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 Gaussian 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 $S_n$=$S_0$+$X(t)$. The primary processing is smoothing, removing the trend - leads to improved time series $x_1$,$x_2$,$x_n$ ("initial"). Consider $x_1$,$x_2$,$x_n$ as the observed values of some quantity at some moments of time. Let us choose a random process $X(t)$, where $x_n$=$x(t_n)$ . This problem has a controversial solution, because different processes (with different distributions) may have the same trajectories. The subjective criteria for selection of $X(t)$ is as follows.

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]:

$$B^H(t), \quad E[B^H(t)]=0, \quad E[B^H(t)B^H(s)]=\frac{1}{2}(c^{2n}+s^{2n}-|t-s|^{2n})$$

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

$$|X(t)-X(s)|\leq C|t-s|^{H}, \quad H<H_1,$$

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

$$y_n = X\left(\frac{k}{n}\right) = -X\left(\frac{k-1}{n}\right), \quad K=1,2,\ldots,n.$$

The vector

$$Y=(y_1, y_2, \ldots, y_n) - N(0, \Sigma),$$

where the correlation matrix

$$\Sigma_j = \sigma^2 \left(\frac{1}{n} \sum_{k=0}^{n-j} \left(1^{2n} + \frac{1}{2} \left|k-j\right|^{2n} - \left|k-j\right|^{2n} \right) \right) = \frac{\sigma^2}{n} S_j.\]$$

$\{y_n\}$ 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 [2].

We check the method for simultaneous estimation of two unknown parameters $\sigma$ and $H$ 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.

$$R_j = \frac{1}{n} \sum_{k=1}^{n-j} y_j - \sigma \sqrt{\frac{n}{2}} J \frac{j+1}{2} \left(\frac{1}{2}\right)$$

Then calculate the mean

$$E_j = E R_j = \sigma^2 \frac{\left(\frac{1}{2}\right)}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right)$$

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

Theorem
With probability 1

$$\frac{R_j}{E_j} \to 1$$

In particular, when $j=1$, we have

$$\frac{R_1}{E_1} \to 1,$$

with the known value estimate for $\sigma$

$$\hat{\sigma} = \sqrt{n} \frac{\sqrt{\pi}}{2} R_1,$$

and $H$ estimate

$$\hat{H} = \frac{\ln \left(\frac{\pi}{2} \frac{\sigma^2}{R_1} \right)}{\ln n}.$$
\[ E \xi = 0, E(\epsilon, u)(\xi, v) = (u, v) \text{ dim } \epsilon = n. \]

Then \( y = v^T \xi \).

Therefore
\[ n = E(\xi, \tilde{\xi}) = E(U^{--1} y, y) = \frac{n^{2H}}{\sigma^2} E(S^{--1} y, y) \]

And consequently the statistic
\[ \hat{\sigma}^2_n = (n)^{2H}(S^{--1} y, y) \]

and here statistics \((n)^{2H}(S^{--1} y, y)\) is an unbiased estimate of the parameter \( \sigma^2 \).

Now we prove consistency of this estimate. Let us introduce the notation:
\[ \hat{\sigma}^2_n = \sqrt{n^{2H}(S^{--1} y, y)} \] (1)

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

\[ D \hat{\sigma}^2_n = 2\sigma^2 \]

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: for an arbitrary value \( H \in (0; 1) \) let us calculate the estimate \( \sigma_{kn} \), matrix \( S^{--1} \) and the estimate \( \hat{\sigma}_{kn} \). Then we iterate the values \( H \) (matrix \( S \) for different \( H \)) with some step
\[ \hat{\sigma}_{kn} = \frac{0.8}{R_{kn}} \sqrt{(S^{--1} y, y)} \approx 1 \cdot \left| \frac{\hat{\sigma}_{kn}^2}{\hat{\sigma}_{kn}} - 1 \right| \rightarrow \min. \] (2)

The values \( \hat{H} \) of parameter \( H \), which satisfies eqn. (2), is an estimate and
\[ \hat{\sigma} = \frac{\hat{\sigma}_{kn} + \hat{\sigma}_{2n}}{2} \]

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

We proved that the estimate of parameter \( \sigma \)
\[ \hat{\sigma}^2_n = \sqrt{n^{2H}(S^{--1} y, y)} \] (3) is consistent.

