TREIZIÈME COLLOQUE GRETSI - JUAN-LES-PINS DU 16 AU 20 SEPTEMBRE 1991

Spectral Estimation with Stochastic Coefficients Models

DE LIMA-VEIGA Alvaro,
PUC-Rio,
Rua Marquês de São Vicente, 225,
22453 - Rio de Janeiro - RJ,
Brésil.

GRENIER Yves,
ENST,
46, rue Barrault,
75634 Paris, Cédex 13,
FRANCE

Résumé - Cet article présente une application des modèles autorégressifs à coefficients aléatoires à l'estimation spectrale des signaux non-stationnaires. Dans ces modèles, l'évolution temporelle des coefficients est décrite comme la somme d'une combinaison linéaire des fonctions du temps connues *a priori* et un bruit blanc vectoriel. Deux méthodes d'estimation basées sur le maximum de vraisemblance sont présentées et testées sur des signaux synthétiques présentant des variations spectrales lentes et rapides. Dans ces simulations nous explorons deux interprétations possibles pour la partie aléatoire du modèle et les confrontons aux modéles ordinaires sans composantes aléatoires dans les coefficients.

Abstract - This article presents an application of time-varying autoregressif models to the spectral estimation of non-stationary signals. In these models, the trajectories of the coefficients are represented by the sum of linear combination of known time-functions and a vector white noise. We present two maximum likelihood methods and test them with different signals presenting smooth or quick spectral variations. In these simulations we explore two possible interpretations for the stochastic part of the model and compare the performance of the stochastic approach to comparable ordinary models with no stochastic coefficents .

1 Introduction

Discrete-time linear models are maybe the most widely used models in signals and time-series analysis. Its application is supported by a well developed theory for the case where the stochastic process generating the observations (or some transformation of it) is stationary; however this is not the case for most applications. A typical way to overcome this difficult is to perform the analysis over short segments where stationarity can be assumed or to define some kind of fading memory around the instant of analysis.

In recent years there has been a steadly growing interest in models that are non-stationary in their structure. One important class of such models is that of **evolutionary models**, where the time-varying coefficients are described by a linear combination of known time functions [1]. Another class of non-stationary models is that of **stochastic coefficient models** [2, 3,4]. In these models, the vector of coefficients is considered as the realization of a stochastic process, which turns the whole model non-linear. These two classes can in fact be collapsed into one class called **evolutionary with stochastic coefficients**, studied in [5]. In these models, the trajectories of the coefficients are described as a sum of a combination of time functions (as in ordinary evolutionary models) plus a stochastic process.

In this paper we consider the particular case of autoregressif models with stochastic coefficients represented by the sum of a weighted combination of time functions and a vector white noise. The model has been applied to the spectral estimation of non-stationary signals and compared with a ordinary evolutionary model (with no stochastic coefficients). We compare their performances in two experiences with synthetic signals and show that the non-linarity introduced by the stochastic coefficients assumption not only can improve the description power of ordinary models, but also indicate structural changings in the process under study.

The article is organized as follows: section 2 describes the model and how to put it in a state space form, in section 3 we recall two estimation methods [5], based on different approachs to compute the likelihood function and how to maximize it. Finally, section 3 describes the experiences and comment their results.

2 Model Description

Let y_t be a univariate processes described by a linear time-varying model,

$$y_t = x_t'\beta_t + v_t \qquad (1)$$

where v_t is a scalar gaussian white noise with variance σ^2 , x_t is a vector of dimension k containing past values of the signal, $x_t' = [y_{t-1}, ..., y_{t-k}]$ and β_t is a k-dimentional parameter vector containing the time-varying coefficients of the model. Its temporal evolution is represented by a equation with two parts: one deterministc represented by a linear

combination of known time functions and a stochastic one represented by a vectoral gaussian white noise, ϵ_t with an unkowmn covariance matrix Σ,

$$\beta_t = Bz_t + \epsilon_t \tag{2}$$

In this equation B is a $k \times k$ matrix of weights (to be estimated) and z_t is a m-dimensional vector containing known-time functions.

$$z_t = [z_0(t), ..., z_{m-1}(t)]'$$

Using (1) and (2) we can rewrite the model as

$$y_t = x_t' B z_t + x_t' \epsilon_t + v_t \tag{3}$$

The right side of this equation shows three terms. The first one represents an evolutionary model identical to the one studied in [1] and will be refered as the *deterministic component* of the model. The remaining terms define a stochastic error around the first term and will refered as the *stochastic component* of the model.

Since y_t and ϵ_t are both random variables, the determination of the the density of y_t is a difficult task. Nevertheless, we can easily verify that its density is gaussian when conditioned to the vector x_t. In fact, y_t/x_t is a linear transform of β_t which is a multivariate gaussian variable. This is the key property to the formulation of the estimation method described in what follows.

