



MODELES EVOLUTIFS A COEFFICIENTS STOCHASTIQUES

Alvaro DE LIMA-VEIGA, Yves GRENIER

Ecole Nationale des Télécommunications, Département Systèmes et Communications  
46, rue Barrault, 75634 Paris, Cedex 13, FRANCE

Resumé

Abstract

Cet article présente les éléments d'une classe de modèles non-stationnaires, nommés MODELES EVOLUTIFS A COEFFICIENTS STOCHASTIQUES, qui réunit les caractéristiques de deux classes importantes de modèles non-stationnaires: les modèles évolutifs et les modèles à coefficients stochastiques. Ce sont des modèles où la trajectoire des coefficients est décrite par un modèle qui comprend une partie déterministe, représentée par une combinaison linéaire de fonctions du temps, et une partie stochastique représentée par un modèle AR vectoriel.

This paper presents a new class of non-stationary models called STOCHASTIC COEFFICIENTS EVOLUTIONARY MODELS. This new class mixes the characteristics of two important classes of time-varying models: EVOLUTIONARY MODELS and STOCHASTIC COEFFICIENTS MODELS. Our model has the trajectory of its coefficients described by the sum of two parts: a deterministic one expressed as a linear combination of known time-functions and a nondeterministic one represented by a vectorial autoregressive model.

L'identification du modèle se fait par une méthode de maximum de vraisemblance. Nous montrons comment évaluer la fonction de vraisemblance approchée en faisant appel à la décomposition orthogonale du signal donnée par un filtre de Kalman. Une méthode de recherche non-linéaire qui fait usage des dérivées du premier ordre calculées par un filtre semblable et connecté au filtre de Kalman est proposée.

Our identification procedure is based upon the maximisation of an approximate likelihood. We show how to compute the likelihood function using the orthogonal decomposition of the signal given by the Kalman filter. The gradient of this likelihood can also be computed on-line, using filters similar to and connected to this Kalman filter.

1. INTRODUCTION

L'analyse des signaux non-stationnaires se fait le plus souvent par des méthodes dont l'origine se trouve dans la théorie des signaux stationnaires. Ces méthodes considèrent directement ou indirectement le fait que le signal est stationnaire sous une fenêtre glissante. C'est le cas, par exemple, des modèles adaptatifs. Dans ces modèles, le vecteur des paramètres est calculé à chaque instant par la correction du même vecteur à l'instant précédent, ce qui engendre un processus semblable à celui de Markov. Même si cela est fréquemment suffisant pour suivre les évolutions du signal, il reste d'autre part le problème de le représenter parcimonieusement soit pour le transmettre, le synthétiser ou le reconnaître.

Dans la section 2 nous montrons que le modèle du signal, y compris le mode d'évolution des coefficients, peut être mis sous forme d'une représentation d'état. Dans la section 3 la fonction de vraisemblance approchée est évaluée en faisant appel à la décomposition orthogonale du signal donnée par le filtre de Kalman qui s'y applique directement. A la fin de cette section nous proposons une procédure de recherche non-linéaire du type gradient à pas variable qui donne l'estimée de la partie aléatoire du modèle. L'estimation de la partie déterministe se fait comme un sous-produit de cette procédure, connectée au filtre de Kalman.

Depuis quelques années, l'intérêt s'est porté sur des modèles plus directement non-stationnaires et qui comportent dans leur structure le mode de variation du signal. Parmi ces modèles on distingue deux classes principales: d'abord, ceux appelés MODELES EVOLUTIFS où la trajectoire des coefficients s'exprime comme une combinaison linéaire des fonctions du temps connues a priori dont les poids sont à estimer; deuxièmement, les MODELES A COEFFICIENTS STOCHASTIQUES dont les coefficients sont eux mêmes des variables aléatoires. Ces derniers reprennent, dans cet aspect, l'idée d'évolution stochastique sous-jacente aux modèles adaptatifs.

