

Influence des moments d'ordre trois sur la structure d'un algorithme adaptatif linéaire quadratique

About the influence of third order moments upon the structure of an adaptive linear quadratic algorithm



P. DUVAUT

Laboratoire de Traitement des Images et du Signal, ENSEA, Les chênes Pourpres, 95000 Cergy-Pontoise, France

Société Kurtosis Ingénierie, 5 avenue du Parc, Le Campus, 95032 Cergy-Pontoise, France

Ancien Elève de l'ENS de Cachan, agrégé de Sciences Physiques en 1981, Docteur en Sciences de l'Université d'Orsay en 1987, il dirige le Département Mathématiques et Traitement du Signal de l'École Nationale Supérieure de l'Électronique et de ses Applications, ENSEA, à Cergy-Pontoise. Il poursuit des activités de recherche au sein de l'équipe de traitement des Images et du Signal, ETIS, de l'ENSEA. Ses principaux thèmes de recherche concernent le traitement des signaux Radar (en relation avec le Département DEM de la Thomson à Malakoff), l'utilisation des moments d'ordre supérieur en Détection et en Estimation, les relations entre non stationnaire, non linéaire et non gaussien. Animateur du Conseil Scientifique de la société Kurtosis Ingénierie, il co-dirige la collection Traitement du Signal des Éditions Hermès.

RÉSUMÉ

L'implémentation de tout algorithme adaptatif linéaire quadratique nécessite, au préalable, afin d'utiliser au mieux les informations a priori du processus observation, d'estimer l'importance relative des moments d'ordre trois par le calcul du skew et du kurtosis. Si ces moments sont nuls, c'est-à-dire pratiquement faibles, la structure adaptée de tout algorithme adaptatif LQ, LMS ou RLS, standard et rapide, est « découplée », au sens où deux procédures stochastiques indépendantes, de pas et de gain différents, identifient récursivement et respectivement les noyaux linéaire et quadratique du filtre optimal. Si ces mêmes moments sont d'intensité non négligeable, la structure de ces algorithmes devient couplée : une seule procédure stochastique, caractérisée par un seul pas et un gain unique, rafraîchit simultanément et de façon conjointe les noyaux linéaire et quadratique. Ces considérations se déduisent de la façon dont s'interprète tout algorithme stochastique, à savoir comme estimée d'une procédure récursive déterministe du type gradient ou Newton-Raphson. Vu qu'un filtre de Volterra transverse reste une fonction linéaire des paramètres, les propriétés de l'algorithme LMSLQ

demeurent très semblables, en dehors de la participation supplémentaire des moments d'ordre trois et quatre, à celles de l'algorithme LMS classique. En particulier, l'étude de la convergence fondée sur la théorie de la Mindépendance, introduite dans le contexte linéaire, s'étend sans difficulté. En présence de moments d'ordre trois, l'inadéquation de structure d'un algorithme LMSLQ, provenant d'une mauvaise utilisation des informations a priori, se modélise, commodément, par un bruit qui vient se superposer au bruit de modèle, appelé bruit « d'inadéquation ». La détermination de la variance de ce dernier montre à quel point ses effets s'avèrent néfastes sur le régime permanent de l'algorithme.

MOTS CLÉS

Filtrage de Volterra transverse, linéaire quadratique, méthode de la plus grande pente, méthode de Newton-Raphson, algorithme LMS, algorithme RLS, convergence, théorie de la Mindépendance, bruit de modèle, bruit de mauvaise adaptation, moments d'ordre trois.

SUMMARY

In order to fit properly the a priori information about the observation process, the implementation of any linear quadratic adaptive filter requires to estimate the influence of the third order moments, by the means of the skew and the kurtosis. If these moments are vanishing, the proper structure of such algorithms is uncoupled, that is to say, two different algorithms update on the one hand the linear kernel and on the other hand the quadratic. If these same moments are not equal to zero, the suited structure is coupled : a unique algorithm jointly updates both linear and quadratic kernels. These results stems from the fact that any adaptive algorithm may be viewed as a stochastic estimate of a deterministic gradient type or Newton type procedure. Since Volterra filtering is still linear with respect to the parameters, the properties of the linear quadratic LMS algorithm are

very similar to these of the classical. In particular, the approach of the convergence based on the « Mary » independence theory remains mainly unchanged. The mismatch in the structure of the LMSLQ algorithm is modelled, when the third order moments are not vanishing, by a noise called the mismatch noise. The explicit derivation of its variance exhibits how the steady state behavior of the algorithm may be damaged by an unfitness of the structure.

KEY WORDS

Transverse Volterra filtering, linear quadratic, Steepest descent method, Newton-Raphson method, LMS and RLS algorithms, convergence, Mary independence theory, model noise, mismatch noise, third order moments.

1. Introduction

La prise en compte des moments d'ordre trois d'un processus non gaussien stationnaire [1], [2], ou l'utilisation de la covariance d'un signal aléatoire non stationnaire [3], [4], conduit inévitablement à la mise en œuvre de filtres linéaires quadratiques [5], [6]. L'optimalité, au sens d'un critère de distance, de tels dispositifs a déjà été abordée que les données soient réelles [6], [7], ou complexes [8]. Il apparaît, de façon synthétique, que les noyaux linéaire et quadratique du filtre optimal sont solutions de deux équations découplées si et seulement si les moments d'ordre trois du processus observation sont nuls [6], [7]. Les performances d'estimation, quant à elles, sont toujours améliorées dès que le processus inconnu et les données ne sont pas conjointement gaussiens [7]. L'identification et le fonctionnement de tels filtres ne vont pas sans poser de nombreux problèmes dont le plus crucial est, sans doute, la complexité liée à la présence du noyau quadratique. De ce fait, et plus encore que dans le cas linéaire, l'élaboration d'algorithmes récursifs d'implémentation, efficaces, dont la complexité est proportionnelle au nombre de paramètres inconnus, s'avère indispensable. Ce souci est permanent dans les références [9] et [10] qui étendent respectivement les algorithmes LMS et RLS rapide aux filtres transverses linéaires quadratiques. L'approche utilisée dans [9] par Koh et Powers, manque néanmoins, quelque peu, de généralité. Tout d'abord, les données y sont supposées gaussiennes, ce qui diminue partiellement l'intérêt d'utiliser un filtre non linéaire, et conduit, par voie de conséquence, à une structure très particulière d'algorithme dans la mesure où les moments d'ordre trois sont nuls. Une étude de la convergence y est proposée ensuite, mais fondée sur l'hypothèse d'indépendance des vecteurs observations successifs, ce qui est, en dehors du filtrage adaptatif d'antenne à capteurs indépendants [11], tout à fait irréaliste [12]. Dans la référence [10], Mathews et Lee, proposent un algorithme RLS rapide qui ne tient pas compte des modifications éventuelles de structure liées aux moments d'ordre trois, leur approche est d'autre part entièrement algébrique et la sortie du filtre mis en œuvre n'est pas centrée.

Il ressort donc de toutes ces dernières considérations que la structure d'un algorithme adaptatif « linéaire quadratique », qu'il s'agisse d'un LMS ou d'un RLS classique ou rapide, adaptée aux propriétés statistiques des données n'est pas clairement définie et que l'évaluation de la dégradation des performances d'une structure particulière due uniquement à sa mauvaise adaptation, fait totalement défaut.

Le présent article se propose de contribuer à éclaircir ces points. Après une exposition du problème abordé, en tant que question de filtrage optimal non linéaire récursif, en relation avec la référence [7], le 3^e § répertorie les structures possibles de tout algorithme adaptatif linéaire quadratique et précise les conditions du choix de structure à opérer en fonction des moments d'ordre trois des données. Après une étude, fondée sur la théorie de la « Mindépendance » [12], des propriétés de l'algorithme LMS linéaire quadratique, l'inadéquation, éventuelle, de

ses structures possibles avec les propriétés statistiques des données, est modélisée, en présence de moments d'ordre trois, dans le § 4, par un bruit, baptisé « bruit de mauvaise adaptation ». Un calcul de sa variance y est proposé quelles que soient les propriétés statistiques de données. La fin de ce paragraphe est consacrée à la mise en valeur des idées développées tout au long de l'article, notamment le rôle des moments d'ordre trois, sur une application particulière d'identification d'un filtre linéaire quadratique, l'accent étant mis, non pas sur l'inadéquation du modèle mais plutôt sur les effets liés à la seule inadéquation des algorithmes lorsque le modèle recherché est optimal.

