Résumé

Le travail présenté dans cette thèse concerne d'une part, le problème de séparation aveugle de sources et d'autre part, celui d'estimation de paramètres en aveugle. Pour ces deux problèmes, une nouvelle classe de modèles de mélanges est étudiée, celle des mélanges inversibles, décrits par des équations différentielles.
Pour le problème de séparation aveugle de source, nous définissons la notion de séparabilité selon un modèle d'entrées fixé, puis proposons des méthodes de séparation basées sur des statistiques des signaux et de leurs dérivées.
Pour le problème d'estimation aveugle de paramètres, nous définissons la notion d'identifiabilité en aveugle selon un modèle d'entrées fixé et présentons une méthode d'estimation exploitant également les résultats sur les dérivées de signaux.
Les techniques d'analyse par intervalles sont exploitées afin d'obtenir des solutions garanties.
Cette étude, principalement théorique, est illustrée par de nombreux exemples simples.

Mots-clés : Séparation aveugle de sources, Mélange inversible, Estimation aveugle de paramètres, Analyse par intervalles, Dérivées de signaux aléatoires, Séparabilité, Identifiabilité en aveugle

Abstract

This dissertation presents the problem of blind source separation and the problem of blind parameter estimation. For these problems, a new class of mixture models is considered : invertible mixtures, described by differential equations. For the blind source separation problem, respectively blind parameter estimation problem, we define the concept of separability, resp. of blind identifiability, according to an input model. Then, we propose a separation, resp. an estimation, method based on new results concerning random signal derivative. Interval analysis methods are exploited in order to obtain guaranteed solutions. Although our study is mainly theoretical, many illustrative examples are proposed.

Keywords: Blind source separation, Invertible mixture, Blind parameter estimation, Interval analysis, Derivative of random signal, Separability, Blind identifiability

Remerciements

Au moment où s'achève l'écriture de ce manuscrit, vient l'heure des traditionnels remerciements . . .

Mes premiers remerciement vont à mes directeurs de thèse, Christian Jutten et Luc Jaulin, qui ont accepté de coencadrer cette thèse et m'ont témoigné leur soutien et leur confiance malgré des débuts difficiles. Qu'ils trouvent ici l'expression de ma reconnaissance la plus sincère.
Je remercie Christian de m'avoir communiqué son expérience et particulièrement de m'avoir laissé une grande liberté dans la direction de cette thèse.
Je remercie Luc pour ses nombreux conseils avisés et sans qui cette thèse ne serait pas ce qu'elle est aujourd'hui.
Je remercie aussi Vincent Vigneron pour son encadrement toujours dans la bonne humeur et les nombreuses discussions que nous avons eu au LIS.
Je ne vous remercierai jamais assez, tous les trois, de m'avoir initié à la recherche.

Un remerciement tout particulier à J.-L. Ferrier pour son accueil au LISA qui m'a permis d'achever ma thèse dans les meilleures conditions possibles. Merci aussi à P.-Y Coulon pour ses précieux conseils à un moment difficile.

J'exprime toute ma reconnaissance à E. Moreau, Professeur à l'Université de Toulon, et E. Walter, Directeur de recherche CNRS au Laboratoire des Signaux et Systèmes, pour l'honneur qu'ils m'ont fait en acceptant de rapporter mon travail. Je leur suis également très reconnaissant pour les nombreuses remarques et suggestions qu'ils ont apportées.

Je voudrais également remercier D.-T. Pham, Directeur de recherche CNRS à l'Université de Grenoble, et, de nouveau, J.-L. Ferrier, Professeur à l'Université d'Angers, de l'intérêt qu'ils ont porté à mes travaux en acceptant de participer au jury.

Un grand merci à mes collègues doctorants du LISA : Laurent, Nico, Xavier, Samir, Michel, Céline, Massa . . . avec qui j'ai eu le plaisir de partager le bureau et bien d'autres choses.
Mes remerciements s'adressent également à tous les membres des laboratoires LIS et LISA qui, durant toutes ces années, m'ont permis de travailler dans un contexte agréable.
En particulier, un grand merci à Laurent Hardouin pour sa sympathie et la confiance qu'il m'a accordée en me permettant de dispenser de nombreux cours.

J'adresse mes remerciements à toute ma famille et mes amis qui, de près comme de loin m'ont aidé et encouragé au moment opportun.

Enfin et surtout, mille fois MERCI à Mathilde.

à Éloi et Mathilde

Notations

$\mathcal{P}(\Omega)$: ensemble des parties de l'ensemble Ω.

$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

$\mathcal{V}$: ensemble des variables aléatoires.

$\mathbb{R}^{\mathbb{R}}$: ensemble des signaux déterministes défini sur $\mathbb{R}$ à valeur dans $\mathbb{R}$.

$\mathfrak{S}$: ensemble des signaux aléatoires.

$\mathcal{S}$: ensemble des signaux aléatoires stationnaires, ergodiques et lisses.

I^n (ou I_s^n) : ensemble des n signaux aléatoires statistiquement indépendants.

$\mathcal{G}^n$ (ou $\mathcal{G}_s^n$) : ensemble des n signaux aléatoires mutuellement gaussiens.

$\mathcal{M}_{\mathbf{u}}$: ensemble des moments d'un vecteur de signaux aléatoires $\mathbf{u}(.)$.

$\mathcal{A}_{\mathbf{u}}$: algèbre des moments d'un vecteur de signaux aléatoires $\mathbf{u}(.)$.

$\mathcal{F}$: ensemble des fonctions définies sur $\mathbb{R}^{\mathbb{R}}$ à valeurs dans $\mathbb{R}$.

$\mathcal{F}^i$: ensemble des fonctions définies sur $\left(\mathbb{R}^{\mathbb{R}}\right)^i$ à valeurs dans $\mathbb{R}$.

$\mathfrak{F}$: ensemble des fonctions définies sur $\mathbb{R}^{\mathbb{R}}$ à valeurs dans $\mathbb{R}^{\mathbb{R}}$.

$\mathfrak{F}^i$: ensemble des fonctions définies sur $\left(\mathbb{R}^{\mathbb{R}}\right)^i$ à valeurs dans $\mathbb{R}^{\mathbb{R}}$.

$\mathcal{L}$: ensemble des fonctions linéaires définies sur $\mathbb{R}^{\mathbb{R}}$ à valeurs dans $\mathbb{R}$.

$\mathcal{L}^i$: ensemble des fonctions linéaires définies sur $\left(\mathbb{R}^{\mathbb{R}}\right)^i$ à valeurs dans $\mathbb{R}$.

$\tilde{\mathcal{L}}$: ensemble des fonctions affines définies sur $\mathbb{R}^{\mathbb{R}}$ à valeurs dans $\mathbb{R}$.

$\tilde{\mathcal{L}}^i$: ensemble des fonctions affines définies sur $\left(\mathbb{R}^{\mathbb{R}}\right)^i$ à valeurs dans $\mathbb{R}$.

$\mathfrak{L}$: ensemble des fonctions affines définies sur $\mathbb{R}^{\mathbb{R}}$ à valeurs dans $\mathbb{R}^{\mathbb{R}}$.

$\mathbf{Id}_n$: matrice identité.

$\mathfrak{D}$: ensemble des matrices diagonales.

$\mathfrak{P}$: ensemble des matrices de permutation.

$\mathcal{C}^1$: ensemble des fonctions continuement différentiables (continues, dérivables, de dérivée continue) .

S_m : ensemble des permutations de $\{1, \ldots, m\}$.

Table des matières

Introduction

Le traitement du signal est une discipline dont l'objectif est principalement d'extraire les informations utiles sur des phénomènes physiques à partir de mesures généralement perturbées. Le modèle formel utilisé pour représenter les problèmes considérés est dessiné sur la figure 1. Les signaux **u** et **y** sont respectivement les entrées et les sorties du système **H**. Par exemple, les signaux considérés peuvent être des signaux sonores et **H** représente alors le filtrage dû à la propagation dans le milieu.

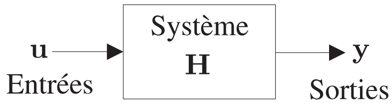


Figure 1 – Système **H** d'entrées **u** et de sorties **y**.

A partir du modèle représenté sur la figure 1, on peut considérer trois problèmes d'estimation statistique :

– Le *problème d'identification* : connaissant les entrées **u** et les sorties **y**, il s'agit d'estimer le système **H**.
– Le *problème de filtrage* : connaissant les entrées **u** et le système **H**, on recherche les sorties **y**.
– Le *problème d'égalisation* (ou déconvolution) s'attache à retrouver les entrées **u** connaissant le système **H** et les sorties **y**.

Dans cette thèse, nous nous intéressons au problème de *séparation aveugle de sources*. C'est un problème transversal qui consiste à estimer les entrées **u** (et éventuellement identifier le système **H**) connaissant uniquement les sorties **y**. En d'autres termes, à partir d'observations d'un mélange de signaux (appelés sources), mélange inconnu (d'où le terme "aveugle"), il consiste à retrouver les sources.

Nous verrons que ce problème ne peut être résolu qu'à l'aide d'hypothèses sur les entrées **u** et le système **H**. Le système le plus simple étudié en séparation aveugle de sources suppose le système **H** linéaire statique et les sources **u** statistiquement indépendantes. Toutefois, le problème de séparation aveugle de sources s'est considérablement enrichi au cours des dernières années et des résultats théoriques et des algorithmes ont été proposés pour des mélanges plus complexes. Ces mélanges constituent la majeure partie des études actuelles. C'est aujourd'hui un problème très général qui se manifeste dans plusieurs domaines et a de nombreuses applications.

0.1 La séparation de source au quotidien

La séparation de sources est un sujet de recherche très actif en raison de son intérêt théorique et de ses nombreuses applications pratiques. Dans la vie courante, nous pouvons citer bon nombre d'applications de séparation de sources correspondant à différents types de signaux rencontrés (signaux de paroles, signaux biomédicaux, etc.). A titre d'exemple :

– Dans le domaine *médical*, elle est utilisée pour l'extraction non invasive de l'ElectroCardioGramme (ECG), l'ElectroMyoGramme (EMG) ou l'ElectroEncéphaloGramme (EEG) [13][14].
Par exemple, l'extraction non-invasive des battements cardiaques du foetus, à partir de signaux ECG enregistés à l'aide d'électrodes placées sur le ventre de la mère (voir figure (2)), est possible grâce à la séparation de sources. Les travaux consistent aujourd'hui à tenter d'extraire la forme

complète de l'ECG du foetus (le complexe PQRST) et non simplement les pics R comme c'est le cas à l'heure actuelle.

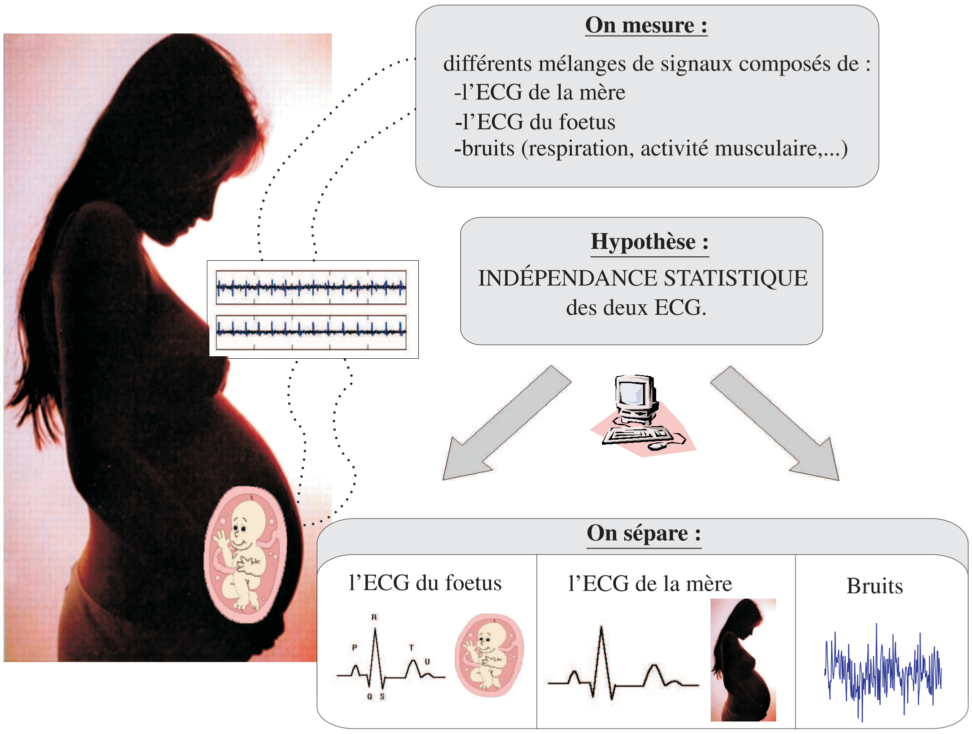


Figure 2 – Extraction non invasive de l'ECG d'un foetus.

– Dans le domaine de la *communication*, le problème le plus connu (en acoustique) est certainement "l'effet cocktail party". Lors d'une fête, il s'agit de séparer la parole de notre interlocuteur parmi les nombreuses voix, la musique et autres bruits de fond.
  On trouve aussi des applications en téléphonie mobile où la parole de notre interlocuteur peut être issue d'un environnement bruyant (musiques, voitures qui passent dans la rue, etc...). Le signal enregistré est parfois considérablement "abimé". Il convient alors de séparer la source utile (parole) de la source perturbante (bruit de fond) pour parvenir ensuite à atténuer ces bruits de fond.
– Dans le domaine de la *sismologie*, on retrouve également le problème de séparation de sources dans le sens où le signal émis, par une vibration ou une explosion, est inconnu et observé au travers d'un mélange reflétant des sources multiples. La séparation de sources doit permettre aux sismologues, à partir d'enregistrements et moyennant des renseignements sur la fonction de transfert du sous-sol, d'obtenir des informations sur la source et de la modéliser.

Cette liste d'applications est non exhaustive. Il y en a bien d'autres, comme en astrophysique [16], pour le contrôle du trafic aérien [23].

Maintenant, présentons plus formellement le problème de séparation aveugle de sources.

0.2 Le problème de séparation aveugle de sources

Dans un contexte général, le problème de séparation aveugle de sources se formule de la manière suivante :

Considérons k signaux aléatoires, appelés *observations* et notés $y_1(.), \ldots, y_k(.)$, issus d'un mélange de l signaux aléatoires appelés *sources* et notés $u_1(.), \ldots, u_l(.)$. On note $\mathcal{T}$ le système liant les observations aux sources de sorte que

$$\mathbf{y}(.) = \mathcal{T}(\mathbf{u}(.)) \tag{1}$$

où $\mathbf{y}(.) = (y_1(.), \ldots, y_k(.))^T$ et $\mathbf{u}(.) = (u_1(.), \ldots, u_l(.))^T$.

Si nous supposons que les observations $\mathbf{y}$ sont connues et que le système $\mathcal{T}$ est connu, alors le problème de reconstruction des sources consiste à inverser le système de mélange $\mathcal{T}$: ceci revient à un problème de déconvolution.

Ici, nous nous plaçons dans un contexte aveugle, c'est-à-dire que seules les observations $\mathbf{y}$ sont connues ($\mathbf{u}$ et $\mathcal{T}$ sont inconnus). Le problème de séparation aveugle de sources consiste à restituer les sources $\mathbf{u}$ à partir des observations $\mathbf{y}$. Il est clair que sans hypothèse supplémentaire, la résolution de ce problème s'avère tout à fait impossible. Il est donc nécessaire de faire des hypothèses sur le mélange ainsi que sur les sources. Par conséquent, le problème de séparation aveugle de sources possède différents niveaux de difficulté selon les hypothèses formulées *a priori*. Plusieurs paramètres sont à prendre en considération :

– Tout d'abord, cela dépend de la **nature du système de mélange** $\mathcal{T}$. Le cas le plus simple consiste à considérer un *mélange linéaire instantané* (ou statique) pour lequel, à chaque instant, les observations sont des combinaisons linéaires des sources aux mêmes instants. Le système (1) se simplifie et on obtient

$$\mathbf{y}(t) = \mathbf{A}\mathbf{u}(t), \tag{2}$$

où $\mathbf{A}$ est une matrice de taille $k \times l$.

Dans la pratique, le *mélange linéaire convolutif* est plus réaliste. Pour ce mélange, à chaque instant, les observations dépendent des sources à différents instants antérieurs :

$$\mathbf{y}(t) = \sum_i \mathbf{A}_i \mathbf{u}(t - \tau_i), \tag{3}$$

où $\forall i$, $\mathbf{A}_i \in \mathbb{R}^{k \times l}$ et $\tau_i \in \mathbb{N}$ (le retard). Le mélange convolutif tient compte de la propagation des signaux dans le milieu considéré. Par exemple, lorsque plusieurs musiciens sont enregistrés par différents microphones dans une salle, la propagation des signaux jusqu'aux microphones est modélisée par une convolution des sources.

Enfin, les mélanges peuvent également être considérés comme *post non-linéaires convolutifs*, par exemple lorsque les microphones enregistrant les musiciens présentent des saturations. Ce type de mélange non-linéaire est constitué d'un mélange linéaire convolutif suivi de non linéarités sur chacune des composantes des sorties :

$$\forall i \in \{1, \ldots, k\}, \; y_i(.) = f_i\left(\sum_{j_1=1}^{l} a_{ij_1} u_{j_1}(t - \tau_j) + \cdots + \sum_{j_n=1}^{l} a_{ij_n} u_{j_n}(t - \tau_{j_n})\right), \tag{4}$$

où f_j est une fonction non linéaire (cette non-linéarité est supposée inversible).

A noter qu'il existe également des *mélanges post non-linéaires* composés d'un mélange linéaire statique suivi de non-linéarités sur chaque composante :

$$\forall i \in \{1, \ldots, k\}, \; y_i(.) = f_i\left(\sum_{j=1}^{l} a_{ij} u_j(t)\right). \tag{5}$$

– La seconde hypothèse à prendre en considération est une information sur la **nature des sources**. Dans la majeure partie des problèmes de séparation aveugle de sources, nous verrons que les sources sont supposées mutuellement indépendantes.
– Enfin, il faut tenir compte également du **nombre de sources** l par rapport au nombre d'observations k. Bien entendu, le cas sur-déterminé ($k > l$) est plus simple à résoudre que le cas sous-déterminé ($k < l$), ce dernier ne pouvant être résolu qu'à l'aide d'importantes informations *a priori* [86][67][29].
– Pour finir, il est important de noter la nécessité de disposer d'observations distinctes. Autrement dit, les mélanges correspondant à chacune des observations doivent être différents. Les observations étant en général issues de capteurs, on peut voir cela comme une nécessité de diversité spatiale. Par exemple, les capteurs sont situés à des endroits différents pour obtenir des mélanges distincts ou encore, les capteurs utilisés sont de natures différentes.

L'objectif de la séparation de sources est de rechercher des signaux aléatoires $\hat{u}_1(.), \ldots, \hat{u}_l(.)$, appelés *sources reconstituées*, qui représentent une estimation des sources $\mathbf{u}(.)$. On note $\mathcal{H}$ le système de séparation :

$$\hat{\mathbf{u}}(.) = \mathcal{H}(\mathbf{y}(.)),$$

où $\hat{\mathbf{u}}(.) = (\hat{u}_1(.), \ldots, \hat{u}_l(.))^T$.

Apparaissent alors deux problèmes qui peuvent être traités différemment. Le premier consiste à estimer la structure du mélange $\mathcal{T}$, le second consiste à estimer directement les sources reconstituées $\hat{\mathbf{u}}(.)$. Bien sûr, dans le cas d'un mélange inversible, ces deux problèmes sont identiques. L'identification de $\mathcal{T}$ conduit directement à la restauration des sources $\hat{\mathbf{u}}(.) = \mathcal{T}^{-1}(\mathbf{y}(.))$.

Après cette présentation générale du problème de séparation de sources, nous présentons maintenant précisément le problème auquel nous nous sommes intéressés dans cette thèse.

0.3 Notre approche

0.3.1 Nature des mélanges étudiés

Dans cette thèse, nous abordons principalement le problème de séparation aveugle de sources pour lequel le modèle de mélange est décrit par des équations différentielles. Pour ces nouveaux systèmes de mélanges, qualifiés d'inversibles, nous proposons une méthodologie de résolution basée sur des résultats concernant les dérivées de signaux aléatoires. Bien que nous présentions en majorité des résultats concernant les systèmes de mélanges linéaires inversibles, notre approche est générale. La méthodologie proposée semble s'étendre sans difficulté aux systèmes de mélanges inversibles non linéaires, c'est ce que nous évoquons en guise de perspectives à la fin du chapitre 3, n'ayant pas eu le temps de le vérifier rigoureusement.

Précisons toutefois que, dans un contexte général, le problème de séparation aveugle de sources pour des mélanges non linéaires est insoluble [34]. En effet, l'exemple ci-dessous montre qu'il existe des transformations non linéaires qui conservent l'indépendance des sorties.

Exemple 0.1. *Soit u_1 une variable qui suit une distribution de Rayleigh de densité de probabilité*

$$f_{u_1}(u_1) = u_1 \exp\left(-\frac{u_1^2}{2}\right),$$

et u_2 une variable aléatoire statistiquement indépendante de u_1, qui suit une distribution uniforme sur $[0, 2\pi]$. Considérons la transformation non linéaire suivante

$$\left\{ \begin{array}{l} y_1 = u_1 \cos(u_2) \\ y_2 = u_1 \sin(u_2) \end{array} \right. ,$$

où y_1 et y_2 sont deux nouvelles variables aléatoires. La densité de probabilité conjointe de y_1 et y_2, notée $p_{y_1 y_2}(y_1, y_2)$, (voir la proposition A.75 pour davantage de détails) est

$$f_{y_1 y_2}(y_1, y_2) = \det J(u_1, u_2) f_{u_1 u_2}(u_1, u_2),$$

avec

$$J(u_1, u_2) = \begin{pmatrix} \cos(u_2) & -u_1 \sin(u_2) \\ \sin(u_2) & u_1 \cos(u_2) \end{pmatrix} \quad et \quad \det J(u_1, u_2) = u_1 \ .$$

Ainsi, on a

$$f_{y_1 y_2}(y_1, y_2) = \exp\left(-\frac{u_1^2}{2}\right)$$

Par conséquent, y_1 et y_2 sont deux variables aléatoires (gaussiennes de moyenne nulle et de variance unitaire) statistiquement indépendantes .

La séparation des mélanges non linéaires étant impossible dans le cas général, ceci justifie notre choix d'étudier une sous classe particulière de mélanges : les systèmes inversibles. A noter que dans la littérature, d'autres classes de mélanges non linéaires ont été considérés, par exemple, les mélanges post non linéaires introduits par Taleb et Jutten dans [87] (voir paragraphe 0.2).

0.3.2 Contributions

Cette thèse comporte deux contributions principales.

La première concerne le problème de séparation de sources pour les systèmes de mélanges inversibles, pour lesquels des résultats sur les dérivées de signaux aléatoires pourront être pleinement exploités afin d'établir des critères de séparation. Ce travail généralise les travaux de Cavassilas, restreints au mélange linéaire instantané. Nous abordons ainsi le problème de séparation de sources sous un angle original.

La seconde contribution concerne le problème plus général (qui englobe le problème de séparation de sources) d'*estimation aveugle de paramètres*. Il s'agit, à partir de l'observation des sorties **y**, d'estimer les paramètres d'un système connu à l'aide d'hypothèses statistiques sur les entrées **u**. Dans cette thèse, nous exploitons exclusivement les hypothèses d'indépendance et de gaussianité des entrées. Ce travail généralise les travaux sur l'identifiabilité de Walter.
A noter que, de la même façon que pour le problème de séparation de source, les systèmes paramétrés considérés sont inversibles.

Pour ces deux problèmes d'estimation en aveugle, les outils d'analyse par intervalles sont largement exploités. Nous verrons les avantages (et les inconvénients) que procurent ces méthodes qualifiées de "garanties".

0.4 Synopsis

Ce document contient quatre chapitres. Les contributions originales sont concentrées dans les chapitres 3 et 4. Le mémoire est organisé de la manière suivante.

1 - État de l'art en séparation aveugle de sources
Dans ce premier chapitre, nous présentons différentes techniques de résolution existantes pour le problème de séparation aveugle de sources. Nous nous attachons principalement à décrire les méthodes de résolution dans le cas des systèmes de mélanges linéaires statiques et distinguons à cette occasion les méthodes basées sur des statistiques d'ordre deux et celles basées sur des statistiques d'ordre supérieur. Nous verrons que le problème de séparation de source, même dans le cas le plus simple d'un mélange linéaire instantané, nous amène à résoudre des systèmes d'équations non linéaires. C'est pourquoi nous avons opté par la suite pour des méthodes d'analyse par intervalles afin de résoudre ce type de problème car elles s'avèrent très efficaces pour la résolution de systèmes non linéaires, notamment en raison de leur capacité à fournir toutes les solutions.

2 - Analyse par intervalles
Dans ce second chapitre, nous présentons les notions fondamentales du calcul par intervalles introduites notamment par Walter et Jaulin. Nous détaillons deux algorithmes utilisés dans la suite : l'algorithme SIVIA de résolution de systèmes d'équations non linéaires et l'algorithme OPTICRIT d'optimisation globale (non linéaire). Ce dernier algorithme est illustré par un exemple concret.

3 - Séparation aveugle de sources
Nous présentons dans ce troisième chapitre notre travail concernant la séparation aveugle de sources.
Dans un premier temps, nous montrons comment les techniques d'analyse par intervalles permettent d'étudier de façon simple et exhaustive les points d'équilibres de l'algorithme Hérault-Jutten ainsi que leur stabilité.
Ensuite, nous présentons une méthodologie de résolution du problème de séparation de sources pour des systèmes de mélanges inversibles. L'originalité de notre approche réside dans l'extension du modèle de mélanges et l'utilisation conjointe des dérivées de signaux aléatoires et du calcul par intervalles pour résoudre les problèmes d'optimisation non convexe qui en découlent.

4 - Estimation aveugle de paramètres
Enfin, nous traitons un problème plus général que celui de la séparation de sources : l'estimation de paramètres en aveugle. Nous nous intéressons dans ce dernier chapitre à l'identifiabilité en aveugle, par analogie à l'identifiabilité classique [90]. Pour finir, nous présentons quelques exemples illustratifs.

Après lecture, les rapporteurs ont souhaité que la partie principale de la thèse soit centrée sur les résultats nouveaux. Pour cette raison, les deux chapitres sur les notions de base concernant la théorie des probabilité et les signaux aléatoires sont proposés en annexe. Ceci explique le volume relativement conséquent des annexes A et B, destinées à l'usage du lecteur non accoutumé à manipuler la théorie des probabilités ou peu familiarisé avec les signaux aléatoires.

Annexes A et B - Notions de Probabilités et signaux aléatoires
Tout d'abord, le problème de séparation de sources met en jeu des phénomènes aléatoires. Dans la mesure où les hypothèses émises sur les sources sont des hypothèses statistiques, nous proposons, dans un

souci purement pédagogique, des rappels élémentaires sur les probabilités.
Nous présentons ensuite une vision originale de la théorie des signaux aléatoires réels, l'originalité étant principalement due à l'introduction de la notation $\mathbb{R}^{\mathbb{R}}$ pour l'ensemble des signaux. On comprendra aisément que $\mathbb{R}^2$ est l'ensemble des couples de $\mathbb{R}$, $\mathbb{R}^{\mathbb{N}}$ l'ensemble des suites réelles et, par suite, $\mathbb{R}^{\mathbb{R}}$ correspond à l'ensemble des fonctions (ou signaux) réelles.
Nous insistons particulièrement sur les propriétés des dérivées de signaux et de leurs moments statistiques, qui sont la base de notre travail de recherche. Cette annexe est également l'occasion d'introduire les définitions pour l'indépendance et pour la gaussianité de signaux aléatoires, utilisées dans les chapitres 3 et 4.
Nous terminons par quelques résultats sur les systèmes entrée-sortie. Nous évoquons notamment le lien entre les propriétés statistiques (à l'ordre deux) des entrées et des sorties.

Annexes C, D, E et F

Le lecteur trouvera dans les quatre annexes suivantes des détails concernant les outils mathématiques que nous avons été amenés à manipuler dans nos travaux : méthode de Newton, décomposition en base de Gröbner, méthode d'estimation des dérivées et algorithme de diagonalisation conjointe de matrices.

CHAPITRE 1

État de l'art en séparation aveugle de sources

La séparation de sources est un problème relativement récent en traitement du signal, qui a été formulé pour la première fois au début des années 1980 par Hérault, Ans et Jutten dans [48] puis [50]. La modélisation d'un phénomène biologique décrit dans [62] est à l'origine des travaux sur la séparation de sources.

Dans ce chapitre, nous nous intéressons au problème de séparation de sources pour le cas particulier des mélanges linéaires statiques pour lesquels nous présentons différentes méthodes de séparation existantes.

1.1 Introduction

Commençons par rappeler que dans le cas d'un mélange linéaire statique, les observations à un instant donné sont une combinaison linéaire des sources au même instant. Supposons que l'on dispose d'autant d'observations y_i que de sources u_i, un tel mélange est alors modélisé par une matrice $\mathbf{A}$ de taille $n \times n$,

$$\mathbf{y}(t) = \mathbf{A}\mathbf{u}(t) \,. \tag{1.1}$$

Si la matrice $\mathbf{A}$ est régulière, alors, il existe une matrice $\mathbf{B}$ permettant d'obtenir une estimation $\hat{\mathbf{u}}$ des sources (voir figure 1.1) telle que

$$\hat{\mathbf{u}}(t) = \mathbf{B}\mathbf{y}(t) \,.$$

La *matrice de transfert* entre les sources et leurs estimations, notée $\mathbf{H}$ vérifie alors

$$\mathbf{H} = \mathbf{B}\mathbf{A}.$$

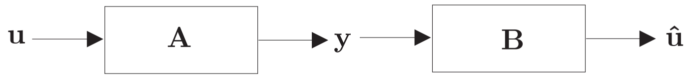


Figure 1.1 – Système de séparation.

La solution idéale consisterait à prendre $\mathbf{B} = \mathbf{A}^{-1}$ mais, en séparation aveugle de sources, seules les observations $\mathbf{y}$ sont supposées connues et cela engendre quelques indéterminations que nous allons préciser dans le paragraphe suivant.

1.2 Indéterminations

Nous notons $\mathbf{A}_j$, $1 \leq j \leq n$, les vecteurs colonnes de la matrice $\mathbf{A}$. On a

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} A_1 & \cdots & A_n \end{pmatrix} \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix}.$$

Si l'on permute deux sources, par exemple $u_j(t)$ et $u_{j+1}(t)$, ainsi que les vecteurs colonnes correspondants de la matrice $\mathbf{A}$, on retrouve exactement le même vecteur d'observations $\mathbf{y}$:

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_j(t) \\ y_{j+1}(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} A_1 & \cdots & A_{j+1} & A_j & \cdots & A_n \end{pmatrix} \begin{pmatrix} u_1(t) \\ \vdots \\ u_{j+1}(t) \\ u_j(t) \\ \vdots \\ u_n(t) \end{pmatrix}. \tag{1.2}$$

Si l'on multiplie chaque source $u_j(t)$ par un scalaire α_j non nul et que l'on divise les vecteurs colonnes correspondants de la matrice $\mathbf{A}$ par ces mêmes scalaires α_j, on ne modifie pas non plus le vecteur d'observations $\mathbf{y}$:

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_j(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{\alpha_1}A_1 & \cdots & \frac{1}{\alpha_j}A_j & \cdots & \frac{1}{\alpha_n}A_n \end{pmatrix} \begin{pmatrix} \alpha_1 u_1(t) \\ \vdots \\ \alpha_j u_j(t) \\ \vdots \\ \alpha_n u_n(t) \end{pmatrix}. \tag{1.3}$$

D'après les égalités (1.2) et (1.3), on voit très nettement qu'à partir de la seule connaissance des observations $\mathbf{y}$, la matrice $\mathbf{A}$ ne pourra être estimée qu'à certaines ambiguïtés près. Le problème de séparation aveugle de sources pour le mélange linéaire statique possède deux indéterminations.

– Une indétermination d'échelle : les sources ne peuvent être retrouvées qu'à un facteur près (voir (1.3)). En d'autres termes, la puissance des sources ne peut être retrouvée.
– Une indétermination sur l'ordre des sources $u_1(.), \ldots, u_n(.)$: les sources ne peuvent être retrouvées qu'à une permutation près (voir (1.2)).

En tenant compte de ces deux indéterminations, on peut donc dire que le mélange est séparé dès lors que la matrice de transfert $\mathbf{H}$ obtenue est de la forme

$$\mathbf{H} = \mathbf{PD},$$

où $\mathbf{P} \in \mathfrak{P}$ est une matrice de permutation et $\mathbf{D} \in \mathfrak{D}$ est une matrice diagonale. Les estimations $\hat{\mathbf{u}}$ des sources vérifient alors

$$\hat{\mathbf{u}}(t) = \mathbf{PDu}(t),$$

ce qui équivaut à

$$\mathbf{B} = \mathbf{PDA}^{-1}.$$

Précisons d'ores et déjà que l'indétermination sur l'amplitude des sources peut être fixée sans perdre en généralité. Pour cela, il y a deux possibilités équivalentes :

– Soit on ajoute des conditions sur les sources afin de fixer les facteurs d'échelle α_j dans l'égalité (1.3). Dans ce cas, les sources sont supposées de puissance unitaire, c'est-à-dire

$$\forall i,\ E\left(u_1^2\right) = 1.$$

– Soit on ajoute des conditions sur la matrice de mélange $\mathbf{A}$ afin de fixer les facteurs d'échelles $(\frac{1}{\alpha j})$ dans l'égalité (1.3). Dans ce cas, les coefficients diagonaux de la matrice $\mathbf{A}$ sont fixés à 1, c'est-à-dire que la matrice $\mathbf{A}$ est de la forme

$$\mathbf{A} = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & & \vdots \\ \vdots & & \ddots & a_{n-1n} \\ a_{n1} & \dots & a_{nn-1} & 1 \end{pmatrix}.$$

1.3 Méthodes de séparation de sources

Nous allons présenter ici les principales techniques de séparation de sources pour les mélanges linéaires statiques. De nombreuses autres techniques sont présentées dans les livres de référence en séparation de sources [51] et [26]. Nous nous limiterons au cas où les signaux sources sont stationnaires. Différentes méthodes pour des signaux non stationnaires sont détaillées dans [75][70].

De plus, dans la mesure où les signaux rencontrés sont échantillonnés, les méthodes que nous allons exposer sont présentées pour des vecteurs aléatoires. La généralisation au cas de signaux aléatoires ne présente pas de difficulté.

Rappelons que la seule hypothèse sur les sources est l'indépendance statistique. Les techniques que nous allons présenter exploitent cette hypothèse de diverses façons. Nous les avons regroupées en deux classes, les méthodes utilisant exclusivement des statistiques d'ordre deux et celles utilisant des statistiques d'ordre supérieur.

Avant de présenter les principes de quelques méthodes de séparation de sources aux ordres supérieurs, nous détaillons, dans le paragraphe suivant, quelques mesures d'indépendance exploitées par la suite. Nous terminons ensuite cet état de l'art en présentant les méthodes de séparation à l'ordre deux.

1.3.1 Mesures d'indépendance

Rappelons qu'un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_n)$ est à composantes indépendantes si et seulement si

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=0}^{n} f_{X_i}(x_i)\,, \tag{1.4}$$

où $f_{\mathbf{X}}$ et f_{X_i} sont respectivement les densités de probabilité conjointe et marginales des X_i (dans la mesure où ces densités existent).
Nous allons voir comment nous pouvons comparer la densité conjointe au produit des densités marginales.

1.3.1.1 Information mutuelle et entropie

Un outil statistique très utile pour mesurer l'écart entre deux densités de probabilité est la divergence de Kullback-Leibner.

Définition 1.1 (Divergence de Kullback-Leibner)**.** Soient $f_{\mathbf{X}}$ et $g_{\mathbf{Y}}$ deux densités de probabilité multidimensionnelles. La divergence de Kullback-Leibner entre $f_{\mathbf{X}}$ et $g_{\mathbf{Y}}$, notée $\delta\left(f_{\mathbf{X}}|g_{\mathbf{Y}}\right)$, est définie par

$$\delta\left(f_{\mathbf{X}}|g_{\mathbf{Y}}\right) = \int_{-\infty}^{+\infty} f_{\mathbf{X}}\left(\mathbf{x}\right) \log \frac{f_{\mathbf{X}}\left(\mathbf{x}\right)}{g_{\mathbf{Y}}\left(\mathbf{x}\right)} d\mathbf{x}.$$

Pour deux vecteurs aléatoires $\mathbf{X}$ et $\mathbf{Y}$, on notera abusivement $\delta\left(\mathbf{X}|\mathbf{Y}\right)$ la divergence de Kullback-Leibner entre les densités de probabilité de $\mathbf{X}$ et $\mathbf{Y}$. Ainsi définie, la divergence de Kullback-Leibner est une grandeur non négative. De plus, elle s'annule si et seulement si les vecteurs $\mathbf{X}$ et $\mathbf{Y}$ ont la même distribution.

Remarque 1.2. *La divergence de Kullback-Leibner n'est pas une distance puisque*

$$\delta\left(\mathbf{X}|\mathbf{Y}\right) \neq \delta\left(\mathbf{Y}|\mathbf{X}\right).$$

La divergence de Kullback-Leibner entre la densité de probabilité conjointe et le produit des densités marginales d'un vecteur aléatoire quelconque coïncide, pour des densités $f_{\mathbf{X}}$ et $g_{\mathbf{Y}}$ particulières, avec une grandeur appelée information mutuelle en théorie de l'information [32].

Définition 1.3 (Information mutuelle)**.** Soit $\mathbf{X}$ un vecteur aléatoire de densité de probabilité conjointe $f_{\mathbf{X}}$ et de densités marginales f_{X_i}. L'information mutuelle du vecteur $\mathbf{X}$, notée $\mathcal{I}\left(\mathbf{X}\right)$, est définie par

$$\begin{aligned}\mathcal{I}\left(\mathbf{X}\right) &= \delta\left(f_{\mathbf{X}}\left(\mathbf{x}\right) | \prod_{i=0}^{n} f_{X_i}\left(x_i\right)\right) \\ &= \int_{-\infty}^{+\infty} f_{\mathbf{X}}\left(\mathbf{x}\right) \log \frac{f_{\mathbf{X}}\left(\mathbf{x}\right)}{\prod_{i=0}^{n} f_{X_i}\left(x_i\right)} dx_1 \ldots dx_n.\end{aligned}$$

L'information mutuelle est par définition une grandeur positive qui s'annule si $\mathbf{X}$ est à composantes indépendantes.

Remarque 1.4. *L'utilisation pratique de l'information mutuelle est toutefois complexe étant donné que les densités de probabilité des sources sont, dans la majorité des cas, inconnues.*

Une autre grandeur est également utilisée, il s'agit de l'entropie de Shannon. Elle permet de donner une nouvelle interprétation à l'information mutuelle.

Définition 1.5 (Entropie de Shannon)**.** Pour un vecteur aléatoire $\mathbf{X}$ de distribution $f_{\mathbf{X}}\left(\mathbf{x}\right)$, l'entropie de Shannon, notée $H\left(\mathbf{X}\right)$, est la quantité

$$H\left(\mathbf{X}\right) = -\int_{-\infty}^{+\infty} f_{\mathbf{X}}\left(\mathbf{x}\right) \log\left(f_{\mathbf{X}}\left(\mathbf{x}\right)\right) d\mathbf{x}.$$

Elle s'écrit également

$$H\left(\mathbf{X}\right) = -E\left(\log\left(f_{\mathbf{X}}\left(\mathbf{x}\right)\right)\right), \tag{1.5}$$

par définition de l'opérateur espérance.

Pour un vecteur aléatoire à n composantes $\mathbf{X} = \left(X_1, \ldots X_n\right)^{\mathrm{T}}$, l'information mutuelle et l'entropie de Shannon sont liées par la relation suivante :

$$\mathcal{I}\left(\mathbf{X}\right) = \sum_{i=1}^{n} H\left(X_i\right) - H\left(\mathbf{X}\right).$$

Ainsi, l'information mutuelle est une mesure de la quantité d'information que les composantes $X_1, \ldots, X_n$ d'un vecteur aléatoire donnent par rapport à l'information "globale" contenue dans le vecteur $\mathbf{X}$.

1.3.1.2 Moments et cumulants

Un grand nombre d'algorithmes de séparation de sources utilisant les statistiques d'ordres supérieurs exploitent indirectement l'indépendance en utilisant des relations entre les moments ou les cumulants. Dès lors, avant de présenter le principe de ces méthodes, nous introduisons quelques notions fondamentales qui y seront exploitées.

Définition 1.6 (Première fonction caractéristique)**.** Soit $\mathbf{X} = (X_1, \ldots, X_n)^T$ un vecteur aléatoire admettant une densité de probabilité $f_{\mathbf{X}}$. On appelle première fonction caractéristique de $\mathbf{X}$, notée $\phi_{\mathbf{X}}(\mathbf{u})$, la valeur moyenne de la fonction $h(\mathbf{X}) = \exp\left(j\mathbf{u}^T\mathbf{X}\right)$. Il s'agit d'une fonction complexe de la variable $\mathbf{u} \in \mathbb{R}^n$, définie par

$$\begin{aligned} \phi_{\mathbf{X}}(\mathbf{u}) &= E\left(\exp\left(j\mathbf{u}^T\mathbf{X}\right)\right) \\ &= \int f_{\mathbf{X}}(\mathbf{x}) \exp\left(j\mathbf{u}^T\mathbf{x}\right) d\mathbf{x}. \end{aligned} \tag{1.6}$$

On peut également définir une seconde fonction caractéristique :

Définition 1.7 (Seconde fonction caractéristique)**.** Soit $\mathbf{X} = (X_1, \ldots, X_n)^T$ un vecteur aléatoire admettant une densité de probabilité $f_{\mathbf{X}}$. On appelle seconde fonction caractéristique de $\mathbf{X}$, la fonction $\psi_{\mathbf{X}}(\mathbf{u})$ définie par

$$\psi_{\mathbf{X}}(\mathbf{u}) = \ln\left(\phi_{\mathbf{X}}(\mathbf{u})\right). \tag{1.7}$$

Un des principaux intérêts des fonctions (1.6) et (1.7) est qu'elles expriment de façon très simple les moments et les cumulants de $\mathbf{X}$, comme le montre la proposition suivante :

Proposition 1.8. *Le moment d'ordre k du vecteur aléatoire $\mathbf{X}$ est donné par*

$$E(X_1, \ldots, X_k) = (-j)^k \left.\frac{\partial^k \phi_{\mathbf{X}}(\mathbf{u})}{\partial u_1 \ldots \partial u_k}\right|_{\mathbf{u}=\mathbf{0}}.$$

Le cumulant[1] *d'ordre k de $\mathbf{X}$ est donné par*

$$Cum(X_1, \ldots, X_k) = (-j)^k \left.\frac{\partial^k \psi_{\mathbf{X}}(\mathbf{u})}{\partial x_1 \ldots \partial x_k}\right|_{\mathbf{u}=\mathbf{0}}.$$

Démonstration. Voir [77]. □

Remarque 1.9. *Il y a évidemment la même information dans les cumulants que dans les moments. A titre d'exemple, voici une relation entre les premiers moments et cumulants [63].*

$$E(X_1, X_2) - E(X_1)E(X_2) = Cum(X_1, X_2).$$

Toutefois, les cumulants sont davantage utilisés que les moments en séparation de sources. En effet, nous allons voir par la suite qu'ils ont la propriété d'être nuls pour des vecteurs indépendants.

[1]Les cumulants peuvent être définis de la façon suivante [63] :

$$Cum(x_1, \ldots, x_k) = E(u_1, \ldots, u_k) - E(g_1, \ldots, g_k)$$

où $\mathbf{g}$ est un vecteur aléatoire gaussien de mêmes moyenne et espérance que $\mathbf{x}$. Cette définition met en évidence le rôle des cumulants comme une mesure d'écart par rapport à la gaussianité.

Nous nous intéressons maintenant aux propriétés des cumulants d'un vecteur aléatoire à composantes indépendantes.

Proposition 1.10. *Les cumulants du vecteur aléatoire* $\mathbf{X}$ *sont tous nuls, si au moins une des composantes de* $\mathbf{X}$ *est indépendante des autres composantes, c'est-à-dire si*

$$\exists i \in \{1, \ldots, n\} \text{ tel que } \forall j \neq i,\ I_v\left(X_i, X_j\right).$$

Démonstration. Voir [63]. □

La proposition 1.10 permet d'affirmer que si tous les cumulants de tout ordre sont nuls, c'est-à-dire si

$$\forall k \in \mathbb{N},\ Cum\left(X_1, \ldots, X_k\right) = 0,$$

l'indépendance est assurée. Ce critère est exploité par certaines méthodes de séparation de sources.

Toutefois, dans la pratique, la manipulation des statistiques de tout ordre est impossible (car trop complexe) et on cherche alors l'ordre suffisant permettant d'obtenir la séparation. Dans la majorité des cas, les statistiques d'ordre quatre sont utilisées (voir paragraphe 1.3.2). En effet, les statististiques d'ordre deux conduisent à des méthodes exploitant la décorrélation (voir paragraphe 1.3.3). D'autre part, les statistiques d'ordre trois sont nulles pour des signaux de densité de probabilité symétrique. Cette restriction amène, dans le cas général, à exclure les statistiques d'ordre trois.

Avant de présenter les méthodes aux ordres supérieurs, nous introduisons la notion de contraste qui sera utilisée dans la quasi totalité de ces méthodes.

1.3.1.3 Fonction de contraste

La notion de contraste a été introduite par Donoho dans [37] puis reprise par Comon dans [28]. Un contraste est une fonction à valeur réelle qui ne dépend que de la densité de probabilité.

Définition 1.11 (Fonction de contraste)**.** Une fonction de contraste J est une application de $\mathcal{V}^n$ dans $\mathbb{R}$, qui, à un vecteur aléatoire $\mathbf{X}$, associe un réel et qui vérifie les deux conditions suivantes :
- $J\left(\mathbf{X}\right)$ dépend uniquement de la loi de probabilité de $\mathbf{X}$.
- Pour tout vecteur aléatoire $\mathbf{X}$ à composantes indépendantes et pour toute matrice $\mathbf{M} \in \mathbb{R}^{n \times n}$, on a l'inégalité suivante

$$J\left(\mathbf{MX}\right) \leq J\left(\mathbf{X}\right).$$

L'égalité $J\left(\mathbf{MX}\right) = J\left(\mathbf{X}\right)$ a lieu uniquement lorsque $\mathbf{M} = \mathbf{DP}$ où $\mathbf{D} \in \mathfrak{D}$ et $\mathbf{P} \in \mathfrak{P}$. Ainsi, le maximum de la fonction de contraste est atteint pour $\mathbf{X} \in \mathcal{I}_n$.

Grâce aux fonctions de contraste, le problème de séparation de sources peut se formaliser sous forme de problème d'optimisation. En effet, en considérant le problème de séparation de sources défini dans le paragraphe 1.1, on cherche la matrice de séparation $\mathbf{B}$ telle que

$$\mathbf{B} = \underset{\mathbf{M}}{\mathrm{Max}} J(\mathbf{My}).$$

Les méthodes présentées dans le paragraphe suivant considèrent différentes fonctions de contraste. Notons que ces méthodes basées sur des statistiques aux ordres supérieures rentrent dans le cadre plus général de l'analyse en composantes indépendantes [51].

1.3.2 Méthodes aux ordres supérieurs : Analyse en composantes indépendantes

Dans ce paragraphe, nous présentons les principales méthodes de séparation de sources basées sur des statistiques d'ordres supérieurs. Bien sûr , la séparation de signaux gaussiens est impossible dans la mesure où leurs statistiques d'ordres supérieurs à deux n'apportent aucune information.

Nous exposerons successivement les fonctions de contrastes basées sur une mesure directe (information mutuelle et infomax [21] [5][74]) puis une mesure indirecte (moments, cumulants [28][81][19][43]). Rappelons que les mesures directes utilisent de façon explicite les densités de probablité conjointes et marginales des signaux alors que les mesures indirectes ne mesurent qu'une indépendance partielle des composantes de vecteurs aléatoires, souvent limitée à l'ordre quatre.
Notons que les mesures indirectes imposent que l'on dispose d'un nombre important de réalisations afin d'estimer les statistiques des signaux. D'ailleurs, quelle que soit la mesure d'indépendance considérée, il est nécessaire d'estimer soit des densités de probabilité, soit des moments ou des cumulants à partir d'une seule observation des signaux de mélange. Ceci explique que l'on suppose que les signaux soient stationnaires et ergodiques afin de pouvoir remplacer les moyennes d'ensemble par des moyennes temporelles.

Notons que les différentes approches présentées peuvent être vue comme la minimisation de l'information mutuelle des signaux reconstruits $\hat{u}$. Cependant, chacune exploite différemment l'information mutuelle.

1.3.2.1 Approche du maximum de vraisemblance

Nous présentons dans ce paragraphe une solution du problème de la séparation de sources, au sens du maximum de vraisemblance. Le seul point délicat est le choix du modèle utilisé pour calculer la vraisemblance.

1.3.2.1.1 Critère de vraisemblance Il est indispensable de définir un modèle de distribution pour les sources et un modèle de mélange. Ici, s'agissant de la séparation d'un mélange linéaire instantané, le modèle est évidemment une matrice $\mathbf{A}$ telle que

$$\mathbf{y} = \mathbf{A}\mathbf{u}.$$

Soit $f(\mathbf{u})$ une distribution hypothétique des sources $\mathbf{u}$ vérifiant la condition d'indépendance (1.4), c'est-à-dire

$$f(\mathbf{u}) = \prod_{i=1}^{n} f_i(u_i).$$

La probabilité du mélange $\mathbf{y}$ (voir proposition A.75 page 155) est donc

$$\begin{aligned} f_{\mathbf{y}}(\mathbf{y}) &= |\det \mathbf{B}|\, f(\mathbf{u}) \\ &= |\det \mathbf{B}| \prod_{i=1}^{n} f_i(u_i), \end{aligned}$$

où $\mathbf{B} = \mathbf{A}^{-1}$ et f_i est la densité hypothétique marginale de la source u_i. En posant $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)^{\mathrm{T}}$, où les $\mathbf{b}_i$ sont les lignes de $\mathbf{B}$, on a

$$f_{\mathbf{y}}(\mathbf{y}) = |\det \mathbf{B}| \prod_{i=1}^{n} f_i\left(\mathbf{b}_i^{\mathrm{T}}\mathbf{y}\right).$$

Supposons que l'on dispose de T échantillons connus, notés $\mathbf{y}(1), \ldots, \mathbf{y}(T)$, des observations $\mathbf{y}$. Les observations étant indépendamment distribuées, la densité de probabilité $f_{\mathbf{y}}(\mathbf{y})$ s'écrit :

$$f_{\mathbf{y}}(\mathbf{y}) = |\det \mathbf{B}| \prod_{t=1}^{T} \prod_{i=1}^{n} f_i \left(\mathbf{b}_i^{\mathrm{T}} \mathbf{y}(t)\right).$$

La vraisemblance de $\mathbf{B}$ et f par rapport à $\mathbf{y}$, notée $L(\mathbf{B}, f)$, est donnée par

$$L(\mathbf{B}, f) = |\det \mathbf{B}| \prod_{t=1}^{T} \prod_{i=1}^{n} f_i \left(\mathbf{b}_i^{\mathrm{T}} \mathbf{y}(t)\right),$$

et le logarithme de la vraisemblance divisé par T (appelé log-vraissemblance normalisé) vaut :

$$\frac{1}{T} \log L(\mathbf{B}, f) = \log |\det \mathbf{B}| + \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^{\mathrm{T}} \mathbf{y}(t)\right)\right). \tag{1.8}$$

Pour séparer le mélange, il faut déterminer la valeur de $\mathbf{B}$ maximisant la log-vraisemblance (normalisée), c'est-à-dire trouver la matrice $\mathbf{B}$ telle que les observations $\mathbf{y}(1), \ldots, \mathbf{y}(T)$ soient les plus probables (pour le modèle f considéré). Les méthodes [5][43][76] exploitent ce contraste.

Remarque 1.12. *Soit $\mathbf{u}_0$ un vecteur aléatoire indépendant et identiquement distribué de densité $f_{\mathbf{u}_0}(\mathbf{u}_0) = f(\mathbf{u})$. Dans [15], il est démontré que*

$$\lim_{T \to +\infty} \frac{1}{T} \log L(\mathbf{B}, f) = -\delta(\mathbf{B}\mathbf{y}|\mathbf{u}_0) + \textit{constante}.$$

L'idée de la démonstration consiste à voir que

$$\begin{aligned} \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right) &= E\left(\sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^T \mathbf{y}\right)\right)\right) \\ &= \int_{-\infty}^{+\infty} f_{\mathbf{u}}(\mathbf{u}) \left(\sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right)\right) d\mathbf{u}. \end{aligned}$$

Comme $f_{\mathbf{u}}(\mathbf{u})$ est la vraie densité de probabilité des sources $\mathbf{u}$, en posant

$$\sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right) = \frac{\sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right)}{f_{\mathbf{u}}(\mathbf{u})} f_{\mathbf{u}}(\mathbf{u}),$$

on obtient :

$$\begin{aligned} \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right) &= -\int_{-\infty}^{+\infty} f_{\mathbf{u}}(\mathbf{u}) \log \left(\frac{f_{\mathbf{u}}(\mathbf{u})}{\sum_{i=1}^{n} \log \left(f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right)}\right) - \int_{-\infty}^{+\infty} f_{\mathbf{u}}(\mathbf{u}) \log \left(f_{\mathbf{u}}(\mathbf{u})\right) d\mathbf{u} \\ &= -\int_{-\infty}^{+\infty} f_{\mathbf{u}}(\mathbf{u}) \log \left(\frac{f_{\mathbf{u}}(\mathbf{u})}{\log \left(\prod_{i=1}^{n} f_i \left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right)}\right) - H(\mathbf{u}). \\ &= -\int_{-\infty}^{+\infty} f_{\mathbf{u}}(\mathbf{u}) \log \left(\frac{f_{\mathbf{u}}(\mathbf{u})}{f_{\mathbf{u}_0}(\mathbf{u}_0)}\right) - H(\mathbf{u}). \end{aligned}$$

où $H(\mathbf{u})$ est une constante (ne dépend pas de $\mathbf{B}$).

$$= -\delta(\mathbf{By}|\mathbf{u}_0) - H(\mathbf{u}).$$

Autrement dit, maximiser le critère (1.8) peut être vu comme minimiser la divergence de Kullback-Leibner de $\mathbf{By}$ et $\mathbf{u}_0$ (c'est-à-dire l'écart entre la densité hypothétique des sources f et la densité de $\hat{\mathbf{u}} = \mathbf{By}$).

1.3.2.1.2 Algorithme du maximum de vraisemblance Il peut sembler délicat de déterminer le modèle f pour les distributions des sources mais fort heureusement, dans le cas d'un mélange linéaire, le problème de séparation de sources s'avère robuste vis-à-vis d'un mauvais choix de f. En effet, dans [21][51], il est montré que même avec un modèle f erroné, maximiser la vraisemblance permet tout de même d'obtenir des estimations correctes des sources.

L'algorithme le plus simple exploitant le maximum de vraisemblance maximise la vraisemblance (1.8) par une méthode de gradient [5]. La dérivée de $\log L(\mathbf{B}, f)$ par rapport à $\mathbf{B}$ est donné par

$$\frac{\partial L}{\partial \mathbf{B}} = \left(\mathbf{B}^{\mathrm{T}}\right)^{-1} + E\left(\mathbf{g}(\mathbf{By})\,\mathbf{y}^{\mathrm{T}}\right), \tag{1.9}$$

où $\mathbf{g}(\mathbf{u}) = (g_1(u_1), \ldots, g_n(u_n))$ est une fonction vectorielle et les fonctions g_i, $i \in \{1, \ldots, n\}$, définies par

$$g_i = (\log f_i)' = \frac{f_i'}{f_i},$$

sont appelées *fonctions scores* du signal u_i.

Ainsi, en fixant une distribution connue, le critère $\log L(\mathbf{B}, f)$ peut être maximisé par une méthode de gradient en utilisant (1.9).

Cependant, cet algorithme nécessite l'inversion d'une matrice $\mathbf{B}$ à chaque itération et sa convergence est donc lente. Elle peut être améliorée en utilisant une méthode de gradient naturel (ou relatif) qui simplifie la maximisation. Le principe du gradient naturel est décrit dans [17]. On obtient :

$$\nabla\mathbf{B} = \left(\mathbf{Id}_n + E\left(\mathbf{g}(\hat{\mathbf{u}})\,\hat{\mathbf{u}}^{\mathrm{T}}\right)\right)\mathbf{B}.$$

Remarque 1.13. *Ce nouvel algorithme, basé sur le gradient naturel du maximum de vraisemblance peut être interprété comme une **décorrélation non-linéaire** dont le principe sera davantage détaillé dans le paragraphe 1.3.2.3. En effet, il suffit de voir que la condition d'annulation du gradient relatif s'écrit*

$$E\{\mathbf{g}(\hat{\mathbf{u}})\hat{\mathbf{u}}^T\} = \mathbf{Id}_n,$$

et, sachant que les fonctions scores $\mathbf{g}$ sont non-linéaires, on voit que le critère du maximum de vraisemblance revient à annuler $E\left(g_i(\hat{u}_i)\,\hat{u}_j^T\right)$, pour tout $i \neq j$, c'est-à-dire à réaliser une décorrélation non-linéaire des observations.

1.3.2.2 Approche par information mutuelle

Nous avons vu précédemment que l'information mutuelle est une mesure de dépendance qui, par conséquent, peut être utilisée pour estimer la matrice $\mathbf{B}$ telle que

$$\hat{\mathbf{u}} = \mathbf{By}. \tag{1.10}$$

Le problème de séparation de sources est alors ramené à la minimisation de l'information mutuelle

$$\mathcal{I}(\hat{\mathbf{u}}) = \sum_{i=1}^{n} H(\hat{u}_i) - H(\hat{\mathbf{u}})$$

des sources estimées.

A la différence du maximum de vraisemblance, ici, aucun modèle de distribution des sources n'est présupposé. Toutefois, notons que dans la pratique, l'utilisation de cette méthode est très complexe puisque le calcul de l'information mutuelle nécessite d'estimer la densité de probabilité conjointe $f_{\hat{\mathbf{u}}}$ des signaux. En général, on préfère exploiter l'information mutuelle sous une forme différente. En effet, d'après l'égalité (1.10) et quelques résultats élémentaires[2] sur l'entropie, on a

$$\mathcal{I}(\hat{\mathbf{u}}) \quad = \quad \sum_{i=1}^{n} H(\hat{u}_i) - H(\mathbf{y}) - \log|\det \mathbf{B}|. \tag{1.11}$$

Et, en remarquant que $H(\mathbf{y})$ est constant, afin d'éviter l'estimation de la densité multivariable de $\hat{\mathbf{u}}$ (et celle de $\mathbf{y}$), on préfère souvent minimiser le contraste suivant

$$J(\mathbf{B}) \quad = \quad \sum_{i=1}^{n} H(\hat{u}_i) - \log|\det \mathbf{B}|, \tag{1.12}$$

plutôt que l'information mutuelle $\mathcal{I}(\hat{\mathbf{u}})$.

A noter qu'il existe également d'autres méthodes exploitant différemment l'information mutuelle [5][52]. Ainsi, dans la plupart des cas, les méthodes exploitant l'information mutuelle aboutissent à un critère qui permet de choisir la matrice de séparation et n'exploitent jamais directement l'information mutuelle [27].

Remarque 1.14. *Dans [21], il est montré que le critère du maximum de vraisemblance et le critère de minimum d'information mutuelle sont très liés. Afin de s'en convaincre, voici l'idée générale. Pour montrer la ressemblance entre ces deux critères, nous considérons la valeur moyenne (par rapport à T) de la log-vraisemblance (1.8) :*

$$E(\log L(\mathbf{B}, f)) = \log|\det \mathbf{B}| + \sum_{i=1}^{n} E\left(\log\left(f_i\left(\mathbf{b}_i^T \mathbf{y}(t)\right)\right)\right).$$

[2]Soient $\mathbf{x} \in \mathcal{V}^n$, considérons la transformation $\mathbf{f}$ inversible telle que $\mathbf{y} = \mathbf{f}(\mathbf{x})$. On a

$$f_{\mathbf{y}}(\mathbf{y}) = f_{\mathbf{x}}(\mathbf{f}^{-1}(\mathbf{y}))|\det J_f(\mathbf{f}^{-1}(\mathbf{y}))|^{-1}.$$

D'autre part, en notant $H(\mathbf{y}) = -E(\log(p_{\mathbf{y}}(\mathbf{y})))$, on a

$$\begin{aligned} E(\log(p_{\mathbf{y}}(\mathbf{y}))) &= E(\log(p_{\mathbf{x}}(\mathbf{f}^{-1}(\mathbf{y}))|\det J_f(\mathbf{f}^{-1}(\mathbf{y}))|^{-1})) \\ &= E(\log(p_{\mathbf{x}}(\mathbf{x})|\det J_f(\mathbf{x})|^{-1})) \\ &= E(\log(p_{\mathbf{x}}(\mathbf{x}))) - E(\log(|\det J_f(\mathbf{x})|)), \end{aligned}$$

et donc, on obtient

$$H(\mathbf{y}) = H(\mathbf{x}) + E(\log(|\det J_f(\mathbf{x})|)).$$

Dans le cas particulier d'une transformation linéaire $\mathbf{y} = \mathbf{M}\mathbf{x}$, on obtient

$$H(\mathbf{y}) = H(\mathbf{x}) + E(\log(|\det \mathbf{M}|)).$$

D'autre part, d'après la définition 1.5 et quelques résultats élémentaires sur l'entropie, on a

$$\mathcal{I}(\hat{\mathbf{u}}) = \sum_{i=1}^{n} H(\hat{u}_i) - H(\mathbf{y}) - \log|\det \mathbf{B}|.$$

Et, d'après (1.5),

$$\mathcal{I}(\hat{\mathbf{u}}) = -\sum_{i=1}^{n} E\left(\log f_{\hat{u}_i}(\hat{u}_i)\right) - H(\mathbf{y}) - \log|\det \mathbf{B}|.$$

Ainsi, on peut donc voir que $E(\log L(\mathbf{B}, f))$ est une approximation (au signe près et à la constante $H(\mathbf{y})$ près) de $\mathcal{I}(\hat{\mathbf{u}})$ dans la mesure où les modèles hypothétiques des sources f_i sont des approximations des véritables densités des sources $f_{\hat{u}_i}$.

1.3.2.3 Approche par moments ou cumulants

Ces méthodes, exploitant de manière explicite les moments d'ordre supérieur, sont basées sur le résultat de la proposition 1.10 et sont appelées *méthodes de déconvolution non linéaire*.

Elles utilisent des fonctions non linéaires des signaux pour exploiter l'indépendance statistique au delà de l'ordre deux : c'est le cas par exemple du premier algorithme de séparation de sources, présenté dans [49] et [59], qui minimise le critère

$$E\left(f(u_i)\, g(u_j)\right),$$

où f et g sont des fonctions impaires ($f(u) = u^3$ et $g(u) = u$). Les performances et la stabilité de l'algorithme de Jutten et Hérault ont été étudiées notamment par Sorouchyari dans [84]. Nous reviendrons sur cette étude dans le chapitre 3 (voir la section 3.1 page 53) afin de concrétiser les techniques présentées. Citons également l'agorithme JADE basé sur la minimisation du critère

$$\sum_{ijkl \neq iikl} Cum\left(u_i, u_j, u_k, u_l\right)^2.$$

L'avantage de ce critère est qu'il peut s'exprimer sous forme d'un problème de diagonalisation conjointe (voir annexe F), pour lequel des algorithmes efficaces de résolution existent.

Il est important de noter que les méthodes de déconvolution non linéaire présentées ici peuvent être vue comme dérivées de la minimisation de l'information mutuelle avec des approximations particulières des densités de probabilité. Par exemple en utilisant un développement de Gram-Charlier à l'ordre quatre des densités [2] ou un developpement de Edgeworth [28].

1.3.3 Décorrélation ou méthode à l'ordre 2

Nous avons vu que les méthodes aux ordres supérieurs sont possibles uniquement si au plus une source est gaussienne. En effet, pour des signaux gaussien, les statistiques d'ordres supérieurs (à deux) n'apportent aucune information.

Maintenant, nous allons nous intéresser aux méthodes de séparation à l'ordre deux, basées sur les matrices d'autocorrélation. Rappelons que pour un signal aléatoire $\mathbf{u}(.) \in \mathcal{S}$, la matrice d'autocorrélation $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}$ est définie par :

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau) = E\left(\mathbf{u}(t)\,\mathbf{u}(t-\tau)^{\mathrm{T}}\right) - E(\mathbf{u}(t))\, E(\mathbf{u}(t))^{\mathrm{T}}.$$

A défaut d'assurer l'indépendance statistique des sources estimées, les méthodes à l'ordre deux se limitent à leur décorrélation. Des hypothèses supplémentaires de coloration des sources sont alors indispensables pour la séparation.

Nous allons détailler le principe de séparation. Il se déroule en deux étapes distinctes. Tout d'abord une étape de blanchiment spatial, puis l'estimation d'une matrice orthogonale. Pour davantage de précisions, le lecteur pourra consulter [6].

1.3.3.1 Première étape : le blanchiment spatial

Le blanchiment spatial consiste à trouver une matrice $\mathbf{W}$ qui, appliquée aux observations $\mathbf{y}$, fournit des signaux $\mathbf{z} = \mathbf{W}\mathbf{y}$ tels que

$$\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}\left(0\right) = \mathbf{Id}_n,$$

c'est-à-dire

$$\mathbf{W}\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(0\right)\mathbf{W}^{\mathrm{T}} = \mathbf{Id}_n. \tag{1.13}$$

Remarque 1.15. *Notons que la matrice $\mathbf{W}$ n'est pas unique. En effet, en multipliant chaque membre de l'égalité par une matrice orthogonale quelconque $\mathbf{V}$, l'égalité (1.13) est conservée. On a*

$$\mathbf{V}\mathbf{W}\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(0\right)\mathbf{W}^{T}\mathbf{V}^{T} = \mathbf{V}\mathbf{Id_n}\mathbf{V}^{T} = \mathbf{Id}_n.$$

La construction d'une matrice $\mathbf{W}$ vérifiant (1.13) s'obtient de la manière suivante.
La matrice $\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(0\right)$ étant symétrique et non négative (voir proposition B.11 page 164), on a la décomposition en éléments propres :

$$\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(0\right) = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}},$$

où $\mathbf{U}$ est une matrice orthogonale composée des vecteurs propres de $\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}(= \mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(0\right))$ et $\mathbf{D}$ est une matrice diagonale composée des valeurs propres de $\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}$. Vérifions maintenant que la matrice $\mathbf{W}$ définie par $\mathbf{W} = \mathbf{D}^{-\frac{1}{2}}\mathbf{U}^{\mathrm{T}}$ est une matrice de blanchiment (c'est-à-dire $\mathbf{W}$ vérifie (1.13)). En effet, on a

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}} &= \mathbf{W}\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\mathbf{W}^{\mathrm{T}} \\
&= \mathbf{D}^{-\frac{1}{2}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{D}^{-\frac{1}{2}} \\
&= \mathbf{Id}_n.
\end{aligned}$$

Après blanchiment, on obtient des signaux $\mathbf{z}$ décorrélés et normés. Cependant, la séparation n'est toujours pas réalisée. Pour s'en convaincre, considérons la décomposition en valeur singulière de $\mathbf{A}$,

$$\mathbf{A} = \mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{Q}^{\mathrm{T}},$$

avec $\mathbf{D}$ une matrice diagonale composée des vecteurs propres de $\mathbf{A}$, $\mathbf{U}$ et $\mathbf{Q}$ des matrices orthogonales. On a

$$\begin{aligned}
\mathbf{z} &= \mathbf{W}\mathbf{A}\mathbf{u} \\
&= \mathbf{D}^{-\frac{1}{2}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{Q}^{\mathrm{T}}\mathbf{u} \\
&= \mathbf{Q}^{\mathrm{T}}\mathbf{u}.
\end{aligned}$$

On voit que $\mathbf{z}$ est un mélange par une matrice orthogonale $\left(\mathbf{Q}^{\mathrm{T}}\right)$ des sources $\mathbf{u}$. Or, une matrice orthogonale n'étant pas, en général, une matrice de permutation, il reste à estimer cette matrice orthogonale $\mathbf{Q}^{\mathrm{T}}$ pour séparer le mélange.

Remarque 1.16. *Une autre façon de voir que la séparation n'est pas réalisée consiste à considérer le problème de séparation de sources comme un problème d'estimation des $n^2 - n$ coefficients de la matrice $\mathbf{A}$ (on rappelle que ses coefficients diagonaux sont fixés à 1). La contrainte de blanchiment des observations $\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(0) = \mathbf{Id}_n$ nous donne $\frac{n^2-n}{2}$ équations, ce qui est insuffisant comme nous l'avons vu.*

Le blanchiment spatial est une étape préliminaire de nombreuses méthodes de séparation de sources [6] [28].

Remarque 1.17. *Le blanchiment spatial et l'analyse en composantes principales (voir [60]) sont deux problèmes identiques. En effet, le blanchiment correspond à une analyse en composantes principales des observations $\mathbf{y}$ et les composantes principales sont les signaux $\mathbf{z}$.*

1.3.3.2 Seconde étape : l'estimation de Q

Nous allons ici détailler la méthode d'estimation de la matrice orthonormale $\mathbf{Q}$ reconstruisant une estimation de sources primaires $\mathbf{u}$. Cette méthode, présentée dans l'algorithme SOBI [6], utilise les autocorrélations des observations pour des retards non nuls. Ainsi, la fonction d'autocorrélation des observations blanchies est

$$\begin{aligned}\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(\tau) &= E\left(\mathbf{z}(t)\,\mathbf{z}(t-\tau)^{\mathrm{T}}\right)\\ &= \mathbf{WA}E\left(\mathbf{u}(t)\,\mathbf{u}(t-\tau)^{\mathrm{T}}\right)\mathbf{A}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\\ &= \mathbf{Q}^{\mathrm{T}}\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)\,\mathbf{Q}. \end{aligned} \tag{1.14}$$

Sachant que les sources sont indépendantes, la matrice $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ est une matrice diagonale pour tout retard τ (voir la proposition B.46 page 181). Ainsi, d'après l'égalité (1.14), il est clair que la matrice $\mathbf{Q}$ est composée des vecteurs propres de $\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(\tau)$ (à τ fixé). En conséquence, la matrice $\mathbf{Q}$ peut être identifiée en diagonalisant $\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(\tau)$.

Remarque 1.18. *Si deux termes diagonaux sont identiques dans $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$, alors deux colonnes de $\mathbf{Q}$ ne peuvent pas être estimées. En effet, les deux vecteurs colonnes associés à cette valeur propre de multiplicité deux sont choisis orthogonaux et normés dans un espace vectoriel de dimension deux. Il y a bien entendu une infinité de choix possibles pour ces deux vecteurs, dont une seule correspond à la matrice $\mathbf{Q}$ recherchée.*

En conséquence de la remarque 1.18, le choix du retard doit se faire de sorte que les termes diagonaux de $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ soient tous différents. Ainsi, une condition nécessaire de succès pour cette méthode est l'existence d'un retard $\tau \neq 0$ tel que les fonctions d'autocorrélation $\Gamma_{u_i,u_i}(\tau)$ des sources en τ soient non nulles et distinctes, c'est-à-dire $\forall i, j,\, i \neq j$,

$$\Gamma_{u_i,u_i}(\tau) \neq \Gamma_{u_j,u_j}(\tau).$$

Dans la pratique, on utilise plusieurs retards $\tau_i, i \in I \subset \mathbb{N}$, non nuls et on diagonalise simultanément (ou conjointement) les matrices $\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(\tau_i)$, c'est-à-dire qu'on recherche la matrice $\hat{\mathbf{Q}}$ telle que

$$\forall \tau_i,\ \hat{\mathbf{Q}}\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(\tau_i)\,\hat{\mathbf{Q}}^{\mathrm{T}} \in \mathfrak{D}.$$

L'annexe F détaille les fondements des méthodes de diagonalisation conjointe d'un ensemble de matrices. C'est un problème d'optimisation qui consiste à minimiser la somme quadratique des éléments hors diagonaux des matrices $\hat{\mathbf{Q}}\mathbf{\Gamma}_{\mathbf{z},\mathbf{z}}(\tau_i)\,\hat{\mathbf{Q}}^{\mathrm{T}}$ [18]. Un algorithme de diagonalisation conjointe est présenté et illustré dans cette annexe.

La séparation des sources (voir figure 1.2) est alors obtenue en posant

$$\mathbf{B} = \hat{\mathbf{Q}}\mathbf{W},$$

où $\hat{\mathbf{Q}}$ est l'estimée de $\mathbf{Q}$.

$$\mathbf{u} \longrightarrow \boxed{\mathbf{A}} \longrightarrow \mathbf{y} \longrightarrow \boxed{\mathbf{W}} \longrightarrow \mathbf{z} \longrightarrow \boxed{\mathbf{Q}} \longrightarrow \hat{\mathbf{u}}$$

Figure 1.2 – Système de séparation à l'ordre deux.

1.3.4 Méthode géométrique

Citons, avant de conclure, qu'il existe également des méthodes de séparation basées sur des critères géométriques. Dans [79][78] et [3] sont proposées des idées originales de séparation basées sur la représentation géométrique des mélanges. Cependant, ces méthodes restent souvent limitées à un faible nombre de sources.

1.4 Bilan des méthodes de séparation du mélange linéaire instantané

Le seul cas où l'ensemble des méthodes précédemment citées est inefficace est celui où les sources sont de distribution gaussienne et possèdent des autocorrélations identiques. Par exemple, la séparation de signaux gaussiens blancs est impossible puisque $\Gamma_{u_i,u_i}(\tau) = 0$ (pour $\tau \neq 0$). La raison de cette inefficacité est simple :

Pour deux sources gaussiennes circulaires[3] u_1 et u_2 décorrélées (et donc indépendantes), tout produit de $\mathbf{u}$ par une matrice orthogonale (ou unitaire) $\mathbf{Q}$ fournit des signaux $\mathbf{Qu}$ de densité de probabilité identique à $\mathbf{u}$ (voir la proposition B.65 page 187), c'est-à-dire que le signal $\mathbf{Qu}$ est également gaussien circulaire et décorrélé.
Il est donc impossible de les séparer, ni avec une méthode d'ordre deux, ni avec aucune autre méthode (puisque les signaux gaussien sont entièrement caractérisés par leurs statistiques d'ordre un et deux).

En conclusion, on peut dire que l'ensemble des mélanges linéaires instantanés, dont au moins deux sources sont gaussiennes circulaires, est non séparable.
Pour pouvoir séparer, il faut qu'une des deux conditions suivantes soit vérifiée :

- Au plus une source est gaussienne. Les méthodes exploitant l'indépendance statistique à l'aide de statistiques d'ordre supérieur sont alors à utiliser (voir paragraphe 1.3.2)
- Les autocorrélations des sources sont différentes. L'utilisation des méthodes exploitant l'indépendance statistique à l'aide de statistiques d'ordre deux est possible (voir paragraphe 1.3.3).

Bien sûr, si les deux conditions précédentes sont vérifiées alors toutes les méthodes présentées sont envisageables.

[3]c'est-à-dire de fonctions d'autocorrélation identiques.

1.5 Conclusion

Comme nous l'avons montré tout au long de ce chapitre présentant les principales techniques de séparation de sources, le problème de séparation du mélange linéaire instantané consiste, dans la plupart des cas, à optimiser une fonction objectif (ou fonction de contraste) non linéaire de manière à assurer l'indépendance des entrées. Il est bien entendu que la séparation de mélanges plus complexes constitue également un problème d'optimisation [61][31]. Ainsi, dans la majeure partie des cas, les méthodes de séparation de sources exploitent des techniques de descente de gradient afin d'optimiser un critère non convexe, souvent mal conditionné et pouvant contenir une infinité de minimiseurs. Il est dès lors nécessaire de réaliser une étude préalable afin de s'assurer de la convergence de la méthode vers la solution.

Nous avons choisi dans le cadre de cette thèse, d'utiliser des techniques d'optimisation différentes, basées sur le calcul ensembliste, qui sont plus fiables, ne sont pas piégées par des minima locaux et sont peu sensibles à de mauvais conditionnements. Notons qu'à notre connaissance, il s'agit de la première fois où le calcul ensembliste a été employé sur des critères statistiques. Nous verrons que ces techniques présentent quelques avantages non négligeables (mais hélas aussi quelques inconvénients). Notre démarche est d'autant plus justifiée que nous nous intéressons, principalement dans le chapitre 4, à des systèmes de mélanges non linéaires qui nécessitent une méthode de résolution efficace.

CHAPITRE 2

Analyse par intervalles

On ne peut pas dater précisément la naissance de l'arithmétique par intervalles. Certains l'attribuent, en 1966, à Ramon Moore dans [72], pour d'autres on la trouve dès 1958 dans [85]. Jusqu'à la fin des années 1970, le calcul scientifique ne dispose pas encore d'énormément de puissance, les débuts de l'arithmétique par intervalles sont timides.

Les progrès de l'arithmétique en matière d'arrondis dans les années 70 ont favorisé le développement de l'arithmétique par intervalles. Grâce aux intervalles, on obtient un résultat garanti (au lieu de rechercher une solution d'un problème, on recherche un intervalle contenant la solution ou l'ensemble des solutions d'un problème) mais contrairement à ce qu'on aurait pû penser, il ne suffit pas pour cela de remplacer un arrondi par un intervalle. Ainsi, quelques déceptions, notamment concernant la finesse des intervalles solutions ont ralenti l'essor de l'arithmétique par intervalles, qui eut parfois du mal à convaincre les utilisateurs potentiels de son intérêt.

Toutefois, à partir du début des années 1990, on trouve de nouvelles applications de l'arithmétique par intervalles. Elle devient un puissant outil déterministe pour déterminer l'optimum global d'une fonction continue, l'image directe ou inverse d'un ensemble par une fonction, les zéros d'une fonction, etc.

Dans ce chapitre, après une présentation de l'arithmétique des intervalles, nous étudions quelques exemples d'applications du calcul ensembliste. Dans un premier temps, nous nous intéressons à la résolution de systèmes d'équations non linéaires puis, nous présentons une méthode d'optimisation globale ensembliste de systèmes non linéaires.

2.1 Introduction à l'arithmétique par intervalles

En arithmétique par intervalles, on ne manipule plus des nombres, qui approchent plus ou moins fidèlement une valeur, mais des intervalles contenant cette valeur.
Par exemple, on peut tenir compte d'une erreur de mesure en remplaçant une valeur mesurée x avec une incertitude ε par l'intervalle $[x-\varepsilon, x+\varepsilon]$. Ainsi, le nombre π peut être remplacé par $[3.14, 3.15]$.
Si l'on désire obtenir un résultat valide pour tout un ensemble de valeurs, on utilise un intervalle contenant toutes ces valeurs.
L'objectif de l'arithmétique par intervalles est de fournir des résultats qui contiennent à coup sûr la valeur ou l'ensemble de valeurs cherché ; on parle de *résultats garantis*.

Dans cette première partie, nous commençons par des rappels élémentaires sur la théorie des ensembles avant de détailler la notion d'intervalles et d'introduire le vocabulaire qui sera utilisé par la suite. Puis, nous abordons l'arithmétique des intervalles en insistant particulièrement sur la notion fondamentale de fonction d'inclusion qui sera illustrée par de nombreux exemples.

2.1.1 Théorie des ensembles

2.1.1.1 Vocabulaire

Soient deux ensembles $\mathbb{X}$ et $\mathbb{Y}$, on rappelle que

$$\begin{aligned}\mathbb{X}\cap\mathbb{Y} &\triangleq \{x|x\in\mathbb{X} \text{ et } x\in\mathbb{Y}\} \text{ (intersection)}\\ \mathbb{X}\cup\mathbb{Y} &\triangleq \{x|x\in\mathbb{X} \text{ ou } x\in\mathbb{Y}\} \text{ (union)}\\ \mathbb{X}\backslash\mathbb{Y} &\triangleq \{x|x\in\mathbb{X} \text{ et } x\notin\mathbb{Y}\} \text{ (différence)}\\ \mathbb{X}\times\mathbb{Y} &\triangleq \{(x,y)\,|x\in\mathbb{X}, y\in\mathbb{Y}\} \text{ (produit cartésien)}\end{aligned}$$

Ces quatre opérations ensemblistes sont illustrées par la figure 2.1.

