

3-1 Introduction :

L'objet de ce chapitre est de préciser les modèles de mélanges linéaire instantane de sources sur lesquels nous allons travailler et nous avons donnés quelques algorithmes de séparation aveugle.

il y a deux grande famille de mélange ils est le mélange linéaire et le mélange non-linéaire , nous avons étudié le 1^{er} modèle qui continue l'autre à trois modèle mélange instantané, mélange spectrale et le mélange convolutif, Dans ce chapitre on a basé sur le mélange instantané.

3-2 Définition de mélange instantané :

Le mélange linéaire instantané est le mélange le plus simple dans lequel l'hypothèse est faite que les signaux sources arrivent en même temps sur tous les capteurs mais avec des intensités différentes, et ce quelles que soient les positions des sources par rapport aux capteurs. Les J observations du mélange s'expriment alors en fonction des I signaux sources de la façon suivante :

$$x_j(t) = \sum_{i=1}^I a_{ji}s_i(t), j = 1, \dots, J \dots\dots\dots (3.1)$$

ou encore pour l'ensemble des observations et des sources, selon la formulation matricielle suivante :

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_J(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1I} \\ a_{21} & a_{22} & \cdots & a_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ a_{J1} & a_{J2} & \cdots & a_{JI} \end{pmatrix} \cdot \begin{pmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_I(t) \end{pmatrix} \dots\dots\dots (3.2)$$

noté de façon plus compacte $x = As$. La matrice A est définie par $J \times I$ coefficients a_{ji} constants. La séparation de sources à partir de mélanges instantanés passe d'abord par l'identification, pour chaque source s_i , des facteurs d'atténuation a_{ji} .

3-2-1 Présentation du modèle :

Le mélange instantané de sources réelles sans bruit additif. C'est en effet la situation minimale commune à l'ensemble des méthodes évoquées ici. L'extension au cas de signaux complexes ne devrait pas poser de grandes difficultés. La présentation de la méthode *SOBI* peut servir d'exemple pour cette situation. Par contre, la présence du bruit additif constitue un problème délicat. Il est traité par les techniques de sous-espaces au

cours de l'étape de blanchiment des observations dans les méthodes basées sur la diagonalisation. Nous reviendrons sur ce point en fin de paragraphe. Nous allons donc préciser les hypothèses pour le modèle à temps discret :

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), n = 1, 2, \dots \quad (3.3)$$

Nous rappelons ensuite le problème classique des indéterminations concernant la matrice de mélange $\mathbf{A}$ et surtout celui de son identifiabilité par les méthodes à l'ordre 2.

3-2-2 Les hypothèses du modèle :

Dans ce paragraphe nous avons donné les hypothèses nécessaires de mélange instantané, Les signaux sources sont indépendants et identiquement distribués et mutuellement indépendants :

- Le nombre de capteurs est supérieur ou égal au nombre d'émetteurs.
- La matrice de canal $\mathbf{H}$ est de rang complet M .
- Le bruit $b(n)$ est additif indépendant des signaux sources.
- Au plus un seul signal source a une distribution gaussienne.
- Les signaux sources sont indépendants et identiquement distribués et mutuellement indépendants.

3-2-3 La séparation :

Le principe général de la séparation aveugle de sources consiste à trouver une transformation $\mathbf{W}$ qui permette d'obtenir, à partir des signaux reçus, des signaux mutuellement indépendants, ou en d'autres termes de calculer l'inverse de la matrice de mélange $\mathbf{H}$. Alors, on a :

$$\mathbf{z}(n) = \hat{\mathbf{s}}(n) = \mathbf{W}^T \mathbf{y}(n) \quad (3.4)$$

où $\mathbf{W}$ est la matrice de séparation de dimension $N \times M$ permettant de récupérer les signaux sources à un facteur d'échelle et une permutation près. Alors, dans la suite une matrice $\mathbf{W}$ est dite matrice de séparation, si et seulement si elle satisfait, en l'absence de bruit, l'égalité suivante :

$$\mathbf{W}^T \mathbf{y}(n) = \mathbf{P}_s(n) \quad (3.5)$$

ou encore :

$$\mathbf{W}^T \mathbf{H} = \mathbf{P} \mathbf{\Lambda} \quad (3.6)$$

où $\mathbf{P}$ est une matrice de permutation et $\mathbf{\Lambda}$ est une matrice diagonale non singulière.

