

TRAITEMENT DU SIGNAL ET SES APPLICATIONS

Nice 7 au 12 mai 1973

---

COMPLEMENTS A LA THEORIE  
DE WIENER KOLMOGOROFF.

par Paul KREE  
Département de Mathématiques  
Université de Paris 6

---

**RESUME**

Formulation de la théorie de l'estimation en termes de processus généralisés. Application à la prédiction de processus non markoviens, à la construction de champs markoviens, et à l'estimation de ces champs.

**SUMMARY**

Estimation theory for generalized stochastic processes.  
Applications : prediction of non markovian processes,  
construction and estimation of markovian fields.

Dans notre précédent exposé [1] à ce colloque, nous avons indiqué comment la théorie des processus linéaires et des probabilités cylindriques permet d'interpréter les problèmes de statistiques relatifs à des processus, comme des problèmes de statistiques relatifs à une famille infinie de v.a. Le vocabulaire étant le suivant :

statistique usuelle	statistique des processus
donnée $\vec{X}$ de n variables $X_1 \dots X_n$	donnée d'un processus $(X_t, t \in \mathbb{R})$
loi conjointe des n variables (sur un espace euclidien $E^n$ )	probabilité cylindrique (sur un espace de Hilbert)
donnée d'un échantillon $x_1 \dots x_n$	donnée d'une trajectoire $t \mapsto X_t(\omega)$
Décomposition de X suivant les composantes principales	décomposition de Karhunen Loeve du processus
analyse discriminante (test d'homogénéité)	détection d'un signal
...	

Cette façon de procéder conduit à l'utilisation de nombreuses méthodes mathématiques, dont certaines d'ailleurs ont déjà été appliquées dans d'autres techniques.:

espaces de Sobolev, méthode de Galerkin, théorie des noyaux reproduisants, formes quadratiques sur des espaces de Hilbert...

Ce point de vue commence à faire l'objet de travaux théoriques très récents en traitement du signal, mécanique statistique, et théorie des champs : voir par exemple les références [3], [4].



De plus, du point de vue numérique, étant donné qu'un ordinateur ne peut traiter que des ensembles finis de nombres, les théorèmes d'existence et d'unicité (que la théorie permet éventuellement de prouver en dimension infinie) n'ont d'intérêt pratique que si la théorie permet de construire des procédures efficaces d'approximation numérique des solutions, ces procédures permettant de se ramener à des situations connues en statistique. Ceci nous a conduits à faire le point de toutes ces techniques pratiques d'analyse des données : voir [2]. Il reste naturellement à faire la jonction entre ces nouvelles théories et les applications pratiques. Nous voudrions donner dans cet exposé à propos des problèmes d'estimation un aperçu de ce point de vue en montrant les limites et les insuffisances de la théorie classique.

### §1 - LES DONNEES DU PROBLEME.

On rappelle d'abord le principe heuristique de la théorie de Wiener-Kolmogoroff et quelques notations de la théorie des processus linéaires. On montre (en utilisant un exemple) comment se pose le problème de l'approximation linéaire optimale. Soit  $(\Omega, \mathcal{B}, P)$  un espace probabilisé et soit  $L^2(\Omega)$  l'espace de Hilbert réel des variables aléatoires (v.a.) du second ordre. Soit  $T$  une partie de  $\mathbb{R}^n$  et soit  $a$  un point de  $\mathbb{R}^n$  disjoint de  $T$ . Soit  $(X_u)_{u \in \mathbb{R}^n}$  un champ stochastique continu du second ordre sur  $\mathbb{R}^n$ , supposé observé sur  $T$ , et qu'on cherche à estimer au point  $a$ . On suppose connue la fonction de cohérence  $R$  (ou d'autocorrélation) du processus

$$\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

(1)

$$(t ; t') \mapsto R(t, t') = \mathcal{C}(X_t X_{t'})$$

On cherche un opérateur linéaire d'estimation  $E_0$ , défini sur un

espace vectoriel topologique  $F$  de fonctions  $f$  définies sur  $T$ , à valeurs dans  $\mathbb{R}$  ;  $E_0$  étant défini par une "fonction"  $K_0$  définie sur  $T$  de la manière suivante

$$(2) \quad \begin{array}{ccc} & E_0 & \\ F & \xrightarrow{\quad} & \mathbb{R} \\ f & \xrightarrow{\quad} & E_0 f = \int_{t \in T} K_0(t)u(t)dt \end{array}$$

