NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

FILTRAGE ADAPTATIF DE CANAUX NON STATIONNAIRES

Odile MACCHI * Eweda EWEDA **

* Laboratoire des Signaux et Systèmes. CNRS-ESE. Plateau du Moulon. 91190 GIF-SUR-YVETTE
** Military Technical College. LE CAIRE. EGYPTE.

RESUME

On étudie la poursuite adaptative du vecteur de filtrage optimal par l'algorithme du gradient stochastique à pas d'incrémentation μ constant, dans le cas non stationnaire et pour des observations successives corrélées. On montre que l'erreur quadratique moyenne de poursuite est la somme de deux termes. Le premier terme, proportionnel à μ est l'erreur en régime stationnaire. Quant à l'erreur en régime transitoire due à la non-stationnarité elle n'est pas, comme l'ont indiqué des auteurs antérieurs, proportionnelle à $1/\mu$ mais à $1/\mu^2$. Dans le cas d'observations indépendantes, on définit quantitativement le concept, jusqu'ici assez flou, de "non-stationnarité lente" et l'on calcule précisément le pas d'incrémentation optimal correspondant au compromis entre finesse de convergence et vitesse de poursuite respectivement reflétés par les deux termes précédents.

SUMMARY

Adaptive tracking of the non-stationary optimal filtering is studied when the algorithm is governed by the stochastic gradient algorithm with constant step-size μ and correlated successive data. It is shown that the mean square tracking error is the sum of two contributions. The first one, proportional to μ, is the steady-state error ; the other one is the transient part of the error, due to nonstationarity. It is not proportional to $1/\mu$, as some previous authors have indicated, but to $1/\mu^2$. The idea, rather vague until now, of a slowly varying channel is given an appropriate quantitave definition in the case of independent observations.

Then the optimum step-size is evaluated : it gives the numerical trade-off between convergence accuracy and tracking speed, each of them being respectively reflected by the previous two contributions.

I. LE FILTRAGE ADAPTATIF NON STATIONNAIRE

Nous étudions le problème classique d'un filtre linéaire, représenté par le vecteur H de ses coefficients et optimisé dans le but de minimiser l'erreur quadratique moyenne (EQM)

$$e(H) \triangleq E(|a-H^T X|^2) \qquad (1)$$

entre un signal utile a et son estimation linéaire $H^T X$ sur la base d'une observation vectorielle X corrélée avec a et de même dimension N que H.

Dans le cas envisagé la correspondance $a_k \to X_k$ est variable avec le temps, représenté par l'indice k. Ainsi la suite des couples $(a_k X_k)$ est non stationnaire. De ce fait l'EQM (1) est elle-même fonction de k et pas seulement de H et son minimum est atteint pour un vecteur optimal $\tilde{H}_k$ variable avec le temps. Il est bien connu que $\tilde{H}_k$ satisfait l'équation, variable avec k,

$$E(X_k X_k^T)\ \tilde{H}_k = E(a_k X_k), \qquad (2)$$

et qu'il peut être approché par un algorithme récursif tel que l'algorithme dit du "gradient stochastique"ou encore algorithme de Widrow [1]

$$H_{k+1} = H_k + \mu X_k (a_k - X_k^T H_k), \qquad (3)$$

dont l'incrément est l'opposé du gradient stochastique de l'EQM (1), aux fins de minimisation récursive. Dans (3), le pas d'incrémentation μ est une quantité positive fixe.

L'objet de cet article est l'étude de l'erreur de poursuite de l'algorithme (3) en régime permanent. Elle reste aléatoire et sera caractérisée en moyenne, par exemple à l'aide de la déviation quadratique moyenne (DQM)

$$\varepsilon(\mu) \triangleq \limsup_{k \to \infty} E (|H_k - \tilde{H}_k|^2). \qquad (4)$$

Dans ce contexte, notre étude apporte deux contributions nouvelles.