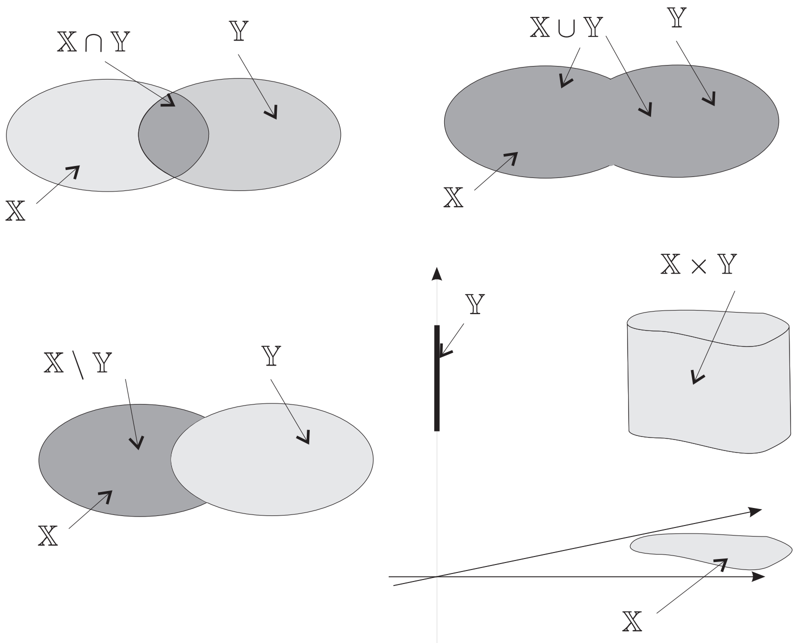


Figure 2.1 – Opérations $\cap$, $\cup$, $\backslash$ et $\times$ sur les ensembles.

L'inclusion de l'ensemble $\mathbb{X}$ dans l'ensemble $\mathbb{Y}$ est définie par

$$\mathbb{X}\subset\mathbb{Y}\Leftrightarrow\{x\in\mathbb{Y}|x\in\mathbb{X}\},$$

l'égalité des ensembles $\mathbb{X}$ et $\mathbb{Y}$ est définie par

$$\mathbb{X}=\mathbb{Y}\Leftrightarrow(\mathbb{X}\subset\mathbb{Y} \text{ et } \mathbb{Y}\subset\mathbb{X}).$$

2.1.1.2 Opérations supplémentaires

Soient $\mathbb{X}$ et $\mathbb{Y}$ deux ensembles et $f:\mathbb{X}\to\mathbb{Y}$ une fonction. Si $\mathbb{X}_1\subset\mathbb{X}$, *l'image directe* de $\mathbb{X}_1$ par f est

$$f(\mathbb{X}_1)\triangleq\{f(x)\,|x\in\mathbb{X}_1\}.$$

Si $\mathbb{Y}_1 \in \mathbb{Y}$, *l'image réciproque* de $\mathbb{Y}_1$ par f est

$$f^{-1}(\mathbb{Y}_1) \triangleq \{x \in \mathbb{X} | f(x) \in \mathbb{Y}_1\}.$$

On a donc,

$$f(\emptyset) = f^{-1}(\emptyset) = \emptyset.$$

De plus, on montre facilement que, si $\mathbb{X}_1, \mathbb{X}_2 \in \mathbb{X}$ et $\mathbb{Y}_1, \mathbb{Y}_2 \in \mathbb{Y}$, alors

$$\begin{array}{lcl} f(\mathbb{X}_1 \cap \mathbb{X}_2) & \subset & f(\mathbb{X}_1) \cap f(\mathbb{X}_2), \\ f(\mathbb{X}_1 \cup \mathbb{X}_2) & = & f(\mathbb{X}_1) \cup f(\mathbb{X}_2), \\ f^{-1}(\mathbb{Y}_1 \cap \mathbb{Y}_2) & = & f^{-1}(\mathbb{Y}_1) \cap f^{-1}(\mathbb{Y}_2), \\ f^{-1}(\mathbb{Y}_1 \cup \mathbb{Y}_2) & = & f^{-1}(\mathbb{Y}_1) \cup f^{-1}(\mathbb{Y}_2), \\ f\left(f^{-1}(\mathbb{Y})\right) & \subset & \mathbb{Y}. \end{array}$$

Après ces précisions élémentaires, présentons la théorie des intervalles.

2.1.2 Les intervalles

2.1.2.1 Vocabulaire

Définition 2.1 (Intervalle)**.** On appelle intervalle réel, noté $[x]$, un ensemble connexe fermé de $\mathbb{R}$. L'ensemble des intervalles de $\mathbb{R}$ est noté $\mathbb{IR}$.

Exemple 2.2. $[1,6], \{3\},]-\infty, 2], \mathbb{R}, \emptyset$ *sont des intervalles alors que* $]2,3[, [2,1]$ *et* $[-2,-1] \cup [1,6]$ *ne le sont pas.*

Définition 2.3 (Eléments caractéristiques d'un intervalle)**.** Soit $[x]$ un intervalle.
- La borne inférieure de $[x]$, notée x^-, est définie par

$$x^- = \inf\{x | x \in [x]\}.$$

- La borne supérieure de $[x]$, notée x^+, est définie par

$$x^+ = \sup\{x | x \in [x]\}.$$

- La taille de l' intervalle $[x]$, notée $w([x])$, est définie par

$$w([x]) = x^+ - x^-.$$

Notons que la taille d'un intervalle n'est rien d'autre que sa mesure de Lebesgue.
- Le centre de l'intervalle $[x]$, noté $m([x])$ ou encore x_c, est défini par

$$m([x]) = x_c = \frac{x^+ + x^-}{2}.$$

La valeur absolue de l'intervalle $[x]$, notée $|[x]|$, est définie par

$$|[x]| = \max\left(\left\{\left|x^+\right|, \left|x^-\right|\right\}\right).$$

2.1.2.2 Les opérations ensemblistes sur $\mathbb{IR}$

Un intervalle étant par définition un ensemble, nous pouvons lui associer les opérateurs ensemblistes classiques. Commençons par l'intersection de deux intervalles :

Définition 2.4 (Intersection)**.** Soient $[x]$ et $[y]$ deux intervalles, l'intersection de $[x]$ et $[y]$, notée $[x] \cap [y]$, est définie par

$$[x] \cap [y] = \begin{cases} \emptyset \text{ si } x^+ < y^- \text{ ou } x^- > y^+, \\ [\max(x^-, y^-), \min(x^+, y^+)] \text{ sinon.} \end{cases}$$

Remarque 2.5. *Notons qu'ainsi définie, l'intersection $[x] \cap [y]$ est un intervalle. Par contre, si on considère l'union de deux intervalles, cette dernière n'est, en général, pas un intervalle car l'union de deux intervalles disjoints n'est pas connexe, d'où l'intérêt d'introduire la notion d'intervalle enveloppe.*

Définition 2.6 (Intervalle enveloppe)**.** Soit $\mathbb{X}$ un sous ensemble de $\mathbb{R}$. L'intervalle enveloppe de $\mathbb{X}$, noté $[\mathbb{X}]$, est le plus petit intervalle contenant $\mathbb{X}$.

Exemple 2.7. $[[1,3] \cup [5,8]] = [1,8]$.

Définition 2.8 (Union au sens des intervalles)**.** Soient $[x]$ et $[y]$ deux intervalles, l'union, au sens des intervalles, de $[x]$ et $[y]$, notée $[x] \sqcup [y]$, est définie par

$$[x] \sqcup [y] = [[x] \cup [y]] .$$

Ainsi définie, l'union de deux intervalles est bien un intervalle.

Définition 2.9 (Inclusion)**.** Les tests d'inclusion stricte ou large d'un intervalle $[x]$ dans un intervalle $[y]$ s'écrivent respectivement

$$[x] \subset [y] \Leftrightarrow \left(y^- < x^- \text{ et } x^+ < y^+\right),$$
$$[x] \subseteq [y] \Leftrightarrow \left(y^- \leq x^- \text{ et } x^+ \leq y^+\right).$$

2.1.2.3 Les vecteurs intervalles réels

Définition 2.10 (Vecteur intervalle)**.** On appelle vecteur intervalle, ou encore boîte ou pavé, un sous ensemble de $\mathbb{R}^n$, noté $[\mathbf{x}]$, défini comme le produit cartésien de n intervalles, c'est-à-dire un vecteur dont chaque composante est un intervalle :

$$[\mathbf{x}] = [x_1] \times \cdots \times [x_n] = ([x_1], \ldots, [x_n]) .$$

L'ensemble de toutes les boites de $\mathbb{R}^n$ est noté $\mathbb{IR}^n$.

La Figure 2.2 illustre le fait qu'un vecteur intervalle est un ensemble parallélépipédique, nous utiliserons indifféremment les termes de vecteur intervalle, de pavé ou de boîte, voir même d'intervalle.

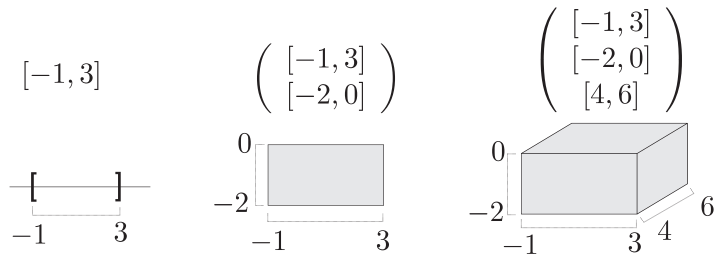


Figure 2.2 – Exemples d'éléments de $\mathbb{IR}$, $\mathbb{IR}^2$ et $\mathbb{IR}^3$.

Un grand nombre des notions que nous avons introduites pour un intervalle peuvent être reprises pour un vecteur intervalle.

Définition 2.11 (Eléments caractéristiques d'un vecteur intervalle)**.** Soit $[\mathbf{x}] \in \mathbb{IR}^n$ un vecteur intervalle.
- On appelle centre de $[\mathbf{x}]$ le vecteur, noté $m([\mathbf{x}])$ ou $\mathbf{x}_c$, défini par

$$m([\mathbf{x}]) = \mathbf{x}_c = (m([x_1]), \ldots, m([x_n])) \in \mathbb{R}^n.$$

- On appelle taille de la boîte $[\mathbf{x}]$ la longueur $w([\mathbf{x}])$ de son plus grand coté

$$w([\mathbf{x}]) = \max_{i\in\{1,\ldots n\}} w([x_i]).$$

- La norme vectorielle de la boîte $[\mathbf{x}]$ est définie par

$$\|[\mathbf{x}]\| = \max_{i\in\{1,\ldots n\}} |[x_i]|.$$

2.1.2.4 Les opérations ensemblistes sur $\mathbb{IR}^n$

De même que pour les intervalles, nous allons définir l'intersection et l'union de vecteurs intervalles. A noter que ces définitions sont identiques à leur version ensembliste. Nous avons cependant choisi de les redéfinir pour davantage de clarté.

Définition 2.12 (Intersection et union de boîtes)**.** Soient $[\mathbf{x}]$,$[\mathbf{y}] \in \mathbb{IR}^n$, deux vecteurs intervalles.
- L'intersection de $[\mathbf{x}]$ et $[\mathbf{y}]$, notée $[\mathbf{x}] \cap [\mathbf{y}]$, est définie par

$$[\mathbf{x}] \cap [\mathbf{y}] = \begin{cases} \emptyset \text{ si il existe } i \in \{1, \ldots n\} \text{ tel que } [x_i] \cap [y_i] = \emptyset, \\ ([x_1] \cap [y_1], \ldots, [x_n] \cap [y_n]) \text{ sinon.} \end{cases}$$

- L'union, au sens des intervalles, de $[\mathbf{x}]$ et $[\mathbf{y}]$, notée $[\mathbf{x}] \sqcup [\mathbf{y}]$, est définie par

$$[\mathbf{x}] \sqcup [\mathbf{y}] = ([[x_1] \cup [y_1]], \ldots, [[x_n] \cup [y_n]]).$$

Notons qu'on a bien $[\mathbf{x}] \cap [\mathbf{y}] \in \mathbb{IR}^n$ et $[\mathbf{x}] \sqcup [\mathbf{y}] \in \mathbb{IR}^n$.

Définition 2.13 (Inclusion)**.** L'inclusion d'un vecteur intervalle $[\mathbf{x}] \in \mathbb{IR}^n$ dans un vecteur intervalle $[\mathbf{y}] \in \mathbb{IR}^n$, notée $[\mathbf{x}] \subset [\mathbf{y}]$, est définie par

$$[\mathbf{x}] \subset [\mathbf{y}] \Leftrightarrow (\forall i, 1 \leq i \leq n, [x_i] \subset [y_i]).$$

Définition 2.14 (Bissection)**.** On appelle plan principal d'une boîte $[\mathbf{x}] \in \mathbb{IR}^n$, le plan perpendiculaire au plus grand coté de $[\mathbf{x}]$, qui partage la boîte en deux boîtes symétriques La bissection d'une boîte $[\mathbf{x}] \in \mathbb{IR}^n$ consiste à la scinder en deux nouvelles boîtes séparées par le plan principal.

Exemple 2.15. *La figure 2.3 illustre la bissection d'un pavé de* $\mathbb{IR}^3$.

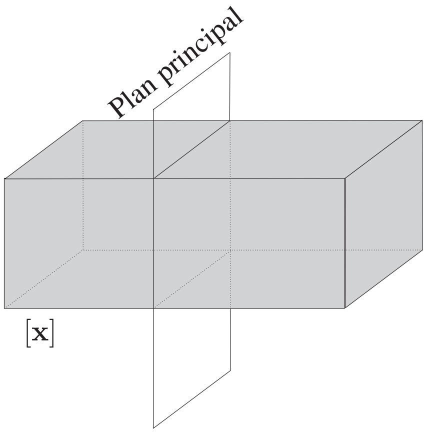


Figure 2.3 – Bissection d'un pavé $[\mathbf{x}]$ de $\mathbb{IR}^3$ suivant son plan principal.

Exemple 2.16. *La bissection de* $[\mathbf{x}] = [1,3] \times [-1,5] \in \mathbb{IR}^2$ *conduit à deux boîtes* $[\mathbf{x}_1] = [1,3] \times [-1,2]$ *et* $[\mathbf{x}_2] = [1,3] \times [2,5]$.

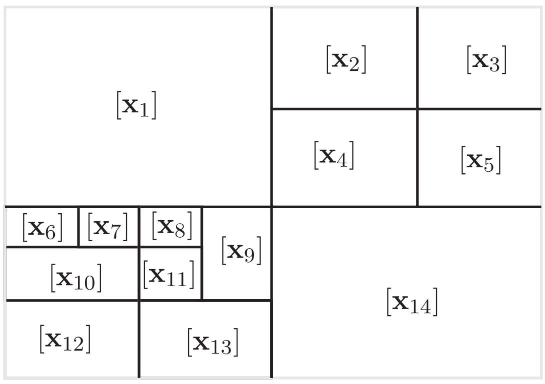
2.1.2.5 Sous-pavages

Définition 2.17 (Sous-pavage d'ensemble)**.** On appelle sous-pavage de l'ensemble $\mathbb{X} \in \mathbb{IR}^2$ un ensemble de boîtes $\{[\mathbf{x}_i]\}_{i \in \mathbb{N}}$ de $\mathbb{R}^n$ qui vérifient les conditions suivantes :

$$\begin{cases} \bigcup_{i \in \mathbb{N}} [\mathbf{x}_i] = \mathbb{X}, \\ \forall i, j \in \mathbb{N}, i \neq j,\ w\left([\mathbf{x}_i] \cap [\mathbf{x}_j]\right) = 0. \end{cases}$$

c'est-à-dire que $\{[\mathbf{x}_i]\}_{i \in \mathbb{N}}$ est un recouvrement de $\mathbb{X}$.

Exemple 2.18. *La figure 2.4 représente un sous-pavage d'un pavé* $\mathbb{X} \in \mathbb{R}^2$.



$\mathbb{X}$

Figure 2.4 – Exemple de sous-pavage d'un pavé $[\mathbf{x}]$ de $\mathbb{IR}^2$.

Proposition 2.19. *Tout ensemble compact*[1] *plein*[2] $\mathbb{X} \subset \mathbb{R}^n$ *peut être encadré par un sous-pavage intérieur* $\mathbb{X}^-$ *et un sous-pavage extérieur* $\mathbb{X}^+$ *tels que*

[1] Fermé et borné
[2] de mesure non nulle

$$\mathbb{X}^{-} \subset \mathbb{X} \subset \mathbb{X}^{+}.$$

Démonstration. Admis. □

Nous allons maintenant nous intéresser à l'arithmétique des intervalles. La notion de fonction d'inclusion y sera abordée. Nous lui consacrons une attention toute particulière étant donnée son importance dans la suite.

2.1.3 Arithmétique des intervalles

2.1.3.1 Les relations

En tant que parties de $\mathbb{R}$, les éléments de $\mathbb{IR}$ héritent naturellement des relations d'égalité, d'infériorité et de supériorité de $\mathbb{R}$.

Définition 2.20 (Egalité, infériorité)**.** Soit $[x]$,$[y] \in \mathbb{IR}$.
L'égalité des deux intervalles $[x]$ et $[y]$, notée $[x] = [y]$, est définie par

$$[x] = [y] \;\; \Leftrightarrow \;\; x^{-} = y^{-} \text{ et } x^{+} = y^{+}.$$

Les relations d'infériorités (stricte et large) sont définies par

$$[x] < [y] \Leftrightarrow x^{+} < y^{-},$$
$$[x] \leq [y] \Leftrightarrow x^{+} \leq y^{-}.$$

Ces deux relations se transposent trivialement aux symboles $>$ et $\geq$.

Remarquons que les relations d'ordre $\leq$ et $\geq$ sont partielles dans $\mathbb{IR}$, c'est-à-dire que, si $[x]$ n'est pas inférieur (resp. supérieur) à $[y]$, alors $[y]$ n'est pas forcément inférieur (resp. supérieur) à $[x]$ étant donné que deux intervalles peuvent avoir une intersection non vide.

Remarque 2.21 (Extension sur $\mathbb{IR}^n$)**.** *Les définitions d'égalité, d'infériorité et de supériorité sur $\mathbb{IR}$ peuvent être étendues trivialement sur $\mathbb{IR}^n$, dans le sens où chaque composante du vecteur intervalle devra vérifier la relation. Par exemple, l'égalité de deux vecteurs intervalles $[\mathbf{x}]$, $[\mathbf{y}] \in \mathbb{IR}^n$ sera définie par :*

$$[\mathbf{x}] = [\mathbf{y}] \Leftrightarrow \forall i, 1 \leq i \leq n, \;\; [x_i] = [y_i] \, .$$

2.1.3.2 Les opérations binaires

Les opérations classiques de l'arithmétique des réels, telles l'addition (+), la soustraction (-), la multiplication ($\times$), la division (/), le maximum ($\max$) et le minimum ($\min$), peuvent également être étendues aux intervalles.

Définition 2.22 (Opérations sur $\mathbb{IR}$)**.** Soient $\star \in \{+, -, \times, /, \max, \min\}$ et $[x] \in \mathbb{IR}$, $[y] \in \mathbb{IR}$ deux intervalles. On a

$$[x] \star [y] \triangleq [\{x \star y | x \in [x] \, , y \in [y]\}] \, .$$

Ainsi défini, $[x] \star [y]$ est le plus petit intervalle contenant tous les résultats possibles de l'opération $\star$ appliquée à tous les éléments x de $[x]$ et tous les éléments y de $[y]$. Précisons un peu l'ensemble $[x] \star [y]$ pour chacune des opérations usuelles :

- $[x^-, x^+] + [y^-, y^+] = [x^- + y^-, x^+ + y^+]$,
- $[x^-, x^+] - [y^-, y^+] = [x^- - y^+, x^+ - y^-]$,
- $[x^-, x^+] \times [y^-, y^+] = [\min(x^-y^-, x^-y^+, x^+y^-, x^+y^+), \max(x^-y^-, x^-y^+, x^+y^-, x^+y^+)]$,
- $1/[y^-, y^+] = \begin{cases} \left[\min\left(\frac{1}{y^-}, \frac{1}{y^+}\right), \max\left(\frac{1}{y^-}, \frac{1}{y^+}\right)\right] & \text{si } \{0\} \notin [y^-, y^+] \\ [-\infty, +\infty] & \text{si } \{0\} \in [y^-, y^+] \end{cases}$
- $[x^-, x^+]/[y^-, y^+] = [x^-, x^+] \times (1/[y^-, y^+])$
- $\max([x^-, x^+], [y^-, y^+]) = [\max(x^-, y^-), \max(x^+, y^+)]$,
- $\min([x^-, x^+], [y^-, y^+]) = [\min(x^-, y^-), \min(x^+, y^+)]$.

Voici concrètement quelques exemples :

- $([1, 2.2] \times [0, 2]) + [1, 3] = [0, 4.4] + [1, 3] = [1, 7.4]$,
- $1/[-2, 2] = [-\infty, +\infty]$,
- $[3, 4]/[0, 0] = \emptyset$.

Nous pouvons d'ores et déjà constater que les opérations définies ci-dessus ne présentent pas les propriétés algébriques de leurs contreparties ponctuelles. En voici quelques contre-exemples :

- Tout d'abord, la soustraction n'est pas la réciproque de l'addition. En effet, $(\mathbb{IR}, +)$ n'est pas un groupe.
  Nous avons l'inclusion $\{0\} \subset [x] - [x]$ mais il ne s'agit pas, en général, d'une égalité. En effet,
  $$[x] - [x] = \{x - y | x \in [x], y \in [x]\} \supset \{x - x | x \in [x]\} = \{0\}.$$
  Par exemple, si $[x] = [2, 3]$, $[x] - [x] = [2, 3] - [2, 3] = [-1, 1] \neq 0$.
- De la même façon, la division n'est pas la réciproque de la multiplication.
  Nous avons seulement l'inclusion $\{1\} \subset [x]/[x]$ et non l'égalité.
  Par exemple, si $[x] = [2, 3]$, $[x]/[x] = [2, 3]/[2, 3] = [2/3, 3/2] \neq 1$.
- De plus, la multiplication d'un intervalle par lui-même ne correspond pas à l'élévation au carré. Prenons par exemple l'intervalle $[x] = [-3, 2]$, on a $[x] \times [x] = [-3, 2] \times [-3, 2] = [-6, 9]$, alors que $[x]^2 = \{x^2 | x \in [x]\} = [0, 9]$.
- Enfin, la multiplication n'est pas distributive par rapport à l'addition.
  Pour $[x] = [-2, 3]$, $[y] = [1, 4]$ et $[z] = [-2, 1]$, on a
  $$\begin{aligned} [x] \times ([y] + [z]) &= [-2, 3] \times ([1, 4] + [-2, 1]) \\ &= [-2, 3] \times [-1, 5] \\ &= [-10, 15], \end{aligned}$$
  $$\begin{aligned} [x] \times [y] + [x] \times [z] &= [-2, 3] \times [1, 4] + [-2, 3] \times [-2, 1] \\ &= [-8, 12] \times [-6, 4] \\ &= [-14, 16]. \end{aligned}$$
  Comme l'illustre cet exemple, la multiplication est seulement *sous-distributive* par rapport à l'addition, c'est-à-dire
  $$[x] \times ([y] + [z]) \subset [x] \times [y] + [x] \times [z],$$
  où
  $$\begin{cases} [x] \times ([y] + [z]) = \{x \times (y + z) | x \in [x], y \in [y], z \in [z]\}, \\ [x] \times [y] + [x] \times [z] = \{[x] \times [y] + [x'] \times [z] \, | x, x' \in [x], y \in [y], z \in [z]\}. \end{cases}$$

Remarque 2.23 (Extension sur $\mathbb{IR}^n$)**.** *Pour étendre ces règles de calcul à $\mathbb{IR}^n$, il suffit de reprendre les définitions dans le cas réel et de remplacer les variables et les opérations par leurs équivalents intervalles. Par exemple, le produit scalaire de deux vecteurs intervalles $[\mathbf{x}], [\mathbf{y}] \in \mathbb{IR}^n$ est défini par*

$$[\mathbf{x}] \cdot \left[\mathbf{y}^T\right] = \sum_{i=0}^{n} [x_i] [y_i] .$$

De la même façon, nous pouvons créer l'ensemble des matrices intervalles $\mathbb{IR}^m \times \mathbb{IR}^n$ de taille m lignes et n colonnes et définir des règles de calcul d'algèbre linéaire simples comme l'addition et la multiplication.

2.1.3.3 Les fonctions élémentaires

Nous allons maintenant définir les fonctions élémentaires ($\sin, \exp, \coth, \ldots$) prenant des intervalles pour paramètres.

Définition 2.24 (Fonctions élémentaires sur $\mathbb{IR}$)**.** Soit $f \in \{\cos, \sin, \exp, \log, \ldots\}$ une fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$ et soit $[z] \in \mathbb{IR}$, on définit :

$$f([z]) \triangleq [\{f(z) | z \in [z]\}] .$$

Autrement dit, $f([z])$ est le plus petit intervalle contenant tous les résultats possibles de f, appliquée à tous les éléments z de $[z]$.

Exemple 2.25. *Pour une fonction f monotone, on peut calculer aisément l'image $f([z])$ d'un intervalle $[z]$. Voici quelques exemples :*

$$\begin{aligned} &\bullet\ \exp([a,b]) = [\exp(a), \exp(b)] \\ &\bullet\ ([-1,3])^2 = [0,9] \\ &\bullet\ abs([-7,1]) = [1,7] \\ &\bullet\ \sqrt{[-10,4]} = [0,2] \\ &\bullet\ \log([-2,-1]) = \emptyset \end{aligned} \tag{2.1}$$

Il est également possible d'établir des algorithmes de calcul pour des fonctions périodiques à partir du moment où l'on connaît ces fonctions pour les réels. Par exemple,

$$\sin([\pi/3, \pi]) = [0, 1].$$

2.1.3.4 Evaluation d'une expression

Puisque nous savons calculer le résultat d'une opération arithmétique ou d'une fonction élémentaire quand les variables prennent pour valeurs des intervalles, on sait également calculer le résultat d'une expression mêlant opérations arithmétiques ou algébriques et fonctions élémentaires sur des intervalles. Voici quelques exemples pour illustrer.

– Considérons l'expression $A = [x]^3 - 2\,[x]^2 + [x] - 3$ et cherchons à calculer A pour $[x] = [-5, 2]$:

$$\begin{aligned}
A &= [-5,2]^3 - 2[-5,2]^2 + [-5,2] - 3 \\
A &= [-125,8] - 2[0,25] + [-5,2] - 3 \\
A &= [-125,8] - [0,50] + [-5,2] - 3 \\
A &= [-183,7].
\end{aligned}$$

– Considérons l'expression à trois variables $B = \sin([x]) + 2\,[x]\exp([y]) - [y]^2\,\sqrt{[z]}$ et cherchons le résultat pour $x = \left[-\pi, \frac{\pi}{4}\right]$, $y = [-1,1]$ et $z = [1,4]$:

$$\begin{aligned}
B &= \sin\left[-\pi, \frac{\pi}{4}\right] + 2\left[-\pi, \frac{\pi}{4}\right]\exp[-1,1] - [-1,1]^2 \times \sqrt{[1,4]} \\
B &= \sin\left[-\pi, \frac{\pi}{4}\right] + 2\left[-\pi, \frac{\pi}{4}\right]\exp[-1,1] - [-1,1]^2 \times \sqrt{[1,4]} \\
B &= \left[-1, \frac{\sqrt{2}}{2}\right] + \left[-2\pi, \frac{\pi}{2}\right] \times \left[\frac{1}{e}, e\right] - [0,1] \times \ [1,2] \\
B &= \left[-1, \frac{\sqrt{2}}{2}\right] + \left[-2\pi e, \frac{\pi e}{2}\right] - [0,2] \\
B &= \left[-3 - 2\pi e, \frac{\sqrt{2}}{2} + \frac{\pi e}{2}\right].
\end{aligned}$$

Intéressons-nous maintenant à l'utilisation des vecteurs d'intervalles dans un cadre fonctionnel, c'est-à-dire par rapport à des fonctions intervalles vectorielles définies sur $\mathbb{IR}^n$ à valeurs dans $\mathbb{IR}^m$. Nous allons introduire la notion de fonction d'inclusion.

2.1.3.5 Les fonctions d'inclusion

Définition 2.26 (Fonction d'inclusion)**.** Soit $\mathbf{f}$ une fonction définie sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^m$. Une fonction ensembliste $[\mathbf{f}] : \mathbb{IR}^n \to \mathbb{IR}^m$ est une fonction d'inclusion de $\mathbf{f}$ si les deux conditions suivantes sont vérifiées :
i) $[\mathbf{f}]$ est monotone par inclusion, c'est-à-dire :

$$\forall\,[\mathbf{y}]\,,[\mathbf{x}] \in \mathbb{IR}^n, [\mathbf{y}] \subseteq [\mathbf{x}] \Rightarrow [\mathbf{f}]\,([\mathbf{y}]) \subseteq [\mathbf{f}]\,([\mathbf{x}])\,. \tag{2.2}$$

ii) $[\mathbf{f}]$ est une extension intervalle de la fonction $\mathbf{f}$, c'est-à-dire :

$$\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{f}\,(\mathbf{x}) = [\mathbf{f}]\,(\mathbf{x})\,. \tag{2.3}$$

Théorème 2.27. *Soit* $\mathbf{f}$ *une fonction de* $\mathbb{R}^n$ *dans* $\mathbb{R}^m$ *et* $[\mathbf{f}]$ *une fonction d'inclusion pour* $\mathbf{f}$, *définie sur* $\mathbb{IR}^n$ *à valeurs dans* $\mathbb{IR}^m$, *alors*

$$\forall\,[\mathbf{x}] \in \mathbb{IR}^n, \mathbf{f}\,([\mathbf{x}]) \subset [\mathbf{f}]\,([\mathbf{x}])\,.$$

La figure 2.5 illustre cette inclusion.

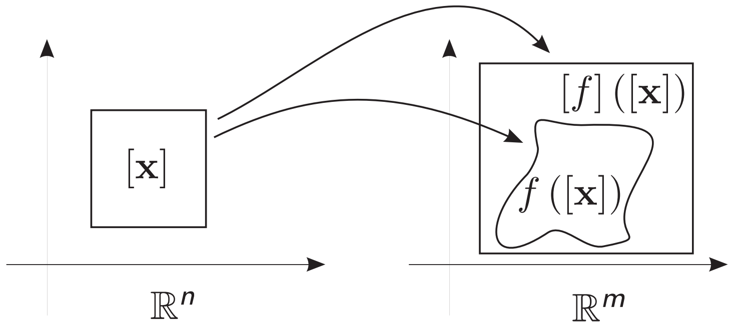


Figure 2.5 – Illustration du théorème 2.27.

Démonstration. Par définition,

$$\mathbf{f}([\mathbf{x}]) \triangleq \{\mathbf{f}(\mathbf{x})|\mathbf{x} \in [\mathbf{x}]\}.$$

D'après l'égalité (2.3), on peut écrire

$$\mathbf{f}([\mathbf{x}]) = \{[\mathbf{f}](\mathbf{x})|\mathbf{x} \in [\mathbf{x}]\}.$$

Or, d'après (2.2), on a

$$\forall \mathbf{x} \in [\mathbf{x}], \ \ [\mathbf{f}](\mathbf{x}) \subseteq [\mathbf{f}]([\mathbf{x}]).$$

On a donc bien l'inclusion

$$\mathbf{f}([\mathbf{x}]) \subset [\mathbf{f}]([\mathbf{x}]).$$

□

Quelle que soit la fonction $\mathbf{f}$, nous pouvons trouver une fonction d'inclusion $[\mathbf{f}]$ pour $\mathbf{f}$. En effet, il suffit, par exemple, de substituer dans l'expression de $\mathbf{f}$ les opérations sur les réels par les mêmes opérations sur les intervalles. En voici deux exemples :

Exemple 2.28. *Soit la fonction $f(x) = x^2 + 2x + 4$. On obtient une fonction d'inclusion pour f en posant*

$$[f] : \begin{cases} \mathbb{IR} & \rightarrow & \mathbb{IR} \\ [x] & \rightarrow & [x]^2 + 2[x] + 4. \end{cases}$$

Ainsi, pour $[x] = [-3, 4]$, nous obtenons

$$\begin{array}{rcl} [f]([-3,4]) & = & [-3,4]^2 + 2[-3,4] + 4 \\ & = & [-2, 28]. \end{array}$$

Notons que l'image par f de l'intervalle $[-3, 4]$ est $f([-3, 4]) = [3, 28]$ qui est bien un sous ensemble de $[f]([-3, 4]) = [-14, 28]$.

Exemple 2.29. *Soit $\mathbf{f}$ la fonction définie par*

$$\mathbf{f} : \begin{cases} \mathbb{R}^2 & \rightarrow & \mathbb{R}^3 \\ (x_1, x_2) & \rightarrow & \left(x_1 x_2, x_1^2, x_1 - x_2\right) \end{cases}.$$

On obtient une fonction d'inclusion $[\mathbf{f}]$ de $\mathbf{f}$ en posant

$$[\mathbf{f}] : \begin{cases} \mathbb{IR}^2 \rightarrow & \mathbb{IR}^3 \\ ([x_1], [x_2]) \rightarrow & \left([x_1] \times [x_2], [x_1]^2, [x_1] - [x_2]\right) \end{cases}.$$

Nous pouvons calculer $[\mathbf{f}]([-1, 1], [-1, 1]) = ([-1, 1], [0, 1], [-2, 2])$.

Dans la suite, nous nous intéressons à l'élaboration de fonctions d'inclusion efficaces (au sens d'un critère) pour une fonction vectorielle $\mathbf{f}$. Avant de rentrer dans les détails, précisons pour commencer que la construction d'une fonction d'inclusion $[\mathbf{f}]$ pour $\mathbf{f} = (f_1, \ldots, f_m)$ se ramène à l'élaboration de m fonctions d'inclusion $[f_1], \ldots, [f_m]$ pour ses fonctions coordonnées $f_1, \ldots, f_m$. En effet, une fonction d'inclusion pour $\mathbf{f}$ sera donnée par

$$[\mathbf{f}]([\mathbf{x}]) = ([f_1]([\mathbf{x}]), \ldots, [f_m]([\mathbf{x}])).$$

Nous pouvons donc nous focaliser uniquement sur le calcul de fonctions d'inclusion d'une fonction à valeur dans $\mathbb{R}$. Plus particulièrement, nous nous intéresserons à celles qui sont construites par compositions des fonctions élémentaires usuelles (sin, exp, $\sqrt{}, \ldots$) et des opérateurs $(+, -, \times, /)$.

A priori, la méthode la plus simple pour trouver une fonction d'inclusion d'une fonction algébrique $f : \mathbb{R}^n \to \mathbb{R}$ consiste à remplacer les occurrences des variables scalaires $(x_1, x_2, ...)$ dans l'expression de f par les variables intervalles correspondantes $([x_1], [x_2], \ldots)$, ainsi que toutes les fonctions élémentaires et les opérateurs par leurs extensions intervalles. Nous obtenons ainsi la *fonction d'inclusion naturelle* de f. Cependant, bien souvent, cette fonction d'inclusion naturelle n'est pas minimale au sens de l'inclusion.

Définition 2.30 (Fonction d'inclusion minimale)**.** La fonction d'inclusion minimale $[f^\star]$ de f vérifie :

$$\forall \mathbf{x} \in [\mathbf{x}], \quad [f^\star]([\mathbf{x}]) = [f([\mathbf{x}])].$$

Comme l'illustre l'exemple suivant qui est extrait de [64], la fonction d'inclusion naturelle est rarement minimale. Kieffer propose une comparaison des efficacités de différentes fonctions d'inclusion d'une même fonction f.

Exemple 2.31. *Comparons les performances des fonctions d'inclusions établies à partir de quatres formulations d'une même fonction f.*

$$\begin{aligned}
f_1(x) &= x(x+1) \\
f_2(x) &= x \times x + x \\
f_3(x) &= x^2 + x \\
f_4(x) &= \left(x + \frac{1}{2}\right) - \frac{1}{4}.
\end{aligned}$$

Pour $[x] = [-1, 1]$, on obtient

$$\begin{aligned}
[f_1]([x]) &= [x]([x]+1) = [-2, 2] \\
[f_2]([x]) &= [x] \times [x] + [x] = [-2, 2] \\
[f_3]([x]) &= [x]^2 + [x] = [-1, 2] \\
[f_4]([x]) &= \left([x] + \frac{1}{2}\right) - \frac{1}{4} = \left[-\frac{1}{4}, 2\right].
\end{aligned}$$

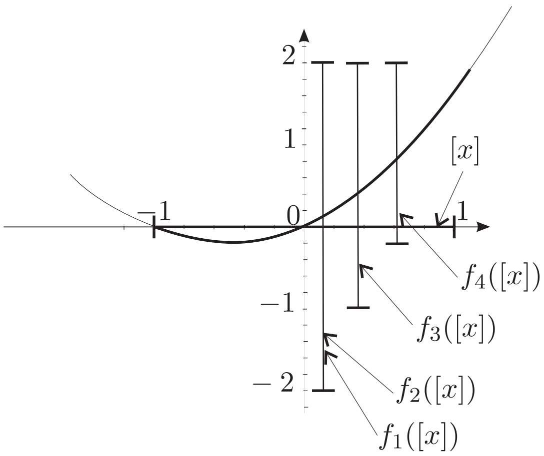


Figure 2.6 – Comparaison des efficacités des fonctions d'inclusion $[f_1]$, $[f_2]$, $[f_3]$ et $[f_4]$.

Selon l'expression choisie pour la fonction f, l'encadrement obtenu s'avère plus ou moins grossier (voir figure 2.6). Cela vient du fait que les deux expressions $[x] \times [x]$ et $[x]^2$ ne sont pas équivalentes. En effet, dans l'expression $[x] \times [x]$, chacune des occurrences de x peut varier indépendamment. L'ensemble image $f([-1, 1])$ de la fonction initiale est $\left[-\frac{1}{4}, 2\right]$. $[f_4]$ est donc la fonction d'inclusion minimale de la fonction f.

Une fonction d'inclusion $[f]$ qui fournit un intervalle image différent de l'évaluation intervalle de la fonction réelle (c'est-à-dire $[f]([\mathbf{x}]) > f([\mathbf{x}])$) sur l'intervalle considéré est dite *pessimiste*. Il n'existe pas de méthode systématique permettant, pour une fonction donnée, d'obtenir une fonction d'inclusion minimale. Cependant, on peut constater empiriquement que plus une variable apparaît fréquemment dans l'expression de f, plus le pessimisme a des chances de devenir important. Cela peut s'interpréter par le fait que deux occurrences d'une même variable sont considérées comme variant indépendamment l'une de l'autre. Ainsi, les incertitudes se combinent pour un résultat plus mauvais que si l'on avait considéré les occurrences comme variant de façon identique.

Remarque 2.32. *Notons qu'il existe d'autres types de fonctions d'inclusion faisant intervenir par exemple, des développements en série de la fonction initiale (pour plus de détails, voir [45]). Elle donne des résultats en général moins pessimistes que la fonction d'inclusion naturelle dans le cas où $[x]$ est de faible longueur et lorsque le problème de dépendance pour f se pose.*

Nous allons nous intéresser maintenant à deux types de problèmes pour lesquels le calcul ensembliste permet d'obtenir des résultats probants, à savoir la résolution des systèmes d'équations non linéaires puis l'optimisation globale non linéaire.

2.2 Résolution de système d'équations non linéaires.

Avant de présenter une méthode de résolution d'un système d'équations non linéaires, qui sera l'occasion de présenter l'algorithme SIVIA, il nous paraît important de présenter quelques autres méthodes de résolution pour ce même problème. Ceci nous permettra en particulier de comparer les performances des différentes approches.

2.2.1 Introduction

Le problème considéré dans cette partie est la résolution d'un système de n équations non linéaires à n inconnues, qui s'écrit sous la forme

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \text{ pour } \mathbf{x} \in [\mathbf{x}_0] \subset \mathbb{R}^n. \tag{2.4}$$

Par résolution, nous entendons la recherche de toutes les solutions réelles de ce système. Parmi les multiples approches de ce problème, nous distinguerons les méthodes de résolution numériques, et celles dites formelles (ou encore algébriques).

2.2.1.1 Les méthodes numériques

Les méthodes numériques sont les premières à avoir été développées. Aujourd'hui encore, elles sont les plus utilisées. Dans cette première famille, on peut distinguer plusieurs genres.

2.2.1.1.0.1 Les méthodes de type Newton Le principe de base de toute méthode dérivée de l'algorithme de Newton consiste à construire une suite de points $\mathbf{x}_i$ initialisée en un point $\mathbf{x}_0$ choisi par l'utilisateur, et qui converge, si possible, vers une solution du système. Ainsi, dans le cas de la résolution d'une seule équation à une variable $f(x) = 0$, le processus le plus classique pour construire cette suite s'écrit

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)},$$

Ce procédé se généralise sans grande difficulté dans le cas de n équations à n inconnues. Pour davantage de détails concernant cet algorithme, le lecteur pourra se référer à l'annexe C.

L'approche de Newton possède toutefois des défauts non négligeables. Tout d'abord, la convergence de la suite $(\mathbf{x}_k)_k$ vers une solution est difficile à garantir et peut dépendre fortement du choix du point initial $\mathbf{x}_0$. En outre, pour obtenir plusieurs solutions, il est indispensable de répéter la méthode décrite ci-dessus en changeant les points initiaux mais, pour avoir l'assurance d'avoir trouvé toutes les solutions réelles du système, ceci suppose d'en connaître *a priori* le nombre.

2.2.1.1.0.2 Les méthodes d'optimisation globale Remarquons que la résolution d'un système d'équations non linéaires peut également être vu comme un problème d'optimisation globale. Le critère pourra par exemple s'écrire

$$h(\mathbf{x}) = \sum_{i=1}^{n} f_i(\mathbf{x})^2.$$

La recherche des solutions du système (2.4) est alors équivalente à celle des minimas de la fonction h.

Cette méthode possède toutefois des inconvénients. Même dans le cas où f possède peu de zéros, la fonction h peut posséder beaucoup de minima locaux. Aussi cette méthode n'est-elle utilisée que dans les rares cas où elle permet une simplification du problème.

2.2.1.1.0.3 Les méthodes par homotopie Le principe des méthodes par homotopie consiste à tenter de converger vers les solutions du système. Pour cela, on résoud successivement des systèmes approchés. Considérons la fonction $\mathbf{h}$ définie par

$$\mathbf{h}(\mathbf{x}, t) = (1-t)\,\mathbf{g}(\mathbf{x}) + t\mathbf{f}(\mathbf{x})$$

où $\mathbf{f}$ est la fonction dont on cherche les zéros, $\mathbf{g}$ une fonction arbitraire dont on connaît les zéros, et t un scalaire inclus dans $[0, 1]$. Lorsque $t = 0$, les zéros de $\mathbf{h}$ sont égaux à ceux de $\mathbf{g}$, qui, rappelons le, sont connus. La méthode consiste à accroître progressivement la valeur de t et à résoudre chaque nouveau système rencontré à partir des solutions du précédent, ceci, jusqu'à atteindre la valeur $t = 1$. Lorsque $t = 1$, les zéros de $\mathbf{h}$ correspondent alors à ceux de la fonction $\mathbf{f}$.

Notons tout d'abord que la convergence du processus n'est pas garantie. D'autre part, un inconvénient majeur de cette approche réside dans le fait qu'il est nécessaire de travailler dans le plan complexe $\mathbb{C}$, des racines initialement complexes pouvant aboutir à des solutions réelles. Pour un problème possédant de nombreuses solutions complexes et peu de solutions réelles, ceci est un véritable handicap. A l'heure actuelle, cette approche est principalement utilisée dans le cadre de la résolution de systèmes d'équations polynomiales.

2.2.1.1.0.4 La méthode ensembliste C'est cette dernière approche numérique qui nous intéresse. Nous la présentons en détails dans la partie 2.2.2 suivante.

2.2.1.2 Les méthodes algébriques

Contrairement aux méthodes numériques, les méthodes formelles ne s'appliquent que pour des systèmes d'équations polynomiales. Elles sont basées sur une manipulation des équations et tendent à retarder au maximum l'étape de résolution numérique.
Le principal outil des méthodes algébriques est l'utilisation des bases de Gröbner. Ces dernières, issues du travail de thèse de Buchberger, permettent une réécriture du système initial sous une forme plus simple à résoudre. Les personnes intéressées par davantage de détails concernant la définition et la construction d'une base de Gröbner pourront consulter l'annexe D.

L'utilisation des bases de Gröbner permet une réécriture du système (2.4) sous forme triangulaire

$$\begin{cases} f_1\left(x_n\right) & = 0 \\ f_2\left(x_{n-1}, x_n\right) & = 0 \\ \vdots & \vdots \\ f_t\left(x_1, x_2, \ldots, x_n\right) & = 0. \end{cases} \tag{2.5}$$

La résolution de l'équation en x_n fournit numériquement l'ensemble des solutions du système pour x_n. Ces dernières, reportées dans l'équation suivante permettent d'obtenir les solutions associées pour x_{n-1} et ainsi de suite...

Toutefois, les équations du système (2.5) sont généralement de degré relativement élevé et le fait de reporter successivement des estimations de solutions contribue à la prolifération des erreurs numériques. D'autre part, seul le nombre de solutions complexes est précisément connu et nous n'avons aucune information sur le nombre de solutions réelles.

Enfin, en plus des inconvénients classiques concernant les inévitables étapes de calcul numérique, le principal défaut des méthodes algébriques provient de la complexité des expressions générées par les manipulations des équations. La construction d'une base de Gröbner est en général très longue. Par ailleurs, le nouveau système obtenu peut être de grande dimension (jusqu'à doublement exponentielle au nombre de variables), avec des coefficients complexes et des degrés élevés. Sauf dans certains cas particuliers, cette méthode de résolution est très coûteuse.

2.2.1.3 Perspectives

A partir de maintenant, nous consacrons notre attention aux techniques intervalles pour la résolution de systèmes d'équations non linéaires.

Afin de mettre en évidence l'intérêt du raisonnement intervalle face à ce problème connu comme difficile, nous avons choisi de présenter ce qui peut se faire quasiment de plus simple, sachant que cela est déjà loin d'être inefficace. Nous présentons notamment l'algorithme intervalle SIVIA, dédié au problème de l'inversion ensembliste, problème qui dépasse largement le cadre des systèmes d'équations. Ne nécessitant aucune contrainte forte sur la nature des équations, SIVIA s'avère performant sur les systèmes de dimension modeste.

Pour pouvoir traiter des problèmes plus ambitieux, il est préférable de choisir d'autres méthodes comme par exemple, la généralisation intervalle de la méthode de Newton. Le lecteur intéressé pourra consulter [72] pour plus de détails.

2.2.2 L'inversion ensembliste

Dans ce paragraphe, nous présentons l'algorithme SIVIA qui résoud le problème très général de l'inversion ensembliste. Ses performances sur le problème plus spécifique des sytèmes d'équations non linéaires seront ensuite illustrées sur un exemple de séparation de sources dans la section 3.1.
Commençons d'abord par définir la notion d'inversion ensembliste.

2.2.2.1 Introduction

Considérons un ensemble $\mathbb{X}$ de $\mathbb{R}^n$ et une fonction $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. Soit l'ensemble $\mathbb{Y} \subset \mathbb{R}^m$ défini par la relation

$$\mathbb{Y} = \mathbf{f}(\mathbb{X}). \tag{2.6}$$

Le problème auquel nous nous intéressons est celui de l'inversion ensembliste défini comme suit :

Définition 2.33 (Inversion ensembliste)**.** On appelle inversion ensembliste, la caractérisation de l'ensemble $\mathbb{X}$, comme image réciproque par la fonction $\mathbf{f}$ de l'ensemble $\mathbb{Y}$, autrement dit,

$$\mathbb{X} = \mathbf{f}^{-1}(\mathbb{Y}) = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{f}(\mathbf{x}) \in \mathbb{Y}\}.$$

Remarque 2.34. *Pour $\mathbb{Y} = \mathbf{0}$, nous retrouvons le problème précédent de la résolution du système d'équations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.*

Exemple 2.35. *Caractériser un des sous ensembles suivants*

$$\begin{aligned}
\mathbb{X}_1 &= \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 \le 1\right\} \\
\mathbb{X}_2 &= \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\right\} \\
\mathbb{X}_3 &= \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 \le 1 \textit{ et } x_1 - x_2 \in [0, 1]\right\}
\end{aligned}$$

est un problème d'inversion ensembliste. En effet, nous pouvons réécrire ces ensembles

$$\begin{aligned}
\mathbb{X}_1 &= \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 \in [-\infty, 1]\right\} \\
\mathbb{X}_2 &= \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 \in [1, 1]\right\} \\
\mathbb{X}_3 &= \left\{(x_1, x_2) \in \mathbb{R}^2 | \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 - x_2 \end{pmatrix} \in [-\infty, 1] \times [0, 1]\right\}
\end{aligned}$$

et en posant

$$\mathbf{f}: \left\{ \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x_1, x_2) & \rightarrow & x_1^2 + x_2^2 \end{array} \right. \quad \textit{et} \quad \mathbf{g}: \left\{ \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ (x_1, x_2) & \rightarrow & \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 - x_2 \end{pmatrix} \end{array} \right.$$

nous obtenons

$$\begin{array}{l} \mathbb{X}_1 = \mathbf{f}^{-1}\left([-\infty, 1]\right) \\ \mathbb{X}_2 = \mathbf{f}^{-1}\left([1, 1]\right) \\ \mathbb{X}_3 = \mathbf{g}^{-1}\left([-\infty, 1] \times [0, 1]\right). \end{array}$$

Avant de présenter l'algorithme d'inversion ensembliste SIVIA, précisons un peu le vocabulaire qui sera utilisé par la suite.

Définition 2.36 (Pavé acceptable, inacceptable, ambigu)**.** Un pavé $[\mathbf{x}]$ est dit acceptable si $[\mathbf{x}] \in \mathbb{X}$.
Un pavé $[\mathbf{x}]$ est dit inacceptable si $[\mathbf{x}] \cap \mathbb{X} = \emptyset$.
Un pavé $[\mathbf{x}]$ est dit ambigu si il n'est ni acceptable, ni inacceptable.
On parle de pavé indéterminé lorsqu'on ne peut dire si il est acceptable, inacceptable ou ambigu. Bien qu'il soit forcément dans une des trois catégories, l'état actuel des connaissances ne permet pas toujours de le classer.

D'après la propriété fondamentale des fonctions d'inclusions (voir théorème 2.27), on a

$$\forall \, [\mathbf{x}] \, , \mathbf{f}\left([\mathbf{x}]\right) \subset [\mathbf{f}]\left([\mathbf{x}]\right).$$

ainsi, nous pouvons obtenir des conditions suffisantes d'acceptablité (resp. d'inacceptabilité) d'un pavé $[\mathbf{x}]$.

Proposition 2.37. *On a :*
i) $[\mathbf{f}]\left([\mathbf{x}]\right) \subset \mathbb{Y} \Rightarrow [\mathbf{x}]$ *est acceptable.*
ii) $[\mathbf{f}]\left([\mathbf{x}]\right) \cap \mathbb{Y} = \emptyset \Rightarrow [\mathbf{x}]$ *est inacceptable.*

Démonstration. i) $[\mathbf{f}]\left([\mathbf{x}]\right) \subset \mathbb{Y} \Rightarrow \mathbf{f}\left([\mathbf{x}]\right) \subset \mathbb{Y} \Rightarrow [\mathbf{x}] \subset \mathbb{X} \Rightarrow [\mathbf{x}]$ est acceptable.
ii)$[\mathbf{f}]\left([\mathbf{x}]\right) \cap \mathbb{Y} = \emptyset \Rightarrow \mathbf{f}\left([\mathbf{x}]\right) \not\subset \mathbb{Y} \Rightarrow [\mathbf{x}] \cap \mathbb{X} = \emptyset \Rightarrow [\mathbf{x}]$ est inacceptable. □

Nous allons maintenant présenter l'algorithme SIVIA (*Set Inverter Via Interval Analysis*) et plus spécialement l'approche consistant à encadrer l'ensemble solution $\mathbb{X}$ par des sous-pavages.

2.2.3 SIVIA

L'algorithme SIVIA a été présenté dans [56][57][55]. L'objectif de cet algorithme est d'obtenir, connaissant $\mathbb{Y}$, un encadrement, aussi précis que l'on souhaite, de l'ensemble solution $\mathbb{X} = \mathbf{f}^{-1}\left(\mathbb{Y}\right)$, par deux sous-pavages $\mathbb{X}^-$ et $\mathbb{X}^+$ de sorte que $\mathbb{X}^- \subset \mathbb{X} \subset \mathbb{X}^+$. Notons que le sous-pavage $\mathbb{X}^-$ peut éventuellement être vide.
Pour obtenir cet encadrement, nous avons besoin d'une caractérisation appropriée de l'ensemble $\mathbb{Y}$, d'une fonction d'inclusion $[\mathbf{f}]$ de la fonction $\mathbf{f}$, et d'une valeur ε fixant la précision de l'encadrement. Nous supposons que l'on dispose d'un pavé initial $[\mathbf{x}_0]$ connu, suffisamment grand, de sorte $\mathbb{X} \subset [\mathbf{x}_0]$.

2.2.3.1 Principe

Nous allons chercher à partager le pavé initial de recherche $[\mathbf{x}_0]$ en trois classes de pavés : les pavés acceptables, les pavés inacceptables et les pavés indéterminés. Nous ferons en sorte que ces pavés soient de taille inférieure ou égale au seuil de tolérance ε fixé au départ.

– Le sous-pavage $\mathbb{X}^-$ évoqué précédemment sera l'ensemble des pavés dont on a pu prouver qu'ils étaient acceptables.
– Le sous-pavage $\mathbb{X}^+$ sera la réunion du sous-pavage $\mathbb{X}^-$ et de l'ensemble, noté $\Delta\mathbb{X}$, des pavé indéterminés.

Nous présentons maintenant l'algorithme SIVIA, appliqué à un pavé $[\mathbf{x}]$, initialement égal au pavé de recherche $[\mathbf{x}_0]$.

2.2.3.2 L'algorithme SIVIA (Set Inversion Via Interval Analysis)

Entrées

$[f]$	:	fonction d'inclusion pour f.
$[\mathbf{x}_0]$	:	pavé de recherche.
ε	:	seuil de largeur des pavés résultats.

Sorties

$\mathbb{X}^-$	:	ensemble des pavés acceptables.
$\Delta\mathbb{X}$	:	ensemble des pavés indéterminés.

Algorithme $[\mathbf{x}] = [\mathbf{x}_0]\,, \mathbb{X}^- := \emptyset$, $\Delta\mathbb{X} := \emptyset$ (initialisation).

Si $[\mathbf{f}]([\mathbf{x}]) = \mathbf{0}$ alors $\mathbb{X}^- := \mathbb{X}^- \cup [\mathbf{x}]$

Sinon **Si** $[\mathbf{f}]([\mathbf{x}]) \not\ni \mathbf{0}$ alors rien ($[\mathbf{x}]$ est inacceptable, donc sans intérêt)

Sinon **Si** $w([\mathbf{x}]) \leq \varepsilon,$ alors $\Delta\mathbb{X} := \Delta\mathbb{X} \cup [\mathbf{x}]$

Sinon diviser $[\mathbf{x}]$ en deux sous-pavés,
puis réitérer le traitement sur les 2 sous-pavés obtenus.

Fin Si

Fin Si

Fin Si

Précisons que la section du pavé $[\mathbf{x}]$ en deux sous-pavés se fait selon le plan principal. Il s'agit d'une bissection (voir définition 2.14).

Remarque 2.38. *Dans le cadre de la résolution du système d'équations (2.4), il n'existe aucun pavé* $[\mathbf{x}]$ *vérifiant* $[\mathbf{f}]([\mathbf{x}]) = \mathbf{0}$*, à moins que la fonction* $\mathbf{f}$ *ne s'annule sur des domaines entiers, aussi le sous-pavage* $\mathbb{X}^-$ *restera-t-il vide. Par conséquent, la démarche est simplifiée puisqu'on se limite à trouver le sous-pavage* $\Delta\mathbb{X}$*. Il suffit seulement de déterminer si* $\mathbf{0}$ *appartient ou non à* $[\mathbf{f}]([\mathbf{x}])$*, ce qui peut s'écrire :*

$$\mathbf{0} \in [\mathbf{f}]([\mathbf{x}]) \Leftrightarrow \forall i, 0 \in [f_i]([\mathbf{x}]) = \left[f_i^-, f_i^+\right] \Leftrightarrow \forall i, f_i^- \leq 0 \leq f_i^+.$$

Comme le montre cette dernière remarque, le problème des systèmes d'équations non linéaires de la forme (2.4) ne représente donc pas le meilleur exemple d'application de SIVIA puisque nous perdons notamment l'intérêt du sous-pavage $\mathbb{X}^-$. Dans les chapitres 3 et 4, nous verrons que pour des problèmes complexes, l'algorithme SIVIA s'avère très efficace. Précisons toutefois, étant donné la simplicité de l'algorithme, que la taille des problèmes à résoudre doit rester modeste, sous peine d'explosion des calculs.
Un solveur plus complet, basé sur SIVIA est téléchargeable à l'adresse http://www.istia.univ-angers.fr/~ dao/.

2.3 Optimisation globale non linéaire

Nous allons présenter dans ce paragraphe, une autre utilisation possible de l'analyse par intervalles. Nous venons de voir que les techniques d'analyse par intervalles permettent de résoudre le problème d'inversion ensembliste et en particulier des systèmes d'équations non linéaires, nous allons maintenant nous intéresser à l'utilisation du calcul par intervalles pour le problème plus délicat de l'optimisation globale non linéaire d'une fonction réelle. Avant de présenter un algorithme permettant de répondre au problème, nous précisons les termes employés par la suite.

2.3.1 Introduction

Le problème consiste à minimiser une fonction scalaire $f : \mathbb{R}^n \to \mathbb{R}$, appelée *fonction coût* lorsque les variables $(x_1, x_2, \ldots, x_n)$ décrivent un pavé $[\mathbf{x}_0]$ de $\mathbb{R}^n$. Le problème d'optimisation globale est le suivant :

$$\text{Minimiser } f(\mathbf{x}) \text{ pour } \mathbf{x} \in [\mathbf{x}_0]. \tag{2.7}$$

Bien entendu, plusieurs valeurs distinctes peuvent correspondre au minimum de la fonction coût, chacune de ces solutions est appelée *minimiseur global* de f.

2.3.2 Algorithme

Nous détaillons ici l'algorithme qui a été présenté par Skelboe dans [83] puis raffiné dans [46]. Nous avons nommé OPTICRIT cet algorithme qui permet de trouver le minimum global de la fonction f, noté $f^\star$. Bien entendu, ce minimum peut être atteint en différents points.

Soient $[f]$ une fonction d'inclusion pour f et ε une valeur (aussi petite que l'on veut) qui fixe la précision de l'encadrement. On recherche les pavés de $\mathbb{IR}^n$ dont la taille est inférieure à ε, sur lesquels f est susceptible d'atteindre son minimum.

Commençons par préciser les notations qui seront utilisées.

Notation 2.39. *Soit* $[\mathbf{x}]$ *un intervalle quelconque de* $\mathbb{R}^n$.
- On note $\mathbf{x}_c$ *la valeur du centre de ce pavé.*
- On note $f_{\mathbf{x}}^+$ *(resp.* $f_{\mathbf{x}}^-$*) la borne supérieure (resp. la borne inférieure) du pavé* $[f]([\mathbf{x}])$. *Ainsi, on a* $[f]([\mathbf{x}]) = [f_{\mathbf{x}}^-, f_{\mathbf{x}}^+]$.

2.3.2.1 Principe

On considère le pavé initial de recherche $[\mathbf{x}_0]$. On calcule l'image intervalle $\left[f_{\mathbf{x}_0}^-, f_{\mathbf{x}_0}^+\right]$ de $[\mathbf{x}_0]$ puis, on évalue la valeur de l'image $[f](\mathbf{x}_{0c})$ du centre de l'intervalle $[\mathbf{x}_0]$. La valeur $f_{\mathbf{x}_{0_c}}$ correspond à une borne maximum du minimum $f^\star$ de la fonction f.

On bissecte ensuite la boîte $[\mathbf{x}_0]$ en deux boîtes $[\mathbf{x}_1]$ et $[\mathbf{x}_2]$ puis, on itère le même procédé pour ces deux nouvelles boîtes. On calcule leur image respective par intervalle puis les images de leur centre, ce qui donne éventuellement une nouvelle borne maximum pour $f^\star$ (si $f^\star > f_{\mathbf{x}_{i_c}}$). Ensuite, on supprime les boîtes $[\mathbf{x}_i]$ dont la borne minimum de l'image $f^-_{\mathbf{x}_i}$ est supérieure à la borne maximum de $f^\star$.

On bissecte les boîtes restantes et on réitère le procédé jusqu'à ce que les pavés restants soient de tailles inférieures à la précision recherchée ε.

2.3.2.2 L'algorithme OPTICRIT

Entrées

$[f]$	:	fonction d'inclusion pour f.
$[\mathbf{x}_0]$	:	pavé de recherche.
ε	:	seuil de largeur des pavés résultats.

Sorties

$F^\star$	:	borne supérieure de $f^\star$.
$\mathcal{L}$	:	liste des pavés succeptibles de contenir un minimiseur global de f (en attente d'être traités).

Algorithme $\mathcal{L} := \{[\mathbf{x}_0]\}$, $F^\star = f^+_{\mathbf{x}_{0_c}}$ (initialisation)

Tant que $\mathcal{L}$ contient un pavé dont la taille est supérieure à ε **faire**

Sortir le plus grand pavé $[\mathbf{x}]$ de $\mathcal{L}$
$[f^-_{\mathbf{x}}, f^+_{\mathbf{x}}] := [f]([\mathbf{x}])$

Si $f^-_{\mathbf{x}} \leq F^\star$, alors $\mathcal{L} := \mathcal{L} \cup [\mathbf{x}]$
$F^\star := \min(F^\star, f^+_{\mathbf{x}_{1_c}})$
Eliminer tous les pavés $[\mathbf{x}]$ de $\mathcal{L}$ tels que $f^-_{\mathbf{x}} > F^\star$
Fin Si
Bissecter ce pavé $[\mathbf{x}]$ en deux sous-pavés $[\mathbf{x}_1]$ et $[\mathbf{x}_2]$
Ajouter $[\mathbf{x}_1]$ et $[\mathbf{x}_2]$ dans la liste $\mathcal{L}$

Fin tant que

Cet algorithme permet d'estimer le minimum $f^\star$ de f sur le pavé d'étude $[\mathbf{x}_0]$. A la fin de cet algorithme, la liste $\mathcal{L}$ contient tous les pavés $[\mathbf{x}]$ qui vérifient, d'une part, $w([\mathbf{x}]) \leq \varepsilon$ et d'autre part, $f^-_{\mathbf{x}} \leq F^\star \leq f^+_{\mathbf{x}}$. Or, par définition, le minimum global de la fonction f vérifie pour tout intervalle $[\mathbf{x}]$, $f^\star \leq F^\star \leq f^+_{\mathbf{x}}$, nous en déduisons que

$$\min_{\mathbf{x} \in \mathcal{L}} f^-_{\mathbf{x}} \leq f^\star \leq F^\star.$$

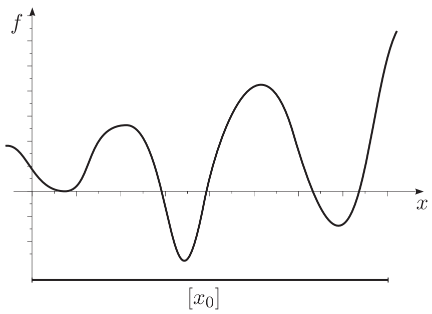
Nous avons donc bien obtenu un encadrement du minimum global. Tout pavé restant dans la liste $\mathcal{L}$ est susceptible de contenir un minimiseur de f. Attention, le nombre d'éléments de $\mathcal{L}$ ne correspond pas au nombre de minimiseurs de f car un pavé restant dans $\mathcal{L}$ peut ne pas contenir de minimiseur de f ou bien en contenir plusieurs.
Le résultat sera plus ou moins intéressant selon la fonction d'inclusion choisie pour f et la précision ε recherchée.

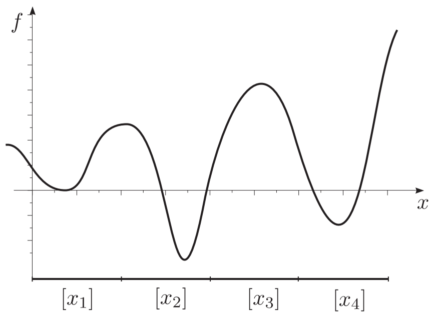
Remarque 2.40. *Notons que l'avantage de l'utilisation du calcul ensembliste pour résoudre ce type de problème d'optimisation est la certitude d'avoir obtenu un minimum global contrairement aux méthodes de gradient (comme la méthode de Newton) qui ne donnent, dans la meilleur des cas, qu'un minimum local de la fonction.*

2.3.2.3 Illustration

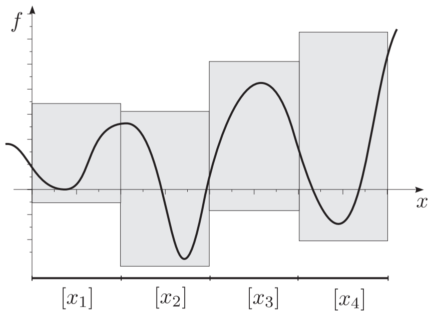
Pour finir, nous présentons une illustration graphique de l'algorithme OPTICRIT.

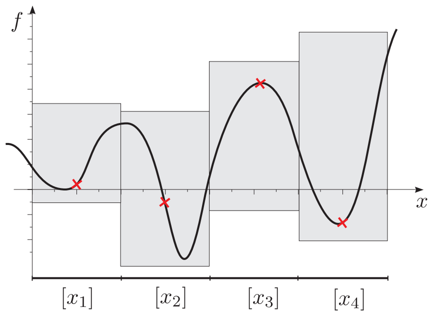
- On considére la fonction f, représentée ci-dessous, à optimiser sur l'intervalle $[x_0]$.



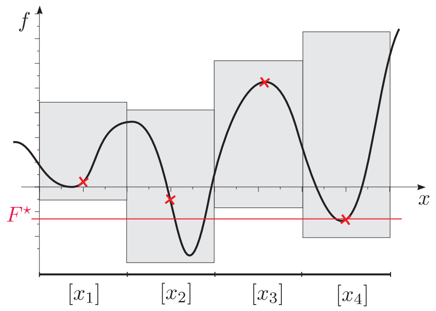


- On considère directement un sous-pavage de $[x_0]$ constitué de quatre pavés $[x_1]$, $[x_2]$, $[x_3]$ et $[x_4]$.
- On calcule l'image de chacun des pavés $[x_i]$ par la fonction d'inclusion $[f]$.



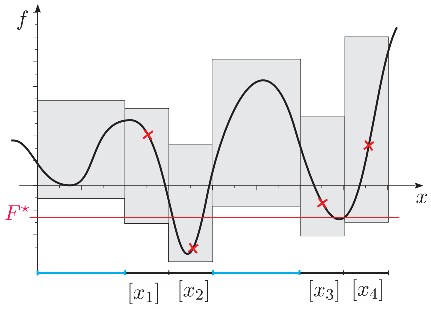


- On évalue l'image du centre pour chaque pavé $[x_i]$.
- On obtient ainsi une borne maximum $F^\star$ du minimum de la fonction f.

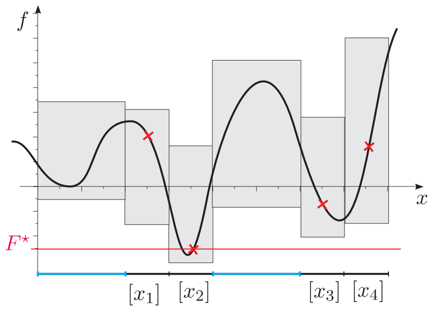


- On "supprime" les intervalles $[x_1]$ et $[x_3]$ dont les bornes inférieures des images sont supérieures à $F^\star$.

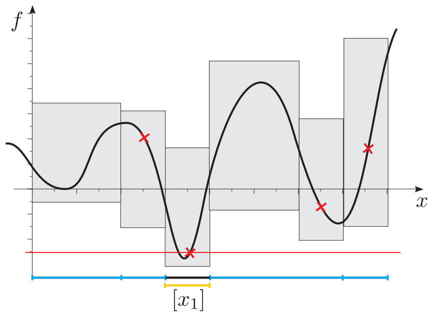
- On bissecte les deux intervalles restants $[x_2]$ et $[x_4]$.



- On obtient quatre nouveaux intervalles $[x_1]$, $[x_2]$, $[x_3]$ et $[x_4]$ dont on évalue l'image du centre.



- On obtient une nouvelle borne maximum $F^{\star}$ du minimum de la fonction f.
- On "supprime" tous les intervalles dont les bornes inférieures des images sont supérieures à $F^{\star}$.



- L'intervalle $[x_1]$ étant le seul intervalle restant dont la borne inférieure est inférieur à $F^{\star}$, il contient à coup sûr un minimiseur global de f.
- En itérant le procédé, on peut obtenir un intervalle aussi précis que l'on veut contenant un minimiseur de f.

CHAPITRE 3

Séparation aveugle de sources

Dans ce chapitre, nous nous intéressons au problème de séparation aveugle de sources. Il est constitué de deux sections principales totalement indépendantes l'une de l'autre.

La première s'intéresse au cas du mélange linéaire statique et illustre les avantages des techniques par intervalles (et notamment l'algorithme SIVIA) sur l'étude de la stabilité de l'algorithme Hérault-Jutten.

Le seconde section présente une méthodologie de résolution du problème de séparation de sources pour des systèmes de mélanges inversibles décrits par des équations différentielles. Nous introduisons notamment la notion de séparabilité.

3.1 Étude de la stabilité de l'algorithme de Hérault-Jutten

Dans cet exemple, nous proposons une application de l'algorithme SIVIA à un problème de séparation de sources. Ceci a fait l'objet d'une publication dans [89].

Ici, nous nous intéressons à la séparation du mélange linéaire instantané tel qu'il a été formulé pour la première fois par Jutten. Nous utilisons un critère de déconvolution non linéaire afin d'exploiter l'indépendance des entrées (voir paragraphe 1.3.1.2 page 19).

3.1.1 Énoncé du problème

On considère le réseau de neurones de Hérault-Jutten présenté sur la figure 3.1 où les observations y_1 et y_2 sont des combinaisons linéaires des sources u_1 et u_2, c'est-à-dire :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{3.1}$$

Nous supposons que la matrice de mélange $\mathbf{A} = (a_{ij})_{i,j}$ est régulière et que les sources u_1 et u_2 sont stationnaires, indépendantes et de moyennes nulles (hypothèse classique en séparation de sources), c'est-à-dire

$$E(u_1) = E(u_2) = 0.$$

Les sorties $\hat{u}_1$ et $\hat{u}_2$ du système (voir figure 3.1), estimations respectives des sources u_1 et u_2, vérifient

$$\begin{cases} \hat{u}_1 = y_1 - w_{12}\hat{u}_2 \\ \hat{u}_2 = y_2 - w_{21}\hat{u}_1 \end{cases} \tag{3.2}$$

où w_{12} et w_{21} sont les poids adaptatifs du réseau qui doivent être déterminés de sorte que les deux conditions suivantes soient vérifiées :

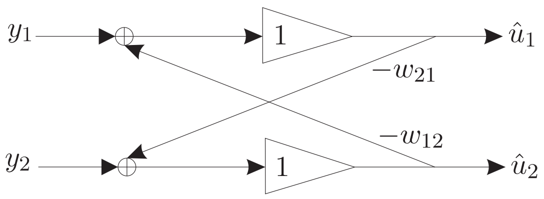


Figure 3.1 – Réseaux de neurones Hérault-Jutten.

i) Chaque sortie $\hat{u}_i$ est proportionnelle à une seule source u_i.

ii) Les sorties $\hat{u}_1$ et $\hat{u}_2$ sont indépendantes.

En remplaçant, dans le système (3.2), y_1 et y_2 par leurs expressions en fonction de u_1 et u_2 (voir équation (3.1)), nous obtenons un nouveau système

$$\left\{ \begin{array}{l} \hat{u}_1 = a_{11}u_1 + a_{12}u_2 - w_{12}\hat{u}_2 \\ \hat{u}_2 = a_{21}u_1 + a_{22}u_2 - w_{21}\hat{u}_1 \end{array} \right. .$$

Nous pouvons donc exprimer $\hat{u}_1$ et $\hat{u}_2$ en fonction des sources u_1 et u_2 :

$$\left\{ \begin{array}{l} \hat{u}_1 = \frac{1}{1-w_{12}w_{21}} \left(\left(a_{11} - w_{12}a_{21} \right) u_1 + \left(a_{12} - w_{12}a_{22} \right) u_2 \right) \\ \hat{u}_2 = \frac{1}{1-w_{21}w_{12}} \left(\left(a_{21} - w_{21}a_{11} \right) u_1 + \left(a_{22} - w_{21}a_{12} \right) u_2 \right) \end{array} \right. . \tag{3.3}$$

L'indépendance des signaux u_1 et u_2 implique qu'ils sont décorrélés, c'est-à-dire

$$E\left(u_1u_2\right) = 0,$$

et donc

$$E\left(\hat{u}_1\hat{u}_2\right) = 0.$$

Cependant, la décorrélation n'est pas suffisante pour trouver les w_{12} et w_{21} (voir paragraphe 1.3.3 page 25).

Par ailleurs, l'indépendance des sorties $\hat{u}_1$ et $\hat{u}_2$ implique que pour toutes fonctions f et g arbitraires, on a

$$E\left(f\left(\hat{u}_1\right) g\left(\hat{u}_2\right)\right) = 0.$$

Par exemple, nous pouvons prendre les fonctions f et g suivantes :

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ u & \longrightarrow & u^3, \end{array} \right.$$

$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ u & \longrightarrow & u. \end{array} \right.$$

A partir de ces résultats, Hérault et Jutten ont proposé de déterminer les coefficients w_{12} et w_{21} de sorte que

$$\left\{ \begin{array}{l} E\left(\hat{u}_1^3\hat{u}_2\right) = 0 \\ E\left(\hat{u}_2^3\hat{u}_1\right) = 0. \end{array} \right. \tag{3.4}$$

Ils proposent donc pour w_{12} et w_{21} la loi d'adaptation suivante :

$$\frac{\partial w_{ij}}{\partial t} = \mu\hat{u}_i^3\hat{u}_j,$$

où μ est une constante positive, et obtiennent alors un algorithme itératif estimant w_{12} et w_{21}.

3.1.2 Conditions de stabilité

Une analyse rigoureuse du réseau a été effectuée par Sorouchyari dans [84]. Il utilise la théorie de la stabilité de Lyapounov pour montrer qu'il existe quatre points d'équilibre stable pour l'algorithme proposé par Hérault et Jutten. En effet, Comon [30] et Sorouchyari [84] ont étudié les propriétés de convergence de cette méthode et déduit des conditions de stabilité dans le cas de deux sources u_1 et u_2 et de deux observations y_1 et y_2. Ils ont montré qu'un seul des quatre points stables est une solution acceptable pour la séparation de sources.

Pour que l'algorithme converge, il faut que les conditions suivantes, appelées *conditions d'acceptabilité*, soient vérifiées (voir [84] pour davantage de détails) :

$$w_{12}w_{21} > 1, \tag{3.5}$$

$$E\left(u_1^4\right) E\left(u_2^4\right) > 9\left(E\left(u_1^2 u_2^2\right)\right)^2. \tag{3.6}$$

Sans hypothèse *a priori* sur les sources, il semble alors difficile de prévoir la convergence ($\mathbf{u}$ étant inconnu). Nous allons voir que les techniques ensemblistes permettent d'obtenir à coup sûr la solution du problème en utilisant les conditions d'acceptabilité (3.5) et (3.6).

3.1.3 Simulation

Considérons les observations y_1 et y_2 générées à partir du modèle discret suivant

$$\begin{pmatrix} y_1(k) \\ y_2(k) \end{pmatrix} = \begin{pmatrix} 1 & 0.6 \\ 0.3 & 1 \end{pmatrix} \begin{pmatrix} u_1(k) \\ u_2(k) \end{pmatrix}, \tag{3.7}$$

où $u_1(k) = \sin(7,3kT_e)$, $u_2(k) = \sin(4kT_e)$ et $T_e = 0,1$ et ajoutons aux observations l'intervalle $\left[-10^{-2}, 10^{-2}\right]$ modélisant l'incertitude de mesure. Déterminons l'ensemble des coefficients (w_{12}, w_{21}) compatible avec la relation de stabilité (3.6).

En remplaçant $\hat{u}_1$ et $\hat{u}_2$ par leurs expressions respectives (voir (3.3)) dans le système (3.4), on obtient un système non linéaire de deux équations à deux inconnues w_{12} et w_{21} de la forme

$$\left\{ \begin{array}{l} f_1(w_{12}, w_{21}) = 0 \\ f_2(w_{12}, w_{21}) = 0 \end{array} \right. , \tag{3.8}$$

où $f_1(w_{12}, w_{21}) = E\left(\hat{u}_1^3 \hat{u}_2\right)$ et $f_2(w_{12}, w_{21}) = E\left(\hat{u}_2^3 \hat{u}_1\right)$.

Nous allons utiliser l'algorithme SIVIA pour le résoudre. La démarche est la suivante. Soient $[f_1]$ et $[f_2]$ des fonctions d'inclusion pour les fonctions f_1 et f_2. Notons que les fonctions f_1 et f_2 n'utilisent que des opérations standards d'arithmétique réelle, si bien que leurs fonctions d'inclusions naturelles s'obtiennent immédiatement.

Le problème consiste à résoudre le système d'équations

$$\left\{ \begin{array}{l} [f_1]([w_{12}], [w_{21}]) = 0 \\ [f_2]([w_{12}], [w_{21}]) = 0 \end{array} \right. \tag{3.9}$$

pour

$$[w_{12}] \times [w_{21}] \in [-5, 5]^2.$$

Après moins d'un dixième de seconde, les quatres zones correspondant aux quatres points d'équilibre de l'algorithme de Hérault et Jutten sont déterminées avec une précision fixée $\varepsilon = 10^{-2}$. La figure 3.2 illustre le résultat obtenu dans le pavé de recherche $[-5, 5] \times [-5, 5]$.

Figure 3.2 – Zones solutions pour $([w_{12}], [w_{21}])$

Maintenant, afin d'éliminer les 3 solutions inacceptables, nous allons exploiter les conditions (3.5) et (3.6) d'acceptabilité. Nous posons :

$$f_3 : \left\{ \begin{array}{lcl} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (w_{12}, w_{21}) & \longrightarrow & E\left(\hat{u}_1^4\right) E\left(\hat{u}_2^4\right) - 9\left(E\left(\hat{u}_1^2 \hat{u}_2^2\right)\right)^2. \end{array} \right.$$

Il est clair que si $f_3\left(w_{12}, w_{21}\right) > 0$, ce qui correspond à $E\left(u_1^4\right) E\left(u_2^4\right) > 9\left(E\left(u_1^2 u_2^2\right)\right)^2$, alors la condition (3.6) est vérifiée.

On sait que w_{12} et w_{21} doivent vérifier le système (3.8) ainsi que les deux conditions (3.5) et (3.6) de stabilité. Le seul couple solution acceptable est donc le couple (w_{12}, w_{21}) qui vérifie le système suivant

$$\left\{ \begin{array}{l} f_1\left(w_{12}, w_{21}\right) = 0 \\ f_2\left(w_{12}, w_{21}\right) = 0 \\ w_{12} w_{21} > 1 \\ f_3\left(w_{12}, w_{21}\right) > 0. \end{array} \right. \tag{3.10}$$

Nous utilisons à nouveau l'algorithme SIVIA pour résoudre la version intervalle de 3.10, soit

$$\left\{ \begin{array}{l} [f_1]\left([w_{12}], [w_{21}]\right) = 0 \\ [f_2]\left([w_{12}], [w_{21}]\right) = 0 \\ [w_{12}]\,[w_{21}] \in [1, +\infty[\\ [f_3]\left([w_{12}], [w_{21}]\right) \in [0, +\infty[, \end{array} \right. \tag{3.11}$$

avec

$$[w_{12}] \times [w_{21}] \in [-5, 5]^2$$

et $[f_1]$, $[f_2]$ et $[f_3]$ des fonctions d'inclusion naturelles des f_i.