En pratique,  $f$  est une trajectoire du processus  $X$  restreint à  $T$  :  $f$  est donc un élément aléatoire de  $F$ . Il en résulte que  $E_0 f$  est un nombre aléatoire ; et par conséquent, on a une v.a.

$$\omega \xrightarrow{Y} \int_{\theta \in T} X_\theta(\omega) K_0(\theta) d\theta$$

On cherche  $K_0$  de façon que cette v.a. soit la projection orthogonale de  $X_a$  sur l'adhérence du sous-espace de  $L^2(-\Omega)$  engendré par les v.a.  $X_t$ ,  $t$  décrivant  $T$ . D'où l'équation

$$(3) \quad \forall t \in T, \quad \xi(YX_t) = \xi(X_a X_t)$$

D'où l'équation de Wiener-Hopf (où l'inconnue est  $K_0$ )

$$(4) \quad \forall t \in T \quad \int_{\theta \in T} K_0(\theta) R(\theta, t) d\theta = R(a, t)$$

Une étude théorique supplémentaire est nécessaire pour déterminer le degré d'irrégularité de la "fonction"  $K_0$ . L'équation (4) est résolue et discutée dans des cas très simples seulement, par exemple :

$$(5) \quad n=1, T = \mathbb{R}^-, a > 0, R(t, t') = R(t-t'), \text{ la fonction } u \mapsto R(u) \text{ ayant une transformée de Fourier rationnelle.}$$

Plus particulièrement :



$$(6) \quad R(t, t') = A \exp(-\alpha |t-t'|) \Rightarrow K(t) = e^{-\alpha t} \delta_0(t)$$

On veut étendre ces résultats en utilisant la théorie des processus linéaires et des probabilités cylindriques, théorie dont on rappelle brièvement le formalisme. Cette théorie est par rapport à la théorie des processus stochastiques, ce qu'est la théorie des distributions par rapport à la théorie des fonctions. Et de même que la théorie des distributions est très utile en physique déterministe, la théorie des probabilités cylindriques et des processus linéaires est très utile dans l'étude des phénomènes aléatoires.

(7) Processus linéaires.

Un processus linéaire  $R$  de type 2 basé sur l'espace  $U$  est une application linéaire continue  $R$  de  $U$  dans  $L^2(\Omega)$ ; alors qu'un processus linéaire  $R$  basé sur l'espace vectoriel  $U$  serait seulement une application linéaire  $R$  de  $U$  dans  $L^0(\Omega)$ .

Soit  $V$  un autre espace et soit  $(A) = (D(A), A)$  un opérateur linéaire non nécessairement continu (n.n.c.) de  $V$  dans  $U$  dont le domaine  $D(A)$  est dense dans  $V$ . Si l'application linéaire

$$\begin{aligned} D(A) &\rightarrow L^2(\Omega) \\ v &\mapsto R(A(v)) \end{aligned}$$

se prolonge par continuité en une application linéaire continue  $V \rightarrow L^2(\Omega)$ , le prolongement  $RA$  de  $R \circ A = RA$  est appelé composé du processus linéaire  $R$  avec l'application  $(A)$ . La transformée de Fourier (T.F.)  $\hat{R}$  du processus linéaire  $R$  est par définition l'application

$$\begin{aligned} U &\rightarrow \mathbb{C} \\ u &\mapsto \hat{R}(u) = \mathcal{E}[\exp(-iR(u))] \end{aligned}$$



De ces définitions, il résulte que si  $R$  est composable avec  $(A)$ , la T.F. du processus  $RA$  est le prolongement continu à  $V$  de l'application  $v \mapsto R(Av)$ .