Tout d'abord nous montrons, sous des hypothèses très larges et parfaitement réalistes, que la DQM admet une borne supérieure de la forme

$$\varepsilon(\mu) \leqslant \beta\mu + \beta' \frac{d^2}{\mu^2}, \quad \forall \mu \in]o, \mu_2[, \qquad (5)$$

où μ_2, β et β' sont des constantes positives, tandis que d représente l'incrément maximum de l'estimateur optimal $\tilde{H}_k$:

$$d = \sup_k |\tilde{H}_k - \tilde{H}_{k+1}|. \qquad (6)$$

Dans la borne (5), le premier terme est une mesure de la finesse de convergence de l'agorithme (3). Naturellement il diminue lorsque μ décroit. Il demeure seul lorsque $\tilde{H}_k$ est constant (d=o). Le second terme mesure la rapidité de poursuite des variations de $\tilde{H}_k$ en non-stationnaire. Il diminue lorsque μ croit et ne dépend que du rapport (d/μ), ainsi qu'il est conforme à l'intuition. Avec ses deux termes variant en sens inverse cette borne traduit l'idée naturelle simple qu'il existe un pas d'incrémentation optimal μ_{opt} réalisant le compromis entre finesse de convergence et vitesse de poursuite.

Notre deuxième apport est l'étude, sous une hypothèse plus restrictive d'indépendance entre les observations successives,du cas très usuel des "variations lentes". Nous donnons la première définition rigoureuse de cette expression assez floue. Ceci permet un calcul exact du pas d'incrémentation optimal μ_{opt} pour lequel la DQM est minimum.

II. RESULTATS ANTERIEURS ET DISCUSSION

L'étude de la DQM (4) pour caractériser la convergence de l'algorithme (3) a déjà été abordée par de nombreux auteurs, par exemple [1]-[5]. Depuis fort longtemps aussi les utilisateurs de ce type d'algorithme savent que le choix du pas d'incrémentation doit résulter d'un compromis entre finesse de convergence et vitesse de poursuite, chacune d'elles correspondant à une contribution dans la DQM dont les sens de variation vis-à-vis de μ sont opposés, comme c'est le cas dans (5). Néanmoins il est extrêmement difficile de chiffrer quantitativement ces contributions. C'est pourquoi les auteurs ont eu recours à des hypothèses simplificatrices, qui ne sont pas toujours réalistes ni faciles à légitimer. Ces hypothèses sont de deux natures différentes selon qu'elles concernent le type de corrélation qui gouverne la suite des observations (a_k, X_k), ou le degré de non-stationnarité de cette suite, c'est-à-dire l'évolution plus ou moins rapide de l'estimateur optimal $\tilde{H}_k$.

II. 1- CORRELATION DES OBSERVATIONS

Il est bien connu que le calcul de la DQM se heurte à la présence d'une très forte corrélation entre les observations successives X_k. En effet dans les applications de filtrage les vecteurs X_k correspondent à une fenêtre temporelle glissante

$$X_k^T = (x(kT), \ldots , x((k-N+1)T)) \qquad (7)$$

d'une suite x(kT) d'échantillons régulièrement espacés et provenant d'une même fonction continue x(t). Ainsi X_k et X_{k+1} partagent (N-1) composantes et sont très corrélés. L'hypothèse d'indépendance entre les X_k successifs faite par beaucoup d'auteurs [1]-[3], n'est donc pas justifiée.

