

Identification adaptative en blocs à Structure Prédictive: Applications à l'Annulation d'Echos Acoustiques

A. Perrier, E. Moulines et J. Prado

ENST, Télécom Paris, Département Signal, 46 rue rault, 75634 Paris Cedex 13, France.

E-mail: perrier@sig.enst.fr, Fax: 1 45887935.

RÉSUMÉ

Dans cette contribution, nous présentons deux nouveaux algorithmes de type Quasi-Newton par blocs, fondés sur la factorisation de l'inverse de la matrice de corrélation d'un processus AR. La complexité de cet algorithme est marginalement supérieure à celle du LMS. Nous présentons ensuite une version équivalente à un algorithme Newton-LMS (adaptation échantillon par échantillon), utilisant les idées développées dans [1]. Des simulations préliminaires démontrent que ces deux algorithmes ont des caractéristiques de convergence supérieures aux algorithmes LMS et BLMS, pour des entrées corrélées stationnaires ou non-stationnaires (parole).

ABSTRACT

In this contribution, we present two new block Newton type algorithm based on the particular factorization of the inverse of the correlation matrix of an A.R. process. The computational load is only slightly increased with respect to the standard LMS algorithm. A version equivalent to the Newton-LMS (where the adaptation of the filter coefficients is performed on a sample-by-sample rather than a block-by-block basis) is then presented, based on the ideas developed in [1]. The simulations show superior convergence characteristics when dealing with correlated stationary as well as non-stationary input signals (such as speech).

1 Introduction

L'annulation d'échos acoustiques est un problème qui se pose dans un grand nombre d'applications (téléphonie main-libre, téléconférence...). Les difficultés essentielles rencontrées sont d'une part liées à la durée des réponses impulsionnelles des canaux acoustiques à identifier et d'autre part à la nature des signaux à traiter.

Dans ce contexte, les algorithmes classiques s'avèrent inappropriés. En effet, les algorithmes de type LMS temporel convergent trop lentement et ceux de type MCR sont trop complexes [2, 3]. La convergence du LMS est principalement affectée par la dispersion des valeurs propres de la matrice d'autocorrélation du signal d'entrée. La méthode utilisée généralement afin de s'affranchir de ce problème est de blanchir le signal d'entrée sans toutefois recourir à une structure de type MCR "standard". Des exemples de ce type d'approche sont les implantations fréquentielles du LMS[4], les divers algorithmes à structures prédictives [5, 6] où encore des variantes simplifiées à prédiction de taille réduite [3].

Nous proposons dans cette contribution, un algorithme par blocs de type quasi-Newton fondé sur la factorisation particulière de l'inverse de la matrice de corrélation des processus Auto-Régressifs : le Block Auto-Regressive Normalized-LMS ou BARN-LMS. A la différence des algorithmes à structure prédictive classique, le prédicteur est estimé par bloc éliminant ainsi le couplage induit par les algorithmes adaptatifs de prédiction et d'adaptation qui rendent délicate l'utilisation de telles structures dans des situations pratiques. Nous présentons ensuite une version de type "Nested-LMS", exploitant les idées développées dans [1]: cette version est équivalente à un algorithme Newton-LMS (le filtre est adapté échantillon par échantillon et non pas bloc-par-bloc).

Afin d'évaluer les performances de ces algorithmes, nous

présentons quelques simulations préliminaires démontrant des améliorations significatives de la vitesse de convergence par rapport à des algorithmes LMS et BLMS.

2 Présentation du modèle

L'annulation d'échos acoustiques est un problème d'identification adaptative d'un filtre à réponse impulsionnelle finie (RIF) modélisant le chemin d'écho -prenant en compte, les réponses des transducteurs et de la salle, réponses supposées linéaires-. Les signaux observés sont d'une part le signal d'entrée du canal (mesuré en entrée du système de diffusion), noté $x(n)$ et le signal de référence, noté, $y(n)$ (mesuré en sortie du système d'acquisition).

