

The Fractional Brownian Motion: Estimation and Approximation of Time Series

Bondarenko V*

Kyiv Polytechnic Institute, National Technical University of Ukraine, Ukraine

Abstract

In this paper we propose two problems which related to fractional Brownian motion. First problem- simultaneous estimation of two parameters-Hurst exponent and the volatility, that describes this random process. Numerical tests for the simulated fBm provided an efficient method. Second problem- approximation of the increments of observed time series with power function by increments from the fractional Brownian motion. Approximation and estimation have shown on the example of real data- daily deposit interest rates.

Keywords: Fractional brownian motion; Gaussian random process; Numerical experiment

Introduction

The purpose of research is the construction of mathematical models of observed actual data which characterizing dynamics of financial processes.

Consider x_1, x_2, \dots, x_n as the observed values of some quantity at some moments of time. Let us choose a random process $X(t)$, where

$$x_k = X\left(\frac{k}{n}\right)$$

Fractional Brownian motion is defined as a Gaussian random process with characteristics:

$$B_H(t), EB_H(t) = 0, B_H(0) = 0, EB_H(t)B_H(s) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H})$$

Smoothness of the trajectories of the process $B_H(t)$ is defined by the parameter H : almost all the trajectories satisfy the Holder condition:

$$|X(t) - X(s)| \leq c|t - s|^\alpha, \alpha < H,$$

which generalizes the known Levy's result for the Wiener process [1].

The assumption of non-Gaussian one-dimensional distributions can be easily met if we move on to the process

$$X(t) = g(B_H(t)),$$

where g is an odd increasing function. At the same time, the dimensional distributions $X(t)$ have a density

$$f_{X(t)}(y) = \frac{1}{\sqrt{2\pi t^H}} \exp\left\{-\frac{(g^{-1}(y))^2}{2t^{2H}}\right\} \frac{1}{g'(g^{-1}(y))}$$

Consider $X(t) = \sigma B^H(t)$, then examine the increment

$$y_k = X\left(\frac{k}{n}\right) - X\left(\frac{k-1}{n}\right), k = 1, 2, \dots, n.$$

The vector

$Y = (y_1, y_2, \dots, y_n) \sim N(0; V)$, where the correlation matrix

$$v_{jk} = \frac{\sigma^2}{n^{2H}} \left(\frac{|k-j+1|^{2H} + |k-j-1|^{2H}}{2} - |k-j|^{2H} \right) = \frac{\sigma^2}{n^{2H}} s_{jk};$$

$\{y_k\}$ constitutes a Gaussian stationary sequence. Henceforth, the increments are going to be the subject of consideration. Consider the algorithm proposed for simultaneous estimation of parameters.

We inspect the method for simultaneous estimation of two unknown parameters fBm (H, σ) and propose a method for approximation of the time series by the power function from the increments of fractional Brownian motion [2].

Description of the Algorithm

Consider the absolute random moments of the increments of fractional Brownian motion.

$$R_{jn} = \frac{1}{n} \sum_{k=1}^n |y_k|^j, j - real.$$

Then calculate the mean:

$$E_n(j) = ER_{jn} = \frac{\sigma^j}{n^{jH}} \cdot \frac{2^{j/2} \Gamma\left(\frac{j+1}{2}\right)}{\sqrt{\pi}}.$$

The result was first proved in [3]:

Theorem:

With probability 1

$$\frac{R_{jn}}{E_n(j)} \rightarrow 1$$

In particular, when $j = 1, \sqrt{\frac{\pi}{2}} \cdot \frac{n^H}{\sigma} R_{1n} \rightarrow 1,$

we obtain a consistent estimate for σ

$$\sigma_{1n} = n_H \sqrt{\frac{\pi}{2}} R_{1n} \tag{1}$$

with the known value estimate for H :

*Corresponding author: Bondarenko V, Kyiv Polytechnic Institute, National Technical University of Ukraine, Ukraine, Tel: +380 (44) 2366913; E-mail: valeria_bondarenko@yahoo.com

Received February 05, 2015; Accepted November 23, 2015; Published November 28, 2015