Considérons tout d'abord le cas stationnaire, où $\tilde{H}_k = \tilde{H}$ ne dépend pas de k. Dans ce cas le calcul de la DQM a déjà été présenté en utilisant un modèle justifiable de corrélation entre les X_k [4]-[6]. En particulier notons le modèle utilisé dans [5] car nous l'utiliserons ici aussi, cf. paragraphe III. On suppose que la mémoire des observations est finie (et même uniformément bornée, pour le cas de non-stationnaire) ce qui est traduit par l'hypothèse suivante

Il existe M positif assurant pour tout k l'indépendance entre les suites $\{ (a_j, X_j) : j \leq k \}$ *et* $\{ (a_j, X_j) : j \geq k+M \}$ (A-1)

Notons que [4] utilise un modèle de corrélation plus général mais avec une barrière dans l'algorithme autour du filtre optimal $\tilde{H}$ et que [6] suppose ce filtre optimal parfait en ce sens que son estimation est sans erreur

$$a_k = \tilde{H}^T X_k \quad , \tag{8}$$

tandis que [5] correspond mieux à l'usage normal de l'algorithme (3).

Ce qui est remarquable, en tous cas, c'est que la forme obtenue pour la DQM ne varie pas avec l'hypothèse d'indépendance ou de corrélation pour les observation s successives : dans le cas stationnaire tous les auteurs s'accordent sur le résultat suivant :

$$\varepsilon(\mu) \leq \beta \mu \quad , \forall \mu \leq \mu_2 \tag{9}$$

quelquesoit leur modèle et leur type de démonstration. Nous remarquons que (9) est la forme particulière que prend notre résultat général (5) lorsque $\tilde{H}_k$ est fixé (d=o). Plusieurs auteurs ont tenté une explication de ce phénomène d'invariance du résultat par rapport au modèle de corrélation (cf. [7] par exemple).

Dans le cas non-stationnaire on peut citer l'article [1] qui suppose l'indépendance entre les X_k successifs et les articles [4] et [8] qui prennent en compte leur corrélation. Il faut noter immédiatement que la forme des résultats ne varie pas selon qu'est faite ou non l'hypothèse d'indépendance, tout comme dans le cas stationnaire. On y retrouve toujours les deux contributions variant en sens inverse. Néanmoins les auteurs trouvent généralement la contribution non stationnaire en $\beta'' d^2 | \mu$ et non pas en $\beta' d^2 | \mu^2$, comme dans notre résultat (5). Celui-ci met donc en cause les études antérieures, ainsi que nous allons l'expliquer maintenant.

II. 2- NON-STATIONNARITE LENTE

Même si l'on simplifie l'analyse du problème par l'indépendance des couples successifs (a_k, X_k), il demeure une difficulté tout-à-fait spécifique du cas non-stationnaire et qui n'a pas été résolue de manière satisfaisante jusqu'ici. La difficulté réside dans la définition de ce qu'il est convenu d'appeler "non-stationnarité lente". Ce cas est en effet très intéressant car très fréquent en pratique. Pourtant la plupart des études sont approximatives car elles utilisent deux hypothèses successives [4], [9] du type

$$\lambda\mu \ll 1. \quad , \tag{10}$$

puis

$$d \ll \mu \rho . \tag{11}$$

Dans ces équations ρ est une constante qui n'est généralement pas précisée, sauf par une phrase telle que "le temps de réponse du filtre adaptatif H_k est très court devant le temps de stationnarité" [1]. La constante λ, quant à elle, est définie par

