ESTIMATION OF LONG MEMORY PARAMETER AND ITS CONFIDENCE INTERVALS

Ieva Dasmane, Jānis Valeinis

University of Latvia

28.05.2011

Publications

(1) Confidence intervals for the long memory parameter based on wavelets and resampling, P. L. Conti, L. De Giovanni, S. A. Stoev and M. S. Taqqu. Statistica Sinica 18 (2):559–579, 2008

(2) Estimators of long-memory: Fourier versus wavelets, G. Fay, E. Moulines, F. Roueff, M. S. Taqqu. Journal of Econometrics 151 (2):159–177, 2009

Long memory parameter

Hurst parameter H

- hydrologist Hurst (1951) - Nile River problem
- biophysics, computer networking, finance
- fractal and chaos theory, long memory processes and spectral analysis

H is a measure of the level of self-similarity of a process and exhibits long-range dependence.

Self - similar process

A continuous parameter stochastic process $X = \{X(t) : 0 \leq t < \infty\}$ is self - similar process with self - similarity parameter (Hurst parameter) $0 < H < 1$, if $\{X(at) : 0 \leq t < \infty\}$ and $\{a^H X(t) : 0 \leq t < \infty\}$ has the same finite - dimensional distribution for all $a > 0$.

Long memory process

A discrete stationary process $(X_j; j > 1)$ is long range dependent (LRD or long memory) process, if its correlation coefficients

$$\rho(k) = \sigma^{-2}\{E(X_j X_{j+k}) - (EX_j)^2\}$$

take form

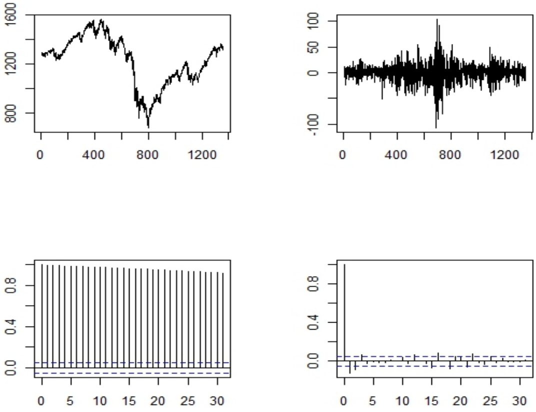
$$\rho(k) \sim c_r k^{-(1-\alpha)}, \tag{1}$$

for some $k \to \infty$ and $0 < \alpha < 1$, where k is the lag of autocorrelation, α is long memory parameter and c_r is absolute constant.

Relation between Hurst and the long memory parameter

$$H = \frac{1+\alpha}{2} \tag{2}$$

Process SP500



att.: Popular index value of United States equities market since 1957. Data taken from 03.01.2006 – 25.05.2011. Top left : SP 500 process. Top right: its increments. Bottom: autocorrelation functions.

Fractional Brownian motion

Gaussian process $B_H = \{B_H(t) : 0 \leq t \leq \infty\}$ with $0 < H < 1$ is fractional Brownian motion (fBm), if the following properties holds:

- $B_H(t)$ has stationary increments;
- $B_H(0) = 0$ and $E(B_H(t)) = 0$ for all $t \geq 0$;
- $E(B_H)^2(t) = t^{2H}$ for all $t \geq 0$;
- $B_H(t)$ has a Gaussian distribution for $t > 0$.

Covariance function

$$\gamma(s,t) = E(B_H(s)B_H(t)) = \frac{1}{2}\{t^{2H} + s^{2H} - (t-s)^{2H}\} \tag{3}$$

for all $0 < s < t$.

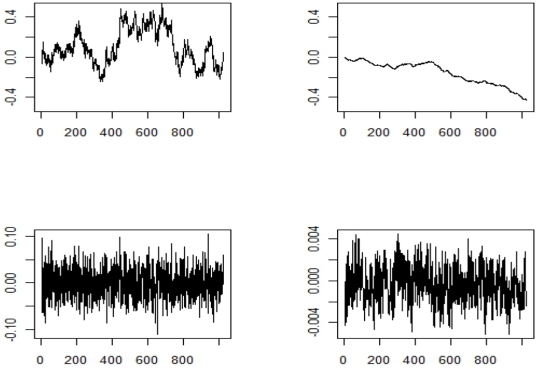
The incremental process $X = \{X_k : k = 0, 1, \dots\}$ of fractional Brownian motion is fracional Gaussian noise

