



Déconvolution autodidacte mono-source/multi-capteurs: Techniques au second ordre.

K. Abed-Meraim, Ph. Loubaton, E. Moulines

Télécom Paris / CNRS URA 820 / GdR TdSI, 46 rue Barrault, 75634 PARIS

email: abedmera@sig.enst.fr

Résumé

Très récemment, un certain nombre de travaux [1, 2, 3, 4] ont mis en évidence que, sous certaines hypothèses relativement réalistes, le fait de disposer de plusieurs capteurs permet de résoudre le problème de l'égalisation autodidacte en utilisant uniquement les statistiques du second ordre des observations. Cet article est consacré à une étude comparative de certaines de ces techniques d'identification. Plus précisément, la comparaison portera sur les performances statistiques et la robustesse à une erreur sur la détermination de l'ordre du modèle. Quelques simulations numériques dans le contexte des communications radio-mobiles seront présentées.

Abstract

Blind channel identification methods based on the multi-channel or oversampled channel output is a problem of current theoretical and practical interest [1, 2, 3, 4]. It is demonstrated that, using multi-output observations, the second order blind identification is possible under some mild hypotheses. In this contribution, we present a comparative study between the subspace method and the linear prediction approach. We will show the performance superiority of the subspace approach, when the exact order is known. The effect of overestimating the channel order is investigated by simulations: it is demonstrated that the prediction error method is "robust" to over-determination, in contrast to the subspace methods.

1 Introduction

Des travaux récents [1, 3] ont donné dernièrement un nouvel essor au problème de la déconvolution autodidacte mono-entrée/multi-sorties. Le problème posé est le suivant: On observe à la sortie d'une antenne (ou simplement par sur-échantillonnage, voir [3]) un processus aléatoire vectoriel de dimension q à temps discret $y(n) = [y_1(n), \dots, y_q(n)]^T$ supposé être la sortie d'un filtre vectoriel causal $q \times 1$ de fonction de transfert polynômiale (de degré M)¹ inconnue $h(z) = [h_1(z), \dots, h_q(z)]^T$ excité par une suite de variables aléatoires scalaires $v(n)$ (signal source), avec éventuellement un bruit additif $w(n)$:

$$y(n) = [h(z)]v(n) + w(n) \quad (1)$$

Il s'agit de restituer $h(z)$ au vu des propriétés statistiques de $y(n)$, ou d'une façon plus réaliste à partir d'un nombre fini d'observations $y(1), \dots, y(T)$. L'identification de $h(z)$ permet alors la reconstruction de la suite excitante $v(n)$, qui dans de nombreux contextes applicatifs tels que l'égalisation, représente l'information à extraire.

Avant d'aborder ce problème, rappelons que celui-ci a déjà fait l'objet d'études approfondies dans le cas scalaire mono-entrée/mono-sortie. En particulier, il est montré que lorsque la fonction de transfert $h(z)$ est supposée être d'inverse causal (ce qui dans le cas rationnel équivaut à $h(z) \neq 0$ pour $|z| \geq 1$) le problème est solvable uniquement à l'ordre 2, puisque $h(z)$ n'est autre que la factorisation causale à phase minimale de la densité spectrale (débruitée) $S_y(z)$ de l'observation y . Malheureusement, cette hypothèse est extrêmement restrictive, et n'est pas vérifiée dans des applications importantes comme l'égalisation autodidacte. Il est alors clair que le problème considéré ne

peut être résolu à partir des statistiques du second ordre de $y(n)$, et lorsque la suite $v(n)$ est non-gaussienne, il faut avoir recours aux statistiques d'ordre supérieurs [5].