2. MODELE EVOLUTIF A COEFFICIENTS STOCHASTIQUES

Soit le signal  $y_t$  mesuré sur l'intervalle  $[1, T]$  modélisé par un modèle autorégressif non-stationnaire d'ordre  $k$ .

$$y_t = x_t^T \beta_t \tag{1}$$

où  $x_t^T = [1, y_{t-1}, \dots, y_{t-k}]$  et  $\beta_t$  est un vecteur dont la première composante est le bruit excitateur, les composantes suivantes sont les coefficients AR à l'instant  $t$ . Par simplicité de notation nous avons réuni dans un seul vecteur ces deux composantes. Cela nous permettra, d'autre part, de considérer d'autres structures que le bruit blanc habituel pour l'excitation.

Cet article se propose de présenter les éléments d'une classe de modèles, appelés EVOLUTIFS A COEFFICIENTS STOCHASTIQUES, qui fait le pont entre ces deux approches: l'évolution certaine et l'évolution aléatoire des coefficients. La dynamique des coefficients du modèle que l'on propose comprend une partie déterministe évolutive (dans le sens décrit ci-dessus) et une partie stochastique représentée par un modèle AR vectoriel. Cette approche a déjà été utilisée par Swamy, Tinsley[1] dans un modèle de régression appliqué à l'économétrie. Cependant, la méthode d'estimation qu'ils proposent ne peut pas être utilisée lorsque le modèle est autorégressif à cause de la non-linéarité qui en découle. D'autre part la méthode est telle que les deux parties du modèle, i.e., déterministe et stochastique, doivent être estimées séparément, l'estimation de l'une supposant que l'autre est connue. La méthode que nous proposons est applicable aux modèles autorégressifs et estime les deux parties simultanément.

Nous avons modélisé la trajectoire de  $\beta_t$  comme la somme des deux parties: l'une déterministe, représentée par une combinaison linéaire des fonctions du temps et l'autre purement stochastique, représentée par un modèle autorégressif vectoriel d'ordre  $p$  stationnaire. Alors,

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

où  $B = \left\{ b_{ij}, i=0, \dots, k \quad j=1, \dots, m \right\}$  est une matrice à estimer, dite partie "déterministe" du modèle. Le vecteur,  $z_t^T = [z_0(t), \dots, z_{m-1}(t)]$  est une base de fonctions connues a priori, qui donne le caractère "évolutif" du modèle. La base de fonctions la plus simple est celle des puissances du temps,  $z_i(t) = \frac{t^i}{i!}$  ou son équivalent après orthogonalisation, où les fonctions  $z_i(t)$



sont des polynômes de Legendre. Une autre base est celle issue de la décomposition en série de Fourier des coefficients  $\beta_i(t)$ . A cet ensemble s'ajoute la base de fonctions sphéroïdales aplaties dont la base théorique et une méthode de calcul sont présentées par Grenier[2]-annexe 5.

La partie stochastique,  $\varepsilon_t$ , est représentée par un modèle autorégressif vectoriel,

$$\varepsilon_t = \sum_{i=1}^p \phi_i \varepsilon_{t-i} + e_t \quad (3)$$

où  $e_t$  est un vecteur aléatoire indépendant centré et de matrice de covariance constante  $\Sigma_e$ .

Le problème de l'identification du modèle est celui de l'estimation de  $B, \phi_1, \dots, \phi_p$  et  $\Sigma_e$ . C'est un problème assez délicat car fortement non-linéaire. D'une part le modèle stochastique doit être estimé sans la connaissance ni de l'entrée ni de la sortie, d'autre part, la partie évolutive suppose la connaissance du modèle stochastique pour pouvoir être calculée. Dans la suite on présente une méthode basée sur le filtre de Kalman qui rend possible le calcul de la fonction de vraisemblance du modèle uniquement en fonction des paramètres du modèle autorégressif vectoriel. L'estimation de la partie déterministe (B) se fera alors comme un sous-produit de l'application du filtre de Kalman.