3 The Identification Problem

The model presents three unknown elements that must be estimated. The weighting matrix B, which defines the deterministic part of the model, and the (co)variances Σ and σ^2 defining the stochastic part of the model. This is a highly non-linear problem since the estimation of one component depends on the knowledge of the other. We examine two variations of a maximum likelihood method based on diferent ways of evaluating the likelihood function defining indiretly the way matrix B is estimated. The first one is based on a **input-output** formulation of the model and the second on a **state-space** formulation, denomination that will be extended to the corresponding estimation methods. Both methods will depend on a non-linear search method to estimate the parameters of the stochastic part of the model. The solution space however is not unrestricted because both Σ and σ^2 must be non-negative by definition. Since the restriction $\Sigma \geq 0$ cannot be explicitily represented, the maximization will be performed in terms of its Cholesky decomposition,

$$\Sigma = DD'$$

where D is a lower triangular matrix. This way, Σ will always be non-negative for all matrix D.

3.1 Input-output formulation:

Let us define u_t as the stochastic component of the model (3);

$$u_t = x_t' \epsilon_t + v_t$$

The random variable u_t condionned to the knowledge of x_t represents the prediction error at time t. It is easy to see that these prediction errors given by,

$$\eta_t = (y_t - x_t' B z_t) \tag{4}$$

form a sequence of independent gaussian random variables with zero mean and variances given by,

$$h_t = x_t' D D' x_t + \sigma^2 \tag{5}$$

Then, the log-likelihood function can be expressed as,

$$L(B, D, \sigma^2) = C - \sum_{t=1}^{T} [\log h_t + \eta_t^2 h_t^{-1}] \tag{6}$$

where C is a constant.

Maximization of the log-likelihood

The value of the matrix B is readly obtained in terms of the parameters of the stochastic component and is given by;[1]

$$vec\ B = \left[\sum_{t=1}^{T} \frac{(z_t \otimes x_t)(z_t \otimes x_t)'}{x_t' \Sigma x_t + \sigma^2}\right]^{-1} \sum_{t=1}^{T} \frac{(z_t \otimes x_t) y_t}{x_t' \Sigma x_t + \sigma^2} \tag{7}$$

The stochastic part can then be obtained by a non-linear search algorithm. The gradient of (6) with respect to the non-null elements of D and σ^2 is easily obtained even though one can only assure local optimality, as this is a multimodal function.

The estimation algorithm consist of two steps: computing B using equation (7) and then maximizing (6) with respect to the parameters D and σ^2 until convergence is achieved.

3.2 State-space formulation

The model (1) - (2) can be cast in a state space form. We define the state vector containing the elements that are needed by the instant t to compute y_t. These elements are the matrice B and the vector process ϵ_t. We define,

$$\alpha_t' = [(vec\ B)', \epsilon_t']$$

Then,

$$\alpha_t = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} 0 \\ I \end{bmatrix}$$

Despite of its unnatural form, this formulation provides the proper framework to obtain the estimated values of vectors coefficient β_t that will be used in the second experience described in section 4. Besides, assuming that the parameters of the stochastic part of the model σ^2 and Σ_a are known, a MAP estimate of the B is given for each time instant by the application of a Kalman filter [6] to the above state space equations. Also, as is well known [7], the Kalman filter performs an orthogonal decomposition of the signal. Because we treat here the gaussian case, the inovation sequence η_t, given by,

$$\eta_t = y_t - (z_t \otimes x_t)' b_{t/t-1} \tag{8}$$

[1] *vec* B stands for the vector obtained from B by stacking its columns of B one on top of the other, in order, from left to right. $A \otimes B$ is the Kronecker product of A and B.

forms a sequence of independent gaussian random variable with zero mean and variance given by,

$$h_t = (z_t \otimes x_t)' P_{t-1}(z_t \otimes x_t) + x_t DD' x_t + \sigma^2 \qquad (9)$$

In the above equations, $b_{t/t-1}$ is the estimated value of $vec\ B$ given the observations until time $t-1$ and P_t is the covariance of the estimation error.

Maximization of the log-likelihood

As in the **input-output** formulation, the log-likelihood can be expressed by (6) with h_t and η_t replaced by the ones in equations (8), (9). Please note that in this case, the likelihood is independent of B. Actually, it depends on its estimated value given by the Kalman filter. The maximization of the log-likelihood needs a non-linear search algorithm depending on the computation of the gradient of (6) in respect to the parameters. This is done by side recursions obtained by direct differentiation of the Kalman filter equations. In practice, we reduce the complexity of the algorithm by using linearizations to search for the optimal step in the unidierectional search step of the non-linear maximization algorithm.