Au bout d'un dixième de seconde, on obtient le résultat illustré par la figure 3.3. On voit qu'il y a une unique zone solution correspondant à la séparation du mélange.

Figure 3.3 – Zone solution pour $([w_{12}], [w_{21}])$

3.1.4 Conclusion

L'analyse par intervalles permet l'étude de la stabilité de l'algorithme de Hérault et Jutten dans un contexte à erreur bornée. Les avantages des techniques par intervalles, et en particulier SIVIA, sont multiples :

- Nous n'avons pas besoin de choisir de valeurs initiales pour w_{12} et w_{21}.
- La convergence de la méthode est assurée.
- Nous obtenons l'ensemble des solutions possibles pour les coefficients w_{12} et w_{21} (aucune solution n'est perdue!).
- L'incertitude ε d'estimation des coefficients w_{12} et w_{21} est connue (ici $\varepsilon = 10^{-2}$).

L'inconvénient est que les conditions de stabilité sont difficiles à calculer, même dans le cas de deux entrées - deux sorties. En général, on a rarement accès à de telles conditions. Toutefois, en ajoutant des équations supplémentaires d'indépendance (comme par exemple $E(u_1 u_2) = 0$), on peut éliminer aisément les solutions inacceptables.

3.2 Séparation des systèmes de mélanges inversibles

Dans cette section, nous présentons une démarche générale pour séparer des mélanges de signaux. Nous restreignons l'étude de séparation de sources à la classe particulière des systèmes de mélange qualifiés d'inversibles que l'on retrouve abondamment dans la nature. L'idée forte sera d'utiliser les dérivées de signaux aléatoires afin de séparer ces mélanges décrits par des équations différentielles. Nos motivations sont nées de l'automatique où certaines méthodes utilisent les dérivées pour reconstruire l'état d'un système [39].
L'objectif de cette section est de répondre à la question suivante : Est-il possible, moyennant des hypothèses statistiques, de retrouver les sources d'un système inversible?
A notre connaissance, la séparation aveugle de tels systèmes régis par des équations différentielles n'a pas été étudiée auparavant. Nous proposons ici une méthode pour résoudre ce type de problème.

Dans un premier temps, nous précisons la notion de système inversible. Ensuite, nous présentons les hypothèses statistiques que nous considérons sur les signaux. Elles nous permettent d'obtenir des informations (appelées *équations d'estimation* [20]) sur le système de mélange.

Dans un second temps, nous évoquons la relation d'équivalence de conservation de la forme d'onde dans le but de définir la notion de séparabilité d'un système inversible. Notre approche est largement inspirée de [88].

Ensuite, nous nous consacrons à la séparation proprement dite des systèmes de mélanges inversibles. Des résultats (présentés en annexe B) concernant les dérivées de signaux aléatoires, à la fois peu connus et peu utilisés, sont exploités.
Nous commençons par la séparation des systèmes de mélanges linéaires instantanés (une sous-classe des systèmes inversibles). Pour ces derniers, des équations d'estimation basées sur les dérivées successives de la matrice d'autocorrélation des sorties sont utilisées [22]. Nous illustrons cette méthode de séparation à l'ordre deux sur différentes simulations dans lesquelles nous utilisons des algorithmes de calcul par intervalles, présentés dans le chapitre 2, afin d'obtenir des résultats probants.
Puis, nous terminons cette section en étendant la méthode proposée à la séparation des systèmes de mélanges linéaires inversibles. Nous exploitons successivement les dérivées puis les décalages temporels de la matrice d'autocorrélation des sorties afin de résoudre le problème de séparation de sources pour ce type de mélange.

Tout au long de cette section, nous insistons particulièrement sur le cas particulier des systèmes de mélanges inversibles linéaires. Cependant, pour terminer, nous évoquons, en guise de perspective, le cas des mélanges non linéaires.

3.2.1 Introduction

Rappelons que le problème de séparation aveugle de sources consiste à retrouver des sources connaissant un mélange de celles-ci, appelé observations.
Nous considérons uniquement la classe des modèles de mélange décrits par des systèmes inversibles. Un système inversible de degré relatif r s'écrit sous la forme

$$\mathbf{u}\left(t\right) = \psi\left(\mathbf{y}\left(t\right), \dot{\mathbf{y}}\left(t\right), \ldots, \mathbf{y}^{(r-1)}\left(t\right), \mathbf{y}^{(r)}\left(t\right)\right), \tag{3.12}$$

où $t \in \mathbb{R}$ est le temps, $\mathbf{u} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$ est le vecteur d'entrées (ou sources), $\mathbf{y}^{(i)} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$ est la i-ième dérivée du vecteur de sorties $\mathbf{y}$ (ou observations) et $\psi : \left(\mathbb{R}^{\mathbb{R}}\right)^{n(r+1)} \mapsto \left(\mathbb{R}^{\mathbb{R}}\right)^n$ est une fonction analytique.

Exemple 3.1. $\mathbf{y}\left(t\right) = \dot{\mathbf{u}}\left(t\right)$ *n'est pas un système inversible.*

Exemple 3.2. $\mathbf{u}(t) = \mathbf{y}(t) + \dot{\mathbf{y}}(t)$ *est un système inversible linéaire de degré relatif un.*

L'avantage des systèmes inversibles est qu'ils modélisent une grande classe de systèmes physiques et sont donc très présents dans la nature. Notamment, tous les systèmes plats[1] sont des systèmes inversibles.

Remarque 3.3. *Les mélanges linéaires statiques, classiquement rencontrés en séparation de sources, sont des systèmes inversibles. En effet, un mélange linéaire statique est modélisé par un système inversible linéaire de degré relatif nul.*

Le problème de séparation aveugle de sources qui nous intéresse dans ce chapitre est le suivant :
On considère le système inversible de degré relatif r décrit par (3.12), c'est-à-dire

$$\mathbf{u}(t) = \psi\left(\mathbf{y}(t), \dot{\mathbf{y}}(t), \ldots, \mathbf{y}^{(r-1)}(t), \mathbf{y}^{(r)}(t)\right),$$

où le vecteur de sources $\mathbf{u}(t)$ et la fonction ψ sont inconnus. Seules les observations $\mathbf{y}(t)$ sont connues.

Le vecteur $\mathbf{u}(t)$ étant inconnu, nous allons lui associer un vecteur de signaux aléatoires $\mathbf{u}(.)$. Ainsi, dorénavant $\mathbf{u}(t)$ est une réalisation de $\mathbf{u}(.)$. Dès lors, $\mathbf{y}(t)$ sera également associé à un vecteur de signaux aléatoires $\mathbf{y}(.)$ de sorte que les observations $\mathbf{y}(t)$ correspondent à une réalisation de $\mathbf{y}(.)$. Ainsi, la fonction ψ associe aux réalisations $\mathbf{y}^{(i)}(t)$ des signaux aléatoires $\mathbf{y}^{(i)}(.)$, $i \in \{0, \ldots, r\}$, la réalisation $\mathbf{u}(t)$ de $\mathbf{u}(.)$.

Ainsi, notre problème consiste, à partir des observations $\mathbf{y}(t)$, à estimer les sources $\mathbf{u}(t)$. Cependant, en remarquant que la connaissance de ψ permet, d'après l'égalité (3.12), de déduire $\mathbf{u}(t)$, nous préférons en pratique l'estimation de la fonction ψ. Le problème de séparation aveugle de sources consiste dès lors à estimer la fonction inconnue ψ. Pour cela nous posons les hypothèses suivantes sur les signaux.

3.2.1.1 Hypothèses sur les signaux

3.2.1.1.1 Sur les signaux de sorties Nous nous plaçons dans un environnement idéal où il n'existe aucun bruit de mesure. Ainsi, le vecteur de sorties $\mathbf{y}(t)$ et ses dérivées $\mathbf{y}^{(i)}(t)$ sont connus. L'annexe E propose (de façon non exhaustive) différentes méthodes d'estimation des dérivées d'un signal. Notons toutefois que, même dans le cadre idéal où $\mathbf{y}(t)$ et ses dérivées sont connues, il n'existe, à notre connaissance, aucune méthode générale pour estimer $\mathbf{u}(t)$.

[1]Dans [40], Fliess définit un système plat comme suit :
Le système dynamique

$$\begin{cases} \dot{\mathbf{x}}(.) &= \mathbf{f}(\mathbf{x}(.), \mathbf{u}(.)), \\ \mathbf{y}(.) &= \mathbf{g}(\mathbf{x}(.)), \end{cases}$$

où $\mathbf{x} \in (\mathbb{R}^{\mathbb{R}})^m$, $\mathbf{u} \in (\mathbb{R}^{\mathbb{R}})^n$ et $\mathbf{y} \in (\mathbb{R}^{\mathbb{R}})^n$ sont respectivement les vecteurs d'état, d'entrée et de sortie, est plat , avec la sortie plate $\mathbf{y}$, si il existe des fonctions $\mathbf{h} : \mathbb{R}^m \times (\mathbb{R}^n)^{r+1} \mapsto \mathbb{R}^n$, $\phi : (\mathbb{R}^n)^r \mapsto \mathbb{R}^m$ et $\psi : (\mathbb{R}^n)^{r+1} \mapsto \mathbb{R}^n$ telles que

$$\begin{aligned} \mathbf{y} &= \mathbf{h}(\mathbf{x}, \mathbf{u}, \dot{\mathbf{u}}, \ldots, \mathbf{u}^{(r)}) \\ \mathbf{x} &= \phi(\mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r-1)}) \\ \mathbf{u} &= \psi(\mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r-1)}, \mathbf{y}^{(r)}). \end{aligned} \tag{3.13}$$

Il existe de nombreux modèles décrivant le fonctionnement (partiel) de l'avion à décollage vertical (PVTOL), la voiture, le bateau à voile (voir [68] et [66] pour davantage d'exemples) qui sont des systèmes plats. Et, d'après l'égalité (3.13), on voit aisément que tous les systèmes plats sont des systèmes inversibles.

3.2.1.1.2 Sur les signaux d'entrées Afin de résoudre le problème de séparation de sources, nous émettons des hypothèses H_i sur le vecteur de signaux $\mathbf{u}(.)$. L'ensemble $\mathbb{M} = \{\mathrm{H}_i\}_i$ de ces hypothèses est appelé *modèle d'entrée* pour $\mathbf{u}(.)$. Par la suite, on notera abusivement

$$\mathbf{u}(t) \in \mathbb{M}.$$

Ainsi, le problème de séparation de sources consiste alors à estimer la fonction ψ telle que

$$\psi\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right) \in \mathbb{M}. \tag{3.14}$$

Le modèle d'entrée $\mathbb{M} = \{\mathrm{H}_0, \mathrm{H}_1, \mathrm{H}_2\}$ que nous considérons est le suivant :

H_0) Le vecteur d'entrées est à composantes non nulles.

H_1) Le vecteur d'entrées $\mathbf{u}(.)$ appartient à $\mathcal{S}^n$, où $\mathcal{S}$ est l'ensemble des signaux aléatoires stationnaires, ergodiques et lisses[2].
Par conséquent, $\dot{\mathbf{u}}(.)$,$\ddot{\mathbf{u}}(.)$,$\ldots, \mathbf{u}^{(k)}(.), \ldots$ appartiennent également à $\mathcal{S}^n$. Notons que cette hypothèse n'implique pas que le vecteur d'observations $\mathbf{y}(.)$ appartienne à $\mathcal{S}^n$. Par exemple, si le système (3.12) est instable, $\mathbf{y}(.)$ est non stationnaire et ne peut donc pas appartenir à $\mathcal{S}^n$.

H_2) Le vecteur d'entrées $\mathbf{u}(.)$ est à composantes indépendantes, c'est-à-dire $I_s(u_1, \ldots, u_n)$, ou encore $\mathbf{u} \in \mathcal{I}^n$.

De ces hypothèses sur les signaux d'entrées, on déduit des relations pour la fonction ψ.

3.2.1.1.3 La fonction ψ D'après le lemme B.46 page 181, les hypothèses H_1 et H_2 du modèle d'entrées $\mathbb{M}$ peuvent s'écrire

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R},\ \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)\right) = \mathbf{0}, \tag{3.15}$$

où $\mathbf{\Gamma}^{(k)}(\tau)$ est la dérivée k-ième de la matrice d'autocorrélation et $\mathbf{d}(.)$ est une fonction qui associe à une matrice le vecteur constitué de tous ses éléments non diagonaux. Et, d'après l'expression de $\mathbf{u}$ dans l'équation (3.12), on a

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R},\ \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\psi\left(\mathbf{y}(t),\ldots,\mathbf{y}^{(r)}(t)\right),\psi\left(\mathbf{y}(t),\ldots,\mathbf{y}^{(r)}(t)\right)}(\tau)\right) = \mathbf{0}. \tag{3.16}$$

D'après la proposition B.28 page 173, on en déduit les équations suivantes

$$\begin{gathered}\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R},\\ \mathbf{d}\left(E\left(\psi^{(k)}\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)\psi\left(\mathbf{y}(t-\tau), \ldots, \mathbf{y}^{(r)}(t-\tau)\right)^{\mathrm{T}}\right) - \mu_{\mathbf{u}^{(k)}}\mu_{\mathbf{u}}^{\mathrm{T}}\right) = \mathbf{0}\end{gathered} \tag{3.17}$$

avec $\mu_{\mathbf{u}} = E\left(\psi\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)\right)$ et $\mu_{\mathbf{u}^{(k)}} = E\left(\psi^{(k)}\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)\right)$.
Les équations (3.17), qui lient les signaux de sorties et leurs dérivées connues, doivent être vérifiées par la fonction ψ. Par analogie avec [20], nous les appelons *équations d'estimation*. Remarquons toutefois que, sans connaissance *a priori* sur la fonction ψ, ces informations ne peuvent être exploitées. Nous allons donc approximer la fonction ψ .

[2]Un signal aléatoire lisse est un signal aléatoire dont la i-ième dérivée (en moyenne quadratique) par rapport à t est définie pour tout $i \in \mathbb{N}$.

3.2.1.2 Approximation de ψ

Sachant que, par définition d'un système inversible, la fonction ψ est analytique, elle peut être approximée par son développement de Taylor à l'ordre q que l'on note $\hat{\psi}\left(\mathbf{y},\ldots,\mathbf{y}^{(r)}\right)$.
Par exemple, pour un système inversible de degré relatif nul ($r = 0$) composé de deux entrées et deux sorties ($n = 2$), l'approximation à l'ordre deux ($q = 2$) de ψ est donnée par

$$\hat{\psi}\left(\mathbf{y}\right) = \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + b_{11}y_1^2 + b_{12}y_1y_2 + b_{13}y_2^2 \\ a_{21}y_1 + a_{22}y_2 + b_{21}y_1^2 + b_{22}y_1y_2 + b_{23}y_2^2 \end{pmatrix} \tag{3.18}$$

où les coefficients a_{ij} et b_{ij} sont inconnus. A noter que l'entier q est choisi afin d'obtenir une estimation convenable de ψ.

Remarquons qu'ainsi définie, $\hat{\psi}\left(\mathbf{y},\ldots,\mathbf{y}^{(r)}\right)$ peut être vue comme une fonction paramétrée en $\mathbf{p}$ où les p_i sont les coefficients des matrices $\mathbf{A}_{i_1},\ldots,\mathbf{Q}_{i_1,\ldots,i_q}$. On peut donc noter

$$\hat{\psi}\left(\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) \equiv \hat{\psi}\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right).$$

Exemple 3.4. *Si l'ordre d'approximation q est égal à 1, le développement de Taylor $\hat{\psi}$ d'un système inversible de degré relatif r est une fonction linéaire de la forme*

$$\hat{\psi}\left(\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) = \sum_{i=0}^{r} \mathbf{A}_i\mathbf{y}^{(i)}, \tag{3.19}$$

où les matrices $\mathbf{A}_i$ sont inconnues. Notons qu'alors, la fonction $\hat{\psi}$ peut être vue comme une fonction paramétrée en $\mathbf{p}$ où les $p_{k_{i,j}}$ sont les coefficients des matrices $\mathbf{A}_k$.
Par exemple, pour $n = 1$ et $\mathbf{y} \in \mathcal{S}^2$, on a

$$\hat{\psi}\left(\mathbf{y},\dot{\mathbf{y}}\right) = \underbrace{\begin{pmatrix} p_{0_{11}} & p_{0_{12}} \\ p_{0_{21}} & p_{0_{22}} \end{pmatrix}}_{\mathbf{A}_0}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \underbrace{\begin{pmatrix} p_{1_{11}} & p_{1_{12}} \\ p_{1_{21}} & p_{1_{22}} \end{pmatrix}}_{\mathbf{A}_1}\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix}.$$

Ainsi, $\hat{\psi}\left(\mathbf{y},\dot{\mathbf{y}}\right)$ est bien une fonction paramétrée et peut donc être notée $\hat{\psi}\left(\mathbf{p},\mathbf{y},\dot{\mathbf{y}}\right)$ avec

$$\mathbf{p} = \left(p_{0_{11}},\ldots,p_{0_{22}},p_{1_{11}},\ldots,p_{1_{22}}\right) \in \mathbb{R}^8.$$

En conclusion, le problème de séparation aveugle de sources consiste maintenant à estimer le vecteur de paramètres $\mathbf{p}$ tel que

$$\hat{\psi}\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) \in \mathbb{M}.$$

Avant de présenter la méthodologie de résolution, nous introduisons la notion de séparabilité d'un système inversible $\hat{\psi}\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right)$.

3.2.2 Séparabilité d'un système inversible

Dans ce paragraphe, nous allons présenter la notion de séparabilité (structurelle) pour un système inversible. Nous nous sommes inspirés des travaux de Tong [88] qui a défini, pour des mélanges linéaires instantanés, cette notion sous le nom d'identifiabilité. Nous l'avons volontairement renommée *séparabilité* dans la mesure où, en automatique, l'identifiabilité a une signification autre que nous serons amenés à utiliser dans le chapitre 4 (voir paragraphe 4.2.1 page 95).

La séparabilité est une propriété qui caractérise la possibilité de résoudre le problème de séparation de sources pour un système de mélange et un modèle d'entrées donnés. Il s'agit de déterminer par avance, avant même la récolte des observations, si les paramètres de la fonction $\hat{\psi}$ peuvent être identifiés en supposant que les entrées $\mathbf{u}$ suivent le modèle $\mathbb{M}$. Avant de définir plus en détails cette notion non triviale, nous introduisons quelques résultats généraux.

3.2.2.1 Généralités

Considérons le système inversible de degré relatif r et paramétré suivant :

$$\mathbf{u}(t) = \hat{\psi}\left(\mathbf{p}, \mathbf{y}(t), \dot{\mathbf{y}}(t), \ldots, \mathbf{y}^{(r)}(t)\right), \tag{3.20}$$

où $\mathbf{p} \in \mathbb{R}^{n_p}$ est le vecteur de n_p paramètres. On appelle $\mathbb{R}^{n_p}$, également noté $\mathcal{P}$, l'*espace d'identification* du système (3.20). Afin de simplifier, on note

$$\psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \hat{\psi}\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right),$$

ainsi $\psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ est une approximation de $\psi\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$.

Commençons par remarquer de façon évidente qu'un élément quelconque $\mathbf{p}'$ de l'espace d'identification $\mathcal{P}$ n'est pas une bonne estimation de $\mathbf{p}$ dans la mesure où $\mathbf{u}'(t) = \psi\left(\mathbf{p}', \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)$ ne possède pas obligatoirement les mêmes caractéristiques que $\mathbf{u}(t) = \psi\left(\mathbf{p}, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)$, telles les formes d'ondes ou certaines propriétés statistiques.

Comme nous l'avons déjà précisé dans le chapitre 1, dans la pratique, nous souhaitons seulement extraire la forme d'onde des signaux d'entrées. Le gain des sources ainsi que leur numéro d'ordre importent peu. Ainsi, si $\mathbf{D} \in \mathfrak{D}$ est une matrice diagonale inversible et $\mathbf{P} \in \mathfrak{P}$ est une matrice de permutation, les signaux $\mathbf{u}'(t) = \mathbf{PDu}(t)$ peuvent être considérés comme étant une estimation satisfaisante des sources $\mathbf{u}(t)$. Cette notion peut se formaliser par la construction d'une relation d'équivalence sur $\mathcal{P}$, dite de c*onservation de la forme d'onde*.

Définition 3.5 (Relation de conservation de la forme d'onde)**.** Soient $\mathbf{p}_1$,$\mathbf{p}_2 \in \mathcal{P}$ et, pour tout $i \in \{1, 2\}$,

$$\mathbf{u}_i(t) = \psi\left(\mathbf{p}_i, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right).$$

$\mathbf{p}_1$ et $\mathbf{p}_2$ vérifient la relation de conservation de la forme d'onde, notée $\mathbf{p}_1 \sim \mathbf{p}_2$ (ou encore $\mathbf{u}_1 \sim \mathbf{u}_2$), si on a

$$\psi\left(\mathbf{p}_1, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right) = \mathbf{PD}\psi\left(\mathbf{p}_2, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right), \tag{3.21}$$

où $\mathbf{D} \in \mathfrak{D}$ est une matrice diagonale inversible et $\mathbf{P} \in \mathfrak{P}$ est une matrice de permutation. Cette relation de conservation de la forme d'onde s'écrit encore

$$\mathbf{u}_1(t) = \mathbf{PDu}_2(t),$$

d'où la notation abusive $\mathbf{u}_1 \sim \mathbf{u}_2$.

Proposition 3.6. *La relation de conservation de la forme d'onde est une relation d'équivalence sur $\mathcal{P}$.*

Démonstration. Soient $\mathbf{p}_1$,$\mathbf{p}_2$ et $\mathbf{p}_3$ trois éléments de $\mathcal{P}$.
- Reflexivité : $\mathbf{p}_1 \sim \mathbf{p}_1$ est trivial.
- Symétrie : $\mathbf{p}_1 \sim \mathbf{p}_2 \implies \mathbf{u}_1 = \mathbf{PDu}_2 \implies \mathbf{D}^{-1}\mathbf{P}^T\mathbf{u}_1 = \mathbf{u}_2 \implies \mathbf{p}_2 \sim \mathbf{p}_1.$
- Transitivité : $\begin{cases} \mathbf{p}_1 \sim \mathbf{p}_2 \\ \mathbf{p}_2 \sim \mathbf{p}_3 \end{cases} \implies \begin{cases} \mathbf{u}_1 = \mathbf{PDu}_2 \\ \mathbf{u}_2 = \mathbf{P}'\mathbf{D}'\mathbf{u}_3 \end{cases} \implies \mathbf{u}_1 = \mathbf{PDP}'\mathbf{D}'\mathbf{u}_3$
$\implies \mathbf{u}_1 = \mathbf{P}''\mathbf{D}''\mathbf{u}_3 \implies \mathbf{p}_1 \sim \mathbf{p}_3.$ □

Intéressons-nous maintenant aux propriétés statistiques de la relation $\sim$.

Proposition 3.7. *Soient $\mathbf{p}_1$ et $\mathbf{p}_2$ deux éléments de $\mathcal{P}$ tels que $\mathbf{p}_1 \sim \mathbf{p}_2$. Soient $u_{1,1}(t), \ldots, u_{1,n}(t)$ et $u_{2,1}(t), \ldots, u_{2,n}(t)$ les composantes des signaux $\mathbf{u}_1(t) = \psi\left(\mathbf{p}_1, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)$ et $\mathbf{u}_2(t) = \psi\left(\mathbf{p}_2, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)$. On a les équivalences suivantes :*
$i)$ Les signaux aléatoires $u_{1,1}, \ldots, u_{1,n}$ sont mutuellement indépendants si et seulement si les signaux aléatoires $u_{2,1}, \ldots, u_{2,n}$ sont mutuellement indépendants.
$ii)$ Pour tout $\tau \in \mathbb{R}$, $\mathbf{\Gamma}_{\mathbf{u}_1\mathbf{u}_1}(\tau)$ est diagonale si et seulement si $\mathbf{\Gamma}_{\mathbf{u}_2\mathbf{u}_2}(\tau)$ est diagonale.

Démonstration. $i)$ On sait que $\mathbf{u}_1(t) = \mathbf{PDu}_2(t)$ avec $\mathbf{D} \in \mathfrak{D}$, $\mathbf{P} \in \mathfrak{P}$. Or, les matrices diagonales et les matrices de permutations sont des transformations préservant l'indépendance (voir proposition B.39 page 178). Donc, $\mathbf{u}_1(t) \in I^n$si et seulement si $\mathbf{u}_2(t) = \mathbf{PDu}_1(t) \in I^n$.

$ii)$ On sait que $\mathbf{u}_1(t) = \mathbf{PDu}_2(t)$ avec $\mathbf{D} \in \mathfrak{D}$ et $\mathbf{P} \in \mathfrak{P}$, donc

$$\mathbf{\Gamma}_{\mathbf{u}_2\mathbf{u}_2}(\tau) = \mathbf{PD\Gamma}_{\mathbf{u}_1\mathbf{u}_1}(\tau)\,\mathbf{D}^{-1}\mathbf{P}^T.$$

Par conséquent, si $\mathbf{\Gamma}_{\mathbf{u}_1\mathbf{u}_1}(\tau)$ est diagonale alors $\mathbf{\Gamma}_{\mathbf{u}_2\mathbf{u}_2}(\tau)$ est diagonale, et inversement. □

Grâce à la relation d'équivalence de conservation de la forme d'onde, nous pouvons quotienter l'espace $\mathcal{P}$ par la relation $\sim$. On note $\mathcal{P}/\sim$. La séparation aveugle du système (3.20) consiste donc à identifier un représentant $\mathbf{p}'$ de la classe d'équivalence à laquelle appartient $\mathbf{p}$. Par suite, si $\mathbf{p}' \sim \mathbf{p}$, c'est-à-dire $\mathbf{p}'$ est une estimation de $\mathbf{p}$, alors $\mathbf{u}'(t)$ est une estimation de $\mathbf{u}(t)$ ($\mathbf{u}' \sim \mathbf{u}$).

Nous allons maintenant définir précisément la notion de séparabilité. Comme nous l'avons évoqué, elle est inspirée des travaux de Tong [88]. Toutefois, la définition que nous proposons dans le paragraphe suivant nous paraît plus appropriée dans la mesure où Tong considère à la fois l'espace paramétrique $\mathcal{P}$ et les sources $\mathbf{u}(t)$ pour définir la séparabilité alors que seule la considération de l'ensemble $\mathcal{P}$ semble suffir. De notre point de vue, notre démarche est plus pédagogique. De surcroît, nous constaterons, dans le chapitre suivant, que la séparabilité ainsi définie peut être vue comme une "extension" de l'identifiabilité structurelle (au sens de l'automatique classique).

3.2.2.2 Séparabilité

Nous supposons que les sources $\mathbf{u}(t)$ suivent un modèle d'entrées $\mathbb{M}$ quelconque (on note $\mathbf{u}(t) \in \mathbb{M}$). La séparabilité (structurelle) d'un système inversible est définie comme suit.

Définition 3.8 (Séparabilité)**.** Soit $\mathbb{M}$ un modèle d'entrées quelconque. Pour $\mathbf{y}$ fixé, le système inversible

$$\mathbf{u}(t) = \psi\left(\mathbf{p}, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)$$

est dit (structurellement) séparable selon $\mathbb{M}$ si, pour presque tout $\mathbf{p} \in \mathcal{P}$,

$$\begin{cases} \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\mathbf{p}', \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{cases} \implies \mathbf{p}' \sim \mathbf{p}. \tag{3.22}$$

ou de façon équivalente

$$\left\{\begin{array}{l}\mathbf{u}(t)\in\mathbb{M}\\ \mathbf{u}'(t)\in\mathbb{M}\end{array}\right. \Longrightarrow \mathbf{u}'(t)\sim\mathbf{u}(t).$$

En d'autres termes, un système inversible est séparable selon le modèle d'entrée $\mathbb{M}$ s'il suffit de trouver $\mathbf{p}'$ tel que $\psi\left(\mathbf{p}',\mathbf{y}\left(t\right),\ldots,\mathbf{y}^{(r)}\left(t\right)\right)\in\mathbb{M}$ pour que $\mathbf{u}'\left(t\right)$ soit une estimation de $\mathbf{u}\left(t\right)$ à une permutation et un gain près.

Il est important de noter que la séparabilité est définie relativement à un modèle d'entrées $\mathbb{M}$ et à un système de mélange ψ.

Remarque 3.9. *Nous reviendrons plus en détails, dans le paragraphe 4.2.1, sur la notation "pour presque tout* $\mathbf{p}\in\mathcal{P}$*". Nous pouvons d'ores et déjà préciser qu'un système peut être séparable même si tous les vecteurs de paramètres* $\mathbf{p}$ *ne vérifient pas (3.22).*

Exemple 3.10. *Considérons un système inversible linéaire de degré relatif nul (système linéaire instantané) paramétré défini par*

$$\mathbf{u}\left(t\right)=\psi\left(\mathbf{p},\mathbf{y}\left(t\right)\right)=\mathbf{A}\mathbf{y}\left(t\right), \tag{3.23}$$

où $\mathbf{A}\in\mathbb{R}^{n\times n}$ *est une matrice dont tous les coefficients sont des paramètres et* $\mathbf{u}\left(t\right)\in\mathbb{M}=\{H_0\}$*, c'est-à-dire que le vecteur d'entrée* $\mathbf{u}\left(t\right)$ *est à composantes non nulles. L'espace d'identification du système (3.23) est équivalent à l'ensemble de toutes les matrices carrés de dimension* n *:*

$$\mathcal{P}=\left\{\mathbf{M}\in\mathbb{R}^{n\times n}\right\}.$$

Intéressons-nous à la séparabilité du système (3.23) selon le modèle d'entrées $\mathbb{M}=\{H_0\}$
Il est clair que toutes les matrices non nulles $\mathbf{M}$ *de* $\mathcal{P}$ *ne sont pas en relation avec* $\mathbf{A}$*. En effet, on peut trouver une matrice* $\mathbf{M}\in\mathcal{P}$*, non nulle (qui vérifie donc* $\mathbf{M}\mathbf{y}\in\mathbb{M}$*), telle que* $\mathbf{M}\not\sim\mathbf{A}$*. Le système (3.23) n'est donc pas séparable pour le modèle d'entrées* $\mathbb{M}=\{H_0\}$ *. D'où la nécessité d'ajouter des hypothèses sur les sources et la matrice de mélange pour rendre le système séparable.*

La notion de séparabilité est fondamentale en séparation de sources et sera à la base des travaux présentés dans le chapitre 4. Nous allons maintenant nous intéresser à des systèmes inversibles pour lesquels la fonction ψ caractérisant le mélange est approximée par son développement de Taylor à l'ordre un. Nous insisterons particulièrement sur le mélange linéaire instantané pour lequel nous présentons une méthode de séparation de sources à l'ordre deux [65] avant de la généraliser aux systèmes inversibles de degré relatif r quelconque.

3.2.3 Le mélange linéaire instantané

Dans cette partie, nous allons nous intéresser au cas des systèmes inversibles linéaires de degré relatif nul, plus communément appelés systèmes de *mélanges linéaires instantanés*. Dans ce contexte, nous verrons que notre démarche nous permet de retrouver les résultats présentées par Cavassilas, Xerri et Chabriel dans [22]. Ces systèmes s'écrivent sous la forme

$$\mathbf{u}\left(t\right)=\psi\left(\mathbf{p},\mathbf{y}\left(t\right)\right)=\mathbf{A}\mathbf{y}\left(t\right), \tag{3.24}$$

où $\mathbf{A}\in\mathbb{R}^{n\times n}$ est une matrice régulière paramétrée en $\mathbf{p}$, $\mathbf{u}\in\left(\mathbb{R}^{\mathbb{R}}\right)^n$ est le vecteur d'entrées inconnu et $\mathbf{y}$ est le vecteur de sorties connu. Soit $\mathbb{M}$ le modèle de signaux ergodiques, stationnaires, lisses et statistiquement indépendants défini au paragraphe 3.2.1.1.2. Nous supposons $\mathbf{u}\in\mathbb{M}$.

Le problème de séparation de sources consiste à estimer une matrice $\hat{\mathbf{A}}$ telle que $\hat{\mathbf{A}} = \mathbf{PDA}$ où $\mathbf{D} \in \mathfrak{D}$ et $\mathbf{P} \in \mathfrak{P}$. Afin de limiter le nombre de paramètres et conformément aux remarques du paragraphe 1.2 (page 16), nous cherchons une matrice $\hat{\mathbf{A}}$ de diagonale unitaire telle que

$$\hat{\mathbf{u}} = \hat{\mathbf{A}}\mathbf{y}(t) \in \mathbb{M}. \tag{3.25}$$

Dans le paragraphe suivant, nous allons établir des équations d'estimation, liées au modèle $\mathbb{M}$, permettant de calculer la matrice $\hat{\mathbf{A}}$.

3.2.3.1 Equations d'estimation

En considérant (3.15), les équations d'estimation du système (3.25) sont

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\hat{\mathbf{u}},\hat{\mathbf{u}}}(\tau)\right) = \mathbf{0}. \tag{3.26}$$

Pour simplifier les équations (3.26), nous utilisons uniquement l'influence des dérivées successives, *i.e.* on pose $\tau = 0$. Notons que l'utilisation des décalages temporels, déjà présentée dans le paragraphe 1.3.3 (page 25), est à la base de l'algorithme SOBI (Second Order Blind Identification) [6].
On obtient les équations d'estimation suivantes

$$\forall k \in \mathbb{N}, \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\hat{\mathbf{u}},\hat{\mathbf{u}}}(0)\right) = \mathbf{0}. \tag{3.27}$$

D'après la remarque B.30, les dérivées d'ordre impair de la matrice $\mathbf{\Gamma}_{\mathbf{uu}}(0)$ sont nulles et n'apportent donc aucune information. On a

$$\forall k \in \mathbb{N}, \mathbf{\Gamma}^{(2k+1)}_{\hat{\mathbf{u}},\hat{\mathbf{u}}}(0) = \mathbf{0}.$$

Nous réduisons les équations (3.27) à

$$\forall k \in \mathbb{N}, \mathbf{d}\left(\mathbf{\Gamma}^{(2k)}_{\hat{\mathbf{u}},\hat{\mathbf{u}}}(0)\right) = \mathbf{0},$$

ce qui équivaut à

$$\forall k \in \mathbb{N}, \mathbf{d}\left(\mathbf{\Gamma}^{(2k)}_{\hat{\mathbf{A}}\mathbf{y},\hat{\mathbf{A}}\mathbf{y}}(0)\right) = \mathbf{0}. \tag{3.28}$$

Le système de mélange étant linéaire statique, le vecteur de sorties $\mathbf{y}$ appartient à $\mathcal{S}^n$, donc

$$\forall k \in \mathbb{N}, \mathbf{d}\left(\hat{\mathbf{A}}\mathbf{\Gamma}^{(2k)}_{\mathbf{y},\mathbf{y}}(0)\,\hat{\mathbf{A}}^T\right) = \mathbf{0}, \tag{3.29}$$

où $\hat{\mathbf{A}}$ est une matrice (de diagonale unitaire) de $n(n-1)$ inconnues. L'entier k doit être choisi de façon à obtenir suffisamment d'équations (au moins autant que d'inconnues), en espérant que ces équations soient indépendantes et fournissent une solution. Ce point délicat reste encore à approfondir.

3.2.3.2 Méthode de séparation à l'ordre deux

Nous allons maintenant voir que les équations d'estimation obtenues permettent de résoudre le problème de séparation de sources. Nous proposons une méthode à l'ordre deux basée sur les dérivées des matrices d'autocorrélation des signaux.

Nous nous intéressons aux deux équations correspondant aux entiers $k = 0$ et $k = 1$ dans le système (3.29), c'est-à-dire,

$$\mathbf{d}\left(\hat{\mathbf{A}}\boldsymbol{\Gamma}_{\mathbf{y},\mathbf{y}}(0)\,\hat{\mathbf{A}}^{T}\right) = \mathbf{0}, \tag{3.30}$$

$$\mathbf{d}\left(\hat{\mathbf{A}}\boldsymbol{\Gamma}^{(2)}_{\mathbf{y},\mathbf{y}}(0)\,\hat{\mathbf{A}}^{T}\right) = \mathbf{0}. \tag{3.31}$$

En d'autres termes, la matrice $\hat{\mathbf{A}}$ diagonalise simultanément les matrices $\boldsymbol{\Gamma}_{\mathbf{y},\mathbf{y}}(0)$ et $\boldsymbol{\Gamma}^{(2)}_{\mathbf{y},\mathbf{y}}(0)$. Pour résoudre le problème, il est possible d'utiliser des algorithmes de diagonalisation conjointe d'un ensemble de matrices (à titre d'exemple, un algorithme est présenté dans l'annexe F) afin d'estimer la matrice $\hat{\mathbf{A}}$ ($\sim \mathbf{A}$). L'estimation des sources $\hat{\mathbf{u}}$ ($\sim \mathbf{u}$) est alors donnée par

$$\hat{\mathbf{u}} = \hat{\mathbf{A}}\mathbf{y}. \tag{3.32}$$

Bien entendu, la robustesse de l'estimation de $\hat{\mathbf{A}}$ peut être améliorée en ajoutant des équations supplémentaires faisant intervenir les dérivées des matrices d'autocorrélation d'ordres supérieurs (équations d'estimation (3.29) pour $k \geq 2$). Dans ce cas, l'estimation de $\hat{\mathbf{A}}$ peut être obtenue par diagonalisation conjointe des matrices $\boldsymbol{\Gamma}^{(2k)}_{\mathbf{yy}}(0)$ considérées [18] (l'annexe F présente un algorithme de diagonalisation conjointe).

3.2.3.2.1 Similitudes avec SOBI Par analogie avec la méthode présentée au paragraphe 1.3.3 (page 25), l'estimation de la matrice $\hat{\mathbf{A}}$ ci-dessus peut être obtenue en deux étapes distinctes qui correspondent aux résolutions successives des deux équations (3.30) et (3.31).

- Pour commencer, la contrainte imposée par l'équation (3.30) correspond précisément au blanchiment spatial des observations. Rappelons que ce dernier consiste à estimer une matrice $\mathbf{W}$ telle que
$$\mathbf{W}\boldsymbol{\Gamma}_{\mathbf{y},\mathbf{y}}(0)\,\mathbf{W}^{\mathrm{T}} = \mathbf{Id}_n.$$
Comme nous l'avons vu, la matrice solution est $\mathbf{W} = \mathbf{D}^{\frac{1}{2}}\mathbf{U}^{\mathrm{T}}$, avec $\mathbf{D}$ une matrice diagonale dont les éléments diagonaux sont les valeurs propres de $\boldsymbol{\Gamma}_{\mathbf{yy}}(0)$ et $\mathbf{U}$ une matrice orthogonale dont les colonnes correspondent aux vecteurs propres de $\boldsymbol{\Gamma}_{\mathbf{y},\mathbf{y}}(0)$. Soit $\mathbf{z} = \mathbf{W}\mathbf{y}$ le vecteur d'observations blanchies. Reste maintenant à estimer une matrice orthogonale $\mathbf{Q}$ vérifiant
$$\boldsymbol{\Gamma}^{(2)}_{\mathbf{z},\mathbf{z}}(0) = \mathbf{Q}\boldsymbol{\Gamma}^{(2)}_{\mathbf{u},\mathbf{u}}(0)\,\mathbf{Q}^{\mathrm{T}}.$$
- En exploitant la contrainte imposée par l'équation (3.31), c'est-à-dire $\boldsymbol{\Gamma}^{(2)}_{\mathbf{u},\mathbf{u}}(0)$ est une matrice diagonale, on trouve que la matrice $\mathbf{Q}$ est composée des vecteurs propres de $\boldsymbol{\Gamma}^{(2)}_{\mathbf{z},\mathbf{z}}(0)$. Et donc, de la même façon que dans le paragraphe 1.3.3, la matrice $\mathbf{Q}$ ne peut être estimée que si les valeurs propres de $\boldsymbol{\Gamma}^{(2)}_{\mathbf{uu}}(0)$ sont distinctes (voir remarque 1.18 page 27).

A la différence de la méthode SOBI qui exploite des décalages temporels de la matrice d'autocorrélation, la méthode proposée ici exploite les dérivées successives (d'ordre pair) de la matrice d'autocorrélation. Ainsi, une condition nécessaire de succès pour cette méthode est l'existence d'un ordre de dérivation k (non nul) tel que les dérivées k-ième des fonctions d'autocorrélation des sources en zéro ($\Gamma^{(2k)}_{u_i,u_i}(0)$) soient non nulles et distinctes, c'est-à-dire $\forall i, j \in \mathbb{N}$ tels que $i \neq j$,

$$\Gamma^{(2k)}_{u_i,u_i}(0) \neq \Gamma^{(2k)}_{u_j,u_j}(0). \tag{3.33}$$

Rappelons que la condition nécessaire de succès de SOBI est l'existence d'un retard τ tel que $\forall i, j \in \mathbb{N}$, $i \neq j$,

$$\Gamma_{u_i,u_i}(\tau) \neq \Gamma_{u_j,u_j}(\tau). \tag{3.34}$$

A ce stade de la compréhension, il est intéressant de se demander s'il existe des signaux séparables par SOBI et non séparables par la méthode proposée, et inversement. En d'autre termes, est-ce que les deux conditions (3.33) et (3.34) sont équivalentes? Nous n'avons pour le moment pas de réponse à cette question. Cependant, on peut remarquer que si les dérivées sont approximées par différences finies, on a pour $k = 1$,

$$\Gamma^{(2)}_{u_i,u_j}(0) = \Gamma_{u_i,u_j}(-1) - 2\Gamma_{u_i,u_j}(0) + \Gamma_{u_i,u_j}(1).$$

Ainsi, notre méthode revient à considérer des décalages temporels de la matrice d'autocorrélation (lorsque les dérivées sont estimées par différences finies). Toutefois, dans un contexte général, à la différence de la méthode SOBI valable pour des signaux à temps discret, notre approche est continue.

Remarque 3.11. *Notons également que la méthode présentée ici, basée sur les dérivées des matrices d'autocorrélation, peut s'étendre à des signaux non stationnaires. Pour cela, il convient d'exploiter des résultats sur les dérivées de signaux aléatoires (voir paragraphe B.4.3 page 170).*

Pour illustrer cette méthode, nous l'appliquons sur un exemple simple de séparation de deux sources.

3.2.3.2.2 Exemple Considérons deux mélanges linéaires statiques y_1 et y_2 de deux sources u_1 et u_2, décrits par le système suivant

$$\mathbf{u}(t) = \mathbf{A}\mathbf{y}(t), \tag{3.35}$$

où $\mathbf{u}(t) = (u_1(t), u_2(t))^{\mathrm{T}}$ et $\mathbf{y}(t) = (y_1(t), y_2(t))^{\mathrm{T}}$. L'estimation $\hat{\mathbf{A}}$ de la matrice $\mathbf{A}$, définie par

$$\hat{\mathbf{A}} = \begin{pmatrix} 1 & p_1 \\ p_2 & 1 \end{pmatrix},$$

doit être solution du système d'équations (3.30) et (3.31). Après développement, on obtient le système de deux équations non linéaires à deux inconnues p_1 et p_2 suivant

$$\begin{cases} p_2 E\left(y_1^2\right) + (1 + p_1 p_2) E(y_1 y_2) + p_1 E\left(y_2^2\right) \\ \quad - (E(y_1) + p_1 E(y_2))(p_2 E(y_1) + E(y_2)) = 0 \\ \\ p_2 E(\ddot{y}_1 y_1) + (1 + p_1 p_2) E(\ddot{y}_1 y_2) + p_1 E(\ddot{y}_2 y_2) = 0, \end{cases} \tag{3.36}$$

où les moments des sorties et leurs dérivées sont connus. Remarquons que des simplifications ont été réalisées compte tenu de la symétrie des matrices $\mathbf{\Gamma}^{(2k)}_{\mathbf{y},\mathbf{y}}(0)$.
La résolution formelle du système (3.36), basée sur la décomposition en base de Gröbner (voir annexe D), donne les deux solutions

$$\hat{\mathbf{A}}_1 = \begin{pmatrix} 1 & p_1 \\ p_2 & 1 \end{pmatrix} \text{ et } \hat{\mathbf{A}}_2 = \begin{pmatrix} 1 & \frac{1}{p_2} \\ \frac{1}{p_1} & 1 \end{pmatrix}, \tag{3.37}$$

où

$$p_1 = \frac{2\gamma}{\delta}, \tag{3.38}$$

$$p_2 = \frac{2\alpha}{\delta}, \tag{3.39}$$

avec

$$
\begin{aligned}
\alpha &= E\left(y_2^2\right) E\left(\ddot{y}_1 y_2\right) - E\left(y_1 y_2\right) E\left(\ddot{y}_2 y_2\right) \\
\beta &= E\left(y_1^2\right) E\left(\ddot{y}_2 y_2\right) - E\left(y_2^2\right) E\left(\ddot{y}_1 y_1\right) \\
\gamma &= E\left(y_1 y_2\right) E\left(\ddot{y}_1 y_1\right) - E\left(y_1^2\right) E\left(\ddot{y}_1 y_2\right) \\
\delta &= \beta + sign\left(\beta\right) \sqrt{\beta^2 - 4\alpha\gamma}.
\end{aligned}
$$

On peut remarquer que ces deux solutions vérifient $\hat{\mathbf{A}}_1 \sim \hat{\mathbf{A}}_2$ puisqu'elles sont toutes deux des estimations de $\mathbf{A}$.

Remarque 3.12. *Cette solution analytique pour $\hat{\mathbf{A}}$ est évidente puisque la diagonalisation conjointe de deux matrices carrées de dimension deux symétriques dont l'une est définie positive est toujours possible (voir interprétation géométrique dans le paragraphe F.4.3.1 de l'annexe F).*

3.2.3.3 Simulations

Dans ce paragraphe, nous allons nous intéresser à différents exemples numériques de séparation de sources. Dans un premier temps, nous appliquons simplement la méthode précédente exploitant les dérivées des signaux afin d'obtenir la séparation de mélanges non bruités. Ensuite, nous utilisons les techniques ensemblistes (SIVIA et OPTICRIT) pour mettre en évidence leur intérêt dans la pratique. Nous reprenons à cet effet le premier exemple, puis considérons le cas particulier d'un mélange non séparable.

3.2.3.3.1 Mélange non bruité de deux sources : solution analytique Dans ce premier exemple simple, les sources et les observations sont liées par un système linéaire statique.
Considérons deux sources gaussiennes indépendantes obtenues à partir du filtrage passe-bas du second ordre d'un bruit blanc gaussien. Le filtrage permet d'obtenir des signaux différentiables en préservant leur caractère gaussien. Ces deux signaux, tracés sur les figures 3.4 et 3.5 (9000 échantillons avec une période d'échantillonage $Te = 0.1$), suivent le modèle

$$
\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 2 & -0.8 \\ 0.9 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{3.40}
$$

La distribution conjointe des sources, représentée sur la figure 3.6, illustre l'indépendance des signaux $u_1(t)$ et $u_2(t)$. D'autre part, la distribution conjointe des dérivées des sources, représentée sur la figure 3.7, montre que les signaux $\dot{u}_1(t)$ et $\dot{u}_2(t)$ sont indépendants (comme énoncé dans la proposition B.45 page 180).

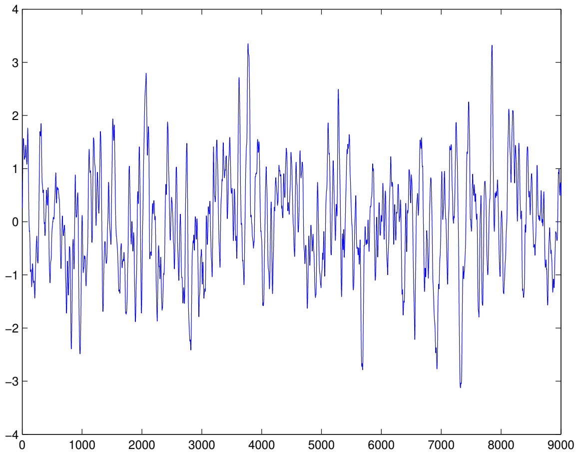


Figure 3.4 – Source $u_1(t)$.

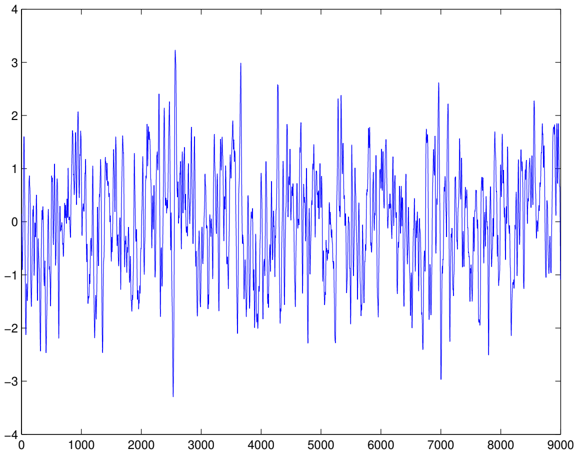


Figure 3.5 – Source $u_2(t)$.

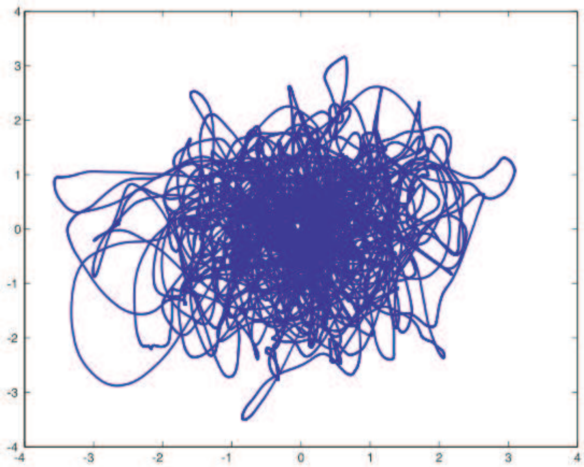


Figure 3.6 – Distribution conjointe des sources $u_1(t)$ et $u_2(t)$.

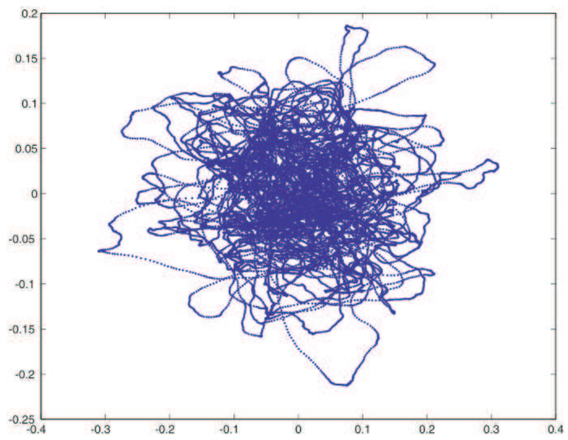


Figure 3.7 – Distribution conjointe des dérivées des sources $u_1(t)$ et $u_2(t)$.

Les signaux observés sont représentés sur les figures 3.8 et 3.9.

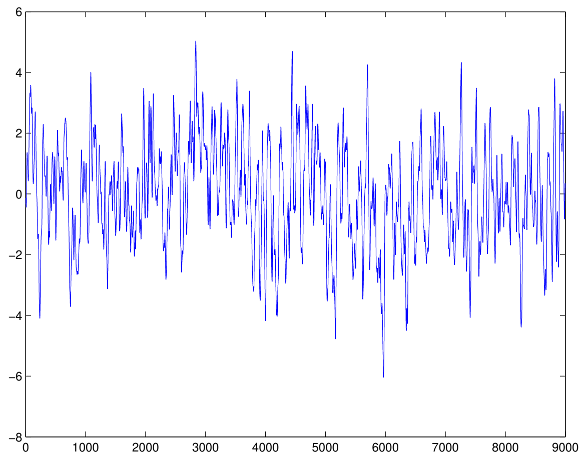


Figure 3.8 – Observation $y_1(t)$.

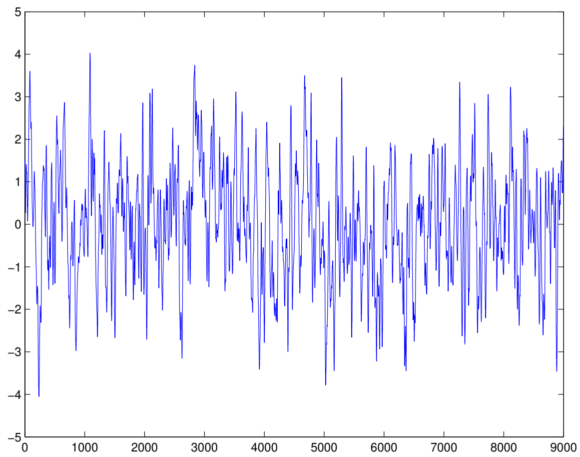


Figure 3.9 – Observation $y_2(t)$.

Les distributions conjointes de $y_1(t)$ et $y_2(t)$ (figure 3.10), respectivement de $\dot{y}_1(t)$ et $\dot{y}_2(t)$ (figure 3.11) illustrent la dépendance statistique des signaux mélangés et de leurs dérivées.

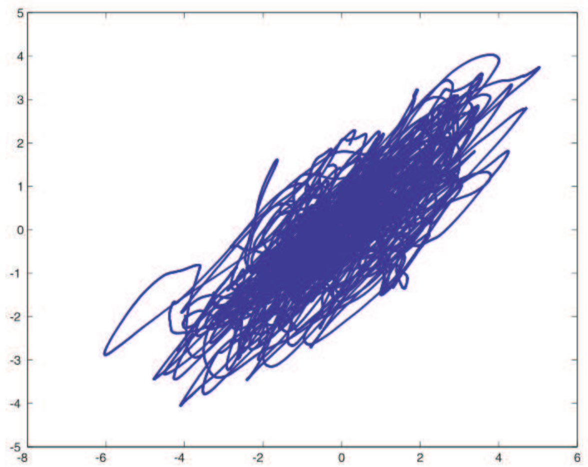


Figure 3.10 – Distribution conjointe des observations $y_1(t)$ et $y_2(t)$.

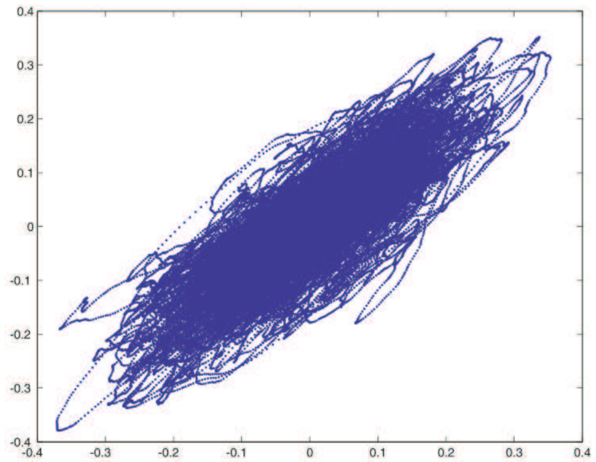


Figure 3.11 – Distribution conjointe des dérivées des observations $y_1(t)$ et $y_2(t)$.

D'après la relation (3.37), la solution analytique pour la matrice $\hat{\mathbf{A}}$ obtenue est

$$\hat{\mathbf{A}} = \begin{pmatrix} 1 & -0.411 \\ -0.295 & 1 \end{pmatrix}. \tag{3.41}$$

Vérifions que $\hat{\mathbf{A}} \sim \mathbf{A}$, c'est-à-dire qu'il existe des matrices $\mathbf{P} \in \mathfrak{P}$ et $\mathbf{D} \in \mathfrak{D}$ telles que $\hat{\mathbf{A}} = \mathbf{PDA}$. Pour cet exemple, si

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ et } \mathbf{D} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}, \tag{3.42}$$

on obtient

$$\mathbf{PDA} = \begin{pmatrix} 1 & -0.4 \\ -0.3 & 1 \end{pmatrix}.$$

Les paramètres estimés de la matrice $\hat{\mathbf{A}}$ sont $\hat{p}_1 = -0.411$ et $\hat{p}_2 = -0.295$ alors que les paramètres idéaux (calculés à partir de $\mathbf{A}$) sont $p_1^\star = -0.4$ et $p_2^\star = -0.3$. Les superpositions des sources estimées $\hat{\mathbf{u}}(t) = \hat{\mathbf{A}}\mathbf{y}(t)$ (pour $\hat{\mathbf{p}}$) et des sources réelles $\mathbf{u}^\star(t) = \mathbf{A}^\star\mathbf{y}(t)$ (pour $\mathbf{p}^\star$) sont tracées sur les figures 3.12 et 3.13.

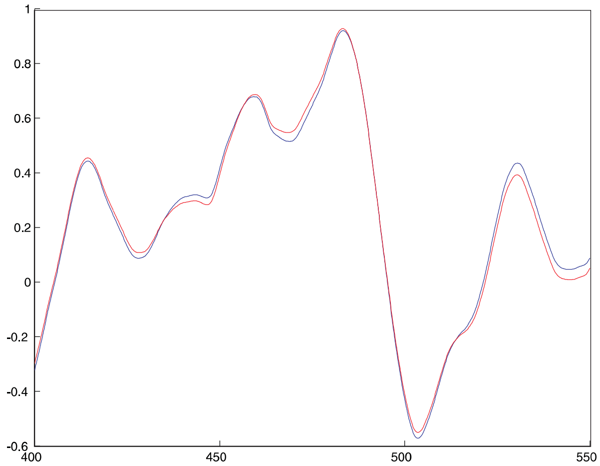


Figure 3.12 – Superposition de la source estimée $\hat{u}_1(t)$ et de la source réelle $u_1^\star(t)$.

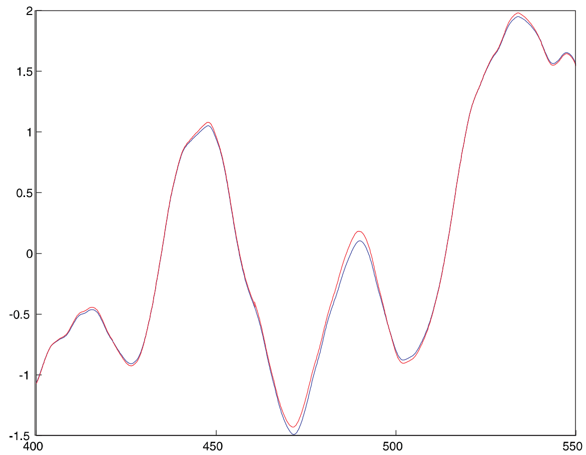


Figure 3.13 – Superposition de la source estimée $\hat{u}_2(t)$ et de la source réelle $u_2^\star(t)$.

Afin d'étudier la robustesse des solutions analytiques pour des mélanges linéaires gaussiens, la figure 3.14 montre l'erreur quadratique moyenne ($E(|\mathbf{u} - \hat{\mathbf{u}}|^2)$) en fonction du nombre d'échantillons, pour des signaux $\mathbf{u}$ et $\hat{\mathbf{u}}$ normés. Pour plus de 2600 échantillons, la solution analytique fournit une solution satisfaisante, avec une erreur inférieure à $-20dB$. Ce nombre important d'échantillons nécessaires n'est pas significatif, c'est une conséquence du sur-échantillonnage des signaux considérés (voir par exemple les figures 3.12 ou 3.13 pour constater le sur-échantillonnage).

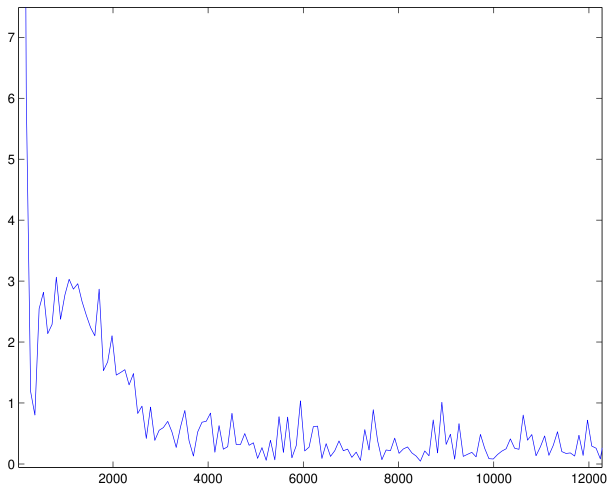


Figure 3.14 – Représentation de l'erreur d'estimation en fonction du nombre d'échantillons.

3.2.3.3.2 Méthode par intervalles

3.2.3.3.2.1 Mélange non bruité de deux sources Dans ce paragraphe, nous reprenons l'exemple précédent. Nous considérons donc le mélange décrit par l'équation (3.40) où les paramètres solutions sont $p_1^\star = -0.4$ et $p_2^\star = -0.3$. A la différence de l'exemple précédent, nous ne nous servons pas de la solution analytique (3.37) mais utilisons les outils ensemblistes du chapitre 2 pour résoudre le système d'équations (3.36).

Considérons un pavé de recherche initial

$$[p_1] \times [p_2] = [-5, 5] \times [-5, 5],$$

et deux fonctions d'inclusion $[f_1]$ et $[f_2]$ définies par

$$\begin{aligned} [f_1]([p_1],[p_2]) = \quad & [p_2]\, E\left(y_1^2\right) + (1 + [p_1][p_2])\, E(y_1 y_2) + [p_1]\, E\left(y_2^2\right) - \\ & (E(y_1) + [p_1]\, E(y_2))\,([p_2]\, E(y_1) + E(y_2)), \end{aligned} \tag{3.43}$$

$$[f_2]([p_1],[p_2]) = \quad [p_2]\, E(\ddot{y}_1 y_1) + (1 + [p_1][p_2])\, E(\ddot{y}_1 y_2) + [p_1]\, E(\ddot{y}_2 y_2). \tag{3.44}$$

Remarque 3.13. *Les fonctions d'inclusions $[f_1]$ et $[f_2]$ sont respectivement les fonctions d'inclusion naturelles correspondant aux deux équations de (3.36).*

Le système d'équations

$$\left\{ \begin{array}{l} [f_1]([p_1],[p_2]) = 0 \\ [f_2]([p_1],[p_2]) = 0 \end{array} \right. \tag{3.45}$$

est une version ensembliste du système d'équations (3.36) qu'il est possible de résoudre en utilisant l'algorithme d'inversion ensembliste SIVIA. En effet, il s'agit ici de caractériser l'ensemble

$$\mathbb{P} = \{(p_1, p_2)\, | \forall i \in \{1,2\}, [f_i]([p_1],[p_2]) \in [0,0]\}.$$

La figure 3.15 représente graphiquement un sous-pavage "extérieur" des solutions de (3.45) pour une précision $\varepsilon = 10^{-2}$.

Figure 3.15 – Solutions p_1 et p_2 (en clair), obtenues par SIVIA après un temps infime pour 2000 échantillons, dans le pavé de recherche $[-5, 5] \times [-5, 5]$.

Les deux zones solutions correspondent aux intervalles solutions

$$[p_1] \times [p_2] = [-0.42, -0.38] \times [-0.31, -0.27], \tag{3.46}$$

et

$$[p_1] \times [p_2] = [-2.46, -2.24] \times [-3.44, -3.30]. \tag{3.47}$$

Remarque 3.14. *A la place de l'algorithme SIVIA, la résolution du système (3.45) peut être effectuée à l'aide de l'algorithme d'optimisation globale OPTICRIT (voir paragraphe 2.3 page 49). La démarche consiste alors à minimiser la fonction coût suivante :*

$$([f_1]([p_1], [p_2]))^2 + ([f_2]([p_1], [p_2]))^2. \tag{3.48}$$

Les intervalles solutions obtenus sont alors identiques.
Supposons que l'hypothèse d'indépendance des sources ne soit pas exactement vérifiée, c'est-à-dire que $[f_1]([p_1], [p_2]) \neq 0$ *et* $[f_2]([p_1], [p_2]) \neq 0, \forall [\mathbf{p}] \in \mathbb{IR}^2$*, l'algorithme SIVIA trouve l'ensemble vide comme solution alors que l'algorithme d'optimisation donne l'ensemble solution le plus satisfaisant (celui qui minimise la fonction coût (3.48)). Les moments étant estimés, les égalités (3.45) peuvent ne pas être vérifiées et dans ce cas l'algorithme OPTICRIT est plus avantageux que SIVIA car il nous fournit les solutions les plus acceptables (selon le critère de minimisation).*

L'avantage d'utiliser des techniques par intervalles pour résoudre ce type de problème est minime. Il est simplement possible de contrôler la précision du résultat via le paramètre ε. La remarque suivante illustre un avantage supplémentaire des techniques par intervalles.

Remarque 3.15. *Sur les estimations des moments* $E(y_1^2),\ E(y_1 y_2), \ldots$ *du système (3.45), il est possible de prendre en compte des incertitudes de la forme*

$$\begin{array}{l} E(y_1^2) \in \left[E(y_1^2) - \epsilon^+, E(y_1^2) + \epsilon^-\right] \\ E(y_1 y_2) \in \left[E(y_1 y_2) - \epsilon^+, E(y_1 y_2) + \epsilon^-\right] \\ \ldots \end{array}$$

Ces incertitudes permettent, par exemple, de modéliser la présence de bruit.
Notons $\left[E(y_1^2)-\epsilon^+, E(y_1^2)+\epsilon^-\right]=[E](y_1^2)$, $[E(y_1y_2)-\epsilon^+, E(y_1y_2)+\epsilon^-]=[E](y_1y_2)$, ...
Ainsi

$$\begin{array}{l} E(y_1^2)\in[E](y_1^2) \\ E(y_1y_2)\in[E](y_1y_2) \\ \ldots \end{array}$$

Une telle prise en compte du bruit conduit, de même que précédemment, à résoudre le système d'équations 3.45 où, cette fois ci, les deux fonctions $[f_1]$ *et* $[f_2]$ *sont définies par :*

$$\begin{aligned} [f_1]([p_1],[p_2]) = & \; [p_2][E]\left(y_1^2\right)+(1+[p_1][p_2])[E](y_1y_2)+[p_1][E]\left(y_2^2\right)- \\ & ([E](y_1)+[p_1][E](y_2))([p_2][E](y_1)+[E](y_2)), \end{aligned} \tag{3.49}$$

$$[f_2]([p_1],[p_2]) = [p_2][E](\ddot{y}_1y_1)+(1+[p_1][p_2])[E](\ddot{y}_1y_2)+[p_1][E](\ddot{y}_2y_2). \tag{3.50}$$

La figure 3.16 illustre l'ensemble des solutions pour une incertitude sur les moments fixée à $\epsilon_i^+=\epsilon_i^-=10^{-2}$.

Figure 3.16 – Ensemble des paramètres solutions dans le pavé de recherche $[-10,10]\times[-10,10]$.

Un solveur, développé en C++, basé sur le calcul par intervalles (SIVIA ou OPTICRIT au choix), est disponible à l'adresse http://www.istia.univ-angers.fr/˜ lagrange/. Il résoud le système (3.45) pour différents choix de sources (cosinus, sinus, bruit blanc gaussien, uniforme, etc.).

Pour mieux se rendre compte de l'intérêt des intervalles, considérons un exemple légèrement différent.

3.2.3.3.2.2 Mélange non bruité de deux sources : Un cas particulier Considérons le problème de séparation de sources pour un mélange de deux sources gaussiennes indépendantes de fonctions d'autocorrélation identiques. Rappelons que pour ce modèle d'entrée, le mélange n'est pas séparable à l'ordre deux (voir paragraphe 1.4 page 28). Les deux sources, tracées sur les figures 3.17 et 3.18 (5000 échantillons avec une période d'échantillonnage $T_e=0.1$), de fonction d'autocorrélation représentées sur la

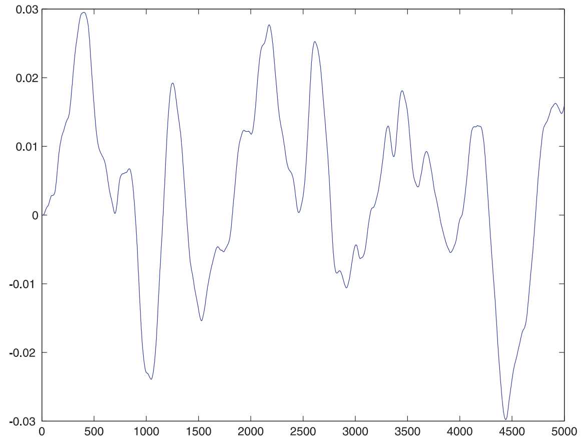


Figure 3.17 – Source $u_1\left(t\right)$.

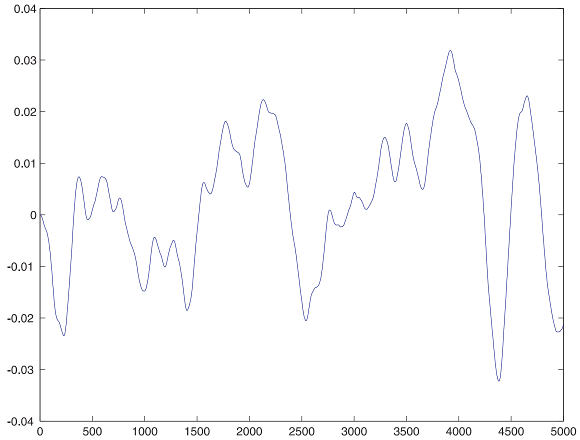


Figure 3.18 – Source $u_2\left(t\right)$.

figure 3.19 suivent le mélange (3.40) rappelé ci-dessous :

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 2 & -0.8 \\ 0.9 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \tag{3.51}$$

Les signaux observés $y_1\left(t\right)$ et $y_2\left(t\right)$ et leurs dérivées $\dot{y}_1\left(t\right)$ et $\dot{y}_2\left(t\right)$, représentés respectivement sur

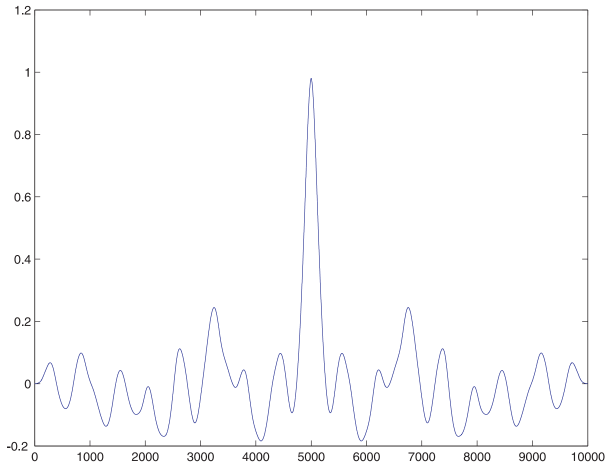


Figure 3.19 – Fonction d'autocorrélation des sources $u_1\left(t\right)$ et $u_2\left(t\right)$.

les figures 3.20, 3.21, 3.22 et 3.23, permettent d'estimer les coefficients du système d'équations (3.45). La figure 3.24 illustre les solutions du système (3.45). On voit qu'il existe une infinité de paramètres solutions et donc que le problème de séparation de sources ne peut être résolu (du moins sans l'ajout d'équations supplémentaires). Contrairement à une technique utilisant la diagonalisation conjointe qui aurait donné une seule et unique solution, l'avantage des techniques par intervalles est d'obtenir l'ensemble des paramètres solutions. On peut ainsi se rendre compte qu'un mélange n'est pas séparable.

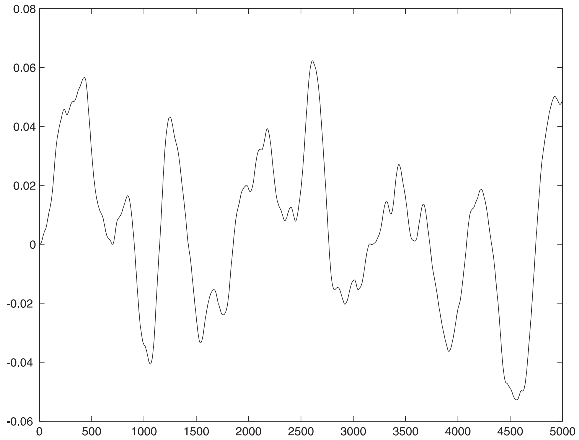


Figure 3.20 – Observation $y_1(t)$.

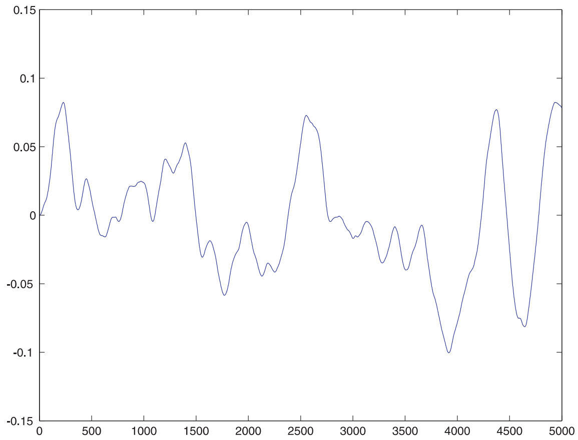


Figure 3.21 – Observation $y_2(t)$.

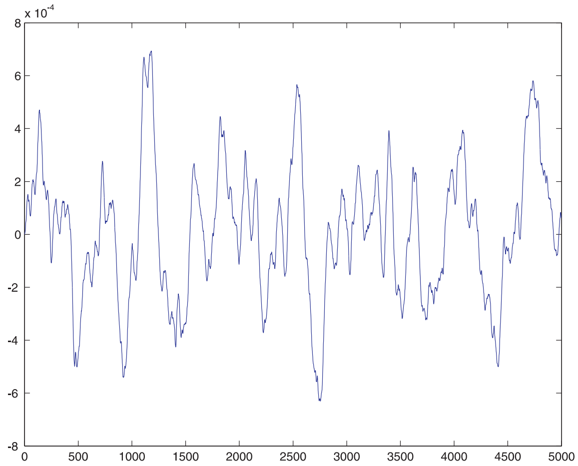


Figure 3.22 – Signal $\dot{y}_1(t)$.

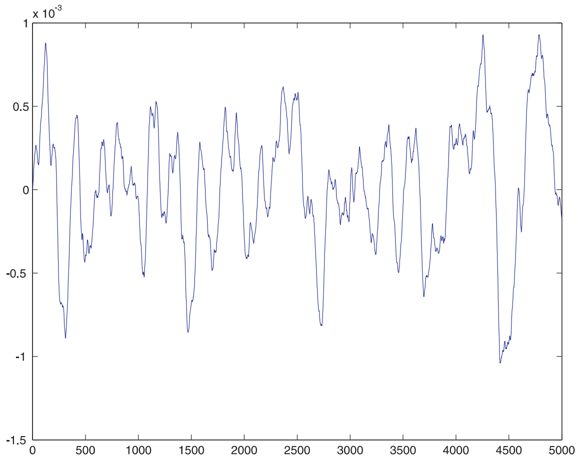


Figure 3.23 – Signal $\dot{y}_2(t)$.

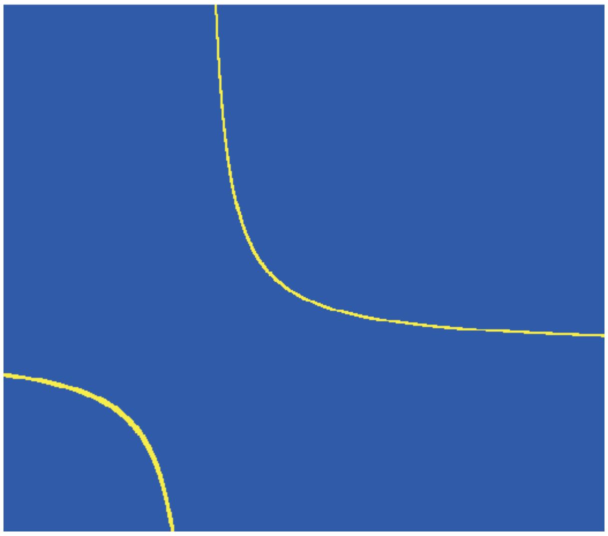


Figure 3.24 – Ensemble des paramètres solutions du système (3.45) dans le pavé de recherche $[-5,5] \times [-5,5]$.

3.2.3.4 Synthèse

En conclusion, nous avons proposé un critère de séparation basé sur les matrices d'autocorrélation des observations et de leurs dérivées. Cette méthode exploitant uniquement des statistiques du deuxième ordre, la séparation de sources gaussiennes est possible. Les seules hypothèses (faibles) sont que les sources inconnues soient différentiables (voir paragraphe B.4 page 168) et qu'il existe un ordre de dérivation k tel que les dérivées k-ième des fonctions d'autocorrélations des sources soient distinctes en zéro.

Dans le cas d'un mélange de deux sources (paragraphe 3.2.3.3.1), on a trouvé une solution analytique cependant, comme toute solution analytique, elle est sensible à l'estimation des moments et donc au bruit. Afin de surmonter ce problème (pour deux sources) et dans le cas général (pour n sources et n observations), la solution peut être estimée en utilisant les algorithmes de diagonalisation conjointe des matrices $E\left(\mathbf{y}\mathbf{y}^T\right)$ et $E\left(\ddot{\mathbf{y}}\mathbf{y}^T\right)$. De surcroît, des matrices d'autocorrélation supplémentaires, basées sur des dérivées d'ordres supérieures ou utilisant différents décalages temporels (en supposant que les sources soient colorées), peuvent être ajoutées afin d'estimer la solution par diagonalisation conjointe.

Dans un second temps (paragraphe 3.2.3.3.2), nous avons présenté des méthodes de résolutions par intervalles qui ont l'avantage majeur de fournir l'ensemble des paramètres solutions. Ainsi, dans un contexte à erreurs bornées, le problème de séparation de source peut être résolu . De plus, il est également possible de se rendre compte de la non séparabilité du mélange de sources considéré et le cas échéant, d'ajouter des équations supplémentaires *a posteriori* afin d'éventuellement briser la non séparabilité.

3.2.4 Le mélange linéaire inversible

Dans cette dernière partie, nous nous intéressons à un nouveau type de mélange modélisé par des systèmes inversibles, nommé mélange linéaire inversible (ou encore mélange linéaire généralisé). Considérons un système inversible de degré relatif r où la fonction ψ est approximée par son développement de Taylor à l'ordre un. On obtient donc un système inversible linéaire de degré relatif r défini par

$$\mathbf{u}(t) = \psi\left(\mathbf{p}, \mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right) = \sum_{i=0}^{r} \mathbf{A}_i \mathbf{y}^{(i)}(t), \tag{3.52}$$

où $\mathbf{A}_i \in \mathbb{R}^{n\times n}$ sont des matrices paramétrées en $\mathbf{p}$, $\mathbf{u} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$ est le vecteur d'entrées inconnu et $\mathbf{y}^{(i)}$ est la i-ième dérivée du vecteur de sorties connu $\mathbf{y}$. Soit $\mathbb{M}$ le modèle de signaux ergodiques, stationnaires, lisses et statistiquement indépendants. Nous supposons ici $\mathbf{u}(t) \in \mathbb{M}$.

Le problème de séparation de sources consiste donc à estimer les matrices $\left\{\hat{\mathbf{A}}_i\right\}_{0\leq i\leq r}$ telles que

$$\hat{\mathbf{A}}_0 = \mathbf{PDA}_0,\ \hat{\mathbf{A}}_1 = \mathbf{PDA}_1,\ \ldots,\ \hat{\mathbf{A}}_r = \mathbf{PDA}_r,$$

où $\mathbf{D} \in \mathfrak{D}$ et $\mathbf{P} \in \mathfrak{P}$. Afin de limiter le nombre de paramètres et conformément aux remarques du paragraphe 1.2 (page 16), nous cherchons les matrices $\left\{\hat{\mathbf{A}}_i\right\}_{0\leq i\leq r}$, avec $\hat{\mathbf{A}}_0$ de diagonale unitaire, telles que

$$\hat{\mathbf{u}}(t) = \sum_{i=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(i)}(t) \in \mathbb{M}. \tag{3.53}$$

3.2.4.1 Equations d'estimation

Dans ce paragraphe, nous allons établir les équations d'estimation des matrices $\hat{\mathbf{A}}_i$ telles que

$$\hat{\mathbf{u}}(t) = \sum_{i=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(i)}(t) \tag{3.54}$$

appartienne à $\mathbb{M}$. D'après l'hypothèse d'indépendance des entrées (voir paragraphe 3.2.1.1.3), on a

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\hat{\mathbf{u}},\hat{\mathbf{u}}}(\tau)\right) = \mathbf{0}. \tag{3.55}$$

D'après l'égalité (3.54), on obtient

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\sum_{i=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(i)}(t), \sum_{i=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(i)}(t)}(\tau)\right) = \mathbf{0}.$$

Par définition de la matrice d'autocorrélation (voir définition B.9 page 163) et d'après la propriété B.28, on a

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(E\left(\sum_{i,j=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(k+i)}(t)\, \mathbf{y}^{(j)\mathrm{T}}(t-\tau)\, \hat{\mathbf{A}}_j^{\mathrm{T}}\right) - \mu_{\mathbf{u}^{(k)}} \mu_{\mathbf{u}}^{\mathrm{T}}\right) = \mathbf{0}, \tag{3.56}$$

où $\mu_{\mathbf{u}^{(k)}}$ et $\mu_{\mathbf{u}}$ sont respectivement les moyennes des signaux $\mathbf{u}^{(k)}$ et $\mathbf{u}$.

