NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

ESTIMATION RECURSIVE EN TRANSMISSION NUMERIQUE

H. KOREZLIOGLU [(1)] et G. MAZZIOTTO [(2)]

(1) E.N.S.T., Dpt. Systèmes et Communications
46, rue Barrault
75 634 - PARIS CEDEX 13

(2) C.N.E.T., PAA/TIM/MTI
38-40, rue du Général Leclerc
92 131 - ISSY LES MOULINEAUX

RESUME

On développe une technique de filtrage non-linéaire récursif pour un modèle où les processus considérés prennent leurs valeurs dans des espaces discrets. Le résultat obtenu a une application immédiate à l'estimation et à la détection d'une classe de signaux codés markoviens, transmis à travers un canal symétrique sans mémoire.

E et F représentent deux espaces discrets, $B = \{B_n ; n \in \mathbb{N}\}$ et $W = \{W_n ; n \in \mathbb{N}\}$ deux suites de variables aléatoires à valeurs respectivement dans E et F, appelés bruits d'état et d'observation, on définit un signal X et une observation Y par les équations suivantes : $X_{n+1} = f_n(X_n, B_n)$ et $Y_n = h_n(X_n, W_n)$ pour $n \geq 0$, où $f_n : E \times E \to E$ et $h_n : E \times F \to F$. On suppose que le couple $(B,W) = \{(B_n, W_n) ; n \in \mathbb{N}\}$ est une suite indépendante à valeurs dans $E \times F$ et indépendante de X_0 et que, pour tout x, l'application $h_n(x,.) : w \to h_n(x,w)$ est bijective.

En utilisant la méthode de la probabilité de référence, obtient une formule récursive exacte pour le calcul de la loi de probabilité conditionnelle de X_n connaissant $Y_0, \ldots, Y_n$. On considère l'extension de la méthode aux divers cas rencontrés dans les applications, en particulier à celui d'un canal de mémoire finie. Dans ce cas, le résultat obtenu suggère une méthode de détection similaire à celle qui aboutit à l'algorithme de Viterbi. Finalement, on étend la méthode aux signaux à valeurs continues et on en déduit un algorithme d'estimation récursive lorsque l'observation n'est connue que par ses valeurs quantifiées.

SUMMARY

We develop a recursive nonlinear filtering technique for a model in which processes take their values in a discrete space. The obtained result has an immediate application to the estimation and to the detection of a class of Markovian coded signals, transmitted through a symmetric memoryless channel.

Given two discrete spaces E and F and two sequences $B = \{B_n ; n \in \mathbb{N}\}$ and $W = \{W_n ; n \in \mathbb{N}\}$ of random variables with values in E and F, called state and observation noises, we define a signal X and an observation Y by the following equations : $X_{n+1} = f_n(X_n, B_n)$ and $Y_n = h_n(X_n, W_n)$, for $n \geq 0$, where $f_n : E \times E \to E$ and $h_n : E \times F \to F$. We suppose that the pair $(B,W) = \{(B_n, W_n) ; n \in \mathbb{N}\}$ is an independent sequence with values in $E \times F$ and independent of X_0 and that, for all x, the mapping $h_n(x,.) : w \to h_n(x,w)$ is one-to-one.

By using the reference probability method we derive an exact recursive formula for the computation of the conditional probability distribution of X_n given $Y_0, \ldots, Y_n$. We consider the extension of the method to various cases encountered in applications, in particular to the case of a channel with finite memory. The corresponding result suggests a method of detection analog to the one leading to the Viterbi algorithm. Finally we extend the method to signals with continuous values and deduce a recursive estimation algorithm when the observation is known only by its quantized values.

ESTIMATION RECURSIVE EN TRANSMISSION NUMERIQUE

H. KOREZLIOGLU - G. MAZZIOTTO

1 - LE MODELE SOUS LA PROBABILITE DE REFERENCE

On utilise ici la méthode de la probabilité de référence qui a permis d'obtenir des résultats importants dans la théorie du filtrage non-linéaire. Nous nous référons à ([13]) pour l'origine de la méthode et à ([3]) et ([10]) pour quelques unes de ses extensions.