$$X_k = B_H(k+1) - B_H(k).$$

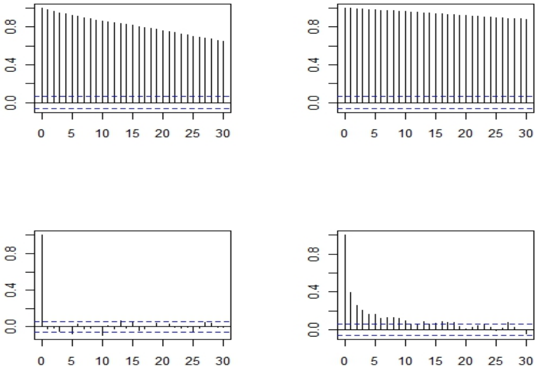
Covariance function

$$\gamma(k) = \frac{1}{2}[|k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H}] \tag{4}$$

for all $k \in \mathbb{Z}$.



att.: Top : left Standart Brownian motion $(H = 0.5)$, right - fractional Brownian motion $(H = 0.9)$. Sample size $N = 1024$. Bottom: increments for the processes on the top.



att.: Top : left - Standart Brownian motion $(\mathrm{H} = 0.5)$, right - fractional Brownian motion $(\mathrm{H} = 0.9)$. Sample size $\mathrm{N} = 1024$. Bottom: autocorrelation functions for increments of the processes on the top.

Hurst parameter estimation in R

- Aggregated variance method
- Periodogramm method
- R/S method
- Wavelet method

Aggregated variance method

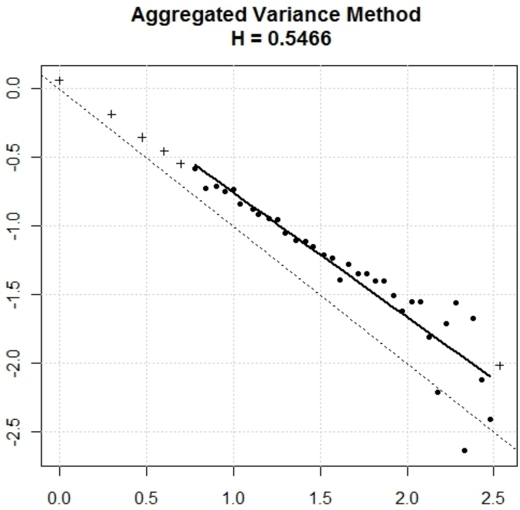
- Agreggated process

$$X_k^{(m)} = \frac{1}{m}(X_{km} + \cdots + X_{(k+1)m-1}), \qquad k = 0, 1, \ldots \quad (5)$$

- $X_k^{(m)}$ self similar with $m^{H-1}X$ for big m and with variance $\mathrm{Var}(X_k^{(m)}) = m^{2H-2}\mathrm{Var}(X_k)$.
- Variance estimator

$$\mathrm{Var}(\hat{X}_k^{(m)}) = \frac{1}{M}\sum_{i=0}^{M-1}(X_i^{(m)} - X^{\overline{(m)}})^2, \quad (6)$$

where $X^{\overline{(m)}}$ is sample average of $X^{(m)}$ and M - the integer part of N/m .



att.: Simulated fractional Gaussian noise process with H = 0.6, length of the process N = 1024. Slope of the line 2H − 2

Periodogramm method

- based on observation that spectral density of the process behaves like $c_f|\lambda|^{1-2H}$ as $\lambda \to 0$.
- Sample autocovariance

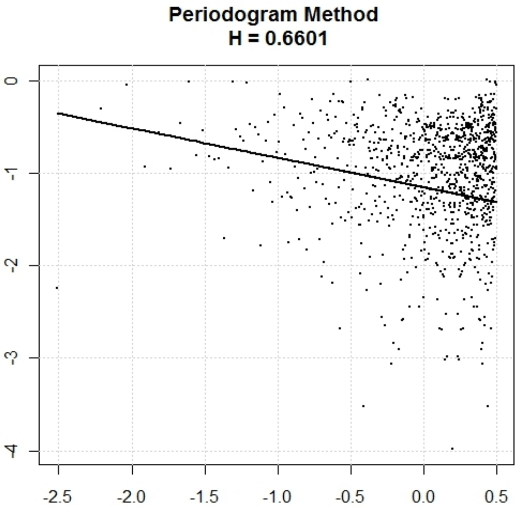
$$\hat{\gamma}(j) = \frac{1}{N} \sum_{k=0}^{N-|j|-1} (X_k - \bar{X})(X_{k+-|j|} - \bar{X}) \tag{7}$$

