



PREDICTION LINEAIRE-QUADRATIQUE, ETUDE D'UNE
RECURSIVITE SUR L'ORDRE, EXTENSION DE L'ALGORITHME
DE LEVINSON

PATRICK DUVAUT

Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Electricité, Plateau du
Moulon, 91190, Gif sur Yvette. Tel : 69 41 80 40.

RESUME

La prédiction à un pas d'un processus aléatoire discret utilisant un filtrage de Volterra discret linéaire-quadratique d'ordre fini de son passé est considérée. Le processus envisagé est à valeurs réelles et de moments finis jusqu'au quatrième ordre. L'approche proposée est fondée sur un cadre de travail totalement nouveau, introduisant la notion de vecteur observation linéaire-quadratique, qui tient totalement compte des spécificités du filtrage quadratique, en particulier, de la redondance qu'il introduit habituellement. Ce cadre conduit à une dérivation des équations normales étendues très proche du cas linéaire. De la même manière, il met en évidence le rôle de couplage joué par les moments du troisième ordre. Il est prouvé qu'un filtrage quadratique améliore toujours les performances d'un prédicteur linéaire, dès que le processus sous-jacent n'est pas gaussien. Une autre utilisation importante de ce cadre réside dans la recherche d'une récursivité sur l'ordre du meilleur prédicteur linéaire-quadratique. Cette récursivité conduit à étendre l'algorithme de Levinson. Bien que débouchant sur une méthode d'orthogonalisation classique, l'algorithme obtenu présente des caractéristiques nouvelles en relation avec le filtrage treillis multidimensionnel et les équations de Chandrasekhar. Pour finir, une mesure, liée au rang de déplacement d'une matrice, de la stationnarité d'un processus jusqu'au quatrième ordre est introduite.

SUMMARY

The problem of predicting a sample drawn from a discrete real valued zero mean process, with finite moments up to the fourth order, using a discrete Volterra linear-quadratic filtering of its past is considered. The derivation is based on a completely new framework, based on linear-quadratic spaces, which takes into account properly the contribution of quadratic observations and their redundancy. Relying on this framework, we present the extended Normal Equation in a fashion very close to the linear. We show how the third order moments introduce a coupling between linear and quadratic terms. Moreover, we justify how we always improve the performance of a linear predictor by adding a quadratic filtering as soon as the underlying process is non gaussian. In keeping with this idea, we give an interesting interpretation, as a quadratic form, of the amount by which the improvement is gained. In order to derive an order recursive solution of linear-quadratic problem, we extend the classical Levinson algorithm, in the case of a stationary process up to the fourth order. As a matter of fact, because of the contribution of quadratic observations, we still obtain an orthogonalization procedure but with new features related to the Chandrasekhar equations and multidimensional lattice filtering.



1. INTRODUCTION

Les approches récentes proposées en traitement du signal non linéaire consistent, pour beaucoup, à trouver des filtres linéaires-quadratiques discrets optimaux en moyenne quadratique. [1], [2], [3], [4], souvent sous des hypothèses restrictives concernant la non linéarité comme un filtrage linéaire suivi d'une quadrature. [1], ou encore supposant les données Gaussiennes. [4]. Il existe peu d'articles s'écartant du cadre Gaussien. La référence [2] donne les équations du filtre linéaire-quadratique optimal pour n'importe quel processus aléatoire réel pourvu que ses moments soient finis jusqu'au quatrième ordre. La référence [3] étend ces résultats au cas complexe. Bien que les présentations de [2] et [3] permettent d'éviter l'hypothèse Gaussienne, elles s'avèrent inutilisables si l'on désire introduire une récursivité sur l'ordre du filtre optimise ou encore écrire explicitement un opérateur dyadique linéaire-quadratique. L'écriture qui est à la base de l'obtention des algorithmes des moindres carrés récursifs, dans le cas linéaire. Pour finir, ces deux approches ne tiennent pas compte d'une forte redondance inhérente au filtrage quadratique. Une récursivité temporelle, fondée sur la méthode du gradient stochastique, est introduite dans [5], [6], mais toujours sous l'hypothèse Gaussienne ce qui rejette une utilisation en présence de moments du troisième ordre, comme c'est parfois le cas en acoustique sous marine, [7]. En dernier lieu, une forme d'extension de l'algorithme de Levinson au filtrage linéaire-quadratique est présentée dans [8], mais la procédure retenue ne fait pas appel directement aux données, comme dans le cas linéaire, mais plutôt à un espace de matrices, déduites des données, par un endomorphisme. Par conséquent, l'algorithme final n'est pas utilisable directement.