Dans ce qui suit  $A \otimes B$  désignera le produit de Kronecker de deux matrices A et B, tandis que l'opérateur  $\text{vec}()$  transforme une matrice en vecteur colonne en rangeant les colonnes de cette matrice les unes au-dessus des autres. De même, l'opérateur  $\text{vech}()$  transforme en vecteur colonne une matrice symétrique en ne prenant dans chacune des colonnes que les éléments de la partie triangulaire gauche de cette matrice.

### 2.1. Représentation d'état pour le Modèle Evolutif à Coefficients Stochastiques

Afin de pouvoir calculer la vraisemblance du modèle défini par les équations(1-3) il est utile de le réécrire sous forme d'une équation d'état. Soient les matrices  $\Phi$  et  $J$  définies par:

$$\Phi = \begin{bmatrix} \phi_1 \phi_2 \dots \phi_p \\ I & 0 & \dots & 0 \\ 0 & I & \dots & 0 \\ \dots & \dots & \dots & 0 \end{bmatrix} \quad \text{et} \quad J = \begin{bmatrix} I \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

Alors l'équation (2) prend la forme:

$$\beta_t = Bz_t + J^T \xi_t$$

où le vecteur  $\xi_t$  est issu de la réécriture de l'équation (3):

$$\xi_t = [\varepsilon_t^T, \varepsilon_{t-1}^T, \dots, \varepsilon_{t-p}^T]^T = \Phi \xi_{t-1} + J e_t$$

Les paramètres du modèle à l'instant t sont contenus dans le vecteur  $\xi_t$  et la matrice B. Définissons alors le vecteur  $\alpha_t$  contenant ces paramètres:

$$\alpha_t = [b^T, \xi_t^T]^T, \quad b = \text{vec}(B)$$

Le modèle s'écrit maintenant sous forme de représentation d'état,

$$\alpha_t = F \alpha_{t-1} + G e_t$$

$$y_t = H_t^T \alpha_t$$

avec

$$F = \begin{bmatrix} I_{m(k+1)} & 0 \\ 0 & \Phi \end{bmatrix} \quad \text{et} \quad G = \begin{bmatrix} 0_{m(k+1), (k+1)} \\ J \end{bmatrix}$$

$$H_t^T = [(z_t^T \otimes x_t^T): x_t^T J]$$

### 2.2. Calcul de la Fonction de Vraisemblance

Le calcul de la vraisemblance  $\text{Log } p(y_1, \dots, y_T / \omega)$  se fait de manière récursive sur T à partir d'un filtrage de Kalman des observations  $y_t$  pour les valeurs  $\omega$  des paramètres du modèle (ici  $\phi_1, \dots, \phi_p, \Sigma_e$ ). Etant données les mesures  $y_t$ ,  $t=1, T$  et les conditions initiales  $x_1, \alpha_0$  et  $P_{0/0}$ , on peut construire le filtre de Kalman correspondant aux équations suivantes:

$$\begin{aligned} \alpha_{t/t-1} &= F \alpha_{t-1/t-1} \\ P_{t/t-1} &= F P_{t-1/t-1} F^T + G \Sigma_e G^T \\ \alpha_{t/t} &= \alpha_{t/t-1} + K_t \eta_t \\ \eta_t &= y_t - H_t^T \alpha_{t/t-1} \\ K_t &= P_{t/t-1} H_t^{-1} \\ h_t &= H_t^T P_{t/t-1} H_t \\ P_{t/t} &= (I - K_t H_t^T) P_{t/t-1} \end{aligned} \quad (4)$$

La représentation d'état, comme SCHWEPPE[3] l'a montré est une façon convenable, car itérative, d'obtenir une décomposition orthogonale de  $y_1, \dots, y_T$ . En effet, sous l'hypothèse gaussienne, les innovations  $\eta_t$  sont des variables aléatoires normales centrées et de variance  $h_t$ . Soit  $\omega^T = [(\text{vec } \phi_1)^T, \dots, (\text{vec } \phi_p)^T, (\text{vec } \Sigma_e)^T]$ . Alors, le logarithme de la vraisemblance s'écrit,

$$L(\omega / Y_0) = -\frac{T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log h_t - \frac{1}{2} \sum_{t=1}^T \eta_t^2 h_t^{-1} \quad (5)$$