$$\lambda = \operatorname*{Inf}_k \lambda_{min} [E(X_k X_k^T)], \tag{12}$$

et l'on fait l'hypothèse

$$\lambda > o \quad , \tag{A-2}$$

que ce minimum est non nul afin de garantir grâce à (2) l'existence et le caractère borné de $\tilde{H}_k$. *L'inégalité (10) exprime que le vecteur de l'algorithme (3) est à variations lentes* ; essentiellement H_k n'a pas varié de manière significative en une seule itération. Il est clair que cette condition n'a rien à voir avec la non-stationnarité $\tilde{H}_k$ que l'on veut poursuivre. On pourra toujours assurer (10) par un choix de μ assez petit. D'après (10), c'est l'algorithme (3) qui est à variations lentes et non nécessairement le phénomène étudié, dont l'évolution ne pourra peut-être pas être suivie par un choix de μ si petit. Quant à elle, *l'inégalité (11) traduit le fait que le vecteur H_k de l'algorithme (3) varie beaucoup plus vite que le vecteur optimal $\tilde{H}_k$*. Naturellement cette inégalité pourra toujours être assurée par un choix de μ assez grand. Mais comme d est imposé par la nature, il n'est pas certain qu'avec ce choix, la condition (10) restera satisfaite. On voit bien dans quelle contradiction le problème reste enfermé. En réalité faire ces hypothèses, c'est déjà supposer le problème en partie résolu. On connait déjà, au moins approximativement, une plage de valeurs μ pour laquelle l'algorithme fonctionne de manière satisfaisante: l'algorithme (3) poursuit bien la valeur $\tilde{H}_k$ grâce à (11) tandis qu'il élimine bien les fluctuations résiduelles grâce à une durée de convergence suffisante - c'est la condition (10) - En d'autres termes utiliser des hypothèses de ce type pour le calcul de la DQM, c'est supposer que μ a été choisi dans la zone de compromis que l'on cherche justement à déterminer en calculant la DQM. De telles hypothèses ne sont donc pas acceptables sans être légitimées, au moins a posteriori en les vérifiant pour le pas d'incrémentation optimal auquel on aboutit.

De notre côté, nous avons adopté une démarche différente. La borne supérieure (5) a été établie sans aucune hypothèse apparentée à (10) ou (11), donc sans supposer nécessairement que les non-stationnarités étudiées sont lentes. Elle a été établie à l'aide d'un modèle réaliste de corrélation entre les observations successives (cf. § III). Néanmoins, sous ces hypothèses générales, (5) ne donne pas directement la possibilité d'optimiser numériquement le pas d'incrementation μ par minimisation de la DQM. Dans ce but nous avons repris le modèle usuel des observations indépendantes afin de pouvoir préciser numériquement les grandeurs du calcul. Dans ce cas nous introduisons une définition claire du concept de "non-stationnarité lente" par l'intermédiaire du "degré de non-stationnarité", noté x. Cette quantité est reliée de manière intrinsèque à la statistique de la suite non stationnaire (a_k, X_k) et à l'estimateur optimal variable $\tilde{H}_k$. Elle n'a rien à voir avec le pas d'incrémentation μ de l'algorithme. La non-stationnarité sera dite lente lorsque

$$x \ll 1 \ . \tag{13}$$

Sous cette hypothèse, le calcul de la DQM peut se faire sur un large intervalle

$$\mu \in \,]o, \mu_2[\quad ; \quad \mu_2 \sim \frac{1}{\lambda} \ . \tag{14}$$

En comparaison avec (10) et (11) on voit que l'intervalle de calcul (14) autorise à la fois les très grandes valeurs ($\mu \simeq \frac{1}{\lambda}$) et les très petites valeurs (μ peut tendre vers 0) qui étaient exclues a priori par les auteurs. En montrant ensuite (cf. § V) que le minimum de la DQM sur tout l'intervalle (14) est effectivement atteint pour une valeur qui vérifie les conditions (10) et (11), nous légitimons donc la manière dont les auteurs antérieurs avaient étudié le problème.

III. CAS GENERAL : UNE BORNE DE LA DQM

Dans ce paragraphe nous énonçons sous forme de théorème, le résultat général, valable pour des observations successives corrélées de mémoire finie, et nous le commentons

Théorème 1 *Sous les hypothèses de mémoire finie et d'uniforme inversibilité (A-1) et (A-2), il existe trois nombres positifs μ_2, β, β' tels que la DQM $\varepsilon(\mu)$ de l'algorithme (3) vérifie l'inégalité (5) dans l'intervalle $\mu \in]o, \mu_2[$, pourvu que les moments soient uniformément bornés conformément à l'hypothèse*

$$\exists A, B \text{ positifs tels que } E(|X_k|^{24M}) \leq B \ ;$$
$$E(|a_k|^{4+4/(6M-1)}) \leq A, \forall k \ . \tag{A-3}$$

