The Power Approximation of Time Series with Using Fractional Brownian Motion

Bondarenko V*
Kyiv Polytechnic Institute, Institute of Applied System Analysis, Technical University of Ukraine, Ukraine

Abstract

We propose the approximating sequence and some of characteristics of this sequence to coincide with the increments of the fractional Brownian motion (fractional Brownian noise) for the observed time series. We study the Hurst parameter estimation algorithm and check the quality of the approximation.

Keywords: The parameters estimation of Fbm; Description of the algorithm; Approximation of real time series

Introduction

We consider a mathematical model for the time series $S\ldots S$.

We consider a mathematical model for the time series $S\ldots S$.

1. Process should possess known characteristics, particularly the Gaussian process should be chosen.

2. $X(t)$ should not be Markov, as the Markov communication does not provide an adequate description of real phenomena.

Fractional Brownian motion is defined as a Gaussian random process with characteristics [1]:

$$B_m(t), E B_m(t) = 0, B_m(t) = \frac{1}{2}(t^{2m} + s^{2m} - |t - s|^{2m})$$

Smoothness of the trajectories of the process $B_m(t)$ is determined 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.

The statement of problem to build a forecast for the initial time series.

The Parameters Estimation of Fbm

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

$$y = X_k - X_{k-1} = X_{k-1} - X_{k-2} = \ldots = X_{k-n} = \ldots = X_0 = X_{-1} = \ldots = X_{-n},$$

The vector

$$Y = (y_1, y_2, \ldots, y_n)^T \sim N(0, Y),$$

where the correlation matrix

$$\sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} = \sigma^2$$

$$\sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} = \sigma^2$$

$$y_i \sim \text{Gaussian stationary sequence. Henceforth, the increments are going to be the subject of consideration. Consider the algorithm proposed for simultaneous estimation of parameters [2].}$$

We check the method for simultaneous estimation of two unknown parameters Fbm ($H, \sigma$) and propose a method for approximation of the time series by the power function from the increments of fractional Brownian motion.

Description of the Algorithm

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

$$R_{\alpha} = \frac{1}{n} \sum_{j=1}^{n} |y_j|$$

Then calculate the mean

$$E\alpha = E R_{\alpha} = \frac{\sigma^2}{\sigma^2} \frac{1}{n} \left( \frac{J + 1}{2} \right)$$

The result was first proved in eqn. (6).

Theorem

With probability 1

$$\frac{R_{\alpha}}{E\alpha} \to 1$$

In particular, when $f = 1 - \sqrt{\frac{\pi}{2} \frac{n}{\sigma}}$, we obtain a consistent estimate for $\sigma$

$$\hat{\sigma} = \frac{\sigma}{\sqrt{\pi}} \frac{n}{R_{\alpha}}$$

with the known value estimate for $H$:

$$\hat{H} = \frac{\ln R_{\alpha}}{\ln n}$$

$\varepsilon$ is the canonical Gaussian vector with the following characteristics:

*Corresponding author: Valeria Bondarenko, Technical University of Ukraine, Kyiv Polytechnic Institute, Institute of Applied System Analysis, Ukraine, Tel: +213 33 31 91 34; E-mail: valeria_bondarenko@yahoo.com

Received April 18, 2017; Accepted April 21, 2017; Published April 30, 2017


Copyright: © 2017 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.
\[ \sigma_{2n}^2 = \langle (S^{-1} y, y) \rangle \]  

(1)

We use the formula for integration by parts for Gaussian measures that leads to the relation (by calculating the dispersion
\[ D \sigma_{2n}^2 = n^{-1/2} E(S^{-1} y, y)^2 - \sigma^2 \]

(2)

where \( \sigma^2 \) is consistent estimate of the parameter \( \sigma \).

Equations (1) and (2) form a system, which is proposed to solve iteratively [3]. The essence of the algorithm is as follows:

Let us calculate the estimate \( \hat{S} \), matrix \( S^{-1} \) and the estimate \( \hat{S}_{2n} \). Then we iterate the values \( H \) (matrix \( S \) for different \( H \)) with some step

\[ \hat{S} = \hat{S}_{1n} + \hat{S}_{2n} \]

(3)

(\( 0.8 \))

(4)

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

We proved that the estimate of parameter \( \sigma \)

\[ \hat{S}_{2n} = \sqrt{n^{2H-1}} \langle S^{-1} y, y \rangle \]

(3) is consistent.

