

Quelques remarques méthodologiques concernant la séparation des mélanges convolutifs par l'approche de décorrélation

⁽¹⁾C. Simon, ⁽¹⁾C. Vignat, ⁽¹⁾Ph. Loubaton, ⁽²⁾C. Jutten, ⁽³⁾G. d'Urso

⁽¹⁾ Equipe Systèmes de Communication, UMLV, 2 rue de la Butte Verte 93166 Noisy le Grand

⁽²⁾ LTIRF/INPG, 46 Avenue Felix Viallet, 38040 Grenoble Cedex

⁽³⁾ EDF/DER, 6 Quai Watier, 78401 Chatou Cedex

RÉSUMÉ

Dans cet article, nous étudions un problème de séparation de mélange convolutif 2 sources / 2 capteurs. Nous étudions la pertinence de la technique dite de décorrélation, consistant à séparer le mélange en adaptant un filtre qui, appliqué aux observations, génère des signaux décorrélés entre eux. Dans le cas d'un filtre de mélange RIF, cette technique permet en théorie de séparer les sources. Notre contribution établit que si le filtre de mélange est RII, les équations exprimant la décorrélation admettent une infinité de solutions dont nous donnons une paramétrisation. Afin de remédier à cet inconvénient, nous suggérons l'utilisation d'une technique basée sur la prédiction linéaire.

ABSTRACT

In this paper, we study the problem of blind separation of a popular and particular convolutive mixture. More specifically, we consider the so-called decorrelation approach. It consists in finding a filter which, driven by the observations, produces decorrelated outputs. It is now well established that this approach allows to separate the sources if the mixing filter is FIR. In this paper, we show that this result is no longer true in the IIR case. In particular, we establish that the decorrelation equations admit infinitely many solutions that we parameterize. In order to overcome this drawback, we propose to use a linear prediction based approach.

1 Introduction

Cet article a pour cadre général le problème de la séparation d'un mélange convolutif de deux sources indépendantes lorsque le nombre de capteurs est aussi égal à deux. Plus précisément, on suppose que deux signaux indépendants $s_1(n)$ et $s_2(n)$ sont observés sur deux capteurs après avoir été mélangés de façon convolutive par un système linéaire inconnu 2 entrées / 2 sorties dont nous noterons $H(z)$ la fonction de transfert. Les signaux y_1 et y_2 prélevés sur les deux capteurs s'écrivent donc :

$$\begin{cases} y_1(n) = [H_{11}(z)]s_1(n) + [H_{12}(z)]s_2(n) \\ y_2(n) = [H_{21}(z)]s_1(n) + [H_{22}(z)]s_2(n) \end{cases}$$

En principe, on définit le problème de la séparation de sources comme celui de la reconstitution des signaux s_1 et s_2 à partir de la seule observation du signal $y(n) = (y_1(n), y_2(n))^T$. Cependant, une analyse rapide de la situation montre que le problème ainsi posé n'a pas de sens puisqu'en l'absence d'information a priori sur les signaux sources, on ne peut espérer reconstituer s_1 et s_2 qu'à un filtre près. En pratique, on se contente donc de chercher à reconstituer la contribution de chaque source sur chaque capteur, i.e. les signaux $[H_{ij}(z)]s_j(n)$ pour les diverses valeurs de i et de j . Un algorithme de séparation de sources permet donc de reconstituer les signaux que l'on observerait si le système inconnu était excité par $[s_1(n), 0]^T$ d'une part et par $[0, s_2(n)]^T$ d'autre part. Puisque les signaux s_1 et s_2 ne sont pas récupérables, on peut les confondre sans restriction notable avec $[H_{11}(z)]s_1(n)$ et $[H_{22}(z)]s_2(n)$ respectivement, ce qui revient à normaliser les termes diagonaux de $H(z)$ à 1.

