



Séparation aveugle de sources

Etudes et implémentation sous

- Matlab ou Python,
- langage C et DSP

F4B101A TP

Année scolaire 2016-2017

1 Introduction

Imaginez que vous vous trouviez dans une pièce où deux personnes parlent simultanément. Vous disposez de deux microphones se trouvant à deux emplacements différents de cette pièce. Ces deux microphones fournissent des signaux $x_1(t)$ et $x_2(t)$. Dans toute la suite, nous supposons que chacun de ces signaux correspond à une somme pondérée des deux signaux sonores $s_1(t)$ et $s_2(t)$ émis par les locuteurs. Ainsi, on a

$$x_1(t) = a_{11} s_1(t) + a_{12} s_2(t) \quad (1)$$

$$x_2(t) = a_{21} s_1(t) + a_{22} s_2(t) \quad (2)$$

Les paramètres a_{11} , a_{12} , a_{21} et a_{22} ne dépendent que des positions relatives entre microphones et locuteurs. Le problème auquel nous allons nous intéresser, connu sous le nom de cocktail party problem, consiste en la reconstitution des signaux s_1 et s_2 à partir de la seule connaissance de x_1 et x_2 et de quelques propriétés statistiques de s_1 et s_2 .

L'analyse en composantes indépendantes (ICA en anglais pour independent component analysis) apporte une solution à ce problème dans le cas où les signaux $s_i(t)$ sont indépendants et non gaussiens. Cette technique a été développée initialement dans un contexte applicatif proche du cocktail party problem mais assez rapidement, elle a trouvé d'autres domaines d'application. Ainsi, en électroencéphalographie, des enregistrements de l'activité du cerveau sont effectués en différents points de la boîte crânienne. Chacun de ces signaux contient un mélange de différents signaux émis par des zones bien précises du cerveau. Il s'agit de les isoler les uns par rapport aux autres. On peut aussi citer des applications en reconnaissance de forme, en compression d'image, etc.

Question 1 : *Rappeler et distinguer les notions d'indépendance et de non corrélation pour deux signaux aléatoires.*

2 Analyse en composantes indépendantes

2.1 Notations

Considérons n signaux x_1, x_2, \dots, x_n correspondant à des mélanges de n signaux indépendants (ou composantes indépendantes)

$$x_j = a_{j1} s_1 + a_{j2} s_2 + \dots + a_{jn} s_n, \quad 1 \leq j \leq n \quad (3)$$

La mention de la dépendance au temps t des signaux a été omise afin d'alléger les notations. Nous supposons que les signaux x_j sont à moyenne nulle. Lorsque ce n'est pas le cas, un changement de variable simple permettra de se ramener à cette situation.

En utilisant une notation vectorielle, il est possible de réécrire (3) sous la forme

$$\mathbf{x} = \sum_{i=1}^n \mathbf{a}_i s_i \quad (4)$$

où encore

$$\mathbf{x} = \mathbf{A} \mathbf{s} \quad (5)$$

avec $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, $\mathbf{s} = [s_1, s_2, \dots, s_n]^T$, $\mathbf{a}_i = [a_{1i}, a_{2i}, \dots, a_{ni}]^T$ et $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$. \mathbf{A} est appelée matrice de mélange. Le problème de l'ICA est d'estimer \mathbf{A} et \mathbf{s} à partir de la connaissance de \mathbf{x} . Essentiellement, le principe de résolution repose sur l'estimation de la matrice \mathbf{A} puis d'une matrice \mathbf{W} telle que $\mathbf{W}\mathbf{A} = \mathbf{I}_n$, où \mathbf{I}_n est la matrice identité de rang n .

Question 2 : *Montrer qu'une fois la matrice \mathbf{W} estimée, on a :*

$$\mathbf{s} = \mathbf{W} \mathbf{x}$$

Par conséquent, les techniques de séparation aveugle de sources proposent une solution au problème de séparation. Cependant cette séparation se fera avec des ambiguïté inhérente d'amplitude et de permutation. Afin de lever l'ambiguïté sur l'amplitude, nous supposons que les composantes indépendantes s_i à estimer sont à variance unité.