### 3. METHODE DU MAXIMUM DE VRAISEMBLANCE

Nous souhaitons maximiser la vraisemblance  $L(\omega / x_1)$  relativement au vecteur  $\omega$  des inconnues. Des méthodes de type Newton-Raphson nécessitant le calcul du gradient et du Hessien du critère à maximiser sont inutilisables dans notre cas, car le calcul du Hessien est trop coûteux. Seules restent accessibles des méthodes reposant uniquement sur le gradient. Soit  $g(\omega)$  le vecteur du gradient de L par rapport à  $\omega$ :

$$g(\omega) \equiv \nabla L(\omega) = \left\{ \frac{\partial L}{\partial \omega_i}, i = 1, (k+1)^2 p + (k+1)k/2 \right\} \quad (6)$$

$$\frac{\partial L}{\partial \omega_i} = \sum_{t=1}^T \frac{\partial h_t}{\partial \omega_i} \frac{1}{h_t} + \sum_{t=1}^T \left[ 2\eta_t \frac{\partial \eta_t}{\partial \omega_i} h_t - \eta_t^2 \frac{\partial h_t}{\partial \omega_i} h_t^{-2} \right]$$

Les valeurs de  $h_t$  et  $\eta_t$  sont calculées par le filtre de Kalman (eq. 4) et ses dérivées par rapport à  $\omega_i$ , notées  $h_t^i$  et  $\eta_t^i$  respectivement, sont calculées par la récursion suivante:

$$\begin{aligned} \alpha_{t/t-1}^i &= F \alpha_{t-1/t-1}^i + F^i \alpha_{t-1/t-1} \\ P_{t/t-1}^i &= F^i P_{t-1/t-1} F^T + F P_{t-1/t-1}^i F^T + F P_{t-1/t-1} F^{iT} \end{aligned} \quad (7a)$$

pour les  $(k+1)^2 p$  premières composantes de  $g(\omega)$ , c'est à dire, celles correspondant aux  $\phi_i$ , et,

$$\begin{aligned} \alpha_{t/t-1}^i &= F \alpha_{t-1/t-1}^i \\ P_{t/t-1}^i &= F^i P_{t-1/t-1} F^T + G I_i G^T \end{aligned} \quad (7b)$$

pour les composantes restantes (celles concernant  $\text{vech}(\Sigma_e)$ ).

Dans les deux cas, la récursion se complète par :

$$\begin{aligned} \alpha_{t/t}^i &= \alpha_{t/t-1}^i + K_t^i \eta_t^i + K_t \eta_t^i \\ \eta_t^i &= -H_t^T \alpha_{t/t-1}^i \\ K_t^i &= P_{t/t-1}^i H_t^{-1} h_t^{-1} - P_{t/t-1} H_t h_t^{-2} h_t^i \\ h_t^i &= H_t^T P_{t/t-1}^i H_t \\ P_{t/t}^i &= -K_t^i H_t^T P_{t/t-1} + (I - K_t H_t^T) P_{t/t-1} \end{aligned} \quad (7c)$$

où  $I_i$  est une matrice avec un seul élément non-nul égal à 1 qui serait la  $i$ -ème composante de  $\text{vec}(I_i)$  et  $F^i$  est une matrice de zéros semblable à F avec un seul élément non-nul égal à 1, à la place de l'élément qui correspond à  $\omega_i$ .