Dans le cas vectoriel qui nous intéresse, la condition sur la fonction de transfert d'être d'inverse causal ($h(z) \neq 0$ pour $|z| \geq 1$) est beaucoup moins restrictive. Il suffit pour la satisfaire que:

$$(H1) \quad h(z) \neq 0 \quad \forall z$$

Autrement dit, que les polynômes $h_i(z)$, $1 \leq i \leq q$ n'aient pas de zéros en commun, ce qui en pratique est toujours vérifié. Partant de cette constatation, nous allons présenter quelques approches récentes utilisant seulement les statistiques d'ordre 2 de l'observation pour réaliser l'identification autodidacte. Dans ce qui suit nous présentons d'abord la technique sous-espace et la technique de prédiction linéaire. Une brève étude de la robustesse à une sur-détermination de l'ordre du modèle est développée par la suite. Une étude par simulations numériques des performances statistiques dans le contexte des communications radio-mobiles, est finalement présentée.

2 La méthode sous-espace

Cette technique d'abord introduite dans [3] puis reformulée dans [6], s'apparente très fortement aux techniques sous-espace du traitement d'antenne. Elle est fondée sur certaines propriétés des sous-espaces propres de la matrice $q(N+1) \times (M+N+1)$ de Sylvester $T_N(h)$ associée à la fonction de transfert polynômiale $h(z)$, matrice définie

¹Par définition $\deg(h(z)) = \max_{1 \leq i \leq q} \deg(h_i(z))$.



par:

$$\mathcal{T}_N^h(h) = \begin{pmatrix} h(0) & \dots & h(M) & 0 & \dots & 0 \\ 0 & h(0) & \dots & h(M) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h(0) & \dots & h(M) \end{pmatrix}$$

Le résultat fondamental sur lequel on s'appuie est le suivant (voir [6]).

Théorème 1 Soit N tel que $N \geq M$. Sous l'hypothèse (H3), nous avons:

1. la matrice $\mathcal{T}_N(h)$ est de rang complet égal à $(M + N + 1)$.
2. Soit $h'(z)$ une fonction de transfert polynômiale $q \times 1$ de degré $M' \geq M$. Les deux conditions suivantes sont équivalentes:

$$\Pi \mathcal{T}_N(f) = 0$$

\Updownarrow

$$f(z) = \alpha(z)h(z), \quad \alpha(z) \text{ polynôme scalaire.}$$

où Π est la matrice de projection orthogonale sur l'orthogonal de l'espace image de $\mathcal{T}_N(h)$, i.e $\text{Im}(\mathcal{T}_N(h))^\perp$.

Basé sur le résultat du théorème 1, on peut procéder de la façon suivante: En utilisant le modèle de signal (1) et en supposant le bruit additif spatialement et temporellement blanc de puissance σ^2 , la matrice de covariance \mathcal{R}_N de $Y_N(n) = [y(n)^T, \dots, y(n-N)^T]^T$ est donnée par:

$$\mathcal{R}_N = \mathcal{T}_N(h) \mathcal{V}_{M+N} \mathcal{T}_N^H(h) + \sigma^2 I_{q(N+1)} \quad (2)$$

où \mathcal{V}_{M+N} est la matrice de covariance de $V_{M+N}(n) = [v(n), \dots, v(n-M-N)]^T$ supposée de rang complet. Sous l'hypothèse (H1), \mathcal{R}_N admet σ^2 comme plus petite valeur propre et Π est la matrice de projection orthogonale sur le sous espace propre associé. Ceci nous conduit à la procédure d'identification suivante (voir [3] pour plus de détails):

- Estimer à partir des observations la matrice de covariance \mathcal{R}_N et par décomposition propre de celle-ci, la matrice de projection Π ,
- Estimer h (à un facteur scalaire près) en minimisant en f le critère quadratique suivant:

$$\|\hat{\Pi} \mathcal{T}_N(f)\|^2 = \text{Tr}(\mathcal{T}_N^H(f) \hat{\Pi} \mathcal{T}_N(f)) \quad (3)$$

sous une certaine contrainte permettant d'éviter la solution triviale $f = 0$ (nous choisirons une contrainte quadratique de la forme $\sum_{k=0}^M f^*(k)f(k) = 1$).