Les hypothèses de ce théorème sont satisfaites dans un très grand nombre d'applications pratiques, ce qui couvre le domaine du filtrage caractérisé par (7). Prenons par exemple le cas des transmissions de données où la correspondance F_k entre la suite utile a_k et la suite observée x(kT) est un filtrage linéaire non stationnaire correspondant au canal de transmission, et où a_k est une donnée numérique donc bornée. Alors, si F_k est uniformément borné, les moments de $|X_k|$ sont tous uniformément bornés, et l'hypothèse (A-3) est satisfaite. L'hypothèse (A-2) signifie que pour toute fréquence dans la bande passante et à tout instant k la fonction de transfert du canal de transmission est (uniformément) bornée inférieurement par λ. Cela est normalement vérifié car même les canaux à distorsion sévère n'annulent pas complétement le gain du canal. Enfin la suite des données a_k étant usuellement indépendante, la mémoire M de l'hypothèse (A-1) du théorème correspond à la durée D de la réponse temporelle du filtre F_k - augmentée de la longueur N du vecteur X_k des observations. Cette durée D est toujours finie dans les applications, généralement égale à N (ou plus faible).

La preuve du théorème 1 a été détaillée dans un article précédent [8] et nous ne la redonnerons pas ici. Nous donnerons seulement une idée approximative de la démonstration, destinée à faire saisir pourquoi la partie non-stationnaire de la DQM est en $\beta'(d/\mu)^2$ et non pas comme on l'a cru en $\beta''d^2/\mu$. Dans ce but, nous introduisons les notations suivantes

$$V_k = H_k - \tilde{H}_k \ ; \quad d_k = \tilde{H}_k - \tilde{H}_{k+1} \tag{15}$$

$$z'_k = a_k X_k - X_k X^T_k \tilde{H}_k \ ; \ z_k = z'_k + \frac{1}{\mu}(\tilde{H}_k - \tilde{H}_{k+1}) , \tag{16}$$

$$U_{j,k} = \begin{cases} (I-\mu X_k X^T_k)(I-\mu X_{k-1} X^T_{k-1}) \ldots (I-\mu X_{j+1} X^T_{j+1}) \ ; & j \leq k \\ I & j \geq k \ . \end{cases} \tag{17}$$

La différence V_k est l'écart de l'algorithme à l'optimalité, dont on veut borner la puissance dans le théorème : la DQM est la limite (supérieure) de $E(|V_k|^2)$. L'incrément d_k de l'estimateur optimal est borné par d et la variable z'_k est centrée, par la définition (2) de $\tilde{H}_k$. De la sorte, en première approximation, la variable z_k peut être bornée par

$$|z_k| \leq (d/\mu) . \tag{18}$$

La matrice $U_{j,k}$ joue un rôle important dans l'étude de la convergence de l'algorithme, car on vérifie aisément que (3) équivaut à

$$V_{k+1} = U_{o,k} V_1 + \sum_{j=1}^{k} \mu \, U_{j,k} \, z_j \ . \tag{19}$$

Ainsi borner $\varepsilon(\mu)$ implique de prouver une décroissance assez rapide vers zéro de la matrice $U_{j,k}$ $(k \to \infty)$ Une telle propriété est démontrée dans [8] sous la forme

$$E(\| U_{j,k} \|^4) \leq C \, (1-\gamma\mu)^{k-j} ; \forall k \geq j ; \forall \mu \leq \mu_1 , \tag{20}$$

où les constantes γ et μ_1 satisfont $\mu_1\gamma < 1$. En première approximation, on peut donc écrire

$$\| U_{j,k} \| \leq C'(1-\delta\mu)^{k-j} \quad , \delta\mu_1 < 1 \; ; \; \mu \leq \mu_1 \tag{21}$$

ce qui implique

$$\| \sum_{j=1}^{k} \mu U_{j,k} \| \leq \frac{C'}{\delta} \quad ; \forall \mu \leq \mu_1 \; . \tag{22}$$

En reportant dans (19) les bornes (18) et (22), on comprend que V_{k+1} soit, comme z_j, en (d/μ). On trouve donc une borne en $(d/\mu)^2$ pour la déviation quadratique.

