FREQUENCY ESTIMATION OF MULTIPLICATIVE ARMA NOISY DATA

Mounir GHOGHO

ENSEEIHT/GAPSE, 2 rue Camichel, 31071 Toulouse, France
Tel: (33) 61 58 83 50 fax: (33) 61 58 82 37
email: Ghogho@len7.enseeiht.fr

ABSTRACT

Parameter estimation for multiplicative noisy data is a pertinent signal processing problem encountered in a wide range of signalling and data-processing applications, including radar, sonar, radio astronomy, seismology and vibroacoustics. The assumption of additive noise is, in these contexts, insufficient for adequate signal modeling. The model considered here incorporates the gaussian amplitude-modulated sinusoids. New algorithms are developed for frequency estimation. The corresponding probability density, prediction, innovation process and ergodicity property are presented. Higher order statistics are used, especially when the process is also contaminated by an additive noise.

I- INTRODUCTION

Statistically, the multiplication of two random processes is a nonlinear operation. The problem of describing the resulting process is one of great complexity and no systematic theory comparable to that for linear transformation is available. In this paper, we suppose that the sinusoid's amplitude fluctuation is regular enough to be represented as a gaussian ARMA process. The resulting spectrum is a convolution of the ARMA spectrum and the two dirac measures concentrated at the frequencies sinusoids. Using this property, which is important only when the autoregressive order is greater or equal to one, [1] estimates the Doppler signal frequency. This paper deals with identification and parameter estimation for a general class of signals, called ARMACOS processes, which can be written as:

$$u(t) = \sum_{i=1}^{m} y_i(t)\cos(\omega_i t + \phi_i) + v(t) = x(t) + v(t),$$

where ϕ_i's are independent random variables uniformly distributed over $[0, 2\pi]$, $v(t)$ is a colored gaussian additive noise and the processes $y_i(t)$, $i = 1, \ldots, m$, are assumed to be independent, stationary and gaussian ARMA(p_i,q_i) i.e.

$$y_i(t) + \sum_{k=1}^{p_i} c_k^i\, y_i(t-k) = e_i(t) + \sum_{k=1}^{q_i} d_k^i\, e_i(t-k),$$ with

$\{e_i(t)\}_{i=1,\ldots m}$ being independent zero-mean white gaussian processes with variance σ_i^2. The processes and the random variables defined in the model are assumed to be independent.

II- SPECTRAL DENSITY OF $x(t)$

Since the ϕ_i's are uniformly distributed over $[0, 2\pi]$, then $x(t)$ is zero-mean stationary with a covariance function:

$$R_x(\tau) = \sum_{i=1}^{m} R_{x_i}(\tau) = \frac{1}{2}\sum_{i=1}^{m} R_{y_i}(\tau)\cos(\omega_i\tau),$$

where $x_i(t) = y_i(t)\cos(\omega_i t + \phi_i)$. The power spectral density (PSD) of each component $x_i(t)$ is:

$$S_{x_i}(z) = \frac{1}{4}\left[S_{y_i}(ze^{j\omega_i}) + S_{y_i}(ze^{-j\omega_i})\right],$$

$$= \frac{\sigma_i^2}{4}\left[\frac{G_i(ze^{j\omega_i})F_i(ze^{-j\omega_i}) + G_i(ze^{-j\omega_i})F_i(ze^{j\omega_i})}{F_i(ze^{j\omega_i})F_i(ze^{-j\omega_i})}\right],$$

where $G_i(z) = D_i(z)D_i^*(\frac{1}{z})$ and $F_i(z) = C_i(z)C_i^*(\frac{1}{z})$.

If z_k is a zero of the numerator, then $\frac{1}{z_k}$ is also a zero of this numerator. The PSD of $x_i(t)$ is then given by:

$$S_{x_i}(z) = \frac{\lambda_i^2\, B_i(z)\, B_i^*(\frac{1}{z^*})}{A_i(z)\, A_i^*(\frac{1}{z^*})}.$$

Where $A_i(z) = C_i(ze^{j\omega_i})C_i(ze^{-j\omega_i})$, $B_i(z)$ is a polynomial of degree $p_i + q_i$, and λ_i^2 is a constant.