Dans cette contribution, nous nous intéresserons principalement à des implantations par blocs des algorithmes. Les coefficients du filtre à identifier sont actualisés par blocs de N échantillons (même si la règle d'adaptation est choisie de façon telle à implémenter un algorithme LMS [1]). Nous notons: L la longueur de la réponse impulsionnelle (supposée connue), $h_s = [h_s(0), \dots, h_s(L-1)]$ les coefficients du filtre à identifier à l'itération s (instant sR) et $y_s = [y_{sN}, \dots, y_{sN+N-1}]^T$ le bloc courant des signaux de référence¹.

2.1 Algorithmes par transformées

Pour s'affranchir de la corrélation du signal d'entrée, l'approche classique consiste à 'normaliser' le gradient stochastique par un gain matriciel Γ approximant l'inverse de la matrice d'autocorrélation de ce signal. La formule

¹Dans cette contribution, les vecteurs seront représentés en caractères minuscules, les matrices en caractères majuscules gras et l'exposant T désigne l'opérateur de transposition.



d'adaptation, dérivée de la version par bloc du gradient stochastique, est alors de la forme:

$$h_{s+1} = h_s + \mu \Gamma \mathbf{X}_s \epsilon_s \quad (1)$$

où μ est le pas d'adaptation, \mathbf{X}_s est la matrice Toeplitz de filtrage $L \times N$ formée à partir des échantillons du signal d'entrée,

$$\mathbf{X}_s = \begin{pmatrix} x(sN) & \dots & x(sN + N - 1) \\ \vdots & \ddots & \vdots \\ & & x(sN) \\ \vdots & & \vdots \\ x(sN - L + 1) & \dots & x(sN - L + N) \end{pmatrix} \quad (2)$$

et ϵ_s est le vecteur $N \times 1$ des erreurs a priori:

$$\epsilon_s = [\epsilon_{sN}, \dots, \epsilon_{sN+N-1}]^T = \mathbf{y}_s - \mathbf{X}_s h_s \quad (3)$$

Le problème principal consiste à trouver une forme du gain matriciel Γ permettant d'obtenir des algorithmes rapides (de l'ordre de complexité du LMS). Dans les algorithmes transformés par blocs, la matrice de normalisation Γ est décomposée de la façon suivante

$$\Gamma = \mathbf{U} \mathbf{A} \mathbf{U}^T \quad (4)$$

où \mathbf{U} est la matrice unitaire de transformation et \mathbf{A} est la matrice diagonale de normalisation. En appelant ν_s les coefficients du filtre dans le domaine transformé $\nu_s = \mathbf{U}^T h_s$ et $\mathbf{Z}_s = \mathbf{U}^T \mathbf{X}_s$ la matrice des signaux d'entrée dans le domaine transformé, la formule d'adaptation prend la forme suivante

$$\begin{aligned} \epsilon_s &= \mathbf{y}_s - \mathbf{X}_s^T h_s = \mathbf{y}_s - \mathbf{Z}_s^T \nu_s \\ \nu_{s+1} &= \nu_s + \mu \mathbf{A} \mathbf{Z}_s \epsilon_s \end{aligned} \quad (5)$$

L'obtention d'algorithmes rapides est lié à l'utilisation de matrices transformations particulières, permettant d'évaluer le produit $\mathbf{U} \mathbf{X}_s$ sous forme d'une transformée par blocs rapides (de type DCT ou DFT). Cette transformation peut, dans certains cas, être aussi utilisée pour évaluer les N échantillons du produit de convolution $\mathbf{X}_s^T h_s$, de manière rapide.

Les performances de l'algorithme bloc par transformés reposent alors sur la validité de l'approximation diagonale de la matrice de corrélation du signal d'entrée dans la base des vecteurs colonnes de la matrice \mathbf{U} .

2.2 algorithmes à structure prédictive (BARN-LMS)

Dans les algorithmes à structure prédictive que nous considérons (BARN-LMS), l'approche utilisée pour approximer l'inverse de la matrice de corrélation du signal d'entrée, est fondée sur la factorisation particulière de l'inverse de la matrice de corrélation d'un processus autorégressif stationnaire.