IV. CAS PARTICULIER : INDEPENDANCE DES OBSERVATIONS

Dans ce paragraphe nous considérons le problème de la vitesse d'évolution de la non-stationnarité. Afin d'obtenir une définition précise et intrinsèque du cas de "non-sationnnarité lente", simplifions l'analyse en supposant que les couples successifs (a_k, X_k) sont indépendants : la mémoire M vaut 1 dans l'hypothèse (A-1)

IV.1- DEGRE DE NON-STATIONNARITE

Pour définir numériquement le degré de non-stationnarité réécrivons les équations (15) et (16) sous la forme

$$a_k X_k = X_k X_k^T \tilde{H}_k + z'_k \; , \tag{16'}$$

$$\tilde{H}_{K+1} = \tilde{H}_k + d_k \tag{15'}$$

En tenant compte du fait que les grandeurs $a_k X_k$ et $X_k X_k^T$ sont connues (c'est le résultat des mesures à chaque itération), la poursuite de $\tilde{H}_k$ apparait comme un problème d'estimation d'une grandeur d'état : $\tilde{H}_k$ est l'état du modèle linéaire $X_k X_k^T \to a_k X_k$, tandis que $(-d_k)$ et z'_k sont respectivement le bruit d'état et le bruit de mesure, dont le rapport des puissances peut être défini par d^2/C avec

$$C = \sup_k C_k ; \quad C_k \triangleq E(|z'_k|^2) = E(|X_k|^2 (a_k - X_k^T \tilde{H}_k)^2) \; . \tag{23}$$

Pour obtenir une quantité x sans dimension, posons

$$A^2 = E\,(|X_k|^4), \tag{24}$$

$$x = Ad / \sqrt{C} \; . \tag{25}$$

On voit qu'alors le paramètre x reflète *la vitesse d'évolution du vecteur optimal* $\tilde{H}_k$, *mesurée relativement au bruit intrinsèque* z'_k *qui entâche la mesure de* $\tilde{H}_k$. Dans la suite, la grandeur x, (25) sera appelée "degré" de non-stationnarité". *Les variations de* $\tilde{H}_k$ *ne sont pas qualifiées de rapides ou lentes dans l'absolu, mais en référence à la qualité de la mesure de* $\tilde{H}_k$ *que constitue l'observation du couple* (a_k, X_k). Ainsi $\tilde{H}_k$ est à variations lentes si son incrément d est faible devant le bruit de mesure z'_k.

L'inégalité

$$x \ll 1 \tag{26}$$

caractérise le cas de "non-stationnarité lente", et ceci indépendamment du pas d'incrémentation μ de l'algorithme. Dans ce qui suit, nous allons voir qu'une hypothèse de même type mais nettement plus large permet théoriquement l'optimisation à l'aide d'un encadrement de la DQM, tandis que l'hypothèse (26) elle-même autorise jusqu'au calcul numérique effectif de l'optimum.