Citation: Bondarenko V (2015) The Fractional Brownian Motion: Estimation and Approximation of Time Series. J Appl Computat Math 4: 269. doi:10.4172/2168-9679.1000269

Copyright: © 2015 Bondarenko V. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

$$\hat{H}_n = \frac{\ln \left(\sqrt{\frac{2 \cdot \sigma}{\pi \cdot R_{1n}}} \right)}{\ln n}$$

ε is the canonical Gaussian vector with the following characteristics:

$$E\varepsilon = 0, E(\varepsilon, u)(\varepsilon, v) = (u, v) \dim \varepsilon = n$$

Then

$$y = V^{\frac{1}{2}}\varepsilon, \text{ therefore}$$

$$n = E(\varepsilon, \varepsilon) = E(V^{-1}y, y) = \frac{n^{2H}}{\sigma^2} E(S^{-1}y, y)$$

And consequently the statistic

$$\hat{\sigma}_{2n}^2 = (n)^{2H-1} (S^{-1}y, y);$$

and here statistics $(n)^{2H-1}(S^{-1}y, y)$ is an unbiased estimate of the parameter σ^2 (has been proved in [4]).

Now we prove consistency of this estimate. Let us introduce the notation:

$$\hat{\sigma}_{2n} = \sqrt{(n)^{2H-1} (S^{-1}y, y)} \quad (2)$$

We use the formula for integration by parts for Gaussian measures that leads to the relation (by calculating the dispersion $D\hat{\sigma}_{2n}^2 = n^{4H-2} E(S^{-1}y, y)^2 - \sigma^4$):

$$D\hat{\sigma}_{2n}^2 = \frac{2\sigma^4}{n}$$

Where $\hat{\sigma}_{2n}$ is consistent estimate of the parameter σ .

Equations (1) and (2) form a system, which is proposed to solve iteratively. The essence of the algorithm is as follows: for an arbitrary value $H \in (0;1)$, let us calculate the estimate $\hat{\sigma}_{1n}$, matrix S^{-1} and the estimate $\hat{\sigma}_{2n}$. Then we iterate the values H (matrix S for different H) with some step

$$\frac{\hat{\sigma}_{2n}}{\hat{\sigma}_{1n}} = \frac{0,8}{R_{1n}} \sqrt{(S^{-1}y, y)} \approx 1, \left| \frac{\hat{\sigma}_{2n}}{\hat{\sigma}_{1n}} - 1 \right| \rightarrow \min \cdot \quad (3)$$

The values \hat{H} of parameter H , which satisfies (3), is an estimate and $\hat{\sigma} = \frac{\hat{\sigma}_{1n} + \hat{\sigma}_{2n}}{2}$

We performed a numerical experiment, which implements the algorithm proposed.

Numerical Experiment

The input data is received by the simulating of modeling values

of fractional Brownian motion, which uses description fBm by the stochastic integral [5]. These values are determined by the formula:

$$B_H \left(\frac{k}{n} \right) = \frac{c}{n^H} \left[\sum_{j=1}^k ((k+j)^\alpha - j^\alpha) \varepsilon_j + \sum_{i=0}^{k-1} (k-i)^\alpha \varepsilon_{n^2+i+1} \right] \equiv \frac{x_k}{n^H}, k = 0, 1, 2, \dots, n$$

and the increments of fBm are as follows:

$$y_k = B_H \left(\frac{k}{n} \right) - B_H \left(\frac{k-1}{n} \right) = \frac{1}{n^H} (x_{k+1} - x_k) = \frac{z_k}{n^H}, k = 1, 2, \dots, n,$$

where the normalized increments are defined with

$$z_k = c_H \sum_{j=1}^k ((k+1+j)^\alpha - (k+j)^\alpha) \varepsilon_j + c_H \sum_{i=0}^{k-1} ((k+1-i)^\alpha - (k-i)^\alpha) \varepsilon_{\frac{n^2}{2}+i+1} + \varepsilon_{\frac{n^2}{2}+k+1} \quad (4)$$

And the random variables

$$\{\varepsilon_r\}_{r=1}^{m+n} \text{ are independent, } \varepsilon_r \sim N(0;1).$$

