X_{ct}$ and X_t have the same distribution. Self-similar stochastic processes are well suited to model physical phenomena that exhibit long memory. The most popular among these processes is the fractional Brownian motion (fBm), because it generalizes the standard Brownian motion and its self-similarity parameter can be interpreted as the long memory parameter.

E-mail addresses: achronop@purdue.edu (A. Chronopoulou), tudor@univ-paris1.fr (C.A. Tudor), viens@purdue.edu (F.G. Viens).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2009.03.026

In this article we study a more general family of processes, the Hermite processes. Every process in this family has the same covariance structure, and thus the same long memory property, as fBm:

$$\mathrm{Cov}(X_t, X_s) = \mathbf{E}[X_t X_s] = 2^{-1}\big(s^{2H} + t^{2H} - |t-s|^{2H}\big), \quad s, t \in [0,1]. \tag{1}$$

A Hermite process can be defined in two ways: as a multiple integral with respect to a standard Wiener process or as a multiple integral with respect to an fBm with suitable H. We adopt the first approach.

Definition 1.1. The Hermite process $(Z_t^{(q,H)})_{t\in[0,1]}$ of order $q \geqslant 1$ and parameter $H \in (\frac{1}{2}, 1)$ is given by

$$Z_t^{(q,H)} = d(H) \int_0^t \cdots \int_0^t \mathrm{d}W_{y_1} \cdots \mathrm{d}W_{y_q} \left(\int_{y_1 \vee \cdots \vee y_q}^t \partial_1 K^{H'}(u, y_1) \cdots \partial_1 K^{H'}(u, y_q)\, \mathrm{d}u \right), \quad t \in [0,1] \tag{2}$$

where W is a standard Wiener process, $K^{H'}$ is the kernel of fBm (see [4, Chapter 5]) and $H' = 1 + \frac{H-1}{q}$.

The constant $d(H) := \frac{(2(2H-1))^{1/2}}{(H+1)H^{1/2}}$ is chosen to match the covariance formula (1). As a multiple Itô integral of order q of a non-random function with respect to Brownian motion, $Z^{(q,H)}$ belongs in the qth Wiener chaos. For $q > 1$, it is far from Gaussian. Like fBm, all Hermite processes $Z^{(q,H)}$ are H-self-similar and have stationary increments and Hölder-continuous paths of any order $\delta < H$. Moreover, they exhibit long-range dependence in the sense that the auto-correlation function is not summable. They encompass the fBm ($q = 1$) and the Rosenblatt process ($q = 2$).

The statistical estimation of H is of great interest and importance, since H describes the memory of the process as well as other regularity properties. Several methodologies to the long-memory estimation problem have been proposed, such as wavelets, variations, maximum likelihood methods (see [1]). Our approach is based on the quadratic variation of the process, by analogy to the techniques which have been used for fBm for many years (see references in [2]), and more recently in [6].

2. Variations of the Hermite process

Let $Z^{q,H}$ be a Hermite process of order q with self-similarity index $H \in (\frac{1}{2}, 1)$ as in Definition 1.1. Assume $Z^{q,H}$ is observed at discrete times $\{\frac{i}{N}\colon i = 0, \ldots, N\}$ and define the centered quadratic variation statistic V_N:

$$V_N = -1 + \frac{1}{N} \sum_{i=0}^{N-1} N^{2H} \big(Z^{(q,H)}_{\frac{i+1}{N}} - Z^{(q,H)}_{\frac{i}{N}}\big)^2. \tag{3}$$

Note that $N^{-2H} = \mathbf{E}[(Z^{(q,H)}_{(i+1)/N} - Z^{(q,H)}_{i/N})^2]$ is a normalizing factor. To compute the variance of V_N we expand V_N in the Wiener chaos. Using Definition 1.1 one sees that $Z^{(q,H)}_{(i+1)/N} - Z^{(q,H)}_{i/N} = I_q(f_{i,N})$, where $I_q(\cdot)$ is the Wiener–Itô integral of order q and $f_{i,N}(y_1, \ldots, y_q)$ is a non-random symmetric H-dependent function of q variables. Using the product formula for multiple Wiener–Itô integrals (see [4, Proposition 1.1.3]), we can write $|I_q(f_{i,N})|^2 = \sum_{l=0}^q l!(C_q^l)^2 I_{2q-2l}(f_{i,N} \otimes_l f_{i,N})$, where the $f \otimes_l g$ denotes the l-contraction of the functions f and g. In this way we obtain the Wiener-chaos expansion of V_N