Extension au cas d'un bruit coloré: Considérons à présent le cas où le bruit $w(n)$ est blanc temporellement, mais de covariance spatiale $E(w(n)w(n)^*) = Q$ totalement inconnue et supposons que la suite $v(n)$ est i.i.d (supposée par convenance, de variance unité). L'identification de l'image de $\mathcal{T}_N(h)$ ne peut se faire directement à partir de l'autocovariance \mathcal{R}_N de $Y_N(n)$. Nous considérons alors la matrice:

$$\mathcal{R}_s = E(Y_{M-1}(n)Y_{M-1}^*(n-M)) + E(Y_{M-1}(n)Y_{M-1}^*(n-M))^H$$

$$= \mathcal{T}_{M-1}(h) \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \mathcal{T}_{M-1}(h)^H \quad (4)$$

où I désigne la matrice identité $M \times M$. Clairement, l'espace image de \mathcal{R}_s coïncide avec celui de $\mathcal{T}_{M-1}(h)$. Modulo une condition légèrement plus forte que (H1) (voir [7]), on peut alors reconstituer $h(z)$ à une constante près à partir de $\text{Im}(\mathcal{T}_{M-1}(h))$. Le projecteur Π peut être estimé ici à partir de la matrice \mathcal{R}_s .

3 La prédiction linéaire.

Dans cette section, nous supposons que (H2) le signal émis $v(n)$ est une séquence de variables aléatoires i.i.d centrée et de variance unité ² et que le bruit additif est temporellement et spatialement blanc. Nous présentons ici une technique basée sur le principe de la prédiction linéaire, celle-ci a été présentée en premier dans [8] dans le cas $q = 2$ puis étendue dans [4].

Pour simplifier la présentation, le cas sans bruit $w(n) = 0$ est d'abord considéré. Une extension immédiate dans le cas bruité est ensuite présentée.

L'idée de base de l'approche de prédiction linéaire est de voir que le processus à moyenne ajustée (MA) $y(n) = [h(z)]v(n)$ est aussi un processus autoregressif (AR) d'ordre fini. Cette propriété provient de l'identité de Bezout généralisée qui assure, sous l'hypothèse (H1), l'existence d'un polynôme vectoriel $1 \times q$, $g(z)$, tel que: $g(z)h(z) = 1$. En appliquant $g(z)$ au processus $y(n)$ il vient que $[g(z)]y(n) = v(n)$: $h(z)$ peut être exactement inversé par un filtre FIR causal. Plus précisément, nous avons le résultat suivant (voir [4]):

Théorème 2 Sous les hypothèses (H1) et (H2), $y(n)$ est un processus autoregressif d'ordre M . Son processus d'innovation est donné par $i(n) = h(0)v(n)$.

$v(n)$ représente le processus d'innovation normalisé de $y(n)$, et un filtre polynômial $1 \times q$ $g(z)$ vérifiant $[g(z)]y(n) = v(n)$ peut être vu comme un *filtre de prédiction*. Un tel filtre peut s'obtenir en résolvant les équations généralisées de Yule-Walker. Enfin la fonction de transfert polynômiale $h(z)$ peut se calculer à partir d'un quelconque filtre de prédiction $g(z)$ en utilisant $h(k) = E(y(n)[g(z)]y(n-k)^*)$.

Le processus d'innovation $i(n)$ est calculé en projetant $y(n)$ sur son passé généré par les variables aléatoires $\{y_i(n-l)/i = 1, q, l = 1, P\}$ où $P \geq M$. Soit $[A(1), \dots, A(P)]$ les $q \times q$ matrices telles que $y(n) + \sum_{k=1}^P A(k)y(n-k) = i(n)$:

$$[A(1), \dots, A(P)] \mathcal{R}_{P-1} = -[r(1), \dots, r(P)] \quad (5)$$

où $r(k) \stackrel{\text{def}}{=} E(y(n+k)y(n)^*)$ et \mathcal{R}_{P-1} est la matrice de covariance de $Y_{P-1}(n) = [y(n), \dots, y(n-P+1)]^T$. Il est à noter que, du fait que la covariance \mathcal{R}_{P-1} n'est pas de rang plein, ces coefficients ne sont pas définis de manière unique. Une solution particulière est donnée par:

$$[A(1), \dots, A(P)] = -[r(1), \dots, r(P)] \mathcal{R}_{P-1}^\# \quad (6)$$

où $A^\#$ désigne la pseudo-inverse de la matrice A . Selon le théorème 2, la matrice de covariance du processus d'innovation $D = h(0)h(0)^* = r(0) + \sum_{k=1}^P A(k)r^H(k)$ est