We now derive the PSD of $x(t)$,

$$S_x(z) = \sum_{i=1}^{m} \frac{\lambda_i^2\, B_i(z)\, B_i^*(\frac{1}{z^*})}{A_i(z)\, A_i^*(\frac{1}{z^*})}$$

$$= \frac{\sum_{i=1}^{m}\left[\lambda_i^2\, B_i(z) B_i^*(\frac{1}{z^*}) \prod_{\substack{k=1,\ldots,m\\ k\neq i}} A_i(z)\, A_i^*(\frac{1}{z^*})\right]}{\prod_{k=1,\ldots,m} A_i(z)\, A_i^*(\frac{1}{z^*})}$$

which can be written as: $S_x(z) = \lambda^2 \dfrac{B(z)\, B^*(\frac{1}{z^*})}{A(z)\, A^*(\frac{1}{z^*})}$,

where $A(z)$ is a polynomial of degree $P = 2\sum_{i=1}^{m} p_i$

0-7803-2431-5/95 $4.00 © 1995 IEEE

defined by $A(z)=\prod_{i=1\ldots m} A_i(z)$, $B(z)$ is a polynomial of degree $Q=\max_i \{p_i+q_i+2\sum_{k=1;k\neq i}^{m} p_k\}$.

III- Probability density, prediction and innovation.

In the following, our study is limited to a single(m=1) ARMACOS process $x(t)=y(t)\cos(\omega_o t+\phi)$, where $y(t)$ is a zero-mean gaussian ARMA process with parameters $(\underline{c},\underline{d},\sigma^2)$. Thus, $x(t)$ is a wide sense ARMA(2p, p+q) process. Let us derive the probability density of the ARMACOS process. A zero-mean gaussian process is completely defined by its covariance function; hence, the a posteriori probability density of $X=(x_1,\ldots,x_T)'$ is given by:

$$f(X\mid\phi) = (2\pi)^{-\frac{T}{2}}\det(R_{x/\phi})^{-\frac{1}{2}}\exp\left\{-\tfrac{1}{2}(X^t R_{x/\phi}^{-1}X)\right\}$$

where $R_{x/\phi}$ is the conditional covariance matrix which is nonsingular with probability 1. Its elements are given by $R_y(i-j)\cos(\omega i+\phi)\cos(\omega j+\phi)$ for $i,j=1,\ldots,T$. The a priori distribution is obtained by averaging, and can be written as:

$$f(X) = \tfrac{1}{2\pi}\int_0^{2\pi} f(X\mid\phi)\,d\phi$$

$$=2(2\pi)^{-\frac{T}{2}-1}\int_0^{\pi}\det(R_{x/\phi})^{-\frac{1}{2}}\exp\left\{-\tfrac{1}{2}(X^t R_{x/\phi}^{-1}X)\right\}d\phi$$

A linear transformation of X; $S=KX+M$, can never be gaussian because its probability density is given by: $f(S)=$

$$2(2\pi)^{-\frac{k}{2}-1}\int_0^{\pi}\det(KR_{x/\phi}K^t)^{-\frac{1}{2}}\exp-\tfrac{1}{2}\left\{(S-M)^t(KR_{x/\phi}K^t)^{-1}(S-M)\right\}d\phi$$

where K is a matrix (k,T) and M is a vector $(k,1)$. Since the spectrum of $x(t)$ is rational, then for any irreducible Fejer-Riesz representation of the spectrum, there exists a zero mean white noise $\varepsilon(t)$ with variance λ^2 which is related to the signal $x(t)$ by the ARMA relation:

$$x(t)+\sum_{k=1}^{P} a_k\; x(t-k) = \varepsilon(t)+\sum_{k=1}^{Q} b_k\;\varepsilon(t-k).$$

The white noise associated with the minimal ARMA representation (i.e. the $B(z)$ roots are also inside the unit circle) is the innovation process. Thus, we establish a bijection between the covariance function of $x(t)$ and the triple $(\lambda^2,\underline{a},\underline{b})$. We use the recursive method based on the innovation algorithm (see [3]) for computing the best linear predictor $\hat{x}(t)$ of the ARMA process $x(t)$ and its normalized mean square error $r(t)$. When T is large enough, $x(t)$ can be passed through the inverse filter $A(z)/B(z)$ for parameter estimation. We note that $\varepsilon(t)$ is not a strict white noise, i.e. not an independent sequence.