The accuracy of the approximation is investigated by Coeurjolly [6]. In particular, numerical experiment has showed that $n \geq 1000$, $a = n^{3/2}$, $m = n^{5/2}$. Let's assume new equality:

$$B_H \left(\frac{k}{n} \right) = \frac{c}{n^H} \left[\sum_{j=1}^k ((k+j)^\alpha - j^\alpha) \varepsilon_j + \sum_{i=0}^{k-1} (k-i)^\alpha \varepsilon_{n^2+i+1} \right] \equiv \frac{x_k}{n^H}, k = 0, 1, 2, \dots, n \quad (5)$$

The increments

$$y_k = B_H \left(\frac{k}{n} \right) - B_H \left(\frac{k-1}{n} \right) = \frac{1}{n^H} (x_{k+1} - x_k) = \frac{z_k}{n^H}, k = 1, 2, \dots, n$$

Let's normalize the statistics R_{jn}

$$r_{jn} = n^{jH} R_{jn} = \frac{1}{n-1} \sum_{k=1}^{n-1} |z_k|^j$$

$$r_{jn} = \frac{\sigma^j 2^{j/2} \Gamma \left(\frac{j+1}{2} \right)}{\sqrt{j\pi}}, n \rightarrow \infty;$$

$$r_{1n} \rightarrow \sigma \sqrt{\frac{2}{\pi}} \approx 0,80; r_{2n} \rightarrow \sigma^2$$

Let's calculate z_{1n}, z_{n-1n} by formula:

$$z_k = c_H \sum_{j=1}^{n^2-k} ((k+1+j)^\alpha) \varepsilon_j + c_H \sum_{i=0}^{k-1} ((k+1-i)^\alpha - (k-i)^\alpha) \varepsilon_{\frac{n^2}{2}+i+1} + \varepsilon_{\frac{n^2}{2}+k+1} \quad (6)$$

$$\hat{\sigma}_{1n} = \sqrt{\frac{\pi}{2}} r_{1n}, \hat{\sigma}_{2n} = \sqrt{\frac{(S^{-1}z, z)}{n}}$$

Table 1 shows the values of the statistics r_{1n}, r_{2n} and r_{1n}^2/r_{2n} for $n = 256; n = 1024$ and for $H \in [0,1; 0,9]$ with step 0,1 [7].

Table 1 also shows the kurtosis $d_n = \frac{r_{1n}^2}{r_{2n}}$. For Gaussian

H	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9
n=256								
r_{1n}	0.62	0.68	0.79	0.77	0.71	0.88	0.76	0.75
r_{2n}	0.79	0.87	1.02	0.97	1.09	1.03	0.91	0.89
d_n	0.49	0.53	0.61	0.61	0.66	0.64	0.63	0.63
n=1024								
r_{1n}	0.72	0.76	0.82	0.83	0.83	0.82	0.82	0.83
r_{2n}	0.8	0.97	1.04	1.05	1.07	1.03	1.06	0.98
d_n	0.64	0.60	0.64	0.65	0.64	0.65	0.63	0.70

Table 1: The values of generated statistics - the Hurst coefficient.

Hk/Hj	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9
n=256								
0.4	1.09	1.10	1.05	1.00	1.12	1.17	1.23	1.28
0.7	1.10	0.92	0.91	0.89	0.95	1.01	1.10	1.17
N=1024								
0.3	0.85	0.89	1.04	1.20	1.4	1.37	1.42	1.48
0.6	1.18	1.12	1.20	1.23	1.04	1.15	1.23	1.25
0.7	1.20	1.18	1.15	1.14	1.13	1.02	0.93	0.89
dn	0.64	0.60	0.64	0.65	0.64	0.65	0.63	0.70

Table 2: The values of efficiency of the estimation algorithm.

$\{z_k\} Ed_n = \frac{2}{\pi} \approx 0,637$, and the values of d_n are close to the theoretical ones [8].