$$V_N = T_{2q} + c_{2q-2} T_{2q-2} + \cdots + c_4 T_4 + c_2 T_2, \tag{4}$$

where $c_{2q-2k} := k!\binom{q}{k}^2$ are the combinatorial constants from the product formula for $0 \leqslant k \leqslant q-1$, and $T_{2q-2k} := N^{2H-1} I_{2q-2k}(\sum_{i=0}^{N-1} f_{i,N} \otimes_k f_{i,N})$. This decomposition allows us to find V_N's precise order of magnitude via its variance's asymptotics, as proved in the following lemma.

Lemma 2.1. *With $c_{H,q} := \frac{4d(H)^4 (H'(2H'-1))^{2q-2}}{(4H'-3)(4H'-2)}$, it holds that*

$$\lim_{N\to\infty} \mathbf{E}\big[c_{H,q}^{-1} N^{2(2-2H')} c_2^{-2} V_N^2\big] = \lim_{N\to\infty} \mathbf{E}\big[c_{H,q}^{-1} N^{2(2-2H')} c_2^{-2} T_2^2\big] = 1.$$

Proof. To establish this result we only need to estimate the L^2-norm of each term appearing in the chaos decomposition, since they are orthogonal in $L^2(\Omega)$. This calculation is achieved by using the so-called isometry property (see [4, Section 1.1.2]) which states that $\mathbf{E}[|I_k(f)|^2] = k!\|f\|^2_{L^2([0,1]^k)}$. It turns out that $\lim_{N\to\infty} \mathbf{E}[c_{H,q}^{-1} N^{(2-2H')(2)} T_2^2] = 1$ and $\mathbf{E}[N^{2(2-2H')} T_{2q-2k}^2] = \mathrm{O}(N^{-2(2-2H')2(q-k-1)})$. Therefore the dominant term in the decomposition is T_2, and the result follows. □

The following theorem gives the precise asymptotic distribution of V_N. Unlike the case $q=1$, when $q \geqslant 2$ there is no range of H for which asymptotic normality holds.

Theorem 2.2. *For $H \in (1/2, 1)$ and $q = 2, 3, 4, \ldots$, let $Z^{(q,H)}$ be a Hermite process of order q and parameter H (see Definition 1.1). Then $c_{H,q}^{-1/2} c_2^{-1} N^{\frac{2-2H}{q}} V_N$ converges in $L^2(\Omega)$ as $N \to \infty$ to a standard Rosenblatt random variable R with parameter $H'' := 2(H-1)/q + 1$; that is, R is the value at time 1 of a Hermite process of order 2 and parameter H''.*

Proof. Let $I_i := [\frac{i}{N}, \frac{i+1}{N}]$, let $H' = 1 + (H-1)/q$, and $a(H') = H'(2H'-1)$. In order to understand the behavior of the renormalized V_N, it suffices to study the limit of the term $N^{2-2H'} T_2$. Indeed, from the proof of Lemma 2.1, the remaining terms in the chaos expansion of $N^{2-2H'} V_N$, i.e. $N^{(2-2H')} T_{2q-2k}$, converge to zero. Since $N^{2-2H'} T_2$ is a second chaos random variable it is now necessary and sufficient to prove that its symmetric kernel converges in $L^2([0,1]^2)$ to $c_{H,q}^{1/2}$ times the kernel of the Rosenblatt process at time 1 (see [4, Section 1.1.2]). Observe that the kernel of $N^{2-2H'} T_2$ can be written as a sum of two terms: $N^{2(H-H')+1} \sum_{i=0}^{N-1} f_{i,N} \otimes_{q-1} f_{i,N} = f_2^N + r_2$, with

$$f_2^N(y,z) := N^{2(H-H')+1} d(H)^2 a(H')^{q-1} \sum_{i=0}^{N-1} 1_{y\vee z \leqslant \frac{i}{N}} \iint\limits_{I_i \times I_i} \mathrm{d}v\, \mathrm{d}u\, \partial_1 K(u,y) \partial_1 K(v,z) |u-v|^{2(H'-1)(q-1)}.$$