4 Results and Conclusions

In this section we present simulation results obtained with two kinds of nonstationary signals. In the first experience, the signals are generated by an evolutionary autoregressif system with stochastic coefficients. In this case, the randomness of the coefficients are attributed to variations of the system produced as the signals are generated. The fidelity of the estimated trajectories will then concentrate only on the determinstc part of the model.

Figure 1 (next page) shows a realization of a signal produced by a evolutionary autoregressif process of order $k = 4$ and with $m = 3$ functions extracted from the Fourier base. In the same figure we show the coefficient trajectories and the corresponding time-frequency representation. In figure 2 (this page), we show the determinstc part of the trajectories (in which we are interest) and its time-frequency representation. Next, in figure 3, we represent the euclidien distance between the estimated and theorical time-frequence representations (TFR) obtained for each iteration of the estimation algorithm, computed over 20 realizations of the signal. We compare the two estimation methods to the results obtained by least square estimation of a evolutionary model with same order and base but with no stochastic component ($\Sigma = 0$). The clear superiority of the estimations given by the stochastic coefficients approach shows its advantages even in the case where only the determinstc component of the trajectories are of interest.

For the second experience, we generate signals presenting spectral jumps causing the trajectories of the coefficients to be step-wise constant. In this case, the stochastic component represent the inability of the determinist component to describe such trajectories with the functions on the base z_t. Figure 4 shows a realization of the signal, the coefficients trajectories and its TFR. Figure 5, shows a example of the trajectories estimated when considering both components of the model. The spectral distances the theorical and estimated TFR's obtained from the 1) determinist trajectories, 2) determinstc+stochastic trajectories and after 3) performing a gliding windowing median-type smoothing of the estimated time-frequency representation, are summarized in table below. The minimum distance obtained (3.6dB) by the smoothed TFR givenm by input-output formulation is comparable(3.4dB) to the one obtained by the short-term least square applied to the same kind of signal [8], which would be normally the best fitted model to this kind of non-stationarity.

	det.	det.+stoch.	smoothed
EVOL.($\Sigma = 0$)	5.2dB	-	-
INPUT-OUTPUT	4.4dB	4.1dB	3.6dB
STATE-SPACE	4.5dB	4.2dB	3.9dB

References

[1] Grenier Y. "Time-dependent ARMA Modeling of Nonstationary Signals" IEEE Trans. on ASSP, vol.31, no. 4, pp.899-911, 1983.

[2] Harvey A.C. "The Estimation of Time-Varying Parameters from Panel Data". Annales de L'INSEE, no. 30-31, pp.204-226, 1978.

[3] Nicholls D.F., Quinn B.G. "Random Coefficient Autoregressive Models: An Introduction". Springer-Verlag, New York Heidelberg Berlin, 1982.

[4] Dembo A., Zeitouni O. "Maximum a Posteriori Estimation of Time-Varying ARMA Proc. from Noisy Obs.". IEEE Trans. on ASSP, vol.35m no. 4, pp.471-476, 1988.

[5] De Lima-Veiga A., "Modèles Non-stationaires à Deux Niveaux". Doctoral thesis, ENST, France, 1989.

[6] Anderson B.D.O.,Moore J.B. "Optimal Filtering". Englewood Cliffs, NJ, Prentice-hall, 1979.

[7] Schweppe F.C. "Evaluation of Likelihood Functions for Gaussian Signals". IEEE Trans. Inform. Theory, vol.11, no.1, pp.61-70, 1964.

[8] Grenier Y., Aboutajdine D., "Comparaison des Représentations Temps-fréquence de Signaux Présentant des Discontinuités Spectrales". Annales des Télécomunications, tome 38, no.11-12, pp. 429-442, 1983.

Fig. 2: The deterministic trajectories to be estimated.

