Fast computation of wavelet coefficients covariance matrix of a generalized long memory process. Application

Olaf KOUAMO

TELECOM ParisTech, France.

Joint work with E. Moulines and F. Roueff

EMS 2010 PIREAUS

August 18, 2010

Outline

1. **New algorithm** We propose a new iterative algorithm for the computation of the covariance matrix of the wavelet coefficients for a process which is not necessarily stationary:
   - either stationary
   - or, it is K-th order stationary.

   Problem: Since the process is not stationary, it is not true that any wavelet functions will provide stationary wavelet coefficients $\Longrightarrow$ their covariance matrix is not defined.

   Originality of our approach: since a K-th order difference of the process X is a stationary process
   - we can find an appropriate wavelet function so that the wavelet coefficients of X are stationary;
   - then, we compute the covariance matrix of these coefficients.

Outline

1. **New algorithm** We propose a new iterative algorithm for the computation of the covariance matrix of the wavelet coefficients for a process which is not necessarily stationary:
   - either stationary
   - or, it is K-th order stationary.

   Problem: Since the process is not stationary, it is not true that any wavelet functions will provide stationary wavelet coefficients $\Longrightarrow$ their covariance matrix is not defined.

   Originality of our approach: since a K-th order difference of the process X is a stationary process
   - we can find an appropriate wavelet function so that the wavelet coefficients of X are stationary;
   - then, we compute the covariance matrix of these coefficients.

Outline

1. **New algorithm**
2. **Application** to the calibration of generalized ARFIMA processes which includes
   - long-memory processes
   - both stationary and non-stationary processes

   by a Maximum-Likelihood (ML) approach in the wavelet domain.

   Problem: when the process is not stationary, the ML approach is untractable since the covariance matrix of the process is not defined.

   Our contribution: for K-th order stationary process, we are able to provide a ML estimation of the ARFIMA coefficients.

Computation of the covariance matrix in the stationary case

When the process X is stationary: there exist iterative algorithms (for example see Moulines [2007]) such that given

- the covariance matrix of the process X
- a wavelet function (with compact support) and the associated quadrature mirror filters h, g see Mallat [1998]

they compute recursively the covariance matrix of the wavelet coefficients among scales of X.

Problem when X is not stationary: its covariance does not exist

$\Longrightarrow$ these usual algorithms do not apply.

Answer: we provide an answer in the case X is K-th order difference stationary.

Computation of the covariance matrix in the stationary case

When the process X is stationary: there exist iterative algorithms (for example see Moulines [2007]) such that given

- the covariance matrix of the process X
- a wavelet function (with compact support) and the associated quadrature mirror filters h, g see Mallat [1998]

they compute recursively the covariance matrix of the wavelet coefficients among scales of X.

Problem when X is not stationary: its covariance does not exist

$\Longrightarrow$ these usual algorithms do not apply.

Answer: we provide an answer in the case X is K-th order difference stationary.

Computation of the covariance matrix in the stationary case

When the process X is stationary: there exist iterative algorithms (for example see Moulines [2007]) such that given

- the covariance matrix of the process X
- a wavelet function (with compact support) and the associated quadrature mirror filters h, g see Mallat [1998]

they compute recursively the covariance matrix of the wavelet coefficients among scales of X.

Problem when X is not stationary: its covariance does not exist $\Longrightarrow$ these usual algorithms do not apply.

Answer: we provide an answer in the case X is K-th order difference stationary.

Computation of the covariance matrix for a K-th order difference stationary process X

Define

- the first order difference process $\mathbf{\Delta}X$: $[\mathbf{\Delta}X]_n \stackrel{\text{def}}{=} X_n - X_{n-1}$
- the K-th order difference process $\mathbf{\Delta}^K$: $\mathbf{\Delta}^K \stackrel{\text{def}}{=} \mathbf{\Delta} \circ \mathbf{\Delta}^{K-1}$

Definition 1

X is a K-th order difference stationary process if $\mathbf{\Delta}^K X$ is stationary.

For a K-th order difference stationary process X

- there exist non-iterative (and thus "greedy") algorithms for the computation of the covariance matrix see e.g. (Percival and al 2000)
- we provide an iterative algorithm.

Algorithm (KMR,2010)

INPUT
- the covariance of the stationary process $\mathbf{\Delta}^K X$.
- a wavelet function with $M \geq K$ vanishing moments
- adequate filters h_M, g_M

OUTPUT covariance matrix of the wavelet coefficients of the process X

In the case $K = 0$, i.e. when X is stationary this algorithm is the iterative algorithm proposed by Moulines [2007].

In the case $K > 0$, how to compute the filters h_M and g_M ?