Supposons que le signal d'entrée soit un processus AR d'ordre p , $p \ll L$. Cette hypothèse est raisonnable dans le contexte de l'annulation d'échos acoustiques car (i) les tailles des filtres à identifier sont longues (de l'ordre de quelques centaines de coefficients pour la téléphonie main libre jusqu'à quelques milliers pour la téléconférence), (ii) le signal de parole peut être raisonnablement modélisé par un processus AR d'ordre 10 à 20.

Soient $a_k(l)$ $1 \leq k \leq l \leq p$ et $\sigma^2(l)$ $1 \leq l \leq p$ les coefficients de prédiction direct et les variances de prédiction aux différents ordres $l \leq p$. L'inverse de la matrice de corrélation peut se factoriser sous la forme :

$$\Gamma = \mathbf{A}^T \mathbf{P}^{-1} \mathbf{A} \quad (6)$$

où \mathbf{P} est une matrice diagonale dont les coefficients sont les inverses des variances de prédiction aux différents ordres

$$\mathbf{P} = \text{diag}[\sigma(p)^{-2}, \dots, \sigma(p)^{-2}, \sigma(p-1)^{-2}, \dots, \sigma(1)^{-2}] \quad (7)$$

et \mathbf{A} est la matrice triangulaire supérieure dont les $L-p$ premières lignes correspondent aux coefficients de prédiction d'ordre p et les p dernières aux coefficients de prédiction aux ordres inférieures :

$$\mathbf{A} = \begin{pmatrix} 1 & a_1 & \dots & a_p & 0 & \dots & 0 \\ 0 & \ddots & & \ddots & \ddots & & \vdots \\ \vdots & \ddots & & & & \ddots & 0 \\ \vdots & & & 1 & a_1 & \dots & a_p \\ & & & & 1 & \dots & a_{p-1} \\ 0 & \dots & \dots & & \ddots & 0 & 1 \end{pmatrix} \quad (8)$$

L'algorithme associé à cette factorisation particulière de la matrice Γ prend donc la forme suivante

$$\begin{aligned} \epsilon_s &= \mathbf{y}_s - \mathbf{X}_s^T h_s \\ h_{s+1} &= h_s + \mu \mathbf{A}^T \mathbf{P}^{-1} \mathbf{A} \mathbf{X}_s \epsilon_s \end{aligned} \quad (9)$$

Par analogie avec les algorithmes par transformés, la matrice de prédiction \mathbf{A} joue le rôle de la matrice de transformation, et la matrice diagonale \mathbf{P} le rôle de la matrice de normalisation. Notons que si l'ordre de prédiction est $p = 0$, l'algorithme proposé devient équivalent à une version normalisée du LMS (i.e., le pas d'adaptation est divisé par l'énergie du signal d'entrée). Cette structure est aussi très similaire aux algorithmes à structure blanchissante symétrique proposés dans [6].

L'obtention d'algorithmes rapides est liée à la structure particulière de la matrice de prédiction: l'évaluation du produit matriciel $\mathbf{A} \mathbf{X}_s$ se réduit essentiellement à calculer les résidus de prédiction d'un modèle AR d'ordre p sur un bloc de $N + L$ échantillons.

3 Implantation du BARN-LMS

Nous décrivons dans cette section une implantation directe des algorithmes BARN-LMS. Des versions accélérées, utilisant des techniques de convolution rapide par blocs, seront présentées à la conférence.

3.1 Evaluation des coefficients de prédiction

En pratique, les coefficients et les variances de prédiction aux différents ordres sont inconnus et évoluent au cours du temps. Il est donc nécessaire de les actualiser pour chaque nouveau bloc de signal d'entrée. Cette procédure est effectuée en 2 étapes:

- *actualisation des coefficients de corrélation*: on utilise un estimateur récursif de la forme

$$\begin{aligned} \tau(s; \tau) &= \lambda \tau(s-1; \tau) + \frac{(1-\lambda)}{N} \sum_{t=sN}^{sN+N-1} \mathbf{x}(n) \mathbf{x}(n-1) \\ 1 &\leq \tau \leq p \end{aligned}$$

où $0 < \lambda < 1$.



