COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 21 JUIN 75

INNOVATION AU SENS LARGE ET FILTRAGE LINEAIRE RECURSIF

WIDE SENSE INNOVATION AND RECURSIVE LINEAR FILTERING

P. HIRSCHLER et H. KOREZLIOGLU

E.N.S.T. - 46 rue Barrault - 75634 PARIS Cedex 13

RESUME

Etant donnée une fonction aléatoire $Z = \{Z(t), t \in [0,1]\}$ du second ordre, nous désignons par $\mathcal{H}_t(Z)$ l'espace de Hilbert engendré par les variables aléatoires $Z(s)$, $s \leq t$. Nous supposons que toutes les fonctions aléatoires considérées ici sont centrées.

Nous considérons le modèle de communication suivant :

Le signal utile X est une fonction aléatoire du second ordre et le signal reçu Y est une fonction aléatoire donnée par $Y(t) = \int_0^t S(u)du + W(t)$, où le signal transmis S est une fonction aléatoire telle que $S(t) \in \mathcal{H}_t(X)$ pour presque tout t et le bruit W est une fonction aléatoire à accroissements homogènes non-corrélés, telle que $W(t) - W(s) \perp \mathcal{H}_s(X)$ pour $s \leq t$.

Soit $\hat{S}(u)$ la projection de $S(u)$ sur $\mathcal{H}_u(Y)$, définie pour presque tout u. Alors nous démontrons que l'innovation au sens large de Y définie par $\nu(t) = Y(t) - \int_0^t \hat{S}(u)du$, est une fonction aléatoire à accroissements homogènes non-corrélés telle que $\nu(t) - \nu(s) \perp \mathcal{H}_s(Y)$ pour $s \leq t$. Si $\mathcal{H}_t(Y) = \mathcal{H}_t(\nu)$ pour tout t, nous disons qu'il existe une équivalence causale entre le signal reçu et son innovation.

Dans le cas où X est une fonction aléatoire continue en moyenne quadratique et markovienne au sens large, nous donnons les conditions nécessaires et suffisantes pour qu'elle satisfasse l'équation : $X(t) = \int_0^t F(u)\, X(u)du + \int_0^t G(u)dB(u)$, où F et G sont des fonctions à carrés sommables et B une fonction aléatoire à accroissements homogènes non-corrélés. Si en plus, dans l'expression du signal reçu Y, $S(u) = H(u)\, X(u)$ où H est une fonction bornée, alors sous l'hypothèse de l'équivalence causale entre Y et ν (l'hypothèse qui est satisfaite dans la plupart des cas d'application), nous démontrons que la projection $X(t)$ sur $\mathcal{H}_t(Y)$ est obtenue par le procédé récursif de Kalman-Bucy.

SUMMARY

Given an arbitrary second order random function $Z = \{Z(t), t \in [0,1]\}$, let $\mathcal{H}_t(Z)$ denote the Hilbert space generaled by the family of random variables $Z(s)$, $s \leq t$. All the random processes considered here have zero expectation.

We consider the following communication model :

The signal X is a second order random function, and the received signal Y is a random function defined by $Y(t) = \int_0^t S(u)du + W(t)$, where the transmitted signal S is a random function such that $S(t) \in \mathcal{H}_t(X)$ for almost all t, and the noise W is a random function with homogeneous uncorrelated increments such that $W(t) - W(s) \perp \mathcal{H}_s(X)$, $s \leq t$.

Let $\hat{S}(u)$ be the projection of $S(u)$ onto $\mathcal{H}_u(Y)$, defined for almost all u. Then we prove that the wide-sense innovation of Y, defined by $\nu(t) = Y(t) - \int_0^t \hat{S}(u)du$, is a random function with homogeneous uncorrelated increments, such that $\nu(t) - \nu(s) \perp \mathcal{H}_s(Y)$, $s \leq t$. We say that there is causal equivalence between the received signal and its innovation if $\mathcal{H}_t(Y) = \mathcal{H}_t(\nu)$ for all t.

In the special case where X is a quadratic mean continuous wide-sense Markov random function, we give necessary and sufficient conditions that it satisfy the equation : $X(t) = \int_0^t F(u)\, X(u)du + \int_0^t G(u)dB(u)$, where F and G are square integrable functions and B is a random function with homogeneous uncorrelated increments. If, in addition, the received signal Y is obtained with $S(u) = H(u)\, X(u)$, for a bounded function H, then, under the condition of causal equivalence between Y and ν (that is satisfied in most cases of application), we prove that the projection of $X(t)$ onto $\mathcal{H}_t(Y)$ is obtained by the Kalman-Bucy recursive method.

INTRODUCTION

Il est bien connu que le modèle classique suivant

$$x(t) = \int_0^t F(u)\, X(u)du + \int_0^t G(u)\, dB(u) \qquad (1)$$

$$Y(t) = \int_0^t H(u)\, X(u)du + W(t) \qquad (2)$$

où B et W sont deux mouvements browniens standards,

X comme solution de (1) est une f.a. markovienne gaussienne, Y est une fonction aléatoire gaussienne et $E\{dB(t)\, dW(\tau)\} = 0$ pour $t \neq \tau$,

conduit aux équations résursives du filtre de Kalman-Bucy (cf. [4]).