On considère deux espaces discrets E et F et on se donne : une v.a. (variable aléatoire) X_0 à valeurs dans E, une suite de v.a. $B = \{B_n ; n \in \mathbb{N}\}$ à valeurs dans E, une suite de v.a. $Y = \{Y_n ; n \in \mathbb{N}\}$ à valeurs dans F. On suppose que la suite $(B,Y) = \{(B_n,Y_n) ; n \in \mathbb{N}\}$ à valeurs dans $E \times F$, est une suite de v.a. mutuellement indépendantes et indépendantes de X_0. On désigne par Ω l'espace $E \times (E \times F)^{\mathbb{N}}$ des suites $x_0, (b_0,y_0), \dots, (b_n,y_n), \dots$ où $x_0 \in E$ et $(b_n,y_n) \in E \times F$, et par $\underline{\underline{A}}$ la plus petite tribu des parties de Ω rendant mesurables les applications coordonnées de Ω. $\mathbb{P}_0$ est alors la probabilité induite sur $\underline{\underline{A}}$, par la suite $X_0, (B_0,Y_0), \dots, (B_n,Y_n), \dots$. En résumé, $(\Omega, \underline{\underline{A}}, \mathbb{P}_0)$ désigne l'espace canonique de cette suite. Nous supposons que toutes les v.a. considérées ici sont définies sur cet espace qui est notre espace de probabilité de référence.

On pose $\rho(x) = \mathbb{P}_0\{X_0 = x\}$, $r_n(y) = \mathbb{P}_0\{Y_n = y\}$

$p_n(b,y) = \mathbb{P}_0\{B_n = b,\ Y_n = y\}$, $q_n(b) = \mathbb{P}_0\{B_n = b\}$, $q_n(b/y) = \mathbb{P}_0\{B_n = b \,/\, Y_n = y\}$. On suppose que les p_n sont strictement positives.

On considère un signal ou un processus d'état défini récursivement par :

$$X_{n+1} = f_n(X_n\ ,\ B_n) \quad ; \ n \in \mathbb{N} \tag{1.1}$$

où f_n est une application de $E \times E$ dans E. Le processus $X = \{X_n ; n \in \mathbb{N}\}$, de valeur initiale X_0, ainsi engendré est une chaîne de Markov à valeurs dans E.

On considère une suite de v.a. $W = \{W_n ; n \in \mathbb{N}\}$ définies comme suit :

$$W_n = w_n(X_n\ ,\ Y_n) \quad ; \ n \in \mathbb{N} \tag{1.2}$$

où w_n est une application de $E \times F$ dans F telle que, pour tout $x \in E$, l'application $y \to w_n(x,y)$ est bijective. Par conséquent, $\forall\, n \in \mathbb{N}$, il existe une application h_n de $E \times F$ dans F telle que

$$Y_n = h_n(X_n\ ,\ W_n) \quad . \tag{1.3}$$

On définit une suite $L = \{L_n ; n \in \mathbb{N}\}$ par :

$$L_n = \prod_{i=0}^{n} \frac{p_i(B_i\ ,\ w_i(X_i,Y_i)\ )}{p_i(B_i\ ,\ Y_i)} \tag{1.4}$$

Comme X_n est une fonction de $X_0, B_0, \dots, B_{n-1}$, alors L_n est une fonction de $X_0, (B_0,Y_0), \dots, (B_n,Y_n)$, que nous désignons aussi par $L_n(X_0,B_0,\dots,B_n,Y_0,\dots,Y_n)$.

On désigne par E_0 l'espérance mathématique sous $\mathbb{P}_0$.

La Proposition suivante établit le fait que L_n est un rapport de vraisemblance, i.e. une densité de probabilité.

2 - PROPOSITION: $E_0(L_n) = 1 \quad \forall\, n \in \mathbb{N}$

Démonstration: Désignons par b^n et y^n respectivement les suites finies $(b_0,b_1,\dots,b_n)$ et $(y_0,y_1,\dots,y_n)$. On a

$$E_0(L_n) = \sum_{x_0,b^n,y^n} \rho(x_0) \prod_{i=1}^{n} p_i(b_i,w_i(x_i,y_i))$$
$$= \sum_{x_0,b^n,u^n} \rho(x_0) \prod_{i=1}^{n} p_i(b_i,u_i) = 1 \quad .$$

Car, pour x_0, b^n fixés, $w_i(x_i,y_i) = u_i$ prend l'ensemble des valeurs prises par Y_i lorsque Y_i parcourt l'ensemble de ses valeurs possibles $y_i \in F$.