Dans la pratique, il est parfois impossible de trouver une expression analytique des équations (3.56). Par exemple :

- Si le système (3.52) est instable, alors le vecteur $\mathbf{y}$ peut être non stationnaire même si $\mathbf{u} \in \mathcal{S}^n$.
- D'autre part, même si le vecteur $\mathbf{y}$ est stationnaire, les espérances introduites dans (3.56) ne peuvent être qu'estimées. En effet, des erreurs, introduites durant l'estimation (principalement dues à l'échantillonnage des $\mathbf{y}^{(i)}$) biaisent l'estimation des espérances.

Nous allons maintenant expliquer comment approximer empiriquement les équations (3.56) par des expressions analytiques.

A partir du moment où les signaux $\mathbf{y}, \dot{\mathbf{y}}, \ldots$ peuvent être non stationnaires, l'opérateur espérance ne se distribue pas dans (3.56). Nous définissons donc l'opérateur *bar* d'un signal $s(.)$ (stationnaire ou non) comme suit :

$$\bar{s} = \frac{1}{N} \sum_{k=0}^{N} s(k\delta),$$

où N est le nombre d'échantillons et δ est la période d'échantillonnage. Cet opérateur est un estimateur empirique de $E(s)$ seulement si s est stationnaire.

Maintenant, à l'aide de l'opérateur *bar*, nous approximons les équations (3.56). Sachant que $\mathbf{u}$ est stationnaire, l'espérance $E\left(\mathbf{u}^{(k)}\mathbf{u}\right)$ existe et est approximée par $\overline{\mathbf{u}^{(k)}\mathbf{u}}$. Ainsi, les équations (3.56) s'approximent par

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\overline{\sum_{i,j=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(k+i)}(t)\, \mathbf{y}^{(j)\mathrm{T}}(t-\tau)\, \hat{\mathbf{A}}_j^{\mathrm{T}}} - \overline{\mathbf{u}^{(k)}}\; \overline{\mathbf{u}^{\mathrm{T}}}\right) = \mathbf{0}.$$

Et, par linéarité de l'opérateur *bar*, on obtient

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\sum_{i,j=0}^{r} \hat{\mathbf{A}}_i \overline{\mathbf{y}^{(k+i)}(t)\, \mathbf{y}^{(j)\mathrm{T}}(t-\tau)}\hat{\mathbf{A}}_j^{\mathrm{T}} - \overline{\mathbf{u}^{(k)}}\; \overline{\mathbf{u}^{\mathrm{T}}}\right) = \mathbf{0}.$$

D'après la proposition B.28 (page 173), les équations d'estimation sont

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\sum_{i,j=0}^{r} (-1)^j\, \hat{\mathbf{A}}_i \overline{\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}^{(k+i+j)}(\tau)}\hat{\mathbf{A}}_j^{\mathrm{T}}\right) = \mathbf{0}, \tag{3.57}$$

où

$$\overline{\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}^{(k+i+j)}(\tau)} = \overline{\left(\mathbf{y}^{(k+i+j)}(t) - \overline{\mathbf{y}^{(k+i+j)}(t)}\right)\left(\mathbf{y}(t-\tau) - \overline{\mathbf{y}(t-\tau)}\right)}$$

est l'estimation de la $(k+i+j)$-ième dérivée matrice d'autocorrélation du vecteur de sortie en τ. Les équations d'estimation (3.57) possèdent $n^2(r+1) - n$ inconnues (*i.e.* tous les coefficients des matrices $\hat{\mathbf{A}}_i$ sachant que la diagonale de $\hat{\mathbf{A}}_0$ est unitaire) et autant d'équations que voulues (*i.e.* $\forall k \in \mathbb{N}$ et $\forall \tau \in \mathbb{R}$). Nous sélectionnerons donc une partie de ces équations pour estimer les matrices $\hat{\mathbf{A}}_i$.

Dans un premier temps, afin de simplifier les calculs, nous posons $\tau = 0$ et n'exploitons que les dérivées successives de la matrice d'autocorrélation en zéro. Ensuite, nous considérons uniquement les décalages temporels ($k = 0$) de la matrice d'autocorrélation avant de comparer les performances respectives de ces deux méthodes.

3.2.4.2 Exploitation des dérivées de la matrice d'autocorrélation

Pour simplifier (3.57), et de la même façon que pour le mélange linéaire statique, nous considérons uniquement l'influence des dérivées successives de $\mathbf{\Gamma}_{\hat{\mathbf{u}},\hat{\mathbf{u}}}^{(k)}(0)$. Les équations (3.55) deviennent

$$\forall k \in \mathbb{N}, \mathbf{d}\left(\sum_{i,j=0}^{r} (-1)^j\, \hat{\mathbf{A}}_i \overline{\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}^{(k+i+j)}(0)}\hat{\mathbf{A}}_j^{\mathrm{T}}\right) = \mathbf{0}. \tag{3.58}$$

Parmi les équations (3.58), nous sélectionnons celles faisant intervenir les ordres de dérivation les plus faibles. La résolution des équations d'estimation (3.58) permet d'obtenir les matrices $\hat{\mathbf{A}}_i$, $i = 0, \ldots, r$, qui reconstruisent des sources estimées $\hat{\mathbf{u}}$ par

$$\hat{\mathbf{u}}\,(t) = \sum_{i=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(i)}\,(t)\,.$$

Présentons un exemple illustrant le système d'équations (3.58) dans le cas d'un mélange simple.

Exemple 3.16. *Considérons un système inversible linéaire de degré relatif un défini par*

$$\mathbf{u}\,(t) = \mathbf{A}_0 \mathbf{y}\,(t) + \mathbf{A}_1 \dot{\mathbf{y}}\,(t)\,.$$

Les équations d'estimation (3.58) pour un tel mélange sont les suivantes :

$$\forall k \in \mathbb{N},$$
$$\mathbf{d}\left(\hat{\mathbf{A}}_0 \overline{\mathbf{\Gamma}^{(k)}_{\mathbf{y},\mathbf{y}}\,(0)} \hat{\mathbf{A}}_0^T - \hat{\mathbf{A}}_0 \overline{\mathbf{\Gamma}^{(k+1)}_{\mathbf{y},\mathbf{y}}\,(0)} \hat{\mathbf{A}}_1^T + \hat{\mathbf{A}}_1 \overline{\mathbf{\Gamma}^{(k+1)}_{\mathbf{y},\mathbf{y}}\,(0)} \hat{\mathbf{A}}_0^T - \hat{\mathbf{A}}_1 \overline{\mathbf{\Gamma}^{(k+2)}_{\mathbf{y},\mathbf{y}}\,(0)} \hat{\mathbf{A}}_1^T\right) = \mathbf{0}. \tag{3.59}$$

Elles relient les $2n^2 - n$ coefficients inconnus des matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$.

3.2.4.2.1 Exemple formel Dans ce paragraphe, nous considérons un exemple formel de séparation de source afin de montrer la séparabilité du système linéaire inversible.

Considérons un système de mélange linéaire inversible de degré relatif un,

$$\mathbf{u}(t) = \mathbf{A}_0 \mathbf{y}(t) + \mathbf{A}_1 \dot{\mathbf{y}}(t), \tag{3.60}$$

où $\mathbf{u}(t) = (u_1(t), u_2(t))^T \in \mathbb{M}$ est le vecteur de sources statistiquement indépendantes et $\mathbf{y}(t) = (y_1(t), y_2(t))^{\mathrm{T}}$ et $\dot{\mathbf{y}}(t) = (\dot{y}_1(t), \dot{y}_2(t))^{\mathrm{T}}$sont respectivement les vecteurs d'observations et des dérivées des observations connues. Le problème de séparation de sources consiste à estimer les deux matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$ telles que

$$\hat{\mathbf{u}}(t) = \underbrace{\begin{pmatrix} 1 & p_1 \\ p_2 & 1 \end{pmatrix}}_{\hat{\mathbf{A}}_0} \mathbf{y}(t) + \underbrace{\begin{pmatrix} p_3 & p_4 \\ p_5 & p_6 \end{pmatrix}}_{\hat{\mathbf{A}}_1} \dot{\mathbf{y}}(t) \in \mathbb{M},$$

où $\mathbf{p} = [p_1, \ldots, p_6] \in \mathbb{R}^6$ est le vecteur de paramétres.

L'objectif de ce paragraphe est de montrer que le mélange considéré est séparable pour le modèle d'entrée $\mathbb{M}$, c'est-à-dire que, la connaissance exacte des dérivées en zéro de la matrice d'autocorrélation des sorties permet l'estimation des paramètres $\mathbf{p}$. Notre démarche est la suivante :

– Dans un premier temps, nous proposons de construire formellement une matrice d'autocorrélation des sorties générées à partir d'un mélange linéaire inversible de sources statistiquement indépendantes. Pour ce faire, nous choisissons une matrice d'autocorrélation des sources connues, ainsi qu'un système inversible connu pour déterminer une matrice d'autocorrélation exacte des sorties.
– Ensuite, à partir de la matrice d'autocorrélation des sorties uniquement, nous montrons que les paramètres $\mathbf{p}$ du système inversible peuvent être retrouvés parfaitement.

Remarque 3.17. *Dans la pratique, les moments des sorties étant biaisés, l'estimation des paramètres* $\mathbf{p}$ *en est rendue complexe. C'est pourquoi, avant de nous intéresser à un cas numérique, nous montrons dans un contexte idéal (formel) que les paramètres peuvent être estimés, moyennant la connaissance exacte des dérivées des matrices d'autocorrélation des sorties.*

Etape 1 : détermination de l' autocorrélation des sorties
Le mélange considéré est

$$\mathbf{u}(t) = \mathbf{A}_0\mathbf{y}(t) + \mathbf{A}_1\dot{\mathbf{y}}(t), \tag{3.61}$$

où

$$\mathbf{A}_0 = \begin{pmatrix} -2 & -\frac{8}{10} \\ \frac{3}{5} & -3 \end{pmatrix},$$
$$\mathbf{A}_1 = \begin{pmatrix} \frac{1}{10} & \frac{6}{10} \\ -\frac{6}{5} & -\frac{3}{10} \end{pmatrix},$$

et $\mathbf{u} \in \mathcal{S}^2$ est un vecteur de sources à composantes indépendantes. Soit $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ sa matrice d'autocorrélation définie par

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau) = \begin{pmatrix} \exp(-\alpha_1^2\tau^2) & 0 \\ 0 & \exp(-\alpha_2^2\tau^2) \end{pmatrix}, \tag{3.62}$$

où $\alpha_1 = \frac{7}{100}, \alpha_2 = \frac{8}{100}$ (voir figures 3.25 et 3.26). Donc,

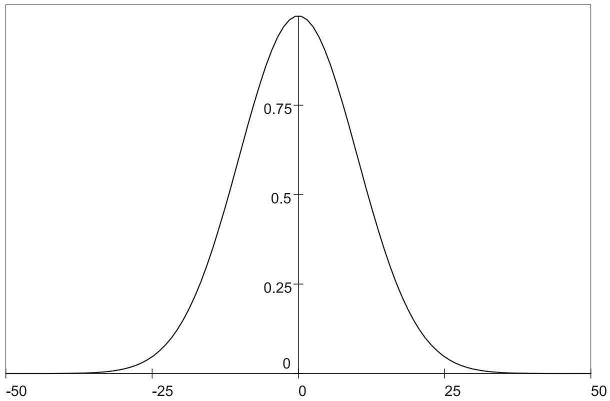


Figure 3.25 – Fonction d'autocorrélation $\Gamma_{u_1,u_1}(\tau) = \exp(-(\frac{7\tau}{100})^2)$.

$$\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(0) = \begin{cases} \begin{pmatrix} \frac{k!\left(-\alpha_1^2\right)^{\frac{k}{2}}}{\left(\frac{k}{2}\right)!} & 0 \\ 0 & \frac{k!\left(-\alpha_2^2\right)^{\frac{k}{2}}}{\left(\frac{k}{2}\right)!} \end{pmatrix} & \text{si } k \text{ pair,} \\ \mathbf{0} & \text{si } k \text{ impair.} \end{cases}$$

Afin de calculer formellement les dérivées en zéro de la matrice d'autocorrélation des sorties, nous utilisons les résultats présentés dans l'annexe B (voir section B.8 page 188) sur la représentation systémique des signaux. En particulier, l'exemple B.69 nous donne

$$\mathbf{\Gamma}^{(k)}_{\mathbf{y},\mathbf{y}}(\tau) = \sum_{i,j=0}^{n} (-1)^{i+k}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i \mathbf{A}_0^{-1}\mathbf{\Gamma}^{(i+j+k)}_{\mathbf{uu}}(\tau)\,\mathbf{A}_0^{-1^T}\left(\mathbf{A}_1^T\mathbf{A}_0^{-1^T}\right)^j,$$

pour n grand ($n = 80$).

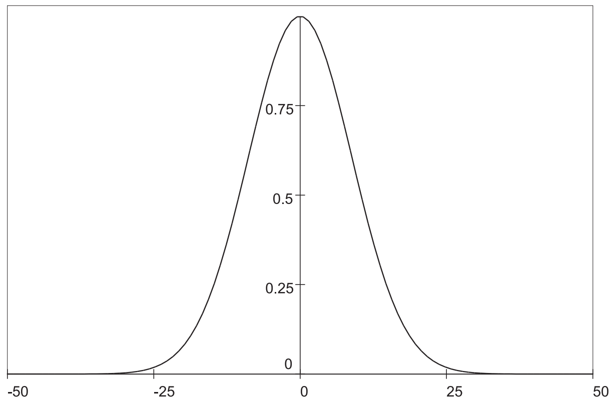


Figure 3.26 – Fonction d'autocorrélation $\Gamma_{u_2,u_2}(\tau) = \exp(-(\frac{8\tau}{100})^2)$.

Etape 2 : séparation des sources

Maintenant, seulement à partir de la connaissance des dérivées en zéro de la matrice d'autocorrélation des sorties $\mathbf{\Gamma}^{(k)}_{\mathbf{y},\mathbf{y}}(0)$, nous allons montrer qu'il est possible de retrouver les paramètres $\mathbf{p}$ des matrices $\hat{\mathbf{A}}_i$ telles que

$$\hat{\mathbf{u}}(t) = \hat{\mathbf{A}}_0\mathbf{y}(t) + \hat{\mathbf{A}}_1\dot{\mathbf{y}}(t) \in \mathbb{M}. \tag{3.63}$$

Les équations d'estimation (3.59) expriment les paramètres du système en fonction des dérivées des matrices d'autocorrélation des sorties en zéro. Après développement, on obtient

$$\begin{aligned}
&p_2\Gamma_{y_1y_1} + (1+p_1p_2)\,\Gamma_{y_1y_2} + p_1\Gamma_{y_2y_2} + (p_3 - p_6 + p_1p_5 - p_2p_4)\,\Gamma^{(1)}_{y_1y_2}\\
&-p_3p_5\Gamma^{(2)}_{y_1y_1} + (-p_3p_6 - p_4p_5)\,\Gamma^{(2)}_{y_1y_2} - p_4p_6\Gamma^{(2)}_{y_2y_2} = 0\\[1em]
&(1-p_1p_2)\,\Gamma^{(1)}_{y_1y_2} + (p_2p_3 - p_5)\,\Gamma^{(2)}_{y_1y_1} + (p_3 - p_6 - p_1p_5 + p_2p_4)\,\Gamma^{(2)}_{y_1y_2}\\
&(p_4 - p_1p_6)\,\Gamma^{(2)}_{y_2y_2} + (p_4p_5 - p_3p_6)\,\Gamma^{(3)}_{y_1y_2} = 0\\[1em]
&p_2\Gamma^{(2)}_{y_1y_1} + (1+p_1p_2)\,\Gamma^{(2)}_{y_1y_2} + p_1\Gamma^{(2)}_{y_2y_2} + (p_3 - p_6 + p_1p_5 - p_2p_4)\,\Gamma^{(3)}_{y_1y_2}\\
&-p_3p_5\Gamma^{(4)}_{y_1y_1} + (-p_3p_6 - p_4p_5)\,\Gamma^{(4)}_{y_1y_2} - p_4p_6\Gamma^{(4)}_{y_2y_2} = 0\\[1em]
&(1-p_1p_2)\,\Gamma^{(3)}_{y_1y_2} + (p_2p_3 - p_5)\,\Gamma^{(4)}_{y_1y_1} + (p_3 - p_6 - p_1p_5 + p_2p_4)\,\Gamma^{(4)}_{y_1y_2}\\
&+ (p_4 - p_1p_6)\,\Gamma^{(4)}_{y_2y_2} + (p_4p_5 - p_3p_6)\,\Gamma^{(5)}_{y_1y_2} = 0\\[1em]
&p_2\Gamma^{(4)}_{y_1y_1} + (1+p_1p_2)\,\Gamma^{(4)}_{y_1y_2} + p_1\Gamma^{(4)}_{y_2y_2} + (p_3 - p_6 + p_1p_5 - p_2p_4)\,\Gamma^{(5)}_{y_1y_2}\\
&-p_3p_5\Gamma^{(6)}_{y_1y_1} + (-p_3p_6 - p_4p_5)\,\Gamma^{(6)}_{y_1y_2} - p_4p_6\Gamma^{(6)}_{y_2y_2} = 0\\[1em]
&(1-p_1p_2)\,\Gamma^{(5)}_{y_1y_2} + (p_2p_3 - p_5)\,\Gamma^{(6)}_{y_1y_1} + (p_3 - p_6 - p_1p_5 + p_2p_4)\,\Gamma^{(6)}_{y_1y_2}\\
&+ (p_4 - p_1p_6)\,\Gamma^{(6)}_{y_2y_2} + (p_4p_5 - p_3p_6)\,\Gamma^{(7)}_{y_1y_2} = 0,
\end{aligned} \tag{3.64}$$

où $\Gamma^{(k)}_{y_iy_j} = \Gamma^{(k)}_{y_iy_j}(0)$. Notons que des simplifications ont été réalisées en tenant compte de la remarque B.30. La résolution formelle (utilisant les bases de Gröbner, voir annexe D) du système (3.64) nous donne les deux vecteurs de paramètres $\hat{\mathbf{p}}$ et $\hat{\mathbf{p}}'$ suivants comme solutions

$$\hat{\mathbf{p}} = \left[-\frac{2}{5}, -\frac{1}{5}, \frac{1}{20}, \frac{3}{10}, \frac{4}{15}, \frac{1}{10}\right],$$

$$\hat{\mathbf{p}}' = \left[-5, -\frac{5}{2}, -\frac{4}{3}, -\frac{1}{2}, -\frac{1}{8}, -\frac{3}{4}\right],$$

correspondant aux coefficients des matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$. On vérifie aisément que ces solutions sont bien des estimations de $\mathbf{A}_0$ et $\mathbf{A}_1$ (c'est-à-dire $\left(\hat{\mathbf{A}}_0, \hat{\mathbf{A}}_1\right) \sim (\mathbf{A}_0, \mathbf{A}_1)$).

En conclusion, les mélanges de sources de matrices d'autocorrélation distinctes sont séparables pour le modèle $\mathbb{M}$. Toutefois, un problème difficile reste ouvert. On peut le formuler en ces termes : A quelles conditions sur les sources l'ensemble (3.64) conduit-il à une solution unique, ou bien, quand conduit-il à des solutions dégénérées ?

3.2.4.2.2 Simulation Maintenant, intéressons-nous à un exemple numérique où les sources et les observations sont liées par un mélange inversible linéaire de degré relatif un.

Considérons deux sources gaussiennes indépendantes et différentiables représentées sur les figures 3.27 et 3.28 (28000 échantillons avec une période d'échantillonage $T_e = 0.1$).

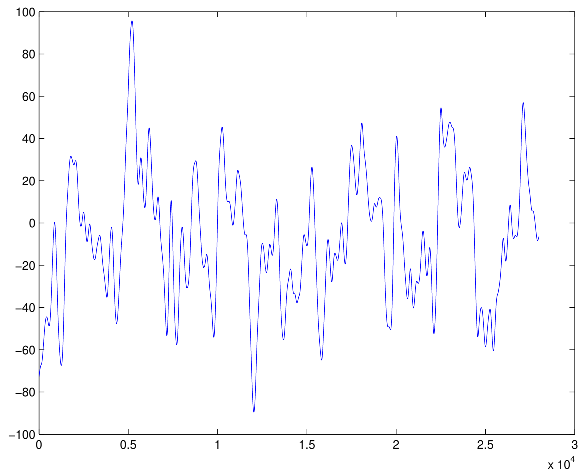


Figure 3.27 – Source $u_1(t)$.

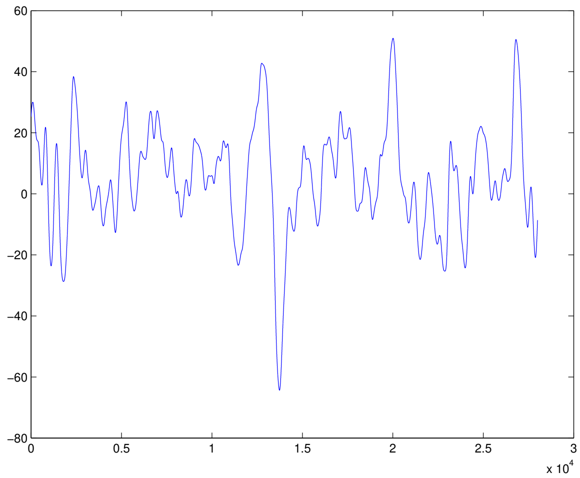


Figure 3.28 – Source $u_2(t)$.

Elles suivent le modèle

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -2 & -\frac{4}{10} \\ \frac{3}{10} & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} -\frac{3}{10} & \frac{6}{10} \\ -\frac{8}{5} & -\frac{3}{10} \end{pmatrix} \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix}. \tag{3.65}$$

Les signaux observés sont représentés sur les figures 3.29 et 3.30.

La résolution du système (3.64) permet d'estimer les matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$ solutions. A l'aide d'une méthode de Newton (voir annexe C), nous trouvons, pour de bonnes conditions initiales (presque toutes marchent), les solutions

$$\hat{\mathbf{A}}_0 = \begin{pmatrix} 1 & -0.6119 \\ -0.0502 & 1 \end{pmatrix} \text{ et } \hat{\mathbf{A}}_1 = \begin{pmatrix} -0.3402 & 0.7047 \\ 0.5471 & -0.0166 \end{pmatrix},$$

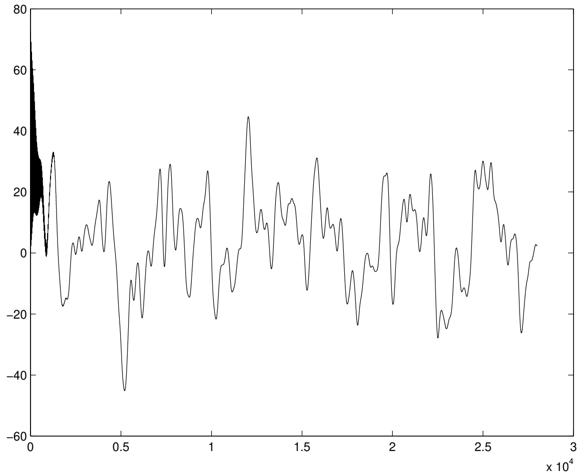


Figure 3.29 – Observation $y_1(t)$.

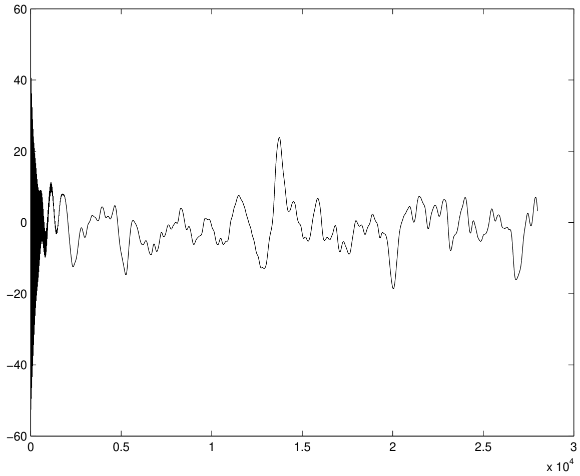


Figure 3.30 – Observation $y_2(t)$.

correspondant aux sources estimées $\hat{\mathbf{u}}$ représentées en superposition avec les sources réelles sur les figures 3.31 et 3.32.

Au premier regard, ces résultats peuvent sembler mauvais. En effet, les paramètres estimés des matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$ ne correspondent pas aux solutions exactes. Cependant, on peut voir, sur les figures 3.31 et 3.32, que les sources sont bien retrouvées à un facteur d'échelle près. En effet, la forme d'onde des signaux est conservée.
Notons toutefois un effet "transitoire" (sur les sources estimées) probablement dû à l'erreur d'estimation des paramètres du système de séparation (qui, rappelons le, est dynamique).

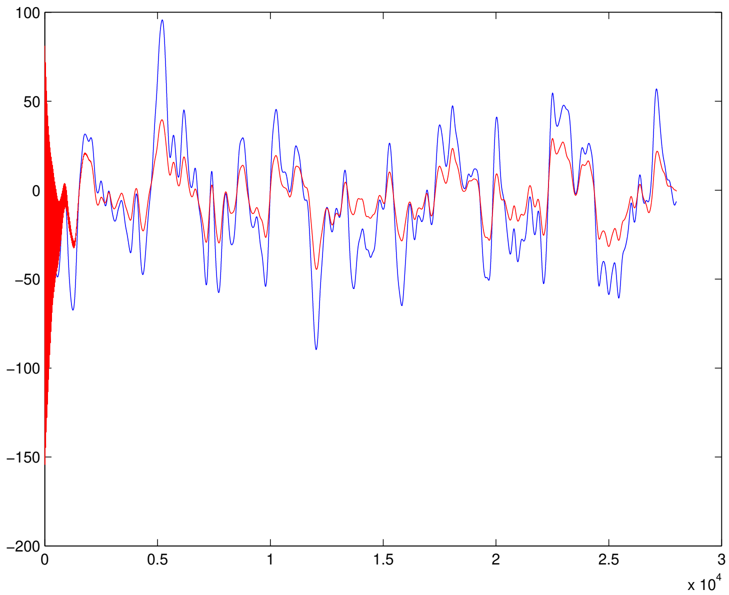


Figure 3.31 – Superposition de la source $u_1(t)$ et de la source estimée $\hat{u}_1(t)$.

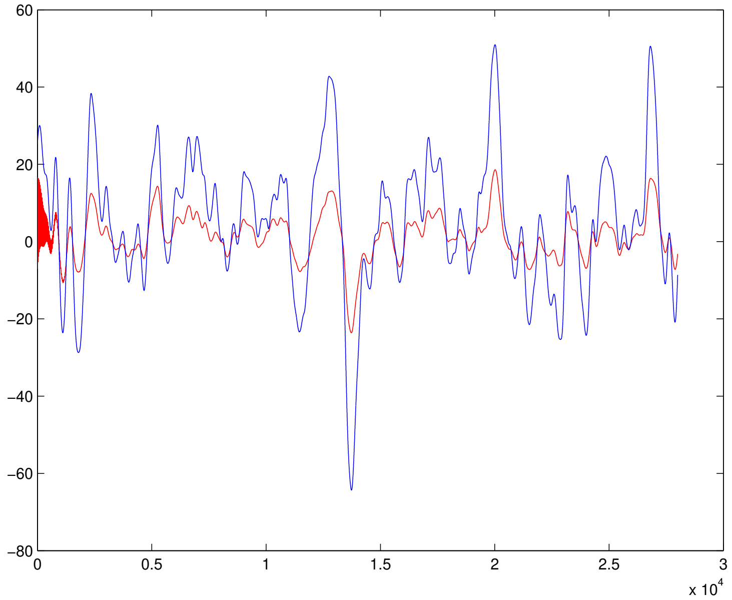


Figure 3.32 – Superposition de la source $u_2(t)$ et de la source estimée $\hat{u}_2(t)$.

Remarque 3.18. *Dans ce paragraphe, nous avons utilisé une méthode de Newton pour la résolution. Les techniques par intervalles (SIVIA ou OPTICRIT) s'avérent peu fructueuses pour ce problème en raison tout d'abord des occurences multiples des paramètres dans (3.64), ne permettant pas d'obtenir une fonction d'inclusion "efficace", puis de la complexité induite par la dimension du problème (ici 6). En effet, il faut obtenir un sous-pavage d'un espace de dimension six. Pour ces deux raisons, nous avons choisi la méthode de Newton (malgré ses inconvénients cités dans l'annexe C) afin d'éviter toute explosion combinatoire des algorithmes par intervalles présentés. Cependant, il existe des techniques ensemblistes plus raffinées, non présentées dans cette thèse, permettant de traiter des problèmes en grande dimension. Le lecteur intéressé pourra se référer à [54].*

3.2.4.3 Exploitation des décalages temporels de la matrice d'autocorrélation

Dans ce paragraphe, nous allons exploiter les décalages temporels de la matrice d'autocorrélation des sorties afin d'utiliser des équations d'estimation différentes. La démarche est similaire à celle du paragraphe 3.2.4.1 dans lequel nous avons utilisé les dérivées successives de la matrice d'autocorrélation des sorties.

En considérant uniquement les décalages temporels (*i.e.* en posant l'ordre de dérivation k nul) dans les équations (3.57), on obtient les équations d'estimation suivantes

$$\forall \tau \in \mathbb{R}, \mathbf{d}\left(\sum_{i,j=0}^{r} (-1)^j \, \hat{\mathbf{A}}_i \overline{\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}^{(i+j)}(\tau)} \hat{\mathbf{A}}_j^{\mathrm{T}}\right) = \mathbf{0}. \tag{3.66}$$

Les équations d'estimation (3.66) possèdent $n^2(r+1) - n$ inconnues et autant d'équations que l'on veut (*i.e.* $\forall \tau \in \mathbb{R}$). Nous sélectionnons donc une partie de ces équations, par exemple celles faisant intervenir les plus petits décalages temporels.

La résolution des équations d'estimation (3.66) permet d'obtenir les matrices $\hat{\mathbf{A}}_i$, $i = 0, \ldots, r$, qui reconstruisent des sources estimées $\hat{\mathbf{u}}$ par

$$\hat{\mathbf{u}}(t) = \sum_{i=0}^{r} \hat{\mathbf{A}}_i \mathbf{y}^{(i)}(t).$$

Nous proposons ci-après un exemple simple illustrant le système d'équations (3.66).

Exemple 3.19. *Soit*

$$\mathbf{u}(t) = \mathbf{A}_0 \mathbf{y}(t) + \mathbf{A}_1 \dot{\mathbf{y}}(t),$$

un système inversible linéaire de degré relatif 1*. Les équations d'estimation (3.66) pour un tel mélange sont*

$$\forall \tau \in \mathbb{R},$$
$$\mathbf{d}\left(\hat{\mathbf{A}}_0 \overline{\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}(\tau)} \hat{\mathbf{A}}_0^T - \hat{\mathbf{A}}_0 \overline{\mathbf{\Gamma}^{(1)}_{\mathbf{y},\mathbf{y}}(\tau)} \hat{\mathbf{A}}_1^T + \hat{\mathbf{A}}_1 \overline{\mathbf{\Gamma}^{(1)}_{\mathbf{y},\mathbf{y}}(\tau)} \hat{\mathbf{A}}_0^T - \hat{\mathbf{A}}_1 \overline{\mathbf{\Gamma}^{(2)}_{\mathbf{y},\mathbf{y}}(\tau)} \hat{\mathbf{A}}_1^T\right) = \mathbf{0}. \tag{3.67}$$

Elles relient les $2n^2 - n$ coefficients des matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$. Pour $n = 2$, le développement des équations (3.67) nous donne, pour $\tau = 0$ puis $\tau = 1$,

$$\begin{aligned}
&p_2 \Gamma_{y_1 y_1}(0) + (1 + p_1 p_2)\Gamma_{y_1 y_2}(0) + p_1 \Gamma_{y_2 y_2}(0) \\
&+ (p_3 - p_6 + p_1 p_5 - p_2 p_4)\Gamma^{(1)}_{y_1 y_2}(0) - p_3 p_5 \Gamma^{(2)}_{y_1 y_1}(0) \\
&+ (-p_3 p_6 - p_4 p_5)\Gamma^{(2)}_{y_1 y_2}(0) - p_4 p_6 \Gamma^{(2)}_{y_2 y_2}(0) = 0 \\
\\
&\Gamma_{y_1 y_2}(1) + p_1 \Gamma_{y_2 y_2}(1) + p_3 \Gamma^{(1)}_{y_1 y_2}(1) + p_4 \Gamma^{(1)}_{y_2 y_2}(1) \\
&+ p_2 \left(\Gamma_{y_1 y_1}(1) + p_1 \Gamma_{y_1 y_2}(-1)\right) + p_2 \left(-p_4 \Gamma^{(1)}_{y_1 y_2}(-1) + p_3 \Gamma^{(1)}_{y_1 y_1}(1)\right) \\
&+ p_6 \left(-\Gamma^{(1)}_{y_1 y_2}(1) - p_1 \Gamma^{(1)}_{y_2 y_2}(1)\right) + p_6 \left(-p_3 \Gamma^{(2)}_{y_1 y_2}(1) - p_4 \Gamma^{(2)}_{y_2 y_2}(1)\right) \\
&+ p_5 \left(+p_1 \Gamma^{(1)}_{y_1 y_2}(-1) - \Gamma^{(1)}_{y_1 y_1}(1)\right) + p_5 \left(-p_3 \Gamma^{(2)}_{y_1 y_1}(1) - p_4 \Gamma^{(2)}_{y_1 y_2}(-1)\right) = 0
\end{aligned} \tag{3.68}$$

Le développement des équations (3.67) pour des décalages τ supérieurs se fait de manière similaire.

Simulation

Reprenons la simulation numérique du paragraphe 3.2.4.2.2. Nous proposons ici d'estimer les matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$ telles que

$$\hat{\mathbf{u}}(t) = \hat{\mathbf{A}}_0 \mathbf{y}(t) + \hat{\mathbf{A}}_1 \dot{\mathbf{y}}(t) \in \mathbb{M}, \tag{3.69}$$

en exploitant uniquement les décalages temporels de la matrice d'autocorrélation des sorties. La résolution du système (3.68), par un algorithme de Newton, nous donne les estimations suivantes pour les matrices $\hat{\mathbf{A}}_0$ et $\hat{\mathbf{A}}_1$:

$$\hat{\mathbf{A}}_0 = \begin{pmatrix} 1 & -0.9219 \\ 0.0777 & 1 \end{pmatrix} \text{ et } \hat{\mathbf{A}}_1 = \begin{pmatrix} -0.4452 & -0.3900 \\ -0.1189 & -0.3115 \end{pmatrix}.$$

Les superpositions des sources réelles $\mathbf{u}^\star(t)$ et des sources estimées $\hat{\mathbf{u}}(t) = \hat{\mathbf{A}}_0 \mathbf{y}(t) + \hat{\mathbf{A}}_1 \dot{\mathbf{y}}(t)$ correspondantes sont représentées sur les figures 3.33 et 3.34.

De même que précédemment, les paramètres estimés ne sont pas satisfaisants, toutefois nous constatons sur les figures 3.33 et 3.34 que la forme d'onde des sources est conservée et qu'ainsi elles sont retrouvées à un facteur d'échelle près.

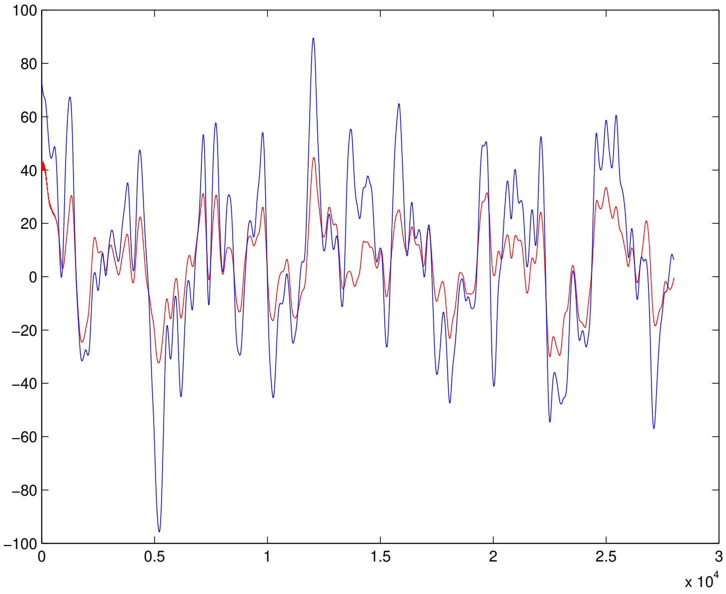


Figure 3.33 – Superposition de la source estimée $\hat{u}_1(t)$ et de la source réelle $u_1^\star(t)$.

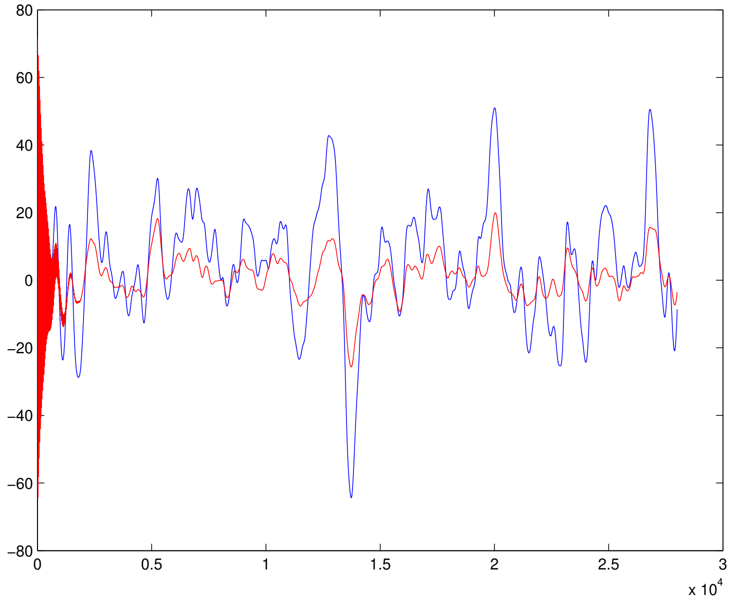


Figure 3.34 – Superposition de la source estimée $\hat{u}_2(t)$ et de la source réelle $u_2^\star(t)$.

3.2.5 Synthèse

Dans cette partie sur les systèmes inversibles linéaires, nous avons présenté une méthode de séparation à l'ordre deux basée sur les décalages temporels ou les dérivées successives de la matrice d'autocorrélation des sorties.

Comme le montrent les simulations, l'estimation des matrices $\mathbf{A}_i$ est complexe du fait de l'erreur

d'estimation introduite par l'opérateur *bar*. Cependant, les sources $\hat{\mathbf{u}}(t)$ estimées fournissent une approximation convenable de $\mathbf{u}(t)$. A noter que nous nous sommes placés dans un contexte difficile où l'on dispose d'autant d'équations que d'inconnues. Bien évidemment, il est possible d'exploiter des équations supplémentaires qui utilisent davantage de décalages temporels ou de dérivées de la matrice d'autocorrélation $\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}(\tau)$.
Une étude de la robustesse reste à réaliser afin de vérifier que l'estimation des matrices $\mathbf{A}_i$ peut être améliorée lorsque le nombre d'équations considérées est plus important.

Nous avons présenté une méthode de résolution élémentaire, toutefois, il serait intéressant de rechercher des estimateurs efficaces et consistants afin de perfectionner la méthode. Il peut être envisageable d'exploiter les travaux [9] sur les mélanges linéaires convolutifs pour obtenir des techniques de séparation plus efficaces (comme par exemple l'algorithme SOBI convolutif). Cette piste reste à explorer.

En perspective, il serait intéressant de vérifier que la démarche proposée se généralise bien pour la séparation de mélanges inversibles non linéaires. On pourrait ainsi étudier des mélanges non linéaires inversibles obtenus en considérant le développement de Taylor de la fonction ψ à l'ordre deux. Par exemple, pour un mélange inversible de degré relatif 0 de deux sources ($n = 2$), on obtient :

$$\begin{aligned} \mathbf{u}(t) &= \psi(\mathbf{y}) \\ &= \mathbf{A}_0\mathbf{y}(t) + \mathbf{B}_0\mathbf{y}^2(t), \end{aligned}$$

où le vecteur $\mathbf{u}(t) = (u_1(t), u_2(t))^{\mathrm{T}}$ est inconnu, les matrices $\mathbf{A}_0 = (a_{ij})_{i,j}$ et $\mathbf{B}_0 = (b_{ij})_{i,j}$ sont inconnues et les vecteurs $\mathbf{y}(t) = (y_1(t), y_2(t))^{\mathrm{T}}$ et $\mathbf{y}^2(t) = \left(y_1^2(t), y_1(t)y_2(t), y_2^2(t)\right)^{\mathrm{T}}$ sont connus.
En développant, on a

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} a_{11}y_1 + a_{12}y_2 + b_{11}y_1^2 + b_{12}y_1y_2 + b_{13}y_2^2 \\ a_{21}y_1 + a_{22}y_2 + b_{21}y_1^2 + b_{22}y_1y_2 + b_{23}y_2^2 \end{pmatrix}.$$

Toutefois, du fait de l'apparition de termes non linéaires dans le modèle de mélange, l'utilisation des statistiques d'ordre supérieur à deux est induite et par conséquent, la séparation de signaux gaussiens (au moins deux) risque d'être impossible.

Un début de réponse à ces questions sera fourni dans le chapitre suivant qui s'intéresse à l'identifiabilité en aveugle de systèmes principalement non linéaires.

CHAPITRE 4

Estimation aveugle de paramètres pour des systèmes inversibles

Dans ce chapitre, nous nous intéressons au problème d'estimation de paramètres en aveugle, c'est-à-dire lorsque les entrées ne sont pas observées [1]. Nous verrons que ce problème présente de fortes analogies avec le problème de séparation de sources étudié dans le chapitre 3. Nous limitons notre étude à la classe des systèmes inversibles qui, comme nous l'avons déjà évoqué, modélisent un grand nombre de phénomènes physiques.

L'objectif de ce dernier chapitre est de répondre à la question suivante : est-il possible, moyennant des hypothèses statistiques sur les entrées, d'identifier en aveugle les paramètres d'un système inversible paramétré ? La notion induite sera appelée identifiabilité en aveugle. Nous montrerons qu'en théorie, pour une sous-classe particulière de systèmes inversibles nommés systèmes inversibles découplés, le problème d'estimation de paramètres en aveugle peut être résolu grâce à l'analyse par intervalles. Notre démarche sera illustrée à l'aide d'exemples concrets.

Avant de présenter notre méthode de résolution et les notions qu'elle requiert, nous commençons par décrire le problème auquel nous nous intéressons dans ce chapitre.

4.1 Problématique

Considérons un système inversible paramétré de degré relatif r

$$\mathbf{u}\left(t\right)=\psi\left(\mathbf{p},\mathbf{y}\left(t\right),\dot{\mathbf{y}}\left(t\right),\ldots,\mathbf{y}^{(r-1)}\left(t\right),\mathbf{y}^{(r)}\left(t\right)\right), \tag{4.1}$$

où

- $t\in\mathbb{R}$ est le temps,
- le vecteur d'entrées $\mathbf{u}\in\mathcal{S}^n$ (avec $\mathcal{S}$, ensemble des signaux aléatoires stationnaires, ergodiques et lisses) est inconnu,
- la i-ième dérivée $\mathbf{y}^{(i)}\in\left(\mathbb{R}^{\mathbb{R}}\right)^n$ du vecteur de sorties $\mathbf{y}$ est connue (consulter l'annexe E pour différentes méthodes d'estimation des dérivées d'un signal aléatoire),
- le vecteur de paramètres $\mathbf{p}\in\mathcal{P}=\mathbb{R}^{n_p}$ est inconnu,
- la fonction analytique ψ définie sur $\mathcal{P}\times\left(\mathbb{R}^{\mathbb{R}}\right)^{n(r+1)}$ à valeurs dans $\left(\mathbb{R}^{\mathbb{R}}\right)^n$ est connue, à l'exception bien sûr du vecteur de paramètres $\mathbf{p}$.

Nous posons uniquement quelques hypothèses statistiques sur les signaux d'entrées inconnues, par exemple l'indépendance ou la gaussianité de $\mathbf{u}$, et notons $\mathbb{M}$ le modèle d'entrées correspondant à ces hypothèses.

Le problème d'*estimation de paramètres en aveugle* consiste à estimer le vecteur de paramètres inconnu $\mathbf{p}$ à partir des signaux observés $\mathbf{y}$ en exploitant les hypothèses statistiques du modèle $\mathbb{M}$. Autrement dit, étant donné le système inversible (4.1) et un modèle statistique $\mathbb{M}$, le problème consiste à caractériser l'ensemble des paramètres solutions suivant :

$$\left\{\mathbf{p} \in \mathcal{P} | \psi\left(\mathbf{p}, \mathbf{y}(t), \ldots, , \mathbf{y}^{(r)}(t)\right) \in \mathbb{M}\right\}. \tag{4.2}$$

Dans un premier temps, nous traduisons en terme d'équations (d'estimation) les hypothèses statistiques sur le vecteur d'entrées grâce aux *fonctions d'estimation* [20]. Nous définissons à cette occasion une nouvelle algèbre nommée *algèbre des moments d'un signal*. Ensuite, nous présentons la démarche que nous avons adoptée pour résoudre le problème d'estimation de paramètres.

4.1.1 Equations d'estimation

Dans ce paragraphe, nous présentons tout d'abord quelques propriétés des moments d'un vecteur de signaux aléatoires puis, précisons la notion de fonction d'estimation fortement liée à celle d'équations d'estimation du paragraphe 3.2.1.1.3.

4.1.1.1 Algèbre des moments

Soit un vecteur $\mathbf{u}(.) = (u_1(.), \ldots, u_n(.))^T$ de n signaux aléatoires quelconques, on appelle $\mathcal{M}_{\mathbf{u}}$ l'ensemble des moments du signal $\mathbf{u}(.)$, dans la mesure où ces derniers sont définis.
Dans ce chapitre, nous considérons uniquement les signaux aléatoires $\mathbf{u}(.) \in \mathcal{S}^n$ (stationnaires, ergodiques et lisses), toutefois les résultats présentés se généralisent sans difficulté à tous les signaux aléatoires pourvu que leurs moments soient définis.

Tous les éléments de l'ensemble $\mathcal{M}_{\mathbf{u}}$ sont des fonctions de $\mathbf{u}(.)$ que nous noterons $\mu(\mathbf{u})$. Par exemple, $\mu : \mathbf{u}(.) \to E(u_1)$ est un moment de $\mathcal{M}_{\mathbf{u}}$.
Rappelons que, d'après la définition A.72 (page 154), un moment vérifie les propriétés suivantes :

$$\begin{aligned} &\mu(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}} && \Longrightarrow \quad \forall a, b \in \mathbb{R}, a\mu(\mathbf{u}) + b \in \mathcal{M}_{\mathbf{u}}, \\ &\left\{ \begin{array}{l} \mu_1(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}} \\ \mu_2(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}} \end{array} \right. && \Longrightarrow \quad \mu_1(\mathbf{u}) + \mu_2(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}}. \end{aligned} \tag{4.3}$$

Voici quelques exemples d'éléments de $\mathcal{M}_{\mathbf{u}}$.

Exemple 4.1. *Soit $\mathbf{u}(.) = (u_1(.), u_2(.))^T$ un vecteur de deux signaux aléatoires. Pour $\tau \in \mathbb{R}$ fixé, on a :*

$$\begin{aligned} &\mu_1(\mathbf{u}) = E(u_1(t)\, u_2(t-\tau)) \in \mathcal{M}_{\mathbf{u}}, \\ &\mu_2(\mathbf{u}) = E\left(u_1^3\right) \in \mathcal{M}_{\mathbf{u}}, \\ &\mu_3(\mathbf{u}) = E(\ddot{u}_2) \in \mathcal{M}_{\mathbf{u}}. \end{aligned}$$

Alors que

$$E(u_1(t)\, u_2(t-\tau))^2 \notin \mathcal{M}_{\mathbf{u}},$$

puisque c'est le carré d'un moment de $\mathbf{u}$*.*

Illustrons les propriétés (4.3). On a bien

$$\forall a, b \in \mathbb{R},\ aE(u_1) + b = E(au_1 + b) \in \mathcal{M}_{\mathbf{u}},$$

$$E(u_1) + E(u_2) = E(u_1 + u_2) \in \mathcal{M}_{\mathbf{u}}.$$

En remarquant que $\mathcal{M}_{\mathbf{u}}$ n'est pas stable par produit ($E(u_1) \times E(u_2) \notin \mathcal{M}_{\mathbf{u}}$), nous introduisons une nouvelle algèbre.

Définition 4.2 (Algèbre des moments)**.** L'algèbre des moments de $\mathbf{u}(.)$, notée $\mathcal{A}_{\mathbf{u}}$, est l'ensemble des fonctions induites par la règle suivante :

$$\left\{ \begin{array}{l} \mu_1(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}} \\ \mu_2(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}} \end{array} \right. \implies \mu_1(\mathbf{u}) \times \mu_2(\mathbf{u}) \in \mathcal{A}_{\mathbf{u}}. \tag{4.4}$$

Ainsi définie, l'algèbre $\mathcal{A}_{\mathbf{u}}$ des moments de $\mathbf{u}(.)$ contient l'ensemble $\mathcal{M}_{\mathbf{u}}$ des moments de $\mathbf{u}(.)$ et tous les produits d'éléments de $\mathcal{M}_{\mathbf{u}}$.

4.1.1.2 Fonction d'estimation

Par analogie avec [20], on définit une fonction d'estimation comme une fonction vectorielle dont chaque composante appartient à l'algèbre des moments introduite précédemment. Plus formellement, on a la définition suivante :

Définition 4.3 (Fonction d'estimation)**.** Une fonction d'estimation est une fonction

$$\mathbf{h} : \left\{ \begin{array}{lll} \mathcal{S}^n & \rightarrow & \mathbb{R}^q \\ \mathbf{u}(.) & \rightarrow & h(\mathbf{u}) = (h_1(\mathbf{u}), \ldots, h_q(\mathbf{u}))^{\mathrm{T}} \end{array} \right. ,$$

avec

$$\forall i, h_i \in \mathcal{A}_{\mathbf{u}},$$

où $\mathcal{A}_{\mathbf{u}}$ est l'algèbre des moments de $\mathbf{u}(.)$.

Exemple 4.4. *La fonction*

$$\mathbf{h} : \left\{ \begin{array}{lll} \mathcal{S}^2 & \rightarrow & \mathbb{R} \\ (u_1(.), u_2(.)) & \rightarrow & E(u_1 u_2) - E(u_1) E(u_2) \end{array} \right. \tag{4.5}$$

est une fonction d'estimation.

Ces fonctions sont construites de façon à s'annuler lorsque les hypothèses statistiques du modèle d'entrées sont satisfaites. Par exemple, si les signaux d'entrées u_1 et u_2 sont supposés décorrélés, la fonction d'estimation (4.5) peut être utilisée.

On voit clairement apparaître l'utilité des fonctions d'estimation. Elles permettent de traduire les hypothèses statistiques sur les entrées $\mathbf{u}$ en équations (d'estimation) du type $\mathbf{h}(\mathbf{u}) = \mathbf{0}$. Ainsi, pour chaque modèle d'entrées $\mathbb{M}$, nous allons construire une fonction d'estimation $\mathbf{h}$ adaptée telle que

$$\mathbf{u} \in \mathbb{M} \implies \mathbf{h}(\mathbf{u}) = \mathbf{0}. \tag{4.6}$$

Remarque 4.5. *L'idéal serait de construire une fonction d'estimation* $\mathbf{h}$ *telle que*

$$\mathbf{u} \in \mathbb{M} \Longleftrightarrow \mathbf{h}(\mathbf{u}) = \mathbf{0}.$$

Cette équivalence est possible, toutefois, dans la pratique, nous avons décidé de l'abandonner puisqu'elle est excessivement coûteuse. En effet, si on suppose l'indépendance, respectivement la gaussianité, des entrées, cela revient à vérifier une infinité d'équations (voir définition B.37 page 177, respectivement définition B.52 page 183) et donc à choisir une fonction d'estimation de dimension infinie.

Grâce aux fonctions d'estimation, le problème d'estimation de paramètres en aveugle se formalise comme suit. Etant donné un système inversible $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}(t), \ldots, , \mathbf{y}^{(r)}(t)\right)$ et une fonction d'estimation $\mathbf{h}$ qui vérifie (4.6) pour un modèle d'entrées $\mathbb{M}$ fixé, il s'agit de caractériser l'ensemble

$$\mathbb{P} = \left\{\mathbf{p} \in \mathcal{P} | \mathbf{h}\left(\psi\left(\mathbf{p}, \mathbf{y}(t), \ldots, , \mathbf{y}^{(r)}(t)\right)\right) = \mathbf{0}\right\}. \tag{4.7}$$

4.1.2 Fil conducteur

Nous proposons dans ce chapitre une nouvelle méthodologie pour résoudre le problème d'estimation de paramètres en aveugle pour des systèmes inversibles, linéaires ou non. Remarquons d'ores et déjà qu'une fois l'estimation de $\mathbf{p}$ effectuée, $\mathbf{u}$ est alors connu via l'égalité (4.1). En ce sens, le problème d'estimation de paramètres en aveugle est proche du problème de séparation aveugle de sources qui consiste à estimer les signaux d'entrées $\mathbf{u}$ inconnus grâce à des hypothèses statistiques sur ces derniers. Ceci explique que les méthodes proposées dans ce chapitre présentent des similitudes avec celles utilisées dans le chapitre 3 pour résoudre le problème de séparation de sources.

Dans un premier temps, nous nous intéressons à certaines propriétés des systèmes inversibles. Nous rappelons d'abord quelques résultats sur l'identifiabilité structurelle [90] qui caractérise, connaissant les entrées et les sorties, les chances de succès de l'estimation de paramètres. En effet, avant même de résoudre de manière numérique le problème d'estimation des paramètres, se pose la question de l'unicité des paramètres solutions. Par exemple, si on considère le système paramétré

$$u = p_1 p_2 y,$$

il est clair que connaissant l'entrée u et la sortie y, les paramètres p_1 et p_2 ne peuvent être estimés. Seul le produit $p_1 p_2$ est "identifiable".

Puis, par analogie avec l'identifiabilité structurelle, nous définissons la notion d'identifiabilité en aveugle, c'est-à-dire la possibilité d'estimer les paramètres $\mathbf{p}$ en exploitant uniquement les sorties et le modèle statistique $\mathbb{M}$ supposé pour les entrées $\mathbf{u}$.

Ensuite, nous construisons des fonctions d'estimation pour un modèle d'entrées $\mathbb{M}$ donné. Notons d'ores et déjà que nous n'étudions en détails les fonctions d'estimation que pour deux types de modèle d'entrées : le modèle indépendant et le modèle gaussien.

Enfin, nous proposons des exemples illustrant notre démarche et terminons par l'estimation en aveugle des paramètres d'un système en automatique : le système des bacs d'eau.

4.2 Identifiabilité

Nous noterons $M(.)$ la structure choisie pour un système paramétré et $M(\mathbf{p})$ le système particulier obtenu en donnant aux paramètres la valeur $\mathbf{p}$ (voir figure 4.1).

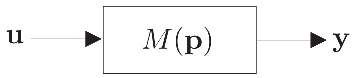


Figure 4.1 – Système entrées-sorties de structure $M(\mathbf{p})$.

Par exemple, dans le cas d'un système inversible, on note

$$M(\mathbf{p}) : \mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right). \tag{4.8}$$

Maintenant, il est important d'étudier les propriétés du système de modèle $M(.)$ le plus indépendamment des valeurs prises par ses paramètres. Nous pourrons ainsi, avant même le recueil des données, détecter d'éventuels problèmes pouvant apparaître lors de l'identification.

Deux propriétés sont particulièrement importantes dans le contexte de l'estimation de paramètres en aveugle :

– L'*identifiabilité structurelle* telle qu'elle est présentée par Walter dans [90].
– L'*identifiabilité en aveugle*, que nous présentons dans cette section, correspondant à une extension de l'identifiabilité structurelle dans le cas où les entrées du modèle de la figure 4.1 sont inconnues.

Dans un premier temps, nous rappelons le concept d'identifiabilité structurelle d'un système paramétré. Puis, par analogie, nous définissons l'identifiabilité en aveugle d'un système paramétré, c'est-à-dire la possibilité de retrouver ses paramètres sans connaître ses entrées.

4.2.1 Identifiabilité structurelle

Dans ce paragraphe, nous présentons l'identifiabilité structurelle définie par Walter et Pronzato. Les résultats de ce paragraphe sont tous issus ou inspirés de [90].

Précisons tout d'abord qu'une propriété est dite structuelle si elle est vraie pour presque toutes les valeurs des paramètres, et éventuellement fausse sur un sous-ensemble de mesure nulle de l'espace d'identification $\mathcal{P}$.

4.2.1.1 Notion d'identifiablité structurelle

Considérons un processus décrit par un modèle $M(.)$ dont on souhaite estimer les paramètres idéaux $\mathbf{p}^\star$. On se place dans le cadre idéal où les données sont non bruitées et où les entrées peuvent être choisies librement. Dans ces conditions, il est toujours possible de trouver un vecteur de paramètres $\hat{\mathbf{p}}$ (il suffit de prendre $\hat{\mathbf{p}} = \mathbf{p}^\star$) de sorte que

$$M(\hat{\mathbf{p}}) = M(\mathbf{p}^\star), \tag{4.9}$$

c'est-à-dire que le comportement entrée-sortie du système de modèle $M(.)$ est identique pour $\mathbf{p}^\star$ et $\hat{\mathbf{p}}$ pour toute entrée en tout temps.
Se pose maintenant le problème d'unicité du vecteur de paramètres $\hat{\mathbf{p}}$ solution. Autrement dit, nous souhaitons savoir si l'égalité (4.9) entraîne forcément l'égalité des paramètres $\hat{\mathbf{p}}$ et $\mathbf{p}^\star$.

Définition 4.6 (Identifiabilité structurelle d'un paramètre)**.** Considérons un vecteur de paramètres

$$\mathbf{p} = \left(p_1, \ldots, p_{n_p}\right) \in \mathcal{P}.$$

Le paramètre p_i est dit structurellement globalement identifiable si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$M(\hat{\mathbf{p}}) = M(\mathbf{p}^\star) \implies \hat{p}_i = p_i^\star. \tag{4.10}$$

Le paramètre p_i est dit structurellement non identifiable si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$, il existe un ensemble non dénombrable de valeurs $\hat{p}_i$ telles que $M(\mathbf{p}^\star) = M(\hat{\mathbf{p}})$.
Le paramètre p_i est dit structurellement localement identifiable si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$, il existe un ensemble dénombrable de valeurs $\hat{p}_i$ telles que $M(\mathbf{p}^\star) = M(\hat{\mathbf{p}})$.

Nous verrons sur un exemple concret (voir remarque 4.9) la nécessité de la restriction *pour presque tout* $\mathbf{p}^\star$.

Par extension de l'identifiabilité structurelle d'un paramètre se définit l'identifiabilité structurelle d'un système paramétré.

Définition 4.7 (Identifiabilité structurelle d'un système)**.** Le système paramétré $M(.)$ est dit *structurellement globalement identifiable* si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$M(\mathbf{p}^\star) = M(\hat{\mathbf{p}}) \implies \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.11}$$

Le système paramétré $M(.)$ est dit *structurellement non identifiable* si au moins un paramètre p_i est structurellement non identifiable.

Remarque 4.8. *Certains paramètres d'un modèle structurellement non identifiable peuvent parfaitement être structurellement globalement identifiables. Nous aurons l'occasion de l'observer dans l'exemple du paragraphe 4.2.1.2.*

Il existe différentes méthodes pour tester algébriquement l'identifiabilité structurelle d'un système paramétré. Nous allons illustrer l'une d'elles sur un exemple. Il s'agit d'une méthode, appelée *approche par transformation de Laplace*, qui s'applique aux systèmes de modèles d'état stationnaires linéaires en entrées[1]. Elle consiste à établir la fonction de transfert $\mathbf{H}(\mathbf{p})$ du système considéré et à résoudre l'équation

$$\mathbf{H}(\mathbf{p}^\star) = \mathbf{H}(\hat{\mathbf{p}}), \tag{4.12}$$

d'inconnue $\hat{\mathbf{p}}$. Ainsi, si, pour presque tout $\mathbf{p}^\star$, l'équation 4.12 admet

- une solution unique pour $\hat{\mathbf{p}}\ (= \mathbf{p}^\star)$, alors le système est structurellement globalement identifiable.
- un ensemble dénombrable de solutions, alors le système est *structurellement localement identifiable*.
- un ensemble non dénombrable de solutions, alors le système est structurellement non identifiable.

Pour davantage de détails concernant les techniques de vérification de l'identifiabilité structurelle d'un système et, particulièrement celle que nous illustrons ci-après, le lecteur pourra consulter [42] et [90].

4.2.1.2 Exemple

L'exemple que nous proposons maintenant est tiré de [42]. Nous considérons le système suivant :

$$\left\{\begin{array}{rcl} \dot{x}_1 & = & -(p_1+p_2)\,x_1 + p_3 x_2 + u \\ \dot{x}_2 & = & p_1 x_1 - p_3 x_2 \\ y & = & x_2. \end{array}\right. \tag{4.13}$$

[1] Un modèle est linéaire en entrées si

$$\forall \lambda_1, \lambda_2 \in \mathbb{R}, y(\mathbf{p}, \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2) = \lambda_1 y(\mathbf{p}, \mathbf{u}_1) + \lambda_2 y(\mathbf{p}, \mathbf{u}_2).$$

Après transformation de Laplace, les équations (4.13) définissant le système s'écrivent

$$\begin{cases} (s+p_1+p_2)\, x_1(s) = p_3 x_2(s) + u(s) \\ (s+p_3)\, x_2(s) = p_1 x_1(s) \\ y(s) = x_2(s) \end{cases} \tag{4.14}$$

En éliminant x_1 et x_2 dans le système (4.14), on obtient

$$(s+p_1+p_2)(s+p_3)\, y(s) = p_1 p_3 y(s) + p_1 u(s)\,, \tag{4.15}$$

ce qui nous permet d'obtenir la fonction de transfert suivante

$$H(s,\mathbf{p}) = \frac{p_1}{s^2 + s(p_1+p_2+p_3) + p_2 p_3}. \tag{4.16}$$

La relation (4.12) se traduit donc par le système d'équations en $\hat{\mathbf{p}}$ suivant :

$$\begin{cases} \hat{p}_1 = p_1^\star \\ \hat{p}_2 + \hat{p}_3 = p_2^\star + p_3^\star \\ \hat{p}_2 \hat{p}_3 = p_2^\star p_3^\star \end{cases} \tag{4.17}$$

qui possède deux solutions

$$\begin{aligned} \hat{\mathbf{p}}_1 &= (p_1^\star, p_2^\star, p_3^\star) \\ \hat{\mathbf{p}}_2 &= (p_1^\star, p_3^\star, p_2^\star)\,. \end{aligned}$$

Ainsi, le paramètre p_1, qui prend la même valeur dans les deux solutions, est structurellement globalement identifiable, alors que les paramètres p_2 et p_3, qui prennent des valeurs différentes, sont seulement structurellement localement identifiables. Autrement dit, à partir de données (les entrées et les sorties) non bruitées, on peut retrouver la valeur exacte du paramètre p_1, mais pour chacun des paramètres p_2 et p_3, on obtiendra deux valeurs possibles et, sans information *a priori*, on ne pourra savoir laquelle est la bonne.

Remarque 4.9. *Cet exemple montre l'intérêt de la précision "pour presque toute valeur de* $\mathbf{p}^\star$*" dans les définitions précédentes. En effet, d'après l'expression (4.16), on voit nettement que si* $p_1^\star = 0$*, la sortie est nulle, quelle que soit l'entrée. Dans ce cas,* p_2 *et* p_3 *sont non identifiables cependant, étant donné que la mesure de l'ensemble* $\{\mathbf{p}^\star | p_1^\star = 0\}$ *est nulle, on considère tout de même* p_2 *et* p_3 *comme structurellement localement identifiables.*

4.2.1.3 Cas particulier des systèmes inversibles

Nous nous intéressons un court paragraphe à l'identifiabilité structurelle dans le cas des systèmes inversibles. La définition 4.7 de l'identifiabilité structurelle devient :

Définition 4.10. Pour $\mathbf{y} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$ fixé, le système inversible paramétré

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$$

est *structurellement globalement identifiable* si pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.18}$$

Remarque 4.11. *Cette définition de l'identifiabilité structurelle pour une sortie $\mathbf{y}$ fixée peut s'étendre naturellement à toutes les sorties $\mathbf{y}$. Il faudrait alors vérifier la condition (4.18) pour presque tout $\mathbf{p} \in \mathcal{P}$ et aussi "pour presque tout $\mathbf{y} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$". Cependant, il convient de vérifier au préalable que $\left(\mathbb{R}^{\mathbb{R}}\right)^n$ est un espace mesurable, ce qui ne semble pas trivial.*

Dans le cas des systèmes inversibles, l'identifiabilité structurelle devient relativement simple à tester, même pour des systèmes non linéaires comme le montre l'exemple suivant. En effet, il n'est plus nécessaire d'utiliser une approche spécifique (par exemple l'approche par transformation de Laplace qui, rappelons le, est uniquement valable pour des systèmes linéaires en entrées).

Exemple 4.12. *Considérons le modèle*

$$\begin{cases} u_1 = -\left(p_1 + p_2\right) y_1 + p_4 \sin(y_2) + p_3 \dot{y}_2 \\ u_2 = p_1 p_2 \dot{y}_1 - p_3 y_2. \end{cases}$$

La relation $\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ se traduit par le système d'équations en $\hat{\mathbf{p}}$ suivant

$$\begin{cases} (p_1^\star + p_2^\star) = (\hat{p}_1 + \hat{p}_2) \\ p_4^\star = \hat{p}_4 \\ p_3^\star = \hat{p}_3 \\ p_1^\star p_2^\star = \hat{p}_1 \hat{p}_2, \end{cases}$$

qui possède une infinité de solutions. On constate que les paramètres p_3 et p_4 sont structurellement globalement identifiables alors que les deux autres paramètres, p_1 et p_2, sont structurellement non identifiables.

Dans ce paragraphe, le concept d'identifiabilité structurelle [90] a été introduit. Nous avons présenté une méthode de test d'identification structurelle sur deux exemples simples. Toutefois, dans la pratique, la vérification de l'identifiabilité structurelle des paramètres d'un système est souvent très complexe et réclame des manipulations algébriques conséquentes. L'utilisation du calcul formel est alors requise (voir annexe D). Dans le cas des systèmes paramétrés non linéaires, l'identifiabilité structurelle est souvent impossible à vérifier tant les calculs sous-jacents sont complexes.

4.2.2 Identifiabilité en aveugle

Nous allons maintenant nous intéresser à l'extension des résultats de la section précédente pour le problème d'estimation de paramètres en aveugle et définir le concept, nouveau, d'identifiabilité (structurelle) en aveugle. Pour des raisons de simplicité, nous présentons ce concept uniquement pour des systèmes paramétrés inversibles, toutefois, il se généralise pour tous les systèmes paramétrés.

Notons d'ores et déjà, que comme pour l'identifiabilité structurelle, il sera en pratique difficile de tester par avance l'identifiabilité en aveugle. Nous l'illustrons néanmoins sur des exemples simples.

4.2.2.1 Définitions

Considérons un système paramétré inversible de la forme

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) \tag{4.19}$$

et $\mathbb{M}$ un modèle statistique pour $\mathbf{u}$ ($\mathbf{u} \in \mathbb{M}$).

Définition 4.13 (Identifiabilité en aveugle d'un paramètre)**.** Considérons un vecteur de paramètres

$$\mathbf{p} = \left(p_1, \ldots, p_{n_p}\right) \in \mathcal{P}.$$

Pour $\mathbf{y} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$ fixé, le paramètre p_i est dit (structurellement globalement) identifiable en aveugle selon le modèle d'entrée $\mathbb{M}$ si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\left\{ \begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array} \right. \Longrightarrow \hat{p}_i = p_i^\star. \tag{4.20}$$

Étendons cette définition au cas des systèmes paramétrés.

Définition 4.14 (Identifiabilité en aveugle d'un système paramétré)**.** Pour $\mathbf{y} \in \left(\mathbb{R}^{\mathbb{R}}\right)^n$ fixé, le système inversible

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$$

est (structurellement globalement) identifiable en aveugle selon le modèle d'entrée $\mathbb{M}$ si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\left\{ \begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array} \right. \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.21}$$

ce qui équivaut à

$$\left\{ \begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array} \right. \Longrightarrow \hat{\mathbf{u}} = \mathbf{u}^\star. \tag{4.22}$$

En d'autres termes, un système est identifiable en aveugle si on peut trouver un seul vecteur de paramètres $\hat{\mathbf{p}}$ $(= \mathbf{p}^\star)$ tel que $\hat{\mathbf{u}} = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M}$.

Remarque 4.15. *Par analogie avec la remarque 4.11, l'identifiabilité en aveugle peut s'étendre "pour presque tout* $\mathbf{y}$*". Par la suite, nous avons choisi, par souci de clarté, de ne plus noter "pour* $\mathbf{y}$ *fixé".*

Notons que la notion d'identifiabilité en aveugle est définie relativement à un modèle d'entrées $\mathbb{M}$.

Remarque 4.16. *Contrairement aux systèmes inversibles séparables selon le modèle* $\mathbb{M}$ *(voir paragraphe 3.2.2.2 page 63), pour lesquels il existe un ensemble de paramètres solutions (tous ceux qui sont en relation avec* $\mathbf{p}$*), signalons que pour les systèmes inversibles identifiables en aveugle selon le modèle* $\mathbb{M}$*, il existe, par définition, un unique vecteur* $\hat{\mathbf{p}}$ $(= \mathbf{p}^\star) \in \mathcal{P}$ *solution tel que* $\hat{\mathbf{u}} \in \mathbb{M}$*.*

Proposition 4.17. *Si le système inversible paramétré*

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$$

est identifiable en aveugle selon le modèle d'entrées $\mathbb{M}$ *alors il est séparable selon* $\mathbb{M}$*.*

Démonstration. Par définition de l'identifiabilité en aveugle, on a, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\left\{ \begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array} \right. \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star.$$

Par définition de la classe d'équivalence $\sim$ qui conserve la forme d'onde (voir paragraphe 3.2.2.1 page 62), on a

$$\hat{\mathbf{p}} = \mathbf{p}^\star \implies \hat{\mathbf{p}} \sim \mathbf{p}^\star.$$

Donc, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\left\{ \begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array} \right. \Longrightarrow \hat{\mathbf{p}} \sim \mathbf{p}^\star.$$

Ainsi, $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$ est séparable selon $\mathbb{M}$. □

De la même façon, on a le résultat suivant.

Proposition 4.18. *Si le système inversible paramétré*

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$$

est identifiable en aveugle pour le modèle $\mathbb{M}$ *alors il est structurellement (globalement) identifiable.*

Démonstration. Par définition de l'identifiabilité en aveugle selon $\mathbb{M}$, on a, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\left\{ \begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array} \right. \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.23}$$

Sachant que $\psi\left(\mathbf{p}^\star, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M}$, on a

$$\psi\left(\mathbf{p}^\star, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) \implies \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M},$$

d'où, d'après (4.23), pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\psi\left(\mathbf{p}^\star, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star,$$

c'est-à-dire que $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$ est structurellement identifiable. □

Remarque 4.19. *Il existe des systèmes structurellement identifiables qui ne sont pas identifiables en aveugle (voir exemple 4.23).*

Nous pouvons maintenant rechercher des conditions d'identifiabilité en aveugle.

4.2.2.2 Conditions d'identifiabilité en aveugle

Pour les systèmes inversibles paramétrés, nous proposons des critères d'identifiabilité en aveugle des paramètres. Ces critères sont complexes et par conséquent, difficilement vérifiables dans la pratique, mais sur des cas particuliers relativement simples, nous verrons qu'ils ont tout de même un intérêt.

Soit $\mathbb{M}$ un modèle d'entrées, l'ensemble des transformations associées à $\mathbb{M}$, noté $\mathcal{T}_{\mathbb{M}}$, est défini par

$$\mathcal{T}_{\mathbb{M}} = \left\{ \tau : \left\{ \begin{array}{ccc} \left(\mathbb{R}^{\mathbb{R}}\right)^n & \rightarrow & \left(\mathbb{R}^{\mathbb{R}}\right)^n \\ \mathbf{u} & \mapsto & \tau(\mathbf{u}) \end{array} \right. | \mathbf{u} \in \mathbb{M} \implies \tau(\mathbf{u}) \in \mathbb{M} \right\}.$$

Autrement dit, $\mathcal{T}_{\mathbb{M}}$ est l'ensemble des transformations conservant le modèle d'entrée $\mathbb{M}$. Dans la pratique, les ensembles $\mathcal{T}_{\mathbb{M}}$ sont excessivement complexes à caractériser. On ne connaît bien souvent qu'un sous-ensemble de $\mathcal{T}_{\mathbb{M}}$.

Exemple 4.20. *D'après la proposition B.54 (page 184), l'ensemble $\mathcal{T}_{\mathcal{G}_s}$ des transformations associées au modèle d'entrées gaussiennes vérifie*

$$\mathfrak{L}^n \subset \mathcal{T}_{\mathcal{G}_s}.$$

D'après la proposition B.39 (page 178), l'ensemble $\mathcal{T}_{I_s}$ des transformations associées au modèle d'entrées indépendantes vérifie

$$\mathfrak{F}^n \subset \mathcal{T}_{I_s}.$$

Théorème 4.21. *Considérons le système inversible*

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \tag{4.24}$$

et un modèle d'entrée $\mathbb{M}$ ($\mathbf{u} \in \mathbb{M}$). Soit $\mathcal{T}_{\mathbb{M}}$ l'ensemble des transformations associées à $\mathbb{M}$. Le système (4.24) est identifiable en aveugle selon $\mathbb{M}$ si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\tau\left(\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star, \tag{4.25}$$

avec $\tau \in \mathcal{T}_{\mathbb{M}}$.

Démonstration. ($\Longrightarrow$) Supposons que, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$

$$\left\{\begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array}\right. \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.26}$$

Or, $\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M}$, donc par définition de $\mathcal{T}_{\mathbb{M}}$, on a

$$\forall \tau \in \mathcal{T}_{\mathbb{M}},\ \tau\left(\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) \in \mathbb{M}.$$

Ainsi, on a

$$\tau\left(\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star.$$

($\Longleftarrow$) Supposons que, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\tau\left(\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star, \tag{4.27}$$

avec $\tau \in \mathcal{T}_{\mathbb{M}}$. On a :

$$\left\{\begin{array}{c} \psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \\ \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \in \mathbb{M} \end{array}\right. \Longrightarrow \exists \tau \in \mathcal{T}_{\mathbb{M}}, \tau\left(\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.28}$$

Donc le système est identifiable en aveugle selon $\mathbb{M}$. $\square$

L'identifiabilité en aveugle consiste à vérifier, pour toutes les fonctions τ associées aux modèle d'entrées $\mathbb{M}$ vérifiant

$$\tau\left(\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right),$$

qu'on a bien l'égalité $\hat{\mathbf{p}} = \mathbf{p}^\star$.

Dans le cas des modèles d'entrées gaussiennes ou indépendantes, évoqués dans l'exemple 4.20, ce théorème est difficilement applicable dans la mesure où seul un sous-ensemble ($\mathfrak{L}^n$ ou $\mathfrak{F}^n$) de transformations inclus dans $\mathcal{T}_{\mathbb{M}}$ est connu. Toutefois, comme nous allons le voir dans les deux paragraphes suivants, pour les modèles gaussiens et indépendants, il "semble" (résultat non formellement démontré) inutile de considérer toutes les transformations de $\mathcal{T}_{\mathbb{M}}$.

Commençons par remarquer que peu de fonctions $\tau \in \mathcal{T}_{\mathbb{M}}$ permettent d'obtenir l'égalité

$$\tau\left(\psi\left(\mathbf{p}^{\star}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right), \tag{4.29}$$

c'est-à-dire que l'influence de la fonction $\tau \in \mathcal{T}_{\mathbb{M}}$ peut rarement être "absorbée" par la fonction ψ moyennant un changement de valeurs des paramètes ($\mathbf{p}^{\star} \rightarrow \hat{\mathbf{p}}$). Notamment, plus la fonction τ est complexe, moins il y a de chances que l'égalité (4.29) soit vérifiée. Les résultats des paragraphes suivants sont fondés sur cette remarque intuitive.

4.2.2.2.1 Modèle d'entrées gaussiennes Dans ce paragraphe, nous nous intéressons à l'identifiabilité en aveugle pour le modèle d'entrées mutuellement gaussiennes.

Considérons le système inversible $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ et le modèle d'entrées gaussiennes $\mathbb{M} = \mathcal{G}_s$. Soit $\mathcal{T}_{\mathcal{G}_s}$ l'ensemble des transformations associées au modèle $\mathcal{G}_s$. Intuitivement, il semble suffisant de vérifier la condition (4.27) du théorème 4.21 uniquement pour les transformations $\tau \in \mathfrak{L}^n \subset \mathcal{T}_{\mathcal{G}_s}$ de la forme

$$\tau : \left\{ \begin{array}{ccc} \left(\mathbb{R}^{\mathbb{R}}\right)^n & \rightarrow & \left(\mathbb{R}^{\mathbb{R}}\right)^n \\ \mathbf{u} & \rightarrow & \mathbf{M}\mathbf{u} + \mathbf{c} \end{array} \right., \tag{4.30}$$

avec $\mathbf{M}$ une matrice quelconque et $\mathbf{c}$ un vecteur constant de $\mathbb{R}^n$. En effet, considérons la transformation $\tau_1 \in \mathcal{T}_{\mathcal{G}_s}$ suivante,

$$\tau_1 : \left\{ \begin{array}{ccc} \left(\mathbb{R}^{\mathbb{R}}\right)^n & \rightarrow & \left(\mathbb{R}^{\mathbb{R}}\right)^n \\ \mathbf{u} & \rightarrow & \dot{\mathbf{u}} \end{array} \right.,$$

On a $\tau_1\left(\psi\left(\mathbf{p}^{\star}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \dot{\psi}\left(\mathbf{p}^{\star}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ et l'égalité (4.29) s'écrit

$$\dot{\psi}\left(\mathbf{p}^{\star}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right).$$

Cette dernière égalité n'est jamais vérifiée (sauf si ψ est nulle!).

Ainsi, nous proposons la conjecture suivante qui découle du théorème 4.21.

Conjecture 4.22. *Le système $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ est identifiable en aveugle selon le modèle d'entrées mutuellement gaussiennes si et seulement si, pour presque tout $\mathbf{p}^{\star} \in \mathcal{P}$,*

$$\mathbf{M}\psi\left(\mathbf{p}^{\star}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) + \mathbf{c} = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^{\star}. \tag{4.31}$$

avec $\mathbf{M}$ une matrice quelconque de dimension $n \times n$ et $\mathbf{c} \in \mathbb{R}^n$.

Montrer l'identifiabilité en aveugle pour le modèle gaussien revient, si cette conjecture est vraie, à vérifier que pour toutes les matrices $\mathbf{M} \in \mathbb{R}^{n \times n}$ et toutes les constantes $\mathbf{c} \in \mathbb{R}^n$ telles que

$$\mathbf{M}\psi\left(\mathbf{p}^{\star}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) + \mathbf{c} = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right),$$

on a bien $\hat{\mathbf{p}} = \mathbf{p}^{\star}$.