Dans ces conditions, le modèle des signaux observés s'écrit :

$$\begin{cases} y_1(n) = s_1(n) + [H_{12}(z)]s_2(n) \\ y_2(n) = [H_{21}(z)]s_1(n) + s_2(n) \end{cases}$$

Dans la suite de cet article, les hypothèses classiques suivantes concernant le filtre $H(z)$ seront supposées vérifiées :

- H1 : les filtres $H_{12}(z)$ et $H_{21}(z)$ sont causaux
- H2 : le filtre $H(z)$ est à phase minimale, c'est-à-dire d'inverse causal, ce qui s'exprime de la façon suivante :

$$1 - H_{12}(z)H_{21}(z) \neq 0 \quad \forall z \quad |z| > 1$$

Commentons rapidement ces hypothèses. Tout d'abord, il convient de mentionner que, compte tenu de la normalisation de $H_{ii}(z)$ à 1 pour $i = 1, 2$, l'hypothèse de causalité n'est pas anodine. L'hypothèse que $H(z)$ est à phase minimale est aussi relativement restrictive mais on la voit très souvent formulée : c'est elle qui permet de supposer que l'on peut reconstituer s_1 et s_2 par un filtrage causal de $y(n)$.

La plupart des approches répertoriées dans la littérature consistent à chercher une matrice $G(z)$ de la forme

$$G(z) = \begin{pmatrix} 1 & -G_{12}(z) \\ -G_{21}(z) & 1 \end{pmatrix}$$

telle que son l'application aux observations $y(n)$ génère un signal $r(n) = [r_1(n), r_2(n)]^T = [G(z)]y(n)$ dont les deux composantes sont indépendantes. La forme de $G(z)$ est justifiée par le fait que la séparation est réalisée (i.e. $r(n) = s(n)$) si et seulement si $G_{ij}(z) = H_{ij}(z)$ $i \neq j$

Cependant, la condition d'indépendance des composantes du signal $r(n)$ est difficile à exprimer analytiquement. Pour

cette raison, on recherche souvent la matrice $G(z)$ de façon à annuler simplement un certain nombre de cumulants croisés des signaux $r_1(n)$ et $r_2(n)$. Plus particulièrement, l'approche par décorrélation consiste à annuler simplement les corrélations entre les signaux $r_1(n)$ et $r_2(n)$. Cette approche a été très étudiée, et il a été montré récemment qu'elle permet apparemment d'obtenir la séparation des sources, sous l'hypothèse supplémentaire que les filtres $H_{ij}(z)$ et $G_{ij}(z)$ sont à réponse impulsionnelle finie (RIF).

Nous nous proposons ici de montrer que, dans le cas plus général où les filtres mélangeants $H_{ij}(z)$ ne sont pas à réponse impulsionnelle finie, l'approche par décorrélation n'a plus d'intérêt : elle ne permet pas de réaliser la séparation des sources. Plus précisément, nous allons montrer que, dans ce cas, cette approche conduit, outre à la solution séparante, à un ensemble infini de solutions non séparantes, ensemble dont nous proposerons une paramétrisation très simple. Ce résultat démontre à l'évidence que la séparation par décorrélation est une approche qui souffre d'un problème de robustesse vis à vis des hypothèses sur le filtre $H(z)$.

2 Les solutions des équations de décorrélation

Nous entendons par équations de décorrélation l'ensemble des relations qui expriment que les deux composantes de $[G(z)]y(n)$ sont décorréliées. Nous savons que la solution séparante de ces équations correspond - au facteur $(1 - H_{12}(z)H_{21}(z))^{-1}$ près - à l'inverse du filtre $H(z)$, lequel est supposé causal et d'inverse causal d'après les hypothèses H1 et H2. Cette solution est donc elle-même un filtre causal et d'inverse causal. Il est donc tout à fait naturel de restreindre notre recherche des solutions des équations de décorrélation à la famille des filtres $G(z)$ de même nature.

Afin de simplifier la présentation des résultats, nous allons supposer que les signaux sources s_1 et s_2 sont à spectre rationnel. Ainsi, le concept de factorisation à phase minimale de leur densité spectrale peut être défini sans le moindre problème technique. Plus précisément, on sait qu'il existe deux bruits blancs de variance unité, notés v_1 et v_2 , et deux filtres causaux à phase minimale $f_1(z)$ et $f_2(z)$ pour lesquels $s_i(n) = [f_i(z)]v_i(n)$ pour $i=1,2$. Il est bien connu que les v_i ne sont rien d'autre que les processus d'innovation normalisés des signaux s_i pour $i = 1, 2$.

