

Le filtrage de Wiener

linéaire-quadratique à horizon fini.

Application à la prédiction

Finite length linear quadratic Wiener filtering,

application to prediction



P. DUVAUT

Laboratoire des Signaux et Systèmes, ESE, Plateau du Moulon, 91190 Gif-sur-Yvette.

Ancien élève de l'École Normale Supérieure de Cachan, agrégé de Sciences Physiques en 1981, il dirige le Département Mathématiques et Traitement du Signal de l'École Nationale Supérieure de l'Électronique et de ses Applications, ENSEA, à Gergy-Pontoise. Il poursuit des activités de recherche au Laboratoire des Signaux et Systèmes, de l'École Supérieure d'Électricité, où il a acquis sa thèse de Doctorat sous la direction du Pr B. Picinbono en février 1987, en Détection Estimation et Filtrage de Volterra. Il a enseigné et effectué des recherches pour la DRET, au Laboratoire d'Electrical Engineering de l'Université de Rhode Island, Kingston, EU, durant l'année 1988. Depuis juin 1989, il exerce une activité conseil en Traitement du Signal, auprès du Département RCM de la Thomson (centre de Malakoff). Ses principaux thèmes de recherche concernent les problèmes de Détection, d'Estimation, de Quantification et de Filtrage non linéaire optimal et adaptatif.

RÉSUMÉ

Une approche du filtrage de Wiener à horizon fini, linéaire-quadratique, dans le cas d'un processus réel de moments finis jusqu'à l'ordre quatre, est proposée puis appliquée à la prédiction de même nature. Le rôle crucial des moments du troisième ordre est mis en avant tant dans la représentation du filtrage que dans la détermination du filtre optimal ou dans l'évaluation de la variance d'estimation. Les performances d'un filtrage linéaire sont nécessairement améliorées en ajoutant un terme quadratique, dès que le processus estimé et le processus observation ne sont pas conjointement gaussiens. Le terme de diminution de la variance de l'erreur d'estimation, est calculé explicitement, quelles que soient les propriétés statistiques du processus sous-jacent. Une famille de processus aléatoires est présentée susceptible de mettre en valeur l'importance du filtrage quadratique dans certaines situations. En particulier, un exemple d'estimation linéairement nulle et quadratiquement singulière est fourni. Pour finir, il est établi que l'annulation des moments d'ordre trois entraîne nécessairement celle de la partie quadratique du prédicteur.

MOTS CLÉS

Filtrage linéaire-quadratique, filtrage de Wiener, filtrage sous contraintes, prédiction, moments d'ordre trois, principe de projection, équations normales étendues, estimation nulle, estimation singulière, performances d'estimation, caractère non gaussien.

SUMMARY

An approach to linear-quadratic Wiener filtering, in the case of a real valued discrete time stochastic process, with finite moments up to the fourth order is given and then extended to the prediction problem. The crucial role played by the third order moments is pointed out through all the derivations ranging from the representation of lq filtering, or the calculation of the optimal filter to the form of the variance of the estimate. It is shown that the performance of a linear filter is always improved by the use of a quadratic term, whenever the observation and the unknown processes are not jointly gaussian. The amount by which the variance decreases is explicitly determined, whatever the statistical properties of the underlying process are. In order to display the opportunity of a quadratic filtering, a particular family of stochastic processes is introduced. A case of null linear estimation and, simultaneously, singular quadratic estimation is raised up. Finally, it is argued that when the third order moments are vanishing, the quadratic part of a predictor vanishes also.

KEY WORDS

Linear-quadratic filtering, Wiener filtering, filtering under constraint, prediction, third order moments, orthogonality principle, extended normal equations, null estimation, singular estimation, performance of estimation, non gaussian features.

1. Introduction

Le problème du filtrage linéaire-quadratique optimal pour la détection, fondée sur le critère de déflexion, [1, 2, 3], a déjà été abordé dans les références [4] et [5]. L'article [4] se limite au cas d'un processus à valeurs réelles, cependant que la référence [5] élargit les résultats au cas d'un processus complexe « bande étroite », fréquemment rencontré en traitement d'antenne, [6, 7]. L'idée essentielle de ces deux articles consiste à utiliser l'information contenue dans les moments d'ordre deux, trois et quatre, lorsqu'ils existent, d'un processus, afin d'améliorer les performances d'un dispositif de détection à contraste, [8], maximal. Cette démarche est justifiée dès que le récepteur optimal, autrement dit, le *rappport de vraisemblance* [9], ne se déduit pas d'un filtrage linéaire par une transformation affine [8]. Les résultats de [4] et [5] sont donc applicables chaque fois que l'on n'a pas à détecter un signal déterministe dans un bruit gaussien [9]. Les développements de [4] et [5] sont fondés sur des outils communs se résumant à trois opérateurs faisant respectivement intervenir les moments d'ordre deux, trois et quatre du processus observation. Ces outils s'avèrent tout à fait adaptés à l'analyse d'une détection « non gaussienne », au sens défini plus haut, mais deviennent inadéquats, pour des raisons précisées ultérieurement, voir paragraphe 2, lorsqu'il s'agit de filtrage optimal à proprement parler, et non de détection.