Nous proposons deux exemples simples pour illustrer ce résultat :

Exemple 4.23 (Système non identifiable en aveugle). *Considérons le système inversible*

$$u = \psi(p, y) = py, \tag{4.32}$$

et le modèle $\mathbb{M} = \mathcal{G}_s$ pour l'entrée u. La condition (4.31) devient, pour presque tout $p^\star \in \mathcal{P}$,

$$\alpha p^\star y + \beta = \hat{p}y \implies \hat{p} = p^\star.$$

avec $\alpha \in \mathbb{R},\ \beta \in \mathbb{R}$. Or

$$\alpha p^\star y + \beta = \hat{p}y \iff \left\{ \begin{array}{c} \hat{p} = \alpha p^\star \\ \beta = 0, \end{array} \right.$$

qui n'implique pas $\hat{p} = p^\star$ (sauf si $p^\star = 0$). Dès lors le système (4.32) est non identifiable en aveugle pour le modèle gaussien.
Notons qu'il est toutefois structurellement identifiable (voir le paragraphe 4.2.1.1).

Exemple 4.24 (Système identifiable en aveugle). *Considérons le système inversible*

$$u = \psi(p, y) = \sin(py), \tag{4.33}$$

et le modèle d'entrées $\mathbb{M} = \mathcal{G}_s$. La condition (4.31) devient pour presque tout $p^\star \in \mathcal{P}$,

$$\sin(\hat{p}y) = \alpha \sin(p^\star y) + \beta \implies \hat{p} = p^\star.$$

avec $\alpha \in \mathbb{R},\ \beta \in \mathbb{R}$. Or

$$\sin(\hat{p}y) = \alpha \sin(p^\star y) + \beta \iff \left\{ \begin{array}{l} \hat{p} = p^\star \\ \alpha = 1 \\ \beta = 0. \end{array} \right.$$

Donc le système (4.33) est identifiable en aveugle pour le modèle gaussien.

Remarque 4.25. *Si on considère le modèle d'entrées deux à deux gaussiennes (plutôt que le modèle d'entrées mutuellement gaussiennes), la condition 4.31 d'identifiabilité en aveugle devient :*

Le système $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ est identifiable en aveugle selon le modèle d'entrées deux à deux gaussiennes si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\mathbf{PD}\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) + \mathbf{c} = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.34}$$

avec $\mathbf{P} \in \mathfrak{P}$, $\mathbf{D} \in \mathfrak{D}$ et $\mathbf{c} \in \mathbb{R}^n$.

4.2.2.2.2 Modèle d'entrées indépendantes Dans ce paragraphe, nous nous intéressons à l'identifiabilité en aveugle pour le modèle d'entrées indépendantes.

Considérons le système inversible $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ et le modèle d'entrées indépendantes $\mathbb{M} = I_s$. Soit $\mathcal{T}_{I_s}$ l'ensemble des transformations associées au modèle I_s. Par analogie avec le paragraphe précédent, il semble intuitivement suffisant de vérifier la condition (4.27) pour les transformations $\tau \in \mathfrak{F}^n \subset \mathcal{T}_{I_s}$ de la forme 4.34. En effet, considérons par exemple les transformations suivantes,

$$\tau_1 : \left\{ \begin{array}{ccc} \left(\mathbb{R}^{\mathbb{R}}\right)^n & \rightarrow & \left(\mathbb{R}^{\mathbb{R}}\right)^n \\ \mathbf{u} & \rightarrow & \mathbf{u}^2 \end{array} \right. ,$$

$$\tau_2 : \left\{ \begin{array}{ccc} \left(\mathbb{R}^{\mathbb{R}}\right)^n & \rightarrow & \left(\mathbb{R}^{\mathbb{R}}\right)^n \\ \mathbf{u} & \rightarrow & \cos(\mathbf{u}) \end{array} \right. .$$

L'égalité (4.29) n'est jamais vérifiée.

Nous proposons donc la conjecture suivante.

Conjecture 4.26. *Le système $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ est identifiable en aveugle selon le modèle d'entrées indépendantes si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,*

$$\mathbf{PD}\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) + \mathbf{c} = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.35}$$

avec $\mathbf{P} \in \mathfrak{P}$, $\mathbf{D} \in \mathfrak{D}$ et $\mathbf{c} \in \mathbb{R}^n$.

Montrer l'identifiabilité en aveugle pour le modèle indépendant revient, si cette conjecture est vraie, à vérifier que pour toutes les matrices $\mathbf{P} \in \mathfrak{P}, \mathbf{D} \in \mathfrak{D}$ et toutes les constantes $\mathbf{c} \in \mathbb{R}^n$ telles que

$$\mathbf{PD}\psi\left(\mathbf{p}^\star, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) + \mathbf{c} = \psi\left(\hat{\mathbf{p}}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right),$$

on a bien $\hat{\mathbf{p}} = \mathbf{p}^\star$.

Nous illustrons ce résultat par un exemple simple.

Exemple 4.27 (Système non identifiable en aveugle)**.** *Considérons le système inversible*

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} y_1 + p_1 y_2 \\ p_2 y_1 + y_2 \end{pmatrix} \tag{4.36}$$

et le modèle d'entrées $\mathbb{M} = I_s$. La condition (4.35) devient, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\left\{ \begin{array}{l} \begin{pmatrix} y_1 + \hat{p}_1 y_2 \\ \hat{p}_2 y_1 + y_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \begin{pmatrix} y_1 + p_1^\star y_2 \\ p_2^\star y_1 + y_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\ \begin{pmatrix} y_1 + \hat{p}_1 y_2 \\ \hat{p}_2 y_1 + y_2 \end{pmatrix} = \begin{pmatrix} 0 & \alpha_1 \\ \alpha_2 & 0 \end{pmatrix} \begin{pmatrix} y_1 + p_1^\star y_2 \\ p_2^\star y_1 + y_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \end{array} \right. \implies \hat{\mathbf{p}} = \mathbf{p}^\star,$$

avec $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ et $(\beta_1, \beta_2) \in \mathbb{R}^2$, ce qui équivaut à

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} y_1 + \hat{p}_1 y_2 = \alpha_1 \left(y_1 + p_1^\star y_2\right) + \beta_1 \\ \hat{p}_2 y_1 + y_2 = \alpha_2 \left(p_2^\star y_1 + y_2\right) + \beta_2 \end{array} \right. \\ \left\{ \begin{array}{l} y_1 + \hat{p}_1 y_2 = \alpha_1 \left(p_2^\star y_1 + y_2\right) + \beta_1 \\ \hat{p}_2 y_1 + y_2 = \alpha_2 \left(y_1 + p_1^\star y_2\right) + \beta_2 \end{array} \right. \end{array} \right. \implies \hat{\mathbf{p}} = \mathbf{p}^\star.$$

Or, en résolvant les deux systèmes d'équations d'inconnues $\hat{\mathbf{p}}$, on a

$$\left\{ \begin{array}{l} y_1 + \hat{p}_1 y_2 = \alpha_1 \left(y_1 + p_1^\star y_2\right) + \beta_1 \\ \hat{p}_2 y_1 + y_2 = \alpha_2 \left(p_2^\star y_1 + y_2\right) + \beta_2 \end{array} \right. \iff \left\{ \begin{array}{l} \alpha_1 = \alpha_2 = 1 \\ \beta_1 = \beta_2 = 0 \\ \hat{p}_1 = p_1^\star \\ \hat{p}_2 = p_2^\star \end{array} \right., \tag{4.37}$$

et

$$\left\{ \begin{array}{l} y_1 + \hat{p}_1 y_2 = \alpha_1 \left(p_2^\star y_1 + y_2\right) + \beta_1 \\ \hat{p}_2 y_1 + y_2 = \alpha_2 \left(y_1 + p_1^\star y_2\right) + \beta_2 \end{array} \right. \iff \left\{ \begin{array}{l} \alpha_1 = p_1^\star, \alpha_2 = p_2^\star \\ \beta_1 = \beta_2 = 0 \\ p_1^\star = \frac{1}{\hat{p}_2} \\ p_2^\star = \frac{1}{\hat{p}_1} \end{array} \right. . \tag{4.38}$$

Ce qui n'implique pas $\hat{\mathbf{p}} = \mathbf{p}^\star$ (voir le système (4.38)). Donc, le système (4.36) n'est pas identifiable en aveugle pour le modèle indépendant.

Remarque 4.28. *Dans la pratique, il n'est pas toujours nécessaire de vérifier la condition d'identifiabilité (4.35) pour toutes les matrices de permutation. On peut souvent se limiter à certaines permutations pour les systèmes ne possédant pas une "bonne symétrie". Par exemple, si on considère le système*

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} y_1 + p_1 y_2 \\ p_2 y_1 + \dot{y}_2 \end{pmatrix}, \tag{4.39}$$

il est clair que l'égalité

$$\begin{pmatrix} 0 & \alpha_1 \\ \alpha_2 & 0 \end{pmatrix} \begin{pmatrix} y_1 + p_1^\star y_2 \\ p_2^\star y_1 + \dot{y}_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} y_1 + \hat{p}_1 y_2 \\ p_2 \hat{y}_1 + \dot{y}_2 \end{pmatrix} \tag{4.40}$$

ne sera jamais vérifiée dans la mesure où les deux équations induites sont du type

$$\begin{aligned} f_1(y_1, \dot{y}_2) &= g_1(y_1, y_2), \\ f_2(y_1, y_2) &= g_2(y_1, \dot{y}_2). \end{aligned} \tag{4.41}$$

Il suffit de vérifier que

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \begin{pmatrix} y_1 + p_1^\star y_2 \\ p_2^\star y_1 + \dot{y}_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} y_1 + \hat{p}_1 y_2 \\ p_2 \hat{y}_1 + \dot{y}_2 \end{pmatrix} \implies \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.42}$$

Or

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \begin{pmatrix} y_1 + p_1^\star y_2 \\ p_2^\star y_1 + \dot{y}_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} y_1 + \hat{p}_1 y_2 \\ p_2 \hat{y}_1 + \dot{y}_2 \end{pmatrix} \Longleftrightarrow \begin{cases} \alpha_1 = \alpha_2 = 1 \\ \beta_1 = \beta_2 = 0 \\ \hat{p}_1 = p_1^\star \\ \hat{p}_2 = p_2^\star \end{cases}, \tag{4.43}$$

le système (4.39) est donc identifiable en aveugle pour le modèle d'entrées indépendantes.

Dans le paragraphe suivant, nous allons présenter une sous-classe de systèmes inversibles particuliers, pour lesquels nous proposons une méthode de résolution du problème d'estimation de paramètres en aveugle. il s'agit des systèmes inversibles découplés.

4.3 Systèmes inversibles découplés

Dans cette partie, nous considérons une sous-classe des systèmes inversibles, appelés systèmes inversibles découplés, pour lesquels nous verrons qu'il est toujours possible d'estimer l'ensemble des paramètres solutions grâce aux techniques d'analyse par intervalles.

Notons que pour ces systèmes inversibles découplés, les paramètres interviennent de façon affine sur les sorties dans le système inversible.

4.3.1 Définitions

Afin de définir un système inversible découplé, nous introduisons la notion de fonction séparable. Attention, le terme séparable utilisé ici provient de la notion de *fonction à variables séparables*, généralement utilisée pour la résolution d'équations différentielles. Il n'a absolument aucun lien avec la notion de séparabilité au sens de la séparation de sources.

4.3.1.1 Fonction séparable

Définition 4.29 (Fonction séparable)**.** On définit par induction l'ensemble $\mathcal{Q}$ des fonctions séparables de sorte que

$$\begin{array}{ll} \text{i)} & f(\mathbf{p}).\phi\left(\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) \in \mathcal{Q}, \\ \text{ii)} & \left\{\begin{array}{l} \varphi_1(.) \in \mathcal{Q} \\ \varphi_2(.) \in \mathcal{Q} \end{array}\right. \Rightarrow \left\{\begin{array}{l} \varphi_1\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) + \varphi_2\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) \in \mathcal{Q} \\ \varphi_1\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) \times \varphi_2\left(\mathbf{p},\mathbf{y},\ldots,\mathbf{y}^{(r)}\right) \in \mathcal{Q} \end{array}\right. , \end{array} \tag{4.44}$$

où f, φ, φ_1 et φ_2 sont des fonctions analytiques.
En d'autres termes, une fonction séparable est une somme ou un produit de fonctions de la forme i).

Exemple 4.30. *La fonction*

$$\varphi\left(\mathbf{p}, y, \ldots, y^{(r)}\right) = \left((p_1 p_2 \exp(y\dot{y}))^2 + \log(p_1)\, y\right)^3 \ddot{y}$$

est une fonction séparable (somme et produit de fonctions séparables) alors que la fonction

$$\varphi(p, y) = \sin(py)$$

ne l'est pas.

Proposition 4.31. *Toute fonction séparable φ s'écrit sous la forme vectorielle :*

$$\varphi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \mathbf{f}(\mathbf{p})^T . \boldsymbol{\phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right). \tag{4.45}$$

Démonstration. Par définition, toute fonction séparable s'écrit comme somme et produit de fonctions séparables, c'est-à-dire

$$\varphi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \sum_i f_i(\mathbf{p}) . \phi_i\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right).$$

Et, en posant

$$\begin{aligned} \mathbf{f}(\mathbf{p}) &= (f_1(\mathbf{p}), \ldots, f_i(\mathbf{p}), \ldots)^{\mathrm{T}}, \\ \boldsymbol{\phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) &= \left(\phi_1\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right), \ldots, \phi_i\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right), \ldots\right)^{\mathrm{T}}, \end{aligned}$$

on obtient l'égalité (4.45). □

Exemple 4.32. *Considérons la fonction*

$$\varphi(p, y, \dot{y}, \ddot{y}) = p^2 y - \sin(\ddot{y}) + \cos(p).$$

Posons

$$\begin{aligned} \mathbf{f}(p) &= \left(p^2, -1, \cos(p)\right)^T, \\ \boldsymbol{\phi}(y, \dot{y}, \ddot{y}) &= (y, \sin(\ddot{y}), 1)^T, \end{aligned}$$

on a

$$\varphi(p, y, \dot{y}, \ddot{y}) = \mathbf{f}(p).\boldsymbol{\phi}(y, \dot{y}, \ddot{y}),$$

donc la fonction φ est séparable.

4.3.1.2 Système découplé

Par extension, nous définissons maintenant un système inversible découplé. Nous avons volontairement choisi le terme découplé (plutôt que séparable) afin d'éviter toute confusion avec la notion de séparabilité. Attirons toutefois l'attention sur le fait que notre définition d'un système découplé n'est pas celle classiquement rencontrée en automatique. Nous avons néanmoins employé ce terme sachant que, dans le contexte d'estimation de paramètres en aveugle, aucune ambiguïté n'est possible.

Définition 4.33 (Système inversible découplé)**.** Un système inversible

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$$

est découplé si toutes les composantes de ψ sont des fonctions séparables.

Par analogie avec la proposition 4.31, on obtient la proposition suivante :

Proposition 4.34. *Tout système inversible découplé ψ s'écrit sous la forme*

$$\psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \mathbf{F}(\mathbf{p})^T . \mathbf{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right).$$

où $\mathbf{F}(\mathbf{p})$ est une matrice dépendant uniquement du vecteur de paramètres $\mathbf{p}$ et $\mathbf{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$ une fonction vectorielle dépendant uniquement des sorties et de leurs dérivées.

Démonstration. Considérons le système inversible découplé

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \begin{pmatrix} \psi_1\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \\ \vdots \\ \psi_n\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \end{pmatrix}.$$

Par définition, on sait que les composantes ψ_i de ψ sont des fonctions séparables. Donc, d'après la proposition 4.31, on a

$$\psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \begin{pmatrix} \mathbf{f}_1(\mathbf{p}) . \boldsymbol{\phi}_1\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \\ \vdots \\ \mathbf{f}_n(\mathbf{p}) . \boldsymbol{\phi}_n\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) \end{pmatrix}.$$

Et, en posant $\mathbf{F}(\mathbf{p}) = \left(\mathbf{f}_1(\mathbf{p}), \ldots, \mathbf{f}_n(\mathbf{p})\right)^{\mathrm{T}}$ une matrice dont les colonnes sont les $\mathbf{f}_i$ et $\mathbf{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \left(\boldsymbol{\phi}_1\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)^{\mathrm{T}}, \ldots, \boldsymbol{\phi}_n\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)^{\mathrm{T}}\right)^{\mathrm{T}}$ un vecteur composé des $\boldsymbol{\phi}_i$, on obtient

$$\psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \mathbf{F}(\mathbf{p})^{\mathrm{T}} . \mathbf{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right).$$

□

Ainsi pour un système inversible découplé, l'influence des paramètres et celle des sorties et de leurs dérivées peuvent toujours être séparées dans la fonction ψ.

Exemple 4.35. *Considérons le système inversible*

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \sin(p_1)\, y_1^3 - \frac{p_1}{p_2}\exp(\ddot{y}_2) \\ p_1^2 p_2 \dot{y}_1^2 - p_2\cos(y_2) \end{pmatrix}.$$

Posons

$$\mathbf{F}(\mathbf{p}) = \begin{pmatrix} \sin(p_1) & -\frac{p_1}{p_2} & 0 & 0 \\ 0 & 0 & p_1^2 p_2 & -p_2 \end{pmatrix}^T,$$

$$\boldsymbol{\Phi}(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}) = \left(y_1^3, \exp(\ddot{y}_2), \dot{y}_1^2, \cos(y_2)\right)^T,$$

on a

$$\psi(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}) = \mathbf{F}(\mathbf{p})^T.\boldsymbol{\Phi}(\mathbf{y}, \dot{\mathbf{y}}, \ddot{\mathbf{y}}).$$

4.4 Estimation ensembliste de paramètres en aveugle pour des systèmes découplés

Dans cette partie, nous proposons une méthodologie de résolution du problème d'estimation de paramètres en aveugle pour des systèmes inversibles découplés.

On considère un système paramétré inversible découplé

$$\mathbf{u} = \mathbf{F}(\mathbf{p})^{\mathrm{T}}.\boldsymbol{\Phi}\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right),$$

où les signaux $\mathbf{y}^{(i)}(t)$ sont connus.
Soit $\mathbb{M}$ un modèle d'entrée pour ce système. D'après les résultats du paragraphe 4.1, le problème consiste à caractériser l'ensemble

$$\mathbb{P} = \left\{\mathbf{p} \in \mathcal{P} | \mathbf{h}\left(\mathbf{F}(\mathbf{p})^{\mathrm{T}}.\boldsymbol{\Phi}\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)\right) = \mathbf{0}\right\}, \tag{4.46}$$

où $\mathbf{h}$ est une fonction d'estimation correspondant au modèle d'entrées $\mathbb{M}$ considéré.

La méthode que nous proposons ici utilise l'analyse par intervalles (notamment l'algorithme SIVIA) pour caractériser l'ensemble $\mathbb{P}$.

Dans un premier temps, nous détaillons la construction de fonctions d'estimation adaptées au modèle d'entrées $\mathbb{M}$ considéré. Ensuite, à l'aide de techniques similaires à celle du paragraphe 3.2.4.1 (page 79), nous approximons ces fonctions dans le but d'obtenir des équations d'estimation exploitables.

Ensuite, nous considérons un exemple simple illustrant la méthodologie proposée. Pour cet exemple, nous supposons d'abord que les entrées sont statistiquement indépendantes puis qu'une entrée est gaussienne afin d'estimer en aveugle les paramètres du système.

Nous terminons ce chapitre par un exemple plus concret montrant l'intérêt de la méthode.

4.4.1 Fonctions d'estimation

Dans ce paragraphe, nous allons détailler la construction de fonctions d'estimation pour différents modèles statistiques d'entrées $\mathbb{M}$. Nous considérons tout d'abord le modèle d'entrées indépendantes $\mathbb{M} = I_s$. Ensuite, nous traitons le modèle d'entrées gaussiennes $\mathbb{M} = \mathcal{G}_s$. Pour mémoire, les fonctions d'estimation $\mathbf{h}$ sont construites de sorte que

$$\mathbf{u} \in \mathbb{M} \implies \mathbf{h}(\mathbf{u}) = \mathbf{0}.$$

4.4.1.1 Modèle d'entrées indépendantes

Tout d'abord, rappelons (voir page 172) quelques propriétés fondamentales des moments d'un vecteur de signaux aléatoires stationnaires, ergodiques et lisses.

Soit $\mathbf{u} = (u_1, \ldots, u_n) \in \mathcal{S}^n$ un vecteur de signaux aléatoires stationnaires, ergodiques et lisses. Ses fonctions d'intercorrélations, définies par

$$\Gamma_{u_i^{(k)}, u_j^{(l)}}(\tau) = E\left(u_i^{(k)}(t)\, u_j^{(l)}(t-\tau)\right) - E\left(u_i^{(k)}(t)\right) E\left(u_j^{(l)}(t-\tau)\right),$$

vérifient

$$i) \quad \Gamma_{u_i^{(k)}, u_j^{(l)}}(\tau) = E\left(u_i^{(k)}(t)\, u_j^{(l)}(t-\tau)\right), \text{ si } k+l \geq 1, \tag{4.47}$$

$$ii) \quad k_1 + l_1 = k_2 + l_2 \implies \Gamma_{u_i^{(k_1)}, u_j^{(l_1)}}(\tau) = (-1)^{l_1 - l_2}\, \Gamma_{u_i^{(k_2)}, u_j^{(l_2)}}(\tau), \tag{4.48}$$

$$iii) \quad \Gamma_{u_i^{(k)}, u_j^{(l)}}(\tau) = 0, \text{ si } k+l \text{ est impaire.} \tag{4.49}$$

Dans ce paragraphe, nous supposons que le modèle pour les signaux d'entrées $\mathbf{u}$ est $\mathbb{M} = I_s$, c'est-à-dire

$$I_s(u_1, \ldots u_n).$$

Alors, pour tout entier $k \geq 0$ et tout entier $l \geq 0$, et pour tout $\tau \in \mathbb{R}$, on a $\forall i, j$ tel que $i \neq j$,

$$\Gamma_{u_i^{(k)}, u_j^{(l)}}(\tau) = 0.$$

Pour simplifier, nous posons $\tau = 0$ afin d'utiliser uniquement l'influence des dérivations successives des fonctions d'intercorrélation et notons

$$\Gamma_{u_i^{(k)}, u_j^{(l)}} = \Gamma_{u_i^{(k)}, u_j^{(l)}}(0).$$

D'après la relation (4.48), on peut voir qu'il existe des relations entre certains moments. Par exemple, $\Gamma_{u_i^{(2)}, u_j} = -\Gamma_{u_i^{(1)}, u_j^{(1)}}$ ou encore $\Gamma_{u_i^{(2)}, u_j} = \Gamma_{u_i, u_j^{(2)}}$. Par conséquent, dans le but d'éliminer les termes redondants, nous choisissons la fonction d'estimation suivante :

$$\mathbf{h} : \begin{cases} \mathcal{S}^n & \rightarrow \quad \mathbb{R}^{\frac{(n-1)n}{2}\frac{(q+1)q}{2}} \\ \mathbf{u} & \mapsto \quad \left(\Gamma_{u_i^{(k)}, u_j}\right)_{i=1,\ldots,n,\ j=i+1,\ldots,n}^{k=0,\ldots,q} \end{cases} \tag{4.50}$$

où l'entier q est l'ordre maximal de dérivation des signaux d'entrées.

Exemple 4.36. *Pour deux signaux d'entrée u_1 et u_2, la fonction $\mathbf{h}$ définie par*

$$\mathbf{h}(\mathbf{u}) = \begin{pmatrix} \Gamma_{u_1, u_2} \\ \Gamma_{u_1^{(1)}, u_2} \\ \Gamma_{u_1^{(2)}, u_2} \end{pmatrix}$$

est une fonction d'estimation pour le modèle d'entrées indépendantes.

En pratique, l'entier q est choisi de façon à ce que le nombre d'équations d'estimation (les composantes de $\mathbf{h}$) soit supérieur ou égal au nombre d'inconnues. Plus q est grand, plus le nombre d'équations est grand par rapport au nombre d'inconnues, et on peut ainsi espérer obtenir une résolution robuste.

Remarquons que, dans les fonctions d'estimation (4.50), nous utilisons uniquement des moments d'ordre deux. Toutefois, il est clair que des moments d'ordre supérieur peuvent être ajoutés.

4.4.1.2 Modèle d'entrées gaussiennes

Nous supposons maintenant que les signaux aléatoires d'entrées sont mutuellement gaussiens, c'est-à-dire

$$\mathcal{G}_s(u_1, \ldots, u_n).$$

La distance par rapport à la gaussianité peut être mesurée en utilisant le kurtosis (voir paragraphe B.6.4 page 186), qui est nul pour un signal gaussien. De plus, d'après les propositions B.62 (page 187)et B.25 (page 170), les dérivées des signaux d'entrée sont centrées et encore gaussiennes. Par conséquent, la fonction d'estimation $\mathbf{h}$ suivante peut être utilisée pour mesurer la gaussianité de $\mathbf{u}$:

$$\mathbf{h} : \begin{cases} \mathcal{S}^n & \rightarrow & \mathbb{R}^{q-1} \\ \mathbf{u} & \mapsto & kurt\left(u_i^{(k)}\right)_{i=1,\ldots,n}^{k=0,\ldots,q} \end{cases}$$

où l'entier q (qui est l'ordre de dérivation maximal) doit être choisi de telle sorte que le nombre d'équations d'estimation soit supérieur ou égal au nombre d'inconnues.

Exemple 4.37. *La fonction* $\mathbf{h}$ *définie par*

$$\mathbf{h}(\mathbf{u}) = \begin{pmatrix} kurt(u_1) \\ \vdots \\ kurt(u_n) \end{pmatrix}$$

est une fonction d'estimation pour le modèle d'entrées gaussiennes.
A noter qu'il est possible de rajouter $kurt(u_1 + u_2), \ldots$

4.4.1.3 Approximation des fonctions d'estimation

Les sorties $\mathbf{y}$ étant connues, rappelons que l'ensemble $\mathbb{P}$ à caractériser est défini à partir de l'équation vectorielle suivante (voir équation (4.46)) :

$$\mathbf{g}(\mathbf{p}) \triangleq \mathbf{h}\left(\mathbf{F}(\mathbf{p})^{\mathrm{T}}.\mathbf{\Phi}\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)\right) = \mathbf{0}.$$

En pratique, seule une estimée de $\mathbf{g}$ peut être obtenue, en effet,

- Sachant que le système $\mathbf{u} = \mathbf{F}(\mathbf{p})^{\mathrm{T}}.\mathbf{\Phi}\left(\mathbf{y}(t), \ldots, \mathbf{y}^{(r)}(t)\right)$ est non linéaire ou instable, le vecteur de signaux de sortie $\mathbf{y}$ est rarement stationnaire, même si $\mathbf{u} \in \mathcal{S}^n$.
- Même si les signaux de sorties $(y_1, \ldots, y_n)$ sont stationnaires, leurs moments peuvent seulement être estimés.

Nous allons voir que, pour les systèmes inversibles découplés, la fonction d'estimation $\mathbf{g}(\mathbf{p})$ (ou encore $\mathbf{h}(\mathbf{u})$) s'approxime par une fonction dépendant uniquement des paramètres $\mathbf{p}$ et des estimations des moments des sorties (connues). Pour cela, comme dans le paragraphe 3.2.3.1 (page 65), nous allons utiliser l'opérateur *bar*. Rappelons que pour un signal stationnaire s, l'estimateur empirique $\bar{s}$ de $E(s)$ est défini par

$$\bar{s} \triangleq \frac{1}{N} \sum_{k=0}^{N} s(k\delta),$$

où N est le nombre d'échantillons et δ la période d'échantillonnage.

Nous proposons le résultat suivant :

Proposition 4.38. *Soient* $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$ *un système découplé et* $\mathbf{h}$ *une fonction d'estimation d'un modèle* $\mathbb{M}$ *fixé. Il existe une fonction* $\hat{\mathbf{g}}$ *qui approxime* $\mathbf{g}(\mathbf{p}) \triangleq \mathbf{h}(\mathbf{u})$ *de la forme*

$$\hat{\mathbf{g}}(\mathbf{p}) = \boldsymbol{\beta}\left(\mathbf{p}, \bar{\mu}_1(\mathbf{y}), \bar{\mu}_2(\mathbf{y}), \ldots, \bar{\mu}_i(\mathbf{y}), \ldots\right). \tag{4.51}$$

où les $\bar{\mu}_i(\mathbf{y})$ *sont des estimations des moments du signal de sortie* $\mathbf{y}$ *(et de ses dérivées).*

Démonstration. Soit $\mathbf{h} : \mathcal{S}^n \to \mathbb{R}^q$ une fontion d'estimation où q est un entier arbitraire. Par définition, les composantes de $\mathbf{h}$ appartiennent à l'algèbre de moments $\mathcal{A}_{\mathbf{u}}$ (*i.e.* les composantes de $\mathbf{h}$ s'écrivent comme des sommes et produits de moments pondérés du signal $\mathbf{u}$). Ainsi,

$$\mathbf{h}(\mathbf{u}) = \boldsymbol{\delta}\left(\mu_1'(\mathbf{u}), \mu_2'(\mathbf{u}), \ldots, \mu_i'(\mathbf{u}), \ldots\right), \tag{4.52}$$

où $\mu_i'(\mathbf{u}) \in \mathcal{M}_{\mathbf{u}}$.
A l'aide de l'opérateur *bar*, la fonction d'estimation $\mathbf{g}(\mathbf{p}) \triangleq \mathbf{h}(\mathbf{u})$ s'approxime par

$$\hat{\mathbf{g}}(\mathbf{p}) = \boldsymbol{\delta}\left(\bar{\mu}_1'(\mathbf{u}), \ldots, \bar{\mu}_i'(\mathbf{u}), \ldots\right), \tag{4.53}$$

où les $\bar{\mu}_i'(\mathbf{u})$ sont des estimations des moments $\mu_i'(\mathbf{u})$ du signal $\mathbf{u}$.
Par ailleurs, le système $\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \dot{\mathbf{y}}, \ldots, \mathbf{y}^{(r)}\right)$ étant séparable, on a

$$\mathbf{u} = \psi\left(\mathbf{p}, \mathbf{y}, \ldots, \mathbf{y}^{(r)}\right) = \mathbf{F}(\mathbf{p})^{\mathrm{T}}.\boldsymbol{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right).$$

Par conséquent, l'expression (4.53) devient

$$\hat{\mathbf{g}}(\mathbf{p}) = \boldsymbol{\delta}\left(\bar{\mu}_1'\left(\mathbf{F}(\mathbf{p})^{\mathrm{T}}.\boldsymbol{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right), \ldots, \bar{\mu}_i'\left(\mathbf{F}(\mathbf{p})^{\mathrm{T}}.\boldsymbol{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right), \ldots\right).$$

Sachant que l'opérateur *bar* est linéaire, on peut écrire, pour tout $i \in \mathbb{N}$,

$$\bar{\mu}_i'\left(\mathbf{F}(\mathbf{p})^{\mathrm{T}}.\boldsymbol{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \mathbf{G}(\mathbf{p})^{\mathrm{T}}.\bar{\mu}_i'\left(\boldsymbol{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right) = \mathbf{G}(\mathbf{p})^{\mathrm{T}}\bar{\mu}_i(\mathbf{y}),$$

où $\mathbf{G}(\mathbf{p})$ est une matrice de dimension identique à $\mathbf{F}(\mathbf{p})$ et

$$\forall i \in \mathbb{N}, \bar{\mu}_i(\mathbf{y}) = \bar{\mu}_i'\left(\boldsymbol{\Phi}\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)\right).$$

Ainsi, on a bien

$$\hat{\mathbf{g}}(\mathbf{p}) = \boldsymbol{\beta}\left(\mathbf{p}, \bar{\mu}_1(\mathbf{y}), \bar{\mu}_2(\mathbf{y}), \ldots, \bar{\mu}_i(\mathbf{y}), \ldots\right).$$

□

Nous allons maintenant observer sur un exemple simple, comment une estimation empirique de $\mathbf{g}$ peut être obtenue.

Exemple 4.39. *Considérons le modèle décrit par la relation*

$$u = \psi(p, y, \dot{y}) = \dot{y} + p\sin(y),$$

et la fonction d'estimation

$$h(u) = E(\dot{u}\dot{u}).$$

Sachant que u est stationnaire,

$$\begin{aligned} g(p) &= h(\psi(p, y, \dot{y})) \\ &= h(\dot{y} + p\sin(y)) \\ &= E[(\ddot{y} + p\dot{y}\cos(y))(\ddot{y} + p\dot{y}\cos(y))] \\ &= E\left[\ddot{y}^2 + 2p\ddot{y}\dot{y}\cos(y) + p^2\dot{y}^2\cos(y)^2\right]. \end{aligned}$$

Contrairement à u, les signaux $\ddot{y}\dot{y}, \ddot{y}\sin(y), \ldots$ ne sont pas nécessairement stationnaires, et il est impossible de distribuer l'opérateur moyenne $E(.)$. En utilisant l'opérateur bar, nous obtenons une estimation $\hat{g}(p)$ de la fonction $g(p)$:

$$\hat{g}(p) = \overline{\ddot{y}^2 + 2p\ddot{y}\dot{y}\cos(y) + p^2\dot{y}^2\cos(y)^2}.$$

Et, par linéarité de l'opérateur bar, $g(p)$ peut s'approximer par

$$\hat{g}(p) = \overline{\ddot{y}^2} + 2p\overline{\ddot{y}\dot{y}\cos(y)} + p^2\overline{\dot{y}^2\cos(y)^2}.$$

Dans cet exemple, la fonction $g(p)$ est approximée par un polynôme du second degré en p $\hat{g}(p)$, dont les coefficients sont calculés à partir des sorties connues.

Remarque 4.40. *On peut vérifier, par un contre exemple, que la proposition 4.38 est fausse pour des systèmes inversibles quelconques (non découplés). Considérons le système*

$$u = \psi(p, y, \ddot{y}) = \sin(py),$$

et la fonction d'estimation

$$h : \left\{ \begin{array}{ccc} \mathcal{S} & \to & \mathcal{A}_{\mathbf{u}} \\ u & \mapsto & E(u) \end{array} \right. .$$

On a

$$\hat{g}(p) = \overline{\sin(py)}$$

qui n'est nullement une fonction d'un nombre fini de $\bar{\mu}_i(\mathbf{y})$ (mais plutôt des $\bar{\mu}_i(\mathbf{py})$) et par conséquent, ne vérifie pas l'égalité (4.51).

4.4.1.4 Synthèse

La proposition 4.38 est fondamentale et justifie à elle seule l'introduction des systèmes inversibles découplés pour l'estimation en aveugle de paramètres. En effet, nous avons montré que, pour les système inversibles découplés, l'équation

$$\mathbf{h}(\mathbf{u}) = \mathbf{0},$$

caractéristique de l'ensemble des paramètres solutions $\mathbb{P}$ (voir (4.46)) peut toujours être approximée par un système d'équations en $\mathbf{p}$ dont les coefficients font intervenir uniquement des estimations des moments de $\mathbf{y}$ et de ses dérivées, c'est-à-dire,

$$\hat{\mathbf{g}}(\mathbf{p}, \bar{\mu}_1(\mathbf{y}), \bar{\mu}_2(\mathbf{y}), \ldots, \bar{\mu}_i(\mathbf{y}), \ldots) = \mathbf{0}.$$

Ainsi, le problème d'estimation de paramètres en aveugle revient à caractériser l'ensemble

$$\mathbb{P} = \{\mathbf{p} \in \mathcal{P} | \hat{\mathbf{g}}(\mathbf{p}, \bar{\mu}_1(\mathbf{y}), \bar{\mu}_2(\mathbf{y}), \ldots, \bar{\mu}_i(\mathbf{y}), \ldots) = \mathbf{0}\}.$$

C'est un problème d'inversion ensembliste (voir paragraphe 2.2.2 page 46) qui peut facilement être résolu par des techniques d'analyse par intervalles, et notamment SIVIA dont c'est précisement le rôle.

4.4.2 Simulations pour un système identifiable en aveugle

Dans cette partie, nous considérons un exemple simple, où les entrées $\mathbf{u}$ et les signaux observés $\mathbf{y}$ sont reliés par un système paramétrique inversible découplé. Nous étudions ce système pour deux modèles d'entrées distincts. Premièrement, nous supposons que les entrées sont indépendantes (modèle d'entrées indépendantes), puis, nous supposons qu'une des entrées est gaussienne. Dans chaque cas, nous construisons une fonction d'estimation adaptée au modèle d'entrées considéré afin d'estimer les paramètres inconnus.

Le système considéré pour les simulations de cette section est le suivant. Nous considérons deux entrées u_1 et u_2 indépendantes $I_s(u_1, u_2)$ et gaussiennes $\mathcal{G}_s(u_1), \mathcal{G}_s(u_2)$ (colorées) obtenues par filtrage d'un bruit blanc gaussien (de $N = 9000$ échantillons). Ces deux signaux vérifient le système non linéaire décrit par les relations suivantes :

$$\begin{cases} u_1 = \frac{1}{p_1}\dot{y}_1 + \frac{p_1}{p_2}\cos(y_2) - y_1 \\ u_2 = \dot{y}_2 + p_1^2 p_2 \sin(y_1) \end{cases} \tag{4.54}$$

où les deux valeurs des paramètres sont $p_1 = -2$ et $p_2 = -0.9$.

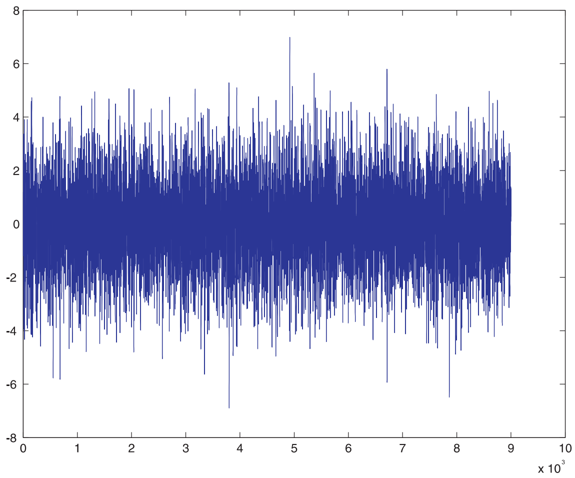


Figure 4.2 – Entrée $u_1(t)$.

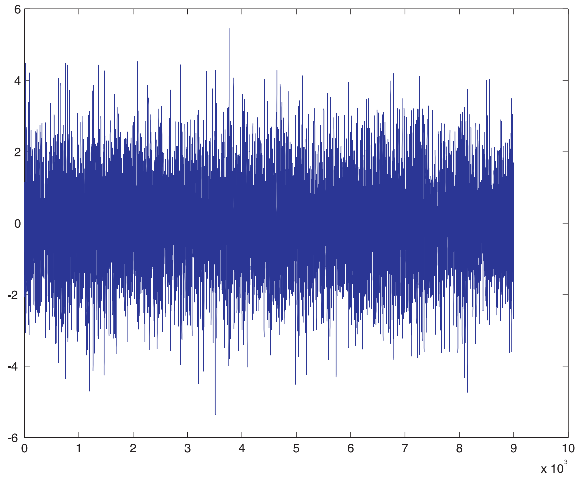


Figure 4.3 – Entrée $u_2(t)$.

Ce système étant instable, $\mathbf{y}$ est non stationnaire (comme le montrent les figures 4.4 et 4.5) bien que $\mathbf{u}$ le soit (voir figures 4.2 et 4.3).

4.4.2.1 Modèle d'entrées indépendantes

Dans ce paragraphe, nous considérons les sorties $\mathbf{y}$ du système (4.54) (et leurs dérivées premières) connues. Les paramètres $\mathbf{p}$ et les entrées $\mathbf{u}$ sont inconnus. Nous supposons uniquement que les entrées sont indépendantes, $I_s(u_1, u_2)$, c'est-à-dire que le modèle d'entrée considéré est $\mathbb{M} = I_s$ (la gaussianité des sources n'est pas exploitée ici). Avant de construire une fonction d'estimation pour ce modèle d'entrées indépendantes qui nous permettra d'estimer les paramètres du système (4.54), nous nous intéressons tout d'abord à l'identifiabilité en aveugle de ce système selon le modèle d'entrées indépendantes.

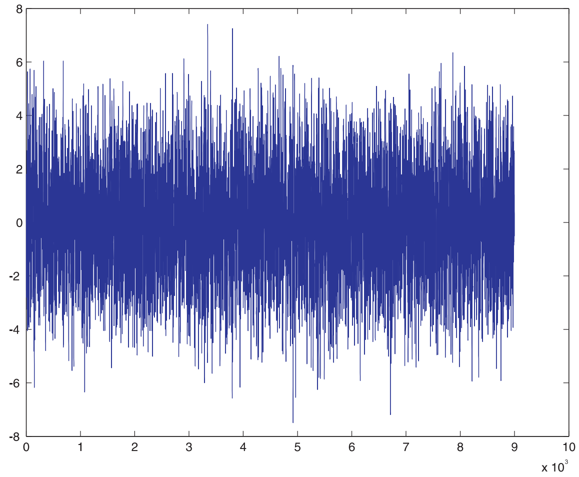


Figure 4.4 – Sortie $y_1(t)$.

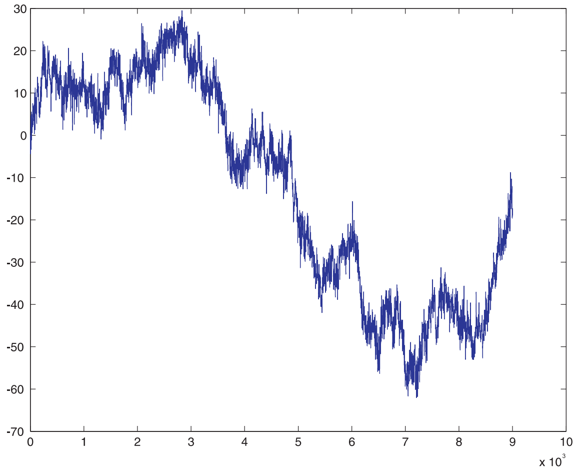


Figure 4.5 – Sortie $y_2(t)$.

4.4.2.1.1 Identifiabilité en aveugle Le système inversible (4.54) est un système découplé puisqu'il s'écrit

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{p_1} & \frac{p_1}{p_2} & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & p_1^2 p_2 \end{pmatrix} \begin{pmatrix} \dot{y}_1 \\ \cos(y_2) \\ y_1 \\ \dot{y}_2 \\ \sin(y_1) \end{pmatrix}.$$

Si la conjecture 4.26 du paragraphe 4.2.2.2 est vraie, le système (4.54) est identifiable en aveugle selon le modèle d'entrées $I_s(u_1, u_2)$ si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$, on a

$$\begin{pmatrix} \alpha_1 \left(\frac{1}{p_1^\star} \dot{y}_1 + \frac{p_1^\star}{p_2^\star} \cos(y_2) - y_1 \right) \\ \alpha_2 \left(\dot{y}_2 + (p_1^\star)^2 p_2^\star \sin(y_1) \right) \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\hat{p}_1} \dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2} \cos(y_2) - y_1 \\ \dot{y}_2 + \hat{p}_1^2 \hat{p}_2 \sin(y_1) \end{pmatrix} \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.55}$$

avec $\alpha_1, \alpha_2 \in \mathbb{R}$ et $\beta_1, \beta_2 \in \mathbb{R}$. En effet, comme le suggère la remarque 4.28, il est inutile de vérifier la condition (4.35) pour les permutations $\mathbf{P} \neq \mathbf{Id}_2$.
La condition (4.55) équivaut à

$$\left\{ \begin{array}{c} \alpha_1 \frac{1}{p_1^\star} \dot{y}_1 + \alpha_1 \frac{p_1^\star}{p_2^\star} \cos(y_2) - \alpha_1 y_1 + \beta_1 = \frac{1}{\hat{p}_1} \dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2} \cos(y_2) - y_1 \\ \alpha_2 \dot{y}_2 + \alpha_2 (p_1^\star)^2 p_2^\star \sin(y_1) + \beta_2 = \dot{y}_2 + \hat{p}_1^2 \hat{p}_2 \sin(y_1) \end{array} \right. \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star.$$

Or,

$$\left\{ \begin{array}{l} \alpha_1 \frac{1}{p_1^\star} \dot{y}_1 + \alpha_1 \frac{p_1^\star}{p_2^\star} \cos(y_2) - \alpha_1 y_1 + \beta_1 = \frac{1}{\hat{p}_1} \dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2} \cos(y_2) - y_1 \\ \alpha_2 \dot{y}_2 + \alpha_2 (p_1^\star)^2 p_2^\star \sin(y_1) + \beta_2 = \dot{y}_2 + \hat{p}_1^2 \hat{p}_2 \sin(y_1) \end{array} \right. \Longleftrightarrow \left\{ \begin{array}{c} \alpha_1 = 1 \\ \alpha_2 = 1 \\ p_1^\star = \hat{p}_1 \\ p_2^\star = \hat{p}_2 \\ \beta_1 = 0 \\ \beta_2 = 0 \end{array} \right. .$$

La conjecture 4.26 étant vraie, le système (4.54) est identifiable en aveugle pour le modèle d'entrées indépendantes.

4.4.2.1.2 Estimation en aveugle Maintenant nous cherchons à estimer en aveugle les paramètres pour le système (4.54). Rappelons que, dans ce paragraphe, nous supposons uniquement que les entrées u_1 et u_2 sont indépendantes. Aucune autre hypothèse n'est pris en compte : en particulier, la gaussianité des signaux u_1 et u_2 n'est pas du tout exploitée. Comme cela est suggéré au paragraphe 4.4.1.1, nous choisissons la fonction d'estimation

$$\mathbf{h} = \begin{pmatrix} \Gamma_{u_1,u_2} \\ \Gamma_{\dot{u}_1,u_2} \end{pmatrix} \tag{4.56}$$

$$= \begin{pmatrix} E(u_1u_2) - E(u_1)E(u_2) \\ E(\dot{u}_1u_2) \end{pmatrix}, \tag{4.57}$$

qui fournit deux équations (autant que de paramètres). Cette fonction d'estimation a été construite de manière à s'annuler lorsque les signaux u_1 et u_2, respectivement les signaux $\dot{u}_1$ et u_2, sont décorrélés. Il est clair que nous pouvons ajouter des composantes supplémentaires pour la fonction $\mathbf{h}$ afin d'améliorer l'estimation et ainsi exploiter plus en profondeur l'hypothèse d'indépendance.

Le système inversible (4.54) considéré étant découplé (et conformément aux résultats du paragraphe 4.4.1.3), la condition $\mathbf{h}(\mathbf{u}) = \mathbf{0}$ peut être approximée, à l'aide de l'opérateur *bar*, par un système d'équations en $\mathbf{p}$ faisant intervenir uniquement les estimations des moments des sorties. Nous obtenons

$$\begin{cases} \overline{u_1u_2} - \overline{u_1}\,\overline{u_2} = 0 \\ \overline{\dot{u}_1u_2} = 0 \end{cases} . \tag{4.58}$$

D'après les expressions (4.54) décrivant le système étudié et sachant que

$$\dot{u}_1 = \frac{1}{p_1}\ddot{y}_1 - \frac{p_1}{p_2}\dot{y}_2\sin y_2 - \dot{y}_1, \tag{4.59}$$

on a

$$\begin{aligned}
\overline{u_1} &= \overline{\tfrac{1}{p_1}\dot{y}_1 + \tfrac{p_1}{p_2}\cos y_2 - y_1} = \tfrac{1}{p_1}\overline{\dot{y}_1} + \tfrac{p_1}{p_2}\overline{\cos y_2} - \overline{y_1}, \\
\overline{u_2} &= \overline{\dot{y}_2 + p_1^2p_2\sin y_1} = \overline{\dot{y}_2} + p_1^2p_2\overline{\sin y_1}, \\
\overline{u_1u_2} &= \tfrac{1}{p_1}\overline{\dot{y}_1\dot{y}_2} + p_1p_2\overline{\dot{y}_1\sin y_1} + \tfrac{p_1}{p_2}\overline{\dot{y}_2\cos y_2} \\
&\quad + p_1^3\overline{\cos y_2\sin y_1} - \overline{y_1\dot{y}_2} - p_1^2p_2\overline{y_1\sin y_1}, \\
\overline{\dot{u}_1u_2} &= \tfrac{1}{p_1}\overline{\ddot{y}_1\dot{y}_2} + p_1p_2\overline{\ddot{y}_1\sin y_1} - \tfrac{p_1}{p_2}\overline{\dot{y}_2^2\sin y_2} \\
&\quad - p_1^3\overline{\dot{y}_2\sin y_2\sin y_1} - \overline{\dot{y}_1\dot{y}_2} - p_1^2p_2\overline{\dot{y}_1\sin y_1}.
\end{aligned} \tag{4.60}$$

En posant $s_i = \sin y_i$, et $c_i = \cos y_i$, le système (4.58) devient

$$\Sigma_I(N) : \begin{cases}
\tfrac{1}{p_1}\overline{\dot{y}_1\dot{y}_2} + p_1p_2\overline{\dot{y}_1s_1} + \tfrac{p_1}{p_2}\overline{c_2\dot{y}_2} + p_1^3\overline{c_2s_1} - \overline{y_1\dot{y}_2} \\
-p_1^2p_2\overline{y_1s_1} - (\tfrac{1}{p_1}\overline{\dot{y}_1} + \tfrac{p_1}{p_2}\overline{c_2} - \overline{y_1})(\overline{\dot{y}_2} + p_1^2p_2\overline{s_1}) = 0 \\
\\
\tfrac{1}{p_1}\overline{\ddot{y}_1\dot{y}_2} + p_1p_2\overline{\ddot{y}_1s_1} - \tfrac{p_1}{p_2}\overline{\dot{y}_2^2s_2} \\
-p_1^3\overline{\dot{y}_2s_2s_1} - \overline{\dot{y}_1\dot{y}_2} - p_1^2p_2\overline{\dot{y}_1s_1} = 0.
\end{cases} \tag{4.61}$$

Ce système d'équations est constitué de deux équations non linéaires à deux inconnues p_1 et p_2. Les coefficients $\overline{\dot{y}_1\dot{y}_2}, \overline{\dot{y}_1s_1}, \overline{c_2\dot{y}_2}, \ldots$ de $\Sigma_I(N)$ dépendent du nombre N d'échantillons.

Nous avons résolu le système $\Sigma_I(N)$ pour différentes valeurs de N en utilisant l'algorithme ensembliste SIVIA (voir paragraphe 2.2.3 page 47). Voici, sur la figure 4.6, l'ensemble des paramètres solutions recherchés dans l'intervalle $[-5, 5]$, obtenus pour $N = 6000$ échantillons. On trouve

$$p_1 \in [-2.19, -2.0] \text{ et } p_2 \in [-0.91, -0.89]. \tag{4.62}$$

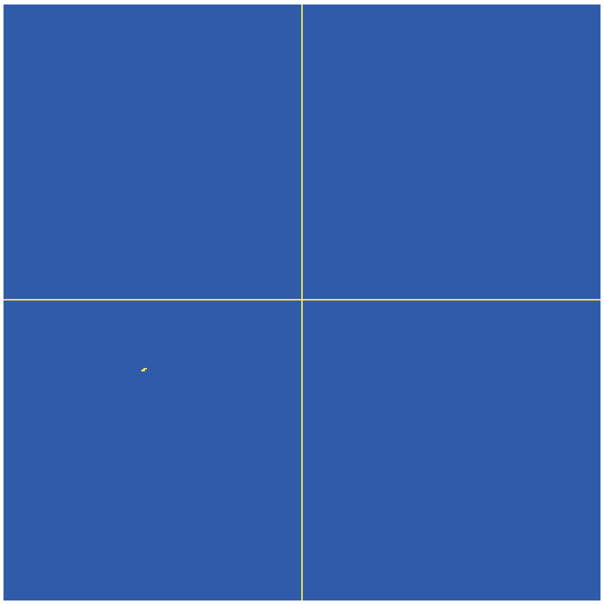


Figure 4.6 – Solutions du système $\Sigma_I(N)$ par SIVIA pour $N = 6000$ échantillons et une précision $\varepsilon = 0.001$.

Pour chaque N, nous trouvons un unique vecteur solution $\hat{\mathbf{p}}(N)$. Le tableau (4.63) représente l'erreur quadratique d'estimation pour différentes valeurs de N. Nous pouvons vérifier que $\hat{\mathbf{p}}(N)$ tend vers $\mathbf{p}^\star$, lorsque N tend vers l'infini.

N	100	600	1000	6000	10000	16000
$(\hat{p}_1(N) - p_1^*)^2$	$4,8.10^{-2}$	$1,1.10^{-2}$	$1,1.10^{-2}$	10^{-4}	$2,5.10^{-5}$	10^{-5}
$(\hat{p}_2(N) - p_2^*)^2$	$7,8.10^{-2}$	$3,5.10^{-2}$	$1,4.10^{-2}$	10^{-3}	9.10^{-5}	3.10^{-6}

(4.63)

4.4.2.2 Une entrée gaussienne

Dans ce paragraphe, l'indépendance des signaux u_1 et u_2 n'est pas du tout exploitée. Nous supposons uniquement que l'entrée u_1 est gaussienne (aucune hypothèse n'est posée sur u_2), c'est-à-dire que le modèle d'entrées considéré est $\mathbb{M} = \mathcal{G}_s(u_1)$.

Afin de suivre le même schéma de résolution que pour le modèle d'entrées indépendantes étudié au paragraphe précédent, intéressons-nous pour commencer à l'identifiabilité du système (4.54) selon le modèle d'entrée $u_1 \in \mathcal{G}_s$.

4.4.2.2.1 Identifiabilité en aveugle Sous l'hypothèse que la conjecture 4.26 du paragraphe 4.2.2.2 est vraie, le système (4.54) est identifiable en aveugle selon le modèle d'entrée $u_1 \in \mathcal{G}_s$ si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$, on a

$$\alpha\left(\frac{1}{p_1^\star}\dot{y}_1+\frac{p_1^\star}{p_2^\star}\cos(y_2)-y_1\right)+\beta=\frac{1}{\hat{p}_1}\dot{y}_1+\frac{\hat{p}_1}{\hat{p}_2}\cos(y_2)-y_1 \implies \hat{\mathbf{p}}=\mathbf{p}^\star. \tag{4.64}$$

La condition (4.64) équivaut à

$$\alpha\frac{1}{p_1^\star}\dot{y}_1+\alpha\frac{p_1^\star}{p_2^\star}\cos(y_2)-\alpha y_1+\beta=\frac{1}{\hat{p}_1}\dot{y}_1+\frac{\hat{p}_1}{\hat{p}_2}\cos(y_2)-y_1 \implies \hat{\mathbf{p}}=\mathbf{p}^\star.$$

Or,

$$\alpha\frac{1}{p_1^\star}\dot{y}_1+\alpha\frac{p_1^\star}{p_2^\star}\cos(y_2)-\alpha y_1+\beta=\frac{1}{\hat{p}_1}\dot{y}_1+\frac{\hat{p}_1}{\hat{p}_2}\cos(y_2)-y_1 \iff \left\{\begin{array}{c}\alpha=1\\ p_1^\star=\hat{p}_1\\ p_2^\star=\hat{p}_2\\ \beta=0\end{array}\right.,$$

avec $\alpha \in \mathbb{R}$ et $\beta \in \mathbb{R}$.
Si la conjecture 4.26 est vraie, le système (4.54) est identifiable en aveugle pour le modèle d'entrée $u_1 \in \mathcal{G}_s$.

4.4.2.2.2 Estimation en aveugle Maintenant, nous cherchons à estimer en aveugle les paramètres pour le système (4.54). Rappelons que seule l'entrée u_1 est supposée gaussienne. Comme il est suggéré dans le paragraphe 4.4.1.2, nous choisissons la fonction d'estimation

$$\mathbf{h}(\mathbf{u})=\begin{pmatrix} kurt\,(u_1)\\ kurt\,(\dot{u}_1)\end{pmatrix} \tag{4.65}$$

$$=\begin{pmatrix} E((u_1-E(u_1))^4)-3\left(E((u_1-E(u_1))^2\right)^2\\ E(\dot{u}_1^4)-3\left(E(\dot{u}_1^2)\right)^2\end{pmatrix}. \tag{4.66}$$

Le système inversible (4.54) considéré étant découplé, la condition $\mathbf{h}(\mathbf{u})=\mathbf{0}$ peut être approximée, à l'aide de l'opérateur *bar*, par le système d'équations en $\mathbf{p}$ suivant

$$\left\{\begin{array}{l}\overline{(u_1-\overline{u_1})^4}-3\overline{(u_1-\overline{u_1})^2}^2=0\\ \overline{\dot{u}_1^4}-3\overline{\dot{u}_1^2}^2=0\end{array}\right. . \tag{4.67}$$

Ainsi, une approximation de l'ensemble $\mathbb{P}$ à caractériser est donnée par

$$\Sigma_G(N):\left\{\begin{array}{l}\overline{u_1^4}-4\overline{u_1}\,\overline{u_1^3}+12\overline{u_1}^2\overline{u_1^2}-6\overline{u_1}^4-3\overline{u_1^2}^2=0\\ \overline{\dot{u}_1^4}-3\overline{\dot{u}_1^2}^2=0.\end{array}\right. \tag{4.68}$$

D'après les expressions (4.54) décrivant le système étudié et l'équation (4.59), on peut écrire le système $\Sigma_G(N)$ comme un système de deux équations non linéaires à deux inconnues p_1 et p_2. Par souci de simplicité, nous ne donnons pas la forme explicite de ce système.

De même que précédemment, les coefficients $\overline{\dot{y}_1\dot{y}_2}, \overline{\dot{y}_1 s_1}, \overline{c_2\dot{y}_2}, \ldots$ du système $\Sigma_G(N)$ dépendent du nombre N d'échantillons et le vecteur de paramètres estimé $\hat{\mathbf{p}}(N)$ converge vers $\mathbf{p}^\star$.

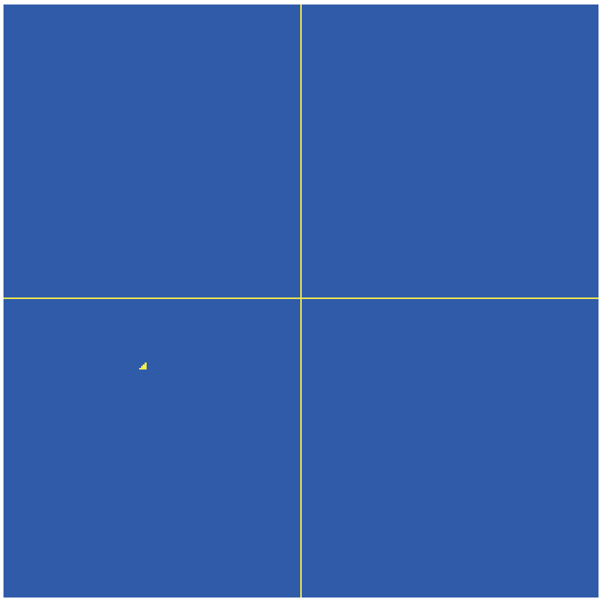


Figure 4.7 – Solutions du système $\Sigma_G(N)$ par SIVIA pour $N = 6000$ échantillons et une précision $\varepsilon = 0.001$.

Le tableau 4.69 et la figure 4.7 (SIVIA) montrent qu'une bonne estimation des paramètres est obtenue à partir de $N = 6000$.

N	100	600	1000	6000	10000	16000
$(\hat{p}_1(N) - p_1^*)^2$	$9,7.10^{-1}$	$8,1.10^{-2}$	$1,4.10^{-2}$	$2,4.10^{-3}$	5.10^{-4}	10^{-5}
$(\hat{p}_2(N) - p_2^*)^2$	$8,9.10^{-1}$	$7,1.10^{-2}$	$1,3.10^{-2}$	$1,3.10^{-3}$	3.10^{-4}	10^{-5}

(4.69)

Nous avons uniquement présenté des exemples pour lesquels les paramètres étaient identifiables en aveugle selon le modèle d'entrées considéré. Cependant, il est clair que cette propriété n'est pas vérifiée par tous les systèmes inversibles. Nous terminons par un exemple pour lequel les paramètres sont non identifiables en aveugle afin de mettre en évidence l'intérêt des techniques ensemblistes.

4.4.3 Estimation des paramètres d'un système non identifiable en aveugle

Nous considérons deux entrées u_1 et u_2 indépendantes, $I_s(u_1, u_2)$, et gaussiennes, $\mathcal{G}_s(u_1)$ et $\mathcal{G}_s(u_2)$, (colorées) obtenues par filtrage d'un bruit blanc gaussien (de $N = 9000$ échantillons). Ces deux signaux vérifient le système non linéaire décrit par les relations suivantes :

$$\begin{cases} u_1 = \frac{1}{p_1}\dot{y}_1 + \frac{p_1}{p_2}\cos y_2 \\ u_2 = \dot{y}_2 + p_2^2 \sin y_1, \end{cases} \tag{4.70}$$

où les deux valeurs des paramètres sont $p_1 = -\frac{10}{9}$ et $p_2 = -\frac{1}{2}$.

Nous allons étudier l'identifiabilité en aveugle des paramètres du système (4.70) d'abord pour le modèle d'entrées indépendantes puis pour le modèle "u_1 gaussien".

4.4.3.1 Modèle d'entrées indépendantes

Dans ce paragraphe, nous supposons uniquement que les entrées u_1 et u_2 sont indépendantes (la gaussianité des signaux u_1 et u_2 n'est pas du tout exploitée), c'est-à-dire que le modèle d'entrées considéré est $\mathbb{M} = I_s(u_1, u_2)$. Nous allons donc chercher l'ensemble des paramètres solutions suivant :

$$\mathbb{P} = \left\{(p_1, p_2) \in \mathbb{R}^2 | I_s(u_1, u_2)\right\}.$$

4.4.3.1.1 Identifiabilité en aveugle Si la conjecture 4.26 du paragraphe 4.2.2.2 est vraie, le système (4.70) est identifiable en aveugle selon le modèle d'entrée $I_s(u_1, u_2)$ si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$, on a

$$\left(\begin{array}{cc} \alpha_1 & 0 \\ 0 & \alpha_2 \end{array}\right)\left(\begin{array}{c} \frac{1}{p_1^\star}\dot{y}_1 + \frac{p_1^\star}{p_2^\star}\cos y_2 \\ \dot{y}_2 + p_2^{\star 2}\sin y_1 \end{array}\right) + \left(\begin{array}{c} \beta_1 \\ \beta_2 \end{array}\right) = \left(\begin{array}{c} \frac{1}{\hat{p}_1}\dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2}\cos y_2 \\ \dot{y}_2 + \hat{p}_2^2\sin y_1 \end{array}\right) \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star,$$

avec $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ et $(\beta_1, \beta_2) \in \mathbb{R}^2$, ce qui équivaut à

$$\left\{\begin{array}{l} \alpha_1\frac{1}{p_1^\star}\dot{y}_1 + \alpha_1\frac{p_1^\star}{p_2^\star}\cos y_2 + \beta_1 = \frac{1}{\hat{p}_1}\dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2}\cos y_2 \\ \alpha_2\dot{y}_2 + \alpha_2 p_2^{\star 2}\sin y_1 + \beta_2 = \dot{y}_2 + \hat{p}_2^2\sin y_1 \end{array}\right. \Longrightarrow \hat{\mathbf{p}} = \mathbf{p}^\star.$$

Or, en résolvant les deux équations d'inconnues $\hat{p}_1$ et $\hat{p}_2$, on a

$$\left\{\begin{array}{l} \alpha_1\frac{1}{p_1^\star}\dot{y}_1 + \alpha_1\frac{p_1^\star}{p_2^\star}\cos y_2 + \beta_1 = \frac{1}{\hat{p}_1}\dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2}\cos y_2 \\ \alpha_2\dot{y}_2 + \alpha_2 p_2^{\star 2}\sin y_1 + \beta_2 = \dot{y}_2 + \hat{p}_2^2\sin y_1 \end{array}\right. \Longleftrightarrow \left\{\begin{array}{l} \alpha_1^2 = 1 \\ \alpha_2 = 1 \\ \beta_1 = \beta_2 = 0 \\ \hat{p}_1 = \pm p_1^\star \\ \hat{p}_2 = p_2^\star \end{array}\right., \tag{4.71}$$

donc, sous l'hypothèse que la conjecture 4.26 est vraie, le système (4.70) n'est pas identifiable en aveugle pour le modèle d'entrée $I_s(u_1, u_2)$. En effet, si le couple $(p_1^\star, p_2^\star)$ est solution, alors le couple $(-p_1^\star, p_2^\star)$ est aussi solution.

4.4.3.1.2 Estimation en aveugle Maintenant, nous cherchons à estimer en aveugle les paramètres du le système (4.70) bien qu'il soit non identifiable en aveugle pour le modèle $\mathbb{M} = I_s$. Rappelons que les entrées u_1 et u_2 sont supposées indépendantes. Nous choisissons la fonction d'estimation

$$\mathbf{h}(\mathbf{u}) = \left(\begin{array}{c} \Gamma_{u_1,u_2} \\ \Gamma_{\dot{u}_1,u_2} \end{array}\right) \tag{4.72}$$

$$= \left(\begin{array}{l} E(u_1u_2) - E(u_1)E(u_2) \\ E(\dot{u}_1u_2) \end{array}\right) \tag{4.73}$$

qui fournit deux équations (autant que de paramètres).

Le système inversible (4.70) considéré étant découplé, la condition $\mathbf{h}(\mathbf{u}) = \mathbf{0}$ peut être approximée, à l'aide de l'opérateur *bar*, par le système d'équations en $\mathbf{p}$ suivant

$$\left\{\begin{array}{l} \overline{u_1u_2} - \overline{u_1}\,\overline{u_2} = 0 \\ \overline{\dot{u}_1u_2} = 0, \end{array}\right. \tag{4.74}$$

c'est-à-dire, en considérant l'expression des entrées dans (4.70),

$$\left\{\begin{array}{l} \frac{1}{p_1}\overline{\dot{y}_1\dot{y}_2} + \frac{p_1}{p_2}\overline{\dot{y}_2\cos y_2} + \frac{p_2^2}{p_1}\overline{\dot{y}_1\sin y_1} + p_1p_2\overline{\sin y_1\cos y_2} - \left(\frac{1}{p_1}\overline{\dot{y}_1} + \frac{p_1}{p_2}\overline{\cos y_2}\right)\left(\overline{\dot{y}_2} + p_2^2\overline{\sin y_1}\right) = 0 \\ \frac{1}{p_1}\overline{\ddot{y}_1\dot{y}_2} - \frac{p_1}{p_2}\overline{\dot{y}_2^2\sin y_2} + \frac{p_2^2}{p_1}\overline{\ddot{y}_1\sin y_1} - p_1p_2\overline{\dot{y}_2\sin y_1\sin y_2} = 0. \end{array}\right. \tag{4.75}$$

L'ensemble $\mathbb{P}$ des paramètres solutions obtenu par résolution est tracé sur la figure 4.8. Nous retrouvons bien deux solutions conformément à l'étude de l'identifiabilité précédente.

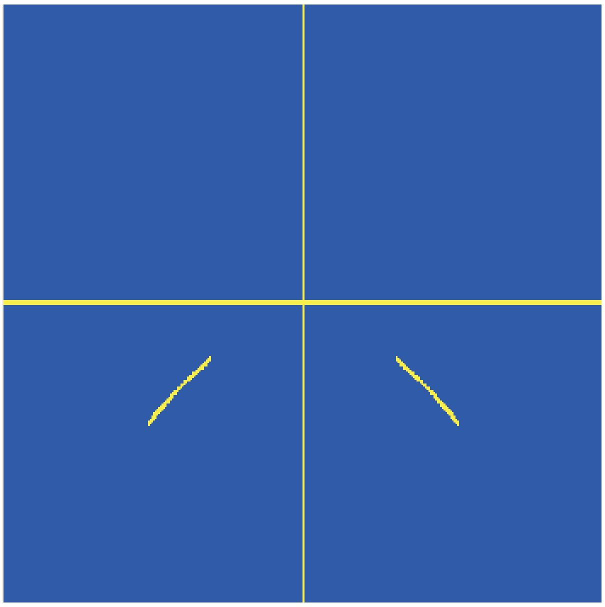


Figure 4.8 – Ensemble des paramètres solutions $(\hat{p}_1, \hat{p}_2) \in [-5, 5] \times [-5, 5]$ obtenus par SIVIA pour $N = 9000$ échantillons.

4.4.3.2 Une entrée gaussienne

Dans ce paragraphe, nous supposons uniquement que l'entrée u_1 est gaussienne (l'indépendance des signaux u_1 et u_2 n'est pas du tout exploitée, aucune hypothèse n'est posée sur u_2) c'est-à-dire que le modèle d'entrées considéré est $\mathbb{M} = \mathcal{G}_s(u_1)$. Nous allons donc chercher l'ensemble des paramètres solutions suivant :

$$\mathbb{P} = \left\{(p_1, p_2) \in \mathbb{R}^2 | u_1 \in \mathcal{G}_s\right\}.$$

4.4.3.2.1 Identifiabilité en aveugle Sous l'hypothèse que la conjecture 4.22 du paragraphe 4.2.2.2 est vraie, le système (4.70) est identifiable en aveugle selon le modèle d'entrée $u_1 \in \mathcal{G}_s$ si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\alpha\left(\frac{1}{p_1^\star}\dot{y}_1 + \frac{p_1^\star}{p_2^\star}\cos y_2\right) + \beta = \frac{1}{\hat{p}_1}\dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2}\cos y_2 \implies \hat{\mathbf{p}} = \mathbf{p}^\star,$$

avec $\alpha \in \mathbb{R}$ et $\beta \in \mathbb{R}$, ce qui équivaut à

$$\alpha\frac{1}{p_1^\star}\dot{y}_1 + \alpha\frac{p_1^\star}{p_2^\star}\cos y_2 + \beta = \frac{1}{\hat{p}_1}\dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2}\cos y_2 \implies \hat{\mathbf{p}} = \mathbf{p}^\star.$$

Or, en résolvant l'équation d'inconnues $\hat{p}_1$ et $\hat{p}_2$, on a

$$\alpha \frac{1}{p_1^\star}\dot{y}_1 + \alpha \frac{p_1^\star}{p_2^\star}\cos y_2 + \beta = \frac{1}{\hat{p}_1}\dot{y}_1 + \frac{\hat{p}_1}{\hat{p}_2}\cos y_2 \Longleftrightarrow \begin{cases} \alpha = \frac{p_1^\star}{\hat{p}_1} \\ \beta = 0 \\ \hat{p}_1 = \frac{p_1^\star}{\alpha} \\ \hat{p}_2 = \frac{p_2^\star}{\alpha^2} \end{cases}, \tag{4.76}$$

donc (si la conjecture 4.22 est vraie) le système (4.70) n'est pas identifiable en aveugle pour le modèle d'entrée $u_1 \in \mathcal{G}_s$. En effet, si le couple $(p_1^\star, p_2^\star)$ est solution, alors les couples $\left(\frac{p_1^\star}{\alpha}, \frac{p_2^\star}{\alpha^2}\right)$, pour tout $\alpha \in \mathbb{R}$, sont également solutions.

La figure 4.9 représente l'ensemble des paramètres solutions $(\hat{p}_1, \hat{p}_2)$ obtenus formellement pour $p_1^\star = -\frac{10}{9}$ et $p_2^\star = -\frac{1}{2}$. Il y en a une infinité.

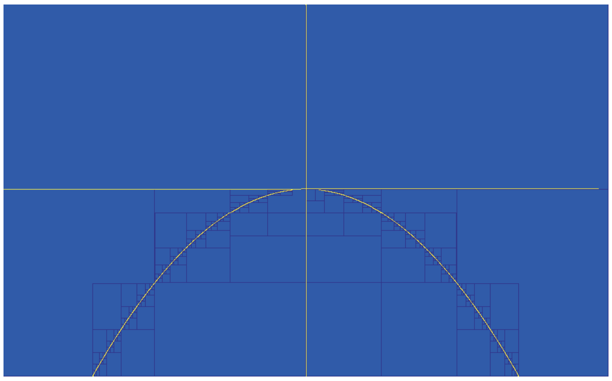


Figure 4.9 – Ensemble des paramètres solutions obtenus formellement pour $p_1^\star = -\frac{10}{9}$ et $p_2^\star = -\frac{1}{2}$ dans le pavé de recherche $[-5, 5] \times [-5, 5]$.

4.4.3.2.2 Estimation en aveugle Maintenant, nous cherchons à estimer en aveugle les paramètres pour le système (4.70). Rappelons que seule l'entrée u_1 est supposée gaussienne. Nous choisissons donc la fonction d'estimation

$$\mathbf{h}(\mathbf{u}) = \begin{pmatrix} kurt\,(u_1) \\ kurt\,(\dot{u}_1) \end{pmatrix}. \tag{4.77}$$

Le système inversible (4.70) considéré étant découplé, la condition $\mathbf{h}\,(\mathbf{u}) = \mathbf{0}$ peut être approximée, à l'aide de l'opérateur *bar*, par le système d'équations en $\mathbf{p}$ suivant

$$\begin{cases} \overline{(u_1 - \overline{u_1})^4} - 3\overline{(u_1 - \overline{u_1})^2} = 0 \\ \overline{\dot{u}_1^4} - 3\overline{\dot{u}_1^2} = 0, \end{cases} \tag{4.78}$$

où u_1 est défini par le système (4.70). L'ensemble $\mathbb{P}$ des paramètres solutions, recherchés dans l'intervalle $[-5, 5]$, obtenu par SIVIA est tracé sur la figure 4.10. Nous retrouvons bien une infinité de paramètres solutions, conformément à l'étude d'identifiabilité en aveugle précédente.

Figure 4.10 – Ensemble des paramètres solutions obtenu par SIVIA dans le pavé de recherche $[-5,5] \times [-5,5]$ pour $N = 9000$ échantillons et une précision $\varepsilon = 10^{-2}$.

4.4.3.3 Bilan

L'intérêt de l'utilisation des intervalles pour ces deux derniers exemples est que la non-identifiabilité du système (4.70) pour le modèle d'entrée considéré ($I_s(u_1, u_2)$ ou $\mathcal{G}_s(u_1)$) est automatiquement décelée. Notons d'ailleurs que, dans les deux exemples présentés, l'ajout d'équations supplémentaires concernant l'indépendance, respectivement la gaussianité, des dérivées des entrées ne permet pas de lever l'ambiguité sur les paramètres. Le système (4.70) est intrinsèquement non identifiable pour les modèles d'entrées considérés.

4.5 Problème des deux bacs d'eau

Dans cette dernière partie, nous illustrons la démarche d'identification de paramètres en aveugle pour un système plus concret. Le problème qui nous intéresse est très classique en automatique, il s'agit d'estimer les paramètres incertains d'un système composé de deux bacs d'eau. Insistons sur l'originalité de la démarche dans la mesure où les paramètres sont estimés sans la connaissance des entrées.

4.5.1 Préliminaires

On considère deux bacs disposés comme sur la figure 4.11.

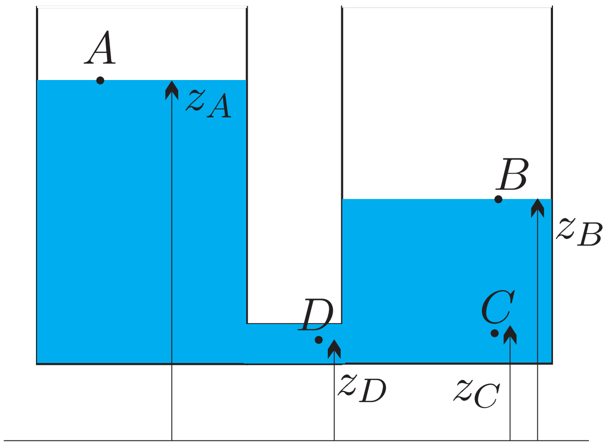


Figure 4.11 – Système hydraulique formé de deux bacs remplis d'eau reliés par un canal.

Dans le bac de gauche, l'eau s'écoule sans frottement en direction du bac de droite. De plus, l'eau s'écoule de façon fluide dans le bac de gauche, contrairement au bac de droite, où des turbulences existent. Ces turbulences absorbent l'énergie cinétique de l'eau et la transforment en chaleur. Sans ces turbulences, nous aurions un éternel mouvement de va-et-vient de l'eau entre les deux bacs. Afin d'appliquer la relation de Bernoulli dans le bac de gauche, nous considérons un tube de courant, c'est-à-dire un tube virtuel (voir figure 4.11) à l'intérieur duquel l'eau possède un mouvement fluide et ne traverse pas les parois. La relation de Bernoulli nous permet alors d'affirmer qu'en tout point du tube, on a

$$P + \rho \frac{v^2}{2} + \rho g z = \text{constante},$$

où P est la pression au point considéré, z son altitude, v la vitesse de l'eau en ce point, ρ la masse volumique de l'eau et g la constante de gravitation. D'après cette loi, nous avons

$$P_D + \rho \frac{v_D^2}{2} + \rho g z_D = P_A + \rho \frac{v_A^2}{2} + \rho g z_A,$$

c'est-à-dire

$$P_D = P_A + \rho g \left(z_A - z_D\right) - \rho \frac{v_D^2}{2}.$$

De plus, nous pouvons supposer que C est loin de la zone de turbulence et que l'eau ne bouge pas. Ainsi, d'après la relation de Bernoulli, nous avons

$$P_C + \rho g z_C = P_B + \rho g z_B,$$

c'est-à-dire

$$P_C = P_B + \rho g \left(z_B - z_C\right).$$

D'autre part, dans la zone de turbulence, l'eau se trouve ralentie, mais nous pouvons supposer que la pression ne varie pas, c'est-à-dire $P_C = P_D$. Ainsi,

$$P_B + \rho g \left(z_B - z_C\right) = P_A + \rho g \left(z_A - z_D\right) - \rho \frac{v_D^2}{2},$$

et, comme $P_A = P_B = P_{atm}$ et $z_C = z_D$, nous avons

$$\rho g\left(z_A - z_B\right) = \rho \frac{v_D^2}{2},$$

ou encore

$$v_D = \sqrt{2g\left(z_A - z_B\right)}.$$

Dans le cas où le niveau du bac droit est supérieur à celui du bac gauche, une étude similaire nous donne

$$v_D = -\sqrt{2g\left(z_A - z_B\right)}.$$

Le signe moins de l'expression indique que le courant circule maintenant du bac droit vers le bac gauche. Donc, la relation générale pour la vitesse de l'eau dans le tunnel est

$$v_D = sign\left(z_A - z_B\right)\sqrt{2g\left|z_A - z_B\right|}.$$

Si a est la section du tunnel, le débit d'eau du bac droit vers le bac gauche est

$$Q_D = a.sign\left(z_A - z_B\right)\sqrt{2g\left|z_A - z_B\right|}. \tag{4.79}$$

4.5.2 Modélisation

Considérons maintenant le système comprenant deux bacs d'eau représenté sur la figure 4.12.

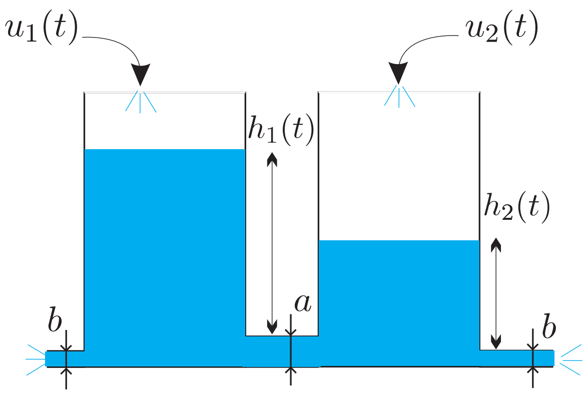


Figure 4.12 – Système constitué de deux bacs d'eau reliés par un canal de section a.