Parmi les travaux récents développant des applications du filtrage linéaire quadratique, on peut citer les références [13] et [14]. La première est consacrée à l'adaptation de « dictionnaire » en reconnaissance de la parole fondée sur les coefficients cepstraux [15], et la seconde évalue l'amélioration, due à un terme supplémentaire quadratique, des performances d'antennes adaptatives sous contrainte.

2. Position du problème

Les notations utilisées sont celles de la référence [7] et plus particulièrement de son § 2. Le problème abordé, dans la suite, présente toutes les caractéristiques « classiques » du traitement du signal non linéaire optimal récursif, dont les schémas de principe sont exposés sur les figures 1 et 2. Entre ces deux figures seul le point de vue change. Le schéma de la figure 1 suppose que toutes les propriétés statistiques, notées symboliquement K dans ce qui suit, des processus mis en jeu sont connues sans erreur ou postulées, il fait référence au point de vue, qu'il convient d'appeler, par nature, « théorique ». A contrario, le schéma de la figure 2 fait référence, par son réalisme, au point de vue « pratique » dans la mesure où les mêmes grandeurs statistiques y sont cette fois estimées et symbolisées, en toute cohérence, par \hat{K} . La distinction n'a rien de fortuit, bien au contraire, puisqu'elle permet, comme il sera vu au § 4, de définir l'optimalité d'une structure spécifique d'algorithme stochastique lq (linéaire quadratique). Il s'agit de rechercher de façon itérative, dans le temps, les $P(P+3)/2$ échantillons de la réponse impulsionnelle optimale, mise sous forme d'un vecteur lq , voir [7] § 2, h_{lqP}^o , d'un filtre de Volterra transverse lq à horizon fini, dont la sortie minimise la distance en moyenne quadratique à un signal « désiré », $d(t)$, centré à valeurs réelles, et dont l'entrée se résume, pour finir, à un processus observation $x(t)$ à valeurs réelles de moments finis jusqu'à l'ordre 4. Les propriétés h_{lqP}^o ainsi que les performances du filtre correspondant sont données dans [7]; La nature du calcul itératif de h_{lqP}^o est conditionnée par le point de vue. Lorsque ce dernier est théorique le processus récursif mis en œuvre est déterministe, au sens où, il génère à l'instant t , un vecteur lq déterministe $h_{lqP}(t)$, voir figure 1, selon un mode caractérisé par une relation de récurrence de la forme

$$(1) \quad h_{lqP}(t) = A(h_{lqP}(t-1), K)$$

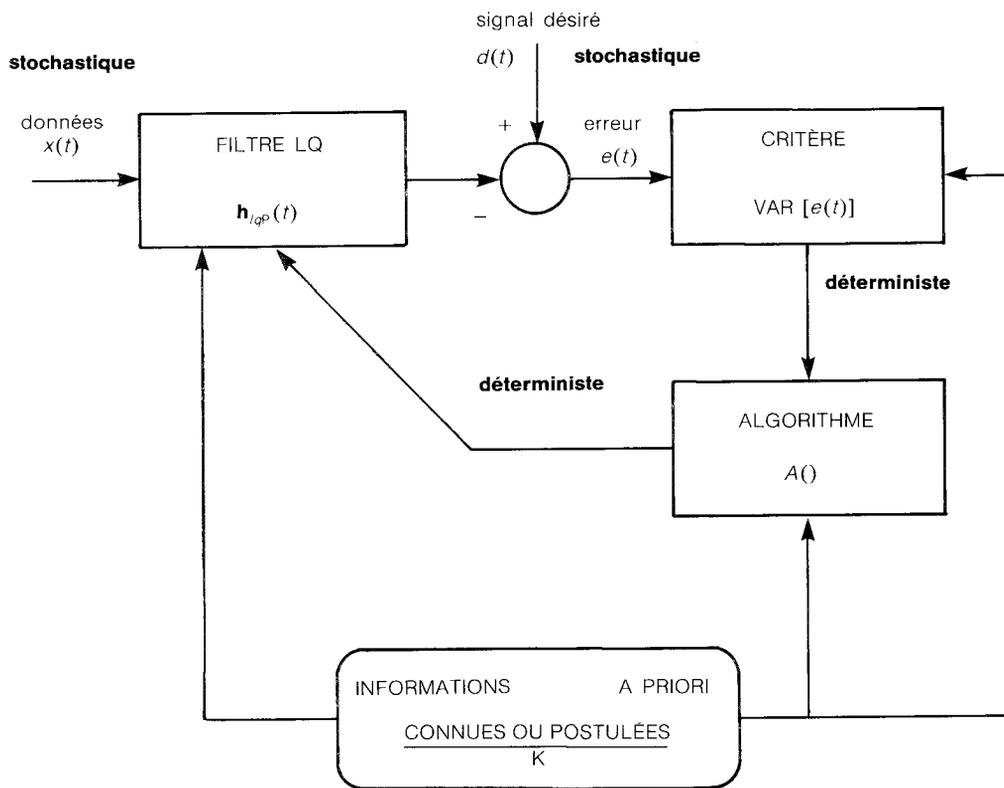


Fig. 1. — Filtrage optimal LQ récursif. Point de vue théorique.

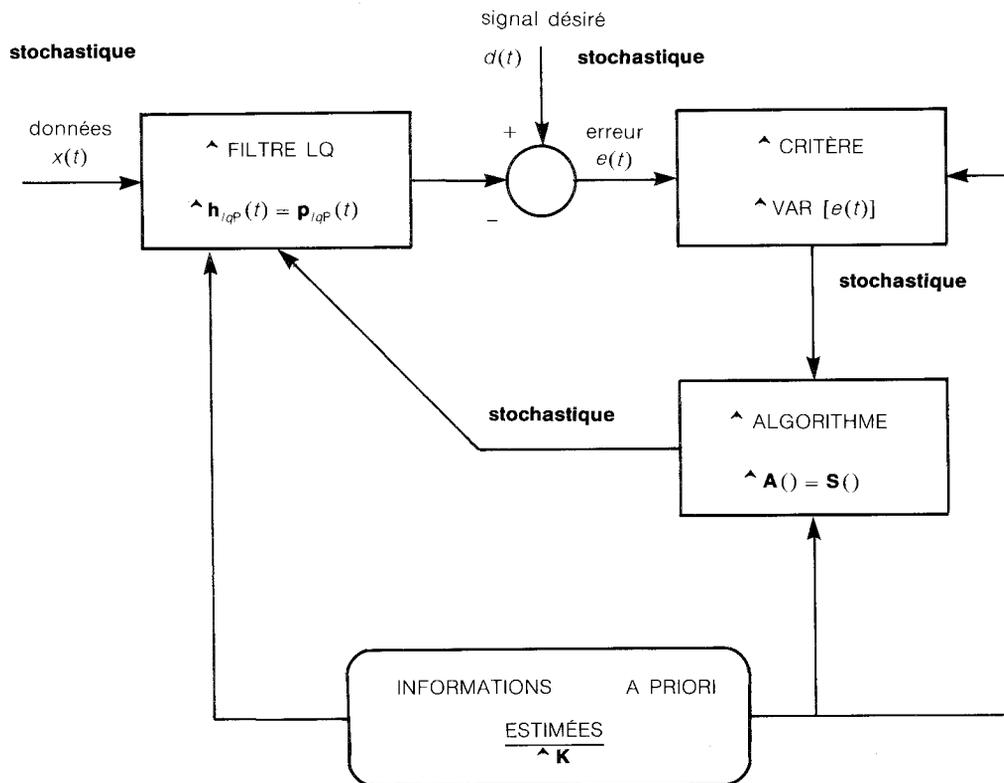


Fig. 2. — Filtrage optimal LQ récursif. Point de vue pratique.