We can show that the remainder term $r_2(y,z)$ converges to zero in $L^2([0,1]^2)$, as $N \to \infty$. Next, for each fixed i, one replaces u and v by the left endpoint of I_i, namely i/N. This approximation results in a function $\check{f}_2^N$ which is pointwise asymptotically equivalent to f_2^N; equivalence in $L^2([0,1]^2)$ is obtained via dominated convergence. The approximant $\check{f}_2^N$ itself is immediately seen to be a Riemann sum approximation, for fixed y, z, of the integral defining the kernel of the Rosenblatt process at time 1, as in Definition 1.1 for $q=2$. To pass from pointwise to $L^2([0,1]^2)$ convergence, dominated convergence is used again, the key point being that one calculates by hand that $\|\,\mathrm{cst}\, \check{f}_2^N\|^2_{L^2([0,1]^2)}$ equals $\sum_{i,j=0}^{N-1} N^{-2} |\int_0^{i\wedge j/N} \partial_1 K^{H'}(u,y) \partial_1 K^{H'}(j/N, y)\,\mathrm{d}y|^2$; bounding this expression by correlations of increments of fBm, one finds an explicit series which is bounded if $H' > 5/8$; this always holds since $q \geqslant 2$ implies $H' \geqslant 3/4$. □

In addition to T_2, it is interesting to explore the behavior of the remaining terms in the chaos expansion of V_N. In the following theorem we study the convergence of the term of greatest order in this expansion, T_{2q}. It turns out this term does have a normal limit when $H < 3/4$; this familiar threshold (see [2]) is the one obtained for normal convergence of V_N in the case of fBm ($q=1$). What we discover here is that when $q=1$, the only term in V_N is to be interpreted as T_{2q}, not T_2; but when $q \geqslant 2$, the term T_{2q} dominates T_2, and therefore V_N cannot converge normally.

Theorem 2.3. *Let $Z^{(q,H)}$ be a Hermite process as in the previous theorem. Let T_{2q} be the term of order $2q$ in the Wiener chaos expansion of V_N. For every $H \in (1/2, 3/4)$, $x_{1,H}^{-1/2} \sqrt{N}\, T_{2q}$ converges to a standard normal distribution, where $x_{1,H}$ is a constant depending only on H.*

Proof. In order to prove this result we use a characterization of the convergence of a sequence of multiple stochastic integrals to a Normal law by Nualart and Ortiz-Latorre (Theorem 4 in [5], which states that if F_N is in the qth chaos and $\mathbf{E}[F_N^2] \to 1$ and $\mathbf{E}[(\|DF_N\|^2_{L^2[0,1]} - 2q)^2] \to 0$ then F_N converges to a normal; see also [3]). Let $F_N = x_{1,H}^{-1/2} \sqrt{N}\, T_{2q}$. Using the same method as in Lemma 2.1, we get $\lim_{N\to\infty} \mathbf{E}[F_N^2] = 1$. Thus, it remains to check that the Malliavin derivative norm $\|DF_N\|^2_{L^2[0,1]} \to 2q$ in $L^2(\Omega)$. Using $\mathbf{E}[F_N^2] \to 1$ and a general immediate calculation, we get $\lim_{N\to\infty} \mathbf{E}\|DF_N\|^2_{L^2[0,1]} = 2q$. The proof is completed by checking that $\|DF_N\|^2_{L^2[0,1]}$

converges in $L^2(\Omega)$ to its mean. To do this, since it is a variable with a finite chaos expansion, it is sufficient to check that its variance converges to 0. The ensuing calculations begin with the explicit computation of $D_r F_N$ as $x_{1,H}^{-1/2}\sqrt{N}\,N^{2H-1}(2q)I_{2q-1}(\sum_{i=0}^{N-1}(f_{i,N}\otimes f_{i,N})(\cdot,r))$, and are similar to those needed to prove Theorem 2.2; their higher complexity reduces via polarization. □

Remark 1. It is possible to give the limits of the terms T_{2q-2} to T_4 appearing in the decomposition of V_N. All these renormalized terms should converge to Hermite random variables of the same order as their indices. This "reproduction" property will be investigated in a subsequent article.