Soit à présent un filtre $G(z)$ causal à phase minimale solution des équations de décorrélation i.e. tel que les 2 composantes r_1 et r_2 du signal $r(n) = [G(z)]y(n)$ sont des signaux décorréliés. Soient $g_i(z)$ $i = 1, 2$ les factorisations causales à phase minimale de ces signaux, et $\mu_i(n)$ les processus d'innovation normalisés correspondants. Alors, la décorrélation de r_1 et r_2 équivaut à celle des bruits blancs

$$\mu_1 \text{ et } \mu_2. \text{ Posons par ailleurs } f(z) = \begin{bmatrix} f_1(z) & 0 \\ 0 & f_2(z) \end{bmatrix},$$

$$g(z) = \begin{bmatrix} g_1(z) & 0 \\ 0 & g_2(z) \end{bmatrix}, \mu(n) = [\mu_1(n), \mu_2(n)]^T \text{ et}$$

$v(n) = [v_1(n), v_2(n)]^T$. Alors,

$$\begin{cases} r(n) = [g(z)]\mu(n) \\ s(n) = [f(z)]v(n) \end{cases}$$

Explicitant les relations de mélange entre $r(n)$ et $s(n)$, nous obtenons $r(n) = [G(z)][H(z)]s(n)$.

Exprimant ensuite les signaux $r(n)$ et $s(n)$ en fonction de leurs processus d'innovation normalisés, nous en déduisons :

$$\mu(n) = \Theta(z)v(n) \quad (1)$$

avec

$$\Theta(z) = g(z)^{-1}G(z)H(z)f(z) \quad (2)$$

Or, μ et v sont tous les deux des bruits blancs de matrice de covariance instantanée égale à l'identité. Par conséquent, $\Theta(z)$ doit vérifier

$$\Theta(z)\Theta(z^{-1})^T = I$$

Par ailleurs, $\Theta(z)$ et son inverse sont les produits de quatre filtres causaux et d'inverses causaux. Or, d'après ce qui précède, l'inverse de $\Theta(z)$ vérifie $\Theta(z)^{-1} = \Theta(z^{-1})^T$, et ne peut donc être causal que si $\Theta(z)$ est réduit à une matrice constante 2×2 orthogonale.

Cette matrice peut être par exemple paramétrée de la façon suivante : $\Theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \cos \theta \begin{bmatrix} 1 & t \\ -t & 1 \end{bmatrix}$ où l'on a posé $t = \tan \theta$. Après simplification, cette paramétrisation, exploitée dans la relation (2), permet d'exprimer le filtre $G(z)$ comme suit :

$$\begin{cases} G_{12}(z) = \frac{f_2(z)H_{12}(z) - t f_1(z)}{f_2(z) - t f_1(z)H_{21}(z)} \\ G_{21}(z) = \frac{t f_2(z) + f_1(z)H_{21}(z)}{t H_{12}(z) + f_1(z)} \end{cases} \quad (3)$$

Nous avons donc établi que si $G(z)$, filtre causal et d'inverse causal, est solution des équations de décorrélation, alors il peut être paramétré sous la forme (3). Bien entendu, il reste, pour caractériser les solutions, à intégrer la contrainte selon laquelle $G(z)$ doit être causal et d'inverse causal. Pour que $G(z)$ soit causal, il faut et il suffit que les dénominateurs de G_{12} et G_{21} soient dépourvus de zéro à l'extérieur du cercle unité. Ceci est vrai pour $t = 0$ (on retrouve alors la solution séparante), mais aussi pour t assez petit. Par ailleurs, il est facile de vérifier que $\det(G(z)) = (1 + t^2)\det(H(z))$, de sorte que, quelle que soit la valeur de t , $\det(G(z)) \neq 0$ pour $|z| > 1$. En conclusion, le filtre $G(z)$ est causal et d'inverse causal si et seulement si il est causal. On en déduit qu'il existe toujours une infinité de solutions aux équations de décorrélation en général. Dans le cas où $H(z)$ est à RIF, on peut imposer à $G(z)$ de l'être également, et une simple inspection de (3) suggère que seul le paramètre $t = 0$ correspond à des filtres G_{12} et G_{21} RIF. On retrouve ainsi le résultat mis en évidence récemment par Lindgren et Broman [3]. Cependant, notre analyse montre clairement que le caractère séparant des solutions des équations de décorrélation est fondamentalement non robuste. Nous illustrerons les effets perceptibles de cet état de fait dans la section consacrée aux simulations.