Ici, nous avons supposé que le nombre de signaux observés n est égal au nombre de composantes indépendantes. Le problème peut se généraliser à un nombre m de signaux observés formés de n composantes indépendantes, avec $m \geq n$.

Question 3 : *Que se passe-t-il lorsque $m < n$?*

2.2 Première illustration

Considérons deux signaux aléatoires indépendants ayant la même densité de probabilité

$$f_{s_i}(x) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{si } |x| \leq \sqrt{3} \\ 0 & \text{sinon} \end{cases}, \quad i = 1, 2 \quad (6)$$

Question 4 : Vérifier que s_1 et s_2 sont de variance unité.

Question 5 : Quelle est la densité de probabilité jointe du couple (s_1, s_2) ? Générer 1000 réalisations de s_1 et s_2 . Représenter graphiquement ces 1000 réalisations dans le plan (s_1, s_2) . Comment se traduit l'indépendance de s_1 et s_2 ?

Deux nouveaux signaux aléatoires x_1 et x_2 sont obtenus par mélange de s_1 et s_2 en utilisant la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \quad (7)$$

Question 6 : Générer 1000 réalisations de x_1 et x_2 à partir de celles de s_1 et s_2 . Représenter graphiquement les 1000 points obtenus dans le plan (x_1, x_2) . Les signaux aléatoires x_1 et x_2 sont-ils encore indépendants ?

Question 7 : Retrouver les coefficients de \mathbf{A} à partir de la représentation graphique de la question 6. Calculez \mathbf{W} et retrouver les réalisations de s_1 et s_2 à partir de celles de x_1 et x_2 . Quelle est la limitation de cette technique d'estimation des coefficients de \mathbf{A} ?

Question 8 : Reprendre les questions précédentes en considérant une densité de probabilité gaussienne de variance unité pour s_1 et s_2 . Est-il encore possible de retrouver s_1 et s_2 à partir de x_1 et de x_2 ?

2.3 Principe

L'analyse en composantes indépendantes repose d'une part sur le fait que les composantes indépendantes ont une densité de probabilité non gaussienne et

d'autre part sur le théorème *central limite*. Ce théorème indique que la densité de probabilité d'une somme de variables aléatoires indépendantes tend (sous certaines conditions) vers une distribution gaussienne. Ainsi, la somme de deux variables aléatoires indépendantes a une densité de probabilité plus proche d'une gaussienne que l'une quelconque de ces deux variables.

Question 9 : *On considère quatre variables aléatoires indépendantes dont la distribution de probabilité est donnée par l'équation (6). Calculer les distributions de probabilité des variables aléatoires $s_1 + s_2$, $s_1 + s_2 + s_3$ et $s_1 + s_2 + s_3 + s_4$. Comparez chacune de ces distributions à une distribution aléatoire gaussienne de même variance.*

Considérons un vecteur \mathbf{x} de n signaux satisfaisant le modèle (5), c'est-à-dire formé d'un mélange de n composantes indépendantes. Nous supposons dans cette partie que chacune des composantes indépendantes a une densité de probabilité identique. Afin d'estimer l'une des composantes s_i , nous allons considérer une combinaison linéaire des x_i notée y

$$y = \mathbf{w}^T \mathbf{x} = \sum_i w_i x_i \quad (8)$$

où \mathbf{w} est un vecteur à déterminer. Si l'on parvient à régler \mathbf{w} de manière à ce qu'il corresponde à une ligne de la matrice inverse de \mathbf{A} , alors y sera égal à l'un des s_i . Il s'agit donc trouver une technique pour déterminer \mathbf{w} .