Dans ce travail, les conditions de validité des équations récursives sont établies lorsque les fonctions aléatoires B, W, X et Y ne sont plus gaussiennes, c'est-à-dire lorsque B et W sont à accroissements homogènes non corrélés, telles que $E\{dB(t)\, dW(\tau)\} = 0$ pour $t \neq \tau$.

Le travail consiste en trois parties. Une première partie est consacrée à l'établissement des propositions nécessaires pour justifier les opérations effectuées par la suite.

Un théorème d'innovation au sens large (théorème 1) constitue le résultat essentiel de la deuxième partie. Sous l'hypothèse de l'équivalence causale, une équation différentielle stochastique au sens large est établie (propositions 4 et 5) pour l'estimation du signal utile lorsque celui-ci est une fonction aléatoire à innovation.

La dernière partie est consacrée au filtrage récursif des fonctions aléatoires markoviennes au sens large. D'abord les conditions nécessaires et suffisantes pour qu'une fonction aléatoire markovienne au sens large satisfasse une équation **différentielle** stochastique au sens large sont données. Finalement, sous l'hypothèse de l'équivalence causale entre le signal reçu et son innovation les formules du filtre linéaire optimal de Kalman-Bucy sont établies.

0. NOTATIONS ET CONVENTIONS

$(\Omega, \mathcal{A}, P)$ est l'espace probabilisé, fixé une fois pour toutes, sur lequel toutes les fonctions et variables aléatoires considérées ici seront définies.

Nous ne considérerons que des fonctions aléatoires (f.a.) et des variables aléatoires (v.a.) à valeurs réelles. Le paramètre temps prendra ses valeurs dans $T = [0\,,1]$. Nous poserons $L^2(T)$ pour $L^2(T, \mathcal{B}_T, \lambda)$, où $\mathcal{B}_T$ est la tribu borélienne de T et λ la mesure de Lebesgue sur $\mathcal{B}_T$, $L^2(\Omega)$ pour $L^2(\Omega, \mathcal{A}, P)$ et $L^2(T \times \Omega)$ pour $L^2(T \times \Omega, \mathcal{B}_T \otimes \mathcal{A}, \lambda \times P)$.

Si $X = \{X(t), t \in T\}$ est une f.a. du second ordre $\mathcal{H}_t(X)$ désignera le plus petit sous-espace hilbertien de $L^2(\Omega)$ contenant les classes d'équivalence des v.a. $X(s)$, $s \leq t$ et nous écrirons $\mathcal{H}_t$ pour $\mathcal{H}_t(X)$.

Une f.a. $X = \{X(t), t \in T\}$ est dite mesurable si elle est $\mathcal{B}_T \otimes \mathcal{A}$ - mesurable comme application de $T \times \Omega$ dans $\mathbb{R}$. Nous considérerons des f.a. mesurables dont les $\lambda \times P$-classes d'équivalence appartiennent à $L^2(T \times \Omega)$. Quand nous écrirons $X \in L^2(T \times \Omega)$ pour une f.a. X il sera sous-entendu que c'est sa classe d'équivalence qui appartient à $L^2(T \times \Omega)$. Nous écrirons de même $X \in L^2(\Omega)$ pour une v.a. X pour signifier que c'est sa P - classe d'équivalence qui appartient à $L^2(\Omega)$.

Une f.a. $X \in L^2(T \times \Omega)$ est dite en escalier s'il existe une partition $0 = t_0 < t_1 < \dots < t_n = 1$ de $[0, 1]$ et des v.a. $X_0, X_1, \dots, X_{n-1}$ dans $L^2(\Omega)$ telles que

$$X(t) = \sum_{k=0}^{n-2} X_k\, 1_{[t_k, t_{k+1}[}(t) + X_{n-1}\, 1_{[t_{n-1}, t_n]}(t) \qquad (0.1)$$

où 1_A, pour une partie A de $[0, 1]$, est la fonction caractéristique de A.

Rappelons que pour tout $X \in L^2(T \times \Omega)$ il existe une suite $\{X_n, n \geq 1\}$ de f.a. en escalier qui converge vers X dans $L^2(T \times \Omega)$ et que l'intégrale $\int_0^1 X(t)dt$, comme intégrale de Lebesgue des trajectoires, définit une v.a. appartenant à $L^2(\Omega)$.

Une f.a. X à accroissements homogènes non-corrélés est une f.a. du second ordre centrée à accroissements non-corrélés telle que $E[X(t) - X(s)]^2 = |t - s|$ et de valeur initiale nulle.

1. PRELIMINAIRES

PROPOSITION 1

Soient une f.a. $X \in L^2(T \times \Omega)$ et $\{\mathcal{H}_t, t \in T\}$ une famille croissante de sous-espaces hilbertiens de $L^2(\Omega)$, (c'est-à-dire si $t_1 < t_2$ alors $\mathcal{H}_{t_1} \subset \mathcal{H}_{t_2}$). Soit P_t la projection de $L^2(\Omega)$ sur $\mathcal{H}_t$. Alors il existe une f.a. $\hat{X} \in L^2(T \times \Omega)$ telle que $\hat{X}(t) = P_t\, X(t)$ pour presque tout t.

Quand nous écrirons $\int_0^t P_u\, X(u)du$ il sera sous-entendu qu'il s'agit de l'intégrale $\int_0^t \hat{X}(u)du$.