Les objectifs de cet article sont à classer en deux catégories. La première vise à introduire un cadre de travail nouveau et très général concernant le filtrage linéaire-quadratique discret, susceptible d'être utilisé dans de nombreux domaines, aussi facilement que le filtrage linéaire. La seconde s'intéresse à l'application de tels outils à l'étude théorique d'une récursivité sur l'ordre du prédicteur linéaire-quadratique optimal. Cette récursivité doit de plus conduire à une extension de l'algorithme de Levinson utilisant directement les données et par conséquent, très proche du cas classique. L'approche proposée ici est donc tout à fait différente de celle disponible dans [8]. Dans la mesure où l'implémentation efficace de filtres linéaires-quadratiques est désormais un sujet de préoccupation courant, [9], il paraît indispensable de consacrer du temps à l'établissement de bases théoriques pour le **filtrage non linéaire discret optimal et adaptatif**. Cet article peut être envisagé comme une contribution à ce genre de travaux, où la récursivité considérée est relative à l'ordre du filtre linéaire-quadratique. La récursivité temporelle fera, d'ailleurs, l'objet d'un autre article. Les notations, les différentes représentations des filtres linéaires-quadratiques et leurs propriétés statistiques seront détaillées dans un prochain article. De tels outils permettent d'appliquer le principe d'orthogonalité au problème de **prédiction linéaire-quadratique** et d'obtenir ainsi les **équations normales étendues** pour tout processus aléatoire réel de moments finis jusqu'au quatrième ordre.

La résolution du problème connexe de prédiction lq **sous contrainte** est l'occasion de montrer que, si les moments du troisième ordre sont non tous nuls, l'introduction d'un filtrage quadratique améliore nécessairement les performances du prédicteur linéaire. Comme dans le cas linéaire, la recherche d'une récursivité sur l'ordre se fait par **orthogonalisation** de l'espace observation adéquat. Celle-ci s'obtient ici en résolvant un autre problème connexe assimilé à une **prédiction vectorielle linéaire-quadratique à rebours**. La récursion finale met en oeuvre un **filtre treillis multidimensionnel** différent, par beaucoup d'aspects, des filtres treillis du filtrage linéaire multidimensionnel. Celui-ci avec la prédiction vectorielle est d'autre part l'occasion d'introduire une mesure de la stationnarité d'un processus jusqu'au quatrième ordre en généralisant la notion de **rang de déplacement d'une matrice**, [10].

2. PRINCIPALES NOTATIONS et ABBREVIATIONS UTILISEES

- $\mathbf{x}_{lq}(t-1)$ vecteur observation linéaire-quadratique élémentaire, volq, détaille $i+1$
- $\mathbf{h}_{lq}(i)$ vecteur déterministe linéaire-quadratique élémentaire, vdlq, détaille $i+1$
- $\mathbf{C}_{lq}(i, j)$ matrice de covariance de volq, détaille $[i+1, j+1]$
- $\mathbf{x}_{lqp}(t-1)$ vecteur observation linéaire-quadratique, volq, d'ordre P détaille $P(P+3)/2$

- \mathbf{h}_{lqp} vecteur déterministe linéaire-quadratique, vdlq, d'ordre P détaille $P(P+3)/2$
- \mathbf{C}_{lqp} matrice de covariance bloc linéaire-quadratique de volq, de taille $[P(P+3)/2, P(P+3)/2]$ susceptible d'être partitionnée en P^2 sous matrices
- \mathbf{f}_{lqp} vdlq intervenant dans le meilleur prédicteur linéaire-quadratique, plq, d'ordre P
- \mathbf{c}_{lqp}^2 variance de l'innovation du meilleur plq future d'ordre P

3. FILTRES de VOLTERRA LINEAIRES-QUADRATIQUES DISCRETS

3.1. Définition

Si $\mathbf{x}(t-1)$ représente un vecteur observation finissant à l'instant $t-1$ et dont les composantes sont extraites d'un processus aléatoire discret réel, nous définissons un filtrage de Volterra discret linéaire-quadratique, flq, de $\mathbf{x}(t-1)$ par

$$\mathbf{f}_{lq}(t-1) = \mathbf{h} + \mathbf{h}^T \mathbf{x}(t-1) + \mathbf{x}^T(t-1) \mathbf{H} \mathbf{x}(t-1) \quad (1)$$

où \mathbf{h} est un vecteur, \mathbf{H} une matrice et où t désigne la variable discrète "temps", élément de l'ensemble des relatifs, \mathbb{Z} . Si de plus P représente la taille du vecteur observation on obtient, compte tenu de $\mathbf{x}^T(t-1) = [x(t-1) \dots x(t-P)]$,

$$\mathbf{f}_{lq}(t-1) = \mathbf{h} + \sum_{i=1}^P h(i)x(t-i) + \sum_{i=1}^P \sum_{j=1}^P x(t-i)H(i, j)x(t-j) \quad (2)$$