Posons $\mathbf{z} = \mathbf{A}^T \mathbf{w}$. D'après l'équation (5), on a $y = \mathbf{w}^T \mathbf{x} = \mathbf{w}^T \mathbf{A} \mathbf{s} = \mathbf{z}^T \mathbf{s}$. Par conséquent, y est une combinaison linéaire de s_i , avec des poids donnés par les z_i . Sachant qu'une somme de variables aléatoires indépendantes est plus gaussienne que chacune de ses composantes, la variable aléatoire y sera la moins gaussienne lorsqu'un seul z_i sera non nul, c'est-à-dire lorsque y ne sera formé que d'une seule composante s_i .

Le réglage de \mathbf{w} se fait donc de manière à rendre la densité de probabilité de y la moins gaussienne possible. Pour faire ce réglage, il faut encore trouver une mesure $\mathcal{J}(y)$ du caractère non gaussien de la densité de probabilité de y . On peut envisager diverses fonctions mesurant le caractère non gaussien, dont le kurtosis et la néguentropie.

2.3.1 Le kurtosis

C'est une mesure classique du caractère non-gaussien d'une distribution de probabilité. Il correspond au cumulants d'ordre 4, donné par :

$$\text{kurt}(y) = \mathbb{E}[y^4] - 3 \left(\mathbb{E}[y^2] \right)^2 \quad (9)$$

Question 10 : *Montrez que le kurtosis d'une variable aléatoire gaussienne centrée est nul.*

Pour la plupart des variables aléatoires, le kurtosis est non nul. La valeur absolue ou le carré du kurtosis peuvent donc constituer de bonnes mesures du caractère non gaussien d'une distribution de probabilité. Cependant, cette fonction est assez sensible à la présence de bruit venant entacher les mesures des signaux x_i .

2.3.2 La néguentropie

La néguentropie est basée sur la notion d'entropie, qui pour une variable aléatoire y de densité de probabilité f est donnée par

$$H(y) = - \int f(y) \log(f(y)) dy \quad (10)$$

L'entropie est un concept de base de la théorie de l'information. Très schématiquement, l'entropie d'une variable aléatoire peut être interprétée comme la quantité d'information que fournit l'observation de cette variable. Plus cette variable est aléatoire, c'est-à-dire imprédictible, plus son entropie sera importante. Un résultat fondamental de la théorie de l'information indique qu'une variable présentant une distribution de probabilité gaussienne possède la plus grande entropie lorsqu'on la compare à d'autres distributions de variance égale. De ce fait, l'entropie constitue une bonne mesure du caractère gaussien d'une distribution de probabilité. La néguentropie est ensuite construite à partir de l'entropie afin d'obtenir une fonction qui soit nulle pour une variable gaussienne et toujours positive dans les autres cas :

$$\mathcal{J}(y) = H(y_{\text{gauss}}) - H(y) \quad (11)$$

où y_{gauss} est une variable aléatoire gaussienne de même variance que y . La néguentropie, par sa justification théorique, est la mesure optimale du caractère non gaussien d'une distribution. En outre, cette fonction est beaucoup

moins sensible au bruit que le kurtosis, elle est cependant beaucoup plus délicate à calculer numériquement. Pour cette raison, des fonctions d'approximation de la néguentropie sont mises en œuvre.

2.3.3 Approximations de la néguentropie

La fonction d'approximation qui sera utilisée par la suite est

$$\mathcal{J}(y) = (\mathbb{E}[G(y)] - \mathbb{E}[G(y_{\text{gauss}})])^2 \quad (12)$$

où $G(\cdot)$ est une fonction non-linéaire qui peut être choisie comme

$$G(y) = \log \cosh(y) \quad \text{ou} \quad G(y) = -\exp(-y^2/2)$$

2.4 Traitements préalables

Dans les paragraphes précédents, les principes de résolution du problème de l'analyse en composantes indépendantes ont été présentés. Pour construire un algorithme mettant en œuvre ces techniques, quelques précautions préalables doivent encore être prises sur les signaux à analyser.