La suite L permet la définition d'une nouvelle loi de probabilité sur $(\Omega,\underline{\underline{A}})$.

3 - COROLLAIRE: *La quantité*

$$L_n(x_0,b_0,\dots,b_n,y_0,\dots,y_n)\ \rho(x_0) \prod_{i=0}^{n} p_i(b_i,y_i) \tag{3.1}$$

exprime une nouvelle probabilité marginale pour l'événement $A = \{X_0=x_0, (B_0,Y_0)=(b_0,y_0), \dots, (B_n,Y_n)=(b_n,y_n)\}$. *Ces probabilités marginales induisent une mesure de probabilité* $\mathbb{P}$ *sur* $(\Omega,\underline{\underline{A}})$ *telle que la probabilité de l'événement A sous* $\mathbb{P}$ *est donnée par (3.1).*

La démonstration est une application élémentaire du théorème de Kolmogorov, ([9]).

4 - PROPOSITION: *Sous la loi* $\mathbb{P}$, *la suite :* X_0, (B_0,W_0), $\dots$, (B_n,W_n), $\dots$ *est une suite indépendante telle que* $\mathbb{P}\{X_0=x\} = \rho(x)$ *et* $\mathbb{P}\{B_n=b\ ,\ W_n=w\} = p_n(b,w)$.

Démonstration: Sous $\mathbb{P}$, la probabilité de l'événement $\{X_0=x_0, B_0=b_0, \dots, B_n=b_n, W_0=w_0, \dots, W_n=w_n\}$ est

$$\sum_{y^n} \rho(x_0) \prod_{i=0}^{n} p_i(b_i,w_i(x_i,y_i))$$

avec la contrainte $w_i = w_i(x_i(x_0,b_0,\dots,b_{i-1}),y_i)$.

Les quantités x_0, b_0, $\dots$, b_n, w_0, $\dots$, w_n étant fixées, il n'y a qu'une seule valeur possible de $(y_0,\dots,y_n)$. Donc la somme ci-dessus est égale à

$$\rho(x_0) \prod_{i=0}^{n} p_i(b_i,w_i) \quad .$$ D'où l'énoncé.

5 - LE MODELE PHYSIQUE

Sous la probabilité $\mathbb{P}$, le modèle que nous avons décrit ci-dessus coincide bien avec le modèle décrit dans le résumé. En conclusion, étant donné la suite indépendante X_0, (B_0,W_0), ..., (B_n,W_n), ... telle que

$\mathbb{P}\{X_0=x\} = \rho(x)$, $\mathbb{P}\{B_n=b, W_n=w\} = p_n(b,w)$,

$\mathbb{P}\{W_n=w\} = r_n(w)$, $\mathbb{P}\{B_n=b \;/\; W_n=w\} = q_n(b/w)$,

$\mathbb{P}\{B_n=b\} = q_n(b)$, les processus d'état et d'observation sont respectivement définis par les équations (1.1) et (1.3).

En ce qui concerne la dépendance entre B et W et les modes de génération de X et de Y, ce modèle physique est semblable au modèle général discret de Kalman, (cf. par exemple ([2])) dans lequel les v.a. sont à valeurs euclidiennes. Une différence essentielle entre le modèle de Kalman et le modèle considéré ici réside dans le fait que la notion d'espérance conditionnelle pour les v.a. à valeurs discrètes, n'a pas de sens. Par conséquent, nous nous attachons ici à évaluer la loi de probabilité conditionnelle des v.a. que l'on cherche à estimer. Alors l'estimation proprement dite ne sera plus qu'une affaire de décision. Comme nous allons le voir, les éléments de calcul sont aussi étroitement liés à la détection.

Dans ce qui suit $E_0(. \;/\; Y^n)$ et $E(. \;/\; Y^n)$ désignent respectivement l'espérance conditionnelle sous $\mathbb{P}_0$ et $\mathbb{P}$.

Le lemme suivant, dont la démonstration est basée sur la définition de l'espérance conditionnelle, donne la version de la formule de Kallianpur - Striebel ([8]) qui correspond à notre modèle.