où $h(i)$ et $x(t-i)$ sont respectivement les composantes de \mathbf{h} et $\mathbf{x}(t-1)$. $H(i, j)$ est le terme de \mathbf{H} situé à l'intersection de la ligne de rang i et de la colonne de rang j . Les filtres définis par (2) s'interprètent éventuellement comme des développements de Volterra infinis, (lorsqu'il existent, [11]), tronqués, de fonctions non linéaires du vecteur observation comme le Rapport de Vraisemblance en détection, ou encore la Régression en estimation. Cette interprétation doit cependant être considérée avec beaucoup de réserves dans la mesure où, en général, un développement de Volterra n'est pas un développement statistiquement orthogonal. Deux termes différents du développement sont, en effet, rarement orthogonaux, [11]. Par conséquent, la troncature d'une série infinie de Volterra ne représente en rien l'opération de projection orthogonale que requiert la recherche d'une optimalité en moyenne quadratique, [12]. Il est à noter, d'autre part, que les observations $x(t-i)x(t-j)$ avec j variant de 1 à i , ne portent pas d'autres informations que celles contenues dans les observations $x(t-i)x(t-j)$ avec j variant de i jusqu'à P , lorsque i prend toute valeur entre 1 et P . De plus, dans la mesure où il est possible de montrer, [2], que le filtre optimal en moyenne quadratique est nécessairement caractérisé par une matrice \mathbf{H} symétrique dans le cas réel, il n'y a aucune raison de conserver des termes redondants dans (2). Par conséquent, sans perte de généralité, nous ne manipulerons, dans la suite, que la partie symétrique des flq, (2), notés flqs et appelés filtres symétriques. Cette restriction est d'ailleurs faite très souvent, [8].

3.2. filtres symétriques

Pour les raisons évoquées ci-dessus, nous considérons uniquement dans la suite les filtres de la forme ci-dessous

$$\mathbf{sf}_{lqp}(t-1) = \mathbf{h} + \sum_{i=1}^P h(i)x(t-i) + \sum_{i=1}^P \sum_{j=1}^i x(t-i)H(i, j)x(t-j) \quad (3)$$

Il est intéressant de noter que si la taille de la partie linéaire est fixée à P , tout flqs défini par (3) dépend de $P(P+3)/2$ paramètres indépendants. Nous n'avons pas tenu compte ici de la contribution du scalaire h qui est fonction des autres paramètres, dans la pratique, afin de respecter certaines contraintes sur le filtre comme l'absence de biais, si le filtre est utilisé dans comme estimateur. Nous nous intéresserons dans la suite à la détermination des $(P+1)(P+4)/2$ paramètres optimaux du filtre d'ordre $P+1$ en fonction des $P(P+3)/2$ paramètres du filtre optimal d'ordre P . Ceci repose sur une forme plus appropriée de (3).

3.3. Représentation couplée d'un flqs

L'idée de base consiste à étendre la notion de vecteur observation, utilisée d'habitude dans un contexte linéaire, au filtrage linéaire-quadratique. Il est possible de réécrire (3) sous une forme totalement analogue à la sortie d'un filtre linéaire en introduisant des vecteurs linéaires-quadratiques élémentaires,



$\mathbf{x}_{iq}(t-i)$, de taille $i+1$, qui contiennent en plus de $\mathbf{x}(t-i)$ toutes les variables quadratiques construites à partir de $\mathbf{x}(t-i)$ et $\mathbf{x}(t-j)$, j variant entre 1 et i et un vecteur de pondération élémentaire \mathbf{l}_q . $\mathbf{h}_{iq}(i)$ lui aussi de taille $i+1$,

$$\mathbf{s}_{iqP}^t(t-1) = \mathbf{h} + \sum_{i=1}^P \mathbf{h}_{iq}^t(i) \mathbf{x}_{iq}(t-i) \quad (4)$$

Une forme encore plus compacte de (3) est accessible. en tant que produit scalaire Euclidien de deux vecteurs et qui s'avèrera très utile dans la suite

$$\mathbf{s}_{iqP}^t(t-1) = \mathbf{h} + \mathbf{h}_{iqP}^t \mathbf{x}_{iqP}(t-1) \quad (5)$$

4. PREDICTION FUTURE LINEAIRE-QUADRATIQUE

4.1. Introduction

Le propos de ce paragraphe est d'utiliser l'information contenue dans les quatre premiers moments d'un processus aléatoire discret pour prédire un échantillon extrait de ce même processus. Cette démarche se justifie dans la mesure où, lorsque le processus n'est pas Gaussien, le meilleur estimateur en moyenne quadratique d'un échantillon en fonction d'un nombre fini d'échantillons appartenant à son passé, n'a aucune raison d'être une fonction linéaire de ces observations. En fait, cet estimateur n'est autre que l'espérance mathématique de la variable estimée conditionnée par les observations. Par conséquent, chaque fois que la structure linéaire n'est pas optimale en termes d'estimation, on aura intérêt à améliorer ses performances en tentant de se rapprocher de l'optimum. L'utilisation des moments jusqu'au quatrième ordre est une possibilité. Elle suppose l'introduction d'observations quadratiques comme $\mathbf{x}(t-i)\mathbf{x}(t-j)$ dans l'estimateur lui-même. La question subsiste, quant au choix du moment le plus élevé retenu. Quatre peut sembler arbitraire. Il s'agit en fait de l'étape suivant le linéaire et donc de la première venant à l'esprit. Cependant, si la seule information disponible est la négation du caractère Gaussien cette question est cruciale et encore très ouverte.