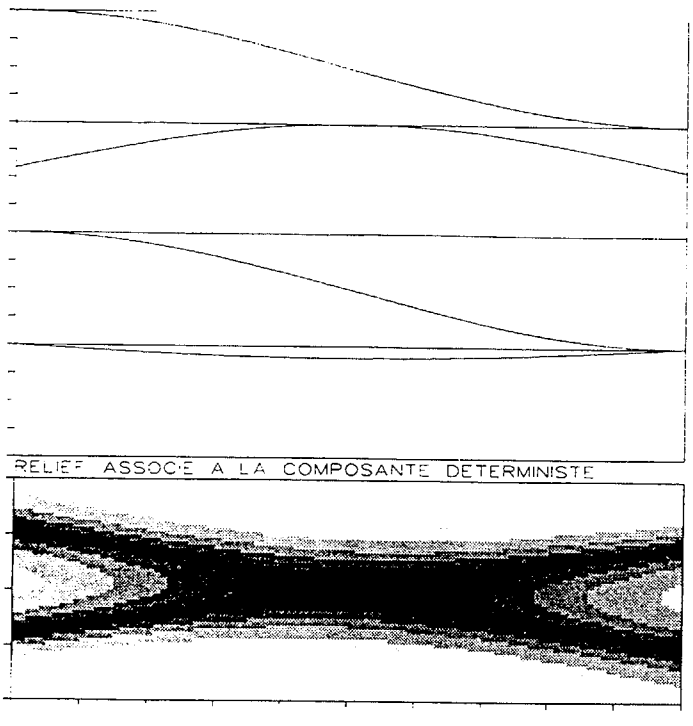

Fig. 1 : A signal produced by an evolutionary AR model with random coefficients.

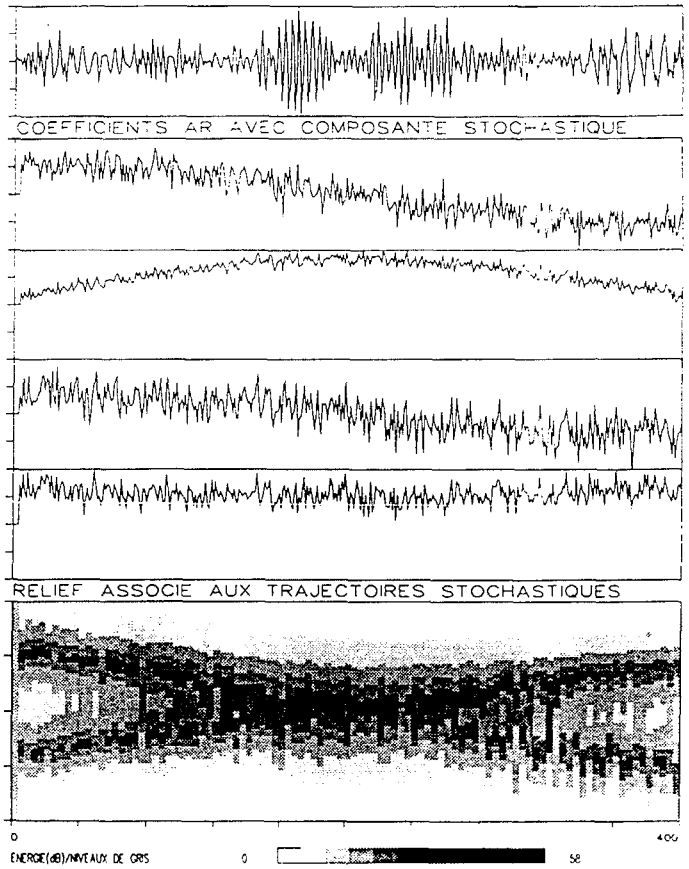

Fig. 3: Spectral distance from estimated to real time-frequence representations.

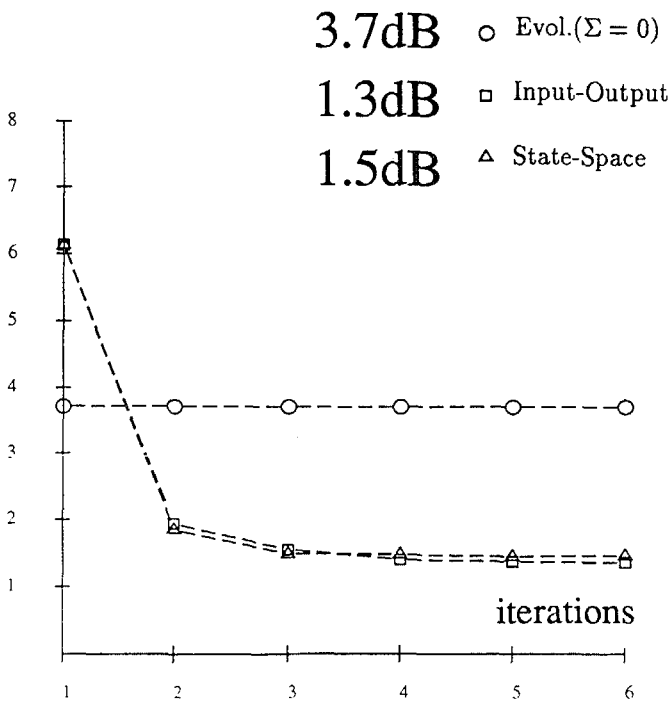

Fig. 4: Signal with spectral jumps.

Fig. 5: Estimated trajectories (determ. + stochastic).