Démonstration

(i) Soit $Z \in L^2(\Omega)$ et soit $Y(t) = P_t Z \in \mathcal{H}_t$. Nous avons pour $\varepsilon > 0$ et $\delta > 0$

$$E\{[Y(t+\varepsilon) - Y(t+\delta)]^2\} = E[Y^2(t+\varepsilon)] + E[Y^2(t+\varepsilon)] - 2E[Y^2(t+\eta)] \qquad (1.1)$$

où $\eta = \inf(\varepsilon, \delta)$. En particulier, pour $\delta = 0$

$$E\{[Y(t+\varepsilon) - Y(t)]^2\} = E[Y^2(t+\varepsilon)] - E[Y^2(t)] \geq 0 \qquad (1.2)$$

La relation (1.2) montre que $F(t) = E[Y^2(t)] \geq 0$ est non-décroissante. D'autre part, $F(t) \leq E(Z^2)$. Donc l'ensemble des points de discontinuité de F est dénombrable. La relation (1.1) montre que la limite à droite de $Y(t)$ existe en m.q.. De même, avec ε et δ négatifs, on peut voir que la limite à gauche de $Y(t)$ existe en m.q.. Posons

$$F(t+) = \lim_{h\downarrow o} F(t+h)\ , \qquad F(t-) = \lim_{h\downarrow o} F(t-h)\ ,$$

$$Y(t+) = \lim_{h\downarrow o} \text{m.q. } Y(t+h), \qquad Y(t-) = \lim_{h\downarrow o} \text{m.q. } Y(t-h).$$

et $a = F(o+)$, $b = F(1-)$.

Soient :

$$\tau_j^n = \inf\{t : F(t+) \geqslant a + j\, n^{-2}\, 2^{-n}(b-a)\}$$

$$S_n = \{\tau_j^n,\ j = 0, 1, \dots, n^2 2^n - 1\} \qquad \text{et}$$

$$\bigcup_{j=1}^{n} S_j = \{t_k^n : k = 0, 1, \dots, k_n - 1\} \text{ et } t_{k_n}^n = 1.$$

Considérons la suite de f.a. $\{Y_n(t),\ n \geqslant 1\}$, définie par

$$Y_n(t) = Y(t_k^n+) \text{ pour } t_k^n < t \leqslant t_{k+1}^n,\ k = o, 1, \dots, k_n - 1,$$
$$Y_n(o) = Y(o).$$

Nous avons pour $t_k^n < t \leqslant t_{k+1}^n$

$$P\{|Y(t-) - Y_n(t)| > n^{-1}\} \leqslant \frac{F(t-) - F(t_k^n+)}{n^{-2}}$$

$$\leqslant \frac{F(t_{k+1}^n-) - F(t_k^n+)}{n^{-2}} \leqslant 2^{-n}(b-a).$$

D'où $\sum_{n=1}^{\infty} P\{|Y(t-) - Y_n(t)| > n^{-1}\} < \infty$.

Le lemme de Borel-Cantelli implique donc que la suite $\{Y_n(t),\ n \geqslant 1\}$ converge p.s. vers $Y(t-)$ pour tout $t \in T$. Posons $\tilde{Y}(t) = \limsup_n Y^n(t)$. Alors $\tilde{Y}$ est une f.a. mesurable et appartient à $L^2(T \times \Omega)$, car elle est la limite supérieure d'une suite de fonctions aléatoires mesurables et

$$\int_o^1 E[\tilde{Y}^2(t)]dt \leqslant \int_o^1 E(Z^2)dt = E(Z^2) < \infty.$$

Comme $\tilde{Y}(t) = Y(t-)$ p.s. et que $Y(t-) = Y(t) = P_t Z$ pour presque tout t, on a $\tilde{Y}(t) = P_t Z$ pour presque tout t. D'autre part, $Y(t-) \in \mathcal{H}_t$ pour tout t; donc $\tilde{Y}(t) \in \mathcal{H}_t$ pour tout t.

(ii) Etant donnée une f.a. en escalier $Z \in L^2(T \times \Omega)$ définie par

$$Z(t) = \sum_{k=0}^{n-2} Z_k\, 1_{[t_k, t_{k+1}[}(t) + Z_{n-1}\, 1_{[t_{n-1}, t_n]}(t)\ ,$$

comme dans la formule (0. 1). Soit $\tilde{Y}_k$ la f.a. correspondant à la v.a. Z_k, construite comme ci-dessus à partir de $P_t Z_k$. Alors

$$\hat{Z}(t) = \sum_{k=0}^{n-2} \tilde{Y}_k(t)\, 1_{[t_k, t_{k+1}[}(t) + \tilde{Y}_{n-1}(t)\, 1_{[t_{n-1}, t_n]}(t)$$

définit une f.a. $\hat{Z}$ dans $L^2(T \times \Omega)$ telle que $\hat{Z}(t) = P_t Z(t)$ pour presque tout t et $\hat{Z}(t) \in \mathcal{H}_t$ pour tout t.