4.2. Formulation du problème

Nous recherchons le meilleur estimateur de $\mathbf{x}(t)$, quel que soit l'instant t , en fonction de son passé "linéaire-quadratique" jusqu'à l'instant $t-P$. Dans la suite, P représentera l'ordre du prédicteur supposé fixe. Nous supposons également que le processus étudié est à valeurs réelles et de moments finis jusqu'au quatrième ordre. Cet estimateur s'écrit comme une combinaison linéaire des éléments de $\mathbf{X}_{iqP}(t-1)$. Par conséquent, nous recherchons un vecteur linéaire-quadratique \mathbf{g}_{iqP} qui rapproche le plus possible, au sens des moindres carrés, l'estimateur

$$\hat{\mathbf{x}}_{iqP}(t) = -\mathbf{g}_{iqP}^t \mathbf{X}_{iqP}(t-1) \quad (6)$$

de $\mathbf{x}(t)$. Le prime signifie que le vecteur correspondant à une valeur moyenne nulle. L'estimateur défini par (6) est un filq particulier de la forme (5) où le scalaire \mathbf{h} est égal à $\mathbf{g}_{iqP}^t E[\mathbf{X}_{iqP}(t-1)]$, afin de supprimer tout biais de (6).

4.3. Equations normales étendues

4.3.1. conséquences du principe d'orthogonalité

Nous présentons ici l'équation donnant le meilleur vdlq, noté \mathbf{f}_{iqP} , qui rapproche (6) le plus possible de $\mathbf{x}(t)$. Nous allons voir que cette équation revêt la même forme que l'équation classique de Yule-Walker. De plus, cette équation est, comme nous le verrons, plus générale dans la mesure où elle s'identifie à la classique lorsque les moments du troisième ordre sont nuls. Il est essentiel de noter également que tous les résultats obtenus dans ce paragraphe ne reposent sur aucune hypothèse particulière concernant la stationnarité du processus ou sa densité de probabilité. Si \mathcal{O}_{iqP} désigne l'espace vectoriel de variables aléatoires, engendré par (6) lorsque \mathbf{g}_{iqP} prend n'importe quelle valeur, alors, le meilleur estimateur linéaire-quadratique de $\mathbf{x}(t)$ n'est autre que la projection orthogonale, au sens statistique, de $\mathbf{x}(t)$ sur cet espace. Par conséquent, tout estimateur de la forme (6) est orthogonal à l'innovation future linéaire-quadratique \mathbf{i}_{iqP} , différence entre $\mathbf{x}(t)$ et le meilleur estimateur,

$$\mathbf{x}_{iqP}^t(t) = \mathbf{x}(t) + \mathbf{f}_{iqP}^t \mathbf{X}_{iqP}(t-1) \quad (7)$$

Ainsi, pour tout \mathbf{g}_{iqP} , nous devons avoir

$$\mathbf{g}_{iqP}^t (E[\mathbf{X}_{iqP}(t-1)\mathbf{x}(t)] + E[\mathbf{X}_{iqP}(t-1)\mathbf{X}_{iqP}^t(t-1)]\mathbf{f}_{iqP}) = 0 \quad (8)$$

Après quoi, (8) conduit à

$$\begin{bmatrix} \xi_{iqP}^t \\ \mathbf{0}_{iqP} \end{bmatrix} = \begin{bmatrix} c(0,0) & \mathbf{c}_{iqP}^t \\ \mathbf{c}_{iqP} & \mathbf{C}_{iqP} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{f}_{iqP} \end{bmatrix} \quad (9)$$

où ξ_{iqP}^t est la variance de (7). \mathbf{C}_{iqP} la matrice de covariance linéaire-quadratique, $\mathbf{c}_{iqP} = E[\mathbf{x}(t)\mathbf{X}_{iqP}(t-1)]$, $c(0,0) = E[\mathbf{x}(t)\mathbf{x}(t)]$, et où finalement $\mathbf{0}_{iqP}$ représente le vecteur nul linéaire-quadratique de taille $P(P+3)/2$. La versatilité de nos outils est mise ici en évidence. Formellement, l'équation (9), est tout à fait semblable aux équations de Yule-Walker du prédicteur linéaire. Cette similitude ne doit pas cacher pour autant la complexité inhérente à la prédiction linéaire-quadratique. \mathbf{C}_{iqP} est une matrice bloc $[P(P+3)/2, P(P+3)/2]$, qui peut être partitionnée en P^2 sous matrices ou interviennent les moments du deuxième, troisième et quatrième ordre. De la même manière, \mathbf{c}_{iqP} , défini ci-dessus, est un vecteur bloc à $P(P+3)/2$ composantes qui peut aussi être partitionné en P sous vecteurs de taille $i+1$

$$\mathbf{c}_{iq}(i,0) = E[\mathbf{x}_{iq}(t-i)\mathbf{x}(t)] \quad (10)$$

dont la $k^{\text{ème}}$ composante, $[\mathbf{c}_{iq}(i,0)]_k$, est égale à $m_3(t-i, t-(i-k+2), t)$, si $2 \leq k \leq i+1$ et $m_2(t-i, t)$ pour $k=1$. Nous déduisons de (9) que si le processus n'est pas parfaitement prédictible au sens linéaire-quadratique, l'inverse de \mathbf{C}_{iqP} existe et

$$\mathbf{f}_{iqP} = \mathbf{C}_{iqP}^{-1} \mathbf{c}_{iqP} \quad (11)$$

$$\xi_{iqP}^t = c(0,0) - \mathbf{c}_{iqP}^t \mathbf{C}_{iqP}^{-1} \mathbf{c}_{iqP} \quad (12)$$