Un des traitements préalables en séparation de sources est de réaliser un blanchiment des signaux mesurés. Cela signifie, que par une opération linéaire, le vecteur \mathbf{x} est transformé en un vecteur $\tilde{\mathbf{x}}$ dont les composantes ne sont pas corrélées et dont la variance est égale à un (le vecteur $\tilde{\mathbf{x}}$ est blanc). Ceci se traduit par

$$\mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] = \mathbf{I}_n$$

Cette opération de blanchiment est toujours possible, par exemple par la décomposition en valeurs propres de la matrice de covariance :

$$\mathbb{E}[\mathbf{x}\mathbf{x}^T] = \mathbf{E} \mathbf{D} \mathbf{E}^T$$

où $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$ est la matrice diagonale contenant les valeurs propres de $\mathbb{E}[\mathbf{x}\mathbf{x}^T]$ et où \mathbf{E} est la matrice orthogonale de ses vecteurs propres. Le calcul de \mathbf{E} et de \mathbf{D} se fait aisément avec la commande "eig" de Matlab. L'opération de blanchiment est alors réalisée par le changement de variables

$$\tilde{\mathbf{x}} = \mathbf{E} \mathbf{D}^{-\frac{1}{2}} \mathbf{E}^T \mathbf{x}$$

Question 11 : Montrez qu'on a bien $\mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] = \mathbf{I}_n$.

Cette opération de blanchiment transforme la matrice de mélange en une nouvelle matrice calculée de la manière suivante

$$\tilde{\mathbf{x}} = \mathbf{E} \mathbf{D}^{-\frac{1}{2}} \mathbf{E}^T \mathbf{A} \mathbf{s} = \mathbf{U} \mathbf{s}$$

L'intérêt principal du blanchiment réside dans le fait que la nouvelle matrice de mélange est orthogonale, Le fait que \mathbf{U} soit orthogonale se répercute sur \mathbf{W} . Une fois les données blanchies, il suffit de trouver une matrice \mathbf{W} dont les vecteurs lignes \mathbf{w}_j sont orthogonaux deux à deux, ce qui simplifie beaucoup la recherche. Dans la suite de cet énoncé, nous supposons que les prétraitements ont été effectués et omettrons la notation tilde.

2.5 Recherche d'une seule composante indépendante

Ce paragraphe est consacré à la détermination d'un seul vecteur \mathbf{w} , c'est-à-dire d'une seule composante indépendante s_i . Nous avons vu au paragraphe 2.3 que le réglage de \mathbf{w} se fait par la recherche de l'optimum d'une fonction $\mathcal{J}(y) = \mathcal{J}(\mathbf{w}^T \mathbf{x})$ mesurant le caractère non gaussien d'une distribution de probabilité. Si on considère l'approximation de la néguentropie donnée par l'équation (12), il faut donc rechercher le maximum de $\mathbb{E}[G(\mathbf{w}^T \mathbf{x})]$, en respectant la contrainte $\mathbb{E}[(\mathbf{w}^T \mathbf{x})^2] = \|\mathbf{w}\|_2^2 = 1$ (on a supposé que les composantes indépendantes avaient une variance unité). En utilisant les conditions de Kuhn-Tucker, ce maximum sous contrainte est obtenu pour des \mathbf{w} satisfaisant

$$\mathbb{E}[\mathbf{x} g(\mathbf{w}^T \mathbf{x})] - \beta \mathbf{w} = 0$$

où $g(\cdot)$ est la dérivée de la fonction $G(\cdot)$. Par la méthode de Newton, il est possible de construire une technique itérative de recherche de \mathbf{w} :

1. Choisir un vecteur \mathbf{w} aléatoirement
2. $\mathbf{w}^+ = \mathbb{E}[\mathbf{x} g(\mathbf{w}^T \mathbf{x})] - \mathbb{E}[g'(\mathbf{w}^T \mathbf{x})] \mathbf{w}$
3. $\mathbf{w} = \frac{\mathbf{w}^+}{\|\mathbf{w}^+\|_2}$
4. tant qu'il n'y a pas convergence, aller à 2

où $g'(\cdot)$ est la dérivée de la fonction $g(\cdot)$.