(iii) Etant donnée une f.a. $X \in L^2(T \times \Omega)$, soit $\{X_n,\ n \geqslant 1\}$ une suite de f.a. en escalier qui converge vers X dans $L^2(T \times \Omega)$. Soit, d'autre part, $\{\hat{X}_n,\ n \geqslant 1\}$ la suite de f.a. dans $L^2(T \times \Omega)$ telles que $\hat{X}_n(t) = P_t X_n(t)$ pour presque tout t et $\hat{X}_n(t) \in \mathcal{H}_t$ pour tout t, construites à partir des X_n comme dans la partie (ii) ci-dessus. Comme $\{X_n,\ n \geqslant 1\}$ est une suite de Cauchy dans $L^2(T \times \Omega)$ il en est de même de $\{\hat{X}_n,\ n \geqslant 1\}$, car

$$\int_o^1 E[\hat{X}_n(t) - \hat{X}_m(t)]^2 dt \leqslant \int_o^1 E[X_n(t) - X_m(t)]^2 dt.$$

Par conséquent $\{\hat{X}_n,\ n \geqslant 1\}$ converge dans $L^2(T \times \Omega)$. Soit $\hat{X}$ un élément de la classe d'équivalence de la limite de $\{\hat{X}_n,\ n \geqslant 1\}$ dans $L^2(T \times \Omega)$. Remarquons que

$$\lim_{n\to\infty} \int_o^1 E[X(t) - X_n(t)]^2 dt = 0.$$

Donc la suite de fonctions $\{E[X(t) - X_n(t)]^2,\ n \geqslant 1\}$ converge en mesure (de Lebesgue) vers 0 et par conséquent il existe une sous-suite $\{E[X(t) - X_{n_k}(t)]^2,\ k \geqslant 1\}$ qui converge vers 0 pour presque tout t ; c'est-à-dire qu'il existe une sous-suite $\{X_{n_k}(t),\ k \geqslant 1\}$ qui converge en m.q. vers $X(t)$ pour presque tout t. Nous avons donc

$$\lim_{n\to\infty} \text{m.q. } P_t X_{n_k}(t) = P_t X(t)$$

pour presque tout t. Comme, d'autre part,

$$\lim_{k\to\infty} \int_o^1 E[\hat{X}(t) - \hat{X}_{n_k}(t)]^2 dt = 0$$

il existe de même une sous-suite $\{\hat{X}_{n_{k_j}}(t),\ j \geqslant 1\}$ qui converge en m.q. vers $\hat{X}(t)$ pour presque tout t. Mais $\hat{X}_{n_{k_j}}(t) = P_t X_{n_{k_j}}(t)$ pour presque tout t. Par conséquent $\hat{X}(t) = P_t X(t)$ pour presque tout t.

PROPOSITION 2

Soient $X \in L^2(T \times \Omega)$, $\mathcal{H}$ un sous-espace hilbertien de $L^2(\Omega)$ et P le projecteur de $L^2(\Omega)$ sur $\mathcal{H}$. Alors il existe une f.a. $\hat{X} \in L^2(T \times \Omega)$ telle que $\hat{X}(t) = PX(t)$ pour presque tout t et que

$$P \int_o^t X(u)du = \int_o^t \hat{X}(u)du \quad \text{p.s.} \qquad (1.3)$$

pour tout t.

Chaque fois que nous écrirons $\int_o^t PX(u)du$ il sera sous-entendu qu'il s'agit de l'intégrale $\int_o^t \hat{X}(u)du$.

Démonstration

Soit encore $\{X_n,\ n \geqslant 1\}$ une f.a. en escalier qui converge vers X dans $L^2(T \times \Omega)$. Posons $\hat{X}_n(t) = PX_n(t)$ et soit $\hat{X}$ un élément de la classe d'équivalence de la limite dans $L^2(T \times \Omega)$ de la suite $\{\hat{X}_n,\ n \geqslant 1\}$. Soit, comme dans la démonstration de la proposition 1, $\{X_{n_k}(t),\ k \geqslant 1\}$ une sous-suite qui converge en m.q. vers $X(t)$ pour presque tout t. Alors nous avons

$$\lim_{n\to\infty} \text{m.q. } \hat{X}_{n_k}(t) = \hat{X}(t)$$

pour presque tout t. Donc $\hat{X}(t) = PX(t)$ pour presque tout t. D'autre part, comme

$$\lim_{n\to\infty} \text{m.q.} \int_0^t X_n(u)du = \int_0^t X(u)du \quad \text{et}$$

$$\lim_{n\to\infty} \text{m.q.} \int_0^t \hat{X}_n(u)du = \int_0^t \hat{X}(u)du \quad \text{nous avons}$$

$$\int_0^t \hat{X}(u)du = \lim_{n\to\infty} \text{m.q.} \int_0^t \hat{X}_n(u)du$$

$$= \lim_{n\to\infty} \text{m.q.}\ P \int_0^t X_n(u)du = P \int_0^t X(u)du.$$

PROPOSITION 3

Soient $X \in L^2(T \times \Omega)$ et $Y \in L^2(\Omega)$. Alors $E[YX(t)] \in L^2(T)$ et

$$E[Y \int_0^t X(u)du] = \int_0^t E[YX(u)]du. \tag{1.4}$$

Démonstration

Il est évident que YX(.) est mesurable et comme $\{E[YX(t)]\}^2 \leqslant E(Y^2)\ E[X^2(t)]$ on a $E[YX(t)] \in L^2(T)$. On peut facilement vérifier que la relation (1.4) est vraie pour toute f.a. en escalier $X \in L^2(T \times \Omega)$. Pour une f.a. $X \in L^2(T \times \Omega)$ arbitraire soit $\{X_n,\ n > 1\}$ une suite de f.a. en escalier qui converge vers X dans $L^2(T \times \Omega)$. D'après l'inégalité

$$\int_0^1 \{E[Y(X(u) - X_n(u))]\}^2 du \leqslant E(y^2) \int_0^1 E\{[X(u) - X_n(u)]^2\} du,$$

$E[YX_n(t)]$ converge vers $E[YX(t)]$ dans $L^2(T)$. Par suite

$$\lim_{n\to\infty} \int_0^t E[YX_n(u)]du = \int_0^t E[YX(u)]du.$$

comme d'autre part,

$$\lim_{n\to\infty} E[Y \int_0^t X_n(u)du] = E[Y \int_0^t X(u)du],$$

on a l'égalité (1.4).