Des solutions d'équations semblables à (9) utilisées dans un contexte de détection sont fournies dans les références [2] et [3] pour différents environnements. Notre propos est autre ici, puisqu'il s'agit de rechercher une récursivité sur l'ordre de la solution optimale. Avant ceci, il convient d'étudier un peu plus en détail la contribution des moments du troisième ordre qui jouent un rôle privilégié.

4.3.2. Rôle joué par les moments du troisième ordre

Il serait intéressant de réécrire (12) comme la somme de deux termes, correspondant d'un côté à la contribution du filtrage linéaire et de l'autre, à la contribution du filtrage quadratique. Ceci n'est pas possible cependant dans la mesure où les moments du troisième ordre sont responsable d'un couplage entre les deux types de filtrage. Un point semblable a déjà été noté dans la référence [2] mais dans un contexte de détection. Afin de détailler davantage ce couplage, il convient d'introduire la seconde représentation des filqs, encore appelée "représentation découplée". Cette représentation ne signifie pas que la matrice intervenant dans (9) a été diagonalisée mais seulement que les contributions de la partie linéaire et de la partie quadratique interviennent séparément.

Il est donc important, à ce stade, de ne pas confondre le découplage de la solution du problème qui suppose, comme nous l'avons dit, la nullité des moments du troisième ordre et le simple découplage de la représentation, qui ne suppose rien sur les moments, mais permet de voir séparément l'intervention du terme linéaire et du terme quadratique dans la formulation du problème. Cette nuance est clairement mise en évidence sur la deuxième forme de (9)

$$\begin{bmatrix} \xi_{iqP}^t \\ \mathbf{0}_P \\ \mathbf{0}_{qP} \end{bmatrix} = \begin{bmatrix} c(0,0) & \mathbf{c}_P^t & \mathbf{c}_{qP}^t \\ \mathbf{c}_P & \mathbf{C}_P & \mathbf{C}_{q|P} \\ \mathbf{c}_{qP} & \mathbf{C}_{q|P} & \mathbf{C}_{qP} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{f}_P \\ \mathbf{f}_{qP} \end{bmatrix} \quad (13)$$

En effet, à la différence de (9), les participations respectives du filtre linéaire \mathbf{f}_P et du filtre quadratique \mathbf{f}_{qP} sont nettement séparées. Les matrices intervenant dans (13) sont explicitées dans la note (A1.26). Nous déduisons de (13) que le couplage de la solution optimale résulte de la matrice $\mathbf{C}_{q|P}$ qui contient seulement des moments d'ordre trois. Si l'on suppose que les deux matrices \mathbf{C}_P



et C_{qp} ne sont pas singulières on obtient une forme intéressante de la variance de l'innovation linéaire-quadratique,

$$\varepsilon_{lqP}^t = \varepsilon_{lP}^t - C_{lqP}^t C_{lqP}^{-1} C_{lqP} C_{lqP}^{-1} C_{lqP} - \gamma_{lqP} - \eta_{lqP} \quad (14)$$

ou

$$\varepsilon_{lP}^t = c(0, 0) - C_{lP}^t C_{lP}^{-1} C_{lP}$$

représente la variance de l'innovation linéaire qui interviendrait en l'absence de moments du troisième ordre. Les deux derniers termes sont des termes de couplage faisant intervenir des moments d'ordre trois.

$$\gamma_{lqP} = C_{lqP}^t C_{lqP}^{-1} C_{lqP} R_P [C_{lqP}^t C_{lqP}^{-1} C_{lqP} - C_P] \quad (15)$$

η_{lqP} se déduit de (15) en permutant le rôle de q et l . R_P and R_{qP} sont les deux matrices situées sur la diagonale principale de la version partitionnée de C_{lqP}^{-1} . On établit que les 3 derniers termes de (14) apportent toujours une contribution négative au premier si les moments d'ordre trois ne sont pas nuls. Cela signifie que le filtrage quadratique améliore les performances de prédiction, dès que le processus sous-jacent n'est pas gaussien. Ce résultat est obtenu en résolvant un problème connexe, celui de la prédiction **linéaire-quadratique sous contraintes**. Il est possible alors de réécrire les trois derniers termes de (14) comme l'opposé d'une forme quadratique, ce qui conduit naturellement au résultat annoncé.