Question 12 : Calculer les fonctions $g(\cdot)$ et $g'(\cdot)$ correspondant à $G(y) = \log \cosh(y)$.

Question 13 : *Mettre en œuvre l'algorithme en essayant d'extraire une des composantes indépendantes des deux signaux se trouvant dans le fichier `SignauxMelange.mat`. Suivant le vecteur \mathbf{w} initial, obtient-on toujours le même résultat ? Le calcul de l'espérance sera approximé par le calcul de la moyenne empirique.*

2.6 Recherche de toutes les composantes indépendantes

L'algorithme précédent fournit des résultats dépendant de la manière dont le vecteur \mathbf{w} a été initialisé. Ceci vient du fait que la fonction $\mathcal{J}(\mathbf{w}^T \mathbf{x})$ présente plusieurs maxima, chacun correspondant à un vecteur permettant de retrouver une composante indépendante différente. Pour retrouver toutes les composantes indépendantes, nous allons nous servir de la propriété d'orthogonalité de la matrice \mathbf{W} , vue au paragraphe 2.4. En effet, lorsque les mesures sont blanchies, la matrice \mathbf{W} à estimer possède des vecteurs lignes orthogonaux deux à deux. Un vecteur \mathbf{w} estimé grâce à l'algorithme du paragraphe précédent est donc une composante de \mathbf{W} , orthogonale à ses autres composantes.

Afin de construire un algorithme de recherche de toutes les composantes $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ de \mathbf{W} , une procédure de construction d'une base orthogonale, basée sur l'orthogonalisation de Gram-Schmidt peut être employée. Cette procédure est la suivante. Admettons que p vecteurs \mathbf{w}_i orthogonaux deux à deux ont déjà été estimés. Grâce à une itération de la procédure pour une seule composante, un nouveau vecteur \mathbf{w}_{p+1} est obtenu, ce vecteur doit être orthogonal aux p vecteurs précédents. Pour cela, on va retirer à \mathbf{w}_{p+1} toutes ses projections sur les autres vecteurs, c'est-à-dire $(\mathbf{w}_{p+1}^T \mathbf{w}_i) \mathbf{w}_i$, $i = 1, \dots, p$. Le vecteur résultant doit encore être normalisé. En résumé, on obtient

1. $\mathbf{w}_{p+1} = \mathbf{w}_{p+1} - \sum_{i=1}^p (\mathbf{w}_{p+1}^T \mathbf{w}_i) \mathbf{w}_i$
2. $\mathbf{w}_{p+1} = \frac{\mathbf{w}_{p+1}}{\|\mathbf{w}_{p+1}\|_2}$

Question 14 : *En utilisant le code `ICA.m` mis à disposition, mettez en œuvre cet algorithme sur les signaux simulés se trouvant dans le fichier `SignauxMelange.mat`. Les signaux originaux sont dans le fichier `SignauxReference.mat`.*

3 Algorithme de séparation de sources basé sur les statistiques du second ordre

Dans le cas où les signaux sources possèdent une cohérence temporelle, il a été montré qu'il est possible de séparer les différents signaux en utilisant les matrices d'intercovariance des signaux capteurs. Ces matrices possèdent une structure simple qui permet de résoudre le problème de l'identification aveugle par la procédure de décomposition propre. Les signaux sources sont supposés stationnaires au second ordre, de moyennes nulles, de puissances finies et mutuellement décorrélés.

Le bruit additif, dans le cas général, $\mathbf{b}(t)$ est modélisé par un processus aléatoire réel stationnaire, temporellement blanc, de moyenne nulle et de matrice de covariance \mathbf{R}_b :

$$\mathbb{E}[\mathbf{b}(t+\tau)\mathbf{b}(t)^T] = \delta(\tau)\mathbf{R}_b = \delta(\tau)\sigma^2\mathbf{I} . \quad (13)$$

où σ^2 représente la puissance du bruit. Ce bruit additif est supposé décorrélé des signaux sources.