Dans ce qui suit, nous ne considérerons plus que des fonctions aléatoires centrées.

DEFINITION 1

Soit $\{\mathcal{H}_t,\ t \in T\}$ une famille croissante de sous-espaces hilbertiens de $L^2(\Omega)$. Une f.a. du second ordre $X = \{X(t),\ t \in T\}$ est appelée une $\{\mathcal{H}_t\}$ - martingale ou une martingale au sens large par rapport à la famille $\{\mathcal{H}_t,\ t \in T\}$ si $\forall t \in T \quad X(t) \in \mathcal{H}_t$ et si pour $s \leqslant t$ $X(t) - X(s) \perp \mathcal{H}_s$.

Remarquer qu'une martingale au sens large est aussi une f.a. à accroissements non-corrélés.

DEFINITION 2

Une f.a. du second ordre $X = \{X(t),\ t \in T\}$, continue en m.q. est appelée une f.a. à innovation s'il existe

(i) une $\{\mathcal{H}_t(X)\}$ - martingale $\nu = \{\nu(t),\ t \in T\}$,

(ii) une f.a. $Z \in L^2(T \times \Omega)$ telle que $Z(t) \in \mathcal{H}_t(X)$ pour presque tout $t \in T$,

telles que

$$X(t) = \nu(t) + \int_0^t Z(u)du \tag{1.5}$$

Nous appellerons ν l'innovation de X.

REMARQUE 1

Noter que la condition : $Z(t) \in \mathcal{H}_t(X)$ pour presque tout t, dans la définition 2, implique que $\int_0^t Z(u)du \in \mathcal{H}_t(X)$. En effet, soit $Y \in L^2(\Omega)$ une v.a. telle que $Y \perp \mathcal{H}_t(X)$. Nous avons d'après la proposition 3

$$E[Y \int_0^t Z(u)du] = \int_0^t E[YZ(u)]du.$$

Comme on a $E[YZ(u)] = 0$ pour presque tout u, on obtient $Y \perp \int_0^t Z(u)du$. Par conséquent

$$\int_0^t Z(u)du \in [\mathcal{H}_t(X)^{\perp}]^{\perp} = \mathcal{H}_t(X).$$

REMARQUE 2

Dans la définition 2, $\int_0^t Z(u)du$ définit une f.a. continue en m.q.. Par conséquent, puisque X est continue en m.q., ν est une f.a. à accroissements non-corrélés et continue en m.q. telle que $\nu(o) = X(o)$.

REMARQUE 3

Si

$$X(t) = \nu'(t) + \int_0^t Z'(u)du$$

est une autre représentation de X où ν' et Z' satisfont les mêmes conditions que ν et Z dans la définition 2, alors on peut démontrer (cf. [5]) que $Z(t) = Z'(t)$ p.s. pour presque tout t et que ν et ν' coïncident p.s. Par conséquent, si une f.a. X possède la représentation (1.5) conformément aux conditions de la définition 2 cette représentation est unique.

2. CANAL DE COMMUNICATION ET PROBLEME DE FILTRAGE

Nous allons considérer ici le modèle suivant concernant la génération du signal utile, le signal transmis et le signal reçu.

(i) Le signal utile est du type

$$\text{(I)} \qquad X(t) = \int_0^t Z(u)du + N(t)$$

avec l'hypothèse

C_1 $\quad X = \{X(t),\ t \in T\}$ est une f.a. continue en m.q. dont l'innovation est $N = \{N(t),\ t \in T\}$, (cf. la définition 2).

(ii) Le signal transmis $S = \{S(t),\ t \in T\}$ est une f.a. satisfaisant à la condition

C_2 $\quad S(t) \in \mathcal{H}_t(X)$ pour presque tout $t \in T$ et $S \in L^2(T \times \Omega)$.

(iii) Le signal reçu est du type

$$\text{(II)} \qquad Y(t) = \int_0^t S(u)du + W(t) \quad \text{où}$$

C_3 $\quad W = \{W(t),\ t \in T\}$, le bruit, est une f.a. à accroissements homogènes non-corrélés telle que $W(t) - W(s) \perp \mathcal{H}_s(X)$ pour $s \leq t$.

Nous ferons en plus l'hypothèse suivante :

C_4 $\quad N(t) - N(s) \perp \mathcal{H}_s(W)$ pour $s \leq t$.

Nous désignerons par (I, II) le modèle ci-dessus.

THEOREME 1

Soit P_t le projecteur de $L^2(\Omega)$ sur $\mathcal{H}_t(Y)$. Alors la f.a. $\nu = \{\nu(t), t \in T\}$ définie par

$$\nu(t) = Y(t) - \int_o^t P_u S(u)\,du \tag{2.1}$$

est une $\{\mathcal{H}_t(Y)\}$ - martingale de même covariance que W.

Démonstration

(i) $\nu(t) \in \mathcal{H}_t(Y)$:

C'est la conséquence immédiate de la relation $\int_o^t P_u(u)du \in \mathcal{H}_t(Y)$ que l'on démontre par un argument identique à celui qui a été utilisé dans la remarque 1.