3 L'approche par prédiction linéaire.

L'approche par prédiction linéaire a été introduite dans [4] dans le cas où le nombre de capteurs est strictement supérieur au nombre de sources. Nous en proposons ici une adaptation au cas plus simple 2 sources 2 capteurs.

Il nous faut tout d'abord rappeler quelques notions et définitions de base. Le processus d'innovation de $y(n)$ est le signal $i(n)$ défini par $i(n) = y(n) + \sum_{k=1}^{+\infty} A_k y(n-k) = [A(z)]y(n)$ où le filtre $A(z) = I + \sum_{k \geq 1} A_k z^{-k}$ est défini par :

$$A(z) = \arg \min_B E \left[\left\| y(n) + \sum_{k=1}^{\infty} B_k y(n-k) \right\|^2 \right]$$

$i(n)$ est un bruit blanc (i.e. $E[i(n+k)i(k)^T] = 0$ pour $n \neq 0$), dont la matrice de covariance instantanée $D = E[i(n)i(n)^T]$ n'est pas nécessairement égale à l'identité : les deux composantes $i_1(n)$ et $i_2(n)$ de $i(n)$ ne sont ni décorréliées, ni de variance unité. On appelle alors processus d'innovation normalisé de $y(n)$ tout bruit blanc $v(n)$ défini à partir de $i(n)$ par $v(n) = L^{-1}i(n)$, où L vérifie¹ $D = LL^T$. Par conséquent, $E[v(n)v(n)^T] = I$. Bien entendu, l'innovation normalisée est définie à une matrice orthogonale près : si $v(n)$ est une innovation normalisée, il est en de même pour $w(n) = Qv(n)$ où Q est une matrice orthogonale. Le principe de l'approche par prédiction linéaire comporte trois étapes :

Etape 1. Puisque le filtre $H(z)$ est par hypothèse à phase minimale, il en est de même pour le filtre $H(z)f(z)$, où on rappelle que $f(z)$ est le filtre diagonal dont les éléments sont les factorisations à phase minimale des densités spectrales des sources s_1 et s_2 . Dans ces conditions, le bruit blanc $v(n)$ de dimension 2, constitué à partir des innovations normalisées des sources, coïncide avec une innovation normalisée de l'observation, et peut donc être extrait, à une matrice orthogonale près, à partir de $y(n)$ par le biais d'un algorithme de prédiction linéaire : après évaluation du filtre de prédiction $A(z)$ grâce aux statistiques du second ordre de $y(n)$, l'innovation $i(n)$ est extraite puis sa matrice de covariance D , enfin $v(n) = L^{-1}i(n)$ où L est une racine carrée de D est généré.

Etape 2. En principe, $v(n) = \Theta v(n)$ où Θ est une matrice orthogonale inconnue. Pour extraire $v(n)$, il suffit de remarquer que $v(n)$ est un mélange instantané des deux composantes v_1 et v_2 de v . Puisque ces deux signaux sont indépendants, on peut les extraire par un algorithme classique de séparation de mélange instantané.

Etape 3. La dernière étape de l'algorithme consiste à reconstituer les contributions de chaque source sur chaque capteur, i.e. les signaux s_1 et s_2 . Pour ceci, posons $K(z) = \Theta^T L^{-1} A(z)$. Alors, il est clair que $H(z)f(z) = K^{-1}(z)$, et qu'en conséquence

$$K^{-1}(z) \begin{bmatrix} v_1(n) \\ 0 \end{bmatrix} = \begin{bmatrix} s_1(n) \\ * \end{bmatrix}$$

Par conséquent, on peut reconstituer s_1 à partir de $K(z)$. La reconstitution de s_2 s'effectue dans les mêmes conditions. Il

convient de remarquer que l'algorithme de séparation instantanée appliqué à $v(n)$ permet de générer $v(n)$, mais à une permutation et à des signes près. Il faut donc tenir compte de cet état de fait au niveau de la reconstitution de s_1 et s_2 , et éventuellement adapter le schéma que nous venons de donner.