(8) Exemples généraux de processus linéaires.

a) Si  $U$  est de dimension finie, il existe une v.a.  $X$  du second ordre  $\Omega \rightarrow U^*$  telle que  $R(u) = (X, u)$  pour tout  $u$  de  $U$ . Donc  $R$  est la T.F. de la loi de  $X$ , et par conséquent,  $R$  est une fonction définie positive.

b) Si  $U$  est de dimension quelconque, notons  $(U_i)_{i \in I}$  la famille des sous-e.v. de dimension finie de  $U$  et pour tout  $i$ , notons  $R_i$  la restriction de  $R$  à  $U_i$ . Soit  $X$  un espace en dualité séparante avec  $U$ ,  $X$  étant muni de la topologie  $\sigma(X, U)$ . On peut identifier  $U_i^*$  à  $X/U_i^{\circ}$ , où  $U_i^{\circ}$  désigne le polaire de  $U_i$ . Vu a), sur tout quotient  $X/U_i^{\circ}$ , on a une loi de probabilité  $m_i$ . Si  $U_i \subset U_j$  ( $i$  et  $j \in I$ ), on en déduit  $U_i^{\circ} \supset U_j^{\circ}$ ; d'où une surjection  $s_{ij} : X/U_j^{\circ} \rightarrow X/U_i^{\circ}$ . Comme  $\hat{R}_i = \hat{R}_j|_{U_i}$ , on a  $m_i = s_{ij}(m_j)$ , d'où un système projectif  $(m_i)_{i \in I}$  de lois de probabilité sur le système projectif  $X/U_i^{\circ}$  d'espaces vectoriels. Un tel système est une probabilité cylindrique sur  $X$ . On appelle vraie probabilité sur  $X$  une mesure de probabilité sur la tribu borélienne de  $X$ : une telle probabilité définit naturellement une probabilité cylindrique sur  $X$ .

c) De la même manière que les opérations sur les distributions se définissent par transposition des opérations linéaires entre espaces de fonction d'épreuve, on peut définir naturellement les opérations sur les probabilités cylindriques par transposition des opérations sur les processus linéaires. Par exemple, si  $(A)$  est un opérateur linéaire  $V \rightarrow U$  tel que  $\bar{R}A$  est défini, et si  $(A)'$  est le transposé sur  $X$  associée au processus linéaire  $R$ , comme étant la probabilité cylindrique sur  $Y$ , associée au processus linéaire  $\bar{R}A$ .



Compléments à la théorie de Wiener-Kolmogoroff

Paul KREE

d) La T.F.  $\hat{R}$  d'un processus linéaire  $R$  basé sur  $U$  est telle que pour tout  $i \in I$ ,  $\hat{R}|_{U_i}$  est continue, définie positive égale à 1 à l'origine. On peut montrer la réciproque (pour  $\hat{R}$  donné à priori,  $R$  n'est défini qu'à une isonomie près). Par exemple, on notera processus gaussien canonique sur un espace de Hilbert, tout processus de T.F.  $u \mapsto \exp(-\frac{1}{2} \|u\|^2)$ .

(9) Vecteur moyen et opérateur de corrélation d'un processus linéaire de type 2.

Soit  $R$  un processus linéaire de type 2 basé sur l'espace  $U$ . Soit  $U'$  le dual fort de  $U$ . Alors la moyenne  $R$  du processus linéaire  $R$  basé sur  $U$  est l'élément de  $U'$  associé à la forme linéaire  $u \rightarrow \xi(Ru)$ , définie sur  $U$ . L'opérateur de corrélation  $C_R = \text{Corr}(R)$  du processus linéaire  $R$  est l'opérateur linéaire continu de  $U$  dans  $U'$  tel que :

$$(10) \quad \forall u_1, u_2 \in U \quad (C_R u_1, u_2) = \xi[(Ru_1)(Ru_2)]$$

Notons que  $C_R$  est un opérateur symétrique positif. Dans l'ensemble des opérateurs linéaires continus symétriques positifs de  $U$  dans  $U'$ , on note  $\ll$  la relation d'ordre naturel :