4.4. Problèmes reliés à la décomposition orthogonale de l'espace observation servant à la prédiction linéaire-quadratique de $x(t)$

Le meilleur estimateur de $x(t)$ en fonction de son passé linéaire-quadratique jusqu'à l'instant $t-P-1$ n'est autre que la projection orthogonale de $x(t)$ sur l'espace $X'_{lqP+1}(t-1)$.

$$X'_{lqP+1}(t-1) = \{ X'_{lqP}(t-1) ; x'_{lq}(t-P-1) \} \quad (24)$$

La formulation récursive ci-dessus signifie que la différence déterminante entre la mise à jour d'espaces observations linéaires et quadratiques vient du fait que dans le premier cas on ajoute un scalaire alors que dans le second un vecteur est nécessaire. Cette difficulté pourrait éventuellement être atténuée en tentant d'établir un lien avec la mise à jour d'espaces observations impliqués dans la prédiction linéaire de processus vectoriels où il est aussi nécessaire d'ajouter un vecteur plutôt qu'un scalaire. L'analogie s'arrête là cependant, car le vecteur intervenant dans (24) voit sa taille reléguée à l'ordre ce qui n'est pas du tout le cas de la prédiction vectorielle linéaire. Par conséquent, ce problème nécessite une attention toute particulière et de nouvelles approches. Nous verrons, en fait, que le "salut" vient d'une question connexe qu'il convient d'appeler **prédiction linéaire-quadratique vectorielle à rebours**.

4.5. Formulation récursive du prédicteur linéaire-quadratique optimal

Nous présentons dans ce paragraphe la formulation récursive, suivant l'ordre, du meilleur prédicteur linéaire-quadratique quand on passe de l'ordre m à l'ordre $m+1$. Il s'agit seulement ici d'une formulation et non d'une récursion complète qui nécessite, comme nous le verrons, quelques étapes supplémentaires. Nous appelons

$$x'_{lq|m}(t-m-1) = x'_{lq}(t-m-1) - \text{PROJ}[x'_{lq}(t-m-1)/X'_{lqm}(t-1)] \quad (25)$$

l'innovation rétrograde d'ordre m du vecteur linéaire-quadratique $x'_{lq}(t-m-1)$. L'indice m indique qu'il faut m vecteurs lq pour écrire la projection sur $X'_{lqm}(t-1)$. Le b quant à lui fait référence à la nature rebours de la prédiction, en ce sens que les vecteurs de $X'_{lqm}(t-1)$ appartiennent au "futur linéaire quadratique" de $x'_{lq}(t-m)$. Un point essentiel à souligner ici, provient du fait que les vecteurs $x'_{lq}(t-i)$, avec $1 \leq i \leq m$, intervenant dans la prédiction de $x'_{lq}(t-m-1)$ ont des tailles différentes. Ce problème sera appelé dans la suite **prédiction vectorielle linéaire-quadratique à rebours**. Il est possible d'écrire en fait

$$\text{PROJ}[x'_{lq}(t-m-1)/X'_{lqm}(t-1)] = -B_{m+1|m}^t x'_{lqm}(t-1) \quad (26)$$

ou

$$B_{m+1|m}^t = [B_{m+1|m}^t(1) \dots B_{m+1|m}^t(i) \dots B_{m+1|m}^t(m)]$$

ce qui conduit à la forme développée

$$x'_{lq|m}(t-m-1) = - \sum_{i=1}^m B_{m+1|m}^t(i) x'_{lq}(t-i) \quad (27)$$

Nous déduisons de (27) que $B_{m+1|m}^t(i)$ est une $[m+2, i+1]$ matrice et que $B_{m+1|m}^t$, sans argument, est de taille $[m+2, m(m+3)/2]$. L'indice m fait donc référence à l'ordre de prédiction alors que $m+1$ précise l'instant prédit. Si m est égal à zéro l'innovation (25) devient égale au vecteur lui-même, à l'instant de la prédiction linéaire. Le principe d'orthogonalité entraîne que (25) est orthogonal à $X'_{lqm}(t-1)$. Par conséquent, il est possible d'annuler une décomposition orthogonale de $x'_{lqm+1}(t-1)$ qui mène à

$$\text{PROJ}[x(t)/X'_{lqm+1}(t-1)] =$$

$$\text{PROJ}[x(t)/X'_{lqm}(t-1)] + \text{PROJ}[x'_{lq|m}(t)/x'_{lq|m}(t-m-1)] \quad (28)$$

Le second terme de (28) s'obtient en écrivant que $x(t)$ peut s'interpréter comme la somme de sa projection sur $X'_{lqm}(t-1)$ et de son innovation d'ordre m . En enlevant (28) de $x(t)$ il reste la formulation récursive annoncée

$$x'_{lq|m+1}(t) = x'_{lq|m}(t) - K_{m+1}^t x'_{lq|m}(t-m-1) \quad (29)$$

$$K_{m+1}^t = E[x'_{lq|m}(t) x'_{lq|m}(t-m-1)] [E[x'_{lq|m}(t-m-1) x'_{lq|m}(t-m-1)]]^{-1} \quad (30)$$

(30) est une propriété classique d'estimation en moyenne quadratique, (30) n'est autre, en fait, que la matrice optimale associant (25) au meilleur estimateur de l'innovation d'ordre m de $x(t)$. Dans ce cas précis, cette matrice est de taille $[1, m+2]$.