IV. 2- ENCADREMENT DE LA DQM

Il est facile de voir que l'algorithme (3) s'écrit à l'aide des notations (15) et (16)

$$V_{k+1} = (I - \mu X_k X_k^T)\, V_k + \mu z'_k + d_k \; . \tag{27}$$

Ainsi l'hypothèse d'indépendance entraîne que

$$(z'_k, X_k) \text{ et } V_k \text{ sont indépendants} \tag{28}$$

En utilisant en outre le caractère centré de z'_k on voit alors que

$$E(V_{k+1}) = [I - \mu E(X_k X_k^T)] E(V_k) + d_k \quad , \tag{29}$$

$$E(|V_{k+1}|^2) = E\{V_k^T [I - 2\mu E(X_k X_k^T) + \mu^2 E(X_k X_k^{T2})] V_k\} + \mu^2 E(|z'_k|^2)$$

$$+ |d_k^2| + 2\mathrm{Re}\{d_k^T [I - \mu E(X_k X_k^T)] E(V_k) - \mu^2 E(z_k'^T X_k X_k^T) E(V_k)\} . \tag{30}$$

L'équation (29) entraîne en même temps

$$\lim_{k \to \infty} \sup |E(V_k)| \leq \frac{d}{\mu\lambda} \; ; \; \forall \mu \leq \frac{2}{\lambda + \lambda'} \quad , \tag{31}$$

$$|d_k^T [I - \mu E(X_k X_k^T)] E(V_k)| < \frac{d^2(1-\mu\lambda)}{\mu\lambda} \quad , \tag{32}$$

où l'on a posé

$$\lambda' = \sup_k \lambda_{max}\{E(X_k X_k^T)\}; \quad \lambda_M = \sup_k \lambda_{max}\{E(X_k X_k^T)^2\} \tag{33}$$

On peut appliquer l'inégalité de Schwartz

$$|E(z_k'^T X_k X_k^T)| \leq \sqrt{E(|z'_k|^2)\, E(|X_k|^4)} \leq A\sqrt{C} \; , \tag{34}$$

au dernier terme de (30), d'où l'on déduit aisément en passant à la limite sur k la borne supérieure suivante $\gamma(\mu)$ pour la DQM :

$$\varepsilon(\mu) \leqslant \gamma(\mu) \triangleq [\lambda\mu(\mu^2 C + d^2) + 2\mu^2 A\sqrt{C}\, d + 2d^2(1-\lambda\mu)][\lambda\mu(2\lambda\mu - \lambda_M\mu^3)]^{-1},$$

à condition que μ vérifie (35)

$$\mu \leqslant \mu'_2 \triangleq \mathrm{Inf}\,\left(\frac{2}{\lambda+\lambda'}, \frac{2\lambda}{\lambda_M}\right) \; . \tag{36}$$

Il est très important en vue de l'optimisation de μ de remarquer que *la borne (35) de la DQM présente un minimum unique* μ_0 *dans l'intervalle* $]0, \mu'_2[$. Ainsi en l'absence d'information supplémentaire sur la DQM, on choisira μ_0 comme valeur du pas d'incrémentation. Il faut aussi remarquer que la borne (35) peut être elle même majorée par une borne de la forme (5) discutée antérieurement, en choisissant les constantes suivantes

$$\mu_2 = \lambda/\lambda_M \; ; \beta = \frac{C}{\lambda}\left(1 + \frac{A^2}{\lambda_M}\right) \; ; \; \beta' = 3|\lambda^2 \tag{37}$$

On voit donc que l'hypothèse d'indépendance faite dans ce paragraphe a permis de mieux appréhender le sens de variation de la DQM. Pour acquérir des informations numériques supplémentaires sur la DQM il faut limiter supérieurement le degré x de non-stationnarité En procédant de manière analogue on obtient alors une borne inférieure croissante $\delta(\mu)$ pour la DQM. L'encadrement correspondant fournit un encadrement numérique précis de la valeur optimale de μ. Nous allons maintenant illustrer notre propos dans le cas particulier d'une "non-stationnarité lente".