7 - <u>LEMME</u>: *Soit une v.a.*

(7.1) $\phi = \phi(X_0, B_0, \ldots, B_n, Y_0, \ldots, Y_n)$. *Alors on a*

$$(7.2) \qquad E(\phi \;/\; Y^n) = \frac{E_0(\phi \; L_n \;/\; Y^n)}{E_0(L_n \;/\; Y^n)} \quad .$$

Par ailleurs nous avons la formulation suivante de $E_0(\phi \;/\; Y^n)$.

8 - <u>LEMME</u>: *Pour la v.a. définie par (7.1), on a*

$$(8.1) \qquad E_0(\phi \;/\; Y^n) = \sum_{x_0, b^n} \phi(x_0, b_0, \ldots, b_n, y_0, \ldots, y_n) \; \rho(x_0) \prod_{i=0}^{n} q_i(b_i/y_i)$$

Ces deux lemmes nous permettrons d'obtenir l'algorithme d'estimation pour notre modèle.

9 - Soit $\{ \Pi_{n/n}(x_n/y^n) \;;\; x_n \in E\}$ la loi de probabilité conditionnelle de X_n étant donnée $Y_0 = y_0, \ldots, Y_n = y_n$ c'est à dire $Y^n = y^n$. On a

$$(9.1) \qquad \Pi_{n/n}(x_n/y^n) = \frac{E_0(L_n \; 1\!\!1_{\{x_n\}}(X_n) \;/\; Y^n=y^n)}{E_0(L_n \;/\; Y^n=y^n)}$$

Posons

$$(9.2) \qquad \sigma_{n/n}(x_n/y^n) = E_0(L_n \; 1\!\!1_{\{x_n\}}(X_n) \;/\; Y^n=y^n)$$

$$(9.3) \qquad \sigma_{n+1/n}(x_{n+1}/y^n) = E_0(L_n \; 1\!\!1_{\{x_{n+1}\}}(X_{n+1}) \;/\; Y^n=y^n) \quad .$$

Alors la formule (9.1) devient :

$$(9.4) \qquad \Pi_{n/n}(x_n/y^n) = \frac{\sigma_{n/n}(x_n/y^n)}{\sum_{x_n \in E} \sigma_{n/n}(x_n/y^n)}$$

Par conséquent, pour évaluer $\Pi_{n/n}(./y^n)$ il suffit de calculer, pour chaque valeur de l'observation y^n, la mesure conditionnelle $\sigma_{n/n}(./y^n)$, que l'on appelle filtre non-normalisé. Dans ce qui suit nous nous intéresserons seulement au calcul de $\sigma_{n/n}(./y^n)$ et de $\sigma_{n+1/n}(./y^n)$. Comme dans le cas du filtre de Kalman, nous verrons que $\sigma_{n+1/n}(./y^n)$ servira d'intermédiaire pour le calcul de $\sigma_{n/n}(./y^n)$.

10 - <u>L'ALGORITHME DE FILTRAGE</u>

Soit

$$(10.1) \qquad K_n(x_n, x_{n+1}/y_n) = \sum_{b_n} 1\!\!1_{\{x_{n+1}\}}(f_n(x_n, b_n)) \; \frac{p_n(b_n, w_n(x_n, y_n))}{p_n(b_n, y_n)} \; q_n(b_n/y_n)$$

On a, pour tout $n \in \mathbb{N}$:

$$(10.2) \qquad \sigma_{n+1/n}(x_{n+1}/y^n) = \sum_{x_n} \sigma_{n/n-1}(x_n/y^{n-1}) \; K_n(x_n, x_{n+1}/y_n)$$

avec $\sigma_{0/-1}(x_0/y^{-1}) = \rho(x_0)$.

Soit

$$(10.3) \qquad V_n(x_n/y_n) = \sum_{b_n} \frac{p_n(b_n, w_n(x_n, y_n))}{p_n(b_n, y_n)} \; q_n(b_n/y_n) \quad .$$

Alors on a

$$(10.4) \qquad \sigma_{n/n}(x_n/y^n) = \sigma_{n/n-1}(x_n/y^{n-1}) \; V_n(x_n/y_n) \quad .$$

Pour établir cet algorithme, on décompose les expressions (9.2) et (9.3) de $\sigma_{n+1/n}(x_{n+1}/y^n)$ et de $\sigma_{n/n}(x_n/y^n)$, en utilisant la formule du Lemme 8.