IV- Cumulants estimation and ergodicity

In what follows, we will use the moments $M(.)$ and cumulants $C(.)$ of order two and four. From a sample $(u_1,\ldots,u_T)^t$, We estimate these as follow:

for $\tau\geq 0$ and $\tau_3\geq\tau_2\geq\tau_1\geq 0$,

$$\hat{C}_{2u}(\tau)=\hat{R}_u(\tau)=\tfrac{1}{T-\tau}\sum_{t=1}^{T-\tau}u(t)u(t+\tau),$$

$$\hat{M}_{4u}(\tau_1,\tau_2,\tau_3)=\frac{1}{T-\tau_3}\sum_{t=1}^{T-\tau_3}u(t)u(t+\tau_1)u(t+\tau_2)u(t+\tau_3),$$

$$\hat{C}_{4u}(\tau_1,\tau_2,\tau_3)=\hat{M}_{4u}(\tau_1,\tau_2,\tau_3)-\hat{R}_u(\tau_1)\hat{R}_u(\tau_2-\tau_3)-\hat{R}_u(\tau_2)\hat{R}_u(\tau_1-\tau_3)-\hat{R}_u(\tau_3)\hat{R}_u(\tau_1-\tau_2)$$

For the ARMACOS process $\{x_t\}$, we have:

$$\bar{x}=\frac{1}{T}\sum_{t=1}^{T}x_t \xrightarrow[T\to\infty]{} 0,$$

$$\hat{R}_x(\tau) \xrightarrow[T\to\infty]{} R_x(\tau),$$

$$\hat{C}_{4x}(\tau_1,\tau_2,\tau_3)\xrightarrow[T\to\infty]{} C_{4x}(\tau_1,\tau_2,\tau_3),$$

with probability 1.

Proof (see appendix)

V- ARMACOS PARAMETER ESTIMATION ($p\geq 1$)

For order estimation, methods based on the covariance matrix singularity, such as the corner method ([2]), can be used. Even if X is non gaussian, the analytical expression of its joint distribution is given. Thus, the maximum likelihood estimator of ω_0 is the one that maximizes $f(X)$. However, this method has two drawbacks; first, it needs a numerical integration at each step in the maximization procedure; second, it can be numerically unstable due to the matrix inversion. And even if $f(X\mid\phi)$ can be expressed without matrix inversion by using the two parameter vectors $\underline{c}$ and $\underline{d}$, the numerical instability subsists. Thus, the ARMA representation of the ARMACOS process makes parameter estimation possible. It is to be noted that there is no bijection between the ARMACOS process and its spectral density, i.e. several ARMACOS have the same covariance function. Therefore, for the identification of each pair of poles $(\rho_k\, e^{i(\lambda_k+\omega_0)},\ \rho_k\, e^{i(-\lambda_k+\omega_0)})$, where $\rho_k\, e^{\pm i(\lambda_k)}$, $k=1,\ldots,p$, are $C(z)$ roots, we should have some a priori knowledge process. For example, ambiguity vanishes if we assume that $\omega_0<\frac{\pi}{2}$ and

$\lambda_k < \omega_0$. In the following, we suppose that there is no ambiguity identification. In the ARCOS case, for Doppler frequency estimation, [1] uses the centroid of the estimated AR frequencies. Since there is not only spectral equivalence between ARMACOS and a special ARMA, as mentioned in [1], but also a temporal equivalence, an intuitive estimation procedure is to minimize the weighted sum of the squares:

$$S(\omega_0, \underline{c}, \underline{d}) = \sum_{t=1}^{T} (x_t - \hat{x}_t)^2 / r_{t-1}$$

with respect to ω_0, $\underline{c}$ and $\underline{d}$,. The algorithm, which will be referred to as the "least squares" estimator is :

Algorithm 1 :

1) initialization:

* $\underline{a}^{(0)}$ is the Solution of the modified Yule-Walker (MYW) equations.