3. Estimation of the long-memory parameter H

Assume that we observe a Hermite process of order q and self-similarity index H in discrete time. From these data we can compute the quadratic variation $S_N := \frac{1}{N}\sum_{i=0}^{N-1}(Z^{(q)}_{(i+1)/N} - Z^{(q)}_{i/N})^2$. We can immediately relate S_N to the scaled quadratic variation V_N: we have $1+V_N = N^{2H}S_N$. By Lemma 2.1, $\lim_{N\to\infty} V_N = 0$ in $L^2(\Omega)$; since V_N has a finite Wiener chaos decomposition, the convergence also holds in any $L^p(\Omega)$. Taking p large enough, the Borel–Cantelli lemma implies that $V_N \to 0$ almost surely. Therefore, taking logarithms, $2H\log N + \log S_N \to 0$ almost surely. We have thus proved the following.

Proposition 3.1. *Let $\hat{H}_N := -\frac{\log S_N}{2\log N}$; it is a strongly consistent estimator for H: $\lim_{N\to\infty}\hat{H}_N = H$ a.s.*

The next step is to determine the asymptotic distribution of $\hat{H}_N$. It turns out that we have convergence to a Rosenblatt random variable in $L^2(\Omega)$, according to the following theorem.

Theorem 3.2. *There is a standard Rosenblatt random variable R with parameter $2H'-1$ such that*

$$\lim_{N\to\infty}\mathbf{E}\big[\big|2N^{2-2H'}(H-\hat{H})\log N - c_2 c_{H,q}^{1/2}R\big|^2\big] = 0.$$

Proof. By definition of $\hat{H}_N$ in Proposition 3.1, and the relation $1+V_N = N^{2H}S_N$, we have

$$2(H-\hat{H}_N)\log N = \log(1+V_N). \tag{5}$$

From Theorem 2.2 we already know that a standard Rosenblatt r.v. R with parameter $2H'-1$ exists such that $\lim_{N\to\infty}\mathbf{E}[|N^{2-2H'}V_N - cR|^2] = 0$. From (5) we immediately get

$$\mathbf{E}\big[\big|2N^{2-2H'}(H-\hat{H})\log N - cR\big|^2\big] \leqslant 2\mathbf{E}\big[\big|N^{2-2H'}V_N - cR\big|^2\big] + 2N^{4-4H'}\mathbf{E}\big[\big|V_N - \log(1+V_N)\big|^2\big].$$

The theorem follows by showing that $\mathbf{E}[|V_N - \log(1+V_N)|^2] = \mathrm{o}(N^{4H'-4})$, which is easily obtained. Indeed, this expectation is of order $\mathbf{E}[V_N^4]$, which, since V_N has a finite chaos expansion, is of order $(\mathbf{E}[V_N^2])^2 = \mathrm{O}(N^{8H'-8})$ by Lemma 2.1. □

References

[1] J. Beran, Statistics for Long-Memory Processes, Chapman and Hall, 1994.

[2] J.F. Coeurjolly, Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths, Statist. Inference Stoch. Process. 4 (2001) 199–227.

[3] I. Nourdin, G. Peccati, Stein's method on Wiener chaos, Probab. Theory Related Fields, 44 pages, available online, doi:10.1007/s00440-008-0162-x.

[4] D. Nualart, Malliavin Calculus and Related Topics, second ed., Springer, 2006.

[5] D. Nualart, S. Ortiz-Latorre, Central limit theorems for multiple stochastic integrals and Malliavin calculus, Stochastic Process. Appl. 118 (4) (2008) 614–628.

[6] C.A. Tudor, F. Viens, Variations and estimators for the selfsimilarity order through Malliavin calculus, Ann. Probab., in press.