- *évaluation des coefficients de prédiction*: on met en oeuvre un algorithme de type Split-Levinson. Le nombre d'opérations induit par ce calcul est de l'ordre de $O(p^2)$ opérations (ce nombre d'opérations est négligeables par rapport aux nombres d'opérations liées au filtrage et à l'adaptation, qui sont de l'ordre de $O(LN)$ opérations pour un bloc de taille N).

Pour des tailles de blocs courtes, cette procédure peut s'avérer inefficace. Elle sera alors remplacée par une estimation récursive des coefficients de prédiction, utilisant des algorithmes de type MCR rapides.

3.2 calcul de $\mathbf{A}_s \mathbf{X}_s$

Les $L - N$ premières lignes du produit $\mathbf{A}_s \mathbf{X}_s$ correspondent aux échantillons (ordonnés de façon appropriés) du produit de convolution du bloc de signal $[x(sN - L + 1), \dots, x(sN + N - 1)]$ par le filtre prédicteur d'ordre p , $\mathbf{a}_s(p) = [1, a_{s1}(p), \dots, a_{sp}(p)]^T$. Cette convolution peut soit être implantée en utilisant une structure directe (le coût de calcul est alors de $O(p(L+N))$ par blocs), soit une structure de convolution rapide.

Les N dernières lignes correspondent aux résultats (ordonnés de façon appropriée) de la convolution du bloc de signal $[x(sN - L + 1), \dots, x(sN + N - 1)]$ par les prédicteurs d'ordre $\leq p$, $\mathbf{a}_s(1), \dots, \mathbf{a}_s(p-1)$. Les résultats de ces convolutions peuvent être obtenus directement en utilisant une implantation en treillis du filtre de prédiction d'ordre p et en prélevant les échantillons de l'erreur directe aux différents étages du treillis: la complexité globale de cette opération est de l'ordre de $O(2pN)$.

3.3 Evaluation du gradient

Le calcul du vecteur d'erreur a priori ϵ_s (Eq:9) est le même que dans le cas de l'algorithme BLMS. En implantant le filtre h_s sous sa forme directe, ce calcul requiert NL multiplications et NL additions (NL opérations sur un processeur possédant une opération multiplication/accumulation). Ce nombre d'opérations peut être significativement réduit en utilisant des structures de filtrage par blocs rapides:

- soit dans le domaine temporel pour les blocs courts ($N \leq 64$) -structure du type Fast Running Filter-
- soit dans le domaine fréquentiel pour des blocs plus longs -convolution du type Overlap-Save ou Overlap-Add-

La complexité totale de l'algorithme ramené à l'échantillon est donc de l'ordre de $O(p^2)/N + 2O(L) + 2O(\frac{pL}{N})$. Elle est donc comparable à celle d'un algorithme LMS, dès que la taille du bloc est 2 à 3 fois plus grande que l'ordre de prédiction retenu.

4 Nested-AR-NLMS

Dans l'implantation du BARN-LMS décrite ci-dessus, les coefficients du filtre à adapter sont constants pendant toute la durée d'un bloc d'échantillons (i.e, les coefficients sont adaptés bloc-par-bloc). L'erreur a priori n'est estimée qu'à chaque nouveau bloc de données. En s'inspirant des structures introduites dans [1], nous dérivons maintenant un nouvel algorithme qui est un équivalent exact d'un algorithme normalisé à l'échantillon: l'erreur a priori que nous

calculons correspond exactement à celle que nous aurions obtenu en implantant un algorithme LMS normalisé par la matrice Γ_s . Toutefois, à la différence d'une implémentation standard, les coefficients du filtre ne sont explicitement évalués que tous les N échantillons.