Let us calculate the estimate \( \hat{\sigma}_{kn} \), matrix \( S^{--1} \) and estimate \( \hat{\sigma}_{kn} \).
\[ \hat{\sigma}_{kn} = \frac{0.8}{R_{kn}} \sqrt{(S^{--1} y, y)} \approx 1 \cdot \left| \frac{\hat{\sigma}_{kn}^2}{\hat{\sigma}_{kn}} - 1 \right| \rightarrow \min. \] (4)

\[ \hat{\sigma} = \frac{\hat{\sigma}_{kn} + \hat{\sigma}_{2n}}{2} \]

\[ F_n = \frac{1}{n - 1} \sum_{k=1}^{n-1} \text{sgn}(y_k y_{k+1}) ; \]

\[ Q_n = \frac{1}{n - 1} \sum_{k=1}^{n-1} |y_k| \]

The generated increments are defined by
\[ z_n = n^H y^k, \quad n = 256; n = 1024 \]

\[ \eta_n = \frac{1}{n} \sum_{k=1}^{n} |z_k| = n^H R_n \]

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

| \( H \) | \( H^1 \) | 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| 1.20| 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 |

Table 1: The values of efficiency of the estimation algorithm.
Then
\[
d_{n} = \frac{1}{n-1} \sum_{k=1}^{n-1} |x_k|^{2}.
\]

Let's assume
\[
d_{n} = \frac{1}{n-1} \sum_{k=1}^{n-1} |x_k|^{2}.
\]

and approximating \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]

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

To check the properties of increments \(z_k\) allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let's approximate \(f\) by the exponential function:
\[
\frac{s + 1}{n^{2H}} + \frac{s - 1}{n^{2H}} - s^{2H},
\]

and approximate \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]

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

To check the properties of increments \(Z_k\) allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let's approximate \(f\) by the exponential function:
\[
\frac{s + 1}{n^{2H}} + \frac{s - 1}{n^{2H}} - s^{2H},
\]

and approximate \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]

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

To check the properties of increments \(Z_k\) allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let's approximate \(f\) by the exponential function:
\[
\frac{s + 1}{n^{2H}} + \frac{s - 1}{n^{2H}} - s^{2H},
\]

and approximate \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]

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

To check the properties of increments \(Z_k\) allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let's approximate \(f\) by the exponential function:
\[
\frac{s + 1}{n^{2H}} + \frac{s - 1}{n^{2H}} - s^{2H},
\]

and approximate \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]

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

To check the properties of increments \(Z_k\) allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let's approximate \(f\) by the exponential function:
\[
\frac{s + 1}{n^{2H}} + \frac{s - 1}{n^{2H}} - s^{2H},
\]

and approximate \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]

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

To check the properties of increments \(Z_k\) allow limit theorems for fractional Brownian motion proved in I. Nourdin.

Let's approximate \(f\) by the exponential function:
\[
\frac{s + 1}{n^{2H}} + \frac{s - 1}{n^{2H}} - s^{2H},
\]

and approximate \(x_k\) by the fractional Brownian motion:
\[
z_k = \sigma^H \left( B_{rac{k}{n}} - B_{rac{k-1}{n}} \right),
\]
\[
\frac{1}{n} \sum_{k=1}^{n} a_k^2 \xi_k \rightarrow \frac{3}{2}, \quad H \in (0; \frac{1}{2})
\]

\[
\frac{1}{n^{1+2H}} \sum_{k=1}^{n} a_k^2 \xi_k \rightarrow 3\eta, \quad H \in (0; \frac{1}{2})
\]

\[\eta \sim N(0; \frac{1}{2H+2})\]