Let us calculate the estimate \( \hat{S}_{1n} \), matrix \( S^{-1} \) and estimate \( \hat{S}_{2} \).

\[ \hat{S}_{2n} = \frac{0.8}{R_{1n}} \sqrt{n^{-2} \langle S^{-1} y, y \rangle \approx 1. \hat{S}_{2n}/\hat{S}_{1n} - 1} \rightarrow \min. \]

(4)

The generated increments are defined by
\[ z_n = n^{-1/2} y_n, \quad n = 256; n=1024 \]

The numerical results verified by effectiveness of the proposed estimation algorithm are summarized in Tables 1-3.

Approximation of Real Time Series

Let’s \( S_1, S_2 \) be observed time series data of an arbitrary nature. Consider the procedure for its approximation of fractional Brownian motion. The first step of the algorithm – initial data processing is to consider the correlation matrix corresponding the Hurst exponents \( H_j, H_k \).

Consider the procedure for its approximation of fractional Brownian motion. The first step of the algorithm – initial data processing is to calculate the selection of parameter \( H_j \) with step \( \Delta H = 0.1 \). The generation of \( Z_n \) is performed for the values
\[ n=256; H_i = 0.4; H_i = 0.7; \]
\[ n=1024; H_i = 0.3; H_i = 0.6; H_i = 0.7. \]

Analysis of data in the table shows that for each \( H_j \)

\[ |c_{ij} - 1| \rightarrow \min, \quad H_j = H_i, \quad H_k = H_i, \]

which shows effectiveness of the algorithm.

Table 1: The values of efficiency of the estimation algorithm.
The general idea of approximation is in dimensional functional transformation $g$ for each increment $y_i$, $g$ is the increasing odd function.

$$z_i = g(y_i)$$

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

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

$$\rho = \rho_{k-j} = \frac{1}{\sqrt{2\pi}} \mathbb{E} y_k y_j.$$  

If $\rho > 0$, then $f(x_1, y_1) \geq f(x_2, y_2)$ if $|x_1| = |y_2| > |y_1|, x_1 y_1 > 0, x_2 y_2 < 0$.

$s\mathbb{E}(g(y_k) g(y_j)) = \mathbb{E} \rho$.

Assume that a two-dimensional density distribution $f_{x,y}$ of increments $(y_i; y_j)$ of converted time series satisfies the following conditions:

1) \( \varphi(x) = \int f_{x,y}(x,y) dy \)

2) \( \mathbb{E}(g(y_k) f_{x,y}(x,y) dx dy) = \mathbb{E}(g(y_k) f_{x,y}(x,y) dx dy) \)

$$\lim_{n \to \infty} \left( \frac{1}{n} \sum_{i=1}^{n-1} |y_i|^2 \right) = \frac{2}{\pi},$$

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

Let’s assume

$$z_k = \text{sgn} y_k |y_k|^\frac{1}{2}, y_k = \text{sgn} y_k |y_k|^\frac{1}{2}, \lambda > 0.$$  \hspace{1cm} (6)

Then

$$d_m = \left( \frac{1}{n-1} \sum_{i=1}^{n-1} |y|^\lambda \right)^2,$$

$$\frac{1}{n-1} \sum_{i=1}^{n-1} |y|^\lambda,$$

$$\mathbb{E}[|y|^\lambda] = \left( \frac{2^\frac{\lambda}{2}}{\sqrt{\pi}} \sigma^\lambda \Gamma(\frac{\alpha+1}{2}) \right)^2 \mathbb{E}(f(x) \sigma dy),$$

$$d = \frac{1}{2} \frac{\Gamma(\lambda + 1)}{\Gamma(\lambda + \frac{1}{2})}.$$  

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

$$X_k = \sum_{j=1}^{k} \text{sgn} y_j |y_j|^\frac{1}{2}.$$  

From the assumption of stationarity the sequence $\{y_t\}$ follows the stationarity of $\{Z_t\}$:

$$\mathbb{E} z_k = \mathbb{E} \left( \sigma\sum_{j=1}^{k} \text{sgn} y_j \left( B_n \left( \frac{k-j}{n} \right) - B_n \left( \frac{k-l}{n} \right) \right) \right).$$

