SPECTRUM ANALYSIS WITH RECURSIVE LEAST SQUARES
ESTIMATES ON AMPLIFIED G-HARMONICS

Benoît TRUONG-VAN

Département de Mathématiques – CNRS – UA 1204 – Université de Pau
Avenue de l'Université, 64000 PAU, FRANCE.

RESUME

Nous proposons une nouvelle méthode efficace
pour obtenir un estimateur convergent du spectre
mixte de signaux pouvant être représentés par un
modèle non-harmonique général. Cette méthode est
fondée sur une régression récursive sur des G-
harmoniques amplifiés. Nous en donnons quelques
résultats d'application.

MOTS CLES : estimateur convergent, modèle non-
harmonique, ARMA.

SUMMARY

We present a computationally efficient new method
to obtain consistent estimate of the mixed spectrum
of signals that are represented by a general non-
harmonic model. This technique is based upon re-
cursive least squares on amplified G-harmonics.
A description of the method and some numerical
results are given.

KEYWORDS : Consistent estimate, non-harmonic
model, ARMA.

1 - Introduction

The problem considered herein is the estimation of
the spectrum of a signal x_t that can be additi-
vely decomposed into an oscillatory component H_t
and an "irregular" component z_t from part of a
realization of the signal, say x_t , t = 1,...,N .
We assume that the H_t component can be written
as $H_t = \sum_{j=1}^{m} C_j \cos(\omega_j t + \Phi_j)$ (1), where the ampli-
tudes C_j , the phases Φ_j, the angular frequencies
or pulsations ω_j and the number m are unknown,
with the ω_j's being arbitrary in $[0,\pi]$. The
"irregular" part is assumed to be a stationary pu-
rely non-deterministic ARMA process, that we
shall briefly call regular ARMA. (The case where
z_t is some non-stationary ARMA has been studied
in [4]). In other words, the signals considered
herein are assumed to be represented by the follo-
wing general non-harmonic model $x_t = H_t + z_t$ (2)
where H_t satisfies (1) and z_t is a discrete re-
gular ARMA (p,q) process. By extension, each sinu-
soid in (1) will be termed G-harmonic.

The non-harmonic model that is commonly adopted in
most of recent works is a special case of (2), in
which z_t is a white noise. For this model, two
techniques for estimating line spectrum are well-
known, that are the Pisarenko harmonic decomposi-
tion and the extended Prony method (see [2]). The
first one requires complex calculations and is com-
putationally expensive. The second method consists
in estimating the model parameters by least squares

(LS) criterion and in solving this non-linear LS
problem in several steps. The deficiency of this
technique is that the LS estimates of the autore-
gressive (AR) parameters in its AR filtering
step are inconsistent. Another technique for estima-
ting mixed spectrum is the Whittle method (see [1])
that gives consistent estimates. However, when the
number of G-harmonics is greater than two, it leads
to intractable calculations. We present here a new
efficient method to estimate the mixed spectrum of
signals of the form (2). The theoretical aspect and
an algorithm of the method are first described, then
some numerical results and illustrations are given.

2 - Least squares estimates on amplified G-harmonics

Since several methods are now available (see e.g.
[3]) to get consistent estimates for the parameters
of regular ARMA processes, the problem of estima-
ting the mixed spectrum of a signal x_t of the form
(2) amounts to that of identifying the G-harmonics
in (2). In this aim, we introduced in [5] the no-
tion of G-harmonic amplifiers, that is stated below.

Proposition . For each j = 1,...,m , let α_j =
$\cos(\omega_j)$ and let β_j be some real. Define $\xi_j(t)$ as
the unique solution of the following equation

$$\xi_j(t) - 2\beta_j \xi_j(t-1) + \xi_j(t-2) = x_t \qquad (3)$$

with the initial conditions $\xi_j(t_o-1) = \xi_j(t_o)=0$.(4)
If $\beta_j = \alpha_j$, then $\xi_j(t+t_o) = U_j(t+t_o) + Z_j(t+t_o)$,

t ≥ 1 where

$$U_j(t+t_o) = (2 \sin \omega_j)^{-1} C_j [t \sin\{\omega_j(t+1)+\Phi_j\} -$$
$$(\sin \omega_j)^{-1}(\sin \phi_j)\sin(\omega_j(t)] \qquad (5)$$

$$Z_j(t+t_o) = (\sin \omega_j)^{-1} \sum_{k=0}^{t-1} \left[\{\sin(k+1)\omega_j\} \times \right.$$
$$\left. \sum_{h=1,h\neq j}^{m} C_h \cos\{\omega_h(t-1-k)+\Phi_h\} + z_{t-1-k}\right] \cdot (6)$$