3.1. Recherche unidirectionnelle

Plutôt que d'utiliser une méthode de gradient avec un pas prédéterminé, il est possible de rechercher dans la direction du gradient le déplacement optimal. Etant donné un point de départ  $\omega^k$  et une direction de recherche  $g(\omega^k)$  on cherche alors le réel  $d$  positif tel que,

$$L(\omega^k + dg(\omega^k)) = \max_d L(\omega^k + dg(\omega^k))$$

On pose :

$$F^d = F|_{\omega=\omega^k} + dF_g \text{ et } \Sigma_e^d = \Sigma_e|_{\omega=\omega^k} + d\Sigma_e^g$$

où  $F_g$  est une matrice de la même forme que  $F$  avec les éléments  $\phi_i$  remplacés par les valeurs correspondant extraites de  $g_k$  et les matrices identités remplacées par 0. La matrice  $\Sigma_e^g$  est telle que  $\text{vec} \Sigma_e^g = [g_{(k+1)2p+1}, \dots, g_{(k+1)2p+k(k+1)/2}]^T$ . Donc,

$$\frac{\partial F^d}{\partial d} = F_g \text{ et } \frac{\partial \Sigma_e^d}{\partial d} = \Sigma_e^g$$

On construit assez facilement une récursion qui nous donne les valeurs des dérivées de  $L(\omega^k + dg^k)$  par rapport à  $d$ .

$$\frac{\partial L}{\partial d}(\omega^k + dg^k) = \sum_{t=1}^T \frac{\dot{h}_t}{h_t} + \sum_{t=1}^T [2\eta_t \dot{h}_t h_t - \eta_t^2 \dot{h}_t^2 h_t] \quad (8)$$

où  $\eta_t$  et  $h_t$  sont calculés par le filtre de Kalman (eq. 4) avec  $F$  et  $\Sigma_e$  remplacés respectivement par  $F^d$  et  $\Sigma_e^d$ , les dérivées de  $\eta_t$  et  $h_t$  par rapport à  $d$ , étant notées,

$$\frac{\partial \eta_t}{\partial d} = \dot{\eta}_t \text{ et } \frac{\partial h_t}{\partial d} = \dot{h}_t$$

Ces valeurs sont données par la récursion ci-dessous:

$$\begin{aligned} \dot{\alpha}_{t/\tau-1} &= F_g \alpha_{t-1/\tau-1} + (F-dF_g) \dot{\alpha}_{t-1/\tau-1} \\ \dot{P}_{t/\tau-1} &= F_g P_{t-1/\tau-1} (F-dF_g)^T + (F-dF_g) \dot{P}_{t-1/\tau-1} (F-dF_g)^T \\ &\quad + (F-dF_g) P_{t-1/\tau-1} \dot{F}_g^T + G \Sigma_e^g G^T \\ \dot{\alpha}_{t/\tau} &= \dot{\alpha}_{t/\tau-1} + K_t \dot{\eta}_t + K_t \dot{h}_t \quad (9) \\ \dot{\eta}_t &= -H_t^T \dot{\alpha}_{t/\tau-1} \\ \dot{K}_t &= \dot{P}_{t/\tau-1} H_t h_t^{-1} - P_{t/\tau-1} H_t h_t^{-2} \dot{h}_t \\ \dot{h}_t &= H_t^T \dot{P}_{t/\tau-1} H_t \\ \dot{P}_{t/\tau} &= K_t H_t^T P_{t/\tau-1} + (I - K_t H_t^T) \dot{P}_{t/\tau-1} \end{aligned}$$

La méthode présente donc la structure usuelle des méthodes de recherche non-linéaire. On fixe un point de départ  $\omega^0$ , des valeurs initiales du jeu de paramètres de  $[\phi_1, \dots, \phi_p, \Sigma_e]$ , qui sera recalculé à chaque tour de l'algorithme.