$$(11) \quad D \ll C \iff \forall u \in U \quad (Du, u) \leq (Cu, u)$$

Notons que si le processus  $R$  basé sur  $U$  est composable avec l'opérateur n.n.c.  $(A)$  de  $V$  dans  $U$  de domaine dense dans  $V$ , alors l'application bilinéaire :

$$(12) \quad D(A) \times D(A) \rightarrow \mathbb{R}$$

$$(v_1, v_2) \mapsto \xi[(RAv_1)(RAv_2)] = (C_R Av_1, Av_2)$$

se prolonge par continuité en une application bilinéaire



continue définie sur  $V$ . Ceci entraîne que  $A' C_R A$  est défini, et que  $\text{Corr}(RA)$  prolonge  $A' C_R A$ .

(13) Donnée simultanée de deux processus linéaires de type 2.

Soient  $R : U \rightarrow L^2(\Omega)$  et  $S : V \rightarrow L^2(\Omega)$  deux processus linéaires de type 2 basés sur les espaces  $U$  et  $V$ . Soit  $T$  le processus linéaire de type 2 basé sur  $W = U \oplus V$  tel que  $T(u \oplus v) = R(u) + S(v)$  quel que soit  $w = u \oplus v \in W$ . On écrit encore  $T = R \oplus S$ . Alors  $C_T : W \rightarrow W'$  est associé à la forme bilinéaire

$$w_1, w_2 \mapsto (C_T w_1, w_2) = \mathcal{E} [(Ru_1 + Sv_1)(Ru_2 + Sv_2)]$$

En écrivant ce dernier terme comme une somme de quatre espérances, on voit que  $C_T$  admet la décomposition par blocs :

$$(14) \quad C_T = \begin{pmatrix} C_R & C_{RS} \\ C_{SR} & C_S \end{pmatrix}$$

avec  $C_{RS} : V \rightarrow U'$  tel que  $(C_{RS} v_1, u_2) = \mathcal{E} [(Ru_2)(Sv_1)]$

et  $C_{SR} : U \rightarrow V'$  tel que  $(C_{SR} u_1, v_2) = \mathcal{E} [(Ru_1)(Sv_2)]$

(15) Un exemple heuristique.

Les physiciens sont bien habitués à l'emploi de l'espace de Hilbert des "fonctions" de carré sommable. Par conséquent, on pourrait croire qu'on peut systématiquement représenter le processus  $(X_t)_t$  défini sur la partie  $T$  de  $\mathbb{R}^n$  par le processus linéaire

$$\begin{aligned} L^2(T) &\rightarrow L^2(\Omega) \\ \varphi &\mapsto \int_T X_t \varphi(t) dt \end{aligned}$$



L'opérateur de corrélation étant alors l'opérateur intégral de  $L^2(T)$  dont le noyau est la fonction  $R(t, t')$ . Or, si l'on faisait ainsi, on obtiendrait un opérateur de corrélation non inversible, ce qui est très gênant en prédiction. Par conséquent, on doit donc représenter les processus stochastiques donnés par des processus linéaires basés sur des espaces plus généraux que  $L^2(T)$ . Ces espaces sont les espaces de Sobolev. Une définition élaborée de ces espaces fait intervenir la théorie mathématique de l'interpolation des applications linéaires : voir [7]. Mais en fait, on peut définir directement ces espaces de la manière suivante (voir exposé dans [8]).

Pour tout nombre réel  $\alpha$ , on pose

$$H^\alpha(\mathbb{R}^n) = \left\{ \text{distributions } T ; \int (1+|u|^2)^\alpha \hat{T}(u)^2 du < \infty \right\}$$

De plus, si  $T$  est un ouvert de  $\mathbb{R}^n$ , on définit  $H^\alpha(T)$  comme l'ensemble des restrictions à  $T$  des distributions de  $H^\alpha(\mathbb{R}^n)$ . Si  $F$  est un fermé de  $\mathbb{R}^n$ , on définit  $H_0^\alpha(F)$  comme étant l'ensemble des distributions  $T$  de  $H^\alpha(\mathbb{R}^n)$  qui sont portées par le fermé  $F$ .