Algorithm (KMR,2010)

INPUT
- the covariance of the stationary process $\mathbf{\Delta}^K X$.
- a wavelet function with $M \geq K$ vanishing moments
- adequate filters h_M, g_M

OUTPUT covariance matrix of the wavelet coefficients of the process X

In the case $K = 0$, i.e. when X is stationary this algorithm is the iterative algorithm proposed by Moulines [2007].

In the case $K > 0$, how to compute the filters h_M and g_M ?

How to compute the filters h_M and g_M ?

Intuition: When X is stationary, given a wavelet function with M vanishing moments, there exist quadrature mirror filters g_0 and h_0 such that

Pyramidal algorithm Mallat,[1998]

$$A_{0,k} = X_k, \quad k \in \mathbb{Z} ,$$

$$A_{j,k} = \left[\downarrow^2 (h_0 * A_{j-1,\cdot})\right]_k , \quad k \in \mathbb{Z}, j \geq 1 ,$$

$$W_{j,k} = \left[\downarrow^2 (g_0 * A_{j-1,\cdot})\right]_k , \quad k \in \mathbb{Z}, j \geq 1 ,$$

and $\{W_{j,k}, j \geq 0, k \in \mathbb{Z}\}$ are the wavelet coefficients of X.

How to compute the filters h_M and g_M ?

Intuition: When X M-th order difference stationary , given a wavelet function with at least M vanishing moments

adapted Pyramidal algorithm (KMR, 2010)

$$A^{(M)}_{0,k} = [\boldsymbol{\Delta}^M X]_k, \quad k \in \mathbb{Z} ,$$

$$A^{(M)}_{j,k} = \left[\downarrow^2 \left(h_M * A^{(M)}_{j-1,\cdot}\right)\right]_k, \quad k \in \mathbb{Z},\, j \geq 1 ,$$

$$W_{j,k} = \left[\downarrow^2 \left(g_M * A^{(M)}_{j-1,\cdot}\right)\right]_k, \quad k \in \mathbb{Z},\, j \geq 1 ,$$

where $\quad g_0 = \boldsymbol{\Delta}^M g_M \qquad h_M = \sum_{s=0}^{M} \binom{M}{s} h_0[k-s]$

and $\{W_{j,k}, j \geq 0, k \in \mathbb{Z}\}$ are the wavelet coefficients of X.

Indeed,

$$\begin{aligned} W_{j,k} &= \left[\downarrow^2 (g_0 * A_{j-1,\cdot})\right]_k \\ &= \left[\downarrow^2 \left(\mathbf{\Delta}^M g_M * A_{j-1,\cdot}\right)\right]_k = \left[\downarrow^2 \left(g_M * \mathbf{\Delta}^M A_{j-1,\cdot}\right)\right]_k \\ &= \left[\downarrow^2 \left(g_M * A^M_{j-1,\cdot}\right)\right]_k \end{aligned}$$

Conclusion

INPUT

- the covariance of the stationary process $\boldsymbol{\Delta}^K X$.
- a wavelet function with $M \geq K$ vanishing moments
- filters h_M, g_M computed by the formula

$$g_0 = \boldsymbol{\Delta}^M g_M \qquad h_M = \sum_{s=0}^{M} \binom{M}{s} h_0[k-s]$$

where h_0, g_0 are the quadrature mirror filters associated to the wavelet function and computed through the Pyramidal algorithm of Mallat [1998]

- K the number of differentiation of X

OUTPUT covariance matrix of the wavelet coefficients of the M-th order difference stationary process X.

Maximum Likelihood method in the wavelet domain

For the estimation of the parameters of a M-th order difference stationary or stationary ARFIMA process X,

1. in the parametric case
2. in the semi-parametric case

we apply algorithms that rely on the covariance of the wavelet coefficients of X. We thus compute this matrix by applying our algorithm to estimate the parameters in the ARFIMA (p, d, q) model.

In the semi-parametric case, we also run for comparison the *Local Whittle Wavelet* (LWW) estimator of the memory parameter d.

Parametric case

Two experiences :

1. $n = 1024$ samples of a ARFIMA$(0, d, 0)$
2. $n = 1024$ samples of a ARFIMA$(1, d, 1)$, $\phi = 0.8$ and $\theta = 0.5$

- for different values of d: include both stationary and non stationary cases
- we estimate the parameters (d, ϕ, θ) and the innovation variance σ^2 of the process.

We compute 1000 independent estimated of the parameters and report in the tables

- the mean value of the estimators over the 1000 replications.
- the estimated variance of these estimators (by a Monte Carlo method over the 1000 replications).

Parametric case

Two experiences :

1. $n = 1024$ samples of a ARFIMA$(0, d, 0)$
2. $n = 1024$ samples of a ARFIMA$(1, d, 1)$, $\phi = 0.8$ and $\theta = 0.5$

- for different values of d: include both stationary and non stationary cases
- we estimate the parameters (d, ϕ, θ) and the innovation variance σ^2 of the process.

