

TP Séparation de sources

Le but de ce TP est de mettre en oeuvre une méthode élémentaire de séparation de deux sources. La méthode est basée sur un blanchiment suivi de la maximisation d'une fonction de contraste.

1 Génération des sources et du mélange

Afin de mener l'étude des paragraphes 2 et 3, nous souhaitons générer des sources présentant certaines caractéristiques et un mélange de ces dernières.

1.1 Sources

On rappelle que Matlab dispose d'un générateur de variables aléatoires gaussiennes (`randn`) et uniformes (`rand`).

1. – Générer un vecteur de T réalisations d'une variable aléatoire centrée, de variance unité et distribuée uniformément sur un intervalle (choisir une valeur de T convenable).
 - Calculer les valeurs théoriques des quatre premiers moments et cumulants d'une telle variable aléatoire. Quelle est la valeur du kurtosis (défini comme $C_4\{s\} \triangleq E\{s^4\} - 3E\{s^2\}^2$) ?
 - Mettre en oeuvre les estimateurs de la moyenne empirique des moments et s'en servir pour estimer les moments et cumulants calculés théoriquement ci-dessus. Vérifier la cohérence des résultats.
2. Mêmes questions avec une variable aléatoire gaussienne.
3. On souhaite générer une variable aléatoire de kurtosis positif. On prendra une variable aléatoire de Laplace (ou encore distribution double exponentielle), centrée et de variance unité. Quelle est cette loi ? Comment la générer ? Reprendre les questions en 1 ci-dessus.

1.2 Mélange

Générer T réalisations d'un vecteur aléatoire de deux sources indépendantes que l'on pourra stocker dans une matrice $\mathbf{s} \in \mathbb{R}^{2 \times T}$. Tirer une matrice $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ aléatoirement et réaliser le mélange $\mathbf{x} = \mathbf{A}\mathbf{s}$ correspondant. Représenter les signaux et donner une image des distributions conjointes des sources et des observations.

2 Etude du blanchiment

1. Calculer la covariance empirique et théorique des sources et des observations ; vérifier la cohérence des résultats.
2. Rappeler le principe du blanchiment et comment il peut se réaliser à l'aide de la décomposition en valeurs singulières. Mettre en oeuvre un préblanchiment (utiliser la commande `svd`). Vérifier le résultat en calculant la covariance empirique des données blanchies. Donner une image de la densité de probabilité conjointe des données après blanchiment.

3 Comparaison de quelques contrastes

L'étape de blanchiment effectuée, il reste à déterminer une matrice orthogonale. Nous comparons quelques contrastes, valables sous contrainte de blanchiment, et qui permettent d'effectuer cette tâche. Si $C_4\{\cdot\}$ désigne le cumulants d'ordre quatre, on considérera :

$$\begin{aligned} \mathcal{C}_1(\mathbf{y}) &\triangleq \sum_{i=1}^N |C_4\{y_i\}|, & \mathcal{C}_2(\mathbf{y}) &\triangleq - \sum_{i=1}^N C_4\{y_i\}, \\ \mathcal{C}_3(\mathbf{y}) &\triangleq \sum_{i=1}^N (C_4\{y_i\})^2, & \mathcal{C}_4(\mathbf{y}) &\triangleq \sum_{i=1}^N \kappa_i C_4\{y_i\}. \end{aligned}$$

où $\kappa_i = C_4\{s_i\}$ pour tout i (les cumulants des sources sont alors supposés connus). Nous nous limiterons au cas $N = 2$ sources. Soit la matrice orthogonale paramétrée par l'angle θ :

$$\mathbf{Q}(\theta) \triangleq \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

1. Générer des sources uniformes et le mélange $\mathbf{y} = \mathbf{Q}(\theta)\mathbf{s}$.
2. Mettre en œuvre les estimateurs des fonctions de contraste ci-dessus sur une grille discrète de valeurs de θ (à choisir judicieusement en fonction du nombre T d'échantillons afin de ne pas imposer une charge de calcul trop importante).
3. Tracer la courbe des différents contrastes en fonction de θ .
4. Prendre des sources différentes (par exemple choisies dans celles suggérées dans la partie 1). Comparer les différents contrastes, vérifier leur validité dans différents cas et commenter.

4 Application à des signaux sonores

1. Mettre en œuvre une séparation complète avec les différents types de sources proposées dans la partie 1. On effectuera un blanchiment suivi d'une détermination de l'angle de rotation. Pour la maximisation du contraste, on pourra se baser sur l'estimation de sa valeur réalisée précédemment sur une grille discrète (on suggère d'utiliser les commandes `max` et `find`).
2. Télécharger les signaux sonores disponibles sur ma page web¹. Les mélanger par une matrice tirée aléatoirement, les séparer et comparer les sorties du séparateur aux sources (compte tenu du nombre d'échantillons, veiller à ne pas effectuer un trop grand nombre d'estimations du contraste).

¹<http://www-public.it-sudparis.eu/~castella/index.php?page=enseignement>