# DETECTION ESTIMATION ET FILTRES DE VOLTERRA

**P. DUVAUT et B. PICINBONO**

**Laboratoire des signaux et systèmes, E.S.E.**
**Plateau du Moulon 91190, Gif sur Yvette.**

**Résumé. Dans des problèmes de détection et d'estimation en moyenne quadratique il n'est pas toujours possible de mettre en oeuvre les dispositifs optimaux à savoir le rapport de vraisemblance et la régression. On peut alors les approximer au sens des moindres carrés par des filtres de Volterra discrets du vecteur observation. Le filtre optimal est solution d'un système d'équations linéaires extensions des équations normales, quelles que soient les propriétés statistiques du vecteur observation. Le filtre optimal tend vers le développement de Volterra de la régression ou du rapport de vraisemblance lorsque son ordre croit indéfiniment. Par contre la troncature du développement ne coïncide pas en général avec un filtre optimal de rang correspondant.**

**Summary.It is not always easy to compute the likelihood ratio or the regression in some detection or estimation problems. However, it is possible to try and approximate them in some mean square sense by non linear discrete Volterra filters of the observation vector. We show that whatever the statistical properties of the observation, the optimal filter is obtained as a solution of a linear system which may be viewed as generalized normal equations. As the Volterra expansion of the regression is unique, the optimal filter converges towards this expansion when its order increases indefinitely. The truncated expansion is not at all the optimal filter derived for the same order.**

## 1. INTRODUCTION

Le dispositif optimal correspondant à la détection de deux hypothèses simples $H_0$, bruit contre $H_1$, signal plus bruit calcule le rapport de vraisemblance, RV, $L(\mathbf{x})$ et le compare à un seuil, [1], [2].

La régression, RG, $R(\mathbf{x})$, où $R(\mathbf{x})=E[y|\mathbf{x}]$, représente le meilleur estimateur en moyenne quadratique, MEMQ, d'une grandeur y à partir de l'observation $\mathbf{x}$, [2], [3].

Un manque d'information ou une trop grande complexité rend parfois délicate la réalisation du RV et de la RG. En effet en dehors de la détection d'un signal déterministe dans un bruit gaussien stationnaire et d'une estimation impliquant un couple $(y,\mathbf{x})$ gaussien, $L(\mathbf{x})$ et $R(\mathbf{x})$ sont en général des fonctions non linéaires de $\mathbf{x}$, [2], [3].

Dans ces conditions il paraît judicieux de s'approcher le plus possible de ces grandeurs au moyen de filtres non linéaires de $\mathbf{x}$ et les filtres de Volterra discrets d'ordre p, FVDP, semblent tout à fait appropriés.

Quelques auteurs ont déjà présenté des travaux analogues comme Koh et Powers, [4], mais en se limitant à des filtres linéaires quadratiques et sous l'hypothèse gaussienne. La référence [5] propose une étude détaillée des filtres d'ordre 2 dans un contexte absolument quelconque.

L'objet de cet article est d'étendre certains résultats de [5], à un ordre quelconque, cette démarche étant justifiée par l'amélioration des performances qui suit l'incrémentation de l'ordre, [6].

## 2. ESPACE DE HILBERT DES FILTRES DE VOLTERRA DISCRETS D'ORDRE P

### 2.1. Filtres de Volterra discrets d'ordre p

Dans toute la suite, le vecteur observation VO, $\mathbf{x}(t)$ à valeurs dans $R^n$ évalué à l'instant discret t, $t \in Z$, admettra pour composantes $x(t-i)$ où i évolue dans l'ensemble d'indices $I_n$ de taille n ; sauf spécifications contraire, on omettra l'argument t de $\mathbf{x}$ afin d'alléger les notations.

On appelle FVDP la grandeur $F_p(\mathbf{x})$ définie de la manière suivante

$$(2.1) \qquad F_p = \sum_{k=1}^{p} \sum_{\mathbf{j}_k \in D_{n,k}} H(\mathbf{j}_k)\; X(\mathbf{t}_k - \mathbf{j}_k) \quad ,$$

où $\mathbf{j}_k$ représente un vecteur indice de taillé k

$$(2.2) \qquad \mathbf{j}_k^{\,t} = (j_1, \ldots\ldots, j_l, \ldots\ldots\ldots\ldots, j_k) \quad ,$$

et où pour tout l dans (1............k), $j_l \in I_n$. Dans ces conditions il vient évidemment

(2.3) $\qquad D_{n,k} = I_n{}^k$

Les termes $H(\mathbf{j}_k)$ composent les noyaux scalaires d'ordre k de $F_p$. Les $X(\mathbf{t}_k - \mathbf{j}_k)$ font partie des observations scalaires de rang k telles que

(2.4)

$$X(\mathbf{t}_k - \mathbf{j}_k) = \prod_{l=1}^{k} x(t-j_l) - E[\prod_{l=1}^{k} x(t-j_l)]$$

Il est clair que $\mathbf{t}_k$ se définit par une relation semblable à (2.2).