Nous reproduisons dans le paragraphe suivant, l'algorithme qui se simplifie dans le cas où B et W sont indépendants.

11 - CAS OU B ET W SONT INDEPENDANTS

Dans ce cas, on a:

$$(11.1)\qquad L_n = \prod_{i=0}^{n} \frac{r_i(w_i(X_i,Y_i))}{r_i(Y_i)}$$

On suppose les fonctions r_i strictement positives sur F. Soit

$$(11.2)\qquad \Gamma_n(x_{n-1},x_n/y_n) = \frac{r_n(w_n(x_n,y_n))}{r_n(y_n)} \cdot \sum_{b_{n-1}} 1\!\!1_{\{x_n\}}(f_{n-1}(x_{n-1},b_{n-1}))\; q_{n-1}(b_{n-1})$$

On a, pour $n \geq 1$,

$$(11.3)\qquad \sigma_{n/n}(x_n/y^n) = \sum_{x_{n-1}} \sigma_{n-1/n-1}(x_{n-1}/y^{n-1}) \cdot \Gamma_n(x_{n-1},x_n/y_n) \;.$$

avec

$$\sigma_{0/0}(x_0/y_0) = \rho(x_0)\; \frac{r_0(w_0(x_0,y_0))}{r_0(y_0)}$$

12 - REMARQUE

Dans le cas où l'équation (1.1) permet de déterminer B_n uniquement en fonction de X_n et X_{n-1}, c'est à dire si

$$(12.1)\qquad B_n = b_n(X_n,X_{n+1}) \;,$$

les formules (10.1) et (11.2) se simplifient. On a alors

$$(12.2)\qquad K_n(x_n,x_{n+1}/y_n) = \frac{p_n(b_n(x_n,x_{n+1}),w_n(x_n,y_n))}{p_n(b_n(x_n,x_{n+1}),y_n)} \cdot q_n(b_n(x_n,x_{n+1})/y_n)$$

$$(12.3)\qquad \Gamma_n(x_{n-1},x_n/y_n) = \frac{r_n(w_n(x_n,y_n))}{r_n(y_n)} \cdot q_{n-1}(b_{n-1}(x_{n-1},x_n)) \;.$$

13 - CAS OU E ET F SONT DES GROUPES COMMUTATIFS ET LES BRUITS SONT INDEPENDANTS

Nous supposons que E et F sont des groupes commutatifs discrets et additifs et nous considérons le modèle suivant qui est très courant dans les applications:

$$(13.1)\qquad X_{n+1} = F_n(X_n) + B_n$$

$$(13.2)\qquad Y_n = H_n(X_n) + W_n$$

où $F_n : E \to E$ et $H_n : E \to F$.

Pour simplifier l'écriture nous supposons que B et W sont indépendants.

Nous avons alors les formules suivantes.

$$(13.3)\qquad \Gamma_n(x_{n-1},x_n/y^n) = \frac{r_n(y_n - H_n(x_n))}{r_n(y_n)}\; q_{n-1}(x_n - F_{n-1}(x_{n-1}))$$

$$(13.4)\qquad \sigma_{0/0}(x_0/y_0) = \rho(x_0)\; \frac{r_0(y_0 - H_0(x_0))}{r_0(y_0)}$$

L'algorithme du paragraphe 11 s'applique alors aisément.

14 - SITUATION A LA VITERBI

L'algorithme qui découle du paragraphe 13 s'adapte bien aux schémas de transmission à travers des canaux symétriques sans mémoire. Cette méthode peut s'étendre au cas de canaux avec mémoire et bruit additif, entrainant l'extension de l'algorithme classique de Viterbi([12]).

Nous supposons donc que E, F, B et W sont comme au paragraphe 13. Etant donnée une suite $\{s_n ; n \in \mathbb{N}\}$, nous désignons par s_m^n le vecteur-colonne de composantes $s_m, s_{m+1}, \ldots, s_n$, avec $m < n$.

Considérons les processus X et Y définis par

$$(14.1)\qquad X_n = f_{n-1}(X_{n-1},\ldots,X_{n-a}) + B_{n-1} \quad , \; a > 1,\; n \geq 1$$

$$(14.2)\qquad Y_n = h_n(X_n,\ldots,X_{n-a+1}) + W_n$$

Il est sous-entendu que, pour $0 \leq n < a-1$, f_n et h_n sont bien définies comme fonctions des X d'indices non-négatifs.