The numerical results verified by effectiveness of the proposed estimation algorithm are summarized in Table 2.

$$c_{kj} = \frac{0,8}{r_1} \sqrt{\frac{(S_j^{-1} z_k, z_k)}{n}}$$

z_k is generated vector of increments fBm; S_j is normalized correlation matrix corresponding the Hurst exponents H_k, H_j . For each H_k (in the fixed line) c_{kj} are calculated the selection of parameter H_j with step $\Delta H_j = 0,1$. The generation of z_k is performed for the values [9]:

$$n = 256; H_k = 0,4; H_k = 0,7; n = 1024; H_k = 0,3; H_k = 0,6; H_k = 0,7.$$

Analysis of data in the table shows that for each H_k

$$|c_{kj} - 1| \rightarrow \min, H_j = H_k, \hat{H}_k = H_j,$$

which shows effectiveness of the algorithm.

The approximation of real time series with using fBm

We should choose an adequate model of the corresponding random process $X(t)$ for actually observed time series. We consider empirical method of checking the hypothesis about normality and if this hypothesis is rejected, than approximation of real time series with using fBm.

Let's assume $S_1..S_n$ as observed time series data of an arbitrary nature. Consider the procedure for its approximation of fractional Brownian motion [10]. The first step of the algorithm - initial data processing, is to bring them to a time series with zero trend $x_k = s_k - M(k)$, $M(k)$ is an approximation of the trend.

Let us select an adequate model of the corresponding random process for the actually observed time series:

$x_0 = 0, x_1..x_n$ - the observed values, $y_k = x_k - x_{k-1}, k = 2..n, Ex_k = 0$. The criterion of the value of the Gaussian increments will be the "excess coefficient"

$$d_n = \frac{\left(\frac{1}{n-1} \sum_{k=2}^n |y_k|\right)^2}{\frac{1}{n-1} \sum_{k=2}^n y_k^2}$$

If d_n is significantly different from $2/\pi$, let's replace the time series $y_1..y_{n-1}$ by the new sequence $z_1..z_{n+1}$. The general idea of approximation is in dimensional functional transformation g for each increment y_k , g - is the increasing odd function.

$$z_k = g(y_k)$$

Consider two-dimensional distribution $(y_k; y_j)$ with Gaussian density:

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma^2} \exp\left\{-\frac{x^2+y^2}{2\sigma^2(1-\rho^2)}\right\} \exp\left\{-\frac{\rho xy}{\sigma^2(1-\rho^2)}\right\},$$

$$\rho = \rho_{k-j} = \frac{1}{\sigma^2} E y_k y_j$$

$$\text{If } \rho > 0, \text{ then } f(x_1, y_1) \geq f(x_2, y_2), |x_1| = |x_2|, |y_1| = |y_2|, x_1 y_1 > 0, x_2 y_2 < 0$$

$$\sigma y_n E(\gamma(\Psi_k) \gamma(\Psi_j)) = \sigma \gamma \gamma \rho$$

Assume that a two-dimensional density δ ιστριβυτιον $\varphi_{k-j}(\xi, y)$ of increments $(y_k; y_j)$ of converted time series satisfies the following conditions:

$$1) \varphi(x) = \int_{-\infty}^{\infty} f_{k-j}(x, y) dy, E y_k = 0;$$

$$2) \text{sgn} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x) g(y) f_{k-j}(x, y) dx dy = \text{sgn} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{k-j}(x, y) dx dy.$$

$$\lim_{n \rightarrow \infty} \frac{\left(\frac{1}{n-1} \sum_{k=1}^{n-1} |g^{-1}(y_k)|\right)^2}{\frac{1}{n-1} \sum_{k=1}^{n-1} (g^{-1}(y_k))^2} = \frac{2}{\pi}$$

$z_k = \gamma^{-1}(y_k)$ is assumed to be a Gaussian random variable. Let's demonstrate the algorithm by a power function g .