On note d'après (2.4) que $F_p$ est centré. Ceci ne restreint pas la généralité du problème, mais évite d'avoir à retirer la valeur moyenne de $F_p$ lors du calcul de sa variance.

Nous supposons que les moments du VO sont finis jusqu'à l'ordre 2p. Aucune autre hypothèse n'est retenue concernant leur nature, en particulier nous ne faisons pas l'hypothèse gaussienne.

La forme (2.1) de $F_p$ est lourde à manipuler. Aussi introduisons nous dans la suite une représentation géométrique de $F_p$ sous forme de produit scalaire Euclidien de 2 "vecteurs blocs colonnes de rang p" $[H_p]$, $[X_p(t)]$

(2.5) $\qquad F_p = [H_p]^t [X_p(t)] \qquad ,$

tels que

(2.6) $[H_p]^t = [\mathbf{H}_1{}^t, \ldots\ldots, \mathbf{H}_k{}^t, \ldots\ldots\ldots, \mathbf{H}_p{}^t] \qquad ,$

où $\mathbf{H}_k$ est un " vecteur d' ordre k" dont les $n^k$ composantes sont les termes $H(\mathbf{j}_k)$, $\mathbf{j}_k \in D_{n,k}$, $[X_p(t)]$ s'obtenant à l'aide d'une relation semblable à (2.6) ; sauf nécessité, on ne précisera plus l'argument t de $[X_p]$.

La version compacte (2.5) de $F_p$ constitue l'une des originalités de ce travail dans la mesure où elle étend la représentation linéaire traditionnelle des filtres de l'observation atteinte lorsque p=1, à des structures non linéaires quelconques. Elle ne doit pas cacher pour autant la complexité de l'outil manipulé.

Afin de fixer les idées notons que

(2.7) $\qquad F_1 = \mathbf{H}_1{}^t \mathbf{X}_1 = \mathbf{h}^t \mathbf{x}_c \qquad ,$

où $\mathbf{H}_1 = \mathbf{h}$ et $\mathbf{X}_1 = \mathbf{x}_c = \mathbf{x} - E[\mathbf{x}]$. On obtient ainsi la structure linéaire couramment introduite dans les problèmes classiques de filtrage de Wiener et d'estimation linéaire en MQ, [3]. De même

(2.8) $\qquad F_2 = [H_2]^t [X_2] = \mathbf{H}_1{}^t \mathbf{X}_1 + \mathbf{H}_2{}^t \mathbf{X}_2$

soit encore en posant en plus des notations (2.7), $H(j_1, j_2) = M(j_1, j_2)$, $\Gamma = E[\mathbf{x}\mathbf{x}^t]$

(2.9) $\qquad F_2 = \mathbf{h}^t \mathbf{x}_c + \mathbf{x}^t M \mathbf{x} - tr[M\Gamma]$

L'expression (2.9) définit un filtre linéaire quadratique fonction d'un vecteur et d'une matrice. Cette catégorie de filtres est étudiée en détail dans [5]. Avant de passer au problème d'optimisation à proprement parler il est indispensable de préciser davantage notre cadre de travail.

### 2.2. Espaces de Hilbert des FVDP

Soit $\mathcal{F}_p$ l'espace de Hilbert des filtres $F_p$ muni du produit scalaire

(2.10) $\qquad \langle F_p, G_p \rangle_K = [H_p]^t |K| [G_p] = E[F_p G_p]$

où $|K| = E[[X_p][X_p]^t]$ est un opérateur bloc défini positif qui associe à un vecteur bloc un autre vecteur bloc. S'il admet une valeur propre nulle alors il existe des filtres $F_p$ tels que $\langle F_p, F_p \rangle_K = 0$, sans que nécessairement $[H_p] = [0_p]$. De tels filtres sont "dégénérés" au sens où ils sont presque sûrement égaux à zéro car $\langle F_p, F_p \rangle_K$ n'est autre que la variance de $F_p$. Dans toute la suite nous travaillerons à l'intérieur de $\mathcal{F}_p{}'$ obtenu en supprimant de $\mathcal{F}_p$ les filtres dégénérés. On trouvera dans [5] des exemples de filtres dégénérés d'ordre 2 et leur utilisation en détection.