4.6. Mise à jour du prédicteur vectoriel linéaire-quadratique à rebours

Nous proposons ici une mise à jour de (25). Pour se faire nous considérons

$$x'_{lq|n}(t-m-1+n) =$$

$$x'_{lq}(t-m-1+n) - \text{PROJ}[x'_{lq}(t-m-1+n)/\{x'_{lq}(t-m-1) \dots x'_{lq}(t-m-n-2)\}] \quad (31)$$

l'innovation vectorielle future d'ordre n associée au vecteur $x'_{lq}(t-m-1+n)$, avec $1 \leq n \leq m$. Comme à l'accoutumée, lorsque n est nul l'innovation s'identifie au vecteur. Prenant en compte ce nouveau problème et introduisant la décomposition orthogonale des espaces observations appropriés, il vient

$$x'_{lq|n}(t-m-1) = x'_{lq|n-1}(t-m-1) - K_{n-1}^b(m+1) x'_{lq|n-1}(t-m-1+n)$$

$$x'_{lq|n}(t-m-1+n) = x'_{lq|n-1}(t-m-1+n) - K_n^b(m+1) x'_{lq|n-1}(t-m-1) \quad (32)$$

ou

$$K_n^b(m+1-n) = R_{n-1}(m+1-n) [E_{n-1}^b(m+1-n)]^{-1} \quad (33)$$

$$R_{n-1}(m+1-n) = E[x'_{lq|n-1}(t-m-1) x'_{lq|n-1}(t-m-1+n)] \quad (34)$$

$$E_{n-1}^b(m+1-n) = E[x'_{lq|n-1}(t-m-1+n) x'_{lq|n-1}(t-m-1+n)] \quad (35)$$

$$K_n^b(m+1) = R_{n-1}^b(m+1-n) [E_{n-1}^b(m+1)]^{-1} \quad (36)$$

$$E_{n-1}^b(m+1) = E[x'_{lq|n-1}(t-m-1) x'_{lq|n-1}(t-m-1)] \quad (37)$$



Les équations (32) ont exactement la même forme que celles intervenant dans la formulation réursive de la prédiction linéaire vectorielle. En d'autres termes, la cellule de base de la récursion (32) n'est autre que la cellule élémentaire d'un filtre treillis multidimensionnel, caractérisée par des matrices dites de "réflexion", définies par (33) et (36). L'innovation disponible à la sortie de la cellule n-1 fait intervenir l'instant t-m-1+n-1 et non l'instant t-m-1+n. Or la taille d'un vecteur linéaire-quadratique change avec son argument, par conséquent, une récursivité totale exige le calcul de

$$\tilde{x}_{1q, n-1}^f(t-m-1+n) \text{ en termes de } \tilde{x}_{1q, n-1}^f(t-m-1+n-1) \quad (38)$$

Ce calcul doit être vu comme une mise à jour temporelle seulement, dans la mesure où les deux innovations de (38) sont du même ordre. Cette dérivation est possible simplement, si l'on considère que le processus étudié est stationnaire jusqu'au quatrième ordre. En effet, cette hypothèse ajoutée à la forme générale des innovations,

$$\tilde{x}_{1q, n-1}^f(t-m-1+n-1) =$$

$$\tilde{x}_{1q}^f(t-m-1+n-1) + \sum_{i=1}^{n-1} A_{m+2-n, n-1}(i) \tilde{x}_{1q}^f(t-m-1+n-1-i) \quad (39)$$

(Il suffit de remplacer n-1 par n dans les observations ci-dessous et m+2-n par m+1-n pour obtenir l'autre innovation. On note que $A_{m+2-n, n-1}(i)$ est une matrice [m+3-n, m+3+i-n]. Il suffit aussi de décrémenter le nombre de lignes et de colonnes précédents pour obtenir la taille de $A_{m+1-n, n-1}(i)$, conduit à la mise à jour suivante

$$A_{m+1-n, n-1}(i) = U_{m+2-n} A_{m+2-n, n-1}(i) U_{m+2-n}^t + D_{m+1-n, n-1}(i) \quad (40)$$

où $U_\alpha = [I_{\alpha, \alpha} \quad 0_\alpha]$. ($I_{\alpha, \alpha}$ est la matrice identité de taille $[\alpha, \alpha]$). Cela signifie qu'à un terme additif près, on calcule les matrices intervenant dans la prédiction d'ordre n-1 de $\tilde{x}_{1q}(t-m-1+n)$ en enlevant la dernière ligne et la dernière colonne des matrices caractérisant l'innovation de $\tilde{x}_{1q}(t-m-1+n-1)$. Ceci est une conséquence de la stationnarité du processus. Si par hasard le processus n'est pas stationnaire, il n'existe pas de relation aussi simple que (40), mais il est possible alors d'introduire une mesure de la stationnarité d'un processus jusqu'au quatrième ordre étendant ainsi la notion de rang de déplacement d'une matrice. Pour finir on déduit, toujours sous la stationnarité, une mise à jour des matrices variances des innovations, qui n'est pas sans rappeler (40).

$$E_{n-1}^f(m+1-n) = U_{m+2-n} E_{n-1}^f(m+2-n) U_{m+2-n}^t - W_{n-1}(m+1-n) \quad (41)$$

Dans la suite, nous noterons symboliquement la mise à jour temporelle traduite par (40) et (41) de la façon suivante

$$\tilde{x}_{1q, n-1}^f(t-m-1+n) = Z_{1q} \tilde{x}_{1q, n-1}^f(t-m-1+n-1) \quad (42)$$

on déduit alors le schéma de principe de la récursion globale, concernant le problème initial de prédiction réursive de x(t) voir figure 1.