* $\omega_o^{(0)}$ is the centroid of $A^0(z)$ frequencies,

$\underline{c}^{(0)}$ is such that $C^{(0)}(ze^{i\omega_o^{(0)}})\, C^{(0)}(ze^{-i\omega_o^{(0)}}) = A^{(0)}(z)$,

$\underline{d}^{(0)}$ is a random vector.

2) minimization : $(\hat{\omega}_0, \hat{\underline{c}}, \hat{\underline{d}}) = \arg\min S(\omega_o, \underline{c}, \underline{d})$.

In step 2, a priori knowledge about ARMA parameters is used. It is interesting to note that, after initialization, minimization is very quick, especially in the ARCOS case because step one give a good initialization for both $\underline{a}$ and $\underline{b}$. In this case, Figure 1 represents the mean square errors (m.s.e.), based on 1000 realizations, of the frequency estimation ($f_o = \omega_o / 2\pi$) as a function of the number of samples, for both our algorithm and the one proposed in [1], which is based on least squares modified Yule-Walker equations (LSMYW) and the centroid idea. in the latter, the number of correlation points M was set to 100.

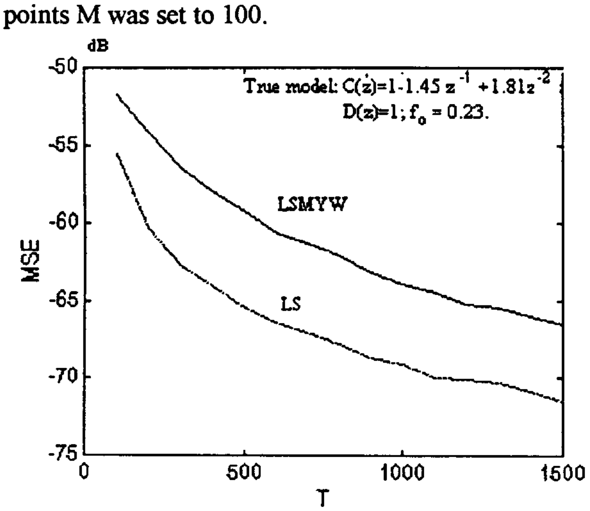

Fig. 1

VI- ADDITIVE NOISY CASE
($p \geq 1$)

When $x(t)$ is embedded in an additive colored gaussian noise $v(t)$, the diagonal slice of the fourth-order cumulant is used :

$$\begin{aligned} C_{4u}(\tau) &= C_{4u}(\tau,\tau,\tau) \\ &= C_{4x}(\tau,\tau,\tau) + C_{4v}(\tau,\tau,\tau) \\ &= C_{4x}(\tau,\tau,\tau) = \tfrac{3}{4} R_y(0) R_x(\tau). \end{aligned}$$

In order to estimate ω_0, the LSMYW method is applied to $\hat{C}_{4u}(\tau)$;

Algorithm 2 :

- Estimate $\underline{a}$ by solving the LSMYW equations:

$$\begin{pmatrix} \hat{C}_{4u}(p+q) & \hat{C}_{4u}(p+q-1) & \cdots & \hat{C}_{4u}(q-p+1) \\ \hat{C}_{4u}(p+q+1) & \hat{C}_{4u}(p+q) & \cdots & \hat{C}_{4u}(q-p+2) \\ \vdots & \vdots & \cdots & \vdots \\ \hat{C}_{4u}(M-1) & \hat{C}_{4u}(M-2) & \cdots & \hat{C}_{4u}(M-2p) \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_{2p} \end{pmatrix} = -\begin{pmatrix} \hat{C}_{4u}(p+q+1) \\ \hat{C}_{4u}(p+q+2) \\ \vdots \\ \hat{C}_{4u}(M) \end{pmatrix}$$

- Find the $\hat{A}(z)$ frequencies $\hat{\omega}_k$.