(ii) $\nu(t) - \nu(s) \perp \mathcal{H}_s(Y)$ pour $s \leq t$:

De (II) et (2.1) nous obtenons

$$\nu(t) = W(t) + \int_o^t [S(u) - P_u S(u)]du \tag{2.2}$$

et d'après la proposition 2 et l'hypothèse H_3 nous voyons que :

$$P_s[\nu(t)-\nu(s)] = P_s[W(t)-W(s)] + \int_s^t P_s[S(u)-P_uS(u)]du$$

$$= P_s[W(t)-W(s)] + \int_s^t [P_sS(u)-P_sS(u)]du = 0$$

Donc ν est une $\{\mathcal{H}_t(Y)\}$ - martingale.

(iii) Pour démontrer que ν et W ont la même covariance, il suffit de montrer que $E[\nu^2(t)] = E[W^2(t)]$.

Posons

$$\int_o^t S(u)du - \int_o^t P_u S(u)du = \int_o^t \overline{S}(u)du$$

et soit $\{0 = t_o^n < t_1^n < \dots < t_{k_n}^n = t,\ n \geq 1\}$ une suite de partitions de plus en plus fines de $[o,t]$ telle que $\lim_{n\to\infty} \max_k (t_{k+1}^n - t_k^n) = 0$. Comme ν est à accroissements non-corrélés et $\nu(o) = 0$, on a

$$E[\nu(t)]^2 = E[\sum_{k=o}^{n-1} \Delta\,\nu(t_k^n)]^2 = \sum_{k=o}^{n-1} E\{[\Delta\nu(t_k^n)]^2\} \tag{2.3}$$

où $\Delta\nu(t_k^n) = \nu(t_{k+1}^n) - \nu(t_k^n)$. $E[\nu(t)]^2$ étant indépendant de n, il est égal à la limite du dernier membre de (2.3) qui, d'après (2.2), peut s'écrire comme

$$\sum_{k=o}^{n-1} E[\Delta W(t_k^n)]^2 + 2\sum_{k=o}^{n-1} E[\Delta W(t_k^n)\int_{t_k^n}^{t_{k+1}^n} \overline{S}(u)du]$$

$$+ \sum_{k=o}^{n-1} E(\int_{t_k^n}^{t_{k+1}^n} \overline{S}(u)du)^2 \tag{2.4}$$

où $\Delta W(t_k^n) = W(t_{k+1}^n) - W(t_k^n)$. D'une part,

$$\sum_{k=o}^{n-1} E[\Delta W(t_k^n)]^2 = E[\sum_{k=o}^{n-1} \Delta W(t_k^n)]^2 = E[W^2(t)],$$

d'autre part,

$$\sum_{k=o}^{n-1} E(\int_{t_k^n}^{t_{k+1}^n} \overline{S}(u)du)^2 \leq \sum_{k=o}^{n-1} (t_{k+1}^n - t_k^n)\int_{t_k^n}^{t_{k+1}^n} E[\overline{S}^2(u)]du$$

$$\leq \max_k (t_{k+1}^n - t_k^n)\int_o^t E[\overline{S}^2(u)]du.$$

Comme $\overline{S}$ $L^2(T\times\Omega)$ nous voyons que la dernière somme dans (2.4) tend vers o. L'inégalité

$$[\sum_{k=o}^{n-1} E(\Delta W(t_k^n)\int_{t_k^n}^{t_{k+1}^n} \overline{S}(u)du)]^2$$

$$\leq \sum_{k=o}^{n-1} E[\Delta W(t_k^n)]^2 \sum_{k=o}^{n-1} E(\int_{t_k^n}^{t_{k+1}^n} \overline{S}(u)du)^2$$

montre que la deuxième somme dans (2.4) tend aussi vers o. Par conséquent $E[\nu^2(t)] = E[W^2(t)]$.

REMARQUE 4

Le théorème précédent exprime le fait que Y est une f.a. à innovation dont l'innovation ν est une f.a. à accroissements homogènes non-corrélés.

REMARQUE 5

L'énoncé du théorème 1 reste valable dans le cas où $S \in L^2(T,\Omega)$, centrée ou non, est définie indépendamment de X, et où W est une f.a. centrée à accroissements non-corrélés quelconque telle que $W(o) = 0$ et pour $s < t$ $W(t) - W(s)$ est orthogonal à l'espace de Hilbert engendré par $\{\int_o^u S(v)dv,\ u \leq s\}$.

PROPOSITION 4

Sous les conditions C_1 à C_4, la f.a. $M = \{M(t),\ t \in T\}$ définie par

$$M(t) = P_tX(t) - \int_o^t P_uZ(u)du \tag{2.5}$$

où P_t est le projecteur de $L^2(\Omega)$ sur $\mathcal{H}_t(Y)$ est une $\{\mathcal{H}_t(Y)\}$ - martingale.

Démonstration

Par application de la proposition 3, comme dans la remarque 1, on peut démontrer que $\int_o^t P_uZ(u)du \in \mathcal{H}_t(Y)$.