En posant

$$X_{n-a+1}^n = (f_{n-1}(X_{n-1},\ldots,X_{n-a})+B_{n-1},X_{n-1},\ldots,X_{n-a+1})^t$$

où $(\ldots.)^t$ désigne le transposé, on peut écrire

$$(14.3)\qquad X_{n-a+1}^n = F_{n-1}(X_{n-a}^{n-1}) + B_{n-1}$$

$$(14.4)\qquad Y_n = H_n(X_{n-a+1}^n) + W_n$$

où $H_n(X_{n-a+1}^n) = h_n(X_n,\ldots,X_{n-a+1})$ et B_{n-1} désigne le vecteur-colonne dont la première composante est la v.a. B_{n-1} de (14.1) et les autres composantes sont toutes nulles. L'application des formules du paragraphe 13 au modèle défini par (14.3) et (14.4) donne:

$$(14.5)\qquad \sigma_{0/0}(x_0/y_0) = \rho(x_0)\; \frac{r_0(y_0 - h_0(x_0))}{r_0(y_0)} \;,$$

pour $1 \leq n \leq a-1$:

$$(14.6)\qquad \sigma_{n/n}(x_0,x_1,\ldots,x_n/y^n) = \sigma_{n-1/n-1}(x_0,\ldots,x_{n-1}/y^{n-1}) \cdot \frac{r_n(y_n - h_n(x_n,\ldots,x_0))}{r_n(y_n)} \cdot q_{n-1}(x_n - f_{n-1}(x_{n-1},\ldots,x_0))$$

et pour $n \geq a$:

(14.7) $\sigma_{n/n}(x_{n-a+1},\dots,x_n/y^n) =$

$$. \frac{r_n(y_n - h_n(x_n,\dots,x_{n-a+1}))}{r_n(y_n)} \; . \sum_{x_{n-a}} \sigma_{n-1/n-1}(x_{n-a},\dots,x_{n-1}/y^{n-1}) \; . \; q_{n-1}(x_n - f_{n-1}(x_{n-1},\dots,x_{n-a}))$$

L'algorithme classique de Viterbi correspond à la détection de la séquence X_0, X_1, ..., X_{a-1}, étant donné $Y^{a-1} = y^{a-1}$. Dans le cas considéré ici, la décision s'appuie sur

(14.8) $\sigma_{a-1/a-1}(x_0,\dots,x_{a-1}/y^{a-1}) =$

$$\prod_{k=0}^{a-1} \frac{r_k(y_k - h_k(x_k,\dots,x_0))}{r_k(y_k)} \; . \; q_{k-1}(x_k - f_{k-1}(x_{k-1},\dots,x_0))$$

avec

$$q_{-1}(x_0 - f_{-1}(.)) = \rho(x_0) \quad .$$

La séquence x_0, x_1,..., x_{a-1} la plus probable est celle pour laquelle l'expression (14.8) prend sa valeur maximale. Remarquons que si $X_n = B_n$ pour $n \geq 1$ et si les v.a. X_0, B_0, ..., B_n étaient uniformément réparties, ceci reviendrait à rechercher le maximum du rapport de vraisemblance:

(14.9) $$\prod_{k=0}^{a-1} \frac{r_k(y_k - h_k(x_k,\dots,x_0))}{r_k(y_k)}$$

comme dans le cas de l'algorithme de Viterbi.

15 - MODELE AVEC ESPACES D'ETAT CONTINUS

Nous considérons ici le cas où $E = \mathbb{R}^s$ et $F = \mathbb{R}^m$. Le modèle est encore le même, à savoir:

(15.1) $X_{n+1} = f_n(X_n,B_n)$

(15.2) $Y_n = h_n(X_n,W_n)$.