L'eau des bacs 1 et 2 communique par un canal de section $a = 0.3$ et peut aussi se déverser vers l'extérieur via deux canaux de section $b = 0.5$ situés aux extrémités de chaque bac. Les entrées du système sont les débits d'eau u_1 et u_2 dans les bacs 1 et 2. D'après la relation (4.79), les débits associés sont donnés par

$$Q_{1ext} = b.\sqrt{2gh_1},$$
$$Q_{2ext} = b.\sqrt{2gh_2}.$$

De même, le débit entre le bac 1 et le bac 2 est donné par

$$Q_{12} = a.sign\left(h_1 - h_2\right)\sqrt{2g\left|h_1 - h_2\right|}.$$

Nous pouvons considérer les hauteurs d'eau dans les bacs comme les variables d'état de ce système. Pour simplifier, nous supposerons que les surfaces des bacs sont toutes égales à $1m^2$. Ainsi, le volume d'eau dans chaque bac sera confondu avec la hauteur. Les équations d'état sont obtenues en écrivant que la variation du volume d'eau dans un bac est égale à la somme des débits entrants moins la somme des débits sortants :

$$\frac{d}{dt}h_1 = -Q_{1ext} - Q_{12} + u_1,$$
$$\frac{d}{dt}h_2 = Q_{12} - Q_{2ext} + u_2.$$

Le système des bacs d'eau est donc régi par les équation d'états suivantes :

$$\begin{pmatrix} \dot{h}_1(t) \\ \dot{h}_2(t) \end{pmatrix} = \begin{pmatrix} -b\sqrt{2gh_1(t)} - a.sign\left(h_1(t) - h_2(t)\right)\sqrt{2g\left|h_1(t) - h_2(t)\right|} \\ -b\sqrt{2gh_2(t)} + a.sign\left(h_1(t) - h_2(t)\right)\sqrt{2g\left|h_1(t) - h_2(t)\right|} \end{pmatrix} + \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}.$$

Par un calcul direct, on montre que ce système est inversible

$$\mathbf{u}(t) = \psi\left(\mathbf{p}, \mathbf{h}(t), \dot{\mathbf{h}}(t)\right).$$

où $\mathbf{u}(t) = (u_1(t), u_2(t))^{\mathrm{T}}$, $\mathbf{h}(t) = (h_1(t), h_2(t))^{\mathrm{T}}$, $\mathbf{p} = (a, b)^T$ et la fonction ψ est définie par

$$\psi\left(\mathbf{p}, \mathbf{h}(t), \dot{\mathbf{h}}(t)\right) = \begin{pmatrix} \dot{h}_1(t) + b\sqrt{2gh_1(t)} + a.sign\left(h_1(t) - h_2(t)\right)\sqrt{2g\left|h_1(t) - h_2(t)\right|} \\ \dot{h}_2(t) + b\sqrt{2gh_2(t)} - a.sign\left(h_1(t) - h_2(t)\right)\sqrt{2g\left|h_1(t) - h_2(t)\right|} \end{pmatrix}. \tag{4.80}$$

4.5.3 Estimation de paramètres en aveugle

Maintentant, considérons le système des bac d'eau de la figure 4.12 décrit par le système inversible (4.80). Nous supposons que

- La section de fuite b de chaque bac est inconnue
- La section a du canal reliant les deux bacs est inconnue.
- Le vecteur d'entrée inconnue $\mathbf{u}(t)$ est à composantes indépendantes.

Seul le vecteur de sortie $\mathbf{h}(t)$ (correspondant aux hauteurs d'eau dans chaque bac) est connu. Le problème que nous proposons de résoudre est l'estimation des paramètres inconnus $\mathbf{p} = (a, b)$ en aveugle simplement à l'aide de l'hypothèse d'indépendance des entrées. Il s'agit d'un problème d'estimation de paramètres en aveugle où l'ensemble à caractériser est le suivant

$$\mathbb{P} = \left\{\mathbf{p} \in \mathbb{R}^2 | \psi\left(\mathbf{p}, \mathbf{h}, \dot{\mathbf{h}}\right) \in I_s^2\right\}.$$

4.5.3.0.1 Identifiabilité en aveugle Nous testons dans un premier temps l'identifiabilité en aveugle des paramètres $\mathbf{p}$ du système (4.80) selon le modèle d'entrées indépendantes. Si la conjecture 4.26 du paragraphe 4.2.2.2 est vraie, le système (4.80) est identifiable en aveugle selon le modèle d'entrées indépendantes si et seulement si, pour presque tout $\mathbf{p}^\star \in \mathcal{P}$,

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \psi\left(\mathbf{p}^\star, \mathbf{h}, \dot{\mathbf{h}}\right) + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \psi\left(\hat{\mathbf{p}}, \mathbf{h}, \dot{\mathbf{h}}\right) \implies \hat{\mathbf{p}} = \mathbf{p}^\star. \tag{4.81}$$

avec $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ et $(\beta_1, \beta_2) \in \mathbb{R}^2$. En effet, comme le suggère la remarque 4.28, il est inutile de vérifier la conditon (4.35) pour les permutations $\mathbf{P}$.

Nous résolvons le système

$$\begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \psi\left(\mathbf{p}^\star, \mathbf{h}, \dot{\mathbf{h}}\right) + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \psi\left(\hat{\mathbf{p}}, \mathbf{h}, \dot{\mathbf{h}}\right), \tag{4.82}$$

et obtenons

$$\begin{cases} \alpha_1 = 1 \\ \alpha_2 = 1 \\ a^\star = \hat{a} \\ b^\star = \hat{b} \\ \beta_1 = 0 \\ \beta_2 = 0 \end{cases},$$

Si la conjecture 4.26 est vraie, le système (4.80) est identifiable en aveugle selon le modèle d'entrées indépendantes.

4.5.3.0.2 Estimation des paramètres solutions Maintenant nous cherchons à estimer en aveugle les paramètres pour le système (4.80). Rappelons que les débits entrées u_1 et u_2 sont supposés indépendants. Comme cela est suggéré au paragraphe 4.4.1.1, nous choisissons la fonction d'estimation

$$\mathbf{h} = \begin{pmatrix} \Gamma_{u_1,u_2} \\ \Gamma_{\dot{u}_1,u_2} \end{pmatrix} \tag{4.83}$$

$$= \begin{pmatrix} E(u_1 u_2) - E(u_1) E(u_2) \\ E(\dot{u}_1 u_2) \end{pmatrix} \tag{4.84}$$

qui fournit deux équations (autant que de paramètres).

Le système inversible (4.80) considéré étant découplé, la condition $\mathbf{h}(\mathbf{u}) = \mathbf{0}$ peut être approximée, à l'aide de l'opérateur *bar*, par un système d'équations en $\mathbf{p} = (a, b)$ faisant intervenir uniquement les estimations des moments des sorties. Nous obtenons

$$\begin{cases} \overline{u_1 u_2} - \overline{u_1}\,\overline{u_2} = 0 \\ \overline{\dot{u}_1 u_2} = 0. \end{cases} \tag{4.85}$$

D'après les expressions (4.54) décrivant le système étudié et sachant que

$$\dot{u}_1(t) = \ddot{h}_1(t) + \frac{\sqrt{g} b \dot{h}_1(t)}{\sqrt{2h_1(t)}} + \frac{\sqrt{g} a \dot{h}_1(t)}{\sqrt{2|h_1(t) - h_2(t)|}}, \tag{4.86}$$

on a

$$
\begin{aligned}
\overline{u_1} &= \overline{\dot{h}_1} + b\sqrt{2g}\,\overline{\sqrt{h_1}} + a\sqrt{2g}\,\overline{sign\left(h_1 - h_2\right)\sqrt{|h_1 - h_2|}}, \\
\overline{u_2} &= \overline{\dot{h}_2} + b\sqrt{2g}\,\overline{\sqrt{h_2}} - a\sqrt{2g}\,\overline{sign\left(h_1 - h_2\right)\sqrt{|h_1 - h_2|}}, \\
\overline{u_1 u_2} &= \begin{array}{l} \overline{\dot{h}_1\dot{h}_2} + b\sqrt{2g}\,\overline{\dot{h}_2\sqrt{h_1}} + a\sqrt{2g}\,\overline{\dot{h}_2 sign\left(h_1 - h_2\right)\sqrt{|h_1 - h_2|}} + \sqrt{2g}b\,\overline{\dot{h}_1\sqrt{h_2}} + 2gb^2\,\overline{\sqrt{h_1 h_2}} \\ +2gab\,\overline{sign\left(h_1 - h_2\right)\sqrt{h_2\left|h_1 - h_2\right|}} - a\sqrt{2g}\,\overline{\dot{h}_1 sign\left(h_1 - h_2\right)\sqrt{|h_1 - h_2|}} \\ -ab\sqrt{2g}\,\overline{\sqrt{2gh_1}\, sign\left(h_1 - h_2\right)\sqrt{2g\left|h_1 - h_2\right|}} - 2a^2 g\,\overline{|h_1 - h_2|} \end{array} \\
\overline{\dot{u}_1 u_2} &= \begin{array}{l} \overline{\ddot{h}_1\dot{h}_2} + \frac{\sqrt{g}b}{2}\overline{\frac{\dot{h}_1\dot{h}_2}{\sqrt{h_1}}} + \frac{\sqrt{g}a}{2}\overline{\frac{\dot{h}_1\dot{h}_2}{\sqrt{|h_1-h_2|}}} + b\sqrt{2g}\,\overline{\ddot{h}_1\sqrt{h_2}} + b^2 g\,\overline{\dot{h}_1\sqrt{\frac{h_2}{h_1}}} \\ +gab\,\overline{\dot{h}_1\sqrt{\frac{h_2}{|h_1-h_2|}}} - a\sqrt{2g}\,\overline{\ddot{h}_1 sign\left(h_1 - h_2\right)\sqrt{|h_1 - h_2|}} \\ -agb\,\overline{\dot{h}_1 sign\left(h_1 - h_2\right)\sqrt{\frac{|h_1-h_2|}{h_1}}} - ga^2\,\overline{\dot{h}_1 sign\left(h_1 - h_2\right)}. \end{array}
\end{aligned}
\tag{4.87}
$$

La résolution de ce système d'équations d'inconnues a et b à l'aide de l'algorithme SIVIA (pour 5000 échantillons , une précision $\varepsilon = 10^{-2}$ et un pavé de recherche initial $[0, 10] \times [0, 10]$), nous donne

$$a \in [0.287, 0.314] \text{ et } b \in [0.49, 0.511].$$

qui contient les valeurs théoriques $a = 0.3$ et $b = 0.5$.

4.6 Conclusion

Dans ce chapitre, nous nous sommes intéressés au problème d'estimation de paramètres lorsque les entrées sont inconnues. Pour cela, nous avons tout d'abord défini la nouvelle notion d'identifiablilité en aveugle d'un système inversible selon un modèle d'entrée fixé, par analogie avec les notions d'identifiablilité structurelle et de séparabilité. Puis, nous avons présenté une méthode originale d'estimation de paramètres en aveugle pour des systèmes inversibles découplées. Même dans un cadre idéal (où les dérivées des signaux de sorties sont connues), il n'existait, à notre connaissance, aucune méthode permettant d'estimer les paramètres inconnus sans la connaissance des entrées.

Bien que les entrées soient inconnues, nous avons montré qu'uniquement à partir d'hypothèses statistiques sur ces dernières, l'estimation des paramètres était, dans certains cas, rendue possible. Malheureusement, nous ne sommes pas encore capables de présenter des conditions générales de succès de notre démarche dans la mesure où il existe sans doute des distributions de signaux $\mathbf{u}$ (pathologiques) pour lesquelles elle échoue. Toutefois, la méthode proposée étant basée sur l'analyse par intervalles, elle nous permet systématiquement d'estimer l'ensemble des paramètres compatibles avec les équations d'estimation considérées. L'intérêt de l'utilisation des outils ensemblistes est que même dans le cas où l'identifiablilité en aveugle d'un système ne peut être testée (pour des systèmes complexes), l'analyse par intervalles nous montre *a posteriori* que le système était identifiable. Inversement, si nous obtenons un ensemble de plusieurs paramètres $\mathbf{p}$ solutions, nous pouvons en déduire que le système est intrinsèquement non identifiable en aveugle pour le modèle d'entrées considéré ou bien que les équations d'estimation utilisées ne fournissent pas suffisamment d'informations. Dans ce dernier cas, il est alors possible d'ajouter des équations supplémentaires, par exemple, faisant intervenir des statistiques d'ordre supérieur ou des dérivées d'ordre supérieur.

Pour conclure, le travail présenté ici ne vise aucune application pratique particulière. Il s'agit simplement d'une étude théorique afin de convaincre le lecteur que l'estimation de paramètres en aveugle est possible. De nombreux points restent encore à éclaircir. En particulier, la condition d'identifiablilité proposée au paragraphe 4.2.2.2 n'est pas prouvée. Notons cependant que nous n'avons, pour le moment, encore jamais rencontré de système vérifiant cette condition et n'étant pas identifiable en aveugle.

Conclusion

Bilan

Séparation aveugle de sources

Dans cette thèse nous nous sommes tout d'abord intéressés au problème de séparation en aveugle de sources pour une nouvelle classe de modèles de mélanges décrits par des systèmes inversibles

$$\mathbf{u} = \psi\left(\mathbf{y}, \ldots, \mathbf{y}^{(r)}\right)$$

Ces modèles étendent les problèmes de séparation de sources à des mélanges décrits par des équations différentielles.

La méthode de séparation proposée consiste, dans un premier temps, à paramétrer la fonction ψ inconnue. Deux solutions sont alors envisageables : soit un modèle pour ψ est connu (par exemple, ψ est construite à partir des équations de la physique), soit aucune structure pour ψ n'est connue et une paramétrisation peut alors être obtenue par un développement de Taylor de ψ. L'ordre du développement doit être choisi suffisamment grand de manière à obtenir une approximation convenable de ψ (à noter que le développement de Taylor existe toujours puisque la fonction ψ est, par hypothèse, analytique). La question du choix de l'ordre du développement n'a pas été abordée dans cette thèse.

Dans un second temps, notre démarche repose sur la construction d'équations d'estimation originales, exploitant des résultats sur les dérivées de signaux aléatoires apparaissant dans [7] qui permettent d'estimer les paramètres inconnus de ψ. Le principal résultat étant que si deux signaux aléatoires sont statistiquement indépendants alors leurs dérivées sont statistiquement indépendantes (quel que soit l'ordre de dérivation). Ainsi, les dérivées peuvent être vues comme des observations supplémentaires ("virtuelles") qui nous permettent d'obtenir des informations supplémentaires sur le système. Ces résultats sur les dérivées de signaux, peu exploités jusqu'alors (seulement pour les mélanges linéaires statiques dans [22]), sont particulièrement intéressants pour des mélanges inversibles dans lesquels interviennent les dérivées.

La méthode de séparation proposée consiste à annuler les dérivées successives des matrices d'autocorrélation des sources reconstruites en zéro. Pour cette méthode, basée sur des statistiques d'ordre deux, nous avons montré qu'une condition nécessaire de succès est qu'il existe au moins un ordre de dérivation k tel que la dérivée k-ième de la matrice d'autocorrélation en zéro ait des valeurs propres distinctes. Cependant, cette condition concernant les sources qui sont, par hypothèse, inconnues, est difficilement exploitable.

Précisons enfin que notre démarche est continue et ne possède, en théorie, aucun inconvénient relatif à l'échantillonnage des signaux. Cependant, en pratique, une implémentation numérique impose, bien entendu, un échantillonnage.

Par ailleurs, nous avons défini la séparabilité de système inversible selon un modèle d'entrée. Cette nouvelle notion, inspirée de [88], est une propriété structurelle du système de mélange qui caractérise la possibilité d'identifier les paramètres du système qui reconstruisent une image des sources à une indétermination d'échelle et une permutation près.

A ce propos, nous avons pû nous rendre compte d'un avantage non négligeable de l'utilisation des techniques d'analyse par intervalles. Elles permettent en effet de vérifier, *a posteriori*, la séparabilité du mélange selon le modèle d'entrée considéré.

Estimation de paramètres en aveugle

Nous avons ensuite étudié le problème d'*estimation de paramètres en aveugle,* toujours pour des systèmes inversibles. Comme nous l'avons déjà expliqué, ce problème présente de fortes similitudes avec le problème de séparation aveugle de sources et nous aurions pû le nommer *séparation de sources exacte* dans le sens où l'identification du vecteur de paramètres conduit à l'estimation des sources (via l'équation du système inversible) exacte, c'est-à-dire sans indétermination d'échelle ni de permutation. A la différence du problème de séparation de sources, nous nous sommes plus particulièrement attardés sur le cas des systèmes inversibles non linéaires.

Nous avons présenté des résultats sur l'estimation de paramètres en aveugle pour ces systèmes et, notamment montré que l'observation des entrées n'est pas toujours nécessaire pour estimer les paramètres. Des hypothèses statistiques sur les entrées peuvent être suffisantes.

Dans un premier temps, nous avons introduit la notion nouvelle d'identifiabilité en aveugle selon un modèle d'entrée fixé, fortement inspirée de l'identifiabilité structurelle présentée par Walter dans [90]. Nous avons en particulier étudié les modèles d'entrées indépendantes et gaussiennes, pour lesquels nous avons conjecturé une condition formelle (empirique) d'identifiabilité en aveugle. Concernant le problème d'estimation de paramètres en aveugle, l'information initiale est tellement pauvre qu'il nous a paru intéressant d'aller chercher des renseignements supplémentaires dans les dérivées des signaux de sortie. Nos motivations sont nées de l'automatique où elles sont classiquement utilisées.
Il est fondamental de constater que nous n'aurions jamais obtenu de telles conditions d'identifiabilité en aveugle sans tenir compte des résultats sur les dérivées de signaux aléatoires.

Dans un second temps, nous avons proposé une démarche de résolution du problème d'estimation de paramètres en aveugle que nous avons illustrée sur de nombreux exemples. De façon analogue au problème de séparation de sources, notre méthode de résolution repose sur la construction d'équations d'estimation (exploitant les résultats sur les dérivées de signaux aléatoires) qui permettent d'estimer les paramètres.

Techniques d'analyse par intervalles

Pour résoudre les problèmes de séparation de sources et d'estimation de paramètres en aveugle, nous avons utilisé des techniques d'analyse par intervalles. A notre connaissance, c'est la première fois que ces techniques ont été utilisées sur des critères statistiques.

Le premier avantage des intervalles est la garantie de fournir toutes les solutions. En effet, nous obtenons toujours l'ensemble des solutions compatibles avec un système d'équations fixé et comme nous l'avons vu à plusieurs reprises, cette information est loin d'être négligeable. Un atout majeur de ces outils ensemblistes est que la séparabilité ou l'identifiabilité en aveugle d'un système selon un modèle d'entrée fixé peuvent être vérifiées *a posteriori*.
En plus de la garantie des solutions, les techniques d'analyse par intervalles offrent une estimation convenable des solutions dans la mesure où elles sont peu sensibles à de mauvais conditionnements.

Comme nous avons eu l'occasion de le préciser, les techniques ensemblistes perdent en efficacité lorsque le problème considéré augmente en dimension. En général, SIVIA est efficace jusqu'à quatre ou cinq paramètres recherchés et devient obsolète au delà de la dimension six.
Par contre, il existe des techniques par intervalles plus élaborées (Contraction, Méthode de Newton par intervalles,. . .), non présentées dans cette thèse, mais qui s'avèreraient plus efficaces pour des problèmes en grande dimension. Les développements actuels laissent d'ailleurs supposer que ces techniques vont encore se perfectionner, ne serait-ce que grâce à l'amélioration perpétuelle des processeurs d'ordinateur de plus en plus puissants. Cependant, nous n'avons pas utilisé ces techniques plus élaborées et leur avons préféré la méthode de Newton, plus simple à mettre en oeuvre. L'inconvénient est que nous retrouvons des problèmes dûs au mauvais conditionnement des équations d'estimation, ainsi que les risques liés à

l'existence de minima locaux, etc.

Perspectives

Nous avons globalement obtenu des résultats encourageants pour les deux problèmes d'estimation en aveugle étudiés :

– séparation aveugle de sources
– estimation aveugle de paramètres.

Pour ces deux problèmes, nous avons établi des équations d'estimation basées sur les dérivées de signaux aléatoires. Cependant, la mise en oeuvre pratique de ces équations (théoriques) nécessite le remplacement des moments par leurs estimations (via l'opérateur *bar*). Il est alors intéressant de s'interroger sur les conséquences de ces approximations sur les paramètres estimés. De surcroît, on peut se demander si l'ajout d'équations améliore la précision de l'estimation de ces paramètres. En effet, en pratique, nous nous sommes placés dans un contexte difficile, en exploitant systématiquement un nombre d'équations égal à celui du nombre d'inconnues, notre objectif étant simplement de montrer la possibilité d'estimer en aveugle des paramètres.
Comme le suggèrent ces remarques, une étude de robustesse reste donc à effectuer. Cette étude est d'autant plus importante que les résultats présentés (lors des simulations) sont associés à de grands nombres d'échantillons. Elle permettra de vérifier, éventuellement, l'obtention de bons résultats même pour de petits nombres d'échantillons.

Bien entendu, d'autres points restent encore à préciser et de nouvelles pistes de recherche sont également apparues :

– Comme nous l'avons évoqué dans le chapitre 3, il serait intéressant de voir laquelle des méthodes exploitant les décalages temporels ou les dérivées successives de la fonction d'autocorrélation signaux aléatoires, est la plus performante et sous quelles conditions. Nous avons vu que, lorsque les dérivées sont estimées par différences finies, les deux méthodes s'avèrent identiques. Existe t-il toujours des relations entres ces deux méthodes si les dérivées sont estimées différemment (par exemple, par des techniques présentées dans l'annexe E).
  Nous avons pris connaissance tardivement des résultats théoriques de Blanc-Lapierre [7] et des travaux de Cavassilas [22] à ce sujet, il serait intéressant, dans un premier temps, d'approfondir ces résultats.
– Concernant l'identifiabilité en aveugle, il paraît important de s'intéresser à la démonstration des conditions d'identifiabilité empiriques obtenues pour les modèles d'entrées indépendantes et gaussiennes. En effet, ces conditions d'identifiabilité, présentées dans le chapitre 4, reposent sur des remarques intuitives. Un approfondissement est nécessaire.
  A ce sujet, il pourrait également être envisageable de considérer des modèles d'entrées supplémentaires. Par exemple, les entrées sont gaussiennes et indépendantes, etc...
– Concernant le problème de séparation de sources, la démarche présentée a été validée sur des systèmes linéaires inversibles, cependant, elle semble s'étendre sans difficulté au cas des systèmes non linéaires. Toutefois, une étude supplémentaire est nécessaire afin d'obtenir la classe des mélanges de signaux séparables. En effet, la non-linéarité du mélange impose l'utilisation des statistiques d'ordres supérieurs des sorties.
– De manière générale, pour les deux problèmes considérés, l'utilisation des techniques par intervalles nous a permis, dans la mesure où l'on obtient une unique solution, de montrer la séparabilité ou l'identifiabilité en aveugle (selon le problème considéré) d'un système de mélange. Inversement

si un ensemble de solutions est obtenu, on ne peut absolument pas conclure quant à la non séparabilité ou la non identifiabilité en aveugle du système considéré. En effet, il est tout à fait possible que, bien qu'un système soit séparable ou identifiable en aveugle, les équations d'estimation choisies conduisent à une infinité de solution (par exemple si les équations sont dépendantes).
Ce point est extrêmement délicat et mériterait, à lui seul, une étude poussée. Il serait en particulier intéressant d'établir (et c'est sans aucun doute très complexe) des conditions d'unicité des solutions pour des équations d'estimation fixées, ou inversement, de rechercher dans quels cas les équations d'estimation considérées ne seront pas suffisantes.

Rappelons pour terminer que les résultats présentés dans cette thèse ne visent, pour le moment, aucune application pratique. Nous avons modestement chercher à montrer, au moins dans un cadre théorique, que la séparation de sources ou l'estimation de paramètres en aveugle étaient possibles pour des systèmes inversibles.

Annexes

ANNEXE A

Probabilités

Comme nous l'avons évoqué dans l'introduction, les problèmes de séparation de sources mettent en jeu des phénomènes aléatoires. Dans une démarche pédagogique, nous présentons dans cette première annexe les fondements des probabilités. Cette annexe, consacrée à l'exposition des notions de base des probabilités, s'appuie essentiellement sur la théorie de la mesure et de l'intégration. Elle a été réalisée principalement à partir de [41][4][77][73][71]. Pour certains résultats que nous avons choisi d'admettre, les personnes intéressées trouveront les preuves dans les deux premiers ouvrages.

Dans un premier temps, nous présentons les espaces sur lesquels sont définies les probabilités, puis, dans la seconde partie, nous introduisons la notion de variable aléatoire avant de nous intéresser dans les parties A.3 et A.4 aux couples de variables aléatoires puis aux vecteurs aléatoires, ce qui sera l'occasion d'introduire principalement les notions d'indépendance et de gaussianité.

A.1 Espace de probabilité

Une expérience ou un phénomène est dit *aléatoire* lorsque la description la plus complète des conditions de l'expérience ne suffit pas à prévoir le résultat qui est soumis au hasard. Ce dernier peut être interprété, soit comme notre incompétence à concevoir, expliciter et utiliser les lois régissant le phénomène physique considéré, soit comme un manque d'information sur les conditions de l'expérience. On peut trouver de nombreux exemples d'expériences aléatoires : le jet d'un dé, le tirage du loto... Au contraire, une expérience dont le résultat peut être prévu avant la réalisation est dite *déterministe*.

La modélisation du calcul des probabilités a été inventée par A.N. Kolmogorov dans son livre "Grundbegriffe der Wahrscheinlichkeitsrechnung" paru en 1933. Cette modélisation, dans le sens où elle induit une simplification des phénomènes aléatoires, permet de faire des calculs pour prédire des résultats. Elle est réalisée à partir du triplet fondamental $(\Omega, \mathcal{A}, P)$ que nous allons décrire dans cette première partie.

Le premier paragraphe sera consacré aux événements relatifs à un phénomène dans lequel intervient le hasard. Ces événements apparaîtront comme des sous-ensembles de l'élément Ω du triplet. Ensuite, pour permettre une manipulation plus aisée, nous introduirons une famille remarquable $\mathcal{A}$ d'événements, obéissant à des règles algébriques précises. Puis, nous étudierons les propriétés de l'élément P (la mesure de probabilité) qui nous permettra de pondérer les événements appartenant à la classe $\mathcal{A}$.

A.1.1 L'espace des observables Ω

L'ensemble Ω du triplet $(\Omega, \mathcal{A}, P)$ est appelé l'*espace des observables*. Un *événement aléatoire* est un sous-ensemble de Ω dont on peut dire, au vu d'une expérience, s'il est réalisé ou non. Par la suite, comme Ω contient tous les événements, nous l'appellerons abusivement l'*espace des événements*.

Les événements aléatoires étant des ensembles, nous pouvons effectuer les opérations logiques habituelles. Voici quelques interprétations de ces dernières.

Pour $A, B \in \Omega$,

– Ω est l'événement certain et $\varnothing$ est l'événement impossible.
– A^c est l'événement contraire à A.
– $A \cap B$ est l'événement « A et B sont réalisés ».
– $A \cup B$ est l'événement « A ou B sont réalisés ».
– $A \backslash B = A \cap (A \cap B)^c$ est l'événement « A privé de B ».
– Si $A \subset B$, la réalisation de l'événement B entraîne la réalisation de A.
– Si $A \cap B = \varnothing$, on dit que A et B sont *incompatibles*. Autrement dit, le résultat d'une expérience ne peut être à la fois dans A et dans B.

Dans la construction du modèle probabiliste, on se donne un ensemble non vide Ω et on cherche conjointement à :

- attribuer une probabilité à tout sous-ensemble de Ω,
- respecter quelques règles de calcul simple (en premier lieu celle de l'additivité).

Il se trouve, pour des raisons mathématiques (venant du fait qu'il existe des sous-ensembles de Ω de nature extrêmement complexe), qu'on ne peut pas satisfaire ces deux exigences à la fois, tout du moins lorsque Ω est continu. L'idée est alors de ne pas attribuer une probabilité à toute partie $A \in \mathcal{P}(\Omega)$, mais seulement à certaines parties appartenant à une classe particulière $\mathcal{A}$ inclue dans $\mathcal{P}(\Omega)$. Lorsque $\mathcal{A}$ est munie de propriétés algébriques bien choisies, on peut remplir les deux conditions.

Les propriétés des *tribus* se sont avérées les plus efficaces. Leurs axiomes et propriétés élémentaires sont donnés dans le paragraphe suivant.

A.1.2 La tribu $\mathcal{A}$

Pour des raisons mathématiques, et afin que la modélisation soit cohérente avec notre intuition, l'ensemble $\mathcal{A} \subset \mathcal{P}(\Omega)$ d'événements aléatoires est défini de la façon suivante :

Définition A.1 (Tribu)**.** $\mathcal{A} \subset \mathcal{P}(\Omega)$ est une tribu (ou une σ-algèbre) de Ω si :
i) $\Omega \in \mathcal{A}$,
ii) $\mathcal{A}$ est stable par passage au complémentaire :

$$\forall A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A},$$

iii) $\mathcal{A}$ est stable par réunion dénombrable : pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$,

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

On déduit de ces trois axiomes que $\mathcal{A}$ est stable par intersection dénombrable et qu'elle contient $\varnothing$.

Exemple A.2. *La plus petite tribu possible sur un espace Ω est l'ensemble $\{\Omega, \varnothing\}$, appelée tribu grossière ou tribu triviale.*

Définition A.3 (Sous-tribu)**.** On dit que $\mathcal{A}'$ est une sous-tribu d'une tribu $\mathcal{A}$ si $\mathcal{A}' \subset \mathcal{A}$ et $\mathcal{A}'$ est une tribu.

Définition A.4 (Tribu engendrée par)**.** Soit $\mathcal{A}$ une tribu, la tribu engendrée par une famille d'événements de $\mathcal{A}$ est la plus petite sous-tribu de $\mathcal{A}$ contenant tous ces événements.

Exemple A.5. *Soit $\mathcal{A}$ une tribu de Ω, la tribu engendrée par un événement $A \in \mathcal{A}$ est l'ensemble $\{\varnothing, \Omega, A, A^c\}$.*

Définition A.6 (Tribu borélienne)**.** Soit Ω un espace topologique (par exemple $\Omega = \mathbb{R}$). On appelle tribu borélienne, notée $\mathcal{B}(\Omega)$, la plus petite tribu contenant tous les ouverts de Ω, appelée aussi tribu engendrée par les ouverts. Dans le cas où $\Omega = \mathbb{R}$ (muni de sa topologie classique), la tribu engendrée par les intervalles ouverts est appelée tribu borélienne.

Définition A.7 (Espace mesurable)**.** On appelle espace mesurable le couple $(\Omega, \mathcal{A})$ où $\mathcal{A}$ est une tribu de Ω.

Nous allons maintenant introduire la notion de probabilité.

A.1.3 La mesure de probabilité P

A.1.3.1 Probabilité sur un espace mesurable

On cherche à définir, pour $A \in \mathcal{A}$, la crédibilité qu'on accorde *a priori* à A (avant le résultat de l'expérience), c'est-à-dire à associer à l'événement A un nombre $P(A)$, compris entre 0 et 1, qui représente la chance qu'il soit réalisé à la suite de l'expérience.

Définition A.8 (Probabilité)**.** Une probabilité (ou mesure de probabilité) sur un espace mesurable $(\Omega, \mathcal{A})$ est une application de $\mathcal{A}$ dans $[0, 1]$, notée P, telle que :
$i)$ $P(\Omega) = 1$
$ii)$ Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements de $\mathcal{A}$, incompatibles deux à deux, c'est-à-dire tels que $\forall m, n \in \mathbb{N}$, $A_n \cap A_m = \varnothing$, on a :

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} P(A_n).$$

Cette dernière propriété s'appelle σ-additivité de P.
Dans le langage de la théorie de la mesure, P est une *mesure positive de masse* 1.

Proposition A.9. *P vérifie les propriétés suivantes :*
$i)$ $\forall A \in \mathcal{A}, P(A^c) = 1 - P(A)$.
$ii)$ $\forall A, B \in \mathcal{A}$, *tels que* $A \subset B, P(A) \leq P(B)$.
$iii)$ $\forall A, B \in \mathcal{A}, P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$iv)$ $\forall (A_n)_{n\in\mathbb{N}}$ *telle que* $\forall n \in \mathbb{N},\ A_n \in \mathcal{A}, P(\bigcup_n A_n) \leq \sum_n P(A_n)$.

Démonstration. i) A^c et A sont disjoints donc $A^c \cap A = \emptyset$. De plus, $A^c \cup A = \Omega$ donc $P(A^c) + P(A) = P(\Omega)$. Comme $P(\Omega) = 1$, on a donc $P(A^c) = 1 - P(A)$.
ii) $B = (B \cap A) \cup (B \cap A^c)$ avec $B \cap A$ et $B \cap A^c$ incompatibles donc $P(B) = P(B \cap A) + P(B \cap A^c)$. De plus, $A \subset B$, donc $B \cap A = A$. On a alors $P(B) = P(A) + P(B \cap A^c)$ avec $P(B \cap A^c) \geqslant 0$, d'où le résultat.
iii) On peut écrire $A \cup B$ comme réunion de sous ensembles disjoints :

$$A \cup B = (A \cap B) \cup (A^c \cap B) \cup (A \cap B^c),$$

d'où

$$P(A \cup B) = P(A \cap B) + P(A^c \cap B) + P(A \cap B^c).$$

De même qu'en ii), on peut écrire:

$$P(B) = P(B \cap A) + P(B \cap A^c) \text{ et } P(A) = P(B \cap A) + P(B^c \cap A),$$

donc

$$P(A \cup B) = P(A \cap B) + P(A) - P(A \cap B) + P(B) - P(A \cap B).$$

On a bien le résultat.
iv) Par récurrence, c'est une conséquence directe du iii). □

Théorème A.10. *Si $P : \mathcal{A} \to [0,1]$ est une probabilité, alors:*
i) Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ telle que $\forall n \in \mathbb{N}$, $A_n \subset A_{n+1}$ et $\lim_{n\to+\infty} A_n = A$, on a

$$\lim_{n\to+\infty} P(A_n) = P(A).$$

ii) Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ telle que $\forall n \in \mathbb{N}$, $A_{n+1} \subset A_n$ et $\lim_{n\to+\infty} A_n = A$, on a

$$\lim_{n\to+\infty} P(A_n) = P(A).$$

Autrement dit, on peut parler de continuité de probabilité pour des suites croissantes ou décroissantes d'éléments de $\mathcal{A}$.

Démonstration. i) Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $\mathcal{A}$ telle que $\lim_{n\to+\infty} A_n = A$. On a donc $\forall n \in \mathbb{N}$, $A_n \subset A_{n+1}$.
Considérons la suite $(B_n)_{n\in\mathbb{N}}$ définie par :

$$B_0 = A_{0,} B_1 = A_1 \backslash A_0, ..., B_n = A_n \backslash A_{n-1}.$$

On vérifie aisément que $\forall n \in \mathbb{N}$, $B_n \in \mathcal{A}$ et que les B_n sont deux à deux disjoints. De plus, $\forall n \in \mathbb{N}$, $A_n = \bigcup_{i=0}^{n} B_i$ et $A = \bigcup_{n=0}^{\infty} B_n$.
On a alors, d'après la σ-additivité de P,

$$P(A_n) = P(\bigcup_{i=0}^{n} B_i) = \sum_{i=0}^{n} P(B_i) \text{ et } P(A) = P(\bigcup_{n=0}^{\infty} B_n) = \sum_{n=0}^{\infty} P(B_n).$$

Or, il est évident (définition d'une somme partielle) que

$$\lim_{n\to+\infty} \sum_{i=0}^{n} P(B_i) = \sum_{n=0}^{\infty} P(B_n).$$

On a donc bien $\lim_{n\to+\infty} P(A_n) = P(A)$.
ii) Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante d'éléments de $\mathcal{A}$ telle que $\lim_{n\to+\infty} A_n = A$.
Alors $(A_n^c)_{n\in\mathbb{N}}$ est une suite croissante d'éléments de $\mathcal{A}$ telle que $\lim_{n\to+\infty} A_n^c = A^c$. Donc, d'après i), $\lim_{n\to+\infty} P(A_n^c) = P(A^c)$.
On a donc, $\lim_{n\to+\infty} 1 - P(A_n) = (1 - P(A))$ soit encore, $\lim_{n\to+\infty} P(A_n) = P(A)$. □

Définition A.11 (Espace de probabilité)**.** Soient Ω l'espace des événements, $\mathcal{A}$ une tribu de Ω et P une probabilité de $\mathcal{A}$. Le triplet $(\Omega, \mathcal{A}, P)$ s'appelle un espace de probabilité ou un espace probabilisé.

Définition A.12. Un événement $A \in \mathcal{A}$ de probabilité nulle (*i.e.* $P(A) = 0$) est dit *négligeable*, et une propriété est dite *vraie P-presque sûrement*, si l'ensemble des $\omega \in \Omega$ tels que cette propriété soit vraie est de probabilité égale à 1.

Considérons deux exemples classiques de probabilités :

Exemple A.13 (Mesure de Dirac)**.** *Pour ω_0 fixé dans Ω, on appelle mesure de Dirac la mesure de probabilité $\delta_{\omega_0} : \Omega \to \mathbb{R}$ définie par*

$$\delta_{\omega_0}(A) = \begin{cases} 1 & si\ \omega_0 \in A, \\ 0 & si\ \omega_0 \notin A. \end{cases} \tag{A.1}$$

Exemple A.14 (Mesure de Lebesgue)**.** *Soit $\Omega = [0,1]$ et $\mathcal{B}([0,1])$ sa tribu borélienne. On appelle mesure de Lebesgue la probabilité $\lambda : [0,1] \to \mathbb{R}$ définie par*

$$\lambda\left(]c,d[\right) = d - c \tag{A.2}$$

Remarque A.15. *On peut montrer (voir [10]) qu'il n'existe pas sur $([0,1], \mathcal{P}([0,1]))$ de probabilité autre qu'une mesure discrète (somme au plus dénombrable de masses de Dirac). En particulier, on peut exhiber une partie de $[0,1]$ ne possédant pas de mesure de Lebesgue. C'est pourquoi on munira naturellement $\mathbb{R}$, et par suite tout espace topologique X, de sa tribu borélienne $\mathcal{B}(X)$.*

A.1.3.2 Fonction de répartition

Dans ce paragraphe, on prendra $\Omega = \mathbb{R}$ (où $\mathbb{R}$ est muni de sa tribu borélienne).

Définition A.16 (Fonction de répartition)**.** On appelle fonction de répartition de la probabilité P, notée F, l'application définie par

$$\forall x \in \mathbb{R}, F(x) = P\left(]-\infty, x]\right).$$

Théorème A.17. *Une fonction $F : \mathbb{R} \mapsto [0,1]$ est une fonction de répartition de la probabilité P sur l'espace $(\Omega, \mathcal{A})$ si et seulement si F vérifie les trois propriétés suivantes :*
i) F est croissante,
ii) F est continue à droite,
iii) $\lim_{x\to+\infty} F(x) = 1$ *et* $\lim_{x\to-\infty} F(x) = 0$.

Démonstration. ($\Longrightarrow$) Soit F une fonction de répartition de la probabilité P.
i) Soit $x, y \in \mathbb{R}$ tels que $x \leqslant y$. On a $]-\infty, x] \subseteq]-\infty, y]$, d'où

$$P(]-\infty, x]) \leq P(]-\infty, y]),$$

c'est à dire $F(x) \leq F(y)$. Donc F est croissante.
ii) Soit $x \in \mathbb{R}$ et soit $(x_n)_{n\in\mathbb{N}}$, une suite décroissante d'éléments de $\mathbb{R}$ telle que $\lim_{n\to+\infty} x_n = x$. On a donc une suite croissante d'éléments de $\mathcal{A}$, $(]-\infty, x_n])_{n\in\mathbb{N}}$ qui tend vers $]-\infty, x]$. D'après le théorème A.10, on a donc

$$\lim_{n\to+\infty} P(]-\infty, x_n]) = P(]-\infty, x]),$$

c'est-à-dire $\lim_{n\to+\infty} F(x_n) = F(x)$. F est donc bien continue à droite.
iii) On a $\lim_{x\to+\infty}]-\infty, x] = \mathbb{R}$. Par continuité de probabilité pour une suite croissante, on a donc

$$\lim_{x\to+\infty} P(]-\infty, x]) = P(\mathbb{R}),$$

d'où $\lim_{x\to+\infty} F(x) = 1$.
De même, $\lim_{x\to-\infty}]-\infty, x] = \emptyset$. Par continuité de probabilité pour une suite décroissante, on a

$$\lim_{x\to-\infty} P(]-\infty, x]) = P(\emptyset),$$

d'où $\lim_{x\to-\infty} F(x) = 0$.
($\Longleftarrow$) Réciproquement, soit F une fonction vérifiant i), ii) et iii). A partir de cette fonction F, nous allons chercher à définir une probabilité P sur la tribu borélienne $\mathcal{B}(\mathbb{R})$. Pour $x, y \in \overline{\mathbb{R}}$, nous posons intuitivement $P(]x, y]) \triangleq F(y) - F(x)$. Il faut maintenant définir $P(A)$ pour tout $A \in \mathcal{B}(\mathbb{R})$. On a

$$\mathcal{B}(\mathbb{R}) = \left\{\text{union finie des intervalles disjoints de type }]x, y],\ x < y \text{ avec } x, y \in \bar{\mathbb{R}}\right\}^{1}.$$

Ceci vient du fait que tout semi-ouvert de $\mathbb{R}$ peut s'écrire comme réunions et complémentaires d'intervalles ouverts de $\mathbb{R}$ ($]x, y] =]x, y[\cup (] - \infty; y[\cup]y; +\infty[)^c$).

Tout élément $A \in \mathcal{A}$ s'écrit donc comme union d'intervalles deux à deux disjoints : $A = \bigcup_{i=1}^{n}]x_i, y_i]$.

Nous posons alors

$$P(A) = \sum_{i=1}^{n} P(]x_i, y_i]) = \sum_{i=1}^{n} \left[F(y_i) - F(x_i)\right].$$

Il faut maintenant vérifier que $P : \mathcal{A} \longrightarrow [0, 1]$ est bien une probabilité.
i) $P\left(] - \infty, +\infty]\right) = F(+\infty) - F(-\infty) = 1$ donc $P(\Omega) = 1$.
ii) Reste à montrer la σ-additivité de P. Nous admettons ici ce résultat (difficile). □

D'après ce théorème, si une fonction F possède les trois propriétés i), ii) et iii), il existe une probabilité P sur $(\Omega, \mathcal{A})$ dont F est la fonction de répartition et inversement. On peut également montrer que cette probabilité est unique. Ainsi, la fonction de répartition F caractérise entièrement la probabilité P de l'espace mesurable $(\Omega, \mathcal{A})$, et donc caractérise entièrement l'espace de probabilité $(\Omega, \mathcal{A}, P)$. Autrement dit, la connaissance de F nous permet d'ignorer totalement l'espace de probabilité $(\Omega, \mathcal{A}, P)$. Ce résultat est intéressant étant donné que l'ensemble des fonctions de répartition est moins abstrait que celui des espaces probabilisés.

Exemple A.18. *La fonction de répartition de la mesure de Dirac définie par (A.1) est :*

$$F(x) = \begin{cases} 0 & \textit{si } x < \omega_0 \\ 1 & \textit{si } x \geq \omega_0. \end{cases}$$

Exemple A.19. *La fonction de répartition de la mesure de Lebesgue définie par (A.2) est :*

$$F(x) = \begin{cases} 0 & \textit{si } x \leq 0 \\ x & \textit{si } 0 \leq x \leq 1 \\ 1 & \textit{si } x \geq 1. \end{cases}$$

[1]Par convention $]x, +\infty] \equiv]x, +\infty[$.

A.1.4 Conditionnement et indépendance

A.1.4.1 Probabilités conditionnelles

Une des plus fructueuses notions de la théorie des probabilités est celle de probabilité conditionnelle. L'idée de base est qu'une information supplémentaire concernant l'expérience modifie la vraisemblance que l'on accorde à l'événement étudié.

Définition A.20 (Probabilité conditionnelle)**.** Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité et B un événement de $\mathcal{A}$ de probabilité strictement positive. L'application

$$P(./B) : \begin{cases} \mathcal{A} & \rightarrow & [0,1] \\ A & \rightarrow & P(A/B) = \frac{P(A \cap B)}{P(B)} \end{cases}$$

définit une nouvelle probabilité sur Ω, appelée probabilité conditionnelle de A sachant B.

A.1.4.2 Notion d'indépendance

Intuitivement, deux événements A et B sont indépendants si le fait de savoir que A est réalisé ne donne aucune information sur la réalisation de B et inversement.
Si B est un événement de probabilité strictement positive, A sera dit indépendant de B si :

$$P(A/B) = P(A) = \frac{P(A \cap B)}{P(B)}.$$

On remarque que cette formule se symétrise et la notion d'indépendance se définit finalement comme suit.

Proposition A.21 (Evénements indépendants)**.** *Soient deux événements aléatoires A et B de $(\Omega, \mathcal{A}, P)$ de probabilité strictement positive. A et B sont indépendants si et seulement si*

$$P(A/B) = P(A) \Leftrightarrow P(B/A) = P(B) \Leftrightarrow P(A \cap B) = P(A)\,P(B).$$

Autrement dit, la probabilité de voir A réalisé ne dépend pas de la réalisation ou non de B (et inversement).

Remarque A.22. *La notion d'indépendance est liée au choix de la probabilité P et n'est pas une notion ensembliste. Cela n'a en particulier rien à voir avec le fait que A et B soient disjoints ou non.*

On peut généraliser la notion d'événements indépendants à une suite finie ou infinie d'événements.

Définition A.23. Une suite $(A_n)_n$ d'événements de $(\Omega, \mathcal{A}, P)$ est dite indépendante, ou mutuellement indépendante, si

$$P(A_{i_1} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k}),$$

pour toute suite finie $(i_1, \ldots, i_k)$ d'entiers deux à deux distincts.

Attention, comme le montrent les exemples suivants, cette définition est un peu délicate.

Exemple A.24. *Pour que la suite d'événements (A_1, A_2, A_3) soit indépendante, il ne suffit pas d'avoir*

$$P(A \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

Considérons un lancer de dé non pipé avec $A_1 = \{1,2,3\}$, $A_2 = \{2,4,6\}$, $A_3 = \{1,2,4,5\}$. On a :

$$\begin{array}{ll} P(A_1 \cap A_2 \cap A_3) = P(\{2\}) = \frac{1}{6}, & P(A_1 \cap A_2) = P(\{2\}) = \frac{1}{6}, \\ P(A_1) = P(A_2) = \frac{1}{2}, & P(A_3) = \frac{2}{3}. \end{array}$$

On a donc bien $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$. Cependant, $P(A_1 \cap A_2) \neq P(A_1)P(A_2)$. Les événements A_1, A_2, A_3 ne sont donc pas mutuellement indépendants.

Exemple A.25. *Pour que la suite (A_1, A_2, A_3) soit indépendante, il ne suffit pas non plus que les événements A_1, A_2 et A_3 soient indépendants deux à deux.*
Considérons deux lancers successifs d'une pièce et les événements $A_1 = \{$face au premier lancer$\}$, $A_2 = \{$face au deuxième lancer$\}$ et $A_3 = \{$les deux tirages donnent le même résultat$\}$. On a

$$\begin{array}{llll} P(A_1 \cap A_2 \cap A_3) = \frac{1}{4}, & P(A_1 \cap A_2) = \frac{1}{4}, & P(A_2 \cap A_3) = \frac{1}{4}, & P(A \cap A_3) = \frac{1}{4}, \\ P(A_1) = \frac{1}{2}, & P(A_2) = \frac{1}{2}, & P(A_3) = \frac{1}{2}. & \end{array}$$

On voit que les événements A_1, A_2 et A_3 sont deux à deux indépendants, mais que la probabilité de leur intersection n'est pas égale au produit des probabilités. Ils ne sont donc pas mutuellement indépendants.

La notion d'événements indépendants est, certes, élémentaire, mais elle est peu maniable et ne reflète pas vraiment la réalité. Intuitivement, on pense plutôt à l'indépendance de groupes d'événements. En effet, par un calcul très simple, on peut vérifier que, si les événements A et B sont indépendants, alors il en est de même des événements A et B^c, A^c et B, A^c et B^c. La bonne notion de "groupe" d'événements étant en fait celle de sous-tribu, on a la généralisation suivante :

Définition A.26 (Tribus indépendantes)**.** Soit $\mathcal{A}_1$, ...,$\mathcal{A}_n$ une famille finie de sous-tribus d'un espace de probabilité $(\Omega, \mathcal{A}, \mathcal{P})$. On dit que cette famille est indépendante si pour tous $B_i \in \mathcal{A}_i$, $i = 1, \ldots, n$, on a

$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \ldots P(B_n).$$

A.2 Variables aléatoires à valeurs réelles

En théorie des probabilités, on préfère travailler sur les variables aléatoires que sur les événements. Attention, la terminologie, consacrée par l'usage, est malencontreuse. En effet, une "variable" aléatoire n'est pas une variable (au sens de l'analyse) mais une application de Ω dans un espace E, (en général $E = \mathbb{R}$ ou $E = \mathbb{R}^d$).

Avant de définir plus précisément une variable aléatoire, commençons par définir une application mesurable.

Définition A.27 (Application mesurable)**.** Soient $(\Omega, \mathcal{A})$ et $(E, \mathcal{E})$ deux espaces mesurables. On dit qu'une application $X : \Omega \longrightarrow E$ est mesurable si pour tout $B \in \mathcal{E}$, on a

$$X^{-1}(B) \triangleq \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}.$$

Exemple A.28. *Soient $\Omega = \{1,2,3\}$, $\mathcal{A} = \{\varnothing, \Omega, \{1,2\}, \{3\}\}$, $E = \Omega$ et $\mathcal{E} = \mathcal{P}(\Omega)$. On considère l'application X définie par $X(\omega) = \omega$.*
On a : $X^{-1}(\{2\}) = \{2\} \notin \mathcal{A}$, donc X n'est pas mesurable.

Exemple A.29. *Soit Ω un ensemble muni de la tribu $\mathcal{A} = \{\varnothing, A, A^c, \Omega\}$ avec $A \in \mathcal{P}(\Omega)$. On considère l'application X définie par*

$$X : \begin{cases} \Omega & \rightarrow & \mathbb{R} \\ \omega & \rightarrow & \begin{cases} a & si & \omega \in A \\ b & si & \omega \in A^c. \end{cases} \end{cases}$$

Pour $B \in \mathcal{B}(\mathbb{R})$, on a

$$X^{-1}(B) = \begin{cases} \varnothing & si & a \notin B & et & b \notin B \\ A & si & a \in B & et & b \notin B \\ A^c & si & a \notin B & et & b \in B \\ \Omega & si & a \in B & et & b \in B. \end{cases}$$

On a $\forall B \in \mathcal{B}(\mathbb{R})$, $X^{-1}(B) \in \mathcal{A}$. Donc X est mesurable.

Remarque A.30. *Toute application $X : (\Omega, \mathcal{P}(\Omega)) \longrightarrow (E, \mathcal{E})$ est mesurable.*

Nous pouvons maintenant définir une variable aléatoire.

Définition A.31 (Variable aléatoire)**.** Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $(E, \mathcal{E})$ un espace mesurable. On appelle variable aléatoire toute application mesurable X de Ω à valeurs dans E.
Par définition même d'une fonction mesurable, une variable aléatoire X est donc caractérisée par la propriété suivante :

$$\forall B \in \mathcal{E},\, X^{-1}(B) \in \mathcal{A}.$$

Remarque A.32. *Si $E = \mathbb{R}$, on parle de variable aléatoire à valeurs réelles.*
Si $E = \mathbb{R}^n$, $X = (X_1, X_2, \ldots, X_n)^T$, on parle de vecteur aléatoire à valeurs réelles.
Si $E = \mathbb{R}^{\mathbb{R}}$, on parle de signal aléatoire à valeurs réelles.

Nous pouvons maintenant "transporter" la structure abstraite du modèle probabiliste $(\Omega, \mathcal{A}, P)$ sur l'espace d'arrivée $(E, \mathcal{B})$. Nous allons pour cela introduire la notion de loi de probabilité.

Définition A.33 (Loi de probabilité)**.** Soit X une variable aléatoire de $(\Omega, \mathcal{A}, P)$ dans $(E, \mathcal{E})$. On appelle loi de probabilité de X, ou loi de X, la probabilité image de P par l'application X, notée P_X et définie par :

$$\forall B \in \mathcal{E}, P_X(B) = P\left(X^{-1}(B)\right) = P\left(\{\omega, X(\omega) \in B\}\right).$$

On vérifie aisément que $P_X(B)$ est une probabilité sur E. Dès lors, $(E, \mathcal{E}, P_X)$ est un espace probabilisé.

Définition A.34 (Fonction de répartition d'une variable aléatoire)**.** On appelle fonction de répartition d'une variable aléatoire X, la fonction de répartion, notée F_X , de sa loi de probabilité P_X :

$$F_X(t) = P_X(]-\infty; t]) = P\left(\{w; X(w) \leq t\}\right).$$

Remarque A.35. *La notion de loi de probabilité est fondamentale dans le sens où, à toute mesure de probabilité P sur l'espace $(E, \mathcal{E})$, on peut associer un espace probabilisé $(\Omega, \mathcal{A}, P)$ et une variable aléatoire $X : (\Omega, \mathcal{A}, P) \rightarrow (E, \mathcal{E})$ de sorte que P soit la loi de probabilité de X. En effet, il suffit de prendre $\Omega = E$, $\mathcal{A} = \mathcal{E}$ et pour X l'application identité de Ω.*
On pourra donc parler d'une variable aléatoire X ayant une loi de probabilité P_X sans mentionner l'espace $(\Omega, \mathcal{A}, P)$ sur lequel X est définie.

Toute l'information probabiliste d'une variable aléatoire est concentrée dans sa loi de probabilité. Ce résultat est intéressant car notre nouvel espace de travail $(E, \mathcal{E}, P_X)$ est beaucoup moins abstrait. On verra, dans le cas où $E = \mathbb{R}$ ou $\mathbb{R}^d$, que l'on peut caractériser les probabilités sur $(E, \mathcal{E})$ par des quantités numériques ou fonctionnelles simples.

A.2.1 Densité de probabilité

Dans cette partie, nous allons introduire la notion de densité d'une variable aléatoire, fortement liée à la notion de fonction de répartition définie précédemment.

Définition A.36 (Densité)**.** On appelle densité ou densité de probabilité sur $\mathbb{R}$ toute fonction $f : \mathbb{R} \to \mathbb{R}$ vérifiant les conditions suivantes :
- f est positive,
- f est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points,
- l'intégrale $\int_{-\infty}^{+\infty} f(t)dt$ converge et vaut 1.

Exemple A.37. *Soit a, un réel strictement positif. Considérons la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par*

$$f(x) = \begin{cases} a\exp(-ax) & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases}$$

On vérifie aisément que f est une densité de probabilité sur $\mathbb{R}$.

Rappelons qu'une variable aléatoire X définie sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est entièrement caractérisée par sa fonction de répartition F_X.

Définition A.38 (Densité d'une variable aléatoire)**.** Une variable aléatoire X est dite absolument continue si il existe une fonction $f : \mathbb{R} \to \mathbb{R}$, positive, continue sauf éventuellement en un nombre fini de points, et telle que pour tout réel x,

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

La fonction f est alors une densité (car $\lim_{x\to+\infty} F_X(x) = 1$), appelée densité de la variable aléatoire X et notée f_X.

Remarque A.39. *D'après la définition A.36, si f est une densité d'une variable aléatoire X, toute fonction positive différant de f en un nombre fini de points est encore une densité de X. On dit alors que la densité d'une variable aléatoire est définie à un nombre fini de points près.*

Théorème A.40. *Toute densité de probabilité sur $\mathbb{R}$ est la densité d'une variable aléatoire absolument continue sur $\mathbb{R}$.*

Démonstration. Soit f une densité sur $\mathbb{R}$. Posons

$$F : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ x & \to & \displaystyle\int_{-\infty}^{x} f(t)dt. \end{cases}$$

On vérifie aisément que la fonction F ainsi définie vérifie toutes les propriétés du théorème A.17, il s'agit donc d'une fonction de répartition. Or, on a vu précédemment qu'on peut trouver une probabilité et, par conséquent, une variable aléatoire X dont F est la fonction de répartition. Cette variable aléatoire X est absolument continue, de densité f. $\square$

A.2.2 Espérance, variance et moments d'une variable aléatoire réelle

A.2.2.1 Intégration

Sans rentrer dans les détails, nous allons définir l'intégrale d'une fonction mesurable.

Définition A.41 (Intégrabilité)**.** Soit X une variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$. On pose $X = X^+ - X^-$ où $X^+ = \sup(X, 0)$ et $X^- = \sup(-X, 0)$.
On dit que X est P-intégrable sur $B \in \mathcal{A}$ si $\int_B |X| \, dP < +\infty$. Dans ce cas, on définit et écrit l'intégrale de X par rapport à P sur B comme suit

$$\int_B X dP = \int_B X^+ dP - \int_B X^- dP.$$

Si $B = \Omega$, on dit que X est P-intégrable, et note $\int X dP = \int_\Omega X dP$.

Définition A.42 (Espace intégrable)**.** Soit p un entier tel que $1 \leq p \leq \infty$. On note $\mathcal{L}^p(\Omega, \mathcal{A}, P)$ l'espace contenant les variables aléatoires X telles que $|X|^p$ soit intégrable, c'est-à-dire que $\int_\Omega |X|^p \, dP < +\infty$.

Proposition A.43. *L'espace $\mathcal{L}^p(\Omega, \mathcal{A}, P)$, $1 \leq p \leq \infty$, est un espace vectoriel.*

Démonstration. On vérifie aisément que $\forall x, y \geq 0$, $(x + y)^p \leq 2^p \sup(x^p, y^p) \leq 2^p(x^p + y^p)$.
Si $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, P)$ alors $f + g$ est mesurable et $0 \leq |f + g|^p \leq 2^p(|f|^p + |g|^p)$.
Donc $|f + g|^p$ est intégrable. Ce qui montre que $f + g \in \mathcal{L}^p(\Omega, \mathcal{A}, P)$. □

Nous énonçons maintenant deux résultats fondamentaux avant d'introduire la notion d'espérance d'une variable aléatoire.

Théorème A.44 (Théorème de "transport")**.** *Soit $X : (\Omega, \mathcal{A}, P) \longrightarrow (E, \mathcal{E}, P_X)$ une variable aléatoire et soit $g : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction mesurable, P_X-intégrable. Alors*

$$\int_\Omega g \circ X(\omega) \, dP(\omega) = \int_E g(x) \, dP_X(x).$$

Démonstration. Admis. Nous pouvons toutefois préciser que pour démontrer ce théorème, on peut d'abord prouver la formule pour des variables aléatoires h simples positives puis, à l'aide de résultats sur les suites croissantes, la généraliser. □

Théorème A.45. *Soit $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire absolument continue, de densité f_X et de loi P_X. Pour toute variable aléatoire $g : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on a l'identité*

$$\int_\mathbb{R} g(x) dP_X(x) = \int_\mathbb{R} g(x) \, f_X(x) \, dx,$$

dans le sens où si l'une de ces deux intégrales existe, l'autre existe aussi et elles sont égales.

Démonstration. Admis. Comme pour le théorème de transport, l'idée de la démonstration est de montrer l'égalité dans un premier temps pour des variables aléatoires g simples positives puis de la généraliser. □

Dans le cas des variables aléatoires admettant une densité, ce dernier théorème nous permettra de faire des calculs en intégrant par rapport à la mesure de Lebesgue plutôt que par rapport à une mesure de probabilité compliquée.

A.2.2.2 Espérance d'une variable aléatoire

Grâce aux résultats précédents, nous pouvons maintenant intégrer les variables aléatoires. L'objectif est de résumer par une valeur numérique le comportement moyen d'une variable aléatoire définie sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}$.

Définition A.46 (Espérance)**.** Soit X une variable aléatoire réelle définie de $(\Omega, \mathcal{A}, P)$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$. Si X est intégrable, on peut définir le nombre réel, appelé l'espérance mathématique (ou l'espérance) de la variable aléatoire X, par

$$\begin{aligned} E(X) &= \int_{\Omega} X(\omega)\, dP(\omega) \\ &= \int_{\mathbb{R}} x dP_X(x). \end{aligned}$$

La deuxième égalité est une application directe du théorème de "transport" A.44 (il suffit de prendre $g(x) = x$).

On dit que X est centrée si elle est intégrable et si $E(X) = 0$.

Proposition A.47. *L'espérance mathématique est une forme linéaire de* $\mathcal{L}^1(\Omega, \mathcal{A}, P)$.

Démonstration. Soient X et Y deux variables aléatoires réelles intégrables (*i.e.* $X, Y \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$) et soient $\alpha, \beta \in \mathbb{R}$, on a

$$\begin{aligned} E(\alpha X + \beta Y) &= \textstyle\int_{\Omega} (\alpha X(\omega) + \beta Y(\omega))\, dP(\omega) \\ &= \textstyle\int_{\Omega} \alpha X(\omega)\, dP(\omega) + \int_{\Omega} \beta Y(\omega)\, dP(\omega) \quad \text{(linéarité de l'intégrale)} \\ &= \alpha \textstyle\int_{\Omega} X(\omega)\, dP(\omega) + \beta \int_{\Omega} Y(\omega)\, dP(\omega) \\ &= \alpha E(X) + \beta E(Y) < \infty, \end{aligned}$$

donc $\alpha X + \beta Y \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$. □

Remarque A.48. *Soit X une variable aléatoire absolument continue, intégrable, de densité f. On a*

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx.$$

Ce résultat est une conséquence directe du théorème A.45. Cette formule importante permet le calcul des espérances mathématiques des variables aléatoires réelles admettant une densité.

A.2.2.3 Moments

Définition A.49 (Moment d'ordre p)**.** Soient $p \geq 1$ un entier et $X \in \mathcal{L}^p(\Omega, \mathcal{A}, P)$. On appelle moment d'ordre p de la variable aléatoire X la quantité

$$E(X^p) = \int_{\Omega} X^p(\omega)\, dP(\omega) = \int_{\mathbb{R}} x^p dP_X(x).$$

Dans la littérature, les moments $E(X^p)$, $p > 2$, sont appelés *moments d'ordre supérieur* (à 2).

Remarque A.50. *Soit $X \in \mathcal{L}^p(\Omega, \mathcal{A}, P)$ une variable aléatoire absolument continue de densité f. On a*

$$E(X^p) = \int_{-\infty}^{+\infty} x^p f(x)\, dx.$$

Cette formule est, là encore, une conséquence du théorème A.45.

A.2.2.4 Variance

Une variable aléatoire étant une fonction qui fluctue en fonction du hasard, elle n'est, en général, pas constante. Dans le paragraphe A.2.2.2, on a associé à X une valeur moyenne $E(X)$ résumant son comportement. On s'intéresse maintenant à une quantification des écarts entre X et $E(X)$. Une distance pratique pour mesurer cet écart est la distance quadratique. La quantité $(X - E(X))^2$ est également une variable aléatoire. On peut s'intéresser, lorsqu'elle existe, à son espérance que l'on appellera variance de X.

Définition A.51 (Variance)**.** Soit $X \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$. On appelle variance de X, notée Var(X) , la quantité positive suivante :

$$\begin{aligned} \text{Var}(X) = E\left((X - E(X))^2\right) &= \int_\Omega (X(\omega) - E(X(\omega)))^2\, dP(\omega) \\ &= \int_\mathbb{R} (x - E(X))^2\, dP_X(x). \end{aligned}$$

On appelle écart-type, noté $\sigma(X)$, la racine carrée de la variance :

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

Remarque A.52. *Si X est une variable aléatoire réelle quelconque, la variable aléatoire $\frac{X-E(X)}{\sigma(X)}$ est toujours une variable centrée[2] réduite[3].*

En utilisant la linéarité de l'espérance, on peut montrer que si X_1 et X_2 sont deux variables aléatoires de carré intégrable, on a :

$$\text{Var}(aX_1 + bX_2) = a^2\text{Var}(X_1) + b^2\text{Var}(X_2) + 2abE((X_1 - E(X_1))(X_2 - E(X_2))).$$

La quantité $E((X_1 - E(X_1))(X_2 - E(X_2)))$ joue un rôle fondamental par la suite.

Définition A.53 (Covariance)**.** Soient X_1,$X_2 \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$. La covariance de X_1 et de X_2, notée Γ_{X_1,X_2} est définie par

$$\Gamma_{X_1,X_2} = E((X_1 - E(X_1))(X_2 - E(X_2))).$$

Elle dépend uniquement de la loi du couple (X_1, X_2). Par un simple développement, on montre que :

$$\Gamma_{X_1,X_2} = E(X_1X_2) - E(X_1)E(X_2).$$

Nous reviendrons sur cette notion de covariance après avoir introduit la notion d'indépendance de deux variables aléatoires.

[2] $E(X) = 0$
[3] $\sigma(X) = 1$

A.2.3 Exemples de lois usuelles

Nous donnons dans cette partie quelques exemples de lois de probabilité utilisées couramment afin d'illustrer les notions que nous venons d'introduire.

A.2.3.1 La loi uniforme

C'est la plus simple des lois admettant une densité.

Définition A.54 (Loi uniforme)**.** Soient a et b, deux réels tels que $a < b$. On dit qu'une variable aléatoire réelle X suit la loi uniforme sur l'intervalle $[a, b]$ si elle admet pour densité la fonction (voir figure A.1)

$$f : x \to \begin{cases} \frac{1}{b-a} & \text{si } a < x < b \\ 0 & \text{sinon.} \end{cases}$$

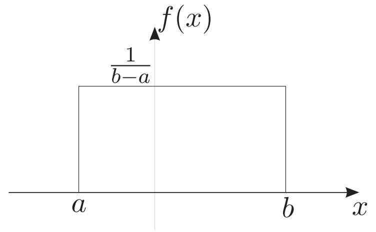


Figure A.1 – Densité de probabilité d'une loi uniforme sur l'intervalle $[a, b]$.

Fonction de répartition : La fonction de répartition F_X de X est la suivante,

$$F_X : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \to \int_{-\infty}^{x} f(t)dt = \begin{cases} 0 & \text{si} & x \leq a \\ \frac{x-a}{b-a} & \text{si} & a < x < b \\ 1 & \text{si} & x \geq b. \end{cases} \end{cases}$$

Espérance et variance : L'espérance et la variance de X sont

$$E(X) = \frac{a+b}{2},$$
$$\text{Var}(X) = \frac{(b-a)^2}{12}.$$

Exemple A.55. *Considérons l'expérience consistant à choisir, au hasard, un réel X dans un intervalle $[a, b]$. La probabilité que X soit dans un sous-intervalle de l'intervalle $[a, b]$ est proportionnelle à la longueur du sous-intervalle. Ainsi définie, X est une variable aléatoire suivant la loi uniforme sur $[a, b]$.*

A.2.3.2 La loi gaussienne

Cette loi, encore appelée loi normale, est fondamentale en théorie des probabilités.

Définition A.56 (Loi gaussienne centrée réduite)**.** On dit qu'une variable aléatoire X suit la loi gaussienne (ou loi normale) centrée réduite si elle admet pour densité la fonction

$$f : x \to \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Fonction de répartition : La fonction de répartition Φ de X est la suivante,

$$\Phi : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \rightarrow & \exp -\frac{x^2}{2}. \end{array} \right.$$

Espérance et variance : Une variable aléatoire de loi gaussienne centrée réduite a une espérance nulle et une variance égale à 1.

Définition A.57 (Loi gaussienne)**.** On dit qu'une variable aléatoire X suit la loi gaussienne (ou loi normale) de paramètres $m_X \in \mathbb{R}$ et $\sigma_X \in \mathbb{R}_+^*$ si elle admet pour densité la fonction (voir figure A.2)

$$f : x \rightarrow \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left(-\frac{(x-m_X)^2}{2\sigma_X^2}\right).$$

On peut vérifier facilement que $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ et donc, qu'ainsi définie, f est bien une densité.

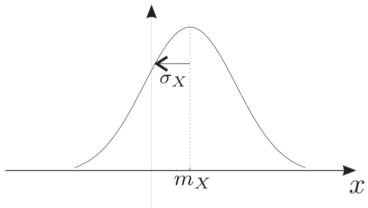


Figure A.2 – Densité de probabilité d'une loi gaussienne.

Fonction de répartition : La fonction de répartition F de la loi gaussienne de paramètres m et σ peut s'écrire, en fonction de la fonction de répartition Φ de la loi gaussienne centrée réduite,

$$F : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \rightarrow & \Phi\left(\frac{x-m}{\sigma}\right). \end{array} \right.$$

Espérance et variance : L'espérance et la variance d'une variable aléatoire X de loi gaussienne de paramètres m_X et σ_X sont

$$E(X) = m_X,$$
$$\text{Var}(X) = \sigma_X^2.$$

Exemple A.58. *Lorsqu'on mesure expérimentalement une grandeur physique dont la valeur exacte est m, on considère souvent que le résultat X de cette mesure est une variable aléatoire suivant la loi gaussienne d'espérance m. Sa variance σ^2 est d'autant plus petite que les instruments de mesure dont on dispose sont performants.*

Remarque A.59. *Les moments d'ordre supérieur à deux d'une variable aléatoire X de loi gaussienne sont des fonctions de sa moyenne m_X et de sa variance σ_X^2.*

A.3 Couple de variables aléatoires

Nous nous intéressons maintenant à un couple de variables aléatoires. Cette partie sera l'occasion d'introduire la notion d'indépendance de deux variables aléatoires.

A.3.1 Espace produit

Soient X_1 et X_2 deux variables aléatoires sur $(\Omega, \mathcal{A}, P)$ à valeurs respectives dans $(E_1, \mathcal{E}_1)$ et $(E_2, \mathcal{E}_2)$. Le couple (X_1, X_2) est défini sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$ où $\mathcal{E}_1 \otimes \mathcal{E}_2$ est la tribu produit sur $E_1 \times E_2$, engendrée par $G = \{A_1 \times A_2, A_1 \in \mathcal{E}_1, A_2 \in \mathcal{E}_2\}$.

Proposition A.60. *Soient X_1 et X_2 deux applications définies sur $(\Omega, \mathcal{A})$ à valeurs respectives dans $(E_1, \mathcal{E}_1)$ et $(E_2, \mathcal{E}_2)$ et $X = (X_1, X_2)$. L'application*

$$X : (\Omega, \mathcal{A}, P) \to (E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$$

est mesurable si et seulement si X_1 et X_2 sont mesurables.

Démonstration. ($\Longleftarrow$) Supposons X_1 et X_2 mesurables. On a :

$$\forall A_1 \in \mathcal{E}_1, \forall A_2 \in \mathcal{E}_2, X_1^{-1}(A_1) \in \mathcal{A}, X_2^{-1}(A_2) \in \mathcal{A}.$$

Pour tout $C \in G = \{A_1 \times A_2, A_1 \in \mathcal{E}_1, A_2 \in \mathcal{E}_2\}$, on a $C = A_1 \times A_2$ avec $A_1 \in \mathcal{E}_1$ et $A_2 \in \mathcal{E}_2$. On a donc

$$\begin{aligned} X^{-1}(C) &= \{\omega \in \Omega : X(\omega) \in C\} \\ &= \{\omega \in \Omega : (X_1(\omega), X_2(\omega)) \in A_1 \times A_2\} \\ &= X_1^{-1}(A_1) \cap X_2^{-1}(A_2) \in \mathcal{A}. \end{aligned}$$

Cette égalité est vraie pour tout $C \in G$. Or, par définition, G engendre $\mathcal{E}_1 \otimes \mathcal{E}_2$. Donc X est mesurable.
($\Longrightarrow$) Réciproquement, si X est mesurable alors,
$\forall A_1 \in \mathcal{E}_1, X_1^{-1}(A_1) = X^{-1}(A_1, E_2) \in \mathcal{A}$, donc X_1 est mesurable,
et, $\forall A_2 \in \mathcal{E}_2, X_2^{-1}(A_2) = X^{-1}(E_1, A_2) \in \mathcal{A}$ donc X_2 est mesurable. □

Désormais, pour deux variables aléatoires X_1 et X_2 définies sur $(\Omega, \mathcal{A}, P)$ à valeurs respectives dans $(E_1, \mathcal{E}_1)$ et $(E_2, \mathcal{E}_2)$ et de lois respectives P_{X_1} et P_{X_2}, on parlera de la variable aléatoire $X = (X_1, X_2)$, de sa loi P_{X_1,X_2}, définie sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $(E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$. Pour $A_1 \in \mathcal{E}_1$ et $A_2 \in \mathcal{E}_2$,

$$\begin{aligned} P_{X_1,X_2}(A_1, A_2) &\triangleq \left(X_1^{-1}(A_1) \cap X_1^{-1} A_2)\right) \\ &= P\left(\{\omega, X_1(\omega) \in A_1, X_2(\omega) \in A_2\}\right). \end{aligned}$$

On pourra également parler, si elle existe, de la densité de probabilité f_{X_1,X_2} du couple $X = (X_1, X_2)$, qui sera dans ce cas une fonction à deux variables.

A.3.2 Loi conjointe, lois marginales

Soient X_1 et X_2 deux variables aléatoires à valeurs respectives dans $(E_1, \mathcal{E}_1)$ et $(E_2, \mathcal{E}_2)$ et $X = (X_1, X_2)$.

Définition A.61 (Loi conjointe)**.** La loi P_{X_1,X_2} de la variable aléatoire $X = (X_1, X_2)$ est appelée loi conjointe du couple $X = (X_1, X_2)$.

Définition A.62 (Lois marginales)**.** Les lois P_{X_1} et P_{X_2} sont appelées lois marginales du couple $X = (X_1, X_2)$.

De façon analogue, on parlera désormais de la densité de probabilité conjointe f_{X_1,X_2} et des densités de probabilité marginales f_{X_1} et f_{X_2}.

Connaissant la loi conjointe P_{X_1,X_2} du couple, on peut en déduire les lois marginales P_{X_1} et P_{X_2}. En effet,

$$\forall A_1 \in \mathcal{E}_1, P_{X_1}(A_1) = P_{X_1,X_2}(A_1, E_2),$$
$$\forall A_2 \in \mathcal{E}_2, P_{X2}(A_2) = P_{X_1,X_2}(E_1, A_2).$$

En revanche, comme le montre l'exemple suivant, la réciproque est fausse.

Exemple A.63. *Soient X_1 et X_2 deux variables aléatoires réelles. Supposons que $P_{X_1} = P_{X_2} = P$, avec*

$$\forall i \in \{1, 2\}, P(X_i = 1) = P(X_i = 0) = \frac{1}{2}.$$

Nous pouvons trouver une infinité de lois conjointes possédant ces lois marginales. En effet, pour tout $\alpha \in [0, \frac{1}{2}]$, les lois P_{X_1,X_2}, définies par

$$P_{X_1,X_2}(0, 0) = P_{X_1,X_2}(1, 1) = \alpha,$$
$$P_{X_1,X_2}(0, 1) = P_{X_1,X_2}(1, 0) = \frac{1}{2} - \alpha,$$

admettent chacune P comme lois marginales.

Nous verrons qu'il existe un cas dans lequel les lois marginales définissent entièrement la loi conjointe : il s'agit de celui où les variables X_1 et X_2 sont indépendantes. Nous allons, dans le paragraphe suivant, introduire la notion d'indépendance de deux variables aléatoires.

A.3.3 Variables aléatoires indépendantes

Avant de définir l'indépendance de deux variables aléatoires, nous avons besoin de la notion de *tribu engendrée par une variable aléatoire*.

Proposition A.64. *Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $(E, \mathcal{E})$. L'ensemble $\mathcal{A}_X$, défini par*

$$\mathcal{A}_X = \left\{X^{-1}(A), A \in \mathcal{E}\right\},$$

est une tribu de Ω. On l'appelle tribu engendrée par X.

Démonstration. Vérifions que $\mathcal{A}_X$ est une tribu de Ω.
i) $\Omega = X^{-1}(E)$ donc $\Omega \in \mathcal{A}_X$.
ii) Si $A \in \mathcal{A}$ alors, $A^c \in \mathcal{A}$. Donc $\left(X^{-1}(A)\right)^c = X^{-1}(A^c) \in \mathcal{A}_X$. $\mathcal{A}_X$ est donc stable par passage au complémentaire.
iii) Enfin, on démontre aisément que $\mathcal{A}_X$ est stable par union dénombrable en utilisant le fait que

$$\bigcup_n X^{-1}(B_n) = X^{-1}\left(\bigcup_n B_n\right).$$

□

Nous pouvons maintenant définir l'indépendance de deux variables aléatoires [73][4].

Définition A.65 (Variables aléatoires indépendantes)**.** Deux variables aléatoires X_1 et X_2, à valeurs respectives dans $(E_1, \mathcal{E}_1)$ et $(E_2, \mathcal{E}_2)$, sont indépendantes si les tribus $\mathcal{A}_{X_1}$ et $\mathcal{A}_{X_2}$ engendrées par ces variables aléatoires sont indépendantes.

Théorème A.66 (Conditions d'indépendance)**.** *Soient X_1 et X_2 deux variables aléatoires sur $(\Omega, \mathcal{A}, P)$ à valeurs respectives dans $(E_1, \mathcal{E}_1)$ et $(E_2, \mathcal{E}_2)$. X_1 et X_2 sont indépendantes si et seulement si une des deux conditions équivalentes suivantes est vérifiée.*
i) Pour tout couple f, g de fonctions mesurables :

$$\begin{aligned} f: &\quad (E_1, \mathcal{E}) \to (E_1, \mathcal{E}), \\ g: &\quad (E_2, \mathcal{E}) \to (E_2, \mathcal{E}), \end{aligned}$$

$f \circ X_1$ et $g \circ X_2$ sont deux variables aléatoires indépendantes à valeurs respectivement dans $(E_1, \mathcal{E})$ et $(E_2, \mathcal{E})$.
ii) La loi de probabilité conjointe P_{X_1,X_2} des variables aléatoires X_1 et X_2 est égale au produit des lois marginales P_{X_1} et P_{X_2} des variables aléatoires X_1 et X_2, c'est-à-dire :

$$P_{X_1,X_2} = P_{X_1} P_{X_2}.$$

Démonstration. i)($\Longleftarrow$) Il suffit de prendre l'identité pour les fonctions f et g et on a directement le résultat.
($\Longrightarrow$) Réciproquement, supposons que les variables aléatoires X_1 et X_2 soient indépendantes. Soit f, g deux fonctions mesurables. On a, pour $A_1 \in \mathcal{E}_1$ et $A_2 \in \mathcal{E}_2$,

$$P\Big((f \circ X_1)^{-1}(A_1) \cap (g \circ X_2)^{-1}(A_2)\Big) = P\left({X_1}^{-1}\left(f^{-1}(A_1)\right) \cap {X_2}^{-1}\left(g^{-1}(A_2)\right)\right),$$

avec $f^{-1}(A_1) \in \mathcal{E}_1$ et $g^{-1}(A_2) \in \mathcal{E}_2$. D'après l'indépendance des variables aléatoires X_1 et X_2, on peut écrire

$$\begin{aligned} P\Big((f \circ X_1)^{-1}(A_1) \cap (g \circ X_2)^{-1}(A_2)\Big) &= P\left({X_1}^{-1}\left(f^{-1}(A_1)\right)\right) P\left({X_2}^{-1}\left(g^{-1}(A_2)\right)\right) \\ &= P\Big((f \circ X_1)^{-1}(A_1)\Big) P\Big((g \circ X_2)^{-1}(A_2)\Big). \end{aligned}$$

Ainsi, les variables aléatoires $f \circ X_1$ et $g \circ X_2$ sont bien indépendantes.
ii) ($\Longrightarrow$) Soient X_1 et X_2 deux variables aléatoires indépendantes. Pour tout $A_1 \in \mathcal{E}_1$ et pour tout $A_2 \in \mathcal{E}_2$, on peut écrire

$$\begin{aligned} P_{X_1,X_2}(A_1, A_2) &= P\left(X^{-1}(A_1, A_2)\right) \text{ où } X = (X_1, X_2) \\ &= P\left({X_1}^{-1}(A_1) \cap {X_2}^{-1}(A_2)\right) \\ &= P\left({X_1}^{-1}(A_1)\right) P\left({X_2}^{-1}(A_2)\right) \\ &= P_{X_1}(A_1) P_{X_2}(A_2). \end{aligned}$$

Or, $\mathcal{E}_1 \otimes \mathcal{E}_2$ est la tribu engendrée par $G = \{A_1 \times A_2, A_1 \in \mathcal{E}_1, A_2 \in \mathcal{E}_2\}$ donc, $P_{X_1,X_2} = P_{X_1} P_{X_2}$.

($\Longleftarrow$) Supposons maintenant que $P_{X_1,X_2} = P_{X_1}P_{X_2}$. La réciproque découle de la définition d'une loi de probabilité. En effet, pour $A_1 \in \mathcal{E}_1$ et $A_2 \in \mathcal{E}_2$, on a

$$\begin{aligned} P\left(X_1^{-1}(A_1) \cap X_2^{-1}(A_2)\right) &\triangleq P_{X_1,X_2}(A_1, A_2) \\ &= P_{X_1}(A_1)\, P_{X_2}(A_2) \\ &\triangleq P(X_1^{-1}(A_1))P(X_2^{-1}(A_2)). \end{aligned}$$

On a donc bien l'indépendance des tribus engendrées par X_1 et X_2 et, donc, par définition, l'indépendance des variables aléatoires X_1 et X_2. $\square$

Dans le cas où les deux variables aléatoires admettent une densité, nous obtenons une nouvelle caractérisation de l'indépendance de deux variables aléatoires.