Nous notons: h_n le vecteur des coefficients du filtre adaptatif à l'instant n (à la différence des notations précédentes, l'indice temporel n réfère à un indice d'échantillon et non à un indice bloc) et η_n l'erreur a priori associée. Considérons les équations d'adaptation de l'algorithme LMS à gain matriciel sur le bloc d'échantillons $[sN, \dots, sN + N - 1]$:

$$\begin{cases} \eta_{sN} & = y_{sN} - X_s^T(0)h_{sN} \\ h_{sN+1} & = h_{sN} + \mu \Gamma_s X_s(0)\eta_{sN} \\ \vdots & \vdots \\ \eta_{sN+N-1} & = y_{sN+N-1} - X_s^T(N-1)h_{sN+N-1} \\ h_{sN+N} & = h_{sN+N-1} + \mu \Gamma_s X_s(N-1)\eta_{sN+N-1} \end{cases} \quad (11)$$

Soit $\epsilon_s = [\epsilon_{sN}, \dots, \epsilon_{sN+N-1}]^T$ le bloc de N échantillons d'erreurs a priori obtenu en figeant les coefficients du filtre adaptatif à l'instant sN :

$$\epsilon_s = \mathbf{y}_s - \mathbf{X}_s^T h_{sN} \quad (12)$$

Il est facile de voir que les deux erreurs a priori sont reliés par l'ensemble de relations suivantes

$$\begin{aligned} \eta_{sN} &= \epsilon_{sN} \\ &\vdots \\ \eta_{sN+N-1} &= \epsilon_{sN+N-1} - \mu X_s^T(N-1)\Gamma_s \sum_{j=0}^{N-2} X_s(j)\eta_{sN+j} \end{aligned} \quad (13)$$

$$\eta_{sN+N-1} = \epsilon_{sN+N-1} - \mu X_s^T(N-1)\Gamma_s \sum_{j=0}^{N-2} X_s(j)\eta_{sN+j} \quad (14)$$

Comme le montre ces équations, le calcul des erreurs a priori à l'intérieur de chaque bloc ne nécessite donc pas d'effectuer le calcul explicite des valeurs successives présent par le filtre adaptatif h_{sN+i} $1 \leq i \leq N-1$. Il suffit en effet de disposer:

- des erreurs a priori $[\epsilon_{sN}, \dots, \epsilon_{sN+N-1}]$ associées au filtrage du signal d'entrée par h_{sN} (le filtre étant constant pendant le durée du bloc, ce filtrage peut-être implanté de façon rapide)
- des valeurs des facteurs de corrections de l'erreur a priori associée au filtrage fixe; on voit aisément que ces facteurs de corrections font intervenir principalement les valeurs de coefficients de corrélations du signal d'entrée blanchi par le filtre prédicteur (ces corrélations peuvent être elles aussi évaluées de façon rapide en utilisant des techniques de calcul dans le domaine fréquentiel).

L'accroissement de complexité par rapport à l'algorithme bloc, est ² de l'ordre de L opérations par échantillons. Différentes implémentations rapides ont été développées et seront présentées lors de la conférence.

²on ne tient pas compte ici des possibilités de rendre ces algorithmes plus rapides en utilisant des structures de calcul de convolution et de corrélation rapides



5 Simulations

Les situations que nous avons retenues sont représentatives de l'annulation d'écho acoustique; la réponse à identifier est un filtre FIR obtenu en tronquant une réponse impulsionnelle mesurée dans une voiture (fréquence d'échantillonnage 8 kHz). Les signaux d'entrée du canal sont (i) soit un signal stationnaire à spectre moyen de parole (bruit USASI défini par le CCITT (Fig:1, longueur du filtre 256 points, taille du bloc 64 points) (ii) soit un signal de parole (Fig:2, longueur du filtre 512 points, taille du bloc 128 points). Dans ces deux situations, l'ordre du modèle AR identifié est de 15. Le pas du BLMS est choisi ici de façon à atteindre le même désajustement (entrée bruit blanc, canal stationnaire) asymptotique que l'algorithme BARN-LMS.

On peut voir sur ces figures Fig:1,2 que les algorithmes BARN-LMS et le Nested-BARN-LMS se comportent mieux que le BLMS, lorsque le signal d'entrée du canal est corrélé, que celui-ci soit stationnaire ou non-stationnaire. Pour le signal de parole, la réjection de l'écho atteint pour les deux méthodes BARN-LMS est en moyenne de l'ordre de 30 dB (elle est de l'ordre de 10 dB pour le BLMS).