D'autre part, nous avons

$$P_s[M(t) - M(s)] = P_s[P_tX(t) - P_sX(s) - \int_s^t P_uZ(u)du]$$

et à l'aide de la proposition 2, nous obtenons

$$P_s[M(t) - M(s)] = P_sX(t) - P_sX(s) - \int_s^t P_sZ(u)du$$

$$= P_s[X(t) - X(s) - \int_s^t Z(u)du]$$

$$= P_s[N(t) - N(s)].$$

Mais comme $\mathcal{H}_s(Y) \subset \mathcal{H}_s(X,W)$ nous avons, d'après la condition C_4, $P_s[N(t) - N(s)] = 0$.

Dans ce qui suit, nous considérons l'hypothèse suivante :

$C_5 \qquad \mathcal{B}_t(Y) = \mathcal{H}_t(\nu)$ pout tout $t \in T$,

que nous appelons l'hypothèse d'équivalence causale entre le signal reçu et son innovation.

PROPOSITION 5

Sous les conditions C_1 à C_5, la $\{\mathcal{H}_t(Y)\}$ - martingale M de la proposition 4 est représentée par

$$M(t) = \int_0^t K(u)\, d\nu(u) \qquad (2.6)$$

où $K \in L^2(T)$.

Démonstration

La condition C_5 entraîne que tout élément de $\mathcal{H}_t(Y)$ peut être représenté par $\int_0^t f(u)d\nu(u)$, où $f \in L^2(T)$. En particulier, on peut écrire :

$$M(t) = \int_0^t g(t,u)d\nu(u).$$

Comme, d'après la proposition 4, $M(t) - M(s) \perp \nu(u)$ pour $u \leqslant s \leqslant t$, nous avons

$$E\{[M(t)-M(s)]\nu(u)\} = \int_0^u [g(t,v)-g(s,v)]dv = 0.$$

Par conséquent, $g(t,v) = g(s,v)$ pour presque tout $v \leqslant u$. En particulier, $g(1,v) = g(s,v)$ pour presque tout $v \leqslant s < 1$.

Donc, on peut prendre $g(s,v) = g(1,v) = K(v)$ pour tout $s \in T$.

La proposition suivante, que nous donnons ici sans démonstration, montre que dans le plupart des cas d'application l'hypothèse C_5 est satisfaite. (Voir [5] pour la démonstration).

PROPOSITION 6

L'hypothèse C_5 de l'équivalence causale entre le signal reçu Y et son innovation ν est satisfaite au moins dans les deux cas suivants :

(i) W est un mouvement brownien et Y est une f.a. gaussienne.

(ii) $E[W(t) \int_0^s S(u)du] = 0$ quels que soient t et s. En particulier, cette condition est satisfaite si X et W sont non-corrélés.

3. FILTRAGE RECURSIF DES FONCTIONS ALEATOIRES MARKOVIENNES AU SENS LARGE

Nous allons étudier dans ce paragraphe le problème de l'estimation linéaire d'une f.a. markovienne au sens large admettant une innovation et qui est observée à travers un canal additif bruité du type II.

Rappelons qu'une f.a. du second ordre est markovienne au sens large si et seulement si sa covariance K satisfait à

$$K(t,u) = K(t,s)\, K^{-1}(s,s)\, K(s,u) \qquad (3.1)$$

pour $u \leqslant s \leqslant t$, où l'on convient de poser $K(t,s)\, K^{-1}(s,s) = 0$ si $K(s,s) = 0$ (cf. [2]).

La proposition suivante fournit une caractérisation des f.a. markoviennes au sens large et à innovation.

PROPOSITION 7

Soit $X \in L^2(T \times \Omega)$ une f.a. continue en m.q. et de covariance K et désignons par $K_{10}(t,s)$ la dérivée $\frac{\partial K(t,s)}{\partial t}$. Alors les trois propositions a, b et c suivantes sont équivalentes.

a) X est markovienne au sens large, $X(o) = 0$ et K satisfait aux conditions ci-dessous :

(i) $K(t,s)$ est absolument continu en t pour $t \geqslant s$,

(ii) $K_{10}(t+,t)\, K^{-1}(t,t) \in L^2(T)$,

(iii) $K(t,t)$ est absolument continu.

b) $X(t) = \int_0^t F(u)\, X(u)du + \int_0^t G(u)\, dB(u)$ p.s.

où $B = \{B(t),\ t \in T\}$ est une f.a. à accroissements homogènes non-corrélés, $F \in L^2(T)$, $G \in L^2(T)$ et $G \geqslant 0$.

c) B, F et G étant comme dans b)

$$X(t) = \varphi(t) \int_0^t \varphi^{-1}(s)\, G(s)dB(s) \qquad \text{p.s.}$$

où

$$\varphi(t) = \exp \{\int_0^t F(u)du\}.$$

Pour la démonstration de cette proposition voir [3] et aussi [1] et [6].

Considérons maintenant le modèle suivant que nous désignerons par (III, IV).