où \mathbf{A} représente l'algorithme choisi. Dans la mesure où malgré la non-linéarité du filtrage, la sortie d'un filtre de Volterra transverse reste « linéaire en paramètre », voir [7], les propriétés élémentaires des procédures du type \mathbf{A} qu'il s'agisse de méthode de la plus grande pente [16], ou de Newton-Raphson [17], resteront semblables au cas linéaire. Seul l'espace des paramètres et la nature de \mathbf{K} changent. En plus des moments d'ordre deux, notés symboliquement m_2 , du processus $x(t)$, requis dans le cas linéaire, le filtrage quadratique nécessite, voir [7], les moments d'ordre trois, m_3 et quatre m_4 . En dernier lieu, des moments mixtes entre $d(t)$ et $x(t)$, notés \mathbf{M} , complètent \mathbf{K} . On écrit alors, de façon synthétique $\mathbf{K}(m_2, m_3, m_4; \mathbf{M})$. Si le point de vue adopté est pratique, les moments évoqués sont estimés et la procédure résultante devient seulement une estimée stochastique, $\hat{\mathbf{A}}$, notée \mathbf{S} , qui génère un vecteur lq aléatoire,

$$(2) \quad \mathbf{p}_{lqP}(t) = \hat{\mathbf{A}} \mathbf{h}_{lqP}(t)$$

estimé de $\mathbf{h}_{lqP}(t)$ à l'instant et vérifiant la relation de récurrence, voir figure 2.

$$(3) \quad \mathbf{p}_{lqP}(t) = \mathbf{S}(\mathbf{p}_{lqP}(t-1), \hat{\mathbf{K}}).$$

Vu sous cet angle, l'arbitraire de la structure de \mathbf{S} est limité par le choix de \mathbf{A} , de son estimation et, en dernier lieu de $\hat{\mathbf{K}}$. C'est cette remarque qui permet de répertorier les différentes structures de bon nombre d'algorithmes adaptatifs lq , et d'en choisir éventuellement une, adaptée aux propriétés statistiques des données.

3. Choix de la structure d'un algorithme adaptatif linéaire quadratique, en fonction des propriétés statistiques des données

3.1. CONSIDÉRATIONS GÉNÉRALES, STRUCTURES COUPLÉES ET DÉCOUPLÉES

Dans la mesure où \mathbf{S} n'est qu'une estimée stochastique de \mathbf{A} , c'est la structure de \mathbf{A} qui détermine totalement celle de \mathbf{S} , sachant que l'estimation contribue simplement à remplacer les paramètres caractéristiques déterministes de la structure choisie par ces mêmes quantités estimées, donc aléatoires, sans modifier, en aucune façon, la structure elle-même. Il suffit, pour s'en convaincre, d'examiner la situation linéaire où l'on passe respectivement en changeant de point de vue, de la méthode de la plus grande pente, sans gain matriciel, à l'algorithme LMS, sans gain matriciel non plus [16], [17], et de l'algorithme de Newton-Raphson, avec gain matriciel, à l'algorithme RLS, nanti, lui aussi d'un gain matriciel [16], [17]. On examine donc ci-après, dans un premier temps, les structures possibles de \mathbf{A} liées à \mathbf{K} . On remarque tout d'abord, en tenant compte des parties linéaire, $\mathbf{h}_{lP}(t)$, de taille P et quadratique, $\mathbf{h}_{qP}(t)$, de taille $P(P+1)/2$, voir [7], de $\mathbf{h}_{lqP}(t)$, que la relation (1) se divise naturellement, voir Annexe 1, en deux sous-procédures :

$$(4) \quad \mathbf{h}_{lP}(t) = \mathbf{A}_l(\mathbf{h}_{lP}(t-1), \mathbf{K}(m_2, m_3; \mathbf{M}))$$

$$(5) \quad \mathbf{h}_{qP}(t) = \mathbf{A}_q(\mathbf{h}_{qP}(t-1), \mathbf{K}(m_3, m_4; \mathbf{M}))$$

où \mathbf{A}_l (resp. \mathbf{A}_q), est la restriction de \mathbf{A} à la partie linéaire de $\mathbf{h}_{lqP}(t)$ (resp. quadratique) obtenue par une simple projection euclidienne. Malgré leur « séparation apparente », les sous-procédures décrites par (4) et (5) n'ont rien de découplées dans la mesure où le calcul de la partie linéaire à l'instant t nécessite la connaissance de la partie quadratique à l'instant $t-1$ et vice versa. Les arguments de \mathbf{K} , dans (4) et (5) sont les seuls nécessités dans les récurrences correspondantes, leur diminution, par rapport aux 4 arguments mis à jour au § 5, m_2, m_3, m_4 et \mathbf{M} , provient de l'opération de projection euclidienne de \mathbf{A} . On dira que les deux sous-procédures (4) et (5) sont découplées si \mathbf{K} est tel que, ce qui revient à dire, ici, que les moments d'ordre trois sont nuls, voir Annexe 1,

$$(6) \quad \mathbf{h}_{lP}(t) = \mathbf{A}_l(\mathbf{h}_{lP}(t-1), \mathbf{K}(m_2, 0; \mathbf{M}))$$

$$(7) \quad \mathbf{h}_{qP}(t) = \mathbf{A}_q(\mathbf{h}_{qP}(t-1), \mathbf{K}(0, m_4; \mathbf{M}))$$

Rendu à ce stade, on constate qu'il existe deux grandes classes de procédures du type \mathbf{A} déterminées par les moments d'ordre trois du processus observation. Lorsqu'ils sont non nuls la procédure est dite couplée, et décrite par (4) et (5), les adaptations, à chaque itération, des parties linéaire et quadratique du filtre optimal sont dépendantes. S'ils sont nuls, la procédure, régie par (6) et (7), revient finalement à mettre en œuvre deux algorithmes tout à fait séparés concernant respectivement le noyau linéaire et le noyau quadratique.

En remplaçant \mathbf{K} par $\hat{\mathbf{K}}$, $\mathbf{h}(t)$ par $\mathbf{p}(t)$ et \mathbf{A} par \mathbf{S} , c'est-à-dire en cohérence avec le discours du § 2, dans (4), (5), (6), (7), on obtient les équations de fonctionnement des deux familles, couplée et découplée, d'algorithmes stochastique lq adaptatif.

3.2. FORME EXPLICITE DES PROCÉDURES lq STOCHASTIQUES COUPLÉES ET DÉCOUPLÉES

En estimant, comme dans le cas linéaire, [17], le gradient intervenant dans les algorithmes \mathbf{A} , par une valeur « instantanée » et la matrice de pondération, linéaire quadratique, \mathbf{M}_{lqP} , dont la structure détermine le type d'algorithme retenu, LMS étendu ou RLS étendu, voir annexe 1, par une forme récursive à l'instant t , $\hat{\mathbf{M}}_{lqP}(t)$, notée $\mathbf{Q}_{lqP}(t)$, on est conduit, voir annexe 1, à la relation de récurrence suivante, qui contient les deux formes possibles, couplée et découplée, de n'importe quel algorithme lq , du type gradient estimé, ou Newton-Raphson estimé, et dont l'utilisation est précisée ci-après,

$$\begin{aligned} \begin{bmatrix} \mathbf{P}_{lP}(t) \\ \mathbf{P}_{qP}(t) \end{bmatrix} &= \begin{bmatrix} \mathbf{P}_{lP}(t-1) \\ \mathbf{P}_{qP}(t-1) \end{bmatrix} + \mu \begin{bmatrix} \mathbf{Q}_{lP}(t) & \mathbf{Q}_{lqP}(t) \\ \mathbf{Q}_{qP}(t) & \mathbf{Q}_{qP}(t) \end{bmatrix} \\ &\times \begin{bmatrix} \mathbf{x}_{lP}(t) \\ \mathbf{x}_{qP}(t) \end{bmatrix} \left(d(t) - [\mathbf{x}_{lP}^T(t) \ \mathbf{x}_{qP}^T(t)] \begin{bmatrix} \mathbf{P}_{lP}(t-1) \\ \mathbf{P}_{qP}(t-1) \end{bmatrix} \right). \end{aligned}$$

Discussion

Choix du type d'algorithme linéaire quadratique adaptatif
Comme dans le cas linéaire pur, [12], [17], c'est la nature

de la matrice de pondération, $\mathbf{Q}_{lqP}(t)$, donnée dans (8) sous forme partitionnée, voir [7], qui conditionne, le type d'algorithme préconisé. Le choix de l'algorithme résulte là encore, du compromis vitesse de convergence et complexité de mise en œuvre, sachant qu'en plus des paramètres des algorithmes linéaires, interviennent ici les moments d'ordre trois et quatre, [7]. Pour l'algorithme LMSLQ, il n'y a pas de rotation du gradient estimé, [16], et $\mathbf{Q}_{lqP}(t) = \mathbf{I}_{lqP}$, matrice identité de taille, pour un ordre P fixé, $[P(P+3)/2, P(P+3)/2]$, voir [7] § 2. Cet algorithme se caractérise par une convergence lente, [12], [16], [17], limitée par le spectre de la matrice de covariance linéaire-quadratique, fonction des moments d'ordre deux trois et quatre, [7], et une complexité proportionnelle, ici, au vu de (8), à $P(P+3)/2$. Les considérations données précédemment se déduisent sans difficulté des études linéaires disponibles dans [12], [16], [17] étant donné qu'un filtre de Volterra LQ transverse reste linéaire en paramètre.