² Au fait il suffit que les variables soient temporellement décorréelées.

de rang 1; la valeur propre non nulle de D est égale à $\lambda_d = \|\hat{h}(0)\|^2$ et le vecteur propre associé est $d = \hat{h}(0)/\|\hat{h}(0)\|$. En prenant $l = d/\sqrt{\lambda_d}$, il vient

Lemme 1 $g(z) = l^*(I + \sum_{k=1}^P A(k)z^{-k})$ est un vecteur polynômial $1 \times q$ tel que $[g(z)]y(n) = v(n)$.

En résumé nous avons la procédure d'identification suivante:

- Estimer les coefficients du filtre de prédiction du processus $y(n)$ par l'équation de Yule-Walker (6).
- Estimer le vecteur $l = \hat{h}(0)/\|\hat{h}(0)\|^2$ par la décomposition propre de la covariance estimée $\hat{D} = \hat{r}(0) + \sum_{k=1}^P \hat{A}(k)\hat{r}^H(k)$. Les coefficients du filtre prédictif $g(z)$ sont alors estimés par:

$$[\hat{g}(0), \dots, \hat{g}(P)] = \hat{l}^*[I, \hat{A}(1), \dots, \hat{A}(P)]$$

- Estimer les coefficients du canal de transmission par:

$$\hat{h}(k) = \sum_{l=0}^P \hat{r}(l+k)\hat{g}(l)^*$$

L'extension au cas bruité est immédiate, il suffit d'estimer la puissance du bruit additif σ^2 et de remplacer là où elle apparaît l'autocovariance $r(0)$ par l'autocovariance débruitée $r(0) - \sigma^2 I$.

4 Robustesse

Comme en pratique l'ordre M du filtre est inconnu, il sera nécessaire de l'estimer avant de procéder à l'identification du canal. On se propose dans cette partie, d'étudier la robustesse des méthodes proposées vis à vis d'une erreur sur l'ordre du modèle.

En ce qui concerne l'approche sous espace, lorsque le degré de $h(z)$ est sur-estimé ($\hat{M} > M$), le théorème 1 montre que la méthode sous-espace estime $h(z)$ à un polynôme scalaire près; $\hat{h}'(z) = \alpha(z)h(z)$. La solution désirée sera donc obtenue par factorisation des zéros communs aux entrées du polynôme vectoriel $\hat{h}'(z)$. En pratique, lorsque la matrice de covariance \mathcal{R}_N est estimée à partir des données, le filtre estimé $\hat{h}'(z)$ sera une approximation d'un certain polynôme $\alpha(z)h(z)$. Il faut alors faire appel à un second algorithme qui décidera combien de zéros sont en commun et factorisera les zéros communs approximatifs aux entrées de $\hat{h}'(z)$. Les performances de l'estimateur obtenu dépendront très fortement de l'efficacité d'un tel algorithme. Une sur-estimation de l'ordre conduira à un estimateur complètement erroné.

D'autre part, on peut voir aisément que le filtre $h(z)$ est exactement estimé par l'approche de prédiction linéaire lorsque la séquence d'autocovariance est exactement connue. En pratique, on ne dispose que des statistiques empiriques et une estimation de l'ordre est nécessaire pour réaliser la pseudo-inverse de la matrice de covariance débruitée $\mathcal{R}_{P-1} - \sigma^2 I$. Toutefois, une erreur sur l'estimation de l'ordre n'affecte pas sérieusement les performances de la méthode. Heuristiquement, les raisons