Etant donné un point  $\omega^k$  on calcule la vraisemblance par (5) en utilisant le filtre de Kalman (eq. 6). Chaque fois que l'on calcule ce filtre on obtient une estimée de  $b$ , donnée par les  $(k+1)m$  premières composantes de  $\alpha_{T/\tau}$ . Si la valeur de la vraisemblance a augmenté on calcule la direction de recherche du nouveau point, donnée par le gradient de  $L$  par rapport à  $\omega^k$  (6,7). Un nouveau point  $\omega^{k+1}$  sera calculé alors par un déplacement  $d$  sur la direction du gradient, c'est à dire;

$$\omega^{k+1} = \omega^k + dg(\omega^k)$$

où  $d$  est calculé par une procédure qui tient compte de la valeur de la dérivée de  $L$  sur la direction de recherche (8-9). On souligne que la fonction  $L$  peut avoir plusieurs maxima et qu'on ne peut pas garantir que le maximum global sera atteint, bien que le cas se soit présenté lors de la plus part de nos essais.

4. VARIANTE DE LA METHODE POUR UN CAS SIMPLE.

Dans les simulations que nous présentons à la fin de l'article, nous utilisons un modèle simplifié où la partie stochastique est représentée comme un bruit blanc vectoriel gaussien. Dans ce cas il est possible d'estimer la partie déterministe par moindres carrés. La fonction de vraisemblance du modèle peut être calculée directement en fonction de  $b$  et  $\Sigma_e$ . Dans la suite nous présentons une méthode d'estimation valable seulement dans ce cas simple, où l'estimée par moindres carrés de  $b$  est utilisée comme solution initiale d'une méthode de maximisation de la vraisemblance, du type gradient.

Soit  $u_t$  l'erreur de prédiction.

$$u_t = y_t - (z_t^T \theta x_t^T) b = x_t^T \varepsilon_t$$

Dans le cas présent,  $\varepsilon_t$  est indépendante de  $x_t^T$  et, par conséquent,  $u_t$  est une variable aléatoire indépendante et centrée de variance  $\sigma_u^2 = x_t^T \Sigma_e x_t$ . L'estimateur  $\hat{b}$  de  $b$  au sens de moindres carrés s'écrit:

$$\hat{b} = \left[ \sum_{t=1}^T (z_t \theta x_t) (z_t^T \theta x_t^T) \right]^{-1} \sum_{t=1}^T (z_t \theta x_t) y_t \quad (10)$$

4.1. Maximisation de la Fonction de Vraisemblance

Sous l'hypothèse gaussienne, le logarithme de la fonction de vraisemblance se calcule, à une constante près,

$$L(b, \Sigma_e) = -\frac{1}{T} \left[ \sum_{t=1}^T \log \sigma_u + \sum_{t=1}^T \frac{u_t^2}{\sigma_u} \right] \quad (11)$$

Dans la méthode présentée ici, nous avons choisi de représenter les paramètres non pas par  $b$  et  $\Sigma_e$  mais par  $b$  et  $D$ , cette dernière matrice, d'éléments  $d_{ij}$ , désignant le facteur gauche de la décomposition de Cholesky de  $\Sigma_e$ . Cela est nécessaire à cause de la difficulté que l'on aurait à définir de manière simple l'espace de recherche pour  $\Sigma_e$  qui doit être une matrice non-négative définie. On calcule les dérivées de (11) par rapport à  $b$  et les éléments non-nuls de  $D$ .

$$b_g = \nabla_b L(b, D) = -\frac{2}{T} \sum_{t=1}^T \frac{(z_t \theta x_t) \eta_t}{(x_t^T \theta x_t^T) s} \quad (12)$$

$$\text{vech}(D_g) = \nabla_{\text{vech}(D)} L(b, D) = -\frac{1}{T} \sum_{t=1}^T \left[ \frac{\dot{\sigma}_u}{\sigma_u} - \frac{\eta_t^2 \dot{\sigma}_u}{\sigma_u^2} \right] \quad (13)$$

où

$$\dot{\sigma}_u^T = \left[ \frac{\partial \sigma_u}{\partial d_{11}}, \dots, \frac{\partial \sigma_u}{\partial d_{mm}} \right] \text{ et } \frac{\partial \sigma_u}{\partial d_{ij}} = 2x_{it} \sum_{k=j}^{p+1} d_{kj} x_{kt}$$