Corollaire A.67. *Soient X_1 et X_2 deux variables aléatoires de densités respectives f_{X_1} et f_{X_2}. On suppose de plus que le couple (X_1, X_2) admet une densité de probabilité f_{X_1,X_2}. Alors, X_1 et X_2 sont indépendantes si et seulement si la densité conjointe f_{X_1,X_2} est égale au produit des densités marginales f_{X_i} des variables aléatoires X_1 et X_2, c'est-à-dire :*

$$f_{X_1,X_2} = f_{X_1} f_{X_2}.$$

Remarque A.68. *D'après la condition i) du théorème A.66, si deux variables X_1 et X_2 sont indépendantes et intégrables et si X_1X_2 est intégrable, alors*

$$E(X_1X_2) = E(X_1)\, E(X_2)\,. \tag{A.3}$$

Il s'ensuit que $\Gamma_{X_1,X_2} = 0$. Mais attention, ceci n'est pas une condition suffisante pour assurer l'indépendance (voir le contre-exemple suivant), excepté dans le cas gaussien que l'on étudiera ultérieurement.

Exemple A.69 (Contre exemple)**.** *Soit une variable $X = (X_1, X_2)$ de loi P_X définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ par*

$$P_X(1,0) = P_X(-1,0) = P_X(0,1) = P_X(0,-1) = \frac{1}{4}.$$

Notons qu'alors les variables X_1 et X_2 suivent la même loi P_{X_i} définie par :

$$P_{X_i}(-1) = P_{X_i}(1) = \frac{1}{4} \text{ et } P_{X_i}(0) = \frac{1}{2}.$$

On a :

$$E(X_1) = E(X_2) = 0 \text{ et } \Gamma_{X_1,X_2} = E(X_1X_2) = 0.$$

Cependant,

$$\begin{aligned} &P_X(1,0) = \tfrac{1}{4}, \\ &P_{X_1}(1)P_{X_2}(0) = \tfrac{1}{4} \times \tfrac{1}{2} = \tfrac{1}{8} \left(\neq \tfrac{1}{4}\right). \end{aligned}$$

Donc, bien que leur covariance soit nulle, les variables aléatoires X_1 et X_2 ne sont pas indépendantes.

En fait, l'égalité (A.3) traduit une propriété plus faible que l'indépendance, la non-corrélation, ou encore la décorrélation.

Définition A.70 (Corrélation)**.** Deux variables aléatoires X_1 et X_2, intégrables, sont dites non corrélées (ou décorrélées) si $E(X_1X_2) = E(X_1)\, E(X_2)$.

Il semble maintenant intéressant de généraliser cette notion d'indépendance pour plus de deux variables aléatoires. Nous allons donc introduire la notion de vecteur aléatoire.

A.4 Vecteurs aléatoires réels

Nous allons généraliser toutes les notions introduites pour une variable aléatoire (espérance, variance, densité, loi de probabilité...) ou pour un couple de variables aléatoires (loi coinjointe, lois marginales, indépendance...) au cas d'une suite finie de variables aléatoires.

Définition A.71 (Vecteur aléatoire)**.** Un vecteur aléatoire réel est une suite finie $(X_1, \ldots, X_n)^{\mathrm{T}}$ de variables aléatoires réelles.
L'ensemble des vecteurs aléatoires réels à n composantes est noté $\mathcal{V}^n$.

A.4.1 Propriétés d'un vecteur aléatoire

Soit $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ un vecteur aléatoire. Par analogie avec les couples de variables aléatoires, on peut parler de la *fonction de répartition* $F_{\mathbf{X}}$, de la *densité* $f_{\mathbf{X}}$ (si elle existe), de la *loi conjointe* $P_{\mathbf{X}}$ et des *lois marginales* $P_{X_1}, \ldots, P_{X_n}$ du vecteur aléatoire $\mathbf{X}$. Les fonctions $F_{\mathbf{X}}$, $f_{\mathbf{X}}$ et $P_{\mathbf{X}}$ sont alors définies sur $\mathbb{R}^n$.

Nous pouvons également généraliser les notions d'espérance et de variance. Commençons par préciser qu'un vecteur $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ est dit intégrable (respectivement de carré intégrable) si chacune de ses composantes $X_1, \ldots, X_n$ est intégrable (respectivement de carré intégrable).

Définition A.72. [Moments d'un vecteur]Soit $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ un vecteur aléatoire intégrable. On appelle espérance de $\mathbf{X}$, le vecteur

$$E(\mathbf{X}) = (E(X_1), \ldots, E(X_n))^{\mathrm{T}} .$$

De même, on définit le moment d'ordre q $(q > 1)$ du vecteur $\mathbf{X}$ par

$$E(\mathbf{X}^q) = (E(X_1^q), \ldots, E(X_n^q))^{\mathrm{T}} .$$

Rappelons (voir proposition A.47) que l'espérance est linéaire, c'est-à-dire, pour deux vecteurs aléatoires $\mathbf{X}$ et $\mathbf{Y}$ et pour $a, b \in \mathbb{R}$, on a $E(a\mathbf{X}+b\mathbf{Y}) = aE(\mathbf{X}) + bE(\mathbf{Y})$.

Définition A.73 (Matrice de variance-covariance)**.** Lorsque $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ est de carré intégrable, on peut, pour chaque couple (X_i, X_j), parler de la covariance Γ_{X_i,X_j}. On appelle matrice de variance-covariance (ou matrice d'autocorrélation) de $\mathbf{X}$, la matrice carrée, notée $\mathbf{\Gamma_X}$ (ou encore $\mathbf{\Gamma_{X,X}}$), définie par

$$\begin{aligned}\mathbf{\Gamma_X} &= E\left((\mathbf{X} - \mathbf{m_X})(\mathbf{X} - \mathbf{m_X})^{\mathrm{T}}\right) = \left(\Gamma_{X_i,X_j}\right)_{1\le i,j\le n} \\ &= E\left(\mathbf{X}\mathbf{X}^{\mathrm{T}}\right) - \mathbf{m_X^{\mathrm{T}}}\mathbf{m_X},\end{aligned}$$

où $\mathbf{m_X} = E(\mathbf{X})$.

Proposition A.74. *i) Toute matrice d'autocorrélation* $\mathbf{\Gamma_X}$ *est une matrice définie non négative* [4]*. Réciproquement, toute matrice* $\mathbf{M}$ *définie non négative, peut être vue comme la matrice d'autocorrélation d'un vecteur aléatoire* $\mathbf{X}$.
ii) Toute matrice d'autocorrélation $\mathbf{\Gamma_X}$ *est symétrique.*

[4]c'est-à-dire $\forall \mathbf{u} \in \mathbb{R}^n, \mathbf{u}^{\mathrm{T}}\mathbf{\Gamma_X}\mathbf{u} \ge 0$

Démonstration. i) $\forall \mathbf{v} \in \mathbb{R}^n$,

$$\begin{aligned}\mathbf{v}^{\mathrm{T}}\boldsymbol{\Gamma}_{\mathbf{X}}\mathbf{v} &= \mathbf{v}^{\mathrm{T}} E\left(\left(\mathbf{X}-\mathbf{m}_{\mathbf{X}}\right)\left(\mathbf{X}-\mathbf{m}_{\mathbf{X}}\right)^{\mathrm{T}}\right)\mathbf{v} \\ &= E\left(\mathbf{v}^{\mathrm{T}}\left(\mathbf{X}-\mathbf{m}_{\mathbf{X}}\right)\left(\mathbf{X}-\mathbf{m}_{\mathbf{X}}\right)^{\mathrm{T}}\mathbf{v}\right) \\ &= E\left(\left|\mathbf{v}^{\mathrm{T}}\left(\mathbf{X}-\mathbf{m}_{\mathbf{X}}\right)\right|^2\right) \geq 0.\end{aligned}$$

Réciproquement, soit $\mathbf{M}$ définie non négative de valeurs propres $\lambda_i \geq 0$ et de vecteurs propres associés $\mathbf{v}_i$. La décomposition spectrale de $\mathbf{M}$ est la suivante

$$\mathbf{M} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^{\mathrm{T}}. \tag{A.4}$$

Considérons un ensemble de variables aléatoires $X_1, \ldots, X_n$ centrées et décorrélées telles que

$$\forall i, E\left(X_i\right) = 0, E\left(X_i^2\right) = \lambda_i \quad \text{et} \quad \forall i \neq j, E\left(X_i X_j\right) = 0.$$

Soit $\mathbf{X}$ le vecteur défini par

$$\mathbf{X} \triangleq \sum_{i=1}^{n} X_i \mathbf{v}_i.$$

On a $\Gamma_{\mathbf{X}} = E\left(\mathbf{X}\mathbf{X}^{\mathrm{T}}\right) - \mathbf{m}_{\mathbf{X}}^{\mathrm{T}}\mathbf{m}_{\mathbf{X}} = E\left(\mathbf{X}\mathbf{X}^{\mathrm{T}}\right) = E\left(\sum_{i,j=1}^{n} X_i \mathbf{v}_i \mathbf{v}_j^{\mathrm{T}} X_j\right) = \sum_{i,j=1}^{n} E\left(X_i X_j\right)\mathbf{v}_i \mathbf{v}_j^{\mathrm{T}} = \sum_{i=1}^{n} E\left(X_i^2\right)\mathbf{v}_i \mathbf{v}_i^{\mathrm{T}} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^{\mathrm{T}}$. Les matrices $\Gamma_{\mathbf{X}}$ et $\mathbf{M}$ sont donc égales (voir l'égalité (A.4)).

ii) Trivial, puisque $\Gamma_{X_i,X_j} = \Gamma_{X_j,X_i}$. □

A.4.2 Transformation d'un vecteur aléatoire

Soit $\mathbf{X} = \left(X_1, \ldots, X_n\right)^{\mathrm{T}} \in \mathcal{V}^n$ un vecteur aléatoire et soient n fonctions $g_1, \ldots, g_n$ définis sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Considérons le vecteur aléatoire $\mathbf{Y} = \left(Y_1, \ldots, Y_n\right)^{\mathrm{T}}$ défini par

$$\mathbf{Y} = \mathbf{g}\left(\mathbf{X}\right),$$

où $\mathbf{g}$ est la transformation définie par

$$\mathbf{g} : \left\{ \begin{array}{ccc} \mathbb{R}^n & \rightarrow & \mathbb{R}^n \\ \mathbf{x} & \mapsto & \left(g_1\left(\mathbf{x}\right), \ldots, g_n\left(\mathbf{x}\right)\right) \end{array} \right. . \tag{A.5}$$

Dans le cas où ces vecteurs aléatoires admettent des densités de probabilité, il semble intéressant de trouver une relation entre les densités conjointes des vecteurs aléatoires $\mathbf{X}$ et $\mathbf{Y}$.

Proposition A.75. *Soit $\mathbf{g} = \left(g_1, \ldots, g_n\right)$ définie par (A.5), bijective, de classe $\mathcal{C}^1$ telle que $\mathbf{g}^{-1}$ soit également de classe $\mathcal{C}^1$. Si $\mathbf{X}$ admet une densité de probabilité $f_{X_1,\ldots,X_n}$ alors, $\mathbf{Y}$ admet la densité de probabilité $f_{Y_1,\ldots,Y_n}$, donnée par la formule*

$$f_{Y_1,\ldots,Y_n}\left(\mathbf{y}\right) = f_{X_1,\ldots,X_n}\left(\mathbf{g}^{-1}\left(\mathbf{y}\right)\right)\left|J_{\mathbf{g}}^{-1}\left(\mathbf{g}^{-1}\left(\mathbf{y}\right)\right)\right|,$$

où

$$J_{\mathbf{g}}(x_1,\ldots,x_n) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}$$

est la matrice jacobienne de la transformation $\mathbf{g}$ *et* $|J_{\mathbf{g}}(x_1,\ldots,x_n)|$ *est son déterminant.*

Démonstration. Ce résultat est admis. □

Exemple A.76. *Soit* $\mathbf{X} = (X_1, X_2)$ *un vecteur aléatoire et soit* $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ *où* $\mathbf{g}$ *est la transformation définie par*

$$\mathbf{g} : \begin{cases} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x_1, x_2) & \mapsto & (x_1, x_1 + x_2). \end{cases}.$$

Ainsi définie, $\mathbf{g}$ *vérifie bien les hypothèses de la proposition A.75. On a*

$$\mathbf{g}^{-1}(x_1, x_2) = (x_1, x_2 - x_1) \text{ et } J_{\mathbf{g}}(x_1, x_2) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Donc

$$\left|J_{\mathbf{g}}^{-1}(x_1, x_2)\right| = \left|\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right| = 1.$$

Ainsi, pour $\mathbf{Y} = (Y_1, Y_2)$, *on obtient*

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_1+X_2}(y_1, y_2) \\ &= f_{X_1,X_2}(x_1, x_2 - x_1) \times 1. \end{aligned}$$

En intégrant par rapport à la première variable, on peut en déduire la densité marginale $f_{X_1+X_2}$.

Il est clair qu'une variable aléatoire peut être vue comme un vecteur aléatoire dégénéré (dimension 1). Ceci implique que l'ensemble des propositions énoncées ci-dessus reste valable pour des variables aléatoires. Maintenant, nous allons étendre la notion d'indépendance de deux variables aléatoires au vecteur aléatoire.

A.4.3 Vecteur aléatoire à composantes indépendantes

Définition A.77 (Vecteur aléatoire à composantes indépendantes)**.** Un vecteur aléatoire $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ est dit, à composantes (mutuellement) indépendantes si les variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes.

Si $\mathbf{X} = (X_1, \ldots, X_n)^T$ est un vecteur aléatoire à composantes indépendantes, noté $I_v(X_1, \ldots, X_n)$, alors ses composantes X_i sont indépendantes deux à deux. La réciproque est fausse, comme le montre l'exemple suivant :

Exemple A.78. *Soient* X_1 *et* X_2 *deux variables aléatoires indépendantes de même loi définie, pour* $i \in \{1, 2\}$, *par* $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$.
Considérons le vecteur aléatoire $\mathbf{X} = (X_1, X_2, X_3)^T$ *avec* $X_3 = X_1 X_2$.
X_3 *suit la même loi que* X_1 *et* X_2. *On peut vérifier aisément que les variables aléatoires* X_1, X_2, X_3

sont deux à deux indépendantes.
Cependant, elle ne sont pas mutuellement indépendantes. En effet,

$$\begin{array}{l} P\left((X_1,X_2,X_3)=(1,1,1)\right)=\frac{1}{4}, \\ P\left(X_i=1\right)=\frac{1}{2}. \end{array}$$

On a : $P\left((X_1,X_2,X_3)=(1,1,1)\right) \neq \prod_{i=1}^{3} P\left(X_i=1\right)$. *Donc, le vecteur aléatoire* $\mathbf{X}=(X_1,X_2,X_3)^T$ *n'est pas à composantes indépendantes.*

Nous pouvons maintenant énoncer plusieurs résultats qui généralisent ceux obtenus dans le cas d'un couple de variables aléatoires.

Proposition A.79. *Un vecteur aléatoire* $\mathbf{X}=(X_1,\ldots,X_n)^T$ *est à composantes indépendantes si et seulement si la loi de probabilité conjointe* $P_{\mathbf{X}}$ *est égale au produit des* n *lois de probabilité marginales :*

$$P_{\mathbf{X}}=\prod_{i=1}^{n} P_{Xi}.$$

Démonstration. Il s'agit d'une généralisation du théorème A.66 ii). □

Nous avons, dans le cas où les variables aléatoires admettent des densités, une autre caractérisation de l'indépendance.

Corollaire A.80. *Soit* $\mathbf{X}=(X_1,\ldots,X_n)^T$ *un vecteur aléatoire dont chaque composante* X_i *admet une densité de probabilité* f_{X_i}. *On suppose de plus que le vecteur* $\mathbf{X}$ *admet une densité de probabilité* f_X. *Alors,* $\mathbf{X}=(X_1,\ldots,X_n)^T$ *est à composantes indépendantes si et seulement si la densité de probabilité conjointe* $f_{\mathbf{X}}$ *est égale au produit des densités de probabilité marginales* f_{Xi}*, c'est-à-dire*

$$f_{\mathbf{X}}=\prod_{i=1}^{n} f_{Xi}. \tag{A.6}$$

Proposition A.81. *Un vecteur aléatoire* $\mathbf{X}=(X_1,\ldots,X_n)^T$ *est à composantes indépendantes si et seulement si, pour toutes fonctions* f_i *mesurables, le vecteur aléatoire (* $f_1\circ X_1,\ldots,f_n\circ X_n$*) est à composantes indépendantes.*

Démonstration. Il s'agit d'une généralisation du théorème A.66 i). □

On en déduit le résultat suivant, qui généralise la remarque A.68.

Corollaire A.82. *Si un vecteur aléatoire* $X=(X_1,\ldots,X_n)^T$ *est à composantes indépendantes et intégrables alors,*

$$E(X_1\ldots X_n)=E(X_1)\ldots E(X_n).$$

On dit que les variables aléatoires X_i *sont décorrélées.*

Attention, en général, cette propriété ne caractérise pas l'indépendance (voir exemple A.69). Nous verrons ultérieurement que, dans le cas gaussien, c'est une condition suffisante.

A.4.4 Le vecteur aléatoire gaussien

Définition A.83 (Vecteur aléatoire gaussien)**.** Soit $\mathcal{L}_n$ l'ensemble des fonctions linéaires de $\mathbb{R}^n$ à valeur dans $\mathbb{R}$. Un vecteur aléatoire $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}} \in \mathcal{V}^n$ est gaussien si pour toute fonction $f \in \mathcal{L}_n$, la variable aléatoire $f(\mathbf{X})$ est gaussienne. En d'autres termes, $\mathbf{X}$ est gaussien, noté $\mathcal{G}_v(X_1 \ldots X_n)$, si

$$\forall (a_1, \ldots, a_n) \in \mathbb{R}^n, \ \sum_{i=1}^{n} a_i X_i$$

est une variable aléatoire gaussienne.

Si un vecteur $\mathbf{X}$ est gaussien alors ses composantes sont gaussiennes, mais la réciproque est fausse comme le montre l'exemple suivant.

Exemple A.84. *Soit X_1 une variable aléatoire gaussienne centrée réduite ($E(X_1) = 0$ et $E\left(X_1^2\right) = 1$) et soit X_2 la variable aléatoire définie par*

$$X_2 = \begin{cases} X_1 & si \ |X| \leq a \\ -X_1 & si \ |X| > a, \end{cases}$$

pour $a > 0$. On vérifie aisément que X_2 est une variable aléatoire gaussienne centrée réduite (par symétrie de la densité de X_1). Cependant, la variable aléatoire $X_1 + X_2$ n'est pas gaussienne (car elle est portée par l'intervalle $[-a, a]$). Ainsi, le vecteur aléatoire (X_1, X_2) n'est pas gaussien bien que ses composantes le soient.

Proposition A.85. *Un vecteur aléatoire $\mathbf{X} = (X_1, \ldots, X_n)^T \in \mathcal{V}^n$ est gaussien si et seulement si sa densité de probabilité conjointe est de la forme*

$$f_{X_1,\ldots,X_n}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{m}{2}} |\mathbf{\Gamma_X}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m_X})^T \mathbf{\Gamma_X}^{-1} (\mathbf{x} - \mathbf{m_X})\right), \tag{A.7}$$

où $\mathbf{m_X}$ et $\mathbf{\Gamma_X}$ sont, respectivement, l'espérance et la matrice de variance-covariance du vecteur $\mathbf{X}$.

Démonstration. Admis. □

Le vecteur gaussien est entièrement déterminé par ses moments d'ordre un et deux (son espérance et sa matrice de variance-covariance).

Proposition A.86. *Un vecteur aléatoire $\mathbf{X} = (X_1, \ldots, X_n)^T \in \mathcal{V}^n$ gaussien est à composantes indépendantes si et seulement si*

$$\forall i \neq j, \ \Gamma_{X_i, X_j} = 0,$$

autrement dit, sa matrice de variance-covariance $\Gamma_{\mathbf{X}}$ est diagonale.

Démonstration. D'après le corollaire A.82, on sait que si $\mathbf{X}$ est à composantes indépendantes alors, $\Gamma_{\mathbf{X}}$ est diagonale.

Réciproquement, si $\Gamma_{\mathbf{X}}$ est diagonale, alors la densité de probabilité (A.7) se factorise comme le produit des densités marginales $f_{X_i}(x_i) = \frac{1}{\sigma_{X_i}\sqrt{2\pi}} \exp\left(-\frac{(x_i - m_{X_i})^2}{2\sigma_{X_i}{}^2}\right)$, ce qui d'après le corollaire A.80 implique que $\mathbf{X}$ est à composantes indépendantes. □

Pour des variables aléatoires gaussiennes, l'indépendance se réduit donc à la décorrélation.

A.4.5 Couple de vecteurs aléatoires

Pour terminer, nous allons nous intéresser à un couple de vecteurs aléatoires. Considérons pour cela deux vecteurs aléatoires $\mathbf{X}$ et $\mathbf{Y}$ appartenant à $\mathcal{V}^n$.

Définition A.87 (Matrice de corrélation)**.** Soient $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ et $\mathbf{Y} = (Y_1, \ldots, Y_n)^{\mathrm{T}}$ deux vecteurs aléatoires de carré intégrable. On appelle matrice de corrélation (ou matrice d'intercorrélation) de $\mathbf{X}$ et $\mathbf{Y}$, la matrice notée $\mathbf{\Gamma}_{\mathbf{X},\mathbf{Y}}$ définie par

$$\begin{aligned}\mathbf{\Gamma}_{\mathbf{X},\mathbf{Y}} &= E\left((\mathbf{X}-\mathbf{m}_{\mathbf{X}})(\mathbf{Y}-\mathbf{m}_{\mathbf{Y}})^{\mathrm{T}}\right) = E\left(\mathbf{X}\mathbf{Y}^{\mathrm{T}}\right) - \mathbf{m}_{\mathbf{X}}\mathbf{m}_{\mathbf{Y}}^{\mathrm{T}}, \\ &= \left(\Gamma_{X_i,Y_j}\right)_{1\leq i,j\leq n} = \left(E\left(X_iY_j\right) - m_{X_i}m_{Y_j}\right)_{1\leq i,j\leq n}.\end{aligned}$$

Si $\mathbf{\Gamma}_{\mathbf{X},\mathbf{Y}} = \mathbf{0}$, on dit que les vecteurs aléatoires $\mathbf{X}$ et $\mathbf{Y}$ sont décorrélés.

Maintenant, définissons l'indépendance de deux vecteurs aléatoires l'un par rapport à l'autre.

Définition A.88 (Indépendance mutuelle de vecteurs aléatoires)**.** Soient $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}} \in \mathcal{V}^n$ et $\mathbf{Y} = (Y_1, \ldots, Y_n)^{\mathrm{T}} \in \mathcal{V}^n$ deux vecteurs aléatoires. $\mathbf{X}$ et $\mathbf{Y}$ sont mutuellement indépendants si et seulement si $\forall g, h : \mathbb{R}^n \to \mathbb{R}$ mesurables, les variables aléatoires $g(\mathbf{X})$ et $h(\mathbf{Y})$ sont indépendantes.
Dans le cas où ces variables admettent des densités, cette condition se traduit par

$$f_{g(\mathbf{X}),h(\mathbf{Y})} = f_{g(\mathbf{X})}f_{h(\mathbf{Y})}. \tag{A.8}$$

Remarque A.89. *Nous avons volontairement choisi cette définition, moins classique que celle proposée dans la littérature, pour l'indépendance. Notre choix est justifié dans l'annexe B sur les signaux aléatoires.*

A.5 conclusion

Dans cette annexe, nous avons présenté les principales notions de la théorie des probabilités nécessaires à la compréhension du mémoire de thèse.
Les définitions de variables et vecteurs aléatoires sont présentées ainsi que les principales grandeurs probabiliste associées telles la moyenne, la variance, la fonction d'autocorrélation, etc. Les définitions fondamentales d'indépendance et de gaussianité sont illustrées sur de nombreux exemples.
Dans l'annexe suivante, nous allons nous intéresser au cas des signaux aléatoires réels et notamment proposer une définition originale de l'indépendance et la gaussianité.

ANNEXE B

Signaux aléatoires

Dans cette annexe, réalisé à partir des ouvrages [58][77][35] et [41] nous introduisons tout d'abord la notion fondamentale de signal aléatoire. Puis, nous nous intéressons aux propriétés de base des signaux, telles la stationnarité et l'ergodisme. Plus particulièrement, nous insistons sur la dérivation de signaux aléatoires, l'originalité de cette thèse étant pour partie d'en exploiter les caractéristiques. Précisons que les signaux utilisés dans cette thèse sont des signaux stationnaires, ergodiques et dérivables, d'où l'intérêt d'étudier de façon précise le sens de ces notions.
Ensuite, nous définissons deux concepts fondamentaux pour les signaux aléatoires : l'indépendance et la gaussianité. Précisons d'ores et déjà que la définition d'indépendance que nous proposons n'est pas classique et mérite une attention toute particulière. C'est pourquoi, dans un souci pédagogique, nous l'expliquerons minutieusement et insisterons sur son originalité.
Enfin, nous terminons cette annexe par des résultats élémentaires sur les systèmes entrée-sortie. Nous évoquerons en particulier le problème suivant : connaissant les propriétés statistiques des entrées, comment en déduire celles des sorties, et inversement.

B.1 Introduction

Les signaux aléatoires sont décrits par des variables aléatoires généralement dépendantes du temps. Il existe deux définitions possibles des signaux aléatoires, correspondant à deux visions différentes :

- La première consiste à voir un signal aléatoire comme une fonction de Ω à valeurs dans $\mathbb{R}^{\mathbb{R}}$ de la même manière qu'une variable aléatoire est une fonction de Ω à valeurs dans $\mathbb{R}$.
- La seconde consiste à voir un signal aléatoire comme un ensemble infini de variables aléatoires, autrement dit, comme un vecteur aléatoire de dimension infinie.

Plus formellement, voici les deux définitions correspondantes :

Définition B.1 (Signal aléatoire (1))**.** Un signal aléatoire $u\,(.)$ est une fonction de l'espace des événements Ω à valeurs dans $\mathbb{R}^{\mathbb{R}}$, qui a un événement élémentaire ω associe une réalisation $u\,(\omega,.)$ de $u\,(.)$. Cette réalisation est une fonction du temps. Plus formellement, on peut écrire

$$u\,(.) : \left\{ \begin{array}{ccl} \Omega & \rightarrow & \mathbb{R}^{\mathbb{R}} \\ \omega & \rightarrow & u\,(\omega,.) : \left\{ \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ t & \mapsto & u\,(\omega,t)\,. \end{array} \right. \end{array} \right.$$

Définition B.2 (Signal aléatoire (2))**.** Un signal aléatoire $u\,(.)$ est une collection de variables aléatoires indexées

$$u(.) = \{u\,(t)\,, t \in \mathbb{R}\},$$

où pour chaque instant t, $u(t)$ est une variable aléatoire.

Ces deux définitions sont équivalentes, et on utilisera l'une ou l'autre selon le problème en étude. On note $\mathfrak{S}$ l'ensemble des signaux aléatoires. Par la suite, afin de ne pas alourdir les notations, on notera abusivement

$$u(.) : \begin{cases} \Omega & \rightarrow & \mathbb{R}^{\mathbb{R}} \\ \omega & \rightarrow & u(\omega, t) \end{cases} \in \mathfrak{S}.$$

Remarque B.3. *Afin d'éviter toute confusion entre la fonction du temps $u(\omega, t) \in \mathbb{R}^{\mathbb{R}}$ et sa valeur $u(\omega, t) \in \mathbb{R}$ pour un t fixé, on précisera systématiquement l'ensemble d'appartenance.*

Exemple B.4. *La figure B.1 illustre la notion de signal aléatoire qui, comme nous l'avons évoqué, peut-être vu de deux façons différentes :*
- Une collection de variables aléatoires $u(t_1), u(t_2)$, etc...
- Une fonction qui à un événement ω_1 associe une réalisation $u(\omega_1, t)$, qui à un événement ω_2 associe une réalisation $u(\omega_2, t)$, etc...

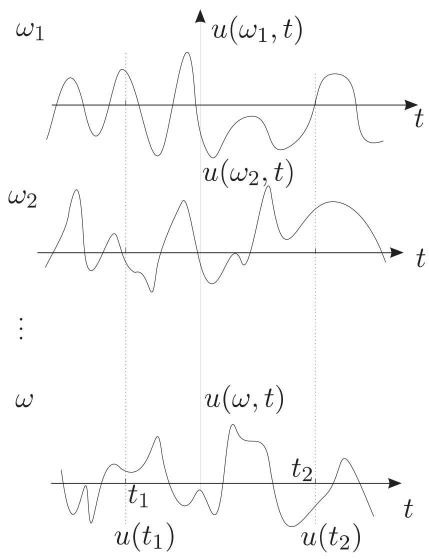


Figure B.1 – Illustration d'un signal aléatoire $u(.)$.

Notons que, lorsque le temps est échantillonné, le signal aléatoire s'exprime comme un ensemble de N variables aléatoires (où N est le nombre d'échantillons), c'est-à-dire un vecteur aléatoire. Les vecteurs aléatoires ont été décrits et étudiés de façon détaillée et complète dans l'annexe A, il semble naturel de chercher à étendre le vocabulaire et les propriétés vus pour ces derniers aux signaux aléatoires. C'est ce que nous proposons dans la suite.

B.1.1 Caractérisation statistique

Définition B.5 (Caractérisation d'ordre N)**.** La caractérisation d'ordre N d'un signal aléatoire consiste à spécifier la densité de probabilité conjointe de tous les ensembles $u(t_1), u(t_2), \ldots, u(t_N)$,

$$\forall t_1, t_2, \ldots, t_N \in \mathbb{R}, \; f_{u(t_1)u(t_2)\ldots u(t_N)}(u_{t_1}, u_{t_2}, \ldots, u_{t_N}).$$

Autrement dit, le signal aléatoire est vu comme une collection de vecteurs aléatoires de dimension N.

Définition B.6 (Caractérisation complète)**.** La caractérisation complète d'un signal aléatoire consiste à spécifier sa caractérisation d'ordre N pour toutes les valeurs N finies ($N < +\infty$).

Dans le cadre de cette thèse, nous utiliserons principalement une caractérisation (partielle) d'ordre deux, c'est-à-dire que nous nous intéresserons à l'ensemble des vecteurs aléatoires :

$$\{u = (u(t_1), u(t_2)) \, | \forall t_1, t_2 \in \mathbb{R}\}.$$

Dans le reste de cette annexe, nous supposerons que toutes les grandeurs statistiques introduites existent. Pour cela, nous resterons dans le cadre de signaux aléatoires "gentils", ceci étant généralement le cas dans la pratique. Par exemple, lorsque nous parlerons de moyenne, nous considérerons un signal composé d'une collection de variables aléatoires intégrables. Bien entendu, cette remarque n'ote rien à la généralité des résultats qui seront présentés.

B.1.2 Moyenne

On peut définir deux types de moyennes :
- les *moyennes statistiques* (ou d'ensemble), effectuées sur l'espace des réalisations sous-jacent à la définition des variables aléatoires qui constituent le signal.
- les *moyennes temporelles*, effectuées sur les fonctions du temps qui sont les réalisations du processus.

Définissons formellement ces moyennes.

Définition B.7 (Moyenne d'ensemble)**.** On appelle moyenne d'ensemble $m_u(t)$ du signal aléatoire $u(.)$ la valeur moyenne de chacune des variables aléatoires qui constituent $u(.)$, c'est-à-dire :

$$m_u(t) = E(u(t)) = \int_{-\infty}^{+\infty} u(t) \, f_{u(t)}(u_t) \, du_t.$$

$m_u(t)$ est une fonction déterministe dépendante du temps t.

Définition B.8 (Moyenne temporelle)**.** On appelle moyenne temporelle $m_u(\omega)$ du signal aléatoire $u(.)$ la valeur moyenne de chacune de ses réalisations :

$$m_u(\omega) = \lim_{T \to +\infty} \frac{1}{2T} \int_{-T}^{T} u(t, \omega) \, dt.$$

$m_u(\omega)$ est une variable aléatoire.

B.1.3 Vecteur de signaux aléatoires

Dans ce paragraphe, par analogie avec les vecteurs (de variables) aléatoires, nous considérons les vecteurs de n signaux aléatoires $\mathbf{u}(.) = (u_1(.), \ldots, u_n(.))^{\mathrm{T}}$. A noter qu'un signal aléatoire peut être vu comme un vecteur (de signaux aléatoire) de dimension un. Ainsi, les définitions suivantes restent valables pour un signal aléatoire.

Définition B.9 (Matrice d'autocorrélation)**.** Soient $\mathbf{u}(.) = (u_1(.), \ldots, u_n(.))^{\mathrm{T}}$ un vecteur de n signaux aléatoires. Considérons deux instants t_1 et t_2, la *matrice d'autocorrélation* $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1, t_2)$ du vecteur $\mathbf{u}(.)$

aux instants t_1 et t_2 est le moment centré d'ordre deux

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2) = E\left((\mathbf{u}(t_1) - E(\mathbf{u}(t_1)))(\mathbf{u}(t_2) - E(\mathbf{u}(t_2)))^{\mathrm{T}}\right), \tag{B.1}$$

$$= E\left(\mathbf{u}(t_1)\mathbf{u}(t_2)^{\mathrm{T}}\right) - E(\mathbf{u}(t_1))\,E(\mathbf{u}(t_2))^{\mathrm{T}}. \tag{B.2}$$

$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2)$ est une fonction déterministe des deux variables réelles t_1 et t_2.

Remarque B.10. *Soit $\mathbf{u}(.) = (u_1(.),\ldots,u_n(.))^T$ un vecteur de n signaux aléatoires, $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2)$ est une matrice carrée, composée des variances et covariances des signaux $u_i(.)$, définie par*

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2) = \begin{pmatrix} \Gamma_{u_1,u_1}(t_1,t_2) & \ldots & \Gamma_{u_1,u_n}(t_1,t_2) \\ \vdots & \ddots & \vdots \\ \Gamma_{u_n,u_1}(t_1,t_2) & \ldots & \Gamma_{u_n,u_n}(t_1,t_2) \end{pmatrix},$$

où $\Gamma_{u_i,u_j}(t_1,t_2)$, $i \neq j$, est appelée fonction d'intercorrélation des signaux u_i et u_j et $\Gamma_{u_i,u_i}(t_1,t_2)$ est appelée fonction d'autocorrélation du signal u_i.

De cette matrice, on peut déduire la variance $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t)$ du vecteur $\mathbf{u}(.)$ à l'instant t, définie par la relation :

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t) = E\left(|\mathbf{u}(t) - E(\mathbf{u}(t))|^2\right).$$

A noter que $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t)$ représente la moyenne quadratique des écarts à la moyenne $E(\mathbf{u}(t))$ des variables aléatoires qui constituent $\mathbf{u}(.)$.

Proposition B.11. *i) Toute matrice d'autocorrélation $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2)$ est définie non négative[1]. Réciproquement, pour toute une matrice définie non négative, on peut toujours trouver un signal d'ordre deux (c'est-à-dire, de valeur de moments d'ordre deux finie), dont la matrice d'autocorrélation est $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2)$.*
ii) Toute matrice d'autocorrélation vérifie (pour des signaux réels) :

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2) = \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_2,t_1)^T.$$

Démonstration. i) Similaire à la preuve de la proposition A.74.
ii) Trivial. □

B.1.4 Couple de vecteurs de signaux aléatoires

Maintenant, intéressons-nous au couple de deux vecteurs de n signaux aléatoires

$$\mathbf{u}(.) = (u_1(.),\ldots,u_n(.))^{\mathrm{T}} \text{ et } \mathbf{y}(.) = (y_1(.),\ldots,y_n(.))^{\mathrm{T}}.$$

Définition B.12 (Matrice d'intercorrélation)**.** Soient deux vecteurs de n signaux aléatoires

$$\mathbf{u}(.) = (u_1(.),\ldots,u_n(.))^{\mathrm{T}} \text{ et } \mathbf{y}(.) = (y_1(.),\ldots,y_n(.))^{\mathrm{T}},$$

[1] c'est-à-dire $\forall \mathbf{v}_i, \mathbf{v}_j, \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{v}_i^{\mathrm{T}} \Gamma_{u,u}(t_i,t_j)\,\mathbf{v}_j \geq 0, \forall t_i,t_j \in \mathbb{R}$

et soient t_1, t_2 deux instants quelconques. La matrice d'intercorrrélation de $\mathbf{u}\left(.\right)$ et $\mathbf{y}\left(.\right)$ aux instants t_1 et t_2 est le moment croisé, centré, d'ordre deux

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{y}}\left(t_1,t_2\right) = E\left(\left(\mathbf{u}\left(t_1\right) - E\left(\mathbf{u}\left(t_1\right)\right)\right)\left(\mathbf{y}\left(t_2\right) - E\left(\mathbf{y}\left(t_2\right)\right)\right)^{\mathrm{T}}\right), \tag{B.3}$$

$$= E\left(\mathbf{u}\left(t_1\right)\mathbf{y}\left(t_2\right)^{\mathrm{T}}\right) - E\left(\mathbf{u}\left(t_1\right)\right)E\left(\mathbf{y}\left(t_2\right)\right)^{\mathrm{T}}. \tag{B.4}$$

$\mathbf{\Gamma}_{\mathbf{u},\mathbf{y}}\left(t_1,t_2\right)$ est une fonction déterministe des deux variables t_1 et t_2.

Proposition B.13. *Soient* $\mathbf{u}\left(.\right) = \left(u_1\left(.\right),\ldots,u_n\left(.\right)\right)^T$ *et* $\mathbf{y}\left(.\right) = \left(y_1\left(.\right),\ldots,y_n\left(.\right)\right)^T$ *deux vecteurs de* n *signaux aléatoires, de matrices d'autocorrélation* $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}\left(t_1,t_2\right)$ *et* $\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(t_1,t_2\right)$*, on a,* $\forall t_1,t_2 \in \mathbb{R}$*,*

$$\mathbf{\Gamma}_{\mathbf{u}+\mathbf{y},\mathbf{u}+\mathbf{y}}\left(t_1,t_2\right) = \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}\left(t_1,t_2\right) + \mathbf{\Gamma}_{\mathbf{u},\mathbf{y}}\left(t_1,t_2\right) + \mathbf{\Gamma}_{\mathbf{y},\mathbf{u}}\left(t_1,t_2\right) + \mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}\left(t_1,t_2\right).$$

Démonstration. Evident par calcul direct. □

Nous allons maintenant nous intéresser à des classes particulières de signaux aléatoires.

B.2 Stationnarité

Dans le langage courant, la notion de stationnarité est utilisée pour décrire un phénomène dont les caractéristiques essentielles ne se modifient pas au cours du temps : on parle de météo stationnaire, d'état stationnaire d'une maladie, etc. . . Le caractère stationnaire d'un signal aléatoire est associé à cette acception du mot, il se définit précisément à l'aide des propriétés d'un signal aléatoire. On dit qu'un signal est stationnaire si ses caractéristiques ne varient pas avec la définition de l'origine du temps, ou encore, si ses caractéristiques statistiques ne varient pas le long du temps. Nous allons définir plusieurs types de stationnarité.

Définition B.14 (Stationnarité au sens strict)**.** Un signal aléatoire $u\left(.\right)$ est stationnaire au sens strict si, pour toute valeur de N, sa caractérisation d'ordre N est invariante par rapport à une translation Δt sur l'axe des temps :

$$f_{u(t_1)\ldots u(t_N)}\left(u_{t_1},\ldots,u_{t_N}\right) = f_{u(t_1+\Delta t)\ldots u(t_N+\Delta t)}\left(u_{t_1+\Delta t},\ldots,u_{t_N+\Delta t}\right).$$

Définition B.15 (Stationnarité à l'ordre deux)**.** Un signal aléatoire $u\left(.\right)$ est stationnaire à l'ordre deux (ou encore stationnaire au sens large) si sa moyenne d'ensemble est constante, c'est-à-dire

$$m_u(t) = m_u,$$

et sa fonction d'autocorrélation ne dépend pas de l'origine du temps.
Comme $\Gamma_u\left(t_1,t_2\right)$ ne dépend que des deux instants du temps t_1 et t_2, cette dernière condition est équivalente à

$$\begin{aligned}\Gamma_{u,u}\left(t_1,t_2\right) &= \Gamma_{u,u}\left(t_1,t_1-\tau\right),\\ &= \Gamma'_{u,u}\left(\tau\right).\end{aligned}$$

avec $\tau = t_1 - t_2$.

La figure B.2 illustre la notion de stationnarité à l'ordre deux. Après un décalage de τ, les propriétés statistiques d'un signal aléatoire à deux instants t_1 et t_2 sont identiques, autrement dit, l'information statistique est conservée.

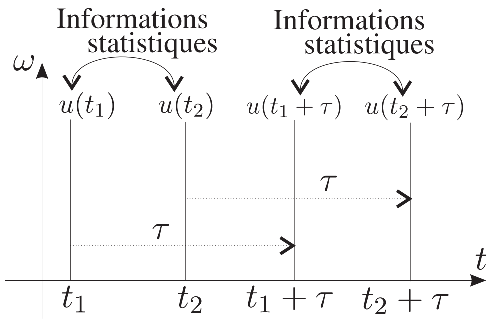


Figure B.2 – Illustration de la stationnarité d'un signal aléatoire $u(.)$.

Par la suite, nous nous intéressons uniquement à la stationnarité à l'ordre deux. On note $\mathcal{S}$ l'ensemble des signaux stationnaires à l'ordre deux.

Remarque B.16 (Signification physique)**.** *La fonction d'autocorrélation est un moyen numérique pour évaluer le caractère aléatoire d'un signal. Ceci est particulièrement clair dans le cas stationnaire. Si la fonction d'autocorrélation est identiquement nulle, c'est-à-dire,*

$$\forall \tau \in \mathbb{R}, \Gamma_{u,u}(\tau) = E\left(u(t)\, u(t-\tau)\right) - m_u^2 = 0,$$

ceci signifie que $u(t)$ et $u(t-\tau)$ sont non corrélés quelles que soient les valeurs de t. Concrètement, le signal à l'instant t a perdu la mémoire de sa valeur à l'instant $t-\tau$.

Exemple B.17. *Un **bruit blanc** au second ordre (ou signal identiquement distribué) est un signal aléatoire $u(.)$ dont la corrélation à deux instants différents est nulle, c'est-à-dire,*

$$\forall t_1, t_2 \in \mathbb{R}, t_1 \neq t_2, \Gamma_{u,u}(t_1, t_2) = 0.$$

Dans le cas où ce bruit blanc est stationnaire, on a

$$\forall \tau \in \mathbb{R}, \tau \neq 0, \Gamma_{u,u}(\tau) = 0,$$

autrement dit, sa fonction d'autocorrélation est nulle partout sauf à l'origine. Elle correspond donc à une impulsion de Dirac (voir figure B.3). On peut dire qu'un bruit blanc est une collection de variables aléatoires décorrélées.

Il est clair que le concept de stationnarité peut s'étendre aux vecteurs de signaux aléatoires. On parlera alors de vecteur aléatoire $\mathbf{u}(.)$ *conjointement stationnaire*.

Définition B.18 (Signaux aléatoires conjointement stationnaires)**.** Un vecteur de signaux aléatoires $\mathbf{u}(.) = (u_1(.), \ldots, u_n(.))^{\mathrm{T}} \in \mathfrak{S}^n$ est conjointement stationnaire si la matrice d'autocorrelation $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1, t_2)$ dépend seulement de la différence $\tau = t_1 - t_2$, c'est-à-dire,

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1, t_2) = \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau).$$

L'ensemble des signaux de $\mathfrak{S}^n$ conjointement stationnaires est noté $\mathcal{S}^n$ ($\mathcal{S}^n \subset \mathfrak{S}^n$).

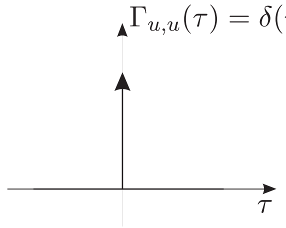


Figure B.3 – Fonction d'autocorrélation $\Gamma_{u,u}(\tau)$ d'un bruit blanc $u(.)$ stationnaire.

Remarque B.19. *Attention, deux signaux aléatoires stationnaires ne sont pas forcément conjointement stationnaires. En effet, on a*

$$\mathcal{S}^n \neq \underbrace{\mathcal{S} \times \cdots \times \mathcal{S}}_{n\,fois}.$$

Dans le cas où $\Gamma_{u_1,u_2}(t_1,t_2) \neq \Gamma_{u_2,u_1}(\tau)$*, on a* $u_1(.), u_2(.) \in \mathcal{S} \not\Rightarrow (u_1(.), u_2(.)) \in \mathcal{S}^2$.

La matrice d'autocorrélation $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ d'un vecteur de signaux $\mathbf{u}(.) \in \mathcal{S}^n$ vérifie la proposition suivante :

Proposition B.20. *La matrice d'autocorrélation* $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ *d'un vecteur de signaux* $\mathbf{u}(.)$ *conjointement stationnaire vérifie :*

$$\forall i, \forall \tau, |\Gamma_{u_i,u_i}(\tau)| \leq \Gamma_{u_i,u_i}(0), \tag{B.5}$$

c'est-à-dire que le maximum absolu d'une fonction d'autocorrélation est atteint pour $\tau = 0$*. et*

$$\forall \tau, \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau) = \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(-\tau)^T, \tag{B.6}$$

Démonstration. L'égalité (B.6) est évidente d'après la proposition B.11 ii) et par définition de la stationnarité.
Remarquons de plus que pour $i = j$, on a

$$\Gamma_{u_i,u_i}(\tau) = \Gamma_{u_i,u_i}(-\tau),$$

c'est-à-dire que la fonction $\Gamma_{u_iu_i}(-\tau)$ est paire.
Pour montrer l'inégalité (B.5), on déduit de l'inégalité de Schwartz[2] appliquée aux variables aléatoires, l'inégalité suivante

$$|\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2)|^2 \leq \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_1)\,\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_2,t_2).$$

Sachant que $\mathbf{u}(.)$ est stationnaire, il résulte que $|\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(t_1,t_2)|^2 \leq \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(0)\,\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(0)$. S'en suit l'inégalité (B.5). $\square$

B.3 Ergodisme

Lorsqu'on utilise l'opérateur d'espérance mathématique E, son application suppose de recourir à ce que l'on appelle une moyenne d'ensemble (voir paragraphe B.1.2). Ainsi, si l'on souhaite estimer la

[2]Pour deux variables aléatoires réelles X_1 et X_2, on a $(E(X_1X_2))^2 \leq E(X_1^2)E(X_2^2)$.

moyenne $m_u(t) = E(u(t))$, à l'instant t, du signal aléatoire $u(.)$ et que l'on dispose de K réalisations $u(\omega_k, t)$ de ce même signal aléatoire, on peut facilement le faire à l'aide la formule suivante :

$$E(u(t)) \approx \hat{m}_u(t) = \frac{1}{K}\sum_{k=0}^{K} u(\omega_k, t). \tag{B.7}$$

Cette forme d'estimation se rapproche de la moyenne d'ensemble. Toutefois, dans la pratique, il arrive souvent qu'on ne dispose que d'une seule réalisation du signal aléatoire. Par contre, il souvent possible d'observer cette réalisation pendant un temps suffisament long. Il est alors assez naturel de vouloir utiliser les moyennes temporelles plutôt que les moyennes d'ensemble. La propriété d'ergodicité régit les relations entre moyennes d'ensemble et moyennes temporelles.

Comme pour la stationnarité, il existe différents types d'ergodisme. Dans ce qui suit, nous nous intéressons uniquement à l'ergodisme d'ordre deux des signaux aléatoires stationnaires.

Définition B.21 (Ergodisme à l'ordre deux)**.** Un signal aléatoire stationnaire est ergodique à l'ordre deux (ou en moyenne quadratique) si les moyennes d'ensemble et les moyennes temporelles suivantes sont identiques

$$E(u(t)) = \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} u(t,\omega)\,dt. \tag{B.8}$$

$$E(u(t)\,u(t-\tau)) = \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} u(t,\omega)\,u(t-\tau,\omega)\,dt. \tag{B.9}$$

Si l'on examine les deux égalités (B.8) et (B.9), on se rend compte que le membre de droite ne dépend pas de t, dès lors, le signal a une moyenne constante qui vérifie :

$$E(u(t)) = m_u.$$

De plus, la fonction d'autocorrélation, qui dépend uniquement de la différence $\tau = t_2 - t_1$, vérifie :

$$\forall t_1, t_2, \Gamma_{u,u}(t_1, t_2) = \Gamma_{u,u}(\tau).$$

B.4 Dérivation

Dans ce paragraphe, nous allons définir la dérivée de signaux aléatoires avant de nous intéresser aux propriétés statistiques des moments d'ordre un et deux de tels signaux. A cette occasion, nous présentons des résultats issues de [7].

B.4.1 Préliminaires

Pour démontrer certains résultats en évitant d'alourdir les calculs, nous introduisons, la grandeur statistique

$$\gamma_{u,y}(t_1, t_2) = E(u(t_1)\,y(t_2)),$$

avec $u(.)$ et $y(.)$ deux signaux aléatoires et $t_1,\ t_2$ deux instants quelconques. Précisons dès maintenant la relation entre $\gamma_{u,y}(t_1, t_2)$ et $\Gamma_{u,y}(t_1, t_2)$.

$$\gamma_{u,y}(t_1, t_2) = \Gamma_{u,y}(t_1, t_2) + E(u(t_1))\,E(y(t_2)).$$

Il ne sera donc pas étonnant que certains résultats vérifiés par $\Gamma_{u,y}(t_1, t_2)$ le soient par $\gamma_{u,y}(t_1, t_2)$ et inversement.

Nous allons maintenant définir précisément la notion de dérivée pour un signal aléatoire.

B.4.2 Définition

Pour caractériser la dérivée de signaux aléatoires, nous posons la définition suivante [44].

Définition B.22 (Dérivée en moyenne quadratique en t_0)**.** Le signal aléatoire $u(.)$ est dérivable en moyenne quadratique en t_0 si la limite

$$\lim_{h\to 0} \frac{u(t_0+h)-u(t_0)}{h}$$

existe.
Lorsqu'elle existe, on la note $\dot{u}(t_0)$. La variable aléatoire $\dot{u}(t_0)$ est appelée dérivée en moyenne quadratique du signal aléatoire $u(.)$ en t_0.

Proposition B.23. *Le signal $u(.)$ a une dérivée en moyenne quadratique à l'instant $t = t_0$ si et seulement si sa matrice d'autocorrélation $\Gamma_{u,u}(t_1,t_2)$ est telle que*

$$\frac{\partial^2 \Gamma_{u,u}(t_1,t_2)}{\partial t_1 \partial t_2}$$

existe en $t_1 = t_2 = t_0$.

Démonstration. Nous allons utiliser le critère de Cauchy pour la convergence en moyenne quadratique. On admet le lemme suivant : "*Une suite de variable aléatoire X_k converge en moyenne quadratique si et seulement si $\forall k,l$, $E(X_kX_l)$ a une limite finie lorsque k et l tendent indépendamment vers plus l'infini*". Ce lemme, appliqué à u, nous dit que $\lim_{h\to 0} \frac{u(t_0+h)-u(t_0)}{h}$ existe (c'est-à-dire $u(.)$ est dérivable en moyenne quadratique en t_0) si et seulement si

$$\lim_{\substack{t_1\to t_0\\ t_2\to t_0}} E\left(\frac{u(t_0+t_1)-u(t_0)}{t_1} \times \frac{u(t_0+t_2)-u(t_0)}{t_2}\right) < \infty. \tag{B.10}$$

En développant le terme (B.10), on obtient la condition

$$\lim_{\substack{t_1\to t_0\\ t_2\to t_0}} \frac{1}{t_1t_2}(E(u(t_0+t_1)\,u(t_0+t_2)) - E(u(t_0+t_1)\,u(t_0))$$
$$-E(u(t_0)\,u(t_0+t_2)) + E(u(t_0)\,u(t_0))) < \infty.$$

Et, par définition de $\Gamma_{u,u}(t_1,t_2)$, ceci équivaut à

$$\lim_{\substack{t_1\to t_0\\ t_2\to t_0}} \frac{1}{t_1t_2}\left(\Gamma_{u,u}(t_0+t_1,t_0+t_2) - \Gamma_{u,u}(t_0+t_1,t_0) - \Gamma_{u,u}(t_0,t_0+t_2) + \Gamma_{u,u}(t_0,t_0)\right) < \infty,$$

qui correspond à l'existence de la dérivée $\frac{\partial^2 \Gamma_{u,u}(t_1,t_2)}{\partial t_1 \partial t_2}$ en (t_0,t_0) . □

Lorsqu'un signal $u(.)$ est dérivable en tout point, sa dérivée constitue un signal aléatoire noté $\dot{u}(.)$. De la proposition B.23, on déduit donc la proposition suivante :

Proposition B.24. *Le signal $u(.)$ est dérivable en moyenne quadratique si et seulement si sa matrice d'autocorrélation $\Gamma_{u,u}(t_1,t_2)$ est telle que*

$$\frac{\partial^2 \Gamma_{u,u}(t_1,t_2)}{\partial t_1 \partial t_2}$$

existe en tout point $t_1 = t_2$.

Nous allons maintenant approfondir et étudier quelques propriétés statistiques essentielles des dérivées d'un signal aléatoire.

B.4.3 Propriétés des dérivées d'un signal aléatoire

Intéressons-nous pour commencer aux statistiques du premier ordre.

B.4.3.1 Statistiques du premier ordre

Proposition B.25. *Pour un signal $u(.)$ dérivable (en moyenne quadratique), la dérivée $\dot{u}(.)$ vérifie*

$$E(\dot{u}(t)) = \frac{d}{dt}E(u(t)).$$

Dans le cas où le signal aléatoire $u(.)$ est stationnaire, on a donc

$$E(\dot{u}(t)) = 0. \tag{B.11}$$

Démonstration. De l'inégalité de Schartz[3], on déduit l'inégalité suivante

$$\left|E\left(\dot{u}(t) - \frac{u(t+h)-u(t)}{h}\right)\right| \leq E\left(\left|\dot{u}(t) - \frac{u(t+h)-u(t)}{h}\right|^2\right)^{\frac{1}{2}}.$$

La convergence en moyenne quadratique implique que le second terme de cette inégalité converge vers 0. Par suite,

$$\lim_{h\to 0} E\left(\dot{u}(t) - \frac{u(t+h)-u(t)}{h}\right) = 0$$

d'où, par linéarité de l'espérance,

$$\lim_{h\to 0}\left(E(\dot{u}(t)) - \frac{E(u(t+h)) - E(u(t))}{h}\right) = 0$$

ou encore

$$\lim_{h\to 0}\frac{E(u(t+h)) - E(u(t))}{h} = E(\dot{u}(t))$$

c'est-à-dire,

$$\frac{d}{dt}E(u(t)) = E(\dot{u}(t)).$$

□

Maintenant, nous présentons les propriétés des statistiques des moments d'ordre deux.

B.4.3.2 Statistiques du second ordre

B.4.3.2.1 Cas des signaux aléatoires quelconques Dans [44], I. Guikham et A. Skorokhod proposent les formules suivantes :

$$\begin{aligned}
\gamma_{u,\dot{u}}(t_1,t_2) &= \frac{\partial\gamma_{u,u}(t_1,t_2)}{\partial t_2}, \qquad (B.12)\\
\gamma_{\dot{u},u}(t_1,t_2) &= \frac{\partial\gamma_{u,u}(t_1,t_2)}{\partial t_1},\\
\gamma_{\dot{u},\dot{u}}(t_1,t_2) &= \frac{\partial^2\gamma_{u,u}(t_1,t_2)}{\partial t_1\partial t_2},
\end{aligned}$$

[3]Pour deux variables aléatoires réelles X_1 et X_2, on a $(E(X_1X_2))^2 \leq E(X_1)^2E(X_2)^2$

qui peuvent évidemment être étendues pour des dérivées d'ordre supérieur. Nous présentons ici un résultat plus général :

Proposition B.26. *Soit* $u(.)$ *et* $y(.)$ *deux signaux aléatoires suffisamment dérivables (en moyenne quadratique), on a*

$$\forall k, l \in \mathbb{N}, \Gamma_{u^{(k)},y^{(l)}}(t_1, t_2) = \frac{\partial^{k+l}\Gamma_{u,y}(t_1, t_2)}{\partial t_1^k \partial t_2^l}.$$

Dans le cas où les signaux aléatoires $u(.)$ *et* $y(.)$ *sont stationnaires, on a*

$$\forall k, l \in \mathbb{N}, \Gamma_{u^{(k)},y^{(l)}}(\tau) = (-1)^l \frac{\partial^{k+l}\Gamma_{u,y}(\tau)}{\partial \tau^{k+l}}. \tag{B.13}$$

Démonstration. Nous présentons la preuve de cette proposition dans le cas $k = 0, l = 1$ pour bien mettre en évidence notre démarche. La preuve dans le cas général ne présente pas de difficultés supplémentaires, elle se fait par récurrence sur les ordres de dérivation.
Soient $u(.)$ et $y(.)$, deux signaux aléatoires. On a , pour $t_1, t_2, h \in \mathbb{R}$,

$$\left| E\left(u(t_1)\left(\dot{y}(t_2) - \frac{y(t_2+h) - y(t_2)}{h}\right)\right)\right| \leq \left| E\left(u(t_1)\left(\dot{y}(t_2) - \frac{y(t_2+h) - y(t_2)}{h}\right)\right)^2\right|^{\frac{1}{2}}.$$

D'après l'inégalité de Schwartz[4] appliquée aux variables aléatoires, on en déduit :

$$\left| E\left(u(t_1)\left(\dot{y}(t_2) - \frac{y(t_2+h) - y(t_2)}{h}\right)\right)\right| \leq \left| E\left(u(t_1)^2\right)\right|^{\frac{1}{2}} \left| E\left(\left(\dot{y}(t_2) - \frac{y(t_2+h) - y(t_2)}{h}\right)^2\right)\right|^{\frac{1}{2}}.$$

La convergence en moyenne quadratique de $y(.)$ implique que le second terme de cette inégalité converge vers 0 lorsque h tend vers 0. D'où

$$\lim_{h\to 0} E\left(u(t_1)\left(\dot{y}(t_2) - \frac{y(t_2+h) - y(t_2)}{h}\right)\right) = 0.$$

Par linéarité de l'espérance, on en déduit que :

$$\lim_{h\to 0}\left(E(u(t_1)\dot{y}(t_2)) - \frac{E(u(t_1)y(t_2+h)) - E(u(t_1)y(t_2))}{h}\right) = 0,$$

puis

$$\lim_{h\to 0}\left(\frac{E(u(t_1)y(t_2+h)) - E(u(t_1)y(t_2))}{h}\right) = E(u(t_1)\dot{y}(t_2)),$$

soit encore

$$\lim_{h\to 0}\left(\frac{\gamma_{u,y}(t_1, t_2+h) - \gamma_{u,y}(t_1, t_2)}{h}\right) = E(u(t_1)\dot{y}(t_2)).$$

En conséquence, $\frac{\partial \gamma_{u,y}(t_1,t_2)}{\partial t_2}$ existe et

$$\frac{\partial \gamma_{u,y}(t_1, t_2)}{\partial t_2} = \gamma_{u,\dot{y}}(t_1, t_2). \tag{B.14}$$

[4]Pour deux variables aléatoires réelles X_1 et X_2, on a $(E(X_1X_2))^2 \leq E(X_1^2)E(X_2^2)$.

Montrons que l'égalité (B.14) est également vérifiée par $\Gamma_{u,y}(t_1,t_2)$. En effet,

$$\Gamma_{u,\dot{y}}(t_1,t_2) = \gamma_{u,\dot{y}}(t_1,t_2) - E(u(t_1))\,E(\dot{y}(t_2)).$$

D'après la proposition B.25 et l'égalité (B.14), on obtient

$$\begin{aligned}
\Gamma_{u,\dot{y}}(t_1,t_2) &= \frac{\partial \gamma_{u,y}(t_1,t_2)}{\partial t_2} - E(u(t_1))\,\frac{\partial E(y(t_2))}{\partial t_2} \\
&= \frac{\partial}{\partial t_2}\left(\gamma_{u,y}(t_1,t_2) - E(u(t_1))\,E(y(t_2))\right) \\
&= \frac{\partial \Gamma_{u,y}(t_1,t_2)}{\partial t_2}.
\end{aligned}$$

Dans le cas stationnaire, on a

$$\Gamma_{u,y}(t_1,t_2) = \Gamma_{u,y}(t_1,t_1-\tau) = \Gamma_{u,y}(\tau),$$

avec $\tau = t_1 - t_2$, donc $d\tau = dt_1 - dt_2$. On obtient aisément le résultat. □

Nous allons maintenant nous intéresser tout particulièrement au cas des signaux stationnaires, puis aux vecteurs de signaux stationnaires.

B.4.3.2.2 Cas des signaux aléatoires stationnaires L'égalité (B.13) concernant les signaux stationnaires peut être étendue au cas vectoriel. Nous commençons par énoncer un lemme.

Lemme B.27. *Soient n signaux aléatoires stationnaires $u_1(.),\ldots,u_n(.)$. On a $\forall k,\ell \in \mathbb{N}$ et $\forall i,j \in \{1,\ldots,n\}$*

$$\Gamma_{u_i^{(k)},u_j^{(l)}}(\tau) = (-1)^{\ell}\,\Gamma_{u_i,u_j}^{(k+l)}(\tau) = (-1)^{k}\,\Gamma_{u_j,u_i}^{(k+l)}(-\tau).$$

Démonstration. D'après l'égalité (B.13), on a

$$\Gamma_{u_i^{(k)},u_j^{(l)}}(\tau) = (-1)^{l}\,\frac{\partial^{k+l}\Gamma_{u_i,u_j}(\tau)}{\partial \tau^{k+l}} = (-1)^{\ell}\,\Gamma_{u_i,u_j}^{(k+l)}(\tau).$$

D'autre part, d'après la symétrie de la fonction d'intercorrelation,

$$\Gamma_{u_i,u_j}(\tau) = \Gamma_{u_j,u_i}(-\tau).$$

Après $k+l$ dérivations, on obtient

$$\Gamma_{u_i,u_j}^{(k+l)}(\tau) = (-1)^{k+l}\,\Gamma_{u_j,u_i}^{(k+l)}(-\tau),$$

qui peut s'écrire

$$(-1)^{l}\,\Gamma_{u_i,u_j}^{(k+l)}(\tau) = (-1)^{k}\,\Gamma_{u_j,u_i}^{(k+l)}(-\tau).$$

□

On en déduit la proposition suivante :

Proposition B.28. *Soient* $\mathbf{u}(.) \in \mathcal{S}^n$ *et* $\mathbf{y}(.) \in \mathcal{S}^n$ *deux vecteurs de* n *signaux aléatoires. On a* $\forall k, \ell \in \mathbb{N}$,

$$\boldsymbol{\Gamma}_{\mathbf{u}^{(k)},\mathbf{y}^{(l)}}(\tau) = (-1)^{\ell}\,\boldsymbol{\Gamma}^{(k+l)}_{\mathbf{u},\mathbf{y}}(\tau) = (-1)^{k}\,\boldsymbol{\Gamma}^{(k+l)}_{\mathbf{y},\mathbf{u}}(-\tau)^{T}.$$

Démonstration. Conséquence directe du lemme B.27, par définition de la matrice d'intercorrélation de vecteurs de signaux. □

Ainsi tous les moments d'ordre deux des dérivées de vecteurs de signaux aléatoires stationnaires peuvent être obtenus par dérivations successives de leur matrice d'intercorrélation, et inversement. Pour les matrices d'autocorrélation, on obtient le résultat suivant :

Corollaire B.29. *Si* $\mathbf{u}(.) \in \mathcal{S}^n$, *alors* $\forall k \in \mathbb{N}$,

$$\begin{array}{ll} i) & \boldsymbol{\Gamma}^{(2k+1)}_{\mathbf{u},\mathbf{u}}(\tau) = -\boldsymbol{\Gamma}^{(2k+1)}_{\mathbf{u},\mathbf{u}}(-\tau)^{T}, \\ ii) & \boldsymbol{\Gamma}^{(2k)}_{\mathbf{u},\mathbf{u}}(\tau) = \boldsymbol{\Gamma}^{(2k)}_{\mathbf{u},\mathbf{u}}(-\tau)^{T}. \end{array} \tag{B.15}$$

Démonstration. i) D'après la proposition B.28, pour $l = 0$, on a

$$\boldsymbol{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau) = (-1)^{k}\,\boldsymbol{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(-\tau)^{\mathrm{T}}, \quad \forall k \in \mathbb{N}. \tag{B.16}$$

Et, pour $k = 2p + 1$ impaire, cette relation donne

$$\boldsymbol{\Gamma}^{(2p+1)}_{\mathbf{u},\mathbf{u}}(\tau) = -\boldsymbol{\Gamma}^{(2p+1)}_{\mathbf{u},\mathbf{u}}(-\tau)^{\mathrm{T}}, \forall p \in \mathbb{N}.$$

ii) De même, pour $k = 2p$, on obtient

$$\boldsymbol{\Gamma}^{(2p)}_{\mathbf{u},\mathbf{u}}(\tau) = \boldsymbol{\Gamma}^{(2p)}_{\mathbf{u},\mathbf{u}}(-\tau)^{\mathrm{T}}.$$

□

Remarque B.30. $\forall p \in \mathbb{N}$, *la matrice d'autocorrélation* $\boldsymbol{\Gamma}^{(p)}_{\mathbf{u},\mathbf{u}}(0)$ *est antisymétrique si* p *est impair et symétrique si* p *est pair.*

B.5 Indépendance

Dans cette thèse, nous avons choisi une définition originale de l'indépendance de deux signaux aléatoires. Cette définition n'étant pas classique, nous détaillons son intérêt et insistons particulièrement sur les nuances qu'elle introduit par rapport à la définition de l'indépendance classiquement rencontrée dans la littérature.

B.5.1 Préliminaires

La définition de l'indépendance que nous proposons fait intervenir un nouvel ensemble de fonctions. Nous allons donc commencer par expliciter ce dernier.

Soit $\mathcal{F}$ l'ensemble des fonctions définies sur $\mathbb{R}^{\mathbb{R}}$ à valeurs dans $\mathbb{R}$,

$$\mathcal{F} = \left\{ f : \mathbb{R}^{\mathbb{R}} \to \mathbb{R} \right\}.$$

Par exemple,

$$f:\left\{\begin{array}{ccc}\mathbb{R}^{\mathbb{R}} & \rightarrow & \mathbb{R}\\ u(t) & \rightarrow & u(0)\end{array}\right. \tag{B.17}$$

est une fonction de $\mathcal{F}$ qui associe à un signal $u(t)$ sa valeur à l'instant 0.

Rappelons qu'une variable aléatoire réelle X est une fonction qui associe à un événement ω une réalisation (voir annexe A) :

$$X:\left\{\begin{array}{ccc}\Omega & \rightarrow & \mathbb{R}\\ \omega & \rightarrow & X(\omega)\end{array}\right. \in \mathcal{V}.$$

Il est particulièrement important de constater que la composition d'un signal aléatoire $u(.)$ avec une fonction $f \in \mathcal{F}$ est une variable aléatoire. En effet,

$$f \circ u(.):\left\{\begin{array}{ccccc}\Omega & \xrightarrow{u(.)} & \mathbb{R}^{\mathbb{R}} & \xrightarrow{f} & \mathbb{R}\\ \omega & \rightarrow & u(\omega,t) & \rightarrow & f(u(\omega,t))\end{array}\right. \in \mathcal{V}.$$

Notons que la fonction f agit sur les réalisations $u(\omega,t)$ de $u(.)$.
Par exemple, considérons la fonction définie en (B.17),

$$f(u(.)):\left\{\begin{array}{ccc}\Omega & \rightarrow & \mathbb{R}\\ \omega & \rightarrow & u(\omega,0)\end{array}\right. \in \mathcal{V} \tag{B.18}$$

est la variable aléatoire $X = f(u(.))$ extraite de $u(.)$ à l'instant $t = 0$.

Bien entendu, dans cet exemple, la fonction f est volontairement simple. Elle a été choisie de façon à illustrer au mieux notre démarche. Il existe évidemment des fonctions beaucoup plus complexes dans l'ensemble $\mathcal{F}$, par exemple

$$f:\left\{\begin{array}{ccc}\mathbb{R}^{\mathbb{R}} & \rightarrow & \mathbb{R}\\ u(t) & \rightarrow & \dot{u}(5)\exp\left(u^2(0)\right) - \sin(u(10)+1) - 8\end{array}\right. \in \mathcal{F},$$

dans la mesure où cette fonction est bien définie (*i.e.* $\dot{u}(5)$ existe).

Par la suite, lorsqu'on parlera de la variable aléatoire $f(u(.))$, il sera sous-entendu que $f(u(.))$ est bien définie.

B.5.2 Indépendance de signaux aléatoires

Dans cette section, nous allons définir l'indépendance de signaux aléatoires. Classiquement, elle est définie comme suit.

Définition B.31 (Indépendance classique)**.** Deux signaux aléatoires $u_1(.)$ et $u_2(.)$ sont indépendants si $\forall t_i \in \mathbb{R}, i \in \{1,\ldots,N\} \subset \mathbb{N}$, les vecteurs aléatoires $(u_1(t_1)\ldots u_1(t_p))$ et $(u_2(t_{p+1})\ldots u_2(t_N))$ sont indépendants.

Cette définition ne nous convenait pas dans la mesure où elle considère une version discrétisée des signaux. Nous proposons donc une définition "continue" de l'indépendance de signaux en cherchant à se ramener à l'indépendance de variables aléatoires, plus faciles à manipuler.

Par extension naturelle de la définition classique, nous posons la définition suivante pour l'indépendance de deux signaux aléatoires :

Définition B.32 (Indépendance de deux signaux aléatoires)**.** Soit $\mathcal{F}$ l'ensemble des fonctions définies de $\mathbb{R}^{\mathbb{R}}$ dans $\mathbb{R}$. Les signaux aléatoires $u_1(.), u_2(.) \in \mathfrak{S}^2$ sont statistiquement indépendants si $\forall f_1, f_2 \in \mathcal{F}$, les variables aléatoires $X_1 = f_1(u_1(.)), X_2 = f_2(u_2(.))$ sont statistiquement indépendantes, *i.e.*, dans le cas où les densités existent,

$$f_{X_1,X_2} = f_{X_1} f_{X_2},$$

avec f_{X_1,X_2} et f_{X_i} sont respectivement les densités de probabilité conjointe et marginales des X_i.
On note I^2 l'ensemble des couples de signaux aléatoires indépendants ($I^2 \subset \mathfrak{S}^2$).

Notation B.33. *On note $I_s(u_1, u_2)$ l'indépendance des signaux aléatoires $u_1(.)$ et $u_2(.)$ au sens de la définition B.32. Ainsi définie, I_s est une relation binaire sur $\mathfrak{S} \times \mathfrak{S}$. On peut remarquer que I_s n'est pas une relation d'équivalence puisqu'elle n'est ni reflexive,*

$$I_s(u_1, u_1) \text{ est faux,}$$

ni transitive,

$$I_s(u_1, u_2) \text{ et } I_s(u_2, u_3) \nRightarrow I_s(u_1, u_3) \text{ (par exemple pour } u_1 = u_3\text{).}$$

Maintenant, précisons les nuances introduites par cette nouvelle définition. Tout d'abord, on a le résultat suivant :

Remarque B.34. *La définition B.32 (que nous proposons) est une plus forte que la définition classique B.31 de l'indépendance.*

Démonstration. Précisons tout d'abord que la définition B.31 de l'indépendance de deux signaux aléatoires au sens classique est équivalente à

$$\begin{aligned} &\forall t_1, \ldots, t_p, t_{p+1}, \ldots, t_N, \\ &\forall g_1 : \mathbb{R}^p \to \mathbb{R}, \forall g_2 : \mathbb{R}^{N-p} \to \mathbb{R}, \\ &I_v(g_1(u_1(t_1), \ldots, u_1(t_p)), g_2(u_2(t_{p+1}), \ldots, u_2(t_N))). \end{aligned} \tag{B.19}$$

Maintenant, considérons deux signaux aléatoires $u_1(.), u_2(.)$ tels que $I_s(u_1, u_2)$. On a

$$\forall f_1, f_2 \in \mathcal{F}, I_v(f_1(u_1), f_2(u_2)).$$

Soient $t_1, \ldots, t_p, t_{p+1}, \ldots, t_N, g_1 : \mathbb{R}^p \to \mathbb{R}, g_2 : \mathbb{R}^{N-p} \to \mathbb{R}$. Considérons le couple de fonctions (f_1, f_2) avec

$$f_1 : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R} \\ u(t) & \to & g_1(u(t_1), \ldots, u(t_p)), \end{array} \right.$$
$$f_2 : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R} \\ u(t) & \to & g_2(u(t_{p+1}), \ldots, u(t_N)). \end{array} \right.$$

Ainsi définies, les fonctions f_i appartiennent bien à $\mathcal{F}$ et vérifient donc $I_v(f_1(u_1), f_2(u_2))$. On a donc la condition (B.19). $\square$

Les définitions B.32 et B.31 de l'indépendance de deux signaux aléatoires sont ressemblantes, à la seule distinction près que la définition classique propose de vérifier l'indépendance sur une collection finie d'instants du temps alors que notre définition considère les signaux aléatoires dans leur globalité (*i.e.* pour tous les instants possibles). La définition B.32 va nous être très utile pour démontrer des propriétés sur les dérivées de signaux aléatoires indépendants. Avant de poursuivre, nous l'illustrons à l'aide d'un exemple simple.

Exemple B.35. *Considérons deux signaux $u_1(.)$ et $u_2(.)$ indépendants et gaussiens de moyenne nulle et d'écart type unitaire. On construit deux variables aléatoires X_1 et X_2, extraites respectivement de $u_1(.)$ et $u_2(.)$, définies par*

$$X_1 = f_1(u_1(.)) = \sin\left(u_1^2(20)\right) + u_1(10),$$
$$X_2 = f_2(u_2(.)) = u_2(15) + \cos(u_2(40)),$$

avec $f_1, f_2 \in \mathcal{F}$.

La figure B.4 montre une réalisation de chacun des signaux $u_1(.)$ et $u_2(.)$ à laquelle correspond respectivement une réalisation de chacune des variables aléatoires X_1 et X_2.

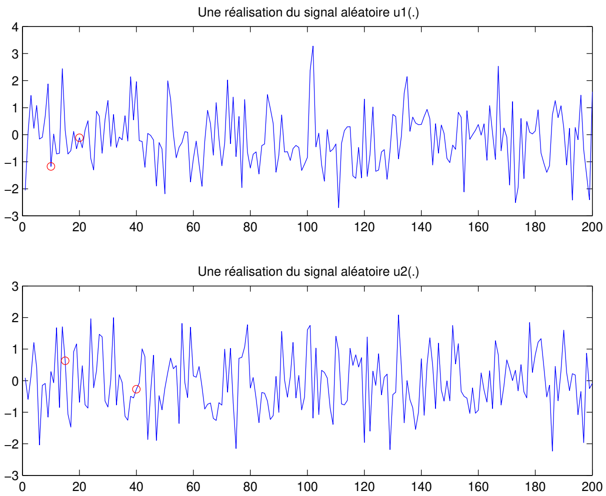


Figure B.4 – Exemple de réalisations pour les signaux aléatoires $u_1(.)$ et $u_2(.)$

La figure B.5 représente la distribution conjointe des variables aléatoires X_1 et X_2 pour 10000 réalisations des signaux aléatoires $u_1(.)$ et $u_2(.)$. On peut voir que cette distribution est caractéristique de deux variables indépendantes puisque la connaissance de X_1 ne donne aucune information sur X_2 et inversement.

Par définition, $u_1(.)$ et $u_2(.)$ sont indépendants si toutes les variables X_1 et X_2, créées à partir de fonctions f_1 et f_2 de $\mathcal{F}$, aussi complexes soient-elles, sont indépendantes.

Remarque B.36. *Tout élément de l'ensemble des réels $\mathbb{R}$ peut être vu comme un élément de $\mathcal{V}$ ($\mathbb{R} \subset \mathcal{V}$). Par exemple, le scalaire $\alpha \in \mathbb{R}$ correspond à une variable aléatoire gaussienne de moyenne α et d'écart type nul (sa densité de probabilité est un pic de Dirac en α). Ainsi, on peut montrer, d'après le théorème A.66 (page 152) que pour tout scalaire α_1 et α_2, on a $I_v(\alpha_1, \alpha_2)$. Il suffit de voir que*

$$\forall f, g, E(f(\alpha_1)\, g(\alpha_2)) = E(f(\alpha_1))\, E(g(\alpha_2)) \quad (= f(\alpha_1)\, g(\alpha_2)).$$

Bien qu'il semble naturel de penser que deux variables aléatoires constantes identiques, par exemple 1 et 1, ne soient pas indépendantes, nous venons de voir que c'est au contraire un cas dégénéré de l'indépendance de variables.
En conséquence, la définition B.32 est bien valable pour toutes les fonctions $f_i \in \mathcal{F}$, même constantes[5].

[5] c'est-à-dire de la forme $f : \mathbb{R}^{\mathbb{R}} \ni u(t) \to c \in \mathbb{R}$

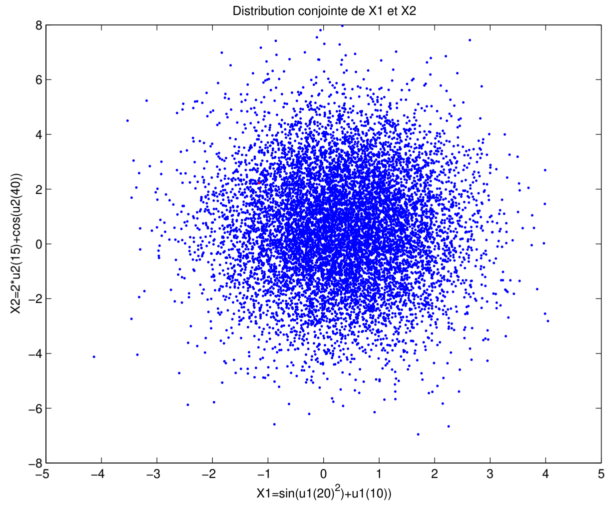


Figure B.5 – Distribution conjointe des variables aléatoires X_1 et X_2.

B.5.3 Généralisation au cas de vecteur de signaux aléatoires

Maintenant, nous proposons de généraliser la définition de l'indépendance au cas d'une famille de signaux aléatoires. Pour cela, on note $\mathcal{F}^i, i \in \mathbb{N}$ l'ensemble des fonctions de $\left(\mathbb{R}^{\mathbb{R}}\right)^i$ à valeur dans $\mathbb{R}$. Par exemple :

$$f : \left\{ \begin{array}{ccc} \left(\mathbb{R}^{\mathbb{R}}\right)^2 & \rightarrow & \mathbb{R} \\ (u_1(t), u_2(t)) & \rightarrow & u_1(0) + u_2(0) \end{array} \right. \in \mathcal{F}^2.$$

Définition B.37 (Indépendance mutuelle de signaux aléatoires)**.** Les n signaux aléatoires $u_1(.), \dots, u_n(.)$ sont mutuellement indépendants (noté $I_s(u_1, \dots, u_n)$) si on a les deux conditions suivantes :

$$\begin{array}{ll} \text{i)} & \forall f_1, \dots, f_n \in \mathcal{F},\ I_v\left(f_1(u_1), \dots, f_n(u_n)\right). \\ \text{ii)} & \forall i \in \{1, \dots, n-1\},\ \forall f \in \mathcal{F}^i, \forall g \in \mathcal{F}^{n-i}, \\ & \forall \sigma \in S_n \text{ tel que } \left\{ \begin{array}{l} \sigma(1) < \sigma(2) < \cdots < \sigma(i), \\ \sigma(i+1) < \cdots < \sigma(n), \\ \sigma(1) < \sigma(i+1), \end{array} \right. \\ & I_v\left(f\left(u_{\sigma(1)}, u_{\sigma(2)}, \dots, u_{\sigma(i)}\right), g\left(u_{\sigma(i+1)}, \dots, u_{\sigma}(n)\right)\right). \end{array}$$

Si on a uniquement la condition i), on dit que les signaux aléatoires $u_1, \dots, u_n$ sont indépendants deux à deux.
On note I^n l'ensemble des vecteurs de n signaux aléatoires mutuellement indépendants ($I^n \subset \mathfrak{S}^n$).