L'objectif de ce papier est de proposer une approche du filtrage de Wiener à horizon fini, linéaire-quadratique, dans le cas d'un processus réel, puis de l'appliquer à la prédiction de même nature. Le paragraphe 2 met en avant le rôle crucial des moments du troisième ordre dans tout filtrage linéaire quadratique et précise du même coup les spécificités et les avantages du cadre de travail introduit par rapport à celles de [4] et [5], dans la situation qui nous intéresse ici. Le paragraphe 3 donne l'occasion d'appliquer ces outils en résolvant le problème posé par le principe de projection [9, 10]. Il est montré que l'on améliore toujours les performances d'un filtrage linéaire en ajoutant un terme quadratique, dès que le processus estimé et le processus observation ne sont pas conjointement gaussiens. Le gain en performance, évaluée en termes de variance de l'erreur d'estimation, est donné explicitement, quelles que soient les propriétés statistiques du processus sous jacent. Ceci est d'ailleurs l'occasion de comparer le filtrage de Wiener linéaire et linéaire-quadratique. Une famille de processus aléatoires est présentée susceptible de mettre en valeur l'importance du filtrage quadratique dans certaines situations. En particulier un exemple d'estimation linéairement nulle et quadratiquement singulière est fourni. Pour finir, les résultats généraux du paragraphe 3, sont appliqués, dans le paragraphe 4, au cas de la prédiction dont une spécificité liée aux moments d'ordre trois est mise en valeur.

Historiquement, les travaux sur le filtrage optimal non linéaire datent de 1958 [11], et 1963 [12]. Plus récemment, Jan Zarzycki, donne dans la référence [13] une méthode d'orthogonalisation d'espaces observation « multilinéaires » en raisonnant dans des espaces

de matrices isomorphes des espaces observations. Depuis quelques années, ce type de filtrage est devenu plus « populaire » dans la mesure où il est désormais possible d'implémenter de manière « efficace » des filtres quadratiques [14]. Le contexte reste néanmoins particulier dans la plupart des travaux qui ont recours à des hypothèses simplificatrices comme un filtrage linéaire suivi d'une quadrature [15], ou encore des données gaussiennes [16]. Certains articles ont déjà été consacré au filtrage adaptatif linéaire-quadratique, là encore sous des hypothèses restrictives liées à l'étude de la convergence des procédures [17, à 20]. Un prochain article, fondé sur le cadre présenté ci-après, qui s'affranchit de toutes ces hypothèses, sera d'ailleurs consacré à ce dernier thème.

2. Le filtrage linéaire-quadratique à horizon fini

2.1. DÉFINITION

On s'intéresse à un processus observation discret, $x(t)$, à valeurs réelles, centré, de moments finis jusqu'à l'ordre 4. On appelle filtrage linéaire-quadratique d'ordre P , flqP, à l'instant t de $x(t)$ l'opération générant le signal

$$(1) \quad y(t) = \sum_{i=1}^P h(i)x(t-i+1) + \sum_{i=1}^P \sum_{j=1}^i h(i, j) \times [x(t-i+1)x(t-j+1) - m_2(t-i+1, t-j+1)]$$

avec

$$(2) \quad m_2(t-i+1, t-j+1) = E\{x(t-i+1)x(t-j+1)\}$$

Au regard des hypothèses de travail, le signal $y(t)$ est centré. On constate, d'autre part, que toutes les observations « redondantes », de la forme $x(t-i+1)x(t-j+1)$ avec $1 \leq i \leq P$ et $i+1 \leq j \leq P$, n'ont pas été considérées dans la mesure où elles n'apportent pas d'informations statistiques supplémentaires. Si les moments d'ordre deux, donnés par (2), sont connus les propriétés du signal $y(t)$ dépendent de $P(P+3)/2$ paramètres réels : les P coefficients « linéaires » $h(i)$ et les $P(P+1)/2$ coefficients « quadratiques » $h(i, j)$.

2.2. RÔLE JOUÉ PAR LES MOMENTS D'ORDRE TROIS DANS LA REPRÉSENTATION DU FILTRAGE LINÉAIRE-QUADRATIQUE

2.2.1. Espaces observation et moments d'ordre trois

L'opération définie par (1) fait intervenir les deux espaces de variables aléatoires $X_P(t)$ et $X_{qP}(t)$ qui regroupent respectivement et sans ambiguïté aucune les observations centrées de la partie linéaire et quadratique du filtrage. Lorsque les moments d'ordre trois sont nuls ces deux espaces sont orthogonaux. Ils jouent alors des rôles séparés dans la contribution aux propriétés statistiques de $y(t)$. Si ces mêmes moments sont non nuls l'orthogonalité disparaît, et ces espaces jouent un rôle conjoint dans la génération

de $y(t)$, leur réunion, notée $x_{iqP}(t)$, forme l'espace observation alors adéquat, qu'il convient d'appeler espace « linéaire-quadratique ». Cette distinction n'a donc rien de fortuit et ne se limite pas à un simple jeu de notation, elle intègre des propriétés statistiques du processus observation, liées aux moments d'ordre trois : la représentation (1) est une décomposition orthogonale ou non.

2.2.2. *Vecteurs observation et moments d'ordre trois*
 a. les moments d'ordre trois sont non nuls, représentation couplée.