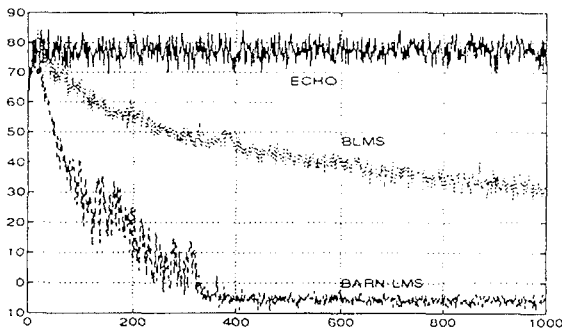


Figure 1: comparaison du BLMS et du BARN-LMS pour un signal de bruit USASI

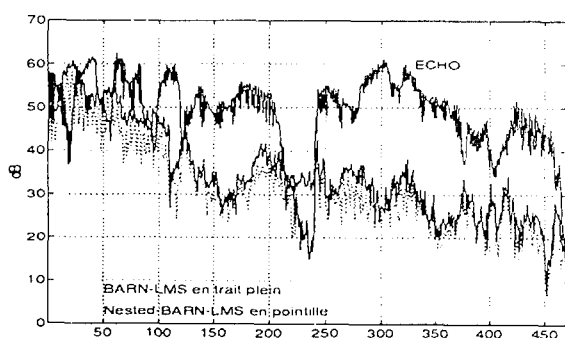


Figure 2: comparaison du Nested-BARN-LMS et du BARN-LMS pour un signal de parole

6 Conclusion

Nous avons présenté dans cette contribution deux nouveaux algorithmes d'identification adaptative de systèmes linéaires à réponse impulsionnelle finie, le BARN-LMS et le Nested-BARN-LMS. Le BARN-LMS est une implémentation d'un algorithme Quasi-Newton BLMS, où inverse

de la matrice de covariance du signal d'entrée est approximée par le produit d'une matrice de prédiction et d'une matrice de normalisation (réminiscentes de la factorisation de Cholevski rapide de l'inverse des matrices de Toeplitz symétriques). Le Nested-BARN-LMS est une implémentation par bloc d'un quasi-Newton-LMS: les erreurs a priori correspondent exactement à l'algorithme adaptatif échantillon-par-échantillon; les coefficients du filtre ne sont explicitement évalués que tout les blocs.

Des simulations préliminaires démontrent que ces deux algorithmes convergent significativement plus vite que le LMS (ou le BLMS), lorsque les entrées du système sont corrélées. Nous présenterons dans une version étendue de ce papier les performances théoriques de ces algorithmes et l'influence de l'estimation des matrices de prédiction et de normalisation sur le désajustement asymptotique et sur la vitesse de convergence.

References

- [1] J. Benesty, *Algorithmes du type gradient à complexité de calcul réduite et vitesse de convergence accrue*. PhD thesis, Université Paris XI, Orsay, 1991.
- [2] O. Macchi and M. Bellanger, "Filtrage adaptatif transverse," *Traitement du Signal*, vol. 5, pp. 115–132, 1988.
- [3] A. Gilloire, "State-of-the-art in acoustic echo cancellation," tech. rep., Publications da Universidade de Vigo, 1991. *adaptive Algorithms: Applications and Non Classical Schemes*.
- [4] O. A. Amrane, *identification de systèmes à réponse impulsionnelle longue par filtrage adaptatif en fréquence: application à l'annulation d'échos acoustiques*. PhD thesis, Ecole Nationale Supérieure des Telecoms, sept 1992.
- [5] E. Hansler and U. Schulteiss, "On-line identification of loudspeaker-enclosure-microphone system for compensating acoustical echoes," in *Proc. IFAC Identification and system parameter estimation*, pp. 1841–1846, 1985.
- [6] M. Mboup, *Identification adaptative par structure prédictives et récursives: applications à l'annulation d'échos acoustiques*. PhD thesis, Université Paris XI, Orsay, 1992.