Pour tout entier  $n$ , et tout  $s$  réel, on note  $G_s(x_1, \dots, x_n)$  la distribution (de Bessel) ayant pour transformée de Fourier

$$(1+u_1^2 + \dots + u_n^2)^{-s/2} = (1+|u|^2)^{-s/2}$$

Par exemple, pour  $n=1$ , on a  $G_1(x) = 2^{-1} \exp(-|x|)$ . Soit  $L$  un processus gaussien canonique basé sur  $L^2(\mathbb{R})$ . Par filtrage de  $L$  à travers un filtre de réponse impulsionnelle  $G_1$ , on obtient un processus  $X_t$  sur  $\mathbb{R}$  qui peut être représenté par le processus  $M = L(G_1 \star)$  basé sur  $H^{-1}(\mathbb{R})$  ;

$$(16) \quad \begin{array}{ccc} L^2(\mathbb{R}) & \xrightarrow{G_1^*} & H^1(\mathbb{R}) \\ C_L = \text{Id} \uparrow & & \uparrow C_U = G_1^* \\ L^2(\Omega) & \xrightarrow{L} & L^2(\mathbb{R}) \end{array}$$

On a  $C_M = (G_1^*(\text{Id}))(G_1^*) = G_2^*$ . Dans le problème (6), on observe  $(X_t)_t$  sur  $\mathbb{R}^-$ , et l'on cherche une estimation de  $X_a(\omega)$ , avec  $a > 0$ . On introduit donc :

. les opérateurs de restriction

$$\begin{array}{ccc} H^1(\mathbb{R}) & \xrightarrow{r_1} & H^1(\mathbb{R}^-) & \text{et} & H^1(\mathbb{R}) & \xrightarrow{r_2} & \mathbb{R} \\ u & \mapsto & u|_{\mathbb{R}^-} & & u & \mapsto & u(a) \end{array}$$

. les transposées de  $r_1$  et  $r_2$  qui sont les injections canoniques de  $U = H^{-1}(\mathbb{R}^-)$  (distribution de  $H^{-1}(\mathbb{R})$  portées par  $\mathbb{R}^-$ ) et de  $V = \mathbb{R} \delta_a$  dans  $H^{-1}(\mathbb{R})$

. les processus  $R = Mi_1$  et  $S = Mi_2$ ,  $T = R \oplus S$ .

Finalement on a le schéma :

$$(17) \quad \begin{array}{ccccc} L^2(\mathbb{R}) & \xrightarrow{G_1^*} & H^1(\mathbb{R}) & \xrightarrow{r_1} & U' = H^1(\mathbb{R}^-) & \xrightarrow{r_2} & V' = \mathbb{R} \\ C_L = \text{Id} \uparrow & & C_U = G_1^* \uparrow & & \uparrow C_S & & \uparrow C_S \\ L^2(\Omega) & \xrightarrow{L} & L^2(\mathbb{R}) & \xrightarrow{G_1^*} & H^1(\mathbb{R}) & \xrightarrow{i_1} & U = H^{-1}(\mathbb{R}^-) & \xrightarrow{i_2} & V = \mathbb{R} \delta_a \end{array}$$



Notons que  $C_T : U \oplus V \rightarrow U' \oplus V'$  est l'isomorphisme coercif  
 $v \mapsto (C_2 * v) (\bar{R} \cup \{a\})$ .

§2 - APPROXIMATION LINEAIRE OPTIMALE.

(18) Le problème de l'approximation linéaire optimale.

Soit  $T = R \oplus S$  un processus linéaire de type 2 basé sur  $W = U \oplus V$   
 Pour tout  $(A) : V \rightarrow U$  tel que  $RA$  soit défini, on pose  
 $e(A) = RA - S$ . Je dirai que  $e(A)$  est l'erreur d'approximation de  $S$   
 en fonction de  $R$ , relatif à l'opérateur d'approximation  $(A)$ . Je  
 cherche une classe  $(\hat{A})$  aussi grande que possible de tels  
 opérateurs, et je cherche  $(A_0) \in (\hat{A})$ , tels que

$$(19) \quad \forall (A) \in (\hat{A}) \quad \text{Corr } e(A_0) \ll \text{Corr } e(A)$$

Notons que  $e(A)$  est le processus linéaire de type 2 basé sur  $V$   
 obtenu en composant  $T = R \oplus S$  avec l'opérateur linéaire  
 $(A) \oplus -\text{Id}_V : V \rightarrow U \oplus V$ .