On vient de voir que si les moments d'ordre trois sont non nuls les observations linéaires et quadratiques jouent un rôle conjoint dont il convient de tenir compte en les faisant intervenir simultanément dans des vecteurs observations « linéaires-quadratiques ». Soient a et b deux entiers relatifs avec, par exemple, $a \leq b$. On introduit le vecteur observation linéaire-quadratique élémentaire volqe,

$$(3) \quad x_{iqb}(a) = [x(a) \ x^2(a) \ x(a)x(a+1) \ \dots \ x(a)x(b)]^T$$

Il regroupe toutes les variables aléatoires non redondantes « linéaires et quadratiques » que l'on peut obtenir avec $x(a)$ jusqu'à l'instant b fixé, la taille de ce vecteur, égale à $b-a+2$, varie avec a , c'est une particularité de ce type de filtrage.

Dans le même ordre d'idée, on définit le vecteur paramètre linéaire-quadratique élémentaire, vplqe, de taille $i+1$, pour $1 \leq i$,

$$(4) \quad h_{iq}(i) = [h(i) \ h(i,i) \ h(i,i-1) \ \dots \ h(i,1)]^T$$

Dans ces conditions, le produit scalaire

$$(5) \quad h_{iq}^T(i) x_{iqr}(t-i+1)$$

a un sens et représente la contribution « linéaire-quadratique de rang i » au filtrage (1) : si l'on s'éloigne progressivement de t , ce qui se produit pratiquement dès que l'on stocke des données en vue de réaliser un filtrage, le vecteur $x_{iqr}(t-i+1)$ contient les nouvelles variables aléatoires intervenant dans le filtrage linéaire-quadratique à l'instant $t-i+1$, sous forme couplée, du fait que les moments d'ordre trois sont non nuls. En faisant varier a entre $b-P+1$ et b dans (3), i entre 1 et P dans (4) on définit des vecteurs linéaires-quadratiques d'ordre P , vlqP, de taille $P(P+3)/2$,

$$(6) \quad x_{iqP}^T(b) = [x_{iqb}^T(b) \ x_{iqb}^T(b-1) \ \dots \ x_{iqb}^T(b-P+1)]$$

$$(7) \quad h_{iqP}^T = [h_{iq}^T(1) \ h_{iq}^T(2) \ \dots \ h_{iq}^T(P)]$$

En conséquence, le produit scalaire

$$(8) \quad h_{iqP}^T x_{iqP}(t)$$

existe et constitue la représentation couplée du signal $y(t)$, voir (1).

Remarque : Sans aucune ambiguïté, on définit des vecteurs « purement linéaires ou quadratiques », analogues, aux vecteurs (4), (3), (6) et (7) qui débouchent sur la représentation découplée du filtrage linéaire-quadratique (1), indispensable lorsque les moments d'ordre trois sont nuls, et parfois très utile, comme il

sera vu dans les paragraphes suivants, si ces moments sont non nuls, afin de mettre en valeur certaines propriétés qui ne sont pas forcément très claires sur la représentation (8),

$$(9) \quad y(t) = [h_P^T \ h_{qP}^T] \begin{bmatrix} x_P(t) \\ x_{qP}(t) \end{bmatrix}$$

2.2.3. *Comparaison avec les outils de travaux antérieurs*

Les opérateurs introduits dans [4] et [5] conduisent à une résolution efficace du problème de détection linéaire-quadratique optimale, au sens de la déflection, même en dehors de l'hypothèse gaussienne. Cette représentation souffre néanmoins de quelques limites dans un contexte de filtrage pur. Elle ne s'affranchit pas, tout d'abord, des observations redondantes, ce qui alourdit inutilement la complexité du filtre mis en œuvre et est donc à bannir si l'on songe à une implémentation. D'autre part, le conditionnement de la représentation (1) par les moments d'ordre trois tel qu'il est présenté au paragraphe 2.2.2 et qui met très clairement en évidence le rôle crucial de couplage joué par ces derniers n'est pas envisagé dans [4] et [5]. Pour finir, la notion de vecteur observation linéaire quadratique, donnée dans le paragraphe 2.2.2, conduit très naturellement à la représentation explicite d'un opérateur dyadique, comme, par exemple,

$$(10) \quad x_{iqP}(t) x_{iqP}^T(t)$$

indispensable à l'étude du filtrage adaptatif linéaire-quadratique, qui sera abordée dans un autre article, n'est pas envisageable dans les développements de [4] et [5].

3. Le filtrage de Wiener linéaire-quadratique à horizon fini

3.1. POSITION DU PROBLÈME

On considère un signal aléatoire centré, à valeurs réelles, inconnu, $d(t)$, relié d'une certaine façon au processus observation $x(t)$. On sait que le meilleur estimateur en moyenne quadratique de $d(t)$ à partir du vecteur $x_P(t)$ n'est autre que la régression

$$(11) \quad R(x_P(t)) = E \{ d(t) | x_P(t) \}$$

espérance conditionnelle du signal à estimer [9, 10].

Lorsque les deux processus invoqués ne sont pas conjointement gaussiens, la régression n'a aucune raison d'être une fonction linéaire du vecteur observation $x_P(t)$ [9], on songe alors à approcher le mieux possible au sens des moindres carrés, la régression, ou $d(t)$, ce qui revient au même, par un filtrage linéaire-quadratique de la forme (1). En effet, lorsque la solution optimale n'est pas linéaire, les performances d'estimation sont nécessairement améliorées par un terme quadratique, ce qui sera justifié quantitativement plus loin. Cet argument tient aussi pour des filtrages de rang plus élevé, cubique, etc., notre

choix est justifié par le fait que la non-linéarité la plus « élémentaire » est quadratique. Il s'agit donc de déterminer le vlqP optimal de taille $P(P+3)/2$, h_{1qP}^0 , qui génère par le processus que décrit (1) le signal $y^0(t)$ « le plus près possible » de $d(t)$. Étant donné qu'*a priori*, on ne sait rien des moments d'ordre trois, le développement qui suit se situe dans le contexte le plus général de tels moments non nuls, voir paragraphe 2. Dans ces conditions, la solution optimale n'est autre que la projection orthogonale de $d(t)$ sur l'espace $X_{1qP}(t)$, voir paragraphe 2,

$$(12) \quad y^0(t) = \text{PROJ}[d(t) | X_{1qP}(t)]$$

relation qui cristallise le *principe d'orthogonalité*, exposé dans [9] et [10], dans le cas linéaire.