5. CONCLUSION

Nous avons introduit dans cet article un nouveau cadre de travail pour l'étude des filtres de Volterra discrets qui conduit à une formulation et une résolution inédite du problème de prédiction linéaire-quadratique, très proche du cas linéaire. De plus, nous avons montré que les performances du prédicteur linéaire sont améliorées, chaque fois que les moments du troisième ordre ne sont pas nuls et que l'amélioration des performances s'interprète toujours en termes d'une forme quadratique déduite de l'étude du problème connexe de prédiction lq sous contraintes. La question cruciale d'une dérivation réursive, sur l'ordre, du meilleur prédicteur lq a été investie conduisant à une extension de l'algorithme de Levinson qui met en évidence des caractéristiques importantes. Parmi celles-ci figure un filtrage treillis

filtrage treillis multidimensionnel à rapprocher du filtrage intervenant dans la prédiction linéaire de processus vectoriels.

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

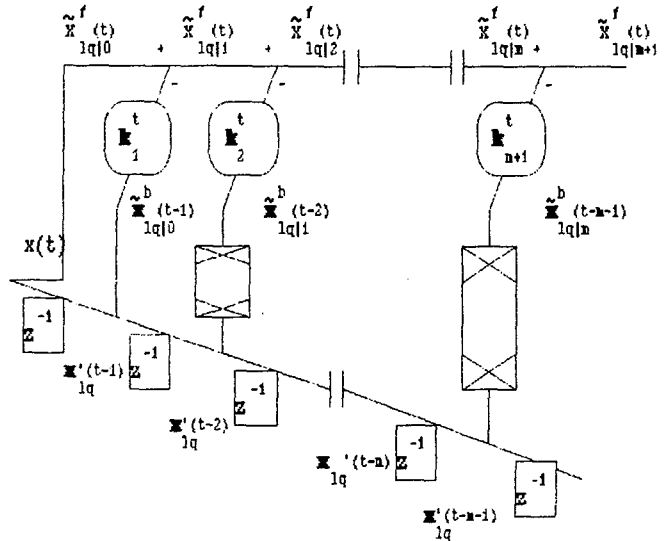


figure 1.

REFERENCES

- [1] T.E. Mc CANNON et N.C. GALLAGHER, "On the design of Nonlinear Discrete-time predictors", IEEE trans. on Inf. theory, Vol. IT 28, No.2, pp. 366-371, March 1982.
- [2] B. PICINBONO et P. DUVAUT, "Linear-quadratic filters for detection and Estimation", IEEE trans. on Inf. theory, Vol. IT, No. 1, pp. 304-311, March 1988.
- [3] B. PICINBONO, "Complex linear quadratic filters for array designing", in Proc. ICASSP-88, New York, N.J., April 88.
- [4] T. KOH et J. POWERS, "Second order Volterra filtering and its applications", Internal report, Univ. Texas at Austin, Dept. El. E., 1984.
- [5] T. KOH et J. POWERS, "An adaptive digital filter with lattice orthogonalization", in Proc. ICASSP-83, Boston, MA, pp. 37-40, 1983.
- [6] T. KOH et J. POWERS, "Second order Volterra filtering and its application to nonlinear System Identification", IEEE Acoust., Speech and Signal Processing, Vol. ASSP-33, No. 6, pp. 1445-1455, December 1985.
- [7] R. J. URICK, Principles of underwater sound, Mc. Graw Hill Book Company, 1975.
- [8] JANZARZYCKI, "Orthogonal ladder-form representations of Nonlinear predictors of the Volterra-Wiener class", in theory and Application of Nonlinear Control Systems, by C. J. Byrnes and A. Lindquist, Elsevier publisher, North Holland, 1986.
- [9] C. NIKIAS et Y. LOU, "VLSI array processing structures of quadratic digital filters with LMS Algorithm", Internal report, Communications and digital signal processing laboratory, Northeastern University, Boston, 1987.
- [10] T. KAILATH et M. MORF, "Displacement rank of a matrix", Bull. Am. Math. Soc., Vol. 1, pp. 769-773, 1979.
- [11] P. DUVAUT et B. PICINBONO, "Détection Estimation et filtres de Volterra", congrès GRETSI, Nice, France, 1987.
- [12] P. DUVAUT, "Contraste et détection application à la quantification et aux filtres de Volterra optimaux pour la détection et l'estimation" thèse de doctorat, University of Paris 11, January 1987.