(20) Nous introduisons l'opérateur linéaire  
 $(A_0) = (D(A_0), A_0) : V \rightarrow U$  (défini si  $C_R$  est injectif) et tel  
 que

$$(21) \quad D(A_0) = \{v \in V, C_{RS} v \in \text{Im } C_R\}, (A_0 v = C_R^{-1} v) \text{ si } v \in D(A_0).$$

Une solution au problème (3) est donnée par la

(22) Proposition. Soient  $R$  et  $S$  deux processus linéaires de  
 type 2, basé respectivement sur les e.v. n. réels  $U$  et  $V$ .

(23) On suppose que  $C_R$  est injectif et que  $D(A_0)$  est dense dans  
 $V$ . Soit  $(\hat{A})$  l'ensemble des opérateurs linéaires n.n.b.

$(A) = (D(A), A)$  de  $V$  dans  $U$  tels que : (i)  $D(A_0) \cap D(A)$  est dense

dans  $V$  et (ii)  $RA$  se prolonge par continuité en un processus linéaire  $RA$  de type 2 basé sur  $V$ .

a) Alors  $(A_0) \in (\mathcal{A})$  et  $(A_0)$  est un opérateur d'approximation linéaire optimale de  $S$  en fonction de  $R$  :

$$(24) \quad \forall(A) \in (\mathcal{A}), \text{Corr}(S-RA_0) \ll \text{Corr}(S-RA).$$

b) Tout autre opérateur d'approximation linéaire optimale  $(A_1)$  de  $S$  en fonction de  $R$ , coïncide avec  $A_0$  sur  $D(A_1) \cap D(A_0)$ .

c) On a la caractérisation suivante (par moindres carrés) de  $A_0 v$ , pour  $v$  donné dans  $V$ . La forme quadratique  $u \rightarrow (C_T(u \oplus v), u \oplus v)$  atteint son minimum (unique) au point  $A_0 v$ .

(25) Remarques sur les hypothèses (7) :

L'hypothèse " $C_R$  injectif" n'est guère restrictive. L'exemple cité dans l'introduction de [9] montre qu'une hypothèse du genre " $D(A_0)$  dense dans  $V$ " est nécessaire.

(25) Propriétés de l'approximation linéaire optimale :

$$a) \quad C_R^{-1} C_{R, S_1} \oplus S_2 = C_R^{-1} C_{RS_1} \oplus C_R^{-1} C_{RS_2}.$$

b) Si  $R_1$  et  $R_2$  sont non corrélés, on a

$$C_{R_1}^{-1} \oplus C_{R_2}^{-1} C_{R_1 \oplus R_2, S} = C_{R_1}^{-1} C_{R_1 S} \oplus C_{R_2}^{-1} C_{R_2 S}.$$

c) Soit  $U_1$  un e.v.n. et soit  $\alpha : U \rightarrow U_1$  un opérateur n.n.b., injectif et surjectif tel que  $R_1 = R \alpha$  soit défini. Si l'on peut approcher de façon optimale  $S$  à l'aide de  $R$ , alors on peut



aussi approcher de façon optimale  $S$  à l'aide de  $R_1$  (mais parfois la deuxième approximation optimale peut exister sans que la première existe).

§3 - APPLICATION A L'ESTIMATION LINEAIRE OPTIMALE DES  
TRAJECTOIRES D'UN PROCESSUS.

Soient  $X$  et  $Y$  deux e.v. en dualité avec  $U$  et  $V$  respectivement ;  $X$  et  $Y$  sont munis des topologies faibles associées à ces dualités.