We compute 1000 independent estimated of the parameters and report in the tables

- the mean value of the estimators over the 1000 replications.
- the estimated variance of these estimators (by a Monte Carlo method over the 1000 replications).

Likelihood experiments

$n = 1024$ ARFIMA$(0, d, 0)$ $\sigma^2 = 1$

d	-0.8	-0.4	0	0.2	0.6	1	1.6	2	2.6	
Sample mean($\widehat{d}$)	-0.79	-0.39	0.002	0.19	0.603	1.007	1.59	1.99	2.588	-
Sample SD($\widehat{d}$)	0.0265.	0.029	0.028	0.026	0.029	0.027	0.027	0.027	0.028	0.024
Sample mean($\widehat{\sigma}^2$)	0.99	0.99	1.002	0.98	1.003	0.99	1.002	1.003	0.99	-
Sample SD($\widehat{\sigma}^2$)	0.047	0.046	0.045	0.045	0.046	0.047	0.045	0.047	0.048	0.044

$n = 1024$ ARFIMA$(0.8, d, 0.5)$ $\sigma^2 = 1$

d	-0.8	-0.4	0	0.2	0.6	1	1.6	2	2.6
Sample mean($\widehat{d}$)	-0.780	-0.38	0.0061	0.18	0.61	1.021	1.609	1.98	2.64
Sample SD($\widehat{d}$)	0.037	0.034	0.031	0.035	0.032	0.038	0.033	0.032	0.039
Sample mean($\widehat{\phi}$)	0.79	0.78	0.81	0.802	0.78	0.76	0.79	0.78	0.75
Sample SD($\widehat{\phi}$)	0.042	0.042	0.041	0.04	0.041	0.042	0.042	0.041	0.043
Sample mean($\widehat{\theta}$)	0.48	0.47	0.503	0.51	0.506	0.47	0.49	0.505	0.41
Sample SD($\widehat{\theta}$)	0.069	0.070	0.071	0.069	0.067	0.069	0.068	0.067	0.071
Sample mean($\widehat{\sigma^2}$)	0.997	1.008	0.99	0.99	1.007	0.99	0.98	0.99	0.98
Sample SD($\widehat{\sigma}^2$)	0.045	0.046	0.045	0.045	0.045	0.046	0.045	0.045	0.048

ARFIMA(p, d, q) model using wavelet domain.

Semi-Parametric case

In this case,

- the parameter of interest is the memory parameter d
- we are only interested at coarse scales

The ML method in the wavelet domain takes into account the wavelet dependence within and between scales

whereas

the Local Whittle Wavelet method Moulines and al [2008] does not.

The following numerical applications show the high performance of the ML method when compared to the LWW method.

Semi-Parametric case

In this case,

- the parameter of interest is the memory parameter d
- we are only interested at coarse scales

The ML method in the wavelet domain takes into account the wavelet dependence within and between scales

whereas

the Local Whittle Wavelet method Moulines and al [2008] does not.

The following numerical applications show the high performance of the ML method when compared to the LWW method.

Experience:

- $n = 4096$ of an ARFIMA$(0, d, 0)$
- $J_2 = 9$ and $J_1 = 3, 4, 5$ for $d \in \{-0.8 \ldots, 1.6\}$
- $J_2 = 8$ and $J_1 = 3, 4, 5$ for $d \in \{2, 2.6\}$
- we estimate the memory parameter d from scale J_1 to scale J_2

We compute 1000 independent estimated of d and report in the table

- the mean value of the estimator over the 1000 replications
- the mean square error (MSE) of each estimator over the 1000 replications