Soit :

$$\mathbf{G}^T = \mathbf{W}^T \mathbf{H} \quad (3.7)$$

la matrice globale du système composé du canal et du séparateur. La séparation parfaite des signaux sources, c'est à dire une séparation sans permutation et sans facteur d'échelle, correspond à :

$$G^T = W^T H = P \Lambda = I \dots\dots\dots (3.8)$$

3-2-4 Indéterminations :

Dans un contexte aveugle où les signaux sources et la matrice de canal sont tous des inconnus, on remarque facilement que le même vecteur $y(n)$ peut être généré à partir d'une infinité de vecteurs $s(n)$. En effet, l'ordre des signaux est arbitraire car toute permutation appliquée sur les signaux sources et sur les lignes de la matrice H correspondantes donne le même vecteur $y(n)$. Donc les signaux sources ne peuvent être récupérés qu'à une permutation près. En plus, l'échange d'un facteur d'échelle entre un signal source et le vecteur colonne de la matrice de mélange H correspondant n'affecte en rien les observations. Ceci est montré à travers l'équation suivante :

$$y(n) = Hs(n) \dots\dots\dots (3.9)$$

Donc, au mieux, les signaux sources sont restitués à une permutation et un facteur d'échelle près. Ces indéterminations sont dues au fait que les signaux sources et la matrice de canal sont tous à la fois des inconnus et ces indéterminations n'influencent pas l'indépendance des signaux que l'on cherche à travers la minimisation de certains critères.

3-3 Algorithme JADE:

L'algorithme JADE utilise un critère à base des cumulants d'ordre quatre. Son principe est résumé de la sorte :

- 1- Constituer la matrice de covariance R_x et calculer une matrice de blanchiment W .
- 2- Constituer les cumulants d'ordre quatre $Cum4(z)$ du processus blanchi

$$z = Wx = U_s + Wb \dots\dots\dots (3.10)$$

(U, W et b sont respectivement matrice unitaire, matrice de blanchiment et bruit blanc), et calculer l'ensemble de ses p plus significatives (ordre croissant en valeur absolue) valeurs et matrices propres $\{\Lambda_r, M_r\}$; $1 \leq r \leq p$ (est le nombre des sources).

3-Diagonaliser conjointement l'ensemble des matrices Λ_r, M_r par une matrice unitaire $\hat{U}$; ou la diagonalisation a été définie par maximisation du critère suivant :

$$C(U, N) = \sum_{r=1}^p |\text{diag}(U^* N_r U)|^2 \dots\dots\dots (3.11)$$

ou $|\text{diag}(\cdot)|$ est la norme du vecteur composé des éléments diagonaux de la matrice argument et $N = \{N_r \mid 1 \leq r \leq p\}$ est l'ensemble à diagonaliser.

- 4- Estimer la matrice de séparation :

$$A = W\#U \dots\dots\dots(3.12)$$

ou le symbole # dénote la matrice pseudo-inverse de Moore-Penrose. L'auteur utilise la technique de Jacobi étendue pour maximiser son critère de diagonalisation conjointe [29].

3-4 Algorithme SOBI:

L'algorithme SOBI est basé sur des statistiques d'ordre deux. Sous l'hypothèse de signaux sources colorés, de spectres différents. Le principe de l'algorithme SOBI est résumé de la sorte :

1. Calculer la matrice de covariance $R_x(\tau)$ des observations.
2. Estimer la matrice de blanchiment W .
3. Blanchir les observations :

$$z(n) = Wx(n) \dots\dots\dots(3.13)$$

4. Estimer la matrice V par la diagonalisation conjointe d'un ensemble de matrices issues de la fonction de corrélation des observations blanchies à différents retard $\tau \neq 0$.
5. Estimer la matrice de mélange par $A = W\#V$ ou le symbole # dénote la matrice pseudo-inverse [30].
6. Estimer les signaux sources par :

$$s(n) = A^H R_z(0)^{-1} x(n) \dots\dots\dots(3.14)$$

3-5 Définition d'Algorithme FastICA :

L'algorithme FastICA est un algorithme très performant de maximisation de la fonction de contraste pour les sources non gaussiennes [31]. Il est basé sur le principe de l'algorithme d'apprentissage itératif de type point fixe.