(26) Le problème de l'estimation linéaire optimale s'énonce ainsi :

Trouver une classe  $(\mathcal{L})$  d'opérateurs linéaires  $X \rightarrow Y$  et  $(E_0) \in (\mathcal{L})$  tels que la probabilité cylindrique sur  $Y$  associée à  $S$ - $RB$  avec  $B : V \rightarrow U$  et  $(B)' = (E_0)$  ait un opérateur de corrélation minimal.

Pour une trajectoire donnée  $x \in X$  du processus  $R$ ,  $E_0 x$  est une estimation de la trajectoire correspondante du processus  $S$ . Le problème de l'estimation linéaire optimale nous apparaît alors comme le "transposé" du problème (18) de l'approximation linéaire optimale, les espaces  $X$  et  $Y$  étant choisis en fonction de la régularité des trajectoires dont on dispose. Nous pouvons prendre pour  $(\mathcal{L})$  l'ensemble  $(\mathcal{L})'$  des transposés des opérateurs n.n.b. de la classe  $(\mathcal{L})$ . En effet, si  $\mathcal{P}$  désigne la probabilité cylindrique sur  $X$  associée à  $R$ , et si  $(A)$  désigne un élément quelconque de  $(\mathcal{L})'$ , on peut définir l'image de  $\mathcal{P}$  par  $(A)'$  comme étant la probabilité cylindrique sur  $Y$  associée à  $\bar{R}A$ .

(27) Proposition.

a) Si les hypothèses (23) sont satisfaites, dans  $(\mathcal{L}) = (\mathcal{L})'$ ,

$(A_0)'$  est un opérateur d'estimation optimale.

b) Si de plus  $C_T$  est injectif, pour tout  $x \in D(A_0')$ ,  $y = A_0'(x)$  est le point de  $\text{Im}C_T \cdot Y$  où la forme quadratique  $y \mapsto (C_T^{-1}(x \oplus y), x \oplus y)$  atteint son minimum.

Notons que (27) est bien connu en dimension finie. Un cas particulier de (27) est présenté dans 10 : les espaces sont de Hilbert, les estimateurs sont bornés et  $\text{Im}C_R$  est supposé fermé.

#### §4 - EXEMPLES ET APPLICATIONS.

(28) En résumé, la théorie exposée aux paragraphes 2 et 3 peut être ainsi résumée.

. On représente les processus donnés par des processus linéaires  $R$  et  $S$  basés sur des espaces fonctionnels  $U$  et  $V$ . Donc on remplace les fonctions d'autocorrélation et d'intercorrélacion par les opérateurs intégraux associés ; et l'on choisit  $U$  et  $V$  de façon que ces opérateurs soient inversibles.

. On considère le problème posé d'estimation, par un problème d'approximation de  $S$  à l'aide de  $R$ .

. Pour  $v$  donné dans  $D(A_0')$ ,  $A_0'v$  peut être caractérisé comme le point où une certaine forme quadratique  $Q$  atteint son minimum.

. Numériquement,  $A_0'v$  peut être défini par la méthode de Galerkin ; c'est-à-dire qu'au lieu de chercher le minimum de  $Q$  sur tout l'espace  $U$ , on cherche le minimum de  $Q$  sur le sous-espace vectoriel de  $U$  engendré par  $N$  éléments donnés de  $U$ .

(29) Exemple heuristique (15).

Comme  $V = \mathbb{R} \delta_a$  est de dimension 1, pour connaître  $A_{\circ}$ , il suffit de connaître  $u_{\circ} = A_{\circ}(\delta_a)$ . Vu (21.c),  $u_{\circ}$  est le point de  $U = H^{-1}(\mathbb{R}^-)$  tel que

$$C_R u_{\circ} = C_{RS} \delta_a$$

soit  $(G_2 * u_{\circ})(t) = G_2(t-a)$  pour tout  $t < 0$ . En appliquant  $G_2 * = -(\frac{d}{dx})^2 + 1$  à chaque membre, on voit que  $u_{\circ}$  est porté par l'origine. Comme de plus  $u_{\circ} \in H^{-1}$ ,  $u$  est proportionnel à  $\delta_{\circ}$ . Finalement, on retrouve (6). Le fait que  $u_{\circ}$  soit porté par l'origine signifie que l'estimation du futur en fonction du passé et du présent ne fait intervenir en fait que le présent (propriété de Markov).