Nous supposons, comme au début, que sous la loi $\mathbb{P}_0$, la suite X_0, (B_0,Y_0), ..., (B_n,Y_n), .. est une suite indépendante, et que les lois de probabilités des v.a. considérées dans le modèle sont définies par leurs densités. Nous désignons par $\rho(x)$ la densité de probabilité de X_0, par $p_n(b,y)$ celle de (B_n,Y_n) et par $q_n(b/y)$ la densité conditionnelle de B_n étant donné $Y_n=y$. Nous supposerons que les p_n sont positives presque partout relativement à la mesure de Lebesgue, et que B_n est uniquement défini par

(15.3) $B_n = b_n(X_n,X_{n+1})$

lorsque X_n et X_{n+1} sont donnés, de même que W_n est uniquement défini par

(15.4) $W_n = w_n(X_n,Y_n)$

lorsque X_n et Y_n sont donnés. Nous supposons de plus que, pour tout x fixé, les fonctions $f_n(x,.)$ et $b_n(x,.)$ sont de classe C^1. Il faut signaler ici que l'hypothèse de réciprocité entre B_n et X_{n+1} pour X_n fixé, n'est pas nécessaire pour l'application de la méthode, mais elle simplifie beaucoup l'exposé.

Nous prenons pour rapport de vraisemblance :

(15.5) $$L_n = \prod_{i=0}^{n} \frac{p_i(B_i,w_i(X_i,Y_i))}{p_i(B_i,Y_i)}$$

La loi de probabilité $\mathbb{P}$ est alors définie par la suite des densités marginales

(15.6) $$\rho(x_0) \prod_{i=0}^{n} p_i(b_i,w_i(x_i,y_i)) \quad .$$

La loi conditionnelle de X_n étant donné $Y^n=y^n$ est donnée par la densité conditionnelle:

(15.7) $$\Pi_{n/n}(x_n/y^n) = \frac{\sigma_{n/n}(x_n/y^n)}{\int_{\mathbb{R}^s} \sigma_{n/n}(x_n/y^n)\, dx_n}$$

où la densité $\sigma_{n/n}(./y^n)$, (le filtre non-normalisé dans ce cas) est obtenue par l'algorithme suivant.

16 - L'ALGORITHME DE FILTRAGE DANS LE CAS CONTINU

Soient

(16.1) $$K_n(x_n,x_{n+1}/y_n) = \frac{p_n(b_n(x_n,x_{n+1}),w_n(x_n,y_n))}{p_n(b_n(x_n,x_{n+1}),y_n)} \; . \; q_n(b_n(x_n,x_{n+1})/y_n) \; . \left|\frac{\partial b_n(x_n,x_{n+1})}{\partial x_{n+1}}\right|$$

où $\left|\frac{\partial b_n(x_n,x_{n+1})}{\partial x_{n+1}}\right|$ est la valeur absolue du déterminant du Jacobien de b_n par rapport à x_{n+1};

(16.2) $\sigma_{n+1/n}(x_{n+1}/y^n) =$

$$\int_{\mathbb{R}^s} \sigma_{n/n-1}(x_n/y^{n-1})\; K_n(x_n,x_{n+1}/y_n)\; dx_n$$

avec

$$\sigma_{0/-1}(x_0/y^{-1}) = \rho(x_0) \quad ,$$

(16.3) $$V_n(x_n/y_n) = \int_{\mathbb{R}^s} \frac{p_n(b_n,w_n(x_n,y_n))}{p_n(b_n,w_n)}\; q_n(b_n/y_n)\; db_n$$

Alors on a

(16.5) $\sigma_{n/n}(x_n/y^n) = V_n(x_n/y_n)\,\sigma_{n/n-1}(x_n/y^{n-1})$.

17 - APPROXIMATION DU CAS CONTINU

Pour des raisons pratiques, il arrive que l'observation soit connue uniquement par ses valeurs quantifiées. Alors le problème d'estimation en fonction des valeurs quantifiées de Y peut se déduire de l'algorithme précédent, comme ci-dessous.

Soit $\{S_n ; n\in\mathbb{N}\}$ une suite des valeurs prises par le processus $\{Y_n ; n\in\mathbb{N}\}$ quantifié. On supposera que la partition de $\mathbb{R}^m$ correspondant à la quantification est fini. Posons :

(17.1) $K_n(x_n,x_{n+1}/S_n) = \int_{S_n} K_n(x_n,x_{n+1}/y_n)\,dy_n$

(17.2) $V_n(x_n/S_n) = \int_{S_n} V_n(x_n,y_n)\,dy_n$

Alors on a

(17.3) $\sigma_{n+1/n}(x_{n+1}/S^n) =$

$$\int_{\mathbb{R}^s} K_n(x_n,x_{n+1}/S_n)\,\sigma_{n/n-1}(x_n/S^{n-1})\,dx_n$$

(17.4) $\sigma_{n/n}(x_n/S^n) = V_n(x_n/S_n)\,\sigma_{n/n-1}(x_n/S^{n-1})$

avec

$$\sigma_{0/-1}(x_0/S^{-1}) = \rho(x_0)$$

et S^n désigne la suite $S_0, \ldots, S_n$.