- Periodogramm

$$I(\lambda) = \sum_{j=-(N-1)}^{N-1} \hat{\gamma}(j) \exp(ij\lambda). \tag{8}$$

- The periodogram and the spectral density f are symmetric around zero. So

$$\lim_{N\to\infty} E[I(\lambda)] = f(\lambda). \tag{9}$$



att.: Simulated fractional Gaussian noise process with $H = 0.6$, length of the process $N = 1024$. Slope of the line $1 - 2H$

R/S method

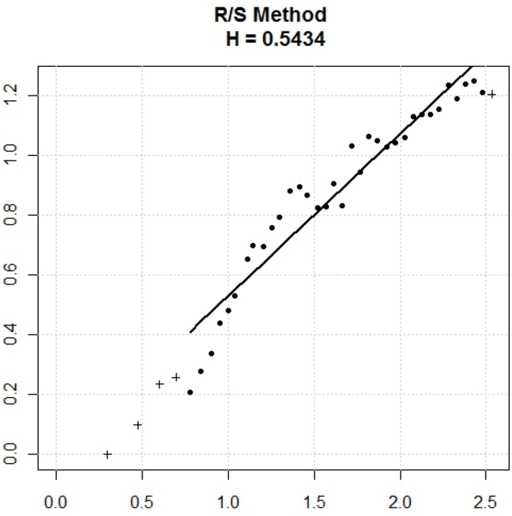
- X_i - water inflow in reservoir at time i
- $Y_j = \sum_{i=1}^{j} X_i$ - total water inflow until time j
- Perfect reservoir cappacity:

$$R(t,k) = \max_{0 \le i \le k} \left[Y_{t+i} - Y_t - \frac{i}{k}(Y_{t+k} - Y_t) \right] - \min_{0 \le i \le k} \left[Y_{t+i} - Y_t - \frac{i}{k}(Y_{t+k} - Y_t) \right]$$

- Standartized $R(t,k)$:

$$S(t,k) = \sqrt{k^{-1} \sum_{i=t+1}^{t+k} (X_i - \bar{X}_{t,k})^2},$$

- $$R/S = \frac{R(t,k)}{S(t,k)}. \qquad (10)$$



att.: Simulated fractional Gaussian noise process with H = 0.6, length of the process N = 1024. Slope of the line is H

Vanishing moments and Fourier transform

Wavelet $\psi(t), t \in \mathbb{R}$ is a function with M vanishing moments, that is,

$$\int t^{\ell}\psi(t)dt = 0, \qquad \text{for} \qquad \ell = 0, \dots, M-1$$

and its Fourier transform

$$\hat{\psi}(\lambda) = \int_{-\infty}^{\infty} \psi(t) \exp^{-i\lambda t} dt \tag{11}$$

decreases as $\lambda \to \infty$.

Scaled and translated wavelet

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k), \; j, k \in \mathbb{Z} \tag{12}$$

Wavelet setting

The wavelet setting involves a scale function $\phi \in \mathrm{L}^2(\mathbb{R})$ and a wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ with associated Fourier transforms.

Properties

- ϕ un ψ are compactly- supported, integrable and

$$\hat{\phi}(0) = \int_{-\infty}^{\infty} \phi(\mathrm{t})\mathrm{dt} = 1 \text{ un } \int_{-\infty}^{\infty} \psi^2(\mathrm{t})\mathrm{dt} = 1.$$

- Function $\sum_{\mathrm{k}\in\mathbb{Z}} \mathrm{k}^{\mathrm{m}}\phi(* - \mathrm{k})$is a polynomial of degree m for all $\mathrm{m} = 0, \ldots, \mathrm{M} - 1$.