Comme dans la section 3.1 nous avons besoin de calculer la dérivée de la fonction de vraisemblance sur la direction du gradient. D'abord on l'écrit en fonction du point de départ  $b$  et  $D$ , du gradient sur ce point et du déplacement  $d$ .

$$L(d) = \frac{1}{T} \sum_{t=1}^T \left[ \log(A_t d^2 - B_t d + C_t) + \frac{(\eta_t + (z_t^T \theta x_t^T) b_g d)^2}{(A_t d^2 - B_t d + C_t)} \right] \quad (14)$$

avec,

$$A_t = x_t^T D D_g D_g^T x_t$$

$$B_t = x_t^T (D D_g^T + D_g D^T) x_t$$

$$C_t = x_t^T D D^T x_t$$



La dérivée par rapport à  $d$  s'écrit alors :

$$\dot{L}(d) = \frac{\partial L(d)}{\partial d} = \frac{1}{T} \sum_{t=1}^T \left[ \frac{2dA_t - B_t}{(A_t d^2 - B_t d + C_t)} \right] \quad (15)$$

$$+ \frac{1}{T} \sum_{t=1}^T \left[ \frac{2(\eta_t + (z_t^T \theta x_t^T) b_g d)(z_t^T \theta x_t^T) b_g d}{(A_t d^2 - B_t d + C_t)} \right]$$

$$- \frac{1}{T} \sum_{t=1}^T \left[ \frac{(2dA_t - B_t)(\eta_t + (z_t^T \theta x_t^T) b_g d)^2}{(A_t d^2 - B_t d + C_t)^2} \right]$$

La recherche non-linéaire se fait alors de la même façon que précédemment dans la section 3. On calcule une estimée initiale de  $b$  par (10) qui intégrera le vecteur de départ  $\omega^0$ . La vraisemblance se calcule directement par l'équation (11). La direction de recherche se calcule par (12) et (13) tandis que le déplacement est calculé par une procédure de recherche unidirectionnelle identique à celle utilisée dans la section 3.

### 5. SIMULATIONS

Nous avons synthétisé un signal évolutif à coefficients stochastiques de longueur  $T=256$ , ordre  $k=4$ . La base utilisée est celle de Fourier  $(1, \cos \frac{\pi t}{T}, \sin \frac{\pi t}{T})$  et la partie stochastique est un bruit blanc vectoriel.

Pour chaque réalisation,  $B$  et  $\Sigma_e$  ont été estimés en faisant varier l'ordre  $k$  et la dimension  $m$  autour des valeurs utilisées dans la synthèse. Ce sont :

$$(k, m) = \left\{ (4, 3), (4, 2), (4, 4), (2, 3), (6, 3) \right\}.$$

Les modèles calculés à chaque itération (jusqu'à 4 au maximum) pour chaque méthode et pour chaque valeur du doublet  $(k, m)$  ont été utilisés pour calculer le relief (succession des spectres instantanés) associé et la distance au relief de référence, celui-ci étant associé au modèle théorique du signal synthétisé. Nous avons testé trois méthodes : La méthode basée sur le filtre de Kalman (FK), celle à deux étapes de moindres carrés (MC) et une version sous optimale de FK, que l'on appellera FS, dans laquelle les matrices de covariance de l'erreur d'estimation du filtre de Kalman utilisé en FK ont été considérées comme étant bande-diagonale de largeur de bande égale à  $m$ , la dimension de la base de fonctions.

la figure 1 montre aussi le résultat dû à l'application d'une méthode d'estimation d'un modèle évolutif ordinaire, décrit en [2].

#### 5.1. Résultats.

Malgré le caractère limité de cette expérience il est possible d'en tirer un certain nombre de conclusions. La première est celle concernant l'expérience où  $k$  et  $m$  ont les mêmes valeurs que dans le modèle théorique. Dans ce cas, FK s'est montrée la meilleure, suivie de près par MC et plus loin par FS. Le résultat inférieur obtenu par la méthode évolutive ordinaire (EVOL) ne surprend point car les variations aléatoires des coefficients ne sont pas comprises dans sa structure. Il est intéressant de remarquer que les méthodes MC et FS ont eu des performances tout à fait comparables pour les trois premières itérations dans presque toutes les expériences. Dans certaines il y a même eu un léger avantage pour la méthode sous-optimale lors de la première itération.