sont les suivantes: d'abord, soit $P > M$ et soit u un vecteur propre 'bruit' de la matrice $\mathcal{R}_{P-1} - \sigma^2 I$: $(\mathcal{R}_{P-1} - \sigma^2 I)u = 0$; de l'équation (5), il vient que $[r(1), \dots, r(P)]u = 0$, les vecteurs propres bruit sont orthogonaux à $[r(1), \dots, r(P)]^H$. Cette relation vraie en statistiques exactes, reste approximativement vérifiée pour les statistiques empiriques. Ensuite, une procédure de détermination de l'ordre du modèle, estime combien de valeurs propres sont égales à σ^2 . Celle-ci échouera lorsqu'une valeur propre 'bruit' λ est 'suffisamment loin' de σ^2 . Ainsi, on s'attend donc à ce que les facteurs $\lambda - \sigma^2$ soient bornés loin de zéro. La contribution d'un vecteur propre 'bruit' dans (6), donnée par $-[\hat{r}(1), \dots, \hat{r}(\hat{M})] \frac{\hat{u}\hat{u}^*}{\lambda - \sigma^2}$ reste donc relativement négligeable. Ceci a été observé sur les simulations numériques qu'on a réalisé.

5 Simulations

Nous décrivons ici des simulations réalisées dans le cadre du modèle GSM des communications radio-mobile [9]. Le signal source est une séquence aléatoire i.i.d de variables discrètes, centrées et de puissance unité, appartenant à une constellation à déplacement de phase à 4 états (MDP4). Le bruit d'observation est une séquence de variables aléatoires gaussiennes indépendantes du signal source. Il est temporellement et spatialement blanc. Le canal à travers lequel ces données seront véhiculées est un canal à trajets multiples affectés chacun d'un retard, d'une atténuation et d'un déphasage. La longueur de l'antenne réceptrice est de 4 capteurs. Au niveau du premier capteur, le canal de transmission est le canal de test à 6 trajets donné par la norme GSM (voir [9]). Nous supposons que chaque trajet suit une loi de propagation plane avec une direction d'arrivée distribuée uniformément dans le secteur angulaire $[0, \pi/3]$. Les canaux de transmission au niveau des autres capteurs sont alors générés en fonction du premier canal et des directions d'arrivée des différents trajets selon la loi de propagation plane (les capteurs sont supposés équidistants). La taille d'échantillons est de 250 et les statistiques sont évaluées sur 100 réalisations.

La figure 1 (respectivement le tableau 1) montre l'erreur quadratique moyenne (MSE) en dB (respectivement le taux d'erreur d'égalisation³) en fonction de la puissance du bruit (en dB)⁴. Elle compare ainsi les performances de la méthode sous espace (SE) et ceux de la prédiction linéaire (PL) pour $P = 5$ et $P = 7$ (on rappelle que P est la longueur du passé sur lequel on fait la prédiction). Ceci montre (i) la supériorité de la méthode sous espace par rapport à la méthode de prédiction linéaire lorsque l'ordre du modèle M est connu, et (ii) le gain en performance lorsqu'on augmente la taille du filtre de prédiction.

Le tableau 2 montre le taux d'erreur sur la détermination de l'ordre du modèle en fonction de σ^2 , pour une fenêtre temporelle de longueur $N = 7$. Le test utilisé pour l'estimation de l'ordre étant le test MDL (Minimum Description Length) [10]. On voit ainsi que les taux d'erreurs

³Ayant choisi des égaliseurs linéaires, l'égalisation pourrait être améliorée si par exemple on choisissait l'égaliseur de Viterbi.

⁴Le SNR à la sortie du canal, tenant compte de la norme des canaux et du nombre de capteurs est obtenu en ajoutant 1.6 dB à la valeur en dB de σ^2 .



sont loins d'être négligeable même pour les faibles puissances du bruit.

Le tableau 3 compare les taux d'erreur d'égalisation pour différentes valeurs de la puissance du bruit et différentes longueurs de la fenêtre temporelles (la ligne $N = 5$ du tableau correspond à l'erreur d'égalisation lorsque l'ordre M est connu). On voit que malgré les forts taux d'erreur sur l'estimation de l'ordre, les performances de la prédiction linéaire restent relativement correctes en particulier pour les bruits de faibles puissances.