Wavelet coefficients

$$\mathrm{W}_{\mathrm{j,k}} = \int_{\mathbb{R}} \mathrm{X}(\mathrm{t})\psi_{\mathrm{j,k}}(\mathrm{t})\mathrm{dt} \tag{13}$$

Wavelet method

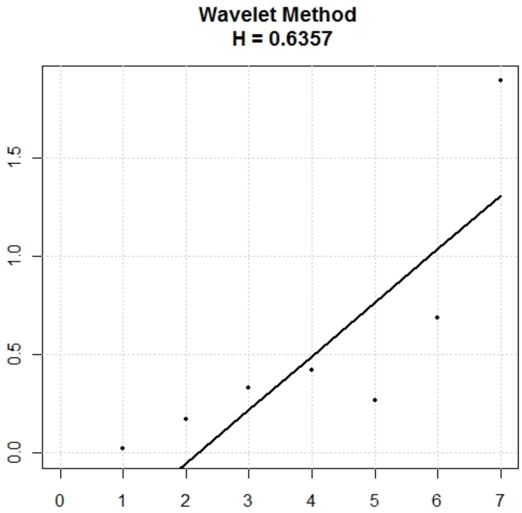
- Related to periodogramm method - we get again estimation from transformed data
- compute wavelet coefficients $W_{j,k}$ at each level j:

$$W_{j,k} = \langle X, \psi_{j,k} \rangle,$$

  where $\langle , \rangle$ is inner product
- variance of coefficients

$$v_j = 2^j \sum_{k=0}^{2^{-j}-1} (W_{j,k})^2. \quad (14)$$

- Plot $\log(v_j)$ against scale index j and fit line by least squares method.



att.: Simulated fractional Gaussian noise process with H = 0.6, length of the process N = 1024. Slope of the line 2H − 1

Confidence intervals using subsampling

- $\hat{H}_n$ - estimator of parameter H from sample with size n
- $\sqrt{\nu_{j_1(n)}(n)}$ - number of wavelet coeficients at scale $j_1(n)$
- define

$$T_n = \sqrt{\nu_{j_1(n)}(n)}(\hat{H}_n - H) \tag{15}$$

- divide sample into i overlapping blocks $B_i = (X_i, \dots, X_{i+l-1})$ with length l, where $i = 1, \dots, N$ and $N = n - l - 1$
- define sub sample copy of T_n based on blocks B_i:

$$\hat{T}_{l,i} = \sqrt{\nu_{j_1(l)}(l)}(\hat{H}_{l,i} - \hat{H}_n) \tag{16}$$

- empirical distribution function based on sub sampling is

$$\hat{Q}_n(x) = \frac{1}{N}\sum_{i=1}^{N} I_{(\hat{T}_{l,i} \leq x)}, \qquad x \in \mathbb{R}. \tag{17}$$

Confidence intervals using subsampling

Confidence interval for Hurst parameter using subsampling

$$\left(\hat{\mathrm{H}}_{\mathrm{n}} + \frac{1}{\sqrt{\nu_{\mathrm{j}_1}}(\mathrm{n})}\hat{\mathrm{Q}}_{\mathrm{n}}^{-1}(\frac{\gamma}{2}), \hat{\mathrm{H}}_{\mathrm{n}} + \frac{1}{\sqrt{\nu_{\mathrm{j}_1}}(\mathrm{n})}\hat{\mathrm{Q}}_{\mathrm{n}}^{-1}(1-\frac{\gamma}{2})\right),$$

where $0 < \gamma < 1$ and $\hat{\mathrm{Q}}_{\mathrm{n}}^{-1}$ is a quantile.

H estimation for FGN

Simulated process: fractional Gaussian noise with length $n = 2^7$ and block size $n = 2^6$. 1000 simulations. Confidence intervals based on subsampling with confidence level $p = 0.95$.

Method	$H = 0.6$			$H = 0.8$		
	Mean	Stds	Cov. prob.	Mean	Stds	Cov. prob.
Wavelets	0.546	0.260	0.996	0.755	0.263	0.994
R/S	0.663	0.164	0.994	0.770	0.165	0.990
Period.	0.643	0.360	0.919	0.915	0.363	0.907
Agg.Var.	0.452	0.236	0.979	0.588	0.229	0.973

- Set of all available wavelet coefficients $W_{j,k}$ from n observations have scale indices between j_0 and j_1:

$$\ell_n(j_0, j_1) = \{(j,k) : j_0 \le j \le j_1, 0 \le k < n_j\}, \qquad (18)$$

where $n_j = \lfloor 2^{-j}(n - T + 1) - T + 1 \rfloor$ for $T \ge 1$.