Let's assume

$$z_k = \text{sgn } y_k |y_k|^{\lambda}, y_k = \text{sgn } y_k |z_k|^{\lambda}, \lambda > 0.$$

$$\text{Then } \left(\frac{1}{n-1} \sum_{k=1}^{n-1} |z_k|^{\lambda}\right)^2, d_n = \frac{\left(\frac{1}{n-1} \sum_{k=1}^{n-1} |z_k|^{\lambda}\right)^2}{\frac{1}{n-1} \sum_{k=1}^{n-1} |z_k|^{2\lambda}}$$

$$E|\xi|^{\alpha} = \frac{2^{\frac{\alpha}{2}}}{\sqrt{\pi}} \sigma^{\alpha} \Gamma\left(\frac{\alpha+1}{2}\right), \text{TO}d = \frac{1}{\sqrt{\pi}} \frac{\Gamma^2\left(\frac{\lambda+1}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right)}$$

Thus, the proposed approximation leads to the following model of original time series:

$$x_k = \sum_{j=1}^k \text{sgn } y_j \cdot |z_j|^{\lambda}.$$

From the assumption of stationary the sequence $\{y_k\}$ follows the stationary of $\{z_k\}$

$$Ez_j z_k = Dz_j \cdot f\left(\frac{k-j}{n}\right), f(0) = 1, f(s) \text{ - is a decreasing function.}$$

Let's approximate f by the exponential function:

$$\frac{\left(s + \frac{1}{n}\right)^{2H} + \left(s - \frac{1}{n}\right)^{2H}}{2} - s^{2H},$$

and approximate $\{x_k\}$ by the fractional Brownian motion:

$$z_k = \sigma^H \left(B_H\left(\frac{k}{n}\right) - B_H\left(\frac{k-1}{n}\right) \right), x_k = \sigma^H \sum_{j=1}^k \text{sgn } y_j \left| B_H\left(\frac{j}{n}\right) - B_H\left(\frac{j-1}{n}\right) \right|^2,$$

The approximation procedure of real time series is based on the value for one parameter d_n .

As an example consider the financial data: 1678 values of interest rate, divided into 9 time windows. The interest rate is given by the following formula in each time window.

$$S(t) = a + b \exp\{X(t)\},$$

a - interest rate of the National Bank of Ukraine,

For the discrete time (Table 3).

$$x_k = \ln \frac{S_k - a}{S_1 - a}, x_1 = 0.$$

The Results of Calculations

For $y_1 \dots y_n$ we calculate the parameter in each time window λ , solving the equation (Table 4).

$$\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma^2\left(\frac{\lambda+1}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right)} = d_n.$$

Now we consider the converted data $(z_1, z_2, \dots, z_{336})$ from the third time window assuming

$$z_k = \text{sgn } y_k \sqrt{|y_k|} = \sigma \left(B_H\left(\frac{k+1}{n}\right) - B_H\left(\frac{k}{n}\right) \right),$$

and estimate the parameters σ, H , using the proposed algorithm:

$$r_n = \frac{1}{n} \sum_{k=1}^n |z_k| = \frac{1}{n} \sum_{k=1}^n \sqrt{|y_k|} \quad (r_{336} = 0,478), \hat{\sigma}_{1n} = \sqrt{\frac{\pi}{2} n^H} r_n, \hat{\sigma}_{2n} = n^H \sqrt{\frac{(S^{-1}z, z)}{n}}$$

$$\frac{\hat{\sigma}_{2n}}{\hat{\sigma}_{1n} r_n} = \frac{0,8}{r_n} \sqrt{\frac{(S^{-1}z, z)}{n}} = 0,0913 \sqrt{(S^{-1}z, z)}$$

Select the step size $\Delta H = 0,05$.

For $H=0,25$ we have:

$$(S^{-1}z, z) = 117,6; \frac{\hat{\sigma}_2}{\hat{\sigma}_1} = 0,99$$

For $H=0,3$:

$$(S^{-1}z, z) = 117,6; \frac{\hat{\sigma}_2}{\hat{\sigma}_1} = 1,007$$

A further search of H gives the values $\frac{\hat{\sigma}_2}{\hat{\sigma}_1}$, which differ from 1, so

$$\hat{H} = 0,3, \hat{\sigma} = 3,42$$

Let's perform the similar calculations for the 8th time window (n = 295):

$$z_k = \text{sgn } y_k |y_k|^{0,476}, k = 1, \dots, 294, r_n = \frac{1}{294} \sum_{k=1}^{294} |z_k| = 0,4;$$

$$\frac{\hat{\sigma}_2}{\hat{\sigma}_1} = \frac{0,8}{0,4} \sqrt{\frac{(S^{-1}z, z)}{294}} = 0,117 \sqrt{(S^{-1}z, z)}$$