L'espace $\mathcal{F}_p{}'$ est isomorphe à l'espace $\mathcal{H}_p$ des vecteurs blocs d'ordre p, si l'on prend la précaution de définir le produit scalaire de 2 vecteurs blocs par

(2.11) $\qquad \langle [H_p], [G_p] \rangle_K = \langle F_p, G_p \rangle_K$

## 3. ESTIMATION EN MOYENNE QUADRATIQUE A STRUCTURE IMPOSEE NON LINEAIRE

Soit y une grandeur aléatoire centrée, il s'agit dans ce paragraphe de trouver le MEMQ de y à partir d'un filtre $F_p$ de $\mathcal{F}_p{}'$. L'inconnue du problème est $[H_p]_o$, vecteur bloc optimal qui fabrique le filtre optimal $F_{p,o}$. Quelles que soient les propriétés de **x**, le principe d'orthogonalité, [3], conduit au système linéaire d'équations suivantes

$$(3.1) \qquad [R_p]=|K|[H_p]_o \quad ,$$

où $[R_p]=E[y[X_p]]$. Le système (3.1) doit être perçu comme une généralisation des équations normales, obtenues traditionnellement dans le cas où p=1, [3]. Si |K| est défini strictement positif la solution optimale est bien d'ordre p. La variance de l'erreur d'estimation s'écrit également sans aucune difficulté

$$(3.2) \qquad \sigma_p{}^2=E[y^2] - [R_p]^t|K|^{-1}[R_p] \quad .$$

La simplicité apparente de la solution réside dans le choix d'un cadre de travail adéquat comme $F_p{}^/$. Ceci ne doit pas occulter pour autant la grande complexité liée à la résolution de (3.1). On se rend mieux compte de la complexité inhérente à cette résolution en considérant une forme développée de (3.1)

$$(3.3) \qquad \forall\, k \in [1........p] \qquad \mathbf{R}_k= \sum_{l=1}^{p} K_{k,l}\, \mathbf{H}_{l,o} \quad .$$

Dans cette équation, $\mathbf{R}_k$ et $\mathbf{H}_{l,o}$ sont respectivement les composantes de $[R_p]$ et $[H_p]_o$ conformément à (2.6), les $K_{k,l}$ quant à eux sont les éléments constitutifs de |K|

$$(3.4) \qquad K_{k,l}=E[\ \mathbf{X}_k\, \mathbf{X}_l{}^t]$$

Si l'on désigne par $f_{k,o}$ le filtre optimal à l'intérieur du sous-espace $f_k$ de $F_p{}^/$, contenant des filtres du type $f_k=\mathbf{H}_k{}^t\mathbf{X}_k$ dont le degré d'homogénéité par rapport à **x** est purement égal à k, il est faux en général d'écrire que

$$(3.5) \qquad F_{p,o}= \sum_{k=1}^{p} f_{k,o}$$

En effet, ces sous-espaces ne sont pas orthogonaux au sens statistique du terme, il existe alors un couplage entre les différents ordres k, couplage provenant du fait que l'on n'a pas toujours $K_{k,l}=0$ si $k\neq l$. Dans le cas p=2 le couplage résulte de l'existence de moments d'ordre 3 de **x** ; cette situation est étudiée dans [5].

## 4. DETECTION OPTIMALE A STRUCTURE IMPOSEE NON LINEAIRE

En dehors de la théorie statistique de la détection, [1], [2], on recense de nombreux critères fondés sur la notion de rapport signal sur bruit dont le plus connu est la déflexion [6], [7]. Ces critères jugent des statistiques de réception indépendemment de la notion de seuil. Une statistique est de bonne qualité, si sa déflexion est élevée ; sous optimale à l'intérieur d'une famille particulère de récepteurs si elle rend maximale la déflexion dans cette même famille. On rappelle que la déflexion d'une statistique S(**x**) du VO s'écrit

$$(4.1) \qquad D[\ S(\mathbf{x})\ ]=[E_1[S]-E_0[S]]^2/\ V_0[S] \quad ,$$

où $E_1[\ ]$, $E_0[\ ]$ désignent les espérances de S sous $H_1$ et $H_0$; $V_0[\ ]$, la variance sous $H_0$.