- Integers L and U are lower and upper scale indices such that $0 \le L < U \le J_n = \max\{j : n_j \ge 1\}$
- $$Var[W^X_{j,0}] = \sigma_j^2(d, f^*) \asymp \sigma^2 2^{2dj}, \text{ kad } j \to \infty,$$
- $$\hat{\sigma}_j^2 = \frac{1}{n_j} \sum_{k=0}^{n_j - 1} (W^X_{j,k})^2,$$

Local Reggresion Wavelets

- LRW is defined as the least squares estimator in the linear regression model

$$\log[\hat{\sigma}_j^2] = \log \sigma^2 + dj\{2\log(2)\} + u_j, \tag{19}$$

where $u_j = \log[\hat{\sigma}_j^2/\sigma^2 2^{2dj}]$.

- d estimator is

$$\hat{d}_n^{LRW}(L, U, w) = \sum_{j=L}^{U} w_{j-L} \log(\hat{\sigma}_j^2), \tag{20}$$

where $w = [w_0, \dots, w_{U-L}]^T$ is a vector of weights that satisfies

$$\sum_{j=0}^{U-L} w_j = 0 \quad \text{and} \quad 2\log(2)\sum_{j=0}^{U-L} jw_j = 1.$$

Local Whittle Wavelets

- wavelet coefficients $W_{j,k}$ and scale indices $\ell_n(j_0, j_1)$
- scaling property holds $\sigma_j^2(d, f^*) \asymp \sigma^2 2^{2dj}$
- Pseudo negative log-likelihood

$$\hat{L}_\ell(\sigma^2, d) = \frac{1}{2\sigma^2} \sum_{(j,k)\in\ell} 2^{-2dj}(W_{j,k}^X)^2 + \frac{|\ell|}{2} \log(\sigma^2 2^{2\langle\ell\rangle d}),$$

where $|\ell|$ denotes the cardinal of ℓ and $\langle\ell\rangle$ is the average scale.

- Define

$$\hat{\sigma}_\ell^2(d) = \arg\min_{\sigma^2>0} \hat{L}_\ell(\sigma^2, d) = |\ell|^{-1} \sum_{(j,k)\in\ell} 2^{-2dj}(W_{j,k}^X)^2.$$

- Pseudo maximum likelihood estimator of d is minimum of the negative log -likelhood

$$\hat{d}^{LWW}(L, U) = \arg \min_{d\in[\Delta_1,\Delta_2]} \hat{L}_{\ell_n(L,U)}(\hat{\sigma}_\ell^2(d), d), \qquad (21)$$

where $[\Delta_1, \Delta_2]$ is the interval of admissible values of d.

d estimation for FBM

- For FBM parameter $H = d - 0.5$

Wavelet	d = 1.1			
	LRW		LWW	
	Mean	Cov.Prob	Mean	Cov.Prob
Haar	0.9751	0.802	1.037	0.955
Daubechies 4	0.961	0.792	1.0161	0.8360
Daubechies 8	0.8513	0.665	0.7079	0.672

Wavelet	d = 1.3			
	LRW		LWW	
	Mean	Cov.Prob	Mean	Cov.Prob
Haar	1.1663	0.801	1.2149	0.957
Daubechies 2	1.1438	0.777	1.2078	0.862
Daubechies 4	1.021	0.650	0.9003	0.662

d estimation for FGN

- For FGN parameter $H = d + 0.5$

Wavelet	d = 0.1			
	LRW		LWW	
	Mean	Cov.Prob	Mean	Cov.Prob
Haar	0.0157	0.840	0.0837	0.925
Daubechies 4	0.0228	0.828	0.0633	0.849
Daubechies 8	-0.0505	0.702	-0.2504	0.6780

Wavelet	d = 0.3			
	LRW		LWW	
	Mean	Cov.Prob	Mean	Cov.Prob
Haar	0.2156	0.840	0.2806	0.938
Daubechies 2	0.2304	0.818	0.2659	0.845
Daubechies 4	0.1674	0.703	0.167	0.703

Thank you for attention!