Enfin, il est clair qu'en pratique le vrai filtre prédictif à horizon infini $A(z)$ ne peut être calculé. On se contente donc d'estimer, à partir d'un nombre fini d'observations, un filtre prédictif à horizon fini $A_N(z) = I + \sum_{k=1}^N A_{k,N} z^{-k}$, avec N suffisamment grand. Dans ces conditions, la procédure que nous venons de décrire s'adapte sans problème méthodologique sérieux, exception faite de la dernière étape. En fait, $K(z)$ est remplacé par un filtre RIF de longueur N , $K_N(z) = \Theta_N^T L_N^{-1} A_N(z)$, où les indices N signifient que Θ_N et L_N sont évaluées sur la base de l'innovation partielle $i_N(n) = [A_N(z)]y(n)$. Pour que cette dernière étape ait un sens, il faut absolument que le filtre $K_N(z)^{-1}$ soit stable, i.e. que $A_N(z)^{-1}$ le soit. Fort heureusement, les coefficients de A_N sont estimés, en pratique, en résolvant une équation de type Yule-Walker basée sur un estimateur \hat{R}_N de la matrice de covariance R_N du vecteur $Y(n) = (y(n)^T, \dots, y(n-N)^T)^T$. Dès que l'estimateur choisi est bloc Toeplitz et défini positif, la stabilité de $A_N(z)^{-1}$ est garantie. En pratique, il est difficile d'imposer simultanément ces deux conditions à l'estimateur \hat{R}_N , mais tout estimateur raisonnable génère un filtre stable pour peu que le nombre d'observations soit suffisant.

Finalement, l'approche par prédiction linéaire semble constituer une alternative satisfaisante aux techniques de décorrélation. Son éventuel désavantage tient au fait qu'elle nécessite implicitement d'approximer $H(z)$ et $f(z)$ par des filtres FIR. Si les signaux sources s_1 et s_2 ont des spectres possédant des zéros, cette dernière étape sera difficile à accomplir, et l'efficacité de la procédure en sera affectée, alors que l'inversion de $H(z)$ peut s'avérer intrinsèquement facile, en particulier en utilisant une technique de décorrélation si $H(z)$ est exactement FIR.

4 Résultats des simulations

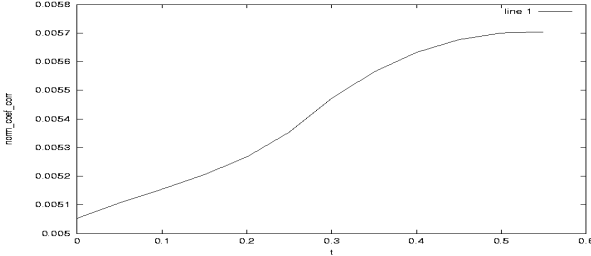
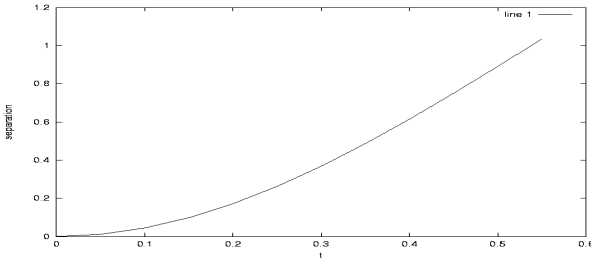
Dans les simulations qui suivent, nous prendrons le filtre $H(z)$ tel que $H_{12}(z) = \frac{1+0.5z^{-1}}{1-0.1z^{-1}}$, $H_{21}(z) = \frac{1+0.2z^{-1}}{1-0.1z^{-1}}$ dans le cas RII et $H_{12}(z) = 1 + 0.2\sqrt{2}z^{-1}$, $H_{21}(z) = -1 + 0.2\sqrt{2}z^{-1}$ dans le cas RIF.

Décorrélation En pratique, on paramètre les filtres $G_{12}(z)$ et $G_{21}(z)$ par des filtres RIF de degrés p fixé. Leurs coefficients sont déterminés de telle sorte que la somme des carrés des $4p+1$ premiers coefficients d'intercorrélations empiriques entre r_1 et r_2 soit minimale. Pour chaque valeur de G_{12} (resp. G_{21}) fixée, la fonction de coût correspondante est quadratique par rapport aux coefficients de G_{21} , de sorte que l'on peut converger vers l'un de ses minima locaux par un algorithme itératif de relaxation très simple. Bien entendu, l'initialisation de l'algorithme est ici cruciale, et rien ne garantit sa convergence vers le minimum global.