Sous ces hypothèses, les matrices d'intercovariance ont la structure simple suivante :

$$\mathbf{R}_x(0) = \mathbb{E}[\mathbf{x}(t)\mathbf{x}(t)^T] = \mathbf{A}\mathbf{R}_s(0)\mathbf{A}^T + \mathbf{R}_b \quad (14)$$

$$\mathbf{R}_x(\tau) = \mathbb{E}[\mathbf{x}(t+\tau)\mathbf{x}(t)^T] = \mathbf{A}\mathbf{R}_s(\tau)\mathbf{A}^T \quad \tau \neq 0 , \quad (15)$$

où $\mathbf{R}_s(\tau) \triangleq \mathbb{E}[\mathbf{s}(t+\tau)\mathbf{s}(t)^T]$ est la matrice d'intercovariance des signaux sources, supposée diagonale. Dans ce qui suit, on se propose de résoudre le problème de l'identification aveugle en utilisant uniquement les matrices d'intercovariance $\mathbf{R}_x(\tau)$ des observations.

Soit \mathbf{U} la matrice de mélange après blanchiment. Soient maintenant les matrices d'intercovariance blanchies $\underline{\mathbf{R}}_x(\tau)$ définies par :

$$\underline{\mathbf{R}}_x(\tau) = \mathbf{U}\mathbf{R}_x(\tau)\mathbf{U}^T \quad \forall \tau \neq 0 , \quad (16)$$

Comme \mathbf{U} est unitaire et $\mathbf{R}_s(\tau)$ diagonale, la relation (16) signifie que les matrices d'intercovariance blanchies se diagonalisent sous la même transformation unitaire \mathbf{U} . Un principe simple d'identification aveugle consiste alors en une diagonalisation conjointe des matrices d'intercovariance blanchies.

Rappelons que si les valeurs propres d'une matrice sont uniques, il n'en est pas de même pour les vecteurs propres. Pour les valeurs propres distinctes, les vecteurs propres normés sont déterminés à une phase et une permutation près. Comme les vecteurs propres des matrices d'intercovariance blanchies sont les colonnes de la matrice \mathbf{U} , on trouve les même indéterminations que dans le problème de la séparation de sources.

La séparation sera plus performante si l'on considère la diagonalisation simultanée d'un ensemble $\{\underline{\mathbf{R}}_x(\tau_i) | i = 1, \dots, K\}$ de K matrices d'intercovariance blanchies. Celles-ci sont simultanément diagonalisées par \mathbf{U} comme il est montré par la relation (16). Ainsi, au lieu de diagonaliser une seule matrice, on détermine la rotation manquante en diagonalisant conjointement un ensemble de matrices d'intercovariance blanchies dans une même base \mathbf{U} . Le schéma général de l'algorithme décrit ci-dessus qui est nommé **SOBI** (*Second Order Blind Identification method*), est le suivant :

1. Estimation de la matrice de blanchiment ,
2. blanchiment des données : $\tilde{\mathbf{x}} = \mathbf{E} \mathbf{D}^{-\frac{1}{2}} \mathbf{E}^T \mathbf{x}$,
3. estimation de K matrices d'intercovariance $\underline{\mathbf{R}}_x(\tau_k)$ de $\tilde{\mathbf{x}}$ à différents retards τ_k , $k = 1, \dots, K$,
4. diagonalisation conjointe approchée des matrices d'intercovariance dans une base \mathbf{U} ,
5. estimation de la matrice de mélange par : $\hat{\mathbf{A}} = \mathbf{E}^T \mathbf{D}^{\frac{1}{2}} \mathbf{E} \mathbf{U}$,
6. estimation des signaux sources.