- $\hat{\omega}_0 = \frac{1}{p} \sum_{k=1}^{p} \hat{\omega}_k$

VII- MACOS case ($p = 0$)

In this case, the covariance function has a cut-off after lag q. Therefore any frequency estimation based on the covariance function is useless, especially when q=0. We use the following cumulant:

$$C_{4u}(0,\tau,\tau) = \tfrac{1}{8} R_y^2(0) \cos(2\omega_0 \tau) + \tfrac{1}{4} R_y^2(\tau).$$

Thus, for our MACOS process,

$$C_{4u}(0,\tau,\tau) = \tfrac{1}{8} R_y^2(0) \cos(2\omega_0 \tau), \text{ for } \tau > q.$$

Therefore, if $\omega_0 < \frac{\pi}{2}$, then it can be estimated using any technique of frequency estimation in additive noise. We can also estimate the frequency by using the fast Fourier transform of the process $\{s_t\} = \{x_t^2\}$. In fact,

$$\begin{aligned} \bar{s}_T(\omega) &= \frac{1}{T}\sum_{t=1}^{T} s_t\, e^{-j\omega t} \\ &= \frac{1}{2T}\sum_{t=1}^{T} y_t^2\, e^{-j\omega t} + \frac{1}{2T}\sum_{t=1}^{T} y_t^2 \cos(2\omega_o t + 2\phi)\, e^{-j\omega t} \\ &= \frac{1}{2T}\sum_{t=1}^{T} (y_t^2 - R_y^2(0))\, (e^{-j\omega t} + \tfrac{1}{2} e^{j((2\omega_o - \omega)t + 2\phi)} + \tfrac{1}{2} e^{-j((2\omega_o + \omega)t + 2\phi)}) \\ &\quad + R_y^2(0) \frac{1}{2T}\sum_{t=1}^{T} (e^{-j\omega t} + \tfrac{1}{2} e^{j((2\omega_o - \omega)t + 2\phi)} + \tfrac{1}{2} e^{-j((2\omega_o + \omega)t + 2\phi)}) \end{aligned}$$

The first term converges to 0 with probability 1 as $T \to \infty$ (see appendix of the ergodicity property). And if $\omega \neq 0, \pm 2\omega_0$ then, The second term is $O(\frac{1+j}{T})$. Thus,

$$\bar{s}_T(\omega) \xrightarrow[T\to\infty]{} \tfrac{1}{2} R_y^2(0)\left[\delta(\omega) + \tfrac{1}{2}\delta(\omega - 2\omega_o) + \tfrac{1}{2}\delta(\omega + 2\omega_o)\right]$$

with probability 1. This result holds for all ARMACOS processes. The method is simple to implement and it is computationally cheap thanks to the fast Fourier algorithm. We can also estimate the mean of s and get rid of the peak at zero in the spectrum. The following figure illustrates this result when y_t is a zero mean gaussian white noise with unit variance:

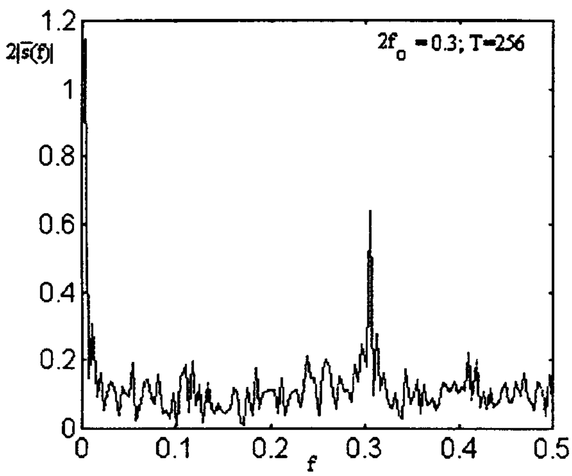

Fig. 2.

