Journal of Multivariate Analysis 168 (2018) 75–104

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Wavelet eigenvalue regression for n-variate operator fractional Brownian motion

Patrice Abry [a], Gustavo Didier [b,*]

[a] Laboratoire de Physique, CNRS, École Normale Supérieure de Lyon, Université Claude Bernard, F-69342 Lyon, France
[b] Department of Mathematics, Tulane University, New Orleans, LA 70118, USA

ARTICLE INFO

Article history:
Received 9 August 2017
Available online 4 July 2018

AMS 2000 subject classifications:
62M10
60G18
42C40

Keywords:
Eigenvalues
Operator fractional Brownian motion
Operator self-similarity
Wavelets

ABSTRACT

In this paper, we extend the methodology proposed in Abry and Didier (2018) to obtain the first joint estimator of the real parts of the Hurst eigenvalues of n-variate operator fractional Brownian motion (OFBM). The procedure consists of a weighted regression on the log-eigenvalues of the sample wavelet spectrum. The estimator is shown to be consistent for any time reversible OFBM and, under stronger assumptions, also asymptotically normal starting from either continuous or discrete time measurements. Simulation studies establish the finite-sample effectiveness of the methodology and illustrate its benefits compared to univariate-like (entry-wise) analysis. As an application, we revisit the well-known self-similar character of Internet traffic by applying the proposed methodology to 4-variate time series of modern, high quality Internet traffic data. The analysis reveals the presence of a rich multivariate self-similarity structure.

1. Introduction

An operator fractional Brownian motion (OFBM) $B_H = \{B_H(t)\}_{t\in\mathbb{R}}$ is a $\mathbb{R}^n$-valued Gaussian stochastic process with stationary increments that satisfies the operator self-similarity relation

$$\{B_H(ct) : t \in \mathbb{R}\} \overset{\mathcal{L}}{=} \{c^H B_H(t) : t \in \mathbb{R}\}, \tag{1}$$

valid for all $c > 0$, where $\overset{\mathcal{L}}{=}$ stands for the equality of finite-dimensional distributions. In relation (1), which generalizes the univariate concept of self-similarity, H is a $n \times n$ matrix called the Hurst matrix, and $c^H = \exp(\ln cH)$, where $\exp A = A^0/0! + A^1/1! + \cdots$ is the usual matrix exponential. If the Jordan form

$$H = PJ_H P^{-1} \tag{2}$$

is diagonalizable with real (Hurst) eigenvalues for a nonsingular P, then the eigenvectors form a coordinate system in which the qth marginal $\{B_H(t)_q\}_{t\in\mathbb{R}}$ of B_H, for $q \in \{1, \ldots, n\}$, is a fractional Brownian motion (FBM) with Hurst scaling index h_q (namely, a Gaussian, self-similar, stationary increment stochastic process); see [41,106]. These coordinate processes need not be independent. It is generally assumed that OFBM is proper, namely, its variance matrix $\mathrm{E}B_H(t)B_H(t)^*$ has full rank for $t \neq 0$.

* Corresponding author.
E-mail address: gdidier@tulane.edu (G. Didier).

https://doi.org/10.1016/j.jmva.2018.06.007
0047-259X/

OFBM is a multivariate fractional process. Univariate fractional processes have been used with great success in the modeling of data sets from many fields of science, technology and engineering; see, e.g., [24,46,61,76,107]. The literature on the probability theory and statistical methodology for univariate fractional processes is now voluminous; see, e.g., [5,12,14,15,25,31,39,47,49,55,77,85–87,94,96,97,103–105], to cite a few.

In modern applications, however, data sets are often multivariate, since several natural and artificial systems are monitored by a large number of sensors. Accordingly, the literature on multivariate fractional processes has been expanding at a fast pace. The contributions include [13,56,57,59,62–65,71,78,98,100] in the time and Fourier domains, and [7,9,29,48,110] in the wavelet domain. See also [58,79,89,99,102] on the related fractional cointegration literature in econometrics.

The framework of operator self-similar (o.s.s.) random processes and fields was originally conceived by Hudson and Mason [60], and by Laha and Rohatgi [66]; it has attracted much attention recently; see, e.g., [16,26,27,35–38,40,50,70,72,80,95,115]. If $H = \operatorname{diag}(h_1, \dots, h_n)$ and $P = I$ in (1), then the latter relation breaks down into simultaneous entry-wise expressions

$$\{B_H(ct) : t \in \mathbb{R}\} \overset{\mathcal{L}}{=} \{(c^{h_1} B_H(t)_1, \dots, c^{h_n} B_H(t)_n)^* : t \in \mathbb{R}\}, \tag{3}$$

where $c > 0$. Relation (3) is henceforth called *entry-wise scaling*. Several estimators have been developed by building upon the univariate-like, entry-wise scaling laws, e.g., the Fourier-based multivariate local Whittle [88,101] and the multivariate wavelet regression [8,110]. However, if H is non-diagonal, then the matrix P mixes together the several entries of B_H. In this case, the univariate-like statistical analysis of each entry of Y will often generate estimates that are undetermined convex combinations of Hurst eigenvalues or, at large scales, estimates of the *largest* Hurst eigenvalue; see, e.g., [4,22,34,108].