3.1 Version simplifiée de l'algorithme SOBI

Dans ce qui suit, nous allons décrire une version simplifiée de l'algorithme SOBI dans le cas particulier de deux signaux sources reçu par deux capteurs $n = 2$. A partir des matrices d'autocorrélation, et sans blanchiment préalable, nous pouvons avoir le système d'équation :

$$\mathbf{R}_{x_1 x_1} = a_{11}^2 \mathbf{R}_{s_1 s_1} + a_{12}^2 \mathbf{R}_{s_2 s_2} + \sigma^2 \mathbf{I} \quad (17)$$

$$\mathbf{R}_{x_2 x_2} = a_{21}^2 \mathbf{R}_{s_1 s_1} + a_{22}^2 \mathbf{R}_{s_2 s_2} + \sigma^2 \mathbf{I} \quad (18)$$

$$\mathbf{R}_{x_1 x_2} = a_{11} a_{21} \mathbf{R}_{s_1 s_1} + a_{12} a_{22} \mathbf{R}_{s_2 s_2}. \quad (19)$$

où $\mathbf{R}_{x_i x_j} = \mathbb{E}[\mathbf{x}_i \mathbf{x}_j^T]$ avec $\mathbf{x}_i = [x_i(1), \dots, x_i(N)]^T$. Nous observons qu'à partir de ce système d'équations nous avons la possibilité d'estimer les coefficients de la matrice de mélange \mathbf{A} . Par conséquent la solution de ce système est donnée par :

$$\hat{\mathbf{A}} = \begin{bmatrix} \beta F_1 - (T_1 - \sigma^2)d_1 & \beta F_{12} - T_{12}d_2 \\ \beta F_{12} - T_{12}d_1 & \beta F_2 - (T_2 - \sigma^2)d_2 \end{bmatrix} \quad (20)$$

où

$$F_1 = \text{off}(\mathbf{R}_{x_1 x_1}) \quad (21)$$

$$F_2 = \text{off}(\mathbf{R}_{x_2 x_2}) \quad (22)$$

$$F_{12} = \text{off}(\mathbf{R}_{x_1 x_2}) \quad (23)$$

$$T_1 = \text{tr}(\mathbf{R}_{x_1 x_1}) \quad (24)$$

$$T_2 = \text{tr}(\mathbf{R}_{x_2 x_2}) \quad (25)$$

$$T_{12} = \text{tr}(\mathbf{R}_{x_1 x_2}) \quad (26)$$

avec les opérateurs $\text{off}(\cdot)$ et $\text{tr}(\cdot)$ sont définis comme suit :

$$\text{off}(\mathbf{M}) = \frac{1}{N(N-1)} \sum_{i \neq j} \sum_{j=1}^N M_{ij} \quad (27)$$

$$\text{tr}(\mathbf{M}) = \frac{1}{N} \sum_{i=1}^N M_{ii} \quad (28)$$

$$\alpha = 2F_{12}T_{12} - (F_1(T_2 - \sigma^2) + F_2(T_1 - \sigma^2)) \quad (29)$$

$$\beta = 2(T_{12}^2 - (T_1 - \sigma^2)(T_2 - \sigma^2)) \quad (30)$$

$$\begin{aligned} \gamma^2 = & (F_1(T_2 - \sigma^2) - F_2(T_1 - \sigma^2))^2 + 4(F_{12}(T_2 \\ & - \sigma^2) - T_{12}F_2)(F_{12}(T_1 - \sigma^2) - T_{12}F_1). \end{aligned} \quad (31)$$

et pour finir, $d_1 = \alpha - \gamma$ et $d_2 = \alpha + \gamma$.

Question 15 : Mettez en œuvre cet algorithme sur les signaux simulés se trouvant dans le fichier `SignauxMelange.mat` en estimant la matrice \mathbf{A} par l'équation (20). Sachant que dans notre cas $\sigma^2 = 0$.

Annexe 1 : Évaluation des performances

Le critère objectif utilisé dans ce projet pour évaluer les performances des algorithmes étudiés est *l'erreur quadratique moyenne normalisée* (EQMN) définie comme suit :

$$\text{EQMN}_i = 1 - \left(\frac{\langle s_i, \hat{s}_i \rangle}{\|s_i\| \|s_i\|} \right)^2 \quad (32)$$