Pour les algorithmes RLSQL, rapides ou non, la nuance intervenant dans la façon dont on calcule le gain, [20], on opère une rotation du gradient instantané, [16], [17], afin d'accélérer, au prix d'une complexité accrue, la convergence, et la matrice de pondération se détermine comme l'inverse d'une estimée récursive de la matrice de covariance LQ,

$$(9) \quad \mathbf{Q}_{lqP}(t) = \left[\sum_{k=0, t} \lambda^{t-k} \mathbf{x}_{lqP}(k) \mathbf{x}_{lqP}^T(k) \right]^{-1}.$$

Dans l'algorithme RLSQL « standard », noté RLSSLQ, on exprime $\mathbf{Q}_{lqP}(t)$ en fonction de $\mathbf{Q}_{lqP}(t-1)$ à l'aide de la formule de Woodbury dont l'application ici est évidente, le lecteur est donc invité à consulter la référence linéaire, [17], et qui débouche sur une complexité proportionnelle à $[P(P+3)/2]^2$, ce qui commence à devenir prohibitif. L'algorithme RLSQL rapide, noté RLSQL, pallie partiellement cet inconvénient, en introduisant une récurrence sur le vecteur gain LQ,

$$(10) \quad \mathbf{g}_{lqP}(t) = \mathbf{Q}_{lqP}(t) \mathbf{x}_{lqP}(t)$$

qui diminue la complexité du RLSSLQ jusqu'à P^3 . Une présentation du RLSQL qui ne tient pas compte de l'influence des propriétés statistiques des données et uniquement algébrique, est disponible dans [10]. La présentation géométrique d'un tel algorithme ne se limite pas à un simple jeu d'écriture comme dans le cas de la version standard citée ci-dessus. Elle suppose l'introduction d'une récurrence sur des opérateurs de projection et transversaux, [21], relatifs à des espaces observations triangulaires LQ, [22], [23], qui s'avère délicate et ne peut reposer uniquement sur le savoir faire linéaire, [21]. Cette approche fera d'ailleurs l'objet d'un prochain article.

Structure du type d'algorithme, imposée par les propriétés statistiques des données

Quelle que soit le type d'algorithme choisi, LMSLQ ou RLSQL, les propriétés statistiques des données, comme il a été vu au § 3.1, imposent, en plus, une structure spécifique liée à l'intensité des moments d'ordre trois, ce qui suppose, pratiquement, d'estimer leur importance

relative, en calculant le skew [18], [19], avant de mettre en œuvre tout algorithme LQ. Ce point, nouveau par rapport au cas linéaire, précise de quelle façon interviennent les informations a priori dans le filtrage adaptatif LQ. Si ces moments sont élevés pratiquement, d'après la valeur du skew, la structure requise adapte globalement et simultanément, à l'aide d'une seule procédure caractérisée par un seul pas μ_{lq} et une seule matrice gain $\mathbf{Q}_{lqP}(t)$, les coefficients linéaires et quadratiques selon la relation (8) et les modifications portant sur le gain matriciel, résultant du choix du type d'algorithme. Si ces mêmes moments revêtent une importance négligeable, la structure devient découplée, voir § 3.1.

Les noyaux linéaires et quadratiques sont alors « rafraîchis », à tout instant, selon deux procédures tout à fait indépendantes caractérisées respectivement par $\{\mu_l, \mathbf{Q}_{lP}(t)\}$ et $\{\mu_q, \mathbf{Q}_{qP}(t)\}$, les relations de récurrence correspondantes étant obtenues en choisissant respectivement $\mathbf{x}_{qP}(t) = \mathbf{O}_{qP}$, $\mu = \mu_l$ et $\mathbf{x}_{lP}(t) = \mathbf{O}_{lP}$, $\mu = \mu_q$ dans (8), voir annexe 1. Ce résultat, n'a somme toute, rien de surprenant, dans la mesure où, lorsque les moments d'ordre trois sont nuls, les noyaux linéaire, \mathbf{h}_{lP}^o , et quadratique, \mathbf{h}_{qP}^o , optimaux, que l'on cherche donc à estimer, ici, de façon récursive, sont solutions de deux équations indépendantes, totalement découplées, voir [7].

Remarque

Si l'on se réfère à la terminologie adoptée au § 2, la relation (8) qui caractérise une procédure stochastique \mathbf{S} fait référence au point de vue dit « pratique ». En conséquence, on a, dans (8), en accord avec les notations de [7] § 2, et (a)

$$(11) \quad \mathbf{x}_{lq}(t-i) = [x(t-i) \wedge [x^2(t-i)]_c = \wedge [x(t-i)x(t-i+1)]_c \dots \wedge [x(t-i)x(t)]_c^T$$

avec

$$(12) \quad \wedge [x(t-i)x(t-j)]_c = x(t-i)x(t-j) - \wedge E[x(t-i)x(t-j)].$$

Il s'avère donc indispensable dans le cas d'algorithmes adaptatifs LQ d'estimer la covariance du processus, par des méthodes tout à fait standard, que le processus sous jacent soit stationnaire ou non, [24]. En toute cohérence, il aurait fallu ajouter un chapeau avant les vecteurs de données intervenant dans (8), ce dernier a été omis afin d'alléger l'écriture et dans la mesure où il n'y a pas d'ambiguïté possible !

4. Effets d'une inadéquation de structure de l'algorithme linéaire-quadratique

4.1. PROPRIÉTÉS DE LA STRUCTURE OPTIMALE EN PRÉSENCE DE MOMENTS D'ORDRE TROIS

Afin de fixer les idées, on s'intéresse seulement dans ce paragraphe à l'algorithme LMSLQ, et plus particulièrement, aux effets d'une inadéquation de sa structure, c'est-

à-dire, au vu du § 3, d'une mauvaise prise en compte des informations a priori. La relation de récurrence de base s'obtient, par conséquent, en substituant, dans (8), la matrice identité à $\mathbf{Q}_{lqP}(t)$. Les sources d'anomalies de structure, apparemment variées, se ramènent, en fait, à deux scénarios : l'utilisation d'une structure découplée à la place d'une couplée et vice versa. Dans la mesure où, de plus, ces situations sont symétriques, seule la première est détaillée dans ce qui suit, elle est donc « pathologique », au sens défini dans le paragraphe précédent, si les moments d'ordre trois du processus sous jacent sont non nuls, ce qui est désormais supposé.