L'algorithme d'apprentissage du réseau cherche à trouver une direction qui est en fait un vecteur w tel que la projection $w^T x$ maximise l'aspect non gaussien. Pour cela, l'algorithme utilise le théorème du point fixe. Si on considère que chaque sortie du réseau est donnée sous forme d'une fonction de type $g(w^T x)$ où g est une fonction scalaire non linéaire, la forme simplifiée de l'algorithme pour l'estimation d'une composante indépendante peut être reformulée de la manière suivante :

- initialiser le vecteur w par de faibles valeurs aléatoires .
- mettre $w^+ = E(x.g(w^T x)) - w.E(g(w^T x))$.
- mettre $w = w^+ / \|w^+\|$.
- si la convergence n'est pas encore atteinte, refaire les étapes 2 et 3.

La convergence signifie que les valeurs de w (anciennes et nouvelles) vont dans la même direction, c'est-à-dire que la différence entre ces deux valeurs est au-dessous du critère de convergence entre deux itérations.

Pour estimer plusieurs composantes, on a besoin d'utiliser l'algorithme précédent avec plusieurs neurones ayant les vecteurs synaptiques $w_1, w_2, \dots, w_N$. Pour que les vecteurs ne convergent pas vers un même maximum, nous devons décorréliser les projections $w^T x$ et $w_i^T x$.

$w_1^T x, w_2^T x, \dots, w_N^T x$ après chaque itération. Il existe plusieurs méthodes de décorrélation.

La performance de l'algorithme FastICA dépend du choix de la fonction g . Les composantes indépendantes peuvent être extraites une à une. Ainsi, on obtient un gain considérable en performance si l'on a seulement besoin de quelques composantes indépendantes à extraire.

3-5-1 Propriétés de l'algorithme FastICA :

Le FastICA L'algorithme des fonctions de contraste sous-jacents ont un certain nombre de propriété souhaitable comparé avec les méthodes existantes pour l'ICA.

1. La convergence est cubique (ou au moins quadratique), sous l'hypothèse du modèle de données de l'ACI (pour preuve [32]. Ceci est en contraste avec les algorithmes basés sur l'ICA ordinaires (stochastique) descente de gradient méthodes, où la convergence est seulement linéaire. Cela signifie une convergence très rapide, comme cela a été confirmé par des simulations et des expériences sur des données réelles [32].
2. Contrairement aux algorithmes basés sur gradient, il y a la taille des pas sans paramètres à choisir. Cela signifie que le algorithme est facile à utiliser.
3. L'algorithme trouve des composants directement indépendants (pratiquement) toute distribution non gaussienne utilisant tout nonlinéarité g . Ceci est en contraste avec de nombreux algorithmes, où une estimation de la distribution de probabilité fonction doit être le premier disponible, et la non-linéarité doit être choisie en conséquence.
4. La performance du procédé peut être optimisé en choisissant une non-linéarité g approprié. En particulier, une peut obtenir des algorithmes qui sont robustes et / ou de la variance minimale. En effet, les deux non-linéarités avoir des propriétés optimales; pour plus de détails [31]
5. Les composants indépendants peuvent être estimés, un par un, ce qui est à peu près équivalent à faire projection poursuite. Ce es utile dans l'analyse des données d'exploration,

et réduit la charge de calcul du procédé les cas où seule une partie des composants indépendants doivent être estimés.

6. Le FastICA a la plupart des avantages des algorithmes neuronaux: Il est parallèle, distribué, informatiquement simple, et nécessite peu d'espace mémoire. méthodes de gradient stochastiques semblent préférables que si rapide adaptativité dans un environnement changeant est nécessaire.

3-6 Conclusion :

Dans ce chapitre nous avons donné une définition sur le mélange instantané et nous avons vu les algorithmes de séparation de sources trouvés dans la littérature (cas de mélanges linéaires instantanés). Fondamentalement, la littérature est riche d'algorithmes de séparation de sources de mélanges instantanés où les critères de dépendance sont variés et les résultats de séparation sont de très bonne qualité.