In the second formula \(f(x) = x^2\),

\[
\frac{1}{n} \sum_{k=1}^{n} a_k^2 \xi_k \rightarrow \frac{1}{2}, \quad H \in (0; \frac{1}{4})
\]

The third formula for \(f(x) = x\) and \(f(x) = 1\):

\[
\frac{1}{n^{3+2H}} \sum_{k=1}^{n} a_k^2 \xi_k \rightarrow \frac{3}{2}B^2(1) \quad H \in (1/2, 1/4))
\]

The third formula for \(f(x) = x\) and \(f(x) = 1\):

\[
\frac{1}{n} \sum_{k=1}^{n} \xi_k \rightarrow 3B(1) \quad H \in (1/2, 1)
\]

With the (3.18)-(3.19) we can check a hypothesis \(T\). The statistics \(z_1, z_2\) are proportional to the increments of fractional Brownian motion. Let’s consider a checking algorithm (with a known \(H\)) [5]:

\[c = \frac{1}{n} \sum_{k=1}^{n} z_k^2\]

\[z_k = c_1 \xi_k = c n^{H} (B(k/n) - B(k))\]

\[v_k = \frac{1}{n} \sum_{j=1}^{k-1} z_j\]

Let’s construct the statistics

\[H \in (0; \frac{1}{2}) \quad ; \quad H \in (0; \frac{1}{2})
\]

\[B_n = \frac{1}{n^{1+2H}} \sum_{k=1}^{n} v_k^2 \xi_k \quad H \in (0; \frac{1}{2})\]

\[C_n = \frac{1}{n} \sum_{k=1}^{n} v_k \xi_k \quad H \in (0; \frac{1}{2})\]

\[D_n = \frac{1}{n} \sum_{k=1}^{n} v_k \xi_k \quad H \in (1/4, 1/2)
\]

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 \(x_1, x_{555}\) with increments \(y_k = x_{i+k} - x_i, k=1,2,365\).

\[R_{\text{a}} = \frac{1}{365} \sum_{k=1}^{365} |y_k| = 0.141; \quad R_{\text{a}} = \frac{1}{365} \sum_{k=1}^{365} y_k^2 = 0.04\]

\[d = 0.5, \quad \hat{\lambda} = 1.4, \quad \gamma = 0.7, \quad k = 1.365
\]

The second example: 1678 values of interest rate. The interest rate is given by the following formula in each time window.

\[S(t) = a + b \exp[X(t)], \quad X(t) = h \frac{S(t) - a}{S(t) - a}, \quad a = 0.085\]

\[x_k = X \frac{k-1}{n} = h \frac{S_k - a}{S_k - a}, \quad x_1 = 0\]

\[y_k = x_{k+1} - x_k\]

\[R_{\text{a}} = 0.425; \quad R_{\text{a}} = 0.541; \quad d = \frac{R_{\text{a}}}{R_{\text{a}}} = 0.334\]

\[\hat{\lambda} = \lambda = 2\]

\[y_k = \text{sgn} y_k |y_k|, \quad z_k = \text{sgn} y_k \sqrt{|y_k|}\]

Applying the estimation algorithm (4), we obtain

\[H = 0.3\]

The third example - monthly data interest rate of Germany Bundesbank (www.bundesbank.de) for 2003-2012, 120 data.

\[Y_k = x_{k+1} - x_k\]

\[R_{\text{a}} = \frac{1}{119} \sum_{k=1}^{119} |y_k| = 1.39; \quad R_{\text{a}} = \frac{1}{119} \sum_{k=1}^{119} y_k^2 = 3.3; \quad d = \frac{R_{\text{a}}}{R_{\text{a}}} = 0.59\]

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

\[\frac{\sigma_{\hat{\lambda}}}{\sigma_\text{H}} = 1.0, \quad H = 0.4; \quad H = 0.4\]

The fourth \(y_k = x_{k+1} - x_k\), \(u_k = 100y_k\) example- \(S_k, S_k\) exchange rate, 1218 data, 2005-2009 (http://www.banque-france.fr).