La procédure adaptée est, par conséquent, globale et calcule récursivement à chaque instant t , simultanément et de façon couplée les noyaux linéaire et quadratique, selon la relation, d'après ce qui a été dit plus haut, et au vu de (8),

$$(13) \quad \mathbf{p}_{lqP}(t) = \mathbf{p}_{lqP}(t-1) + \mu_{lq} \mathbf{x}_{lqP}(t) \times [b_{lqP}(t) + r_{lqP}(t|t-1)]$$

où $b_{lqP}(t)$, appelé, « bruit de sortie linéaire quadratique », représente l'erreur d'estimation du signal « désiré » $d(t)$ à l'aide d'un filtre de Wiener d'ordre P linéaire-quadratique, voir [7],

$$(14) \quad b_{lqP}(t) = d(t) - \mathbf{x}_{lqP}^T(t) \mathbf{h}_{lqP}^o$$

Sa variance, v_{lqP} , donnée dans [7], chiffre tout l'arbitraire d'une méthode paramétrique dans un problème d'estimation : celui du choix de la forme du modèle lorsque l'on sait calculer sans erreur le modèle optimal, soit \mathbf{h}_{lqP}^o ici, indépendamment de la méthode choisie pour estimer ensuite de façon récursive dans le temps, ce même modèle. Dans la suite, on fera référence à ce défaut sous la dénomination « inadéquation du modèle », différente conceptuellement de « l'inadéquation de structure de l'algorithme d'identification » qui, comme il sera vu, peut apparaître même si le modèle retenu est optimal. La quantité $r_{lqP}(t|t-1)$, quant à elle, appelée « erreur résiduelle linéaire quadratique », en cohérence avec le cas linéaire, voir [12] et [17], s'écrit

$$(15) \quad -\mathbf{x}_{lqP}^T(t) \delta \mathbf{p}_{lqP}(t-1)$$

où intervient l'écart, à l'instant $t-1$, entre le produit de l'algorithme et le filtre optimal,

$$(16) \quad \delta \mathbf{p}_{lqP}(t-1) = \mathbf{p}_{lqP}(t-1) - \mathbf{h}_{lqP}^o$$

La variance asymptotique de sa partie fluctuante, [12], [17], qui caractérise le régime permanent de l'algorithme, se calcule sans difficulté, vu l'analogie formelle, à partir du cas linéaire, voir [12] et [17], sous les hypothèses de la « Mindépendance », [12], [17], beaucoup plus réaliste que la théorie de l'indépendance proposée dans [9]. On a, dans ces conditions,

$$(17) \quad \lim_{t \rightarrow \infty} E[r_{lqP}^2(t|t-1)] = \mu_{lq} v_{lqP} E[\mathbf{x}_{lqP}^T(t) \mathbf{x}_{lqP}(t)]/2$$

La variance donnée par (17) chiffre le prix à payer pour

l'adaptation et met en avant les performances de l'algorithme, en tant que tel, une fois qu'un modèle a été choisi. On note, au passage, qu'interviennent, dans les deux derniers termes de (17) les moments d'ordre deux, trois et quatre, voir [17].

4.2. EFFETS DE L'INADÉQUATION DE STRUCTURE D'UN ALGORITHME LMSLQ EN PRÉSENCE DE MOMENTS D'ORDRE TROIS, « BRUIT D'INADÉQUATION »

Notre propos consiste à mettre en avant les effets d'une mauvaise utilisation des informations a priori, non sur le modèle choisi, mais plutôt, sur la structure de l'algorithme LMSLQ, chargé de l'estimer, aussi supposons-nous, dans ce paragraphe, que le modèle LQ entre les données et le signal désiré est optimal, « l'inadéquation du modèle » est donc bien représentée par le signal $b_{lqP}(t)$, (cette précision n'a rien de fortuit, bien au contraire : si, par exemple, le modèle sous jacent optimal était purement quadratique, l'inadéquation du modèle serait matérialisée par le signal

$$(18) \quad d(t) - \mathbf{x}_{qP}(t) \mathbf{h}_{qP}^o$$

et non $b_{lqP}(t)$. Dans (18) \mathbf{h}_{qP}^o représente le filtre de Wiener quadratique optimal différent de la partie quadratique de \mathbf{h}_{lqP}^o lorsque les moments d'ordre trois sont non nuls, [17]). On se limite, de plus, aux conséquences de cette mauvaise utilisation sur le noyau linéaire sachant que les résultats relatifs au noyau quadratique s'obtiendraient par simple permutation d'indices, voir [17]. Si, donc, on ne tient pas compte de l'intensité des moments d'ordre trois en implémentant deux algorithmes indépendants selon une démarche précisée au § 3, la partie linéaire de l'estimée du filtre optimal évolue selon l'équation, voir (8),

$$(19) \quad \mathbf{p}_{lP}(t) = \mathbf{p}_{lP}(t-1) + \mu_1 \mathbf{x}_{lP}(t) \times [b_{lqP}(t) + r_{lP}(t|t-1) + \gamma_{qP}(t)]$$

Les deux premiers termes du crochet de (19) ont déjà été explicités, plus haut, ils caractérisent respectivement l'acuité du modèle et la qualité de l'algorithme, ici linéaire, après convergence. Le troisième est plus inédit puisqu'il n'apparaît ni dans le cas classique, [12], [17], ni dans la récurrence (13) du § 4.1. Egal à

$$(20) \quad \mathbf{x}_{qP}^T(t) [\mathbf{h}_{lqP}^o]_q$$

il quantifie « l'inadéquation » de la structure de l'algorithme LMSLQ, nous l'appellerons, dans la suite, « bruit d'inadéquation de l'algorithme ». Si les moments d'ordre trois étaient nuls, il serait éliminé par l'algorithme lui-même lors des itérations successives, en supposant le processus des données ergodique, au sens faible, d'ordre trois. En effet, à ce moment là, l'identité, liée à l'ergodisme,

$$(21) \quad \lim_{N \rightarrow \infty} (\text{M.Q.}) (1/N) \sum_{k=1, N} \mathbf{x}_{lP}(k) \mathbf{x}_{qP}^T(k) = \mathbf{O}_{l|qP}(t)$$

asservirait la qualité du régime permanent de l'algorithme, décrit par (19), uniquement à la partie fluctuante de

l'erreur résiduelle linéaire $r_{lp}^f(t|t-1)$. Si les moments d'ordre trois sont non nuls, les itérations successives n'éliminent pas $\gamma_{qp}(t)$, bien au contraire, et le prix à payer pour l'adaptation devient « plus élevé ».

La diminution de ses effets « pernicieux » nécessite, alors, de retrancher aux trois termes entre crochets de (19) la quantité $\mathbf{x}_{qp}^T(t) \mathbf{p}_{qp}(t-1)$ afin de remplacer $\gamma_{qp}(t)$ par une erreur résiduelle quadratique $r_{qp}(t|t-1)$ dont la variance asymptotique devient beaucoup plus faible que celle $\gamma_{qp}(t)$, sous réserve d'un couplage, ce que réalise exactement la structure optimale, voir la restriction linéaire de (13). De manière générale, quelles que soient les propriétés statistiques des données, on déduit de (20) et [7],

$$(22) \quad \text{Var}[\gamma_{qp}(t)] = [\mathbf{h}_{lqp}^o]_q^T \mathbf{C}_{qp} [\mathbf{h}_{lqp}^o]_q$$

avec, voir [7],

$$(23) \quad [\mathbf{h}_{lqp}^o]_q = \mathbf{R}_{q|lp} \mathbf{s}_{lp} + \mathbf{R}_{qp} \mathbf{s}_{qp}.$$

Les deux matrices de (23) sont des éléments de la version partitionnée de l'inverse de la matrice de covariance, \mathbf{C}_{lqp} , de $\mathbf{x}_{lqp}(t)$, [7]. Les deux vecteurs, quant à eux, interviennent dans l'intercorrélation entre $d(t)$ et le vecteur observation linéaire quadratique, [7].

Remarque :

Comme il a été précisé plus haut, la « pathologie » due à la mise en œuvre d'une structure couplée à la place d'une découplée se formalise de façon tout à fait symétrique à celle contenue dans le présent paragraphe. Elle intervient donc dans des conditions duales, c'est-à-dire lorsque, cette fois, les moments d'ordre trois sont nuls. Elle se manifeste par une complexité accrue, et non justifiée, liée au couplage. Ce point n'a pas été vu dans la référence [9] où les auteurs, malgré le caractère gaussien des données, c'est-à-dire, en l'absence de moments d'ordre trois ont implémenté, inutilement, une structure couplée.