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

To check the properties of increments $Z_t$ allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let’s approximate $f$ by the exponential function:

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

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

$$z_k = \sigma^\lambda \left( B_n \left( \frac{k-j}{n} \right) - B_n \left( \frac{k-l}{n} \right) \right),$$

$$\mathbb{E} z_k = \mathbb{E} \left( \sigma\sum_{j=1}^{k} \text{sgn} y_j \left( B_n \left( \frac{k-j}{n} \right) - B_n \left( \frac{k-l}{n} \right) \right) \right).$$

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

In the studies mentioned above there have been proved the following limit relations:

1. \( \frac{1}{n^\lambda} \sum_{k=0}^{n} f(B_n \left( \frac{k+j}{n} \right)) \to \int_0^1 f(B(s)) ds, \quad H \in (0, \frac{1}{2}); \)

2. \( \frac{2}{n^\lambda} \sum_{k=0}^{n} f(B_n \left( \frac{k}{n} \right)) \to \int_0^1 f(B(s)) ds, H \in (0, \frac{1}{2}); \)

3. \( \frac{1}{n^\lambda} \sum_{k=0}^{n} f(B_n \left( \frac{k}{n} \right)) \to \int_0^1 \frac{1}{2} f(B(s)) ds, H \in (0, \frac{1}{2}); \)

$$\mathbb{E} \left( \sum_{j=1}^{k} \text{sgn} y_j \left( B_n \left( \frac{k-j}{n} \right) - B_n \left( \frac{k-l}{n} \right) \right) \right) \to 0.$$

$$\alpha_n = n^\lambda B_n \left( \frac{k-j}{n} \right) \to 0.$$

and assuming in the first formula $f(x) = x^2$, let’s write

<table>
<thead>
<tr>
<th>$H$</th>
<th>$n$</th>
<th>$An$</th>
<th>$Bn$</th>
<th>$Cn$</th>
<th>$Dn$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>256</td>
<td>-1.19</td>
<td>0.25</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>-2.1</td>
<td>0.17</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4096</td>
<td>-1.6</td>
<td>0.08</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>256</td>
<td>-1.07</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>-1.85</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4096</td>
<td>-1.48</td>
<td>0.072</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>256</td>
<td>-1.14</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>-1.76</td>
<td>0.082</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4096</td>
<td>-1.54</td>
<td>0.063</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>256</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4096</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>256</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>0.072</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4096</td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>256</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1024</td>
<td>0.019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4096</td>
<td>0.031</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>$H$</th>
<th>$n$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.1</td>
<td>2.1</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
<td>1.3</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>1.01</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4</td>
<td>0.91</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5</td>
<td>0.96</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>0.97</td>
</tr>
<tr>
<td>0.2</td>
<td>0.7</td>
<td>1.02</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Table 3: The values of efficiency of the estimation algorithm.
\[
\frac{1}{n} \sum_{k=1}^{n} a_k z_k^3 \to \frac{3}{2}, \quad H \in (0; \frac{1}{2})
\]
\[
\frac{1}{n^{2H}} \sum_{k=1}^{n} a_k^2 \xi_k^3 \to 0, \quad H \in (0; \frac{1}{2})
\]
\[
\eta \approx N\left(0; \frac{1}{2H+2}\right)
\]
In the second formula \( f(x) = x^2 \),
\[
\frac{1}{n} \sum_{k=1}^{n} a_k^2 \xi_k^3 \to 3\eta, \quad H \in (0; \frac{1}{2})
\]
(3.18)

The third formula for \( f(x) = x \) and \( f(x) = 1 \):
\[
\frac{1}{n^{2H}} \sum_{k=1}^{n} a_k^2 \xi_k \to 3B^2 (l) , \quad H \in (\frac{1}{2}, 1)
\]
(3.19)

Let's consider the application of algorithm.

As a first example consider the data: the daily data of solar activity (366 data, ftp://ftp.ngdc.noaa.gov). We calculated the initial time series \( x_1, x_{365} \) with increments \( y_k = x_{k+1} - x_k, k = 1, 2 \ldots 365 \).