3.2. CONSÉQUENCES DU PRINCIPE D'ORTHOGONALITÉ, ÉQUATIONS NORMALES ÉTENDUES

Dans la mesure où l'espace $X_{1qP}(t)$ est engendré par les composantes du vecteur $x_{1qP}(t)$, voir paragraphe 2, le principe (12) se réécrit

$$(13) \quad s_{1qP} = C_{1qP} h_{1qP}^0$$

avec

$$(14) \quad s_{1qP} = E \{ d(t) x_{1qP}(t) \}$$

$$(15) \quad C_{1qP} = E \{ x_{1qP}(t) x_{1qP}^T(t) \}$$

L'identité (13) signifie que malgré la non-linéarité du filtrage recherché le vecteur paramètre du filtre optimal est solution d'un système linéaire d'équation, qui représente les *équations normales étendues*, quelles que soient les propriétés statistiques des processus observation et inconnu. Un résultat semblable a été obtenu dans [4] et [5] dans un contexte de détection. D'après les résultats du paragraphe 2, la matrice C_{1qP} , appelée *matrice de covariance linéaire-quadratique* est une matrice à $P(P+3)/2$ lignes et $P(P+3)/2$ colonnes. Au regard de (6) et (3) on peut la considérer comme une matrice bloc, à partitionner en P^2 sous matrices, de taille $[i+2, j+2]$,

$$(16) \quad C_{1q}(i, j) = E \{ x_{1qt}(t-i) x_{1qt}^T(t-j) \}$$

avec $0 \leq i, j \leq P-1$, voir paragraphe 2. Si $[C_{1q}(i, j)]_{k, m}$ désigne le terme situé à l'intersection de la ligne de rang k et de la colonne de rang m de $C_{1q}(i, j)$, il vient

$$(17) \quad [C_{1q}(i, j)]_{1, 1} = m_2(t-i, t-j)$$

où m_2 est défini par (6).

$$(18) \quad [C_{1q}(i, j)]_{1, m} = m_3(t-i, t-j, t-(j-m+2)),$$

avec $2 \leq m \leq j+2$, et où m_3 est le moment d'ordre trois des variables x d'arguments correspondant. Pour finir,

$$(19) \quad [C_{1q}(i, j)]_{k, m} = m_4(t-i, t-(i-k+2), t-j, t-(j-m+2)) - m_2(t-i, t-(i-k+2)) m_2(t-j, t-(j-m+2))$$

avec $2 \leq k, m \leq i+2, j+2$, m_4 désignant le moment d'ordre 4 des variables x de rang correspondant. Le

vecteur s_{1qP} , de taille $P(P+3)/2$, d'après (14) et le paragraphe 2, se partitionne également en P sous vecteurs de taille $i+2$, avec $0 \leq i \leq P-1$,

$$(20) \quad s_{1q}(i) = E \{ d(t) x_{1qt}(t-i) \}$$

En conservant le principe des notations utilisées dans [17], [18] et [19], on a, compte tenu du centrage de $d(t)$,

$$(21) \quad [s_{1q}(i)]_1 = E \{ d(t) x(t-i) \}$$

$$(22) \quad [s_{1q}(i)]_k = E \{ d(t) x(t-i) x(t-(i-k+2)) \}$$

avec $2 \leq k \leq i+2$. Si la matrice explicitée par (15) est inversible, le vecteur paramètre optimal s'écrit

$$(23) \quad h_{1qP}^0 = C_{1qP}^{-1} s_{1qP}$$

et les performances d'estimation, évaluées en termes de variance de l'erreur d'estimation revêtent la forme suivante

$$(24) \quad v_{1qP} = E \{ d^2(t) \} - s_{1qP}^T C_{1qP}^{-1} s_{1qP}$$

3.3. RÔLE DES MOMENTS D'ORDRE TROIS DANS LA FORME DU FILTRE OPTIMAL

Afin de voir précisément l'influence des moments d'ordre trois dans la forme de la solution optimale, on choisit la représentation découplée des grandeurs évoquée au paragraphe 2.22 b, que ces moments soient d'ailleurs nuls ou non. Les équations normales étendues deviennent alors

$$(25) \quad \begin{bmatrix} s_P \\ s_{qP} \end{bmatrix} = \begin{bmatrix} C_P & C_{1|qP} \\ C_{q|1P} & C_{qP} \end{bmatrix} \begin{bmatrix} [h_{1qP}^0]_1 \\ [h_{1qP}^0]_q \end{bmatrix}$$