Nous commençons par illustrer l'existence d'un ensemble de solutions "parasites" exprimées en (3). La figure (fig.1)

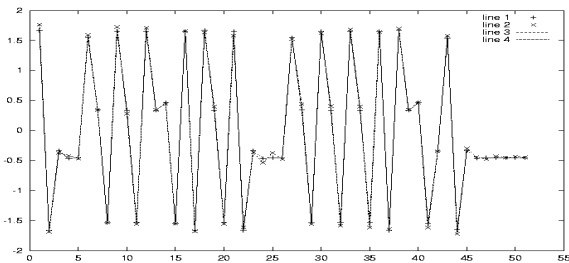
¹ dans le cas qui nous intéresse, la matrice D est inversible

montre la norme ρ du vecteur composé des $4p + 1$ premiers coefficients d'intercorrélation empiriques entre les sorties r_1 et r_2 pour différents filtres $G(z)$ obtenus en calculant la troncature polynômiale de (3) pour diverses valeurs de t . Cette courbe est à rapprocher de celle (fig.2) de l'évolution en fonction de t des taux de restitution définie par $sep = sep_1 + sep_2$ et $sep_i = \frac{\sum_{n=1}^T (s_i - r_i)^2}{\sum_{n=1}^T s_i^2}$ où T représente le nombre d'échantillons.

FIG. 1 — $\rho(t)$ FIG. 2 — $sep(t)$

On remarque, d'après les courbes (fig.1 et 2), que la valeur de ρ est quasiment constante tandis que le taux de restitution augmente notablement. Ceci illustre l'existence d'une infinité de solutions non séparantes aux équations de décorrélation.

Prediction linéaire La courbe ci-dessous (fig.3) montre le signal s_1 (*) et son estimée $r_1(+)$ dans le cas RII. Il

FIG. 3 — $s_1(n)$ et $\hat{s}_1(n)$

apparaît clairement que le taux de reconstruction est tout à fait satisfaisante (cf tableau suivant).

Comparaison des deux méthodes Dans le tableau suivant, nous présentons des résultats comparatifs entre les deux approches proposées précédemment. Le cas où $H(z)$ est RIF est d'abord présenté, le cas où $H(z)$ est RII ensuite. Pour

chaque cas, deux initialisations différentes ont été testées dans la méthode de décorrélation.

	Décorrélation		Prédiction	
	Init. 1	Init. 2	Linéaire	
sep_1	0.0016169	0.92525	0.021393	cas
sep_2	0.0061503	2.9199	0.073438	RIF
sep_1	0.15435	0.58399	0.0015590	cas
sep_2	0.15074	0.55316	0.0013727	RII

Même dans le cas de mélanges RIF, l'approche par décorrélation pose des problèmes de robustesse : on remarque en effet que les résultats sont très sensibles à l'initialisation du filtre $G(z)$, ce qui semble prouver l'existence d'un certain nombre de minima locaux. Ceci est cohérent avec la situation du cas RII dans laquelle les équations de décorrélation ont une infinité de solution. En pratique, dans le cas RII, l'algorithme de décorrélation n'est jamais en mesure de fournir un filtre ayant de bonnes propriétés de séparation, même si on l'initialise proche de la solution optimale.

5 Conclusion

Dans cet article, nous montrons que la résolution des équations de décorrélation pour la séparation de mélanges convolutifs dans le cas de filtres à réponse impulsionnelle infinie n'est pas suffisante pour séparer. Nous proposons ensuite une approche par prédiction linéaire pour résoudre ce problème.

Références

- [1] U. Lindgren & H. Broman. Source separation using a criterion based on second order statistics, soumis à *IEEE Transaction on Signal Processing*.
- [2] N. Delfosse & Ph. Loubaton. Adaptive blind identification of convolutive mixtures, *Proc. Int. Conf. Asilomar 95*, nov.95, pp. 341-345.
- [3] H.L. Nguyen Thi & Ch. Jutten. Blind sources separation for convolutive mixtures, *Signal Processing*, vol. 45, pp. 209-229, 1995.