Since the first term in (5) is an oscillation with pulsation ω_j , the amplitude of which quasi-linearly increases with t whereas the amplitude of the G-harmonics other than the $j^{\underline{th}}$ remains constant in time (see (6)), hence the $\xi_j(t)$ process will be called harmonic amplifier relative to β_j when β_j is near α_j .

The notion of Recursive Least Squares on Amplified Harmonics (RLSOAH) estimates is now defined.

Definition. For each $j = 1,...,m$ let $\{\hat\alpha_{j,n}\}_{n \in \mathbb{N}}$ be the sequence defined over the set $\mathbb{N}$ of non-negative integers as follows :

(i) $\hat\alpha_{j,o} = \bar\alpha_j$ where $\bar\alpha_j$ is some initial estimate for α_j .

(ii) For any $n \in \mathbb{N}$, $\hat\alpha_{j,n+1}$ is the LS estimate for α_j defined by the regression equation

$$\xi_{j,n}(t) + \xi_{j,n}(t-2) = 2\alpha_j \xi_{j,n}(t-1) + e_{j,n}(t) , \qquad t \geq t_o$$

where $e_{j,n}(t)$ is the error term and the $\xi_{j,n}(t)$'s are the unique solution of (3), in which β_j is replaced by $\hat\alpha_{j,n}$, with the initial conditions (4). Note that the $\xi_{j,n}(t)$'s depend on $\hat\alpha_{j,n}$. The $\{\hat\alpha_{j,n}\}$ is called the sequence of RLSOAH estimates for α_j .

The following basic result (see [5]) establishes for each G-harmonic j , the convergence of the RLSOAH estimates $\{\hat\alpha_{j,n}\}$ towards α_j up to $O_p(N^{-1})$.

Theorem. For any $j \in \{1,...,m\}$ and for any given $n \in \mathbb{N}$, let $\tilde\alpha_j = \hat\alpha_{j,n}$ and $\hat\alpha_j = \hat\alpha_{j,n+1}$. Define $\tilde\omega_j = \text{Arcos}(\tilde\alpha_j) \in [0,\pi]$, $\delta_j = \tilde\omega_j - \omega_j$.

If $\delta_j = O(N^{-1})$ and if N is sufficiently large, then either $\hat\alpha_j - \alpha_j = O_p(N^{-1})$ or $|\hat\alpha_j - \alpha_j| = (1-R_j)|\tilde\alpha_j - \alpha_j|$ with $R_j = r_j + o_p(N^{-1})$ and r_j being a real satisfying $0 < r_j < 1$.

The theorem yields a very simple procedure to perform consistent estimates for the pulsations ω_j , since for each G-harmonic, e.g the $j^{\underline{th}}$, it consists of iterations of the following schema, called one RLSOAH operation :

(1) Perform the $\xi_{j,n}(t)$ amplifier relative to $\hat\alpha_{j,n}$.

(2) Compute the $\hat\alpha_{j,n+1}$ by the formula

$$\hat\alpha_{j,n+1} = \left[(\xi_{j,n}(t)+\xi_{j,n}(t-2)) \; \xi_{j,n}(t-1)\right] / 2(\xi_{j,n}(t-1))^2 .$$

The convergence criterion is $|\hat\alpha_{j,n+1} - \hat\alpha_{j,n}| \leq \varepsilon$. Generally, ε can be choosen equal to 10^{-8}. For each G-harmonic, the number of RLSOAH operations necessary to obtain convergence with $\varepsilon = 10^{-8}$ is about 25 when starting with improper initial estimates (see [6]). For series x_t that contain some G-harmonics with great amplitudes and the others with most weaker amplitudes, we associate with the preceding RLSOAH schema, an additional step that reduces the effect of the greater G-harmonics and thus derive an efficient algorithm called recursive regression on amplified harmonics (RROAH) for estimating the G-harmonics in (2). This algorithm can be outlined as follows. (For a detailed description, see [6]).