VIII- CONCLUSIONS

In this paper, an important case of multiplicative processes have been examined. The ARMA representation of the ARMACOS processes allows frequency estimation if the autoregressive order is finite and not zero. When the ARMACOS processes are contaminated by a gaussian additive noise, or the autoregressive order is zero, higher order statistics are used.

Appendix

1- Ergodicity of the mean:

For the stationary ARMACOS process x_t, $R_x(0) < \infty$ and $R_x(\tau) \to 0$ as $|\tau| \to \infty$, then x_t is mean ergodic.

2- Correlation ergodicity:

let $\{z_t\}$ be strictly stationary with zero mean, and its spectral distribution function with no jump at $\omega\ (-\pi \leq \omega \leq \pi)$. It may be shown that the quantity:

$$\bar{z}_T(w) = \frac{1}{T}\sum_{t=1}^{T} z_t\, e^{itw},$$

converges to zero with probability 1, according to the ergodic theorem and the fact that $\bar{z}_T(w)$ converges in mean square to a limit which is 0 with probability 1. A completely analogous result is obtained, for a fixed ϕ, for the quantity :

$$\frac{1}{T}\sum_{t=1}^{T} z_t \cos(\omega t + \phi) = \cos(\phi)\frac{1}{T}\sum_{t=1}^{T} z_t \cos(\omega t) - \sin(\phi)\frac{1}{T}\sum_{t=1}^{T} z_t \sin(\omega t)$$

which is independent of ϕ as $T \to \infty$. Then, the unconditional convergence properties are the same as those of the conditional ones.

For a fixed τ, Let the stationary process $\{z_t\} = \{y_t y_{t+\tau} - R_y(\tau)\}$, then:

$$\hat{R}_x(\tau) = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau} y_t y_{t+\tau} \cos(\omega_o t + \phi)\cos(\omega_o (t+\tau) + \phi)$$

$$= \frac{\cos(\omega_o \tau)}{2}\frac{1}{T-\tau}\sum_{t=1}^{T-\tau} y_t y_{t+\tau} + \frac{1}{2(T-\tau)}\sum_{t=1}^{T-\tau} z_t \cos(\omega_o(2t+\tau) + 2\phi)$$

$$+ \frac{R_y(\tau)}{2(T-\tau)}\sum_{t=1}^{T-\tau}\cos(\omega_o(2t+\tau) + 2\phi)$$

The process z_t verifies the conditions of the previous property. Thus, in the right hand term, the second element converges to 0 with probability 1 and the third one is $O(\frac{1}{T})$. Hence, since y_t is ergodic, $\hat{R}_x(\tau)$ converges to $R_x(\tau)$ with probability 1 as $T \to \infty$.

3- Fourth order cumulant ergodicity:

We apply the same method to the process $\{y_t y_{t+\tau_1} y_{t+\tau_2} y_{t+\tau_3} - M(\tau_1, \tau_2, \tau_3)\}$. Thus, the moments and cumulants are the limits, with probability 1 as $T \to \infty$, of their estimates.

References

[1] O. Besson & F. Castanié, "*On estimating the frequency of a sinusoid in autoregressive multiplicative noise*", Signal processing, Vol. 30, No. 1, January 1993

[2] A. Berlinet & C. Francq, "*Identification of a univariate ARMA model*", Computational statistic(1994), Vol. 9, pp. 117-133.

[3] Brockwell & Davis (1991). *Time series: theory and methods*: Springer-Verlag.

[4] R. F. Dwyer, "*Fourth-order spectra of gaussian amplitude-modulated sinusoids*", J. Acoust. Soc. Am. 90 (2), Pt. 1, August 1991.

[5] E. J. Hannan (1970), " *Multiple time series*", New York, John Wiley

[6] A. Swami, "*Multiplicative noise models: parameter estimation using cumulants*", Signal processing, Vol.36, No. 3, April 1994, pp. 355-373.

[7] A. Swami & M. Mendel, " *Cumulant-based approach to the harmonic retrieval and related problems*", IEEE Trans, Acoust., Speech, Signal processing, Vol. 39, No. 5, May 1991.