où s_P et s_{qP} désignent respectivement les parties linéaire et quadratique du vecteur s_{1qP} , voir (14). Les matrices C_P et C_{qP} sont les analogues linéaires et quadratiques de C_{1qP} , elles regroupent respectivement des moments d'ordre deux et quatre. Les notations concernant le vecteur optimal sont légèrement différentes de celles préconisées dans le paragraphe 2.2. En effet, dans (25), $[h_{1qP}^0]_u$ désigne la partie linéaire (resp. quadratique), du vecteur h_{1qP}^0 , lorsque $u=l$, (resp. $u=q$). Cette précaution est indispensable, si l'on veut distinguer ces deux dernières quantités des filtres de Wiener optimaux purement linéaire, h_P^0 , et purement quadratique, h_{qP}^0 , solutions, d'après une adaptation immédiate de (13), des équations

$$(26) \quad s_P = C_P h_P^0$$

$$(27) \quad s_{qP} = C_{qP} h_{qP}^0$$

On conclut, au regard de (25), (26) et (27) que la partie linéaire (resp. quadratique), du filtre de Wiener optimal linéaire-quadratique coïncide avec le filtre optimal purement linéaire (resp. purement quadratique), si et seulement si la matrice

$$(28) \quad C_{q|1P} = C_{1|qP}^T = E \{ x_{qP}(t) x_P^T(t) \}$$

est nulle, ce qui est le cas, d'après le paragraphe 2, lorsque les moments d'ordre trois sont nuls. En effet, la matrice définie par (28) est une matrice à

$P(P+1)/2$ lignes et P colonnes, à partitionner en P^2 vecteurs colonnes de taille $i+1$.

$$(29) \quad C_{q|l}(i, j) = E \{ x_{qt}(t-i) x(t-j) \}$$

avec $0 \leq i, j \leq P-1$, et où n'interviennent que des moments d'ordre trois, au vu du paragraphe 2. On conclut de ce paragraphe que lorsque les moments d'ordre trois sont nuls le système (25) devient « bloc diagonal » et l'on peut optimiser séparément la partie linéaire et la partie quadratique du filtre global.

3.4. ÉTUDE QUANTITATIVE DE L'AMÉLIORATION DES PERFORMANCES D'ESTIMATION LORSQUE LES PROCESSUS OBSERVATION ET INCONNU NE SONT PAS CONJOINTEMENT GAUSSIENS

3.4.1. Cas gaussien

Lorsque $d(t)$ et $x(t)$ sont conjointement gaussiens, les quantités s_{qp} et $C_{q|l}$ sont nulles, ne faisant intervenir que des moments d'ordre trois, par conséquent, si C_{qp} n'admet pas la valeur propre nulle, la partie quadratique du filtre optimal, $[h_{iqp}^0]_q$, s'annule, la partie linéaire, $[h_{iqp}^0]_l$ revêt ensuite la forme classique, h_p^0 , donnée par (26). Si C_{qp} est singulière, il existe au moins un vecteur quadratique de son noyau, g_{qp} , tel que

$$(30) \quad g_{qp}^T x_{qp}(t) = 0 \quad (\text{presque sûrement})$$

Cela signifie alors que l'espace observation X_{qp} est dégénéré, voir paragraphe 2.1, au sens où au moins l'une de ses observations quadratiques apparaît comme une combinaison linéaire des autres. Un caractère parfaitement prédictible, au niveau quadratique, voit alors le jour. Il suffit, ensuite, de diminuer la taille de $x_p(t)$, donc, par voie de conséquence, celle de $x_{qp}(t)$, pour s'affranchir d'une telle situation. Les liens entre la singularité de C_p et celle de C_{qp} seront examinés dans un autre papier.

3.4.2. Cas non gaussien

On suppose ici que C_{qp} est non nulle et non singulière, de même que C_p . On se propose de relier les variances v_{iqp} , v_{qp} , v_{lp} , des filtres optimaux (23), (26) et (27). On raisonne sur la représentation découplée des grandeurs, que les moments d'ordre trois soient nuls ou non, voir paragraphe 2. Le développement qui suit peut être vu comme une généralisation des travaux de la référence [21]. On appelle H_{lqp} l'espace vectoriel des vecteurs déterministes donnés sous forme découplée, $[h_p^T h_{qp}^T]$. On désigne également par H_p et H_{qp} les deux sous-espaces de H_{lqp} de vecteurs qui, respectivement, finissent par $P(P+1)/2$ zéros et débutent par P zéros. Ces deux sous-espaces sont géométriquement orthogonaux. Tout vecteur de H_p , (resp. H_{qp}), s'assimile à un vecteur de H_{lqp} avec la contrainte d'orthogonalité géométrique à H_{qp} , (resp. H_p). Cette idée est à la base de la démonstration. En effet, le bénéfice de cette vision vient du lien naturel qui s'établit alors entre le filtrage de Wiener lq et le filtrage purement linéaire ou purement quadratique. On établit, ci-

après, que les moments d'ordre trois soient nuls ou non, les deux relations suivantes

$$(31) \quad v_p = v_{iqp} + [h_{iqp}^0]_l^T R_{qp}^{-1} [h_{iqp}^0]_q$$

$$(32) \quad v_{qp} = v_{iqp} + [h_{iqp}^0]_l^T R_p^{-1} [h_{iqp}^0]_l$$

où R_p et R_{qp} désignent respectivement les sous-matrices linéaire et quadratique de la version partitionnée de C_{lqp}^{-1} . Les relations (31) et (32) se déduisent l'une de l'autre par permutation de l et de q , on établit seulement (31).