On constate ensuite, sans surprise, que les méthodes sont plus robustes vis-à-vis d'un sur-dimensionnement (modèles (4,4) et (6,3)) que vis-à-vis d'un sous-dimensionnement (modèles (4,2) et (2,3)) où la méthode sous-optimale FS est la plus sensible. Les deux autres méthodes ont gardé pratiquement le même comportement dans le cas sur-dimensionné, présentant sur tout une convergence plus lente.

Finalement, la figure 1 nous montre que les méthodes basées sur le filtre de Kalman arrivent assez près du maximum en deux itérations tandis que la méthode MC semble avoir une convergence presque linéaire. Cela peut être dû au fait que la fonction de vraisemblance utilisée dans cette dernière méthode ne prend pas en compte

l'incertitude que l'on a sur la valeur de  $b$  et, par conséquent, est moins riche en information que celle utilisée dans les méthodes basées sur le filtre de Kalman.

### 6. CONCLUSIONS

Cet article a présenté les éléments d'une classe de modèles non-stationnaires, nommés modèles évolutifs à coefficients stochastiques, qui réunit les caractéristiques de deux classes importantes de modèles non-stationnaires: les modèles évolutifs et les modèles à coefficients stochastiques.

Le choix d'une forme particulière de représentation d'état nous permet de tourner la non-linéarité du modèle. Nous avons étudié l'application d'une méthode basée sur le filtre de Kalman qui rend possible l'écriture de la fonction de vraisemblance uniquement en fonction de la partie stochastique du modèle. Une autre approche a été appliquée à un cas particulier plus simple de ces modèles. Une série d'expériences avec des signaux synthétisés nous a donné quelques indications sur le comportement de ces méthodes. Les méthodes basées sur le filtre de Kalman se sont montrées plus performantes du point de vue estimation spectrale mais sont plus coûteuses. D'autre part, toutes les méthodes proposées ont été supérieures à la méthode évolutive ordinaire pour ce qui est de reproduire les caractéristiques spectrales des signaux testés.

k	m	i	MC	FK	FS	EVOL
4	3	1	7.382	5.519	5.928	6.239
		2	5.542	2.900	5.074	
		3	4.197	2.670	4.715	
		4	3.465	"	"	
4	2	1	8.717	7.233	6.291	
		2	7.049	5.374	"	
		3	6.131	5.292	"	
		4	5.766	5.255	"	
4	4	1	7.784	5.535	6.156	
		2	6.210	3.046	4.741	
		3	5.022	2.861	4.733	
		4	4.312	"	"	
2	3	1	6.375	6.533	8.278	
		2	"	"	7.704	
		3	"	"	7.629	
		4	"	"	7.593	
6	3	1	9.256	7.907	7.693	
		2	6.511	4.404	7.079	
		3	4.378	3.876	7.066	
		4	3.923	3.867	"	

Fig. 9.1 : - Distance spectrale au modèle théorique, moyenne sur 21 réalisations du signal.

### 7. REFERENCES BIBLIOGRAPHIQUES

- [1] SWAMY P.A.V.B., TINSLEY P.A. "Linear Prediction and Estimation Methods for Regressions Models With Stationary Stochastic Coefficients." JOURNAL OF ECONOMETRICS 12(1980) 103-142.
- [2] GRENIER Y. "Modélisation de Signaux Non-stationnaires." Thèse de Doctorat d'Etat, Paris-Sud, 1984
- [3] SCHWEPPE F.C. "Evaluation of likelihood functions for Gaussian signals," IEEE Trans. Inform. Theory, vol IT-11, no. 1, pp 61-70, Jan. 1965.