\( R_a = 0.425; \quad R_s = 0.541; \quad d_a = \frac{R_a}{R_s} = 0.334 \)
\[
1 - \frac{\Gamma\left(\lambda + \frac{1}{2}\right)}{\sqrt{\pi} \Gamma\left(\lambda + \frac{1}{2}\right)} = 0.334; \quad \lambda = 2
\]
\( y_k = \text{sgn} y_k \left|x_k\right|^{\frac{1}{3}} \); \quad \text{z}_k = \text{sgn} y_k \left[\left|y_k\right|\right]^{\frac{1}{3}}
\]

Applying the estimation algorithm (4), we obtain \( \hat{H} = 0.3 \)

The second example: 1678 values of interest rate. The interest rate is achieved by 

\( Y = X_{k+1} - X_k \)

\( R_a = \frac{1}{119} \sum_{i=1}^{119} |y_i| = 1.39; \quad R_s = \frac{1}{119} \sum_{i=1}^{119} y_i^2 = 3.3; \quad d_s = \frac{R_s}{R_a} = 0.59 \)

Applying the method of Hurst parameter estimation by the formula (4), let's calculate the equation

\[
\hat{\sigma}_2 = \frac{0.8}{n^{\frac{1}{4}}} \left(\frac{\hat{S}^\lambda z^3}{n}\right)
\]

For the converted sequence defined by the equation

\( r_i = \frac{1}{n} \sum z_i = 0.76; \quad r_a = \frac{1}{n} \sum z_i = 0.84; \quad d_a = 0.68 \)

Let's calculate the value of the quantity

\[
\hat{\sigma}_2 = \frac{0.8}{n^{\frac{1}{4}}} \left(\frac{\hat{S}^\lambda z^3}{n}\right)
\]

In this example, the minimum \( \left|\sigma_2 - 1\right| \) achieved by \( H=0.3; 0.8 \).

Let's check the quality of the approximation for each example. For the first example

\[
c = \frac{1}{n} \sum z_i^2 = 0.144
\]

\( D_a = \frac{1}{n} \sum z_i = 0.002; \quad d_a = 0.1 \)

\( F(x) = 2 \Phi\left(\frac{2}{\sqrt{3}} x\right) - 1, x > 0 \)

Let's choose a level of significance 0.1.
The approximation of the time series is satisfactory and hypothesis $T$ is accepted.

For the second example - the values of the bank interest rates

$n = 336, \ H = 0,3; \ c = \frac{1}{n} \sum_{i=1}^{n} z_i^2 = 0,425$

$A_6 = \frac{1}{335} \sum_{i=1}^{335} y_i z_i^2 = -0,16; \ B_6 = \frac{1}{335} \sum_{i=1}^{335} y_i^2 z_i^2 = -0,08$

The theoretical limit value

$$\lim_{n \to \infty} A_6 = -1,5 c^2 = -0,27$$

$$B_6 c^\frac{5}{2} = -0,68$$

$$B_6 c^\frac{5}{2} < \beta_1 = 3,07$$

The hypothesis $T$ is accepted.

For the third example - the market rates of Bundesbank Germany.

$n = 119; \ H = 0,4; \ c = \frac{1}{n} \sum_{i=1}^{n} z_i^2 = 3,3$

$A_6 = -19,1; \ B_6 = -0,8; \ A_6 c^{-2} = -1,75; \ B_6 c^{\frac{5}{2}} = -0,04$

The hypothesis $T$ is accepted.

In the fourth example, we have received the uncertain answer. We calculate the value of statistics

$$A_6 = \frac{1}{n} \sum_{i=1}^{n} y_i z_i^2 = -0,69$$

For $H > \frac{1}{2} \ A_6 \geq 0$, then we should take a decision of hypothesis $H=0,3$ (The theoretical value $A_6$ is equal $-1,5.0,84^2=-1,0$)

**Conclusion**

Thus, we proposed in this research a new method for estimate of parameters $f_{Bm}$, approximation of the time series and types of mathematical models.

We checked the adequacy of models and efficiency of the algorithms on real examples.

**References**