V. PAS D'INCREMENTATION OPTIMAL

Supposons que

$$[\lambda^{-2}\lambda_M \operatorname{Sup}(1,\ 2\sqrt{\lambda_M}/A)]\ x \ll 1\ . \tag{26'}$$

C'est une condition de non-stationnarité lente du même type que (26) et qui lui est pratiquement équivalente, le crochet de gauche dans (26)' étant un nombre sans dimension usuellement de l'ordre de 2 ou 3. Lorsque (26)' est vérifiée, un calcul simple, donné en [8], montre que le minimum de $\gamma(\mu)$ a lieu pour

$$\mu_o = (4d^2/\lambda C)^{1/3} = \left(\frac{4x^2}{A^2\lambda}\right)^{1/3}. \tag{38}$$

L'équation (38) fournit une bonne approximation numérique du pas d'incrémentation optimal à partir de la valeur x du degré de non-stationnarité et des simples moments $E(|X_k|^2)$ et $E(|X_k|^4)$. C'est là, en pratique, une solution simple et efficace au problème de l'optimisation de μ.

Il est très important de remarquer que d'après (26)' et (38), la valeur de μ_o satisfait en même temps les deux inégalités (10) et (11) introduites sans justification par les auteurs antérieurs (cf. § II-2). Ces inégalités prennent ici la forme particulière

$$\lambda\mu_o = (4x^2)^{1/3}\left(\frac{\lambda 2}{A^2}\right)^{1/3} \ll 1;\quad \frac{\mu_o{}^2}{d^2}\cdot\frac{C}{4^{2/3}} = x^{-1/3}\cdot\left(\frac{A}{\lambda}\right)^{2/3} \gg 1, \tag{39}$$

et elles restent vérifiées pour tout μ dans une plage géométrique $[\alpha^{-1}\mu_o, \alpha\mu_o]$, α de l'ordre de quelques unités. C'est la plage de bon fonctionnement autour du minimum de la DQM.

Pour terminer nous noterons que les résultats antérieurs donnant la DQM en $\beta\mu + \beta'' d^2/\mu$ étaient évidemment valables pour les grandes valeurs de μ où la partie stationnaire, qui était correctement calculée, était prépondérante ; mais ils fournissaient une valeur sous-estimée de la DQM non-stationnaire lorsque, μ diminuant, l'hypothèse (11) n'était plus vérifiée. Dans l'ensemble il s'agissait donc d'une sous-estimation du pas d'incrémentation optimal.

REFERENCES :

[1] B. WIDROW, J.M. McCOOL, M.G. LARIMORE, C.R. JOHNSON "Stationary and nonstationary learning characteristics of the LMS adaptive filter", Proc. IEEE, Vol 64, n° 8, 1976, p. 1151-1162.

[2] G. UNGERBOECK, "Theory of the speed of convergence in adaptive equalizers for digital communication", IBM J. Res. and Develop., 1972, p. 546-555.

[3] O. MACCHI, "Résolution Adaptative de l'équation de Wiener-Hopf", Ann. Inst. Henri Poincaré, 14, 1978, p. 357-379.

[4] D.C. FARDEN, "Tracking properties of adaptive signal processing algorithms", IEEE Trans. Acoustics, Speech, and Signal Processing, vol. ASSP 29, n° 3, 1981, p. 439-446.

[5] O. MACCHI, E. EWEDA, "Second-order convergence analysis of stochastic adaptive linear filtering", IEEE Trans. Automatic Control, vol. 28, n° 1, 1983.

[6] R.R. BITMEAD, B.D. ANDERSON, "Performance of adaptive estimation algorithms in dependent random environnements", IEEE Trans. Automatic Control, vol. AC 25, n° 4, 1980, p. 788-794.

[7] J.E. MAZO, "On the Independence Theory of Equalizer Convergence", B.S.T.J., vol. 58, n° 5, 1979, p. 963-993.

[8] O. MACCHI, E. EWEDA, "Tracking properties of adaptive nonstationary filtering", submitted IEEE Trans. on Information Theory, 1982.

[9] A. BENVENISTE, G. RUGET, "A measure of the tracking capability of recursive stochastic algorithms with constant gains", IEEE Trans. on Automatic Control, vol. AC 27, n° 3, 1982.