Pour calculer la densité de probabilité conditionnelle de X_n étant donné $Y^n=S^n$, il suffit de normaliser $\sigma_{n/n}(x_n/S^n)$ comme fonction de x_n.

La dérivation de cet algorithme à partir du précédent est basée sur la définition de l'espérance conditionnelle par rapport à une algèbre de Boole finie.

Lorsque la partition de $F=\mathbb{R}^m$ correspondant à la quantification considérée est de plus en plus fine, le filtre calculé ici s'approche du filtre du paragraphe 16. D'autre part, il est évident qu'en approximant toutes les densités en $x_0, x_1, \ldots, x_n, \ldots$ qui interviennent dans l'algorithme ci-dessus par des fonctions quantifiées sur une base de partition de $E=\mathbb{R}^s$, on obtient un algorithme récursif (approché) pour l'évaluation des probabilités conditionnelles des X_n. Ces problèmes d'approximation seront abordés dans un article ultérieur.

18 - REMARQUE

Le modèle du paragraphe 16 avec bruits additifs et indépendants a été considéré par plusieurs auteurs (cf. par exemple ([1]), ([4]), ([5]), ([6])). A part son intérêt intrinsèque, le modèle a souvent été considéré, sous cette forme, pour servir d'approximation au modèle à temps continu. (Nous nous référons à ([7]) pour ce dernier). Mais le modèle général considéré au paragraphe 16 s'adapte mieux au schéma d'approximation du cas continu, car les meilleures approximations d'une diffusion par échantillonnage font apparaître des équations du type (15.1), qui sont non-linéaires par rapport au bruit (cf. ([11])).

BIBLIOGRAPHIE

([1]) D.L. ALSPACH - H. W. SORENSON: Nonlinear bayesian estimation using gaussian sum approximations. IEEE Vol. AC-17, n° 4, p. 439-448, 1972.

([2]) B.D.O. ANDERSON - J.B. MOORE: Optimal filtering. Prentice Hall, 1979.

([3]) P. BREMAUD - M. YOR: Changes of filtration and of probability measures. Z. Wahrsch. V. Geb. 45, p. 269-295, 1978.

([4]) R.S. BUCY: Bayes theorem and digital realizations for non-linear filters. J. Astronautical Sc. Vol. XVII, n°2, p. 80-94, 1969.

([5]) G.B. DI MASI - W.J. RUNGGALDIER: On measure transformations for combined filtering and parameter estimation in discrete time. Systems and Control Letters 2, p. 57-68, 1982.

([6]) G.B. DI MASI - W.J. RUNGGALDIER: Approximations and bounds for discrete-time nonlinear filtering. Lect. N. Control and Information Sciences 44, p. 191-202, Springer Verlag, 1982.

([7]) M. FUJISAKI - G. KALLIANPUR – H. KUNITA: Stochastic differential equations for the nonlinear filtering problem. Osaka J. Math. 9, p. 19-40, 1972.

([8]) G. KALLIANPUR - C. STRIEBEL: Estimation of stochastic systems: arbitrary system process with additive white noise observation errors. Ann. Math. Stat. 39, p. 785-801, 1968.

([9]) J. NEVEU: Bases mathématiques du calcul des probabilités. Masson, Paris, 1970.

([10]) J. SZPIRGLAS - G. MAZZIOTTO: Modèle général de filtrage non-linéaire et équations différentielles stochastiques associées. Ann. Inst. H. Poincaré Vol. XV, n°2, p. 147-173, 1979.

([11]) D. TALAY: How to discretize stochastic differential equations. Lect. N. of C.I.M.E., 1981.

([12]) A.J. VITERBI - J.K. OMURA: Principes des communications numériques. Dunod, Paris, 1982.

([13]) M. ZAKAI: On the optimal filtering of diffusion processes. Z. Wahrsch. V. Geb. 11, p. 230-243, 1969.