L'objet du présent paragraphe est de trouver le filtre de $F_p{}^/$ optimal au sens de la déflexion. On montre dans [6] que ce filtre est le MEMQ du RV centré sous $H_0$ dans $F_p$. Le vecteur bloc qui constitue le filtre optimal s'obtient alors en remplaçant dans (3.1) y par L(**x**)-1 et E[ ] par $E_0[\ ]$ . Il existe néanmoins une méthode directe fondée sur le produit scalaire (2.11). En effet, d'après les hypothèses de travail on a

$$(4.2) \qquad D[F_p]=\langle[H_p],|K|^{-1}[S_p]\rangle_K{}^2 \,/\, \langle[H_p],[H_p]\rangle_K \ ,$$

où $[S_p]=E_0[\ L(\mathbf{x})[X_p]\ ]=E_1[\ [X_p]\ ]$, vu la définition de L(**x**) comme quotient des densités de **x** sous $H_1$ et $H_0$, [1].

L'inégalité de Schwarz fournit l'équation dont $[H_p]_o$ est solution, $[S_p]=|K|[H_p]_o$, équation (3.1) aux substitions près, évoquées au début de ce paragraphe.

## 5. RELATIONS ENTRE LE FILTRE OPTIMAL ET LE DEVELOPPEMENT DE VOLTERRA DE LA REGRESSION

En vertu de la discussion du paragraphe 4, faire de la détection optimale au sens de la déflexion revient à faire de l'EMQ. Dans ces conditions, il sera uniquement question dans la suite d'estimation ; l'adaptation à la détection consistant à remplacer la grandeur estimée par L(**x**)-1.

Un des développement de Volterra de la RG centrée, $R(\mathbf{x})_c$ peut s'écrire, sous réserve de son existence bien sûr

$$R(\mathbf{x})_c=\sum_{k=1}^{\infty}(1/k!)\{ [\mathbf{x}^t\nabla]^k- E[ [\mathbf{x}^t\nabla]^k ] \} R(\mathbf{z})_{|\mathbf{z}=\mathbf{0}} \quad (5.1)$$

où

$$(5.2) \qquad \mathbf{x}^t\nabla = \sum_{i=1}^{n} x(i)\, \partial\, /\partial z(i) \quad ,$$

avec z(i), composante de rang i de $\mathbf{z}$.

Le développement (5.1) étant unique on a nécessairement $F_{\infty,o} = R_c(\mathbf{x})$. Par contre (5.1) n'étant pas un développement statistiquement orthogonal on a de manière quelconque

$$(5.3) \qquad TR_p[ R_c(\mathbf{x}) ]\neq F_{p,o} \quad ,$$

si $TR_p[R_c(\mathbf{x})\ ]$ désigne la troncature d'ordre p de (5.1).

## 6. CONCLUSION

On a montré dans cet article que les filtres de Volterra optimaux pour la détection au sens de la déflexion et pour l'estimation en MQ s'obtiennent en résolvant un système linéaire d'équations quelles que soient les propriétés statistiques de l'observation . Ce système fait apparaître un couplage entre les différents termes de ces filtres. Dans ces conditions la démarche qui consiste à optimiser séparément chacun des ces termes n'est pas globalement optimale.

De la même manière, contrairement à une idée reçue, la troncature du développement de Volterra de la régression ne coïncide pas en général avec le filtre optimal de rang correspondant dans la mesure où ce développement n'est pas orthogonal.

Finalement, ce développement étant unique le filtre optimal tend vers la régression quand son ordre augmente indéfiniment.

## Références

[1] C.W. HELSTROM, Statistical theory of signal detection, Oxford, Pergamon, 1968.

[2] H.L. VAN TREES, Detection, estimation and modulation theory, part 1, New-York, Wiley, 1968.

[3] A.BLANC-LAPIERRE et B.PICINBONO Fonctions aléatoires, Masson,1981.

[4] T.KOH and E.J.POWERS, "Second order Volterra filtering and its applications", Internal report , Univ.Texas at Austin, Dept.El.E.,1984.

[5] B.PICINBONO and P.DUVAUT, "Linear quadratic receivers for detection and estimation", to be published.

[6] P.DUVAUT, Contraste et détection application à la quantification et aux filtres de Volterra optimaux pour la détection, thèse de doctorat, Paris-Sud, 1987.

[7] B.PICINBONO and P.DUVAUT, "Détection and contrast" in Stochastic processes in underwater acoustics Chap.7, C.BAKER, Springer 1986.