Obtention de la relation (31)

Le vecteur $[h_p^{0T} 0_{qp}^T]^T$ s'envisage comme le vecteur lq optimal sous la contrainte linéaire, dans ces conditions le principe d'orthogonalité, conduit à l'identité,

$$(33) \quad [h_p^T 0_{qp}^T] \begin{bmatrix} s_{lqp} - C_{lqp} \begin{bmatrix} h_p^0 \\ 0_{qp} \end{bmatrix} \end{bmatrix} = 0$$

pour tout $[h_p^{0T} 0_{qp}^T]^T$ dans H_p . En remarquant, dans ces conditions, que le vecteur colonne de droite de (33) appartient nécessairement à H_{qp} , on le note $[0_p^T f_{qp}^T]^T$. Sous réserve que C_{lqp}^{-1} existe, l'orthogonalité géométrique de tout vecteur de H_{qp} à $[h_p^{0T} 0_{qp}^T]^T$, conduit ensuite à la relation, compte tenu de (23),

$$(34) \quad [0_p^T h_{qp}^T] \begin{bmatrix} h_{lqp}^0 - C_{lqp}^{-1} \begin{bmatrix} 0_p \\ f_{qp} \end{bmatrix} \end{bmatrix} = 0$$

pour tout $[0_p^T h_{qp}^T]^T$ dans H_{qp} . Finalement, les formes respectives des variances, la signification de R_{qp} , voir (31) et (32), et l'identité suivante, déduite de la définition de f_{qp} ,

$$(35) \quad [h_{lqp}^0]_q = R_{qp} f_{qp}$$

conduisent à la relation (31).

Discussion

On déduit de (31) et (32) qu'à partir du moment où les processus observation et inconnus ne sont pas conjointement gaussiens, on améliore toujours les performances d'estimation en ajoutant un filtrage quadratique, que les moments d'ordre trois soient nuls ou non, dans la mesure où les matrices R_p et R_{qp} sont définies positives. Si ces derniers moments sont nuls, les expressions (31) et (32) se simplifient compte tenu de, dans ces conditions,

$$(36) \quad R_p^{-1} = C_p$$

$$(37) \quad R_{qp}^{-1} = C_{qp}$$

L'amélioration des performances est due alors aux contributions séparées des filtres optimaux purement linéaire et quadratique.

3.5. EXEMPLE D'APPLICATION, ÉTUDE D'UNE FAMILLE DE PROCESSUS OBSERVATIONS SUSCEPTIBLES DE METTRE EN ÉVIDENCE LE RÔLE DU FILTRAGE QUADRATIQUE

On s'intéresse, dans ce paragraphe, à des exemples d'application possibles du filtrage de Wiener linéaire-quadratique, c'est-à-dire, d'après ce qui précède, à des situations non gaussiennes. Il s'avère que, dans la pratique, l'écart à la loi gaussienne peut provenir de

la perte du caractère gaussien dans une non-linéarité. Aussi, dans ce qui suit, considérerons-nous un processus observation $x(t)$ généré par la chaîne présentée sur la figure 1.

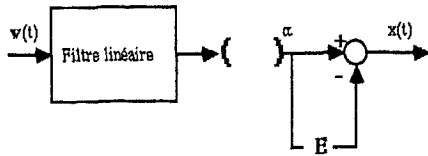


Fig. 1.

Le signal $w(t)$ est un processus aléatoire discret centré, gaussien, blanc, stationnaire et de variance σ^2 . Le filtre linéaire permet, selon les besoins, de modifier la corrélation de $w(t)$. L'élévation à la puissance entière α , avec $2 \leq \alpha$, détruit le caractère gaussien. L'espérance mathématique du signal ainsi obtenu est finalement retranchée afin d'obtenir une observation $x(t)$ centrée, dont les propriétés dépendent, à nature gaussienne d'entrée fixée, de σ , du filtre F , et de la puissance α . Les multiples combinaisons de ces trois paramètres conduisent à un très grand nombre de processus aléatoires dont les propriétés s'avèrent intéressantes, dans un contexte de filtrage linéaire-quadratique. Notre propos n'est pas de détailler les différentes facettes du procédé décrit sur la figure 1, ce qui sera envisagé dans un prochain article consacré au filtrage adaptatif linéaire-quadratique, mais plutôt de sélectionner un jeu particulier afin d'examiner les solutions des équations obtenues dans les paragraphes précédents.

On constate néanmoins sans peine, qu'en l'absence de filtre, une puissance impaire annule les moments d'ordre trois de $x(t)$, donc diagonalise le système (25). Nous nous plaçons sous ces hypothèses dans la suite. On pose

$$(38) \quad \alpha = 2q + 1, \quad \text{avec } 1 \leq q$$

$$(39) \quad \theta(n) = (2n)! / n! 2^n, \quad \text{avec } n \text{ entier tel que } 1 \leq n$$

En conservant les notations des relations (17), (18), (19), on a, ici,

$$(40) \quad m_2(i, k) = \delta(i, k) \sigma^{2(2q+1)} \theta(2q+1)$$

$$(41) \quad m_3(i, k, j) = 0$$

$$(42) \quad m_4(i, k, j, m) = \delta(i, k, j, m) \sigma^{4(2q+1)} (\theta(2(2q+1)) - 3\theta^2(2q+1)) + [\delta(i, k)\delta(j, m) + \delta(i, j)\delta(k, m) + \delta(i, m)\delta(j, k)] \sigma^{4(2q+1)} \theta^2(2q+1)$$