Step 1. Starting with initial estimates $\bar\alpha_1,...,\bar\alpha_m$, we perform RLSOAH operations on the series x_t , then identify the number s of the dominating G-harmonics (called of the first level) and obtain primary estimates $\tilde\alpha_1,...,\tilde\alpha_s$ for the α_j's .

Step 2. An improvement of the primary estimates $\tilde\alpha_j$ and a reduction of the effect of the G-harmonics identified in step 1 are here achieved (if $s \geq 2$) by means of a LS technique (the details of which is given in [6]) that yield final RLSOAH estimates $\hat\alpha_1,...,\hat\alpha_s$ and a series, say $\tilde x_t$, which corresponds to the residuals of the x_t's after removing the G-harmonics identified in step 1. The estimates $\hat\omega_j$ of the ω_j's are then computed by $\hat\omega_j = \text{Arcos}(\hat\alpha_j)$.

Step 3. Since all the remaining G-harmonics are in the residuals $\tilde x_t$, we replace the x_t's by the $\tilde x_t$'s then iterate steps 1 and 2 with initial estimates $\bar\alpha_{s+1},...,\bar\alpha_m$ and identify new G-harmonics, called of the second level. The procedure is thus continued until all the G-harmonics in the x_t's are identified.

3. Some numerical resultats and illustrations

We now illustrate some properties of the RROAH procedure on 4 following simulated series with two G-harmonics and with the same length $N = 98$. They are of the form

$$x_t = C_1 \cos(2\pi t/P_1) + C_2 \cos(2\pi t/P_2) + z_t ,$$

where z_t is an MA(1) process with the parameter $\Theta = 0.7$ and the noise standard deviation $\sigma = 0.2$.

Series B1 : $C_1=15$, $C_2=3$, $P_1=14.4$, $P_2=6.5$
Series B2 : $C_1=15$, $C_2=14$, $P_1=14.4$, $P_2=6.5$
Series B3 : $C_1=15$, $C_2=0.4$, $P_1=14.4$, $P_2=6.5$
Series B4 : $C_1=8$, $C_2=7$, $P_1=6.5$, $P_2=6.87$.

1) The accuracy of the RLSOAH estimates is very satisfying. For example, for series B2, the estimates are $\hat{\alpha}_1 = 0.90617$ ($\hat{P}_1 = 14.389$) and $\hat{\alpha} = 0.56823$ ($\hat{P}_2 = 6.501$).

2) Even when the initial estimates are far from the true values, the RROAH procedure converges to the same limits. For instance, for series B1, with the following initial estimate $\bar{\alpha}_1 = -0.955, 0.148$ or 0.891, we obtain the same value for $\hat{\alpha}_1$, namely 0.90619 ($\hat{P}_1 = 14.39$). (Further results are given in [6] as well as methods to obtain the initial estimates $\bar{\alpha}_j$).

3) For G-harmonics with very weak amplitudes, the accuracy of the RLSOAH estimates remains very satisfying. For example, with series B3, the second G-harmonic of which has an amplitude nearly equal to the noise deviation σ , we obtain $\hat{\alpha}_1 = 0.90618$ ($\hat{P}_1 = 14.39$) at the first level and $\hat{\alpha}_2 = 0.56625$ ($\hat{P}_2 = 6.48$) at the second level.

4) According to the theorem, the RROAH procedure has a good frequency resolution in a certain range of frequencies : Two frequency peaks, say at ω_i and ω_j , such that $|\alpha_i - \alpha_j| > 1/N$, can be easily distinguished by the RLSOAH estimates. Thus, with series B4 that contains two periods $P_1 = 6.5$ and $P_2 = 6.87$ near each other, we obtain $\hat{\alpha}_1 = 0.56424$ ($\hat{P}_1 = 6.47$) and $\hat{\alpha}_2 = 0.61425$ ($\hat{P}_2 = 6.91$).