4.3. EXEMPLE

On illustre ici, sur une application particulière, les idées essentielles relatives au couplage dû à l'existence de moments d'ordre trois. Le schéma de principe est donné sur la figure 3. Le processus observation, $x(t)$, est obtenu par filtrage, à l'aide d'un AR(1, 0.9), puis élévation au carré et centrage, d'un signal gaussien blanc unitaire. Ce procédé confère, par conséquent, à $x(t)$ des moments d'ordre trois non nuls, ce qui, pratiquement correspond à des valeurs respectives du skew, [18], et du kurtosis, [18], de 3,6 et 11,6. Le signal désiré s'écrit directement comme une filtrée, non bruitée, linéaire quadratique de $x(t)$. Cette condition rend optimal le modèle choisi jusque-là, dans l'article, puisque le problème d'estimation de $d(t)$ à l'aide de $x(t)$ devient singulier dès que le filtre d'estimation est LQ, on a, alors, $b_{lqp}(t) = 0$, dans, (13) et (19). Cela signifie, entre autre, que la variance de la partie fluctuante de l'erreur résiduelle linéaire quadratique donnée par (17) est nulle car v_{lqp} est, par construction, également nulle. Ce point est particulièrement évident sur la figure 4, qui représente l'évolution de l'erreur en sortie de l'algorithme couplé décrit par (13). Les procédures découplées s'avèrent donc tout à fait inadaptées, ce que l'on voit clairement sur les représentations des erreurs en sortie des algorithmes purement linéaire, décrit par (19), et purement quadratique, explicité en permutant les indices de (19), données respectivement sur les figures 5 et 6, et où se manifestent respectivement les effets de $\gamma_{qp}(t)$ et $\gamma_{lp}(t)$.

5. Conclusion

L'implémentation de tout algorithme adaptatif linéaire quadratique nécessite, au préalable, afin d'utiliser au mieux les informations a priori du processus observation,

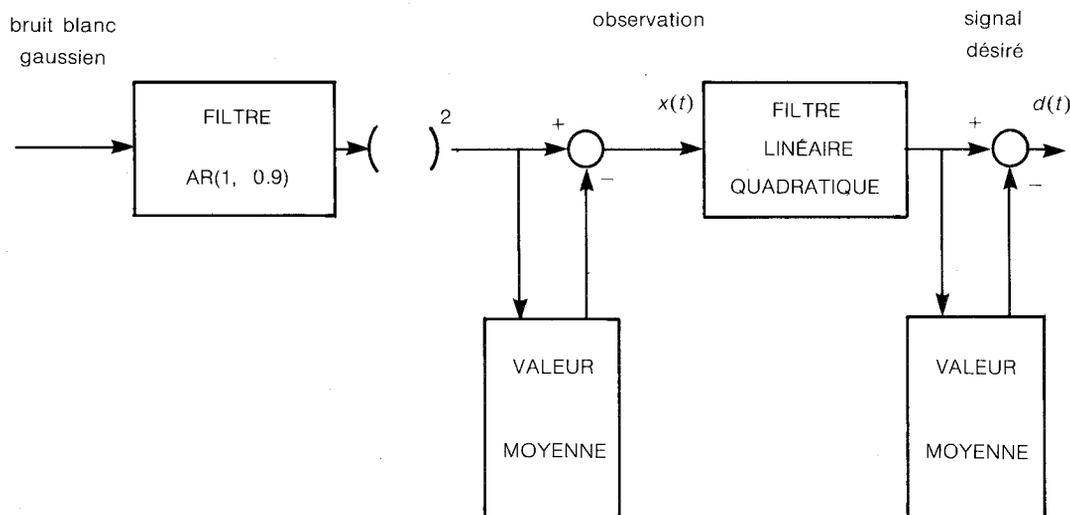


Fig. 3. — Principe de génération du processus observation et du signal désiré.

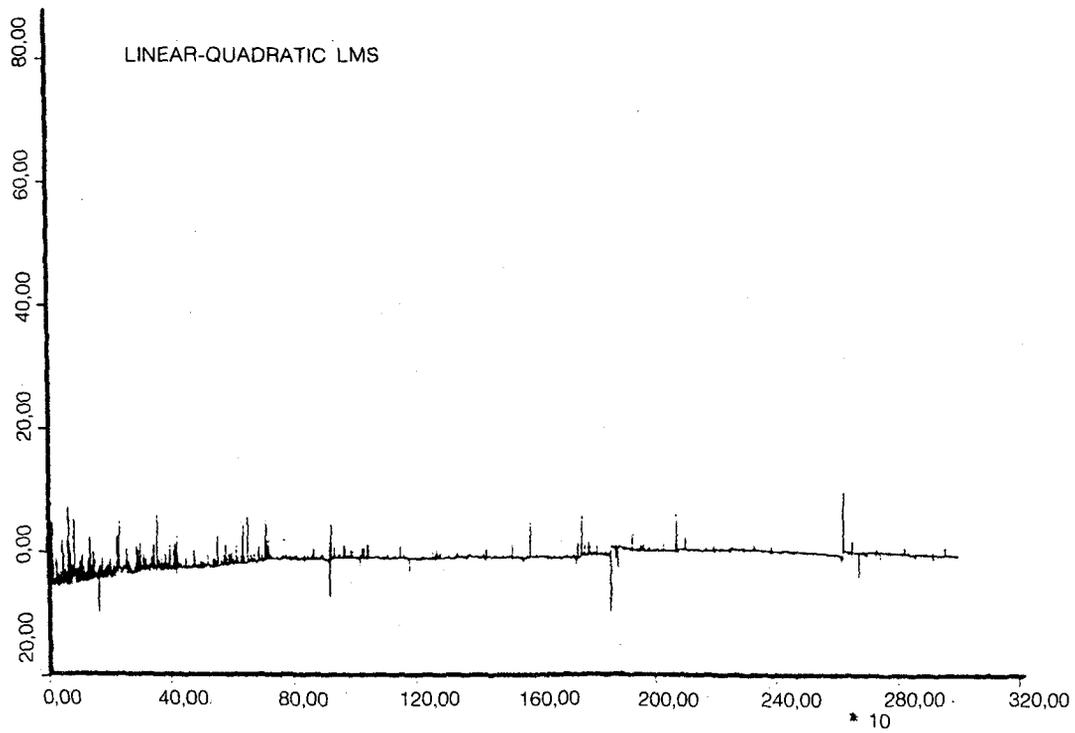


Fig. 4. — Erreur, en sortie de l'algorithme LMSLQ couplé, en présence de moments d'ordre trois.

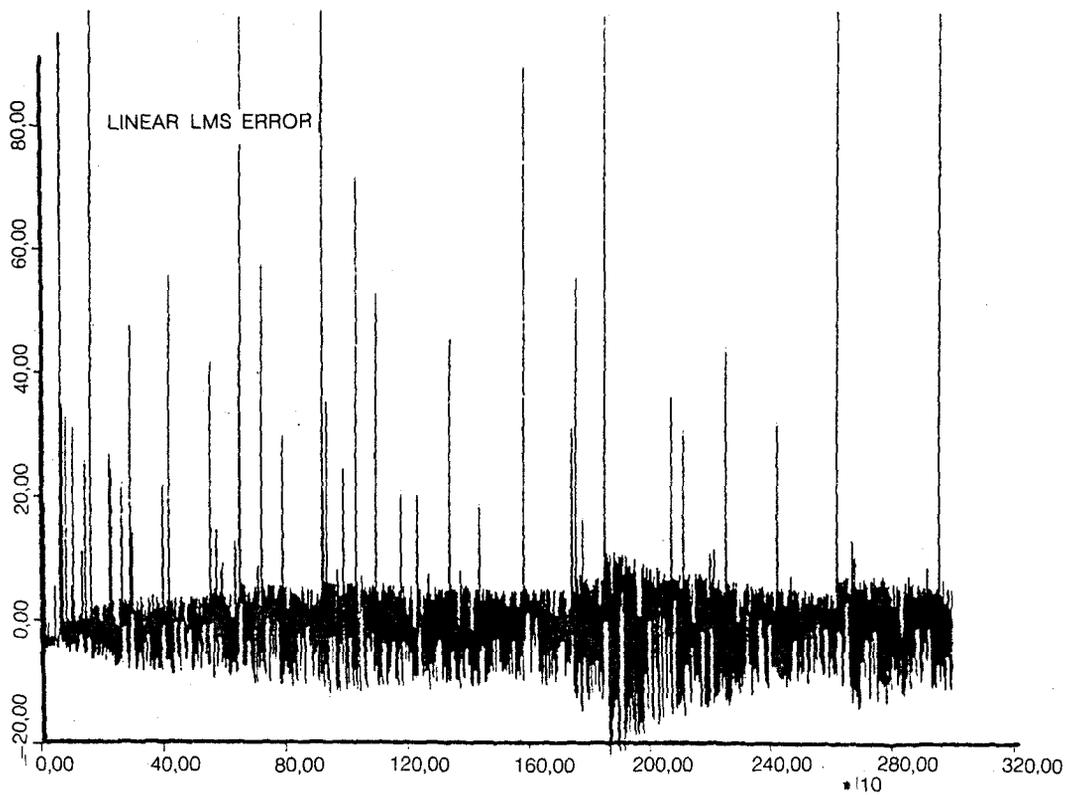


Fig. 5. — Erreur, en sortie de l'algorithme LMSLQ découplé, partie linéaire, en présence de moments d'ordre trois.