L'équation (25) fournit alors, après quelques calculs,

$$(43) \quad h^0(i) = s(i) / \sigma^{2(2q+1)} \theta(2q+1)$$

$$(44) \quad h^0(i, i) = s(i, i) / \sigma^{4(2q+1)} (\theta(2(2q+1)) - \theta^2(2q+1))$$

$$(45) \quad h^0(i, j) = s(i, j) / 2 \sigma^{4(2q+1)} \theta^2(2q+1)$$

L'équation (45) est valable naturellement pour i distinct de j . Les $s(i)$ et $s(i, j)$ sont respectivement les composantes linéaires et quadratiques de s_{1qP} ,

voir (20), (21) et (22). De plus, en appliquant la relation générale (31), l'amélioration des performances grâce au filtrage quadratique est accessible.

$$(46) \quad v_{1qP} = v_P - \gamma_{qP}^2$$

avec

$$(47) \quad \gamma_{qP}^2 = \left(\frac{1}{\sigma^{4(2q+1)} (\theta(2(2q+1)) - \theta^2(2q+1))} \right) \times \sum_{i=0}^{P-1} s^2(i, i) + \left(\frac{1}{2 \sigma^{4(2q+1)} \theta^2(2q+1)} \right) \sum_{i=0}^{P-1} \sum_{j<i} s^2(i, j)$$

L'inégalité de Schwartz, assure que le dénominateur du premier terme de (47) est toujours positif. Notre exemple devient encore plus éloquent quant à l'opportunité d'un filtrage quadratique si l'on considère que le signal $d(t)$ est obtenu par l'opération suivante à partir de $x(t)$, opération qui correspond éventuellement à une non linéarité de canal,

$$(48) \quad d(t) = x^2(t) - E\{x^2(t)\}$$

(on retranche une valeur moyenne dans $d(t)$ car ce signal a été supposé, au départ, centré). Compte tenu de (20), (21), (22), (43) à (48) on montre sans difficulté que

$$(49) \quad h^0(i) = s(i) = 0$$

$$(50) \quad h^0(i, j) = \delta(0, i) \delta(i, j)$$

$$(51) \quad v_{1qP} = v_P - E\{[x^2(t) - E\{x^2(t)\}]^2\}$$

Ce qui prouve que l'on est dans une situation où l'estimation est linéairement nulle et quadratiquement singulière. Le seul terme utile du filtrage, qui conduit, de plus, à une estimation à variance nulle, est quadratique.

4. Introduction à la prédiction linéaire-quadratique

On applique, dans ce paragraphe, les résultats généraux précédents au cas particulier de la prédiction linéaire-quadratique. L'accent est mis sur les spécificités de ce problème. Tous les résultats du paragraphe 3 sont applicables ici dès que l'on prend la précaution de poser $d(t) = x(t+1)$. La forme découplée du système donnant le prédicteur optimal et ses performances devient dès lors, d'après (24) et (25),

$$(52) \quad \begin{bmatrix} v_{1qP} \\ \mathbf{0}_P \\ \mathbf{0}_{qP} \end{bmatrix} = \begin{bmatrix} c(0, 0) & \mathbf{c}_P^t & \mathbf{c}_{qP}^t \\ \mathbf{c}_P & \mathbf{C}_P & \mathbf{C}_{q|1P}^t \\ \mathbf{c}_{qP} & \mathbf{C}_{q|1P} & \mathbf{C}_{qP} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a}_P \\ \mathbf{a}_{qP} \end{bmatrix}$$

avec

$$(53) \quad \mathbf{c}_{1qP} = E\{x(t+1) \mathbf{x}_{1qP}(t)\}$$

$$(54) \quad \mathbf{a}_{1qP} = -\mathbf{h}_{1qP}^0$$

$$(55) \quad c(0, 0) = E\{x^2(t+1)\}$$

Formellement, le système (52) est tout à fait analogue à la réunion des équations (24) et (25). Il existe néanmoins une différence fondamentale entre le filtrage de Wiener lq et la prédiction de même nature qu'il convient de préciser et qui est liée, du reste, au rôle que jouent les moments d'ordre trois. On a vu que l'annulation de ces derniers entraîne le découplage de (25), ou encore, l'optimisation séparée du terme linéaire et du terme quadratique. Des contraintes identiques sont plus « draconiennes » dans un contexte de prédiction. En effet, d'après (52) et (53), on constate alors que, sauf dans le cas singulier examiné au paragraphe 3.41 où la matrice C_{lqP} n'est pas inversible, le terme quadratique du filtrage, a_{qP} , s'annule, ce qui n'a rien de surprenant étant donné que les moments d'ordre trois reflètent la corrélation entre des observations linéaires d'une part et quadratiques de l'autre. Ce point conditionne sévèrement le choix de processus pour lesquels on observera une amélioration sensible des performances grâce à un filtrage quadratique. Les moments d'ordre trois devront être élevés. Si l'on se réfère au procédé donné sur la figure 2, en l'absence de filtrage, la puissance α doit être paire. Une étude plus approfondie de ces spécificités sera proposée dans un prochain article consacrée à l'extension de l'algorithme de Levinson à la prédiction lq.