ML vs LWW

		$n = 4096,\ J_2 = 8, 9$		ARFIMA$(0, d, 0)$		$\sigma^2 = 1$				
d		-0.8	-0.4	0	0.2	0.6	1	1.6	2	2.6
$J_1 = 3$	$\hat{d}^{\mathrm{ML}}$	-0.802	-0.401	0.002	0.200	0.596	1.002	1.599	1.999	2.59
	$\mathrm{MSE}^{\mathrm{ML}}$	0.0014	0.001	$2e{-3}$	$4e{-5}$	0.006	0.001	$1e{-3}$	$7e{-4}$	0.001
	$\hat{d}^{\mathrm{LWW}}$	-0.71	-0.37	0.008	0.204	0.59	0.99	1.59	1.99	2.58
	$\mathrm{MSE}^{\mathrm{LWW}}$	4.14	0.32	0.03	0.007	0.01	0.005	0.003	0.04	0.05
$J_1 = 4$	$\hat{d}^{\mathrm{ML}}$	-0.803	-0.402	$2e{-3}$	0.199	0.602	0.998	1.596	1.999	2.596
	$\mathrm{MSE}^{\mathrm{ML}}$	0.003	0.002	$2e{-5}$	$5e{-5}$	$2e{-3}$	$8e{-4}$	0.006	$3e{-4}$	0.056
	$\hat{d}^{\mathrm{LWW}}$	-0.72	-0.38	0.01	0.205	0.606	0.99	1.59	1.99	2.59
	$\mathrm{MSE}^{\mathrm{LWW}}$	3.0	0.24	0.05	0.01	0.02	0.01	0.02	0.02	0.036
$J_1 = 5$	$\hat{d}^{\mathrm{ML}}$	-0.801	-0.401	$9e{-3}$	0.192	0.599	0.998	1.596	1.999	2.598
	$\mathrm{MSE}^{\mathrm{ML}}$	0.003	$5e{-4}$	$9e{-4}$	0.002	$3e{-4}$	$9e{-4}$	$5e{-3}$	$5e{-4}$	$1e{-3}$
	$\hat{d}^{\mathrm{LWW}}$	-0.72	-0.37	0.01	0.203	0.603	0.99	1.59	1.98	2.58
	$\mathrm{MSE}^{\mathrm{LWW}}$	3.42	0.32	0.001	0.001	$5e{-3}$	0.01	0.02	0.11	0.053

Table: Comparison of the MSE of ML vs LWW on 1000 independent replication of ARFIMA$(0, d, 0)$.

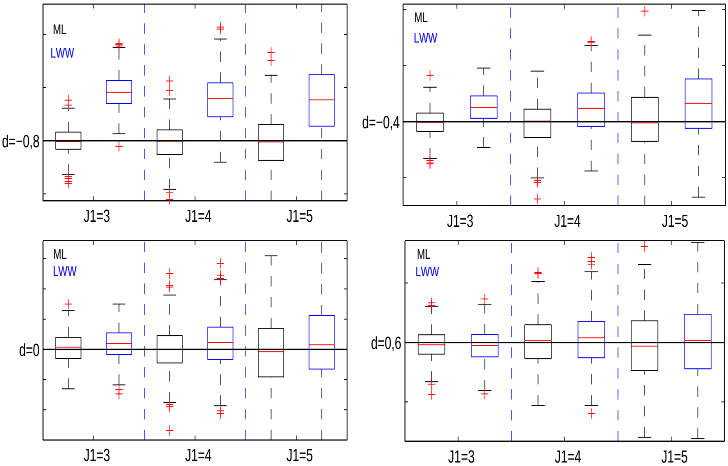

Figure: 1- Comparison of ML and LWW estimators of the memory parameter d of an ARFIMA$(0, d, 0)$ in a semi-parametric frame.

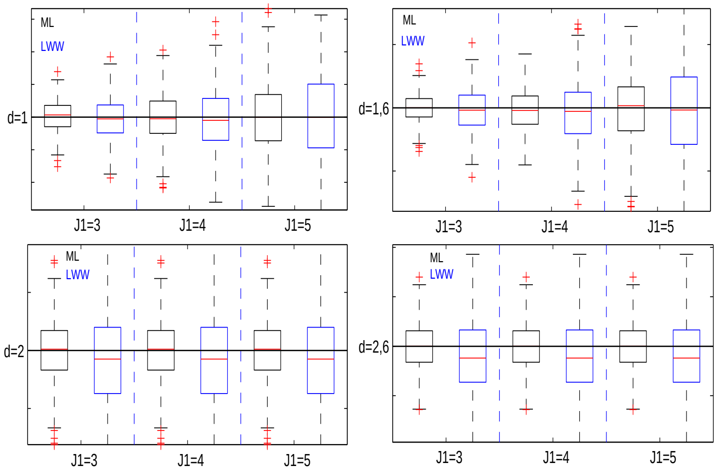

Figure: 2- Comparison of ML and LWW estimators of the memory parameter d of an ARFIMA$(0, d, 0)$ in a semi-parametric frame.

Conclusion

- We have derived an iterative algorithm for the computation of wavelet coefficients covariance matrix that allow us to work beyond the stationary regime.
- When applied to the estimation of the parameters of an generalized ARFIMA(p, d, q) model
  - in the parametric case: we provide an exact maximum likelihood in wavelet domain.
  - in the semi parametric framework, the estimation of d by $\widehat{d}^{\mathrm{ML}}(J_1, J_2)$ yields better results than the one obtain by $\widehat{d}^{\mathrm{LWW}}(J_1, J_2)$ in the first case, we obtain smaller MSE than in the second.

Inference of a generalized long memory process in the wavelet domain.

O. Kouamo E. Moulines and F. Roueff

available upon request

olaf.kouamo@telecom-paristech.fr