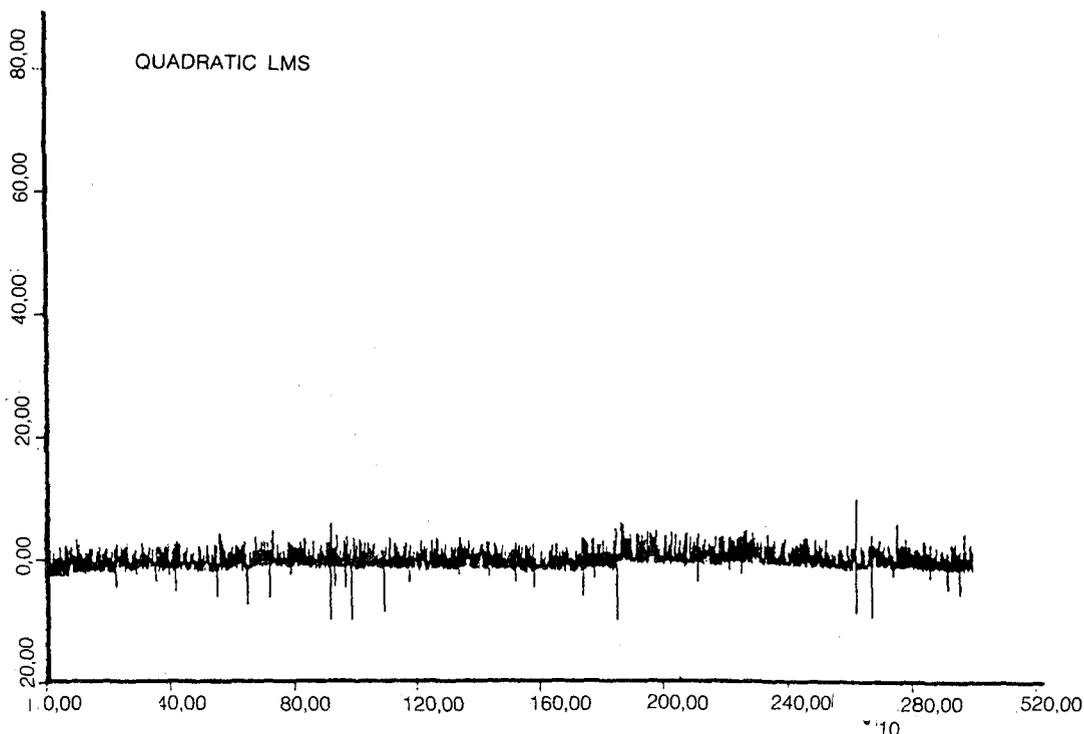


Fig. 6. — Erreur, en sortie de l'algorithme LMSLQ découplé, partie quadratique, en présence de moments d'ordre trois.

d'estimer l'importance relative des moments d'ordre trois par le calcul du skew et du kurtosis. Si ces moments sont nuls, c'est-à-dire pratiquement faibles, la structure adaptée de tout algorithme adaptatif LQ, LMS ou RLS, standard et rapide, est « découplée », au sens où deux procédures stochastiques indépendantes, de pas et de gain différents, identifient récursivement et respectivement les noyaux linéaire et quadratique du filtre optimal. Si ces mêmes moments sont d'intensité non négligeable, la structure de ces algorithmes devient couplée : une seule procédure stochastique, caractérisée par un seul pas et un gain unique, rafraîchit simultanément et de façon conjointe les noyaux linéaire et quadratique. Ces considérations se déduisent de la façon dont s'interprète tout algorithme stochastique, à savoir comme estimée d'une procédure récursive déterministe du type gradient ou Newton-Raphson. Vu qu'un filtre de Volterra transverse reste une fonction linéaire des paramètres, les propriétés de l'algorithme LMSLQ demeurent très semblables, en dehors de la participation supplémentaire des moments d'ordre trois et quatre, à celles de l'algorithme LMS classique. En particulier, l'étude de la convergence fondée sur la théorie de la M indépendante, introduite dans le contexte linéaire, s'étend sans difficulté. En présence de moments d'ordre trois, l'inadéquation de structure d'un algorithme LMSLQ, provenant d'une mauvaise utilisation des informations a priori, se modélise, commodément, par un bruit qui vient se superposer au bruit de modèle, appelé bruit « d'inadéquation ». La détermination de la variance de ce dernier montre à quel point ses effets s'avèrent néfastes sur le régime permanent de l'algorithme. Eu égard à

l'approche développée dans cet article, l'extension à du filtrage de Volterra transverse adaptatif, d'ordre quelconque, devrait se réaliser sans encombre, tous les moments pouvant, cette fois, néanmoins, coupler des noyaux particuliers.

Annexe 1

DICHOTOMIE DE STRUCTURE D'UN ALGORITHME ADAPTATIF LINÉAIRE QUADRATIQUE

On justifie, dans cette annexe, la dichotomie, fondée sur les moments d'ordre trois, des algorithmes déterministes du type A, voir §§ 2 et 3, et donc, par voie de conséquences, de leurs estimées stochastiques, soit, en cohérence avec les notations des §§ 2 et 3, des procédures du type S. On considère le carré de la distance en moyenne quadratique entre un signal $d(t)$ et la sortie d'un filtre transverse d'ordre P, de type « y », et de vecteur réponse impulsionnelle, \mathbf{h}_y ,

$$(A1) \quad \varepsilon_y^2(\mathbf{h}_y) = \gamma_d(0, 0) - 2 \mathbf{h}_y^T \mathbf{s}_y + \mathbf{h}_y^T \mathbf{C}_y \mathbf{h}_y.$$

L'indice « y » prend respectivement les trois valeurs lP, qP et lqP lorsque le filtre mis en œuvre est linéaire, quadratique et linéaire-quadratique. Les propriétés des deux derniers types de filtre sont discutées dans la référence [7]. Au vu de (A1) et de la forme classique obtenue dans le cas linéaire, lorsque $y = lP$, voir [12] et

[17], on constate que seul un jeu d'écriture permet de passer d'un filtre à l'autre, moyennant les notations présentées dans [7]. Bon nombre des propriétés linéaires s'étendent donc immédiatement, aussi, nous contenterons-nous simplement de souligner les différences fondamentales, liées notamment aux questions de couplage, avec l'étude classique et renvoyons sinon le lecteur à la référence [12]. Eu égard à la convexité de $\varepsilon_y^2(\mathbf{h}_y)$, la plupart des procédures déterministes récursives de détermination du vecteur optimal, \mathbf{h}_y^0 , qui permet à $\varepsilon_y^2(\mathbf{h}_y)$ d'atteindre son minimum, v_y , se mettent sous la forme, [16],

$$(A2) \quad \mathbf{h}_y(t) = \mathbf{h}_y(t-1) - (\mu_y/2) \mathbf{M}_y \text{Grad} [\varepsilon_y^2(\mathbf{h}_y)]_{t-1}.$$