(30) Un exemple où la méthode de Wiener-Hopf ne s'applique pas.

Nous présentons ci-après un exemple très naturel de problème de prédiction où la méthode de factorisation ne s'applique pas, (et où notre méthode s'applique). Soit toujours  $M$  un processus généralisé sur  $\mathbb{R}$  correspondant au processus gaussien canonique sur  $L^2(\mathbb{R})$ . On fait subir à ce processus la transformation

$$(M_t)_t \rightarrow (X_t)_t \quad \text{avec} \quad X_t = \int_{t-2}^t M_{\theta}(\omega) d\theta$$

Soit  $a > 0$ . On observe  $(X_t)_t$  sur  $\mathbb{R}^-$  et l'on cherche une estimation de  $X_a$ . La méthode de factorisation ne s'applique pas: si  $\phi(X_t, X_{t'}) = \phi(t-t')$ , la transformée de Laplace de  $\phi$  admet une infinité de zéros. Notre méthode donne l'opérateur suivant  $E_{\circ} : \varphi \rightarrow E_{\circ} \varphi$  d'estimation,

$$\cdot \text{ si } a > 2 : E_{\circ} \varphi = 0$$

(ceci est tout à fait naturel car alors  $X_a$  est indépendant des  $X_t$  pour  $t < 0$ )

. si  $0 < a < 2$  :  $E_0 \varphi = \sum_0^{\infty} \varphi(-2n) - \sum_0^{\infty} \varphi(a-2(n+1))$   
 (ceci est moins naturel!).

Notons que dans le cas où l'on considère un filtrage plus complexe de  $(M_t)_t$  :

$$(M_t) \rightarrow (X_t)_t \quad \text{avec} \quad X_t = \int_{-\infty}^t M(\theta) f(t-\theta) d\theta$$

où  $f$  est une certaine réponse impulsionnelle, alors la théorie ne permet pas de donner toujours une expression analytique de  $E_0 \varphi$ , mais elle permet (Galerkin) de construire des approximations numériques de  $E_0$ .

(31) Les champs markoviens.

Soit  $(X_t)$  un champ stochastique défini sur  $\mathbb{R}^n$ . On dit (intuitivement) que ce champ est markovien si pour tout ouvert régulier  $T$  de  $\mathbb{R}^n$ , la prédiction du champ dans  $T$  à l'aide du champ dans  $\partial T$ , ne fait intervenir que la connaissance du champ sur la frontière de  $T$ . De tels champs interviennent naturellement en mécanique statistique (et Nelson les a introduits très récemment en théorie de représentation des relations de commutation). La théorie ci-dessus permet de construire très facilement de tels champs (reprendre l'exemple heuristique en dimension supérieure). La théorie développée peut être utilisée pour la prédiction d'un champ obtenu comme solution d'une équation des ondes, par exemple

$$\square X_t = Y_t$$



Compléments à la théorie de Wiener-Kolmogoroff

Paul KREE

---

où  $Y_t$  est connu (bruit blanc par exemple). On obtient encore des champs markoviens.



---

BIBLIOGRAPHIE

---

- [1] P. KREE. Gretsia 1971
- [2] P. KREE. Rapport DRME 71-131 - lot 2 - partie 3
- [3] T. KAILATH and D. DUTTWEILER. IEEE Information theory - nov. 1972, vol I - Tome 18 - n° 6 p. 730-744
- [4] DE BRUCQ. Revue du Cethedec
- [5] IM GELFAND et NY VILENKIN. Tome 4. Fonctions analytiques.
- [6] L. SCHWARTZ. Séminaire à l'X. 1971
- [7] J. LIONS et E. MAGENES. Problèmes aux limites non homogènes. Tome 1 - Dunod
- [8] P. KREE. Journal Math. Pures et Appl. 1972
- [9] PAYEN. Thèse - Paris - 1967
- [10] P. KREE. Note C.R. série A - avril 1971