The RROAH estimation procedure was also successfully applied to analyse spectrum of actual data, since with this method, we have shown that the logarithmically transformed series of the well-known Canadian lynx trapping series has a mixed spectrum constituted by two G-harmonics at the periods 9.62 and 5.13 and by the continuous spectrum of an AR(2) process. For the detailed results of this application and further complements, we refer the reader to [6] .
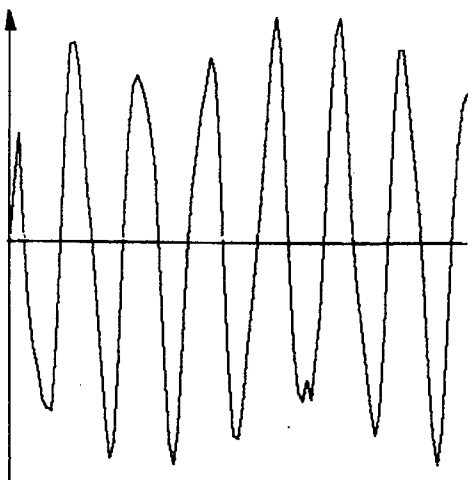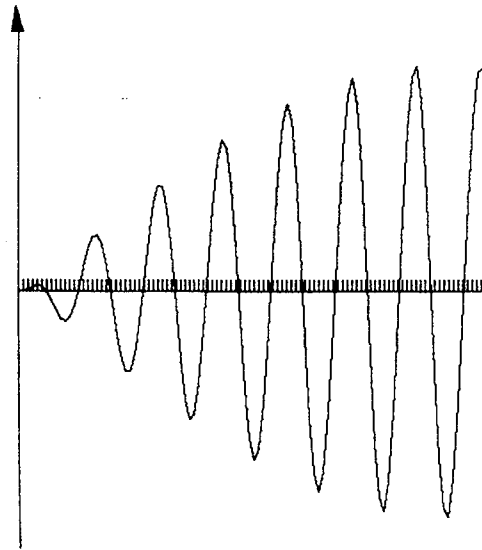

Fig.1.   Series B1

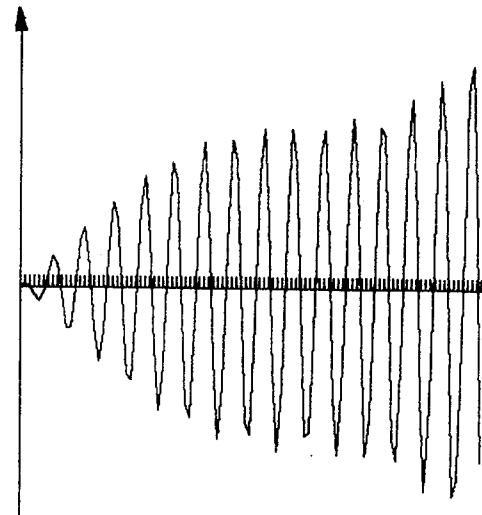Fig.2 $\tilde{\alpha}_1$-amplifier of series B1 with $\tilde{\alpha}_1 = 0.9061965$

Fig.3. $\tilde{\alpha}_2$-amplifier of series B4 with $\tilde{\alpha}_2 = 0.618801$

References

[1] CAMPBELL, M.J. and WALKER, A.M., A survey of statistical work on the Mackenzie River series of annual Canadian lynx trappings for the years 1821-1934 and a new analysis. J.R. Statist.Soc. A, 1977, pp. 411-431.

[2] KAY, S.M. and MARPLE, S.L., Spectrum analysis, A modern perspective. Proceeding of the IEEE, 1981, pp.1380-1419.

[3] TSAY, R.S. and TIAO, G.C., Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary ARMA models. J.A.S.A, 1984, pp. 84-96.

[4] TRUONG-VAN, B., GARIMA processes as an extension of seasonal ARIMA models to time series with harmonic components. Publications de l'Université de Pau, France, 1986.

[5] TRUONG-VAN, B., A new approach to frequency analysis with recursive least squares estimates on amplified harmonics. CNRS-UA 1204, Université de Pau, France, 1986.

[6] TRUONG-VAN, B. Consistent estimation for the spectrum of G-harmonic ARIMA models with application to the lynx series. CNRS-UA 1204, Université de Pau, France, 1986.