No window	1	2	3	4	5	6	7	8	9
a - rate	0.09	0.095	0.085	0.08	0.1	0.12	0.11	0.103	0.078
n - number of data in window	110	160	337	207	91	250	57	295	171
R1n	0.631	0.694	0.425	0.264	0.547	0.345	0.238	0.277	0.137
R2n	0.897	1.098	0.541	0.195	0.788	0.384	0.142	0.248	0.097
dn	0.444	0.439	0.334	0.357	0.380	0.310	0.390	0.309	0.193

Table 3: The results of calculations.

No window	1	2	3	4	5	6	7	8	9
λ	1.55	1.58	2	1.9	1.8	2.1	1.85	2.1	2.8

Table 4: Converted data from window assuming.

H	0.1	0.2	0.3	0.4	0.6	0.7	0.8	0.9
A	0.28	0.22	0.16	0.08	0.1	0.21	0.34	0.43
B	0.64	0.61	0.58	0.54	0.45	0.4	0.33	0.23
n = 256								1.17
r1n	0.5	0.68	0.73	0.77	0.85	0.81	0.80	0.74
r2n	0.51	0.87	0.85	0.97	1.09	1.03	1.02	0.86
dn	0.49	0.53	0.63	0.61	0.66	0.64	0.63	0.64
Fn	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.05
Qn	0.34	0.36	0.36	0.4	0.45	0.48	0.53	0.6
n = 1024								
r1n	0.72	0.76	0.74	0.80	0.79	0.80	0.80	0.81
r2n	0.8	0.97	0.90	1.01	1.03	0.98	1.03	0.94
dn	0.64	0.60	0.61	0.63	0.61	0.65	0.62	0.70
Fn	0.015	0.015	0.02	0.02	0.03	0.03	0.04	0.05
Qn	0.28	0.31	0.26	0.30	0.42	0.46	0.59	0.72

Table 5: Applications 1.

The least deviation of the relation $\frac{\hat{\sigma}_2}{\hat{\sigma}_1}$ from 1 is calculated at $H = 0,3$ (Tables 5 and 6).

$$(S^{-1}z, z) = 75,7; \frac{\hat{\sigma}_2}{\hat{\sigma}_1} = 1,0118 \cdot$$

$$\hat{H} = 0,3, \hat{\sigma} = 2,77 \cdot$$

Conclusion

Simultaneous estimation and approximation of the increments of observed time series with power function by increments from the fractional Brownian motion have been described.

References

1. Barndorff-Nielsen OE (2001) Levy processes: Theory and applications. Birkhauser pp: 421.
2. Mishura Y (2008) Stochastic Calculus for Fractional Brownian Motion and Related Processes. In: Barndorff-Nielsen OE (eds.) Lecture Notes in Mathematics 1929. Springer, Verlag pp: 392.
3. Peltier RF, Levy, Vehel JL (1994) A new method for estimating the parameter of fractional Brownian motion. Rapport de recherche de l'INRIA pp: 27.
4. Bogachev V (1998) Gaussian Measures. AMS Mathematical Surveys and Monographs 62.
5. Mandelbrot BB, Van Ness JW (1968) The Fractional Brownian motion, fractional noises and applications. SIAM Review 10: 422-437.
6. Coeurjolly JF (2000) Simulation and identification of the fractional Brownian motion: A bibliographical and comparative study. Journal of statistical software 5: 1-52.
7. Beran J (1995) Statistics for Long-Memory Processes. Chapman and Hall pp: 315.
8. Mandelbrot BB, Wallis JR (1969) Computer experiments with fractional Gaussian noises I, II, III. Water Resources Research 5: 228-267.
9. Coeurjolly JF (2008) Hurst estimation of locally self-similar Gaussian processes using sample quantiles. The Annals of statistics 36: 1404-1434.
10. Breton JC, Coeurjolly JF (2012) Confidence intervals for the Hurst parameter of a fractional Brownian motion based on concentration inequalities. Statistical Inference for Stochastic Processes 15: 1-26.