Si \mathbf{M}_y est la matrice identité de taille correspondante, la procédure est du type « gradient », [16], sa convergence s'avère limitée, en dehors de la valeur de μ_y , par le spectre de valeurs propres de la matrice \mathbf{C}_y ; si $\mathbf{M}_y = \mathbf{C}_y^{-1}$, la procédure est du type Newton, [16], dont la vitesse de convergence dépend seulement de μ_y . On justifie par l'intermédiaire de (A2) l'expression (1). Le premier caractère inédit, par rapport à la situation linéaire, apparaît, dans (A2), lorsque l'on fait $y = lqP$, et si l'on introduit, de plus, la représentation découplée, utilisable que les moments d'ordre trois de l'observation soient nuls ou non, voir [7], des filtres linéaires quadratiques (on précise ici qu'il convient de faire une différence essentielle entre la représentation découplée d'un vecteur lq qui, suivant l'opportunité, peut intervenir en présence ou non de moments d'ordre trois, et le découplage d'un système d'équation ou encore d'une structure d'algorithme liée uniquement à la nullité de ces mêmes moments, voir [7]). Dans ces conditions, on déduit, en effet, de (A2), une forme moins secrète quant à l'influence des propriétés statistiques des données,

$$(A3) \quad \begin{bmatrix} \mathbf{h}_{lP}(t) \\ \mathbf{h}_{qP}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{lP}(t-1) \\ \mathbf{h}_{qP}(t-1) \end{bmatrix} - \mu_{lP} \left(\frac{1}{2} \right) \begin{bmatrix} \mathbf{M}_{lP} & \mathbf{M}_{l|qP} \\ \mathbf{M}_{q|lP} & \mathbf{M}_{qP} \end{bmatrix} \begin{bmatrix} [\text{Grad} (\varepsilon_{lP}^2)]_l \\ [\text{Grad} (\varepsilon_{lqP}^2)]_{q,t-1} \end{bmatrix}.$$

Deux formes de couplage interviennent explicitement dans (A3), au sens défini par (4) et (5). La première, qui apparaît seulement, par construction, dans les algorithmes du type Newton, provient des deux matrices, transposées l'une de l'autre, voir [7], $\mathbf{M}_{l|qP}$ et $\mathbf{M}_{q|lP}$, dont la nullité est asservie à celle des moments d'ordre trois, [7]. La seconde forme de couplage, commune aux deux familles d'algorithmes, « Newton » et « gradient », émane des parties linéaire et quadratique du gradient de la variance ε_{lqP}^2 , voir (A1) avec $y = lqP$, qui dépendent simultanément des parties linéaire et quadratique du vecteur $\mathbf{h}_{lqP}(t-1)$, dans la mesure où l'on a,

$$(A4) \quad [\text{Grad} (\varepsilon_{lqP}^2)]_{l,t-1} = -2[\mathbf{s}_{lP} - \mathbf{C}_{lP} \mathbf{h}_{lP}(t-1) - \mathbf{C}_{l|qP} \mathbf{h}_{qP}(t-1)]$$

et la formule symétrique déduite en permutant les indices l et q . En remplaçant les grandeurs exactes par des

estimées, selon le procédé décrit au § 2 et l'expression (9), on justifie l'identité (8). D'après ce qui vient d'être dit, l'annulation des moments d'ordre trois modifie la récurrence (A3) à deux niveaux. En ce qui concerne les algorithmes à « gain matriciel », ceci a pour effet de rendre la matrice de pondération bloc diagonale, et de manière conjointe, quel que soit la nature de l'algorithme, les parties linéaire et quadratique du gradient de ε_{lqP}^2 dépendent alors uniquement et respectivement de $\mathbf{h}_{lP}(t-1)$ et $\mathbf{h}_{qP}(t-1)$, ce qui s'écrit en d'autres termes,

$$(A5) \quad [\text{Grad} (\varepsilon_{lqP}^2)]_{l,t-1} = \text{Grad} (\varepsilon_{lP}^2)_{l,t-1}$$

et

$$(A6) \quad [\text{Grad} (\varepsilon_{lqP}^2)]_{q,t-1} = \text{Grad} (\varepsilon_{qP}^2)_{q,t-1}.$$

On déduit alors de tout ceci la forme générale des procédures type « Newton » ou « gradient » en l'absence de moments d'ordre trois, forme, qui illustre les commentaires du § 3.2,

$$\begin{bmatrix} \mathbf{h}_{lP}(t) \\ \mathbf{h}_{qP}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{lP}(t-1) \\ \mathbf{h}_{qP}(t-1) \end{bmatrix} - \mu \left(\frac{1}{2} \right) \times \begin{bmatrix} \mathbf{M}_{lP} & \mathbf{O}_{l|qP} \\ \mathbf{O}_{q|lP} & \mathbf{M}_{qP} \end{bmatrix} \begin{bmatrix} \text{Grad} (\varepsilon_{lP}^2) \\ \text{Grad} (\varepsilon_{qP}^2)_{t-1} \end{bmatrix}.$$

BIBLIOGRAPHIE

- [1] B. FRIEDLANDER & B. PORAT, « Algorithms for optimal estimation of non gaussian processes from high order moments », *Proc. of ICASSP*, pp. 2314-2317, Glasgow, May 1989.
- [2] A. L. SWINDLEHURST & T. KAILATH, « Detection and Estimation using the third Moment Matrix », *Proc. of ICASSP*, pp. 2325-2328, Glasgow, May 1989.
- [3] A. BLANC-LAPIERRE & B. PICINBONO, *Fonctions Aléatoires*, Masson, Paris, 1981.
- [4] B. LEVINE, *Fondements de la radiotechnique statistique*, Editions MIR, Tome 1, Moscou, 1973.
- [5] P. DUVAUT, « Détection et Contraste, Application à la quantification et aux filtres de Volterra optimaux pour la détection et l'estimation », *Thèse de doctorat*, Orsay, février 1987.
- [6] B. PICINBONO & P. DUVAUT, « Linear-quadratic filters for detection and estimation », *IEEE trans. on Inf. Theory*, Vol. IT-34, No. 2, pp. 304-311, March 1988.
- [7] P. DUVAUT, « Le filtrage de Wiener linéaire quadratique à horizon fini, application à la prédiction », en cours de publication, *revue TS*, 1989.
- [8] P. CHEVALIER & B. PICINBONO, « Optimal linear quadratic array for Detection », *Proc. of ICASSP*, Glasgow, May 1989.
- [9] T. KOH & E. J. POWERS, « Second order Volterra filtering and its applications to system identification », *IEEE trans. on ASSP*, Vol. ASSP-33, No. 6, pp. 1445-1455, December 1985.
- [10] V. J. MATHEWS & J. LEE, « A fast least square second order Volterra filter », *Proc. of ICASSP*, New York, April 1988.
- [11] O. L. FROST, « An algorithm for linearly constrained Adaptive Array Processing », *Proc. of IEEE*, 60, No 8, pp. 926-935, August 1972.

- [12] O. MACCHI & E. EWEDA, « Second order convergence analysis of stochastic adaptive linear filtering », *IEEE trans. on A.C.*, AC-28, No. 1, January 1983.
- [13] E. PRUVOST, « Méthodes d'adaptation de dictionnaires fondés sur les coefficients cepstraux », *rapport de stage, ENST*, Paris, juin 1989.
- [14] P. CHEVALIER, Antennes adaptatives sous contraintes, *thèse de doctorat*, Orsay, 1989, en préparation.
- [15] K. CHOUKRI, Several approaches to speaker adaptation in automatic speech recognition systems, *thèse de doctorat*, ENST, Paris, 1987.
- [16] M. MINOUX, *Introduction à la programmation mathématique*, Masson, Paris 1981.
- [17] O. MACCHI & M. BELLANGER, « Le point sur le filtrage adaptatif transverse », *GRETSI*, Nice, 1987.
- [18] M. S. SRIVASTAVA and E. M. CAETER, *An Introduction to applied multivariate Statistics*, North Holland, Elsevier Publisher, New York, 1983.
- [19] C. E. DAVILA, « A second-order adaptive Volterra filter with rapid convergence », *IEEE on ASSP*, ASSP-35, No. 9, September 1987.
- [20] T. KAILATH, « Some alternatives in recursive estimation », *Int. Journal of Aut. Control*, Vol. 32, pp. 311-328, 1980.
- [21] D. ALEXANDER, *Introduction to adaptive filtering*, McGraw-Hill, New York, 1986.
- [22] P. DUVAUT, « Prédiction linéaire quadratique, extension de l'algorithme de Levinson », *GRETSI*, Juan les Pins, 1989.
- [23] P. DUVAUT, « Etude théorique d'une récursivité sur l'ordre de la prédiction linéaire quadratique, extension de l'algorithme de Levinson », en cours de publication, *revue TS*, 1989.
- [24] A. PAPOULIS, *Random Variables and Stochastic processes*, McGraw-Hill, New York, 1984.

Manuscrit reçu le 22 septembre 1988