In [2], the use of the eigenstructure of wavelet variance matrices is proposed for the estimation of the Hurst parameters of OFBM. The main results are obtained in the bivariate context, in which it is shown that wavelet log-eigenvalues – and also wavelet eigenvectors, under assumptions – are consistent and asymptotically normal estimators of the eigenstructure of the Hurst matrix H.

In this paper, we extend this approach by proposing a wavelet eigenvalue regression estimator of the Hurst eigenvalues of n-variate OFBM. The estimator is shown to be consistent for the real parts of the eigenvalues of H for essentially any time reversible OFBM, using one wavelet octave. Under the stronger assumption that Hurst eigenvalues are real and simple (pairwise distinct), we further show that the wavelet eigenvalue regression estimator is asymptotically normal. Establishing the latter properties involves showing that the wavelet log-eigenvalues themselves are a consistent and asymptotically normal estimator of (the real parts of) the eigenvalues of the Hurst matrix; see Theorems 2–3 and Corollary 2—see also Remark 2 for an intuitive discussion of consistency. Under the additional assumption that the matrix of Hurst eigenvectors P (mixing, or coordinates, matrix) in (2) is orthogonal, a consistent sequence of wavelet eigenvectors is also shown to exist; see Corollary 1. With a view toward hypothesis testing, we also investigate conditions for asymptotic normality when all Hurst eigenvalues are equal (Proposition 2). The mathematical framework builds upon the Courant–Fischer variational characterization of eigenvalues (see (A.1)) and is much more general than that in [2], which relies on closed form expressions for eigenvalues and eigenvectors in dimension 2.

In the context of scaling properties, the use of eigenanalysis was first proposed by Meerschaert and Scheffler [83,84] for operator stable laws, and later by Becker-Kern and Pap [13] for o.s.s. processes in the time domain. It has also been used in the cointegration literature; see, e.g., [51,69,93,116]. The wavelet framework has well-documented benefits, such as: (i) robustness with respect to contamination by polynomial trends [30]; (ii) for a sample size ν, the fast wavelet transform may reach computational complexity $O(\nu)$, which is even lower than that of the fast Fourier transform [75]; (iii) quasi-decorrelation of several families of stochastic processes [10,11,44,81,85–87,114], which often leads to Gaussian confidence intervals. By contrast, for example, it is well known that sample covariance matrices are sensitive to contamination by trends, that they can be strongly dependent, and that estimators based on them may be asymptotically non-Gaussian; see [20], Theorem 1; [28], Proposition 1; or [94], Chapter 5.

Computational efficiency is especially important in a multivariate setting. The proposed wavelet eigenvalue approach permits a characterization of scaling laws in data without resorting to an optimization procedure and by using, instead, the fast computation of wavelet transforms and matrix eigenvalues; see, e.g., [21] on computational issues involved in multivariate Gaussian maximum likelihood. Moreover, the methodology is robust with respect to changes of coordinates (Hurst eigenvectors). This reduces the dimension of the relevant parameter space and can further serve to gauge scaling data analysis based on estimators that do account for coordinates or other parameters; see, e.g., [4].

To the best of our knowledge, we are proposing the first provably asymptotically normal eigenanalysis-based estimator of scaling or Hurst eigenvalues in general dimension n, under assumptions. In fact, the operator self-similarity property of the sample wavelet variance $W_a\{a(\nu)2^j\}$ (see Eq. (15)) points to the asymptotic normality of wavelet log-eigenvalues in view of that of the fixed scale sample wavelet variance $W_a(2^j)$. Based on Taylor expansions, establishing the asymptotic distribution of the largest wavelet variance log-eigenvalue (and the smallest, by taking the matrix inverse) is not particularly challenging since the wavelet eigenvalue limit scaling law is always driven by the largest Hurst eigenvalue. However, the scaling behavior of intermediate wavelet log-eigenvalues is typically dominated by higher divergent eigenvalues. Under conditions such as real and simple Hurst eigenvalues, we show that the latter issue can be tackled by a characterization of the asymptotic behavior of wavelet eigenvectors, whose convergence cancels explosive terms; see Proposition 1, part (iv). The most general case of multiple blocks of Hurst eigenvalues with algebraic multiplicity greater than 1 (see Section 2 on terminology) calls