6 Conclusion

Cet article présente une étude comparative des méthodes sous-espace et de prédiction linéaire pour l'identification autodidacte de filtres FIR vectoriels. L'avantage de la méthode sous espace réside dans ses faibles hypothèses d'application, son extensibilité au cas du bruit coloré et ses bonnes performances statistiques lorsque l'ordre du modèle est connu. L'avantage majeure de la prédiction linéaire est, contrairement à la méthode sous espace, sa relative robustesse par rapport à une erreur sur la détermination de l'ordre du modèle.

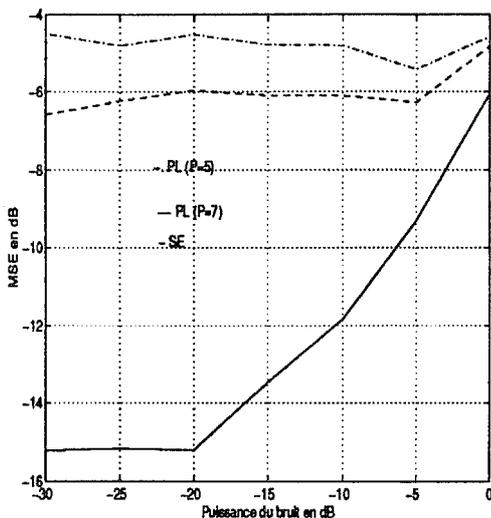


Fig1. EQM en fonction de la puissance du bruit.

σ^2	0	-5	-10	-20	-30
PL P=5	0.2745	0.1574	0.0396	0.0002	0
PL P=7	0.2808	0.1496	0.0319	0.0001	0
SE	0.1900	0.0064	0.0001	0	0

Tab1. Taux d'erreur d'égalisation.

σ^2	0	-5	-10	-15	-20	-30
TE	1.00	0.29	0.38	0.95	0.80	0.78

Tab2. Taux d'erreur sur l'estimation de l'ordre par le test MDL ($N = 7$).

N/σ^2	0	-5	-10	-20	-30
5	0.2745	0.1574	0.0396	0.0002	0
6	0.2676	0.1601	0.0318	0.0038	0.0064
7	0.2707	0.1506	0.0327	0.0026	0.0068
8	0.2762	0.1645	0.0276	0.0048	0.0048
9	0.2777	0.1615	0.0330	0.0044	0.0061
10	0.2746	0.1701	0.0290	0.0085	0.0039

Tab3. Taux d'erreur d'égalisation.

References

- [1] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. of the 25th Asilomar Conference, Pacific Grove, CA*, pp. 856–860, 1991.
- [2] W. Gardner, "A new method of channel identification," *IEEE Tr. on Comm.*, vol. 39, pp. 813–817, August 1991.
- [3] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Tr. on Sig. Proc.*, vol. 43, pp. 516–525, Feb. 1995.
- [4] K. Abed Meraim, P. Duhamel, D. Gesber, P. Loubaton, S. Mayrargue, E. Moulines, D. Slock, "Prediction error methods for time-domain blind identification of multichannel FIR filters," *Proc. ICASSP*, 1995.
- [5] C. Nikias, "Blind deconvolution using higher-order statistics," in *Proc. 2nd Int. Conf. on Higher-Order Statistics* (J.-L. Lacoume, ed.), pp. 49–56, Elsevier, 1992.
- [6] E. M. K. Abed Meraim, P. Loubaton, "A subspace algorithm for certain subspace blind identification problems." submitted to *IEEE trans. on Info. Theory*, 1994.
- [7] K. Abed Meraim, P. Loubaton and E. Moulines, "Subspace method for blind identification of multichannel fir filters in noise field with unknown spatial covariance," in *Proc. ASILOMAR, California, USA.*, Nov. 1994.
- [8] D. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter-banks and multichannel linear prediction," in *Proc. Int. Conf. on Acoust. Speech and Sig. Proc.*, vol. 4, pp. 585–588, 1994.
- [9] J. Proakis, *Digital communications*. New York: Mc Graw-Hill, 1989.
- [10] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Tr. on Sig. Proc.*, vol. 33, pp. 387–392, Apr. 1985.