Le signal utile $X \in L^2(T \times \Omega)$ est une f.a. markovienne au sens large, continue en m.q., satisfaisant à l'équation

$$\text{(III)} \quad X(t) = \int_0^t F(u)\, X(u)du + \int_0^t G(u)\, dB(u)$$

et le signal reçu est donné par

$$\text{(IV)} \quad Y(t) = \int_0^t H(u)\, X(u)du + W(t)$$

où

C_6 $\begin{cases} F,\ G \in L^2(T), \text{ H est une fonction certaine } \mathcal{B}_T\text{ - mesurable et bornée, B et W sont à accroissements homogènes non-correlés et telles que,} \\ \text{pour } s \leqslant t, \\ B(t) - B(s) \perp \mathcal{H}_s(W) \text{ et } W(t) - W(s) \perp \mathcal{H}_s(B). \end{cases}$

Remarquons que le modèle ci-dessus est un cas particulier du modèle (I, II) obtenu pour

$$N(t) = \int_0^t G(u)\, dB(u) \text{ et } S(t) = H(t)\, X(t)$$

et que si les conditions C_6 sont satisfaites, il en est de même de C_1 à C_4. Par conséquent, le théorème 1 et les propositions 4 et 5 sont valables pour le modèle (III, IV). Notons aussi que l'innovation de Y est définie par

$$\nu(t) = Y(t) - \int_0^t H(u)\, P_u X(u)\,du \qquad (3.2)$$

où P_u est le projecteur de $L^2(\Omega)$ sur $\mathcal{X}_u(Y)$. Remarquons enfin que dans les conditions ci-dessus

$$E[B(t)\ W(s)] = \int_0^{t \wedge s} C(u)\,du \qquad (3.3)$$

où $|C(t)| \leq 1$ pour presque tout t.

THEOREME 2

Sous les conditions C_6 du modèle (III, IV) et l'hypothèse C_5 de l'équivalence causale entre Y et ν l'estimation linéaire $P_t X(t)$ de X(t) en fonction de $\{Y(s),\ s \leq t\}$ est la solution de l'équation

$$P_t X(t) = \int_0^t F(u)\, P_u X(u)\,du + \int_0^t K(u)\, d\nu(u) \qquad (3.4)$$

La fonction K, le gain optimal, est donnée par

$$K(t) = P(t)\, H(t) + G(t)\, C(t) \qquad (3.5)$$

où $P(t) = E\{[X(t) - P_t X(t)]^2\}$ est la solution de l'équation de Riccati

$$\frac{dP(t)}{dt} = 2\, F(t) P(t) + G^2(t) - [P(t)\, H(t) + G(t)\, C(t)]^2 \qquad (3.6)$$

Démonstration

L'équation (3.4) n'est qu'une conséquence directe des propositions 4 et 5 appliquées au modèle (III, IV). Pour simplifier l'écriture, posons $\tilde{X}(t) = X(t) - P_t X(t)$. Les équations (III) et (3.4) donnent

$$\tilde{X}(t) = \int_0^t F(u)\, \tilde{X}(u)\,du + \int_0^t G(u)\,dB(u) - \int_0^t K(u)\,d\nu(u)$$

Compte-tenu de (2.2) cette équation peut s'écrire comme

$$\tilde{X}(t) = \int_0^t [F(u) - K(u)H(u)]\, \tilde{X}(u)\,du + \int_0^t G(u)\,dB(u) - \int_0^t K(u)\,dW(u).$$

D'après la proposition 7, nous avons

$$\tilde{X}(t) = \Psi(t) \int_0^t \Psi^{-1}(u) G(u)\,dB(u) - \Psi(t) \int_0^t \Psi^{-1}(u) K(u)\,dW(u)$$

où

$$\Psi(t) = \exp \left\{ \int_0^t [F(u) - K(u)\, H(u)]\,du \right\}.$$

Par conséquent,

$$P(t) = \Psi^2(t) \int_0^t \Psi^{-2}(u) [G^2(u) + K^2(u) - 2\, K(u) G(u) C(u)]\,du$$

Donc

$$\frac{dP(t)}{dt} = 2[F(t) - K(t)H(t)]P(t) + G^2(t) + K^2(t) - 2K(t)G(t)C(t) \qquad (3.7)$$

avec P(o) = 0. La fonction K est telle que P(t) est minimal pour tout t. Un calcul variationnel sur l'intégrale du second membre de (3.7) conduit à (3.5) et en remplaçant K par le second membre de (3.5) dans (3.7) on obtient (3.6).

REMARQUE 6

D'après la proposition 6, l'hypothèse C_5 de l'équivalence causale pour le modèle (III, IV) est satisfaite dans les deux cas suivants

(i) $(B,W) = \{(B(t),\ W(t)),\ t \in T\}$ constitue une f.a. gaussienne à valeurs dans $\mathbb{R}^2$. (Dans ce cas B et W sont des mouvements browniens).

(ii) B et W sont non-corrélés.

Mais nous pouvons conjecturer, dès maintenant, que pour le modèle (III, IV) l'hypothèse de l'équivalence causale est automatiquement satisfaite. Cette question sera étudiée en détail dans le travail [3].

REFERENCES

[1] F.J. BEUTLER, "Multivariate wide-sense Markov Processes and prediction theory", A.M.S., June 1963, Vol.34, n°2, pp.424-438.

[2] J.L. DOOB, "Stochastic Processes", John Wiley & Sons, 1953.

[3] P. HIRSCHLER, H. KOREZLIOGLU, "Filtrage linéaire des processus aléatoires à innovation", Rapport E.N.S.T., Laboratoire Théorie des Communications, 1975.

[4] R. KALMAN, R. BUCY, "New results in linear filtering and prediction theory", J. Basic Engineering, March 1961, pp.95-108.

[5] H. KOREZLIOGLU, "Innovation et factorisation des covariances", Rapport E.N.S.T., Laboratoire Théorie des Communications, 1975.

[6] V. MANDREKAR, "On multivariate wide-sense Markov Processes", Nagoya Math. J., Vol.33, 1968, pp. 7-19.