Let’s calculate the statistics \(R_{\text{a}}, R_{\text{a}}\), \(n=1217\).

\[R_{\text{a}} = \frac{1}{n} \sum_{k=1}^{n} |y_k| = 0.215; \quad R_{\text{a}} = \frac{1}{n} \sum_{k=1}^{n} y_k^2 = 0.52; \quad d = \frac{R_{\text{a}}}{R_{\text{a}}} = 0.089\]

\[\frac{1}{\sqrt{n}} \frac{V(\hat{\lambda} + 1)}{V(\lambda + 1/2)} = 0.089, \quad \lambda = 4\]

For the converted sequence defined by the equation

\[z_k = \text{sgn} y_k |y_k|\]

\[r_{\text{a}} = \frac{1}{n} \sum_{k=1}^{n} z_k = 0.76; \quad r_{\text{a}} = \frac{1}{n} \sum_{k=1}^{n} z_k = 0.84; \quad d = 0.68\]

Let’s calculate the value of the quantity

\[\frac{\sigma_{\hat{\lambda}}}{\sigma_\text{H}} = 0.8 \sqrt{\frac{V(z_k, z_k)}{n}}\]

In this example, the minimum \(\frac{\sigma_{\hat{\lambda}}}{\sigma_\text{H}} = 0.8\) achieved by \(H=0.3, 0.8\).

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

\[c = \frac{1}{n} \sum_{k=1}^{n} z_k = 0.144\]

\[D_n = \frac{1}{n} \sum_{k=1}^{n} v_k z_k = 0.002; \quad D_n = 0.1\]

\[F(x) = 2\Phi^2(\frac{3}{2}x) - 1, x > 0\]

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

$n = 336, \quad H = 0.3; \quad c = \frac{1}{n} \sum_{i=1}^{n} z_i^2 = 0.425$

$A_n = -\frac{1}{335} \sum_{i=1}^{335} y_i z_i^2 = -0.16; \quad B_n = -\frac{1}{335.5} \sum_{i=1}^{335.5} y_i z_i^2 = -0.08$

The theoretical limit value

$\lim_{n \to \infty} A_n = -1.5 c^2 = -0.27$

$B_n c^{-\frac{5}{2}} = -0.68$

$B_n c^{-\frac{5}{2}} \leq \beta = 3.07$

The hypothesis $T$ is accepted.

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

$n = 119; \quad H = 0.4; \quad c = \frac{1}{n} \sum_{i=1}^{n} z_i^2 = 3.3$

$A_n = -19.1; \quad B_n = -0.8; \quad A_n c^{-2} = -1.75; \quad B_n c^{-\frac{5}{2}} = -0.04$

The hypothesis $T$ is accepted.

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

$A_n = \frac{1}{n} \sum_{i=1}^{n} y_i z_i^2 = -0.69$

For $H > \frac{1}{2}, \quad A_n \geq 0$, then we should take a decision of hypothesis $H=0.3$ (The theoretical value $A_n$ is equal $-1.5 \cdot (0.84)^2 = -1.0$)

**Conclusion**

Thus, we proposed in this research a new method for estimate of parameters fBm, 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**

1. Mandelbrot BB, John W, Van Ness (1968) The Fractional Brownian motion, fractional noises and applications. SIAM Review 10: 422-437.

2. Mandelbrot BB (1969) Computer experiments with fractional Gaussian noises. Wat Reso Rese 5: 228-241.

3. Beran J (1994) Statistics for Long-Memory Processes. Chapn Hall pp: 315.

4. Barndorff N, Levy OE (2001) Theory and Applications, pp: 421

5. Mishura Y (2008) Stochastic Calculus for Fractional Brownian Motion and Related Processes. Lecture Notes in Mathematics, pp: 392.