5. Conclusion

Nous avons proposé, dans cet article, une approche du filtrage de Wiener linéaire-quadratique fondée sur un cadre de travail nouveau. Ce cadre est conditionné, par les moments d'ordre trois : leur nullité entraîne l'orthogonalité des espaces observations linéaire et quadratique dont il convient de tenir compte par une représentation découplée, dans le cas contraire, le caractère conjoint conféré à ces deux espaces se cristallise dans une représentation couplée. Tout au long de l'étude, ce rôle crucial des moments du troisième ordre a été mis en avant, tant dans la détermination du filtre optimal que dans l'étude quantitative de l'amélioration des performances. En particulier, il a été vu comment l'annulation de ces derniers découple l'optimisation de la partie linéaire et de la partie quadratique du filtre global et sépare la contribution de ces deux termes à la variance de l'erreur d'estimation. Les résultats obtenus sont généraux, la seule hypothèse retenue concerne l'existence des moments du processus sous jacent jusqu'à l'ordre quatre. On a montré, que lorsque les processus observation et inconnu ne sont pas conjointement gaussiens, la variance de l'erreur d'estimation diminue nécessairement quand on ajoute un filtrage quadratique. Le calcul explicite et général du terme de diminution est fondé sur le filtrage lq sous contrainte. Une famille de processus obtenue par filtrage linéaire puis élévation à une certaine puissance d'un processus gaussien blanc a été introduite afin de préciser la forme du filtre optimal dans un cas particulier, et d'examiner l'opportunité du terme quadratique : il existe des situations d'estimation linéaire nulle et quadratique singulière. Pour finir, la prédiction lq a été introduite comme filtrage de Wiener lq spécifique se détachant néanmoins de manière sensible de ce dernier lorsque les moments d'ordre trois sont nuls, dans la mesure où

cette contrainte annule toujours la partie quadratique du prédicteur.

Ce travail a été soutenu par la DRET sous le contrat ERE/SC/87/1394/DRET/DS/SR.

Manuscrit reçu le 24 mai 1988.

BIBLIOGRAPHIE

- [1] P. DUVAUT, *Contraste et Détection, Application à la Quantification et aux Filtres de Volterra optimaux pour la Détection, Thèse de Doctorat, Univ. Paris-Orsay, janvier 1987.*
- [2] J. L. LAWSON et G. E. UHLENBECK, *Threshold Signals*, New York, McGraw Hill, 1950.
- [3] W. A. GARDNER, A unifying view of second order measures of quality for signal classification, *IEEE, Trans. Comm.*, COM-28, June 1980, p. 807-816.
- [4] B. PICINBONO et P. DUVAUT, Linear-quadratic filters for detection and estimation, *IEEE Trans. on Inf. Theory*, IT-34, n° 2, March 1988, p. 304-311.
- [5] P. CHEVALIER et B. PICINBONO, Optimal linear-quadratic array for detection, *Proc. of ICASSP*, Glasgow, May 1989.
- [6] S. P. APPELBAUM, Adaptive arrays, *IEEE Trans. Ant.*, AP-24, Sept. 1976, p. 585-598.
- [7] H. COX, Optimum arrays and the Schwartz inequality, *Journ. Acout. of America*, 45, janv. 1969, p. 228-232.
- [8] B. PICINBONO et P. DUVAUT, *Contrast and Detection in Stochastic Processes in Underwater Acoustics*, C. BAKER éd., New York, Springer Verlag, 1986, p. 181-203.
- [9] H. L. VAN TREE, *Detection Estimation and Modulation Theory*, part 1, New York, Wiley, 1968.
- [10] A. BLANC-LAPIERRE et B. PICINBONO, *Fonctions Aléatoires*, Paris, Masson, 1981.
- [11] N. WIENER, *Non linear Problems in random Theory*, New York, Wiley, 1958.
- [12] P. EYKHOFF, Some fundamental aspects of parameter Estimation, *IEEE Trans. on Aut. Control*, AC-8, 1963, p. 347-357.
- [13] J. ZARZICKI, *Orthogonal Ladder form of non linear predictors of the Volterra Wiener class*, in *Theory and application of non linear control Systems*, Elsevier Publisher, North Holland, 1986.
- [14] H. CHIANG, Efficient implementation of Quadratic Filters, *IEEE Trans. on ASSP*, ASSP-34, n° 6, Dec. 1986, p. 1511-1528.
- [15] N. C. GALLAGHER, On a class of random processes, exhibiting, optimal non linear one-step predictor, *IEEE Trans. on Inf. Theory*, IT-28, No. 2, March 1982, p. 366-371.
- [16] T. KOH et J. POWERS, Second order Volterra filtering and its application to non linear system identification, *IEEE Trans. on ASSP*, 33, n° 6, Dec. 1985, p. 1445-1455.
- [17] C. E. DAVILA, Second order volterra filter with rapid convergence, *IEEE on ASSP*, 35, n° 9, Sept. 1987.
- [18] V. J. MATHEWS, A fast recursive Least-squares second order Volterra filter, *Proc. of ICASSP*, New York, 1988, p. 156-160.
- [19] V. J. MATHEWS et J. LEE, A fast Least-squares second order filter, *Proc. of ICASSP*, New York, 1988, p. 1583-1586.
- [20] T. KOH et J. POWERS, An adaptive Digital filter with lattice orthogonalization, *Proc. of ICASSP*, Boston, MA, 1983, p. 37-40.
- [21] B. PICINBONO, Détection et Estimation sous contraintes, 10^e Colloque GRETSI, Nice, 1985, p. 137-143.