La définition B.37 peut sembler complexe car les indices des signaux aléatoires ont été rigoureusement choisis afin d'éviter les répétitions inutiles. Nous allons, avant de poursuivre, l'illustrer dans le cas de trois signaux aléatoires ($n = 3$).

Exemple B.38. *Trois signaux aléatoires $u_1(.), u_2(.), u_3(.) \in \mathfrak{S}^3$ sont mutuellement indépendants si*

$$\begin{array}{ll} i) & \forall f_1, f_2, f_3 \in \mathcal{F}, I_v\left(f_1(u_1), f_2(u_2), f_3(u_3)\right) \\ ii) & \forall f_1 \in \mathcal{F}^2, f_2 \in \mathcal{F}, I_v\left(f_1(u_1, u_2), f_2(u_3)\right) \\ iii) & \forall f_1 \in \mathcal{F}^2, f_2 \in \mathcal{F}, I_v\left(f_1(u_1, u_3), f_2(u_2)\right) \\ iv) & \forall f_1 \in \mathcal{F}^2, f_2 \in \mathcal{F}, I_v\left(f_1(u_2, u_3), f_2(u_1)\right) \end{array}$$

Si seule la condition i) est satisfaite, les signaux $u_1(.), u_2(.), u_3(.)$ *sont indépendants deux à deux.*

Précisons évidemment que, par définition, l'indépendance mutuelle est une hypothèse plus forte que l'indépendance deux à deux, on a :

$$\text{Indépendance mutuelle} \implies \text{Indépendance deux à deux.}$$

Dans le cas de deux signaux aléatoires, l'indépendance mutuelle et l'indépendance deux à deux sont équivalentes.

B.5.4 Conservation de l'indépendance

Maintenant, afin de caractériser l'ensemble des fonctions qui préservent l'indépendance statistique des signaux, intéressons-nous à l'ensemble $\mathfrak{F}$ des fonctions définies de $\mathbb{R}^{\mathbb{R}}$ à valeur dans $\mathbb{R}^{\mathbb{R}}$. Ces fonctions qui, à un signal, associe un signal, sont appelées filtres. Par exemple, les fonctions

$$g_1 : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \longrightarrow & \mathbb{R}^{\mathbb{R}} \\ u(t) & \mapsto & \cos(u(t)) \end{array} \right. \quad \text{et} \quad g_2 : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \longrightarrow & \mathbb{R}^{\mathbb{R}} \\ u(t) & \mapsto & u(t)^2 + \exp(\dot{u}(t)) \end{array} \right. \tag{B.20}$$

sont des filtres appartenant à $\mathfrak{F}$.

Avant de commencer, on peut vérifier, par un raisonnement similaire à celui du paragraphe B.5.1, que pour un filtre $g \in \mathfrak{F}$ et un signal aléatoire $u(.) \in \mathfrak{S}$, $g(u(.))$ est bien un signal aléatoire. En effet,

$$g(u(.)) : \left\{ \begin{array}{ccccc} \Omega & \xrightarrow{u(.)} & \mathbb{R}^{\mathbb{R}} & \xrightarrow{g} & \mathbb{R}^{\mathbb{R}} \\ \omega & \mapsto & u(\omega, t) & \mapsto & g(u(\omega, t)) \end{array} \right.$$

associe à un événement ω, une réalisation $g(u(\omega, t))$ (c'est une fonction du temps), il s'agit donc bien d'un signal aléatoire. Par exemple, avec le filtre g_1 défini par (B.20), on a

$$g_1(u(.)) = \cos(u(.)) \in \mathfrak{S}.$$

Proposition B.39 (Conservation de l'indépendance)**.** *Soient n signaux aléatoires $u_1(.), \ldots, u_n(.)$. L'ensemble $\mathfrak{F}$ est compatible avec la relation I_s, c'est-à-dire*

$$I_s(u_1, \ldots, u_n) \Leftrightarrow \forall g \in \mathfrak{F},\, I_s(g(u_1(.)), \ldots, g(u_n(.))). \tag{B.21}$$

De plus (voir figure B.6), on a

$$I_s(u_1, \ldots, u_n) \Leftrightarrow \forall g_1, \ldots, g_n \in \mathfrak{F},\, I_s(g_1(u_1(.)), \ldots, g_n(u_n(.))). \tag{B.22}$$

Démonstration. Pour éviter d'alourdir la démonstration, elle est présentée dans le cas de deux signaux aléatoires, mais peut être étendue aisément au cas de n signaux.
Nous allons commencer par montrer l'assertion (B.22). Soient $g_1, g_2 \in \mathfrak{F}$. Rappelons, d'après la définition B.32, que

$$I_s(g_1(u_1), g_2(u_2)) \Leftrightarrow \forall f_1, f_2 \in \mathcal{F}, I_v(f_1(g_1(u_1)), f_2(g_2(u_2))). \tag{B.23}$$

Pour $f_1, f_2 \in \mathcal{F}$, définissons les fonctions h_1 et h_2 telles que

$$h_i : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \rightarrow & \mathbb{R} \\ u_i(t) & \rightarrow & f_i(g_i(u_i(t))) \end{array} \right., \quad i = 1, 2.$$

Sachant que $I_s(u_1, u_2)$ et $h_i \in \mathcal{F}$, on a $I_v(h_1(u_1), h_2(u_2))$, *i.e.* $I_v(f_1(g_1(u_1)), f_2(g_2(u_2)))$. D'après (B.23), on obtient $I_s(g_1(u_1), g_2(u_2))$.
L'assertion (B.21) se montre trivialement en posant $g = g_1 = g_2$. □

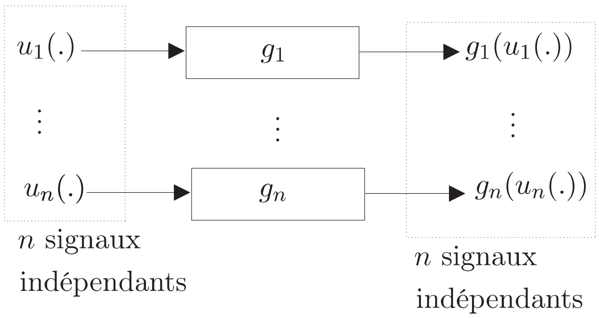


Figure B.6 – Illustration de la conservation de l'indépendance.

Exemple B.40. *Considérons deux signaux aléatoires $u_1(.)$ et $u_2(.)$ indépendants. On a, par exemple, $u_1^2(.)$ et $\cos(u_2(.)) + \dot{u}_2(.)$ indépendants.*

Remarque B.41. *Par analogie avec la remarque B.36, un signal déterministe de $\mathbb{R}^{\mathbb{R}}$ peut être vu comme un élément de $\mathfrak{S}$ $\left(\mathbb{R}^{\mathbb{R}}\subset\mathfrak{S}\right)$. Par exemple, le signal $\cos(t) \in \mathbb{R}^{\mathbb{R}}$ est un signal aléatoire qui à chaque événement associe la même réalisation $\cos(t)$.*
On peut montrer que deux signaux déterministes $u_1(t)$ et $u_2(t)$ sont indépendants. En effet, $\forall f_1, f_2 \in \mathcal{F}$, on a $f_i(u_i(t)) \in \mathbb{R}$ et donc, sachant que deux scalaires sont indépendants, on a $I_v(f_1(u_1(t)), f_2(u_2(t)))$.
C'est un cas dégénéré de l'indépendance de signaux. Par exemple, $\sin(t)$ est indépendant de $\sin(t)$.
Ainsi, la proposition B.39 est bien valable pour tous les filtres $g_i \in \mathfrak{F}$, même constants[6] alors qu'on aurait pu intuitivement supposer le contraire.

Remarque B.42. *Suite à la remarque précédente notifiant que $\mathbb{R}^{\mathbb{R}}\subset\mathfrak{S}$, il est intéressant de noter également que tout signal déterministe (non constant[7]) est non ergodique. En effet, considérons le signal $\cos(t) \in \mathbb{R}^{\mathbb{R}}$, sa moyenne temporelle $m(\cos(t))$ est nulle (sur une période) et sa moyenne d'ensemble $E(\cos(t))$ vaut $\cos(t)$.*

Intéressons-nous maintenant à la structure de l'ensemble $(\mathfrak{F}, \circ)$ qui, comme le montre la proposition B.39, préserve l'indépendance.

Proposition B.43. *L'ensemble $(\mathfrak{F}, \circ)$ est un monoïde[8] dont l'élément neutre est le filtre identité défini par*

$$Id : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R}^{\mathbb{R}} \\ u(t) & \to & u(t) \end{array} \right. .$$

Démonstration. $\circ$ est une loi interne :

$$\forall f, g \in \mathfrak{F},\ f\circ g \in \mathfrak{F}.$$

$\circ$ est associative :

$$\forall f, g, h \in \mathfrak{F},\ (f\circ g) \circ h = f\circ (g \circ h).$$

$\circ$ a un élément neutre :

$$\forall f \in \mathfrak{F}, Id \circ f = f \circ Id = f.$$

□

[6]La fonction $f \in \mathfrak{F}$ est constante si $f(u(t))$ ne dépend pas de $u(t)$. Par exemple, $f(u(t)) = \sin(t)$ est une fonction constante.
[7]Un signal constant ne dépend pas du temps.
[8]L'ensemble $(\mathfrak{F}, \circ)$ est un monoïde si $\circ$ est une loi de composition interne pour $\mathfrak{F}$ qui est associative.

Remarque B.44. *Montrons, à l'aide d'un contre exemple, que $(\mathfrak{F}, \circ)$ n'est pas un groupe. Pour cela, considérons le filtre valeur absolue*

$$abs : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \rightarrow & \mathbb{R}^{\mathbb{R}} \\ u(t) & \rightarrow & abs(u(t)) \end{array} \right. \in \mathfrak{F}.$$

Ce filtre n'étant pas inversible, $(\mathfrak{F}, \circ)$ n'est pas un groupe.
L'ensemble $(\mathfrak{F}, \circ)$ n'est pas non plus un semi-groupe car la loi interne $\circ$ ne commute pas :

$$\exists f, g \in \mathfrak{F}, f \circ g \neq g \circ f.$$

Nous énonçons la proposition suivante énumérant (non exhautivement) quelques fonctions de $\mathfrak{F}$ (filtres) fort utiles préservant l'indépendance.

Proposition B.45. *Pour n signaux aléatoires $u_1(.), \ldots, u_n(.) \in \mathfrak{S}^n$, on a*

$$i) \quad I_s(u_1, \ldots, u_n) \Leftrightarrow I_s(u_n, \ldots, u_1), \tag{B.24}$$

$$ii) \quad I_s(u_1, y), \ldots, I_s(u_n, y) \implies \forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}, I_s(\textstyle\sum_{i=1}^{i=n} \alpha_i u_i, y), \tag{B.25}$$

$$iii) \quad I_s(u_1, \ldots, u_n) \implies \forall (i_1, \ldots, i_n) \in \mathbb{N}^n, I_s\left(u_1^{(i_1)}, \ldots, u_n^{(i_n)}\right). \tag{B.26}$$

Démonstration. i) Trivial.

ii) Supposons que $I_s(u_1, y), \ldots, I_s(u_n, y)$ alors, on a

$$\forall i, \forall f_i \in \mathcal{F}, \forall g_i \in \mathcal{F}, \ I_v(f_i(u_i), g_i(y)).$$

Ceci implique, pour $g = g_1 = \cdots = g_n$,

$$\forall i, \forall f_i \in \mathcal{F}, \forall g \in \mathcal{F}, \ I_v(f_i(u_i), g(y)).$$

Et, pour toute fonction $h : \mathbb{R}^n \to \mathbb{R}$, on a donc

$$I_v(h(f_1(u_1), \ldots, f_n(u_n)), g(y)). \tag{B.27}$$

Maintenant, pour tout filtre $f \in \mathcal{F}$ et pour tout n-uplet $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, il existe des fonctions $f_1, \ldots, f_n, h$ telles que

$$h(f_1(u_1), \ldots, f_n(u_n)) = f(\alpha_1 u_1 + \cdots + \alpha_n u_n). \tag{B.28}$$

Pour vérifier cette égalité, il suffit de poser $f_i(u_i) = \alpha_i u_i$ et $h(x_1, \ldots, x_n) = f(x_1 + \cdots + x_n)$. Par conséquent, la relation (B.27) implique que

$$\forall f \in \mathcal{F}, \forall (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n, \forall g \in \mathcal{F}, I_v(f(\alpha_1 u_1 + \cdots + \alpha_n u_n), g(y)), \tag{B.29}$$

ce qui est équivalent à $I_s(\alpha_1 u_1 + \cdots + \alpha_n u_n, y)$.

iii) Trivial, d'après la proposition B.39 pour des filtres g_j définis par

$$g_j : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \rightarrow & \mathbb{R}^{\mathbb{R}} \\ u_j(t) & \rightarrow & u_j^{(i_j)}(t) \end{array} \right. \in \mathfrak{F}.$$

$\square$

Par la suite, lorsqu'on parlera d'indépendance de n signaux (notée $I_s(u_1,\ldots,u_n)$), il s'agira de l'indépendance deux à deux, à moins que le contraire ne soit explicitement noté. Ainsi,

$$I_s(u_1,\ldots,u_n) \Leftrightarrow \begin{cases} I_s(u_1,u_2) \\ I_s(u_1,u_3) \\ I_s(u_2,u_3) \\ \text{etc}\ldots \end{cases}.$$

B.5.5 Mesures d'indépendance

Dans ce paragraphe, nous étudions les propriétés statistiques des matrices d'intercorrélation de signaux aléatoires indépendants stationnaires. L'objectif est d'obtenir des conditions nécessaires d'indépendance.

Proposition B.46. *Soit $\mathbf{u}(.) \in \mathcal{S}^n$ un vecteur de signaux aléatoires. Si $\mathbf{u}(.)$ est à composantes indépendantes alors,*

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau) \ \textit{est diagonale},$$

ce qui équivaut à

$$\forall k \in \mathbb{N}, \forall \tau \in \mathbb{R}, \mathbf{d}\left(\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)\right) = \mathbf{0}, \tag{B.30}$$

où $\mathbf{d}$ est une fonction qui associe, à une matrice $\mathbf{A}$ de taille $n \times n$, le vecteur constitué de tous les éléments non diagonaux de $\mathbf{A}$. La fonction $\mathbf{d}$ est définie comme suit :

$$\mathbf{d}(\mathbf{A}) = (a_{12},\ldots,a_{1n},a_{21},a_{23}\ldots,a_{2n},a_{n1}\ldots,a_{nn-1}). \tag{B.31}$$

Démonstration. D'après la proposition B.45 ii), si $\mathbf{u}(.)$ à des composantes indépendantes, alors toutes ses dérivées ont également des composantes indépendantes (et donc décorrélées). Donc,

$$\forall k \in \mathbb{N}, \forall i,j = 1\ldots n, i \neq j, E\left(u_i^{(k)}(t)\, u_j(t-\tau)\right) = 0.$$

Dès lors, $\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)$ est diagonale. □

B.6 Gaussianité

Dans cette partie, nous nous intéressons au concept de gaussianité. En effet, de même que pour les variables et les vecteurs aléatoires, la notion de gaussianité se généralise pour des signaux aléatoires.

De façon analogue au paragraphe B.5, notre approche de la gaussianité de signaux aléatoires n'est pas classique dans la mesure où elle est basée sur une définition originale.

B.6.1 Définition

Notons $\mathcal{L}$ l'ensemble des fonctions linéaires de $\mathbb{R}^{\mathbb{R}}$ dans $\mathbb{R}$. Par exemple,

$$f : \begin{cases} \mathbb{R}^{\mathbb{R}} & \rightarrow \quad \mathbb{R} \\ u(t) & \rightarrow \quad \alpha_1 u(10) + \alpha_2 \ddot{u}(0) \end{cases} \ \text{avec } \alpha_i \in \mathbb{R},$$

est une fonction appartenant à $\mathcal{L}$.

Définition B.47 (Signal aléatoire gaussien)**.** Un signal aléatoire $u(.)$ est gaussien si, $\forall f \in \mathcal{L}$, la variable aléatoire $f(u(.))$ est gaussienne.
On note $\mathcal{G}$ l'ensemble des signaux aléatoires gaussiens ($\mathcal{G} \subset \mathfrak{S}$).

Notation B.48. *On note $\mathcal{G}_s(u)$, la gaussianité d'un signal aléatoire $u(.)$. Ainsi définie, $\mathcal{G}_s$ est une relation unaire de $\mathfrak{S}$.*

Remarque B.49. *Comme nous l'avons précisé dans la remarque B.36, toute constante $\alpha \in \mathbb{R}$ peut être vue comme une variable aléatoire gaussienne (de moyenne α et d'écart-type 0). Par exemple, 1 est une variable aléatoire gaussienne qui à chaque réalisation associe 1.*
Ainsi, pour toute fonction constante

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R} \\ u(t) & \to & \alpha \end{array} \right. \textit{ avec } \alpha \textit{ indépendant de } u(t),$$

$f(u(.))$ est une variable aléatoire gaussienne.

Soit $\tilde{\mathcal{L}}$ l'ensemble des fonctions affines de $\mathbb{R}^{\mathbb{R}}$ dans $\mathbb{R}$. Nous entendons par fonction affine, la somme d'une constante (de $\mathbb{R}$) et d'une fonction linéaire de $\mathcal{L}$.

$$\tilde{\mathcal{L}} = \left\{ \begin{array}{l} f : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R} \\ u(t) & \to & \alpha + f_l(u(t)), \end{array} \right. \\ \text{avec } f_l \in \mathcal{L} \text{ et } \alpha \in \mathbb{R} \end{array} \right\}.$$

En conséquence de la remarque B.49, nous énonçons le résultat suivant.

Proposition B.50. *Un signal aléatoire $u(.)$ est gaussien si et seulement si $\forall f \in \tilde{\mathcal{L}}$, la variable aléatoire $f(u(.))$ est gaussienne.*

Pour clarifier ces résultats peu classiques, nous les illustrons par un exemple.

Exemple B.51. *Considérons un signal aléatoire $u(.)$ gaussien et blanc de moyenne nulle et d'écart type unitaire. Construisons la variable aléatoire X, extraite de u, définie par*

$$X = f(u(.)) = 5 + 2u(10) - 4u(3).$$

où $f \in \tilde{\mathcal{L}}$. L'histogramme en bâton B.7 montre la répartition de la variable aléatoire X pour 10000 réalisations de $u(.)$ On peut voir que cet histogramme est caractéristique d'une variable gaussienne.
Par définition, $u(.)$ est gaussien si toutes les variables X créées à partir de fonctions f de $\mathcal{L}$, aussi complexes soient-elles, sont gaussiennes.

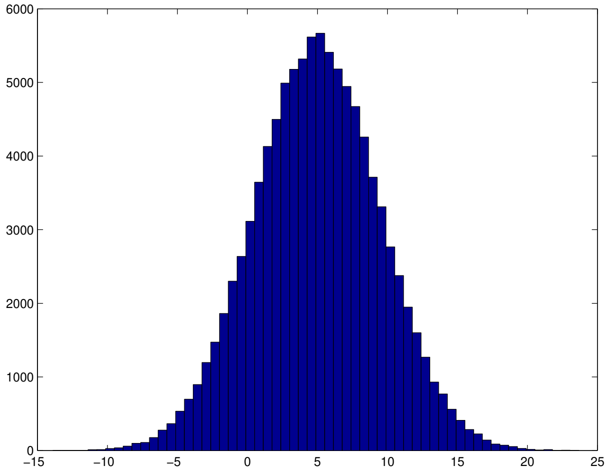


Figure B.7 – Répartition des réalisations de la variable aléatoire $X = 5 + 2u(10) - 4u(3)$.

B.6.2 Généralisation au cas de vecteur de signaux aléatoires

Maintenant, nous proposons de généraliser la notion de gaussianité à une famille de signaux aléatoires.

Considérons l'ensemble $\mathcal{L}^n, n \in \mathbb{N}$, des fonctions linéaires de $\left(\mathbb{R}^{\mathbb{R}}\right)^n$ à valeur dans $\mathbb{R}$. Par exemple,

$$g : \left\{ \begin{array}{lcl} \left(\mathbb{R}^{\mathbb{R}}\right)^2 & \rightarrow & \mathbb{R} \\ (u_1(t), u_2(t)) & \rightarrow & \alpha_1 u_1(0) + \alpha_2 \ddot{u}_2(10), \end{array} \right. \quad \text{avec } \alpha_i \in \mathbb{R},$$

est une fonction de $\mathcal{L}^2$. On définit de même l'ensemble $\tilde{\mathcal{L}}^n, n \in \mathbb{N}$, des fonctions affines de $\left(\mathbb{R}^{\mathbb{R}}\right)^n$ à valeur dans $\mathbb{R}$.

Définition B.52 (Vecteur de signaux aléatoires mutuellement gaussiens)**.** Un vecteur de n signaux aléatoires $\mathbf{u}(.) = (u_1(.), \ldots, u_n(.))^T$ est mutuellement gaussien (noté $\mathcal{G}_s(u_1, \ldots, u_n)$) si, $\forall f \in \mathcal{L}^n$, la variable aléatoire $f(u_1(.), \ldots, u_n(.))$ est gaussienne.
On note $\mathcal{G}^n$ l'ensemble des vecteurs de n signaux aléatoires mutuellement gaussiens ($\mathcal{G}^n \subset \mathfrak{S}^n$).

Par analogie avec la proposition B.50, si un vecteur de signaux aléatoires est mutuellement gaussien alors, pour toutes les fonctions affines de $\tilde{\mathcal{L}}^n$, la variable aléatoire $f(u(.))$ est gaussienne.

Remarque B.53. *Attention, de la même manière que pour la stationnarité (voir remarque B.19), deux signaux gaussiens ne sont pas mutuellement gaussien. En effet, on a*

$$\mathcal{G}^n \neq \underbrace{\mathcal{G} \times \cdots \times \mathcal{G}}_{n\text{-fois}}.$$

Plus précisément,

$$u_1(.), u_2(.) \in \mathcal{G} \nRightarrow (u_1(.), u_2(.)) \in \mathcal{G}^2.$$

Voir l'exemple A.84 d'un vecteur non gaussien à composantes gaussiennes.

B.6.3 Conservation de la gaussianité

Maintenant, afin de caractériser l'ensemble des fonctions g qui préservent le caractère gaussien d'une famille de signaux aléatoires, intéressons-nous à l'ensemble $\mathfrak{L}$ des fonctions affines définies de $\mathbb{R}^{\mathbb{R}}$ à valeur dans $\mathbb{R}^{\mathbb{R}}$. Plus formellement,

$$\mathfrak{L} = \left\{ \begin{array}{l} g : \left\{ \begin{array}{lcl} \mathbb{R}^{\mathbb{R}} & \longrightarrow & \mathbb{R}^{\mathbb{R}} \\ u\left(t\right) & \mapsto & \alpha\left(t\right)u\left(t\right) + \beta\left(t\right), \end{array} \right. \\ \text{avec } \alpha(t), \beta(t) \in \mathbb{R}^{\mathbb{R}}. \end{array} \right\}.$$

Au même titre que précédemment, les fonctions de $\mathfrak{L}$ sont appelées filtres. Par exemple, la fonction

$$g : \left\{ \begin{array}{lcl} \mathbb{R}^{\mathbb{R}} & \longrightarrow & \mathbb{R}^{\mathbb{R}} \\ u\left(t\right) & \mapsto & t^2 u\left(t\right) + \sin\left(t\right) \end{array} \right.$$

est un filtre appartenant à $\mathfrak{L}$.

Proposition B.54. *Soit* $\mathbf{u}\left(.\right) = \left(u_1\left(.\right), \ldots, u_n\left(.\right)\right)^T \in \mathfrak{S}^n$ *un vecteur de* n *signaux aléatoires.* $\mathfrak{L}$ *est compatible avec la relation* $\mathcal{G}_s$*, c'est-à-dire*

$$\mathcal{G}_s\left(u_1\left(.\right), \ldots, u_n\left(.\right)\right) \Leftrightarrow \forall g \in \mathfrak{L}, \mathcal{G}_s\left(g\left(u_1\left(.\right)\right), \ldots, g\left(u_n\left(.\right)\right)\right). \tag{B.32}$$

On a également,

$$\mathcal{G}_s\left(u_1\left(.\right), \ldots, u_n\left(.\right)\right) \Leftrightarrow \forall g_1, \ldots, g_n \in \mathfrak{L}, \mathcal{G}_s\left(g_1\left(u_1\left(.\right)\right), \ldots, g_n\left(u_n\left(.\right)\right)\right). \tag{B.33}$$

Démonstration. Montrons l'assertion (B.33). Soient $g_1, \ldots, g_n \in \mathfrak{L}$. Rappelons, d'après la définition B.52, que

$$\mathcal{G}_s\left(g_1\left(u_1\left(.\right)\right), \ldots, g_n\left(u_n\left(.\right)\right)\right) \Leftrightarrow \forall f \in \mathcal{L}^n, \mathcal{G}_v\left(f\left(g_1\left(u_1\left(.\right)\right)\right), \ldots, \left(g_n\left(u_n\left(.\right)\right)\right)\right). \tag{B.34}$$

Pour $f \in \mathcal{L}^n$, définissons la fonction h telle que

$$h : \left\{ \begin{array}{lcl} \left(\mathbb{R}^{\mathbb{R}}\right)^n & \rightarrow & \mathbb{R} \\ \mathbf{u}\left(t\right) & \rightarrow & f\left(g_1\left(u_1\left(t\right)\right), \ldots, g_n\left(u_n\left(t\right)\right)\right) \end{array} \right. . \tag{B.35}$$

Montrons qu'ainsi définie, la fonction h appartient bien à $\tilde{\mathcal{L}}^n$. Nous le démontrons pour une variable, cela se généralise sans problème pour n variables.

Soit $f \in \mathcal{L}$ et $g \in \mathfrak{L}$. On a $f = c + f_l$, avec f_l linéaire et $g\left(u\left(t\right)\right) = \alpha(t)u\left(t\right) + \beta(t)$ avec $\alpha(t), \beta(t) \in \mathbb{R}^{\mathbb{R}}$.

$$\begin{aligned} f\left(g\left(u\left(t\right)\right)\right) &= f\left(\alpha(t)u\left(t\right) + \beta(t)\right), \\ &= c + f_l\left(\alpha(t)u\left(t\right) + \beta(t)\right), \\ &= c + f_l\left(\alpha(t)u\left(t\right)\right) + f_l\left(\beta(t)\right), \\ &= C + F_l\left(u\left(t\right)\right) \end{aligned}$$

avec $C = c + f_l\left(\beta(t)\right) \in \mathbb{R}$ une constante et $F_l : u\left(t\right) \rightarrow f_l\left(\alpha(t)u\left(t\right)\right)$ linéaire. En effet, par linéarité de f_l, on a

$$\begin{aligned} F_l\left(\lambda u\left(t\right) + \mu y\left(t\right)\right) &= f_l\left(\alpha(t)\left(\lambda u\left(t\right) + \mu y\left(t\right)\right)\right) \\ &= \lambda f_l\left(\alpha(t)u\left(t\right)\right) + \mu f_l\left(\alpha(t)y\left(t\right)\right). \end{aligned}$$

On a donc $f \circ g \in \tilde{\mathcal{L}}$.

Nous venons de vérifier que la fonction h définie par (B.35) appartient bien à l'ensemble $\tilde{\mathcal{L}}^n$. Maintenant, sachant que $\mathcal{G}_s\left(u_1\left(.\right),\ldots,u_n\left(.\right)\right)$, et, d'après la proposition B.50, on a $\mathcal{G}_v\left(h\left(u_1\left(.\right),\ldots,u_n\left(.\right)\right)\right)$, c'est-à-dire $\mathcal{G}_v\left(f\left(g_1\left(u_1\left(.\right)\right),\ldots,g_n\left(u_n\left(.\right)\right)\right)\right)$.
Et, d'après (B.34), il s'en suit que $\mathcal{G}_s\left(g_1\left(u_1\left(.\right)\right),\ldots,g_n\left(u_n\left(.\right)\right)\right)$.

L'assertion (B.32) se montre trivialement en posant $g = g_1 = \cdots = g_n$. □

Exemple B.55. *Si un signal aléatoire $u\left(.\right)$ est gaussien, alors, le signal aléatoire $g\left(u\left(t\right)\right)$ (par exemple, $t^2 u\left(.\right) + \sin\left(t\right)$) est gaussien. La figure B.8 illustre la conservation de la gaussianité.*

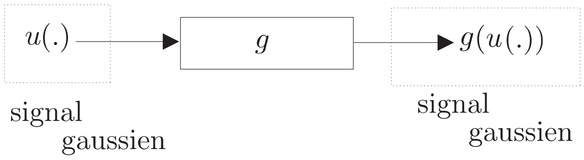


Figure B.8 – Conservation de la gaussianité pour un filtre $g \in \mathfrak{L}$.

Remarque B.56. *Par analogie avec la remarque B.49, un signal déterministe $u\left(t\right)$ de $\mathbb{R}^{\mathbb{R}}$ peut être vu comme un élément de $\mathfrak{S}$ $\left(\mathbb{R}^{\mathbb{R}} \subset \mathfrak{S}\right)$. De plus, $\forall f \in \mathcal{L}$, on a $f\left(u\left(t\right)\right) \in \mathbb{R}$. Sachant qu'un scalaire est gaussien (voir remarque B.49), on a $\mathcal{G}_v\left(f\left(u\left(t\right)\right)\right)$. En conséquence, tout signal déterministe $u\left(.\right)$ est gaussien. C'est un cas dégénéré de la gaussianité de signaux. Par exemple, le signal déterministe $\cos\left(t\right) \in \mathbb{R}^{\mathbb{R}}$ peut être vu comme un signal aléatoire gaussien.*
La proposition B.54 est bien valable pour toutes les filtres $g_i \in \mathfrak{L}$, même constants, bien qu'on ait pu supposer le contraire.

Intéressons-nous maintenant à la structure de $(\mathfrak{L}, \circ)$.

Proposition B.57. *L'ensemble $(\mathfrak{L}, \circ)$ est un monoïde dont l'élément neutre est le filtre identité, défini par*

$$Id : \left\{ \begin{array}{ccc} \mathbb{R}^{\mathbb{R}} & \rightarrow & \mathbb{R}^{\mathbb{R}} \\ u\left(t\right) & \rightarrow & u\left(t\right) \end{array} \right. .$$

Démonstration. $\circ$ est une loi interne :

$$\forall f, g \in \mathfrak{L},\ f \circ g \in \mathfrak{L}.$$

$\circ$ est associative :

$$\forall f, g, h \in \mathfrak{L},\ \left(f \circ g\right) \circ h = f \circ \left(g \circ h\right).$$

$\circ$ a un élément neutre :

$$\forall f \in \mathfrak{L}, Id \circ f = f \circ Id = f.$$

□

Remarque B.58. *L'ensemble $(\mathfrak{L}, \circ)$ n'est pas pas un semi-groupe car la loi interne $\circ$ ne commute pas :*

$$\exists f, g \in \mathfrak{L},\ f \circ g \neq g \circ f.$$

Voici quelques fonctions particulières de $\mathfrak{L}$ (filtres utilisés par la suite) préservant la gaussianité.

Proposition B.59. *Soit* $\mathbf{u}(.) = (u_1, \ldots, u_n)^T$ *un vecteur de n signaux aléatoires de* $\mathfrak{S}^n$. *On a*

$$i) \quad \mathcal{G}_s(u_1, \ldots, u_n) \implies \forall \{j_1, \ldots, j_n\} \subset \{1, \ldots, n\}, \mathcal{G}_s(u_{j_1}, \ldots, u_{j_n}).$$
$$ii) \quad \mathcal{G}_s(u_1, \ldots, u_n) \implies \forall (i_1, \ldots, i_n) \in \mathbb{N}^n, \mathcal{G}_s\left(u_1^{(i_1)}, \ldots, u_n^{(i_n)}\right).$$

Démonstration. i) Trivial

ii) Le filtre dérivation g_j défini par

$$g_j : \begin{cases} \mathbb{R}^{\mathbb{R}} & \to & \mathbb{R}^{\mathbb{R}} \\ u_j(t) & \to & u_j^{(i_j)}(t) \end{cases}$$

est un filtre linéaire de $\mathbb{R}^{\mathbb{R}} \to \mathbb{R}^{\mathbb{R}}$, il appartient donc à $\mathfrak{L}$. D'après la proposition B.39, le résultat est donc trivial. □

Par la suite, lorsqu'on parlera de n signaux aléatoires gaussiens (noté $\mathcal{G}_s(u_1, \ldots, u_n)$), il s'agira de signaux à composantes gaussiennes, à moins que le contraire ne soit explicitement précisé. Ainsi,

$$\mathcal{G}_s(u_1, \ldots, u_n) \Leftrightarrow \forall i, \ \mathcal{G}_s(u_i).$$

B.6.4 Mesure de Gaussianité

La mesure de la gaussianité d'un signal aléatoire stationnaire exploite des moments d'ordre supérieur. Les deux grandeurs suivantes sont généralement utilisées :

Définition B.60 (Skewness)**.** On appelle Skewness d'un signal aléatoire $u(.) \in \mathfrak{S}$, noté $Skew(u)$, la grandeur réelle suivante

$$Skew(u) = \frac{E\left((u(t) - E(u(t)))^3\right)}{E\left((u(t) - E(u(t)))^2\right)^{\frac{3}{2}}}.$$

Définition B.61 (Kurtosis)**.** On appelle Kurtosis d'un signal aléatoire $u(.) \in \mathfrak{S}$, noté $Kurt(u)$, la grandeur réelle suivante

$$Kurt(u) = \frac{E\left((u(t) - E(u(t)))^4\right)}{E\left((u(t) - E(u(t)))^2\right)^2} - 3.$$

Le skewness est une mesure de symétrie de la densité de probabilité, de la même façon que la variance caractérise l'écart à la moyenne. Ainsi, un signal aléatoire stationnaire et ergodique ayant une densité de probabilité symétrique (c'est le cas de loi gaussienne) aura un skewness nul. Au contraire, le skewness prendra des valeurs d'autant plus importantes que la densité de probabilité considérée sera "asymétrique".

Le Kurtosis, quant à lui, est une mesure "d'aplatissement" de la densité de probabilité. Le kurtosis vaut zéro pour une loi gaussienne. A titre d'exemple, on a tracé sur la figure B.9, les densités de probabilités d'une loi gaussienne, d'une loi uniforme et d'une loi de Laplace. Les valeurs du Kurtosis pour ces lois sont données dans le tableau suivant :

Loi	**Kurtosis**
Normale	0
Uniforme	-1.2
Laplace	3

Des remarques précédentes, on déduit aisément la proposition suivante :

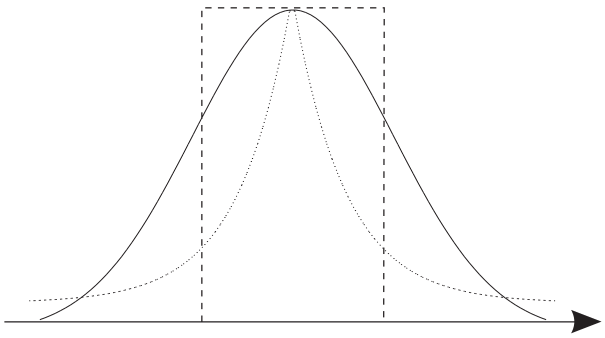


Figure B.9 – Densités des lois uniforme (_ _ _), gaussienne (___) et de Laplace (.....).

Proposition B.62. *Soit* $u(.) \in \mathfrak{S}$ *un signal aléatoire. On a*

$$\mathcal{G}_s(u) \implies \left(\forall i \in \mathbb{N}, \quad Kurt\left(u^{(i)}\right) = 0\right) . \tag{B.36}$$

Démonstration. D'après la proposition B.59 ii), on a $\mathcal{G}_s(u) \implies \mathcal{G}_s\left(u^{(i)}\right)$. Or, le kurtosis d'un signal gaussien est nul donc, $\forall i$, $Kurt\left(u^{(i)}\right) = 0$. □

B.7 Signaux aléatoires gaussiens et indépendants

L'espace des vecteurs de n signaux aléatoires gaussiens et à composantes indépendantes est noté $\mathcal{G}^n \cap I^n$ ($(\mathcal{G}^n \cap I^n) \subset \mathfrak{S}^n$).

Proposition B.63. *L'ensemble* $\mathcal{G}^n \cap I^n$ *des vecteurs de* n *signaux aléatoires gaussiens et indépendants est un espace vectoriel, c'est-à-dire*

$$\begin{aligned}&\forall \mathbf{u}_1(.) \in (\mathcal{G}^n \cap I^n), \forall \mathbf{u}_2(.) \in (\mathcal{G}^n \cap I^n) \text{ et } \forall \alpha_1, \alpha_2 \in \mathbb{R},\\ &\alpha_1 \mathbf{u}_1(.) + \alpha_2 \mathbf{u}_2(.) \in \mathcal{G}^n \cap I^n.\end{aligned} \tag{B.37}$$

Maintenant, intéressons-nous à une classe particulière de vecteurs à la fois gaussiens et indépendants. Nous avons choisi de les nommés signaux aléatoires gaussiens *circulaires*, bien que ce terme soit classiquement réservé à d'autres types de signaux, dans la mesure où aucune ambiguïté n'est possible.

Définition B.64 (Vecteurs de signaux aléatoires gaussiens circulaires)**.** Soit $\mathbf{u}(.) = (u_1(.), \ldots, u_n(.))^T \in \mathfrak{S}^n$ un vecteur de signaux aléatoires gaussiens, de matrices d'autocorrélations respectives $\Gamma_{u_i,u_i}(\tau)$. Le vecteur $\mathbf{u}(.)$ est appelé vecteur gaussien circulaire de corrélation $\Gamma(\tau)$ si il vérifie les deux conditions suivantes

$$\begin{aligned}&i) \quad I_s(u_1, \ldots, u_n),\\ &ii) \quad \forall i \in \{1, \ldots, n\}, \Gamma_{u_i,u_i}(\tau) = \Gamma(\tau).\end{aligned}$$

En d'autres termes, la matrice d'autocorrélation $\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ de $\mathbf{u}(.)$ est une matrice diagonale où tous les éléments diagonaux sont égaux, c'est-à-dire

$$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau) = \alpha \mathbf{Id}_n. \tag{B.38}$$

Théorème B.65. *Si* $\mathbf{u}(.)$ *est un vecteur de signaux aléatoires gaussiens circulaires de corrélation* $\Gamma(\tau)$ *alors, pour toute matrice unitaire* $\mathbf{U}$*, le vecteur* $\mathbf{y}(.) = \mathbf{U}\mathbf{u}(.)$ *est également circulaire gaussien de corrélation* $\Gamma(\tau)$.

Démonstration. On a

$$\begin{aligned}\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}(\tau) &= E\left(\mathbf{y}(t)\mathbf{y}(t-\tau)^{\mathrm{T}}\right)\\ &= E\left(\mathbf{U}\mathbf{u}(t)\mathbf{u}(t-\tau)^{T}\mathbf{U}^{T}\right)\\ &= \mathbf{U}\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)\,\mathbf{U}^{T}.\end{aligned}$$

$\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)$ étant diagonale et d'après l'égalité B.38, on a donc

$$\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}(\tau) = \mathbf{U}\mathbf{U}^{\mathbf{T}}\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau) = \mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}(\tau)\,.$$

□

Remarque B.66. *D'après le théorème B.65, les statistiques du second ordre de* $\mathbf{y}(.)$ *sont identiques à celles de* $\mathbf{u}(.)$ *. Sachant que* $\mathbf{u}(.)$ *est gaussien (i.e. ne comporte que des statistiques d'ordre deux), toutes les propriétés statistiques de* $\mathbf{u}(.)$ *et* $\mathbf{y}(.)$ *sont identiques.*

B.8 Représentation systémique des signaux

Maintenant que les principales notions concernant les signaux aléatoires ont été introduites, nous proposons de nous intéresser, un court paragraphe, à l'étude des systèmes entrée-sortie (voir figure B.10).

Figure B.10 – Représentation d'un système d'entrée $u(t)$ et de sortie $y(t)$.

Le but de ce dernier paragraphe est l'obtention de relations entre les statistiques des signaux d'entrées et de sorties, ceci en fonction des caractéristiques du système considéré. Par exemple, nous avons remarqué précédemment que si le système s'exprime comme une fonction de $\mathfrak{F}$, alors l'indépendance des entrées est conservée pour les sorties.

Nous nous intéressons uniquement aux statistiques d'ordre deux, à savoir les relations entre les matrices de corrélation (autocorrélation et intercorrélation) d'entrées et de sorties. De plus, seuls les systèmes linéaires et invariants seront présentés.

Pour commencer, nous exposons les deux principales représentations des systèmes dynamiques, puis nous établissons les relations entre les corrélations.

Pour davantage de détails, le lecteur pourra se référer aux ouvrages [38] et [8].

B.8.1 Représentation entrée-sortie et représentation d'état

Intéressons nous au cas d'un système dynamique (linéaire et invariant) décrit par une équation différentielle linéaire à coefficients constants.

Définition B.67 (Système linéaire d'ordre N)**.** On appelle système linéaire d'ordre N, tout système décrit par une équation différentielle de la forme

$$\sum_{k=0}^{N}\mathbf{A}_k\mathbf{y}^{(k)}(t) = \sum_{k=0}^{M}\mathbf{B}_k\mathbf{u}^{(k)}(t)\,, \tag{B.39}$$

où $\mathbf{y}(t)$ est la sortie à l'instant t, $\mathbf{u}(t)$ est l'entrée à l'instant t.

La représentation (B.39) est nommée *représentation entrée-sortie*. Dans ce cas, le système est spécifié par une loi qui fait intervenir l'entrée $\mathbf{u}(t)$, la sortie $\mathbf{y}(t)$ ainsi qu'un certain nombre de leurs dérivées. Cependant, les systèmes différentiels peuvent être représentés sous une forme alternative (lorsque $M \leq N$), appelée *représentation d'état*, et donnée par les équations de la forme

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{B.40}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t), \tag{B.41}$$

où l'équation (B.40) est appelée l'*équation d'évolution* et (B.41) l'*équation de sortie*. Cette fois le système est spécifié au moyen d'un signal auxiliaire $\mathbf{x}(t)$, appelé vecteur d'état, lequel vérifie une équation différentielle linéaire à coefficients constants du premier ordre. Dans cette forme générale, le vecteur d'entrée peut être un vecteur de m composantes $u_1, \ldots, u_m$ et le signal de sorties un vecteur de p composantes $y_1, \ldots, y_p$. Dans ce cas, les dimensions des matrices sont $\mathbf{A}(n \times n), \mathbf{B}(n \times m)$ et $\mathbf{C}(p \times n)$.

Toute représentation d'état linéaire peut facilement être mise sous la forme entrée-sortie (on parle de matrice de transfert), et inversement (on parle de forme canonique de contrôlabilité).

B.8.2 Corrélation des sorties

Intéressons-nous à l'expression de la matrice d'autocorrélation des sorties connaissant celles des entrées. Nous exploitons ici les résultats sur les dérivées de signaux aléatoires. Ce que nous proposons est, à notre connaissance, nouveau.

Proposition B.68. *Considérons le modèle d'état suivant*

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \tag{B.42}$$

où $\mathbf{A}$ *est une matrice inversible. On a les relations de récurrence suivantes entre les matrices de corrélation des signaux :*

$$\mathbf{\Gamma}^{(k)}_{\mathbf{x},\mathbf{x}}(\tau) = \mathbf{A}^{-1}\left(-\mathbf{\Gamma}^{(k+2)}_{\mathbf{x},\mathbf{x}}(\tau) + (-1)^{k+1}\mathbf{B}\mathbf{\Gamma}^{(k+1)}_{\mathbf{x},\mathbf{u}}(-\tau)^T - \mathbf{\Gamma}^{(k+1)}_{\mathbf{x},\mathbf{u}}(\tau)\mathbf{B}^T + \mathbf{B}\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)\mathbf{B}^T\right)\mathbf{A}^{-1T},$$

$$\mathbf{\Gamma}^{(k)}_{\mathbf{x},\mathbf{u}}(\tau) = \mathbf{A}^{-1}\left(\mathbf{\Gamma}^{(k+1)}_{\mathbf{x},\mathbf{u}}(\tau) - \mathbf{B}\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)\right),$$

$$\mathbf{\Gamma}^{(k)}_{\mathbf{y},\mathbf{y}}(\tau) = \mathbf{C}\mathbf{\Gamma}^{(k)}_{\mathbf{x},\mathbf{x}}(\tau)\mathbf{C}^T.$$

Démonstration. Considérons le système dynamique

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{B.43}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t). \tag{B.44}$$

D'après la première équation, $\mathbf{A}$ étant inversible, on obtient

$$\mathbf{x}(t) = \mathbf{A}^{-1}\dot{\mathbf{x}}(t) - \mathbf{A}^{-1}\mathbf{B}\mathbf{u}(t),$$

et, après k dérivations, on a

$$\mathbf{x}^{(k)}(t) = \mathbf{A}^{-1}\mathbf{x}^{(k+1)}(t) - \mathbf{A}^{-1}\mathbf{B}\mathbf{u}^{(k)}(t).$$

D'après l'équation (B.11), on a

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{x}^{(k)},\mathbf{x}}(\tau) &= E(\mathbf{x}^{(k)}(t)\,\mathbf{x}(t-\tau)^{\mathrm{T}}) \\
&= E\left(\left(\mathbf{A}^{-1}\mathbf{x}^{(k+1)}(t) - \mathbf{A}^{-1}\mathbf{B}\mathbf{u}^{(k)}(t)\right)\left(\dot{\mathbf{x}}(t-\tau)^{\mathrm{T}}\mathbf{A}^{-1\mathrm{T}} - \mathbf{u}(t-\tau)^{\mathrm{T}}\mathbf{A}^{-1\mathrm{T}}\mathbf{B}^{\mathrm{T}}\right)\right) \\
&= \mathbf{A}^{-1}E\left(\left(\mathbf{x}^{(k+1)}(t) - \mathbf{B}\mathbf{u}^{(k)}(t)\right)\left(\dot{\mathbf{x}}(t-\tau)^{\mathrm{T}} - \mathbf{u}(t-\tau)^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\right)\right)\mathbf{A}^{-1\mathrm{T}} \\
&= \mathbf{A}^{-1}\left(\mathbf{\Gamma}_{\mathbf{x}^{(k+1)},\dot{\mathbf{x}}}(\tau) - \mathbf{B}\mathbf{\Gamma}_{\mathbf{u}^{(k)},\dot{\mathbf{x}}}(\tau) - \mathbf{\Gamma}_{\mathbf{x}^{(k+1)},\mathbf{u}}(\tau)\mathbf{B}^{\mathrm{T}} + \mathbf{B}\mathbf{\Gamma}_{\mathbf{u}^{(\mathbf{k})},\mathbf{u}}(\tau)\mathbf{B}^{\mathrm{T}}\right)\mathbf{A}^{-1\mathrm{T}}.
\end{aligned}$$

En utilisant les résultats de la proposition B.28, nous pouvons simplifier l'expression précédente

$$\mathbf{\Gamma}^{(k)}_{\mathbf{x},\mathbf{x}}(\tau) = \mathbf{A}^{-1}\left(-\mathbf{\Gamma}^{(k+2)}_{\mathbf{x},\mathbf{x}}(\tau) + (-1)^{k+1}\mathbf{B}\mathbf{\Gamma}^{(k+1)}_{\mathbf{x},\mathbf{u}}(-\tau)^{\mathrm{T}} - \mathbf{\Gamma}^{(k+1)}_{\mathbf{x},\mathbf{u}}(\tau)\mathbf{B}^{\mathrm{T}} + \mathbf{B}\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)\mathbf{B}^{\mathrm{T}}\right)\mathbf{A}^{-1\mathrm{T}}.$$

Par ailleurs,

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{x}^{(k)},\mathbf{u}}(\tau) &= E\left(\mathbf{x}^{(k)}(t)\,\mathbf{u}(t-\tau)^{\mathrm{T}}\right) \\
&= E\left(\mathbf{A}^{-1}\left(\mathbf{x}^{(k+1)}(t)\,\mathbf{u}(t-\tau)^{\mathrm{T}} - \mathbf{B}\mathbf{u}^{(k)}(t)\,\mathbf{u}(t-\tau)^{\mathrm{T}}\right)\right) \\
&= \mathbf{A}^{-1}\left(\mathbf{\Gamma}_{\mathbf{x}^{(k+1)},\mathbf{u}}(\tau) - \mathbf{B}\mathbf{\Gamma}_{\mathbf{u}^{(k)},\mathbf{u}}(\tau)\right).
\end{aligned}$$

Donc

$$\mathbf{\Gamma}^{(k)}_{\mathbf{x},\mathbf{u}}(\tau) = \mathbf{A}^{-1}\left(\mathbf{\Gamma}^{(k+1)}_{\mathbf{x},\mathbf{u}}(\tau) - \mathbf{B}\mathbf{\Gamma}^{(k)}_{\mathbf{u},\mathbf{u}}(\tau)\right).$$

Enfin, on a

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{y}^{(k)},\mathbf{y}}(\tau) &= E\left(\mathbf{y}^{(k)}(t)\,\mathbf{y}(t-\tau)^{\mathrm{T}}\right) \\
&= \mathbf{C}\mathbf{\Gamma}_{\mathbf{x}^{(k)},\mathbf{x}}(\tau)\,\mathbf{C}^{\mathrm{T}},
\end{aligned}$$

d'où

$$\mathbf{\Gamma}^{(k)}_{\mathbf{y},\mathbf{y}}(\tau) = \mathbf{C}\mathbf{\Gamma}^{(k)}_{\mathbf{x},\mathbf{x}}(\tau)\,\mathbf{C}^{\mathrm{T}}.$$

□

Les relations de récurrence de la proposition B.68 nous permettent, connaissant la matrice d'autocorrélation de $\mathbf{u}(t)$ et ses dérivées, de calculer les dérivées successives des matrices d'autocorrélation de $\mathbf{x}(t)$ et $\mathbf{y}(t)$. En effet, lorsque les matrices $\mathbf{A}$, $\mathbf{B}$ et $\mathbf{C}$ sont connues, les trois relations de la proposition B.68 traduisent des dépendances linéaires entre les coefficients des matrices $\mathbf{\Gamma}_{\mathbf{x}\mathbf{x}}(\tau)$, $\mathbf{\Gamma}_{\mathbf{x},\mathbf{u}}(\tau)$, $\mathbf{\Gamma}_{\mathbf{y},\mathbf{y}}(\tau)$, ainsi que leurs dérivées.

Nous allons maintenant nous intéresser à un exemple simple, afin de mettre en évidence le processus de récurrence.

Exemple B.69. *Considérons le système dynamique linéaire suivant*

$$\mathbf{u}(t) = \mathbf{A}_0\mathbf{y}(t) + \mathbf{A}_1\dot{\mathbf{y}}(t),$$

où les matrices $\mathbf{A}_0$ et $\mathbf{A}_1$ sont supposées inversibles.
Ce système admet la représentation d'état de la forme suivante

$$\dot{\mathbf{y}}(t) = -\mathbf{A}_1^{-1}\mathbf{A}_0\mathbf{y}(t) + \mathbf{A}_1^{-1}\mathbf{u}(t).$$

Notons qu'il correspond au cas particulier où $\mathbf{A} = -\mathbf{A}_1^{-1}\mathbf{A}_0, \mathbf{B} = \mathbf{A}_1^{-1}$*et* $\mathbf{C} = \mathbf{Id}$ *dans l'expression (B.42).*
La sortie $\mathbf{y}(t)$ *s'écrit*

$$\mathbf{y}(t) = \mathbf{A}_0^{-1}\mathbf{u}(t) - \mathbf{A}_0^{-1}\mathbf{A}_1\dot{\mathbf{y}}(t).$$

Après k *dérivations, cette expression devient :*

$$\mathbf{y}^{(k)}(t) = \mathbf{A}_0^{-1}\mathbf{u}^{(k)}(t) - \mathbf{A}_0^{-1}\mathbf{A}_1\mathbf{y}^{(k+1)}(t).$$

On obtient une relation entre les dérivées successives de $\mathbf{y}(t)$ *et celles de* $\mathbf{u}(t)$*. On en déduit, par récurrence, une expression de* $\mathbf{y}(t)$ *en fontion des dérivées successives de* $\mathbf{y}(t)$ *et* $\mathbf{u}(t)$ *:*

$$\begin{aligned}
\mathbf{y}(t) &= \mathbf{A}_0^{-1}\mathbf{u}(t) - \mathbf{A}_0^{-1}\mathbf{A}_1\dot{\mathbf{y}}(t) \\
\mathbf{y}(t) &= \mathbf{A}_0^{-1}\mathbf{u}(t) - \mathbf{A}_0^{-1}\mathbf{A}_1\mathbf{A}_0^{-1}\dot{\mathbf{u}}(t) - \left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^2\ddot{\mathbf{y}}(t), \\
&\vdots \\
\mathbf{y}(t) &= \sum_{i=0}^{k}(-1)^i\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{u}^{(i)}(t) + (-1)^{k+1}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^{k+1}\mathbf{y}^{(k+1)}(t).
\end{aligned}$$

Supposons qu'à partir d'un certain rang K *très grand (par exemple* $K = 100$*), on ait*

$$\forall k \geq K, \mathbf{y}^{(k)}(t) = \mathbf{0}.$$

Cette hypothèse revient à considérer un système stable, c'est-à-dire

$$\lim_{n\to\infty}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^{k+1} = \mathbf{0}.$$

L'expression des sorties $\mathbf{y}(t)$ *du système peut donc s'écrire uniquement en fontion des dérivées successives des entrées* $\mathbf{u}(t)$*. On a*

$$\mathbf{y}(t) = \sum_{i=0}^{k}(-1)^i\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{u}^{(i)}(t), \text{ pour } k \geq K. \tag{B.45}$$

On peut en déduire une expression de la l*-ième dérivée des sorties en fontion des dérivées successives des entrées* $\mathbf{u}(t)$*.*

$$\mathbf{y}^{(l)}(t) = \sum_{i=0}^{k}(-1)^i\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{u}^{(i+l)}(t), \text{ pour } k \geq K. \tag{B.46}$$

Considérons maintenant les statistiques d'ordre deux des sorties. D'après les égalités (B.45) et (B.46), on a

$$\mathbf{\Gamma}_{\mathbf{y}^{(l)},\mathbf{y}}(\tau) = E\left(\mathbf{y}^{(l)}(t)\,\mathbf{y}(t-\tau)^T\right)$$

qui devient

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{y}^{(l)},\mathbf{y}}(\tau) &= E\left(\sum_{i=0}^{k}(-1)^i\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{u}^{(i)}(t)\,.\left(\sum_{j=0}^{k}(-1)^j\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^j\mathbf{A}_0^{-1}\mathbf{u}^{(j+l)}(t-\tau)\right)^T\right)\\
&= E\left(\sum_{i=0}^{k}(-1)^i\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{u}^{(i)}(t)\,.\left(\sum_{j=0}^{k}(-1)^j\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^j\mathbf{A}_0^{-1}\mathbf{u}^{(j+l)}(t-\tau)\right)^T\right)\\
&= E\left(\sum_{i=0}^{k}\sum_{j=0}^{k}(-1)^{i+j}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{u}^{(i)}(t)\,.\mathbf{u}^{(j+l)}(t+\tau)^T\mathbf{A}_0^{-1T}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^{jT}\right).
\end{aligned}$$

Par linéarité de l'espérance, on obtient

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{y}^{(l)},\mathbf{y}}(\tau) &= \sum_{i,j=0}^{k}(-1)^{i+j}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}E\left(\mathbf{u}^{(i)}(t)\,.\mathbf{u}^{(j+l)}(t+\tau)^T\right)\mathbf{A}_0^{-1T}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^{jT}\\
&= \sum_{i,j=0}^{k}(-1)^{i+j}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{\Gamma}_{\mathbf{u}^{(i)},\mathbf{u}^{(j+l)}}(\tau)\,\mathbf{A}_0^{-1T}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^{jT}.
\end{aligned}$$

On en déduit, d'après la proposition B.28

$$\mathbf{\Gamma}_{\mathbf{y}^{(l)},\mathbf{y}}(\tau) = \sum_{i,j=0}^{k}(-1)^{i+l}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^i\mathbf{A}_0^{-1}\mathbf{\Gamma}_{\mathbf{u},\mathbf{u}}^{(i+j+l)}(\tau)\,\mathbf{A}_0^{-1T}\left(\mathbf{A}_0^{-1}\mathbf{A}_1\right)^{jT}.$$

Cette dernière relation permet bien le calcul de l'autocorrélation des sorties en fonction de l'autocorrélation des entrées.

B.9 conclusion

Dans cette annexe, nous avons présenté la notion de signaux aléatoires de façon originale en partie grâce à l'introduction de l'ensemble $\mathbb{R}^{\mathbb{R}}$. Nous nous sommes intéressés notamment à l'indépendance et la gaussianité. De plus, nous avons exposé quelques résultats peu connus et peu utilisés sur les dérivées de signaux aléatoires qui sont exploités dans cette thèse.
Cette annexe nous paraît importante dans la mesure où elle permet une bonne compréhension des concepts de signaux aléatoires, de dérivées de signaux aléatoires indispensable à la lecture de ce mémoire.

ANNEXE C

Méthode de Newton-Raphson

Les méthodes itératives, et particulièrement la méthode de Newton, figurent parmi les méthodes numériques puissantes permettant la résolution approchée d'équations de toute nature. La méthode de Newton, découverte par Isaac Newton, a été publiée dans *Method of Fluxions* en 1736. L'idée de cette méthode est de partir d'une valeur approchée grossière de la solution, et d'en améliorer la précision par une itération d'un algorithme bien choisi.

Cette annexe a été réalisée à partir de [25] et [36].

C.1 Problème

Soit $\mathbf{f}$ une application à valeurs dans $\mathbb{R}^n$ définie sur un ouvert $\Omega \subset \mathbb{R}^n$ de classe $\mathcal{C}^2$. Nous nous intéressons à la résolution de l'équation suivante

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \Leftrightarrow \begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ f_2(x_1, \ldots, x_n) = 0 \\ \vdots \\ f_n(x_1, \ldots, x_n) = 0 \end{cases}$$

où les applications $f_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq n$, sont connues. On cherche à évaluer numériquement une racine $\mathbf{a}$ de $\mathbf{f}$, connaissant une valeur approchée grossière $\mathbf{x}_0$ de $\mathbf{a}$.

C.2 Principe de la méthode de Newton-Raphson.

L'idée est d'approximer $\mathbf{f}$ par sa partie linéaire au point $\mathbf{x}_0$. En effectuant un développement limité de la fonction $\mathbf{f}$, en un point $\mathbf{x}_0$, on obtient

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0) + \cdots + \frac{\mathbf{f}^{(n)}(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0)^n}{n!} + \circ(\mathbf{x}^n),$$

où

$$\mathbf{f}'(\mathbf{x}_0) = \left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \cdots \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \ddots & \\ \vdots & & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}\right|_{\mathbf{x} = \mathbf{x}_0}$$

En négligeant les termes non linéaires, cela donne

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\left(\mathbf{x} - \mathbf{x}_0\right).$$

Une estimation de la racine de $\mathbf{f}$ peut être obtenue en résolvant l'équation

$$\mathbf{f}(\mathbf{x}_0) + \mathbf{f}'(\mathbf{x}_0)\left(\mathbf{x} - \mathbf{x}_0\right) = 0. \tag{C.1}$$

Dans le cas où $\mathbf{f}'(\mathbf{x}_0)$ est inversible, on obtient une solution unique $\mathbf{x}_1$ qui vérifie

$$\mathbf{x}_1 = \mathbf{x}_0 - \mathbf{f}'(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_0).$$

La méthode de Newton consiste à itérer ce procédé. On définit ainsi une suite $(\mathbf{x}_n)_{n\in N}$, dont le terme général vérifie

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{f}'(\mathbf{x}_k)^{-1}\mathbf{f}(\mathbf{x}_k).$$

On voit immédiatement que chaque itération nous amène à inverser la matrice $\mathbf{f}'(\mathbf{x}_k)^{-1}$, c'est-à-dire à résoudre un système linéaire. Il est fondamental que la matrice des dérivées de $\mathbf{f}$ ne s'annule pas afin que le processus converge.

C.2.1 Interprétation géométrique

La méthode consiste à construire une suite de points $(\mathbf{x}_n)_{n\in N}$ qui converge (si possible) vers une solution du système.
Intéressons-nous au cas particulier d'une fonction $f : \Omega \subset \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^2$. Géométriquement, x_{k+1} est obtenu en intersectant la tangente en M$(x_k, f(x_k))$ à la courbe $\mathcal{C}_f$ représentative de f avec l'axe des abscisses (voir figure C.1).

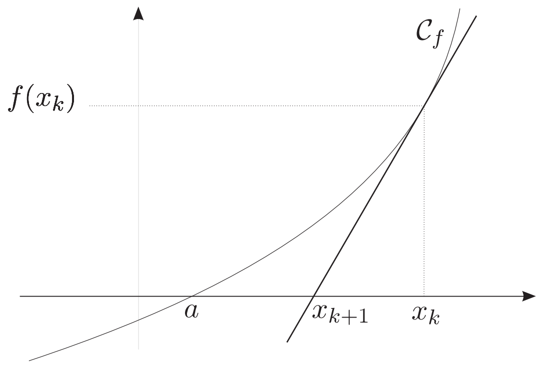


Figure C.1 – Point de vu géométrique - Méthode de Newton.

Dans le cas particulier où l'application $\mathbf{f}$ est affine : $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{A} \in \mathbb{R}^{n\times n}$ et $\mathbf{b} \in \mathbb{R}^n$, la première itération (C.1) se réduit à résoudre le système linéaire $\mathbf{A}\mathbf{x} + \mathbf{b} = \mathbf{0}$. Autrement dit, la méthode converge en une seule itération, ce qui était naturellement prévisible puisqu'une application affine est confondue en tout point avec son plan tangent.

C.2.2 Convergence de la méthode

Nous présentons dans ce paragraphe la convergence de la méthode de Newton pour le cas d'une fonction f définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. La généralisation pour $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ ne présente pas de difficulté.
Si f est une application de classe $\mathcal{C}^2$ et que la matrice des dérivées de f ne s'annule pas, on a une méthode qui converge de manière quadratique. En effet, le développement limité de f au deuxième ordre au point x_i (ce qui suppose bien sûr que la fonction f soit deux fois dérivable au voisinage de la racine) donne

$$f(x) = f(x_i) + f'(x_i)\,(x - x_i) + \frac{f''(x_i)\,(x - x_i)^2}{2}.$$

Et, pour $x = a$, on a

$$0 = f(x_i) + f'(x_i)\,(a - x_i) + \frac{f''(x_i)\,(a - x_i)^2}{2}.$$

Or, la relation de récurrence entre deux estimations successives étant

$$x_{i+1} = x_i - f'(x_i)^{-1} f(x_i),$$

on obtient,

$$x_{i+1} - a = \frac{f'(x_i)^{-1} f''(x_i)}{2}\,(x_i - a)^2\,.$$

On a donc bien une convergence en moyenne quadratique de la suite $(x_n)_{n\in N}$ vers la solution a.

Remarque C.1. *Notons que si la dérivée de la fonction n'est pas connue analytiquement, l'évaluation numérique de sa dérivée est toutefois possible par une formule d'accroissement*

$$f'(x) \simeq \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Dans ce cas, la convergence est toutefois moins rapide que la convergence quadratique.

C.2.3 Inconvénients

Dans la pratique, il est couteux de calculer, à chaque itération, les éléments de la nouvelle matrice $\mathbf{f}'(\mathbf{x}_k)^{-1}$. On peut supposer, dans la mesure où la méthode est convergente, que les vecteurs $\mathbf{x}_k$ consécutifs différent peu, de même que les matrices $\mathbf{f}'(\mathbf{x}_k)^{-1}$ correspondantes. Ces considérations conduisent naturellement à des variantes de la méthode de Newton.

C.3 Variantes de la méthode de Newton

Une variante de la méthode consiste à conserver la même matrice $\mathbf{f}'(\mathbf{x}_k)^{-1}$ pendant p itérations consécutives (p entier supérieur ou égal à 2 fixé). Ainsi, on construit la suite $(\mathbf{x}_n)_{n\in N}$ de la manière suivante :

$$\begin{array}{ll}
\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{f}'(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_k), & 0 \leq k \leq p-1 \\
\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{f}'(\mathbf{x}_p)^{-1}\mathbf{f}(\mathbf{x}_k), & p \leq k \leq 2p-1 \\
\vdots & \\
\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{f}'(\mathbf{x}_{rp})^{-1}\mathbf{f}(\mathbf{x}_k), & rp \leq k \leq (r+1)\,p-1 \\
\vdots &
\end{array}$$

Nous pouvons également conserver toujours la même matrice, ce qui conduit à des itérations du type

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{f}'(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_k),\ \forall k \geq 0,$$

ou même, remplacer la matrice $\mathbf{f}'(\mathbf{x}_0)$ par une matrice inversible $\mathbf{A}_0$ particulièrement simple à inverser :

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{A}_0^{-1}\mathbf{f}(\mathbf{x}_k),\ \forall k \geq 0.$$

Dans le cas des fonction $f : \Omega \subset \mathbb{R} \to \mathbb{R}$, on est facilement convaincu (voir figure C.2) que la convergence peut être obtenue si A_0 est suffisamment voisin de $\mathbf{f}'(\mathbf{x}_0)^{-1}$.

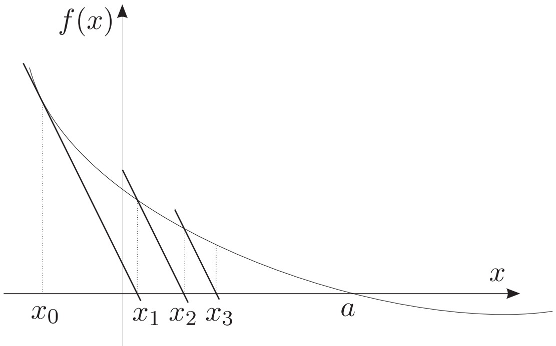


Figure C.2 – Convergence de la méthode dans le cas où A_0 est proche de $\mathbf{f}'(\mathbf{x}_0)^{-1}$.

C.4 Exemples et limites

Dans un premier temps, nous considérons un exemple illustrant l'algorithme de Newton. Puis, nous verrons les limites de la méthode de Newton que nous illustrerons graphiquement.

Exemple C.2. *Considérons le système* (S) *suivant :*

$$(S)\left\{\begin{array}{l} x_1^2 + x_1x_2 - 2x_2^2 = 4 \\ x_1 \exp x_1 + x_2 \exp x_2 = 0. \end{array}\right.$$

Il est de la forme

$$\left\{\begin{array}{l} f_1(x_1, x_2) = 0 \\ f_2(x_1, x_2) = 0 \end{array}\right.$$

où f_1 *et* f_2 *sont les fonctions définies par*

$$f_1 : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x_1, x_2) & \longrightarrow & x_1^2 + x_1x_2 - 2x_2^2 - 4, \end{array}$$
$$f_2 : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x_1, x_2) & \longrightarrow & x_1 \exp x_1 + x_2 \exp x_2. \end{array}$$

Nous allons résoudre le système (S) *à l'aide de la méthode de Newton-Raphson. Pour cela, il nous faut une valeur approchée de chacune des solutions. Nous commençons donc par tracer les courbes représentatives* $(\mathcal{C}_1)$ *et* $(\mathcal{C}_2)$ *des fonctions* f_1 *et* f_2*. Sur la figure C.3, on voit que le système* (S) *admet une unique solution* $(a_1, a_2) \in \mathbb{R}^2$*, proche de* $(-2, 0)$*.*

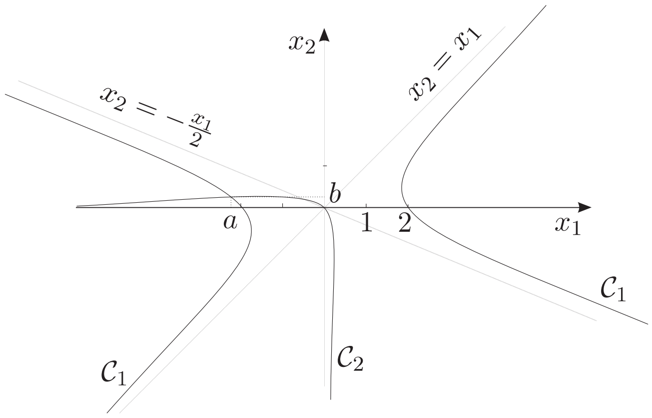


Figure C.3 – Courbes représentatives des fonctions f_1 et f_2.

Nous allons chercher une valeur approchée plus précise de cette solution à l'aide de l'algorithme de Newton. Nous commençons par calculer la matrice des dérivées de $\mathbf{f} = (f_1, f_2)\ldots$

$$\mathbf{f}'(\mathbf{x}) = \begin{pmatrix} 2x_1 + x_2 & x_1 - 4x_2 \\ (x_1+1)\exp x_1 & (x_2+2)\exp x_2 \end{pmatrix}.$$

Cette matrice est inversible au voisinage de (a_1, a_2)*. Son déterminant* $|\mathbf{f}'(\mathbf{x})|$ *vaut*

$$\left|\mathbf{f}'(\mathbf{x})\right| = (x_2+1)(2x_1+x_2)\exp x_2 - (x_1+1)(x_1-4x_2)\exp x_1$$

On a

$$\mathbf{f}'(\mathbf{x})^{-1} = \frac{1}{|\mathbf{f}'(\mathbf{x})|}\begin{pmatrix} (x_2+1)\exp x_2 & -x_1+4x_2 \\ -(x_1+1)\exp x_1 & 2x_1+x_2 \end{pmatrix}.$$

Considérons la suite $(\mathbf{x}_n)_{n\in N}$ *telle que*

$$\begin{cases} \mathbf{x}_0 = (-2, 0) \\ \mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{f}'(\mathbf{x}_k)^{-1}\mathbf{f}(\mathbf{x}_k). \end{cases}$$

Les itérations successives nous donnent :

$$\begin{aligned} &\mathbf{x}_0 = (-2, 0) \\ &\mathbf{x}_1 \simeq (-2.126\,757\,877, 0.253\,515\,753\,3) \\ &\mathbf{x}_2 \simeq (-2.127\,332\,227, 0.208\,205\,741\,5) \\ &\mathbf{x}_3 \simeq (-2.126\,932\,767, 0.206\,281\,505\,3) \\ &\mathbf{x}_4 \simeq (-2.126\,932\,304, 0.206\,278\,155\,6) \\ &\ldots \end{aligned}$$

On voit bien que le processus converge. La solution du système (S) *est* $(a_1, a_2) \simeq (-2.126\,932\,304, 0.206\,278\,155\,6)$. *On peut continuer pour obtenir une valeur plus précise de la solution* (a_1, a_2) *du système* (S).

Dans l'exemple C.2, on obtient de façon rapide une solution du système (S) mais cela vient de la connaissance d'une valeur initiale $\mathbf{x}_0$ proche de la solution du système. La principale difficulté dans la résolution d'équations non linéaires avec cette méthode réside essentiellement dans le choix d'un "bon" point initial $\mathbf{x}_0$, suffisamment voisin d'un zéro de la fonction $\mathbf{f}$. Ceci est très contraignant étant donné qu'en principe on ignore où se trouve les zéros d'une application, le problème étant précisement de les localiser!

En conclusion, un premier inconvénient de la méthode de Newton Raphson est qu'il faut connaître *a priori* une valeur approchée de chacune des solutions du système et donc, de surcroit le nombre de solutions de ce système.

Un second inconvénient est le calcul de la dérivée de la fonction $\mathbf{f}$. Si on a une expression analytique de cette fonction comme dans l'exemple C.2, cela ne pose pas de problème. Cependant, dans la pratique, ce n'est pas toujours le cas et il faut alors évaluer numériquement la dérivée de la fonction $\mathbf{f}$ (voir remarque C.1), ce qui peut s'avérer très délicat. Une autre limite de la méthode est qu'elle ne s'applique plus si on se trouve en un point de dérivée nulle. Numériquement, ce problème apparaît dès qu'une dérivée $\mathbf{f}'(\mathbf{x}_k)$ est voisine de zéro car la nouvelle valeur $\mathbf{x}_{k+1}$ calculée est alors très grande et le point d'intersection de la tangente et de l'axe des abscisses est rejeté à l'infini (voir figure C.4).

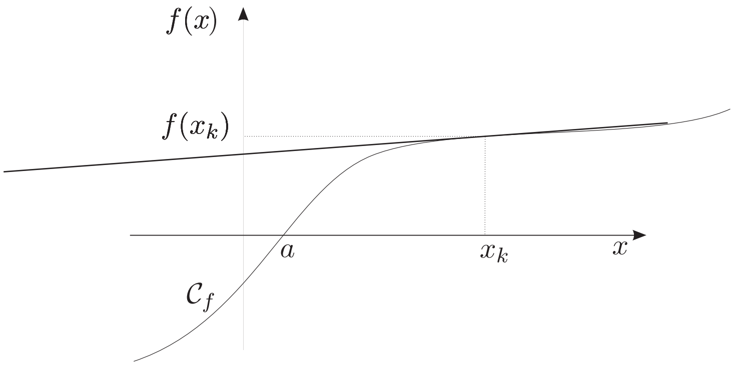


Figure C.4 – Cas particulier où $\mathbf{f}'(\mathbf{x}_k)$ est proche de 0.

Par ailleurs, certaines racines sont quasiment inaccessibles par la méthodes de Newton-Raphson, c'est le cas lorsque la dérivée au voisinage de la solution est très proche de zéro.

Une autre difficulté réside dans le fait que l'algorithme de Newton ne permet d'obtenir à chaque fois qu'une seule solution du système. Dans le cas où le système admet de nombreuses solutions, il faudra donc l'appliquer plusieurs fois en changeant de valeur initiale $\mathbf{x}_0$.

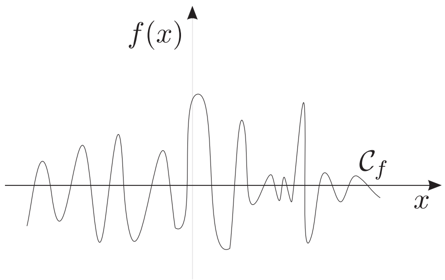


Figure C.5 – Exemple de fonction $f : \mathbb{R} \to \mathbb{R}$ admettant de nombreux zéros.

Lorsque les solutions sont très proches et que la courbe "oscille" beaucoup (voir figure C.5), il n'est pas du tout évident de choisir des points initiaux donnant à chaque fois une solution différente.

ANNEXE D

Bases de Gröbner

Les bases de Gröbner sont à la base d'algorithmes de résolution de nombreux problèmes de la théorie des idéaux de polynômes dont la résolution de systèmes d'équations polynomiales. Leur origine est attribuée à Hironaka dans [47] et surtout à Buchberger qui a le premier proposé un algorithme de calcul les utilisant dans [11].

Nous présentons dans cette annexe les aspects calculatoires des bases de Gröbner. Le lecteur souhaitant davantage d'explications pourra consulter l'article de référence en la matière [12].

D.1 Les polynômes

Commençons par préciser quelques notions élémentaires concernant les polynômes à n variables.

D.1.1 Vocabulaire

Définition D.1 (Monôme)**.** On appelle *monôme* un terme de la forme :

$$M = x_1^{d_1} \times x_2^{d_2} \cdots \times x_n^{d_n}$$

où $x_i^{d_i}$ (pour d_i un entier positif ou nul) est une notation pour $\underbrace{x_i \times \cdots \times x_i}_{d_i \text{ fois}}$. On le note $M = x_1^{d_1} x_2^{d_2} \ldots x_n^{d_n}$.
Le *degré total d'un monôme* M, noté $\deg(M)$, est défini par

$$\deg(M) = d_1 + \cdots + d_n.$$

Précisons qu'un monôme $x_1^{d_1} x_2^{d_2} \ldots x_n^{d_n}$ est *divisible* par un monôme $x_1^{d'_1} x_2^{d'_2} \ldots x_n^{d'_n}$ si pour tout $i \in \{1, \ldots, n\}$, $d'_i \leq d_i$.
Le *plus petit commun multiple* de deux monômes M_1 et M_2, noté $ppcm(M_1, M_2)$, est le plus petit monôme, au sens de l'ordre des degrés totaux, divisible par M_1 et M_2.

Définition D.2 (Polynôme)**.** Etant donné un corps $\mathbb{K}$, on appelle *polynôme* à coefficient dans $\mathbb{K}$ un terme de la forme

$$P = \sum_i a_i M_i,$$

où $a_i \in \mathbb{K}$ et M_i est un monôme.
Le produit $a_i M_i$ est appelé *terme* du polynôme.
Le *degré total du polynôme* est le maximum des degrés des monômes le composant,

$$\deg(P) = \max_i \deg(M_i).$$

L'ensemble des polynômes à n variables $x_1, \ldots, x_n$ sur le corps $\mathbb{K}$ (muni de $+$ et $\times$) est un anneau noté $\mathbb{K}[x_1, \ldots, x_n]$.

Définition D.3 (Idéal)**.** On appelle *idéal* de $\mathbb{K}[x_1, \ldots, x_n]$ un ensemble $\mathfrak{I}$ de polynômes de $\mathbb{K}[x_1, \ldots, x_n]$ qui est vérifie les 3 propriétés suivantes :
$i)$ $\mathfrak{I}$ est stable par addition : $\forall f, g \in \mathfrak{I}, f + g \in \mathfrak{I}$,
$ii)$ $\mathfrak{I}$ est absorbant pour la multiplication : $\forall f \in \mathfrak{I}, \forall g \in \mathbb{K}[x_1, \ldots, x_n], fg \in \mathfrak{I}$,
$iii)$ $\mathfrak{I}$ contient le polynôme nul.

Définition D.4 (Idéal engendré par un ensemble de polynôme)**.** Soit $F = \{f_1, \ldots, f_t\}$ un ensemble de polynômes de $\mathbb{K}[x_1, \ldots, x_n]$. On appelle idéal engendré par F l'ensemble de polynômes défini par :

$$\langle f_1, \ldots, f_t \rangle = \left\{ \sum_{i=1}^{t} h_i f_i | h_1, \ldots, h_t \in \mathbb{K}[x_1, \ldots, x_n] \right\}.$$

Ainsi défini, $\langle f_1, \ldots, f_t \rangle$ est bien un idéal de $\mathbb{K}[x_1, \ldots, x_n]$. $\{f_1, \ldots, f_t\}$ est appelé système générateur ou famille génératrice de $\langle f_1, \ldots, f_t \rangle$.

D.1.2 Relation d'ordre sur les polynômes de $\mathbb{K}[x_1, \ldots, x_n]$

Pour ordonner les polynômes de $\mathbb{K}[x_1, \ldots, x_n]$, il faut se donner une relation d'ordre $\succ$ sur les monômes, stable par multiplication. Un polynôme $f = \sum_{i=1}^{i=t} a_i M_i$ est alors ordonné par $\succ$ si

$$\forall j \in \{1, \ldots, t-1\}, M_j \succ M_{j+1}$$

Dans ce cas, $a_1 M_1$ est appelé *terme maximal* de f par rapport à $\succ$, noté $Mt(f)$ et M_1 est le monôme maximal de f, noté $Mm(f)$.

Par la suite, pour simplifier, nous supposerons, quitte à diviser par a_1, que le terme maximal est unitaire.

La relation d'ordre la plus utilisée sur les monômes est l'*ordre lexicographique*. Il consiste à fixer une précédence des variables, par exemple $x_1 > \cdots > x_n$. On distingue l'ordre lexicographique, noté $\succ_{lex}$, et l'ordre lexicographique homogène, noté $\succ_{lex_h}$, définis comme suit :

Pour deux monômes $M_1 = x_1^{d_1} x_2^{d_2} \ldots x_n^{d_n}$ et $M_2 = x_1^{d'_1} x_2^{d'_2} \ldots x_n^{d''_n}$,
– $M_1 \succ_{lex} M_2 \Leftrightarrow \exists i$ tel que $\left[d_i > d'_i \text{ et } \forall j \in \{1, \ldots, i-1\} \left(d_j = d'_j \right) \right]$.
– $M_1 \succ_{lex_h} M_2 \Leftrightarrow [\deg(M_1) > \deg(M_2)]$ ou $[\deg(M_1) = \deg(M_2)$ et $M_1 \succ_{lex} M_2]$.

Ainsi définies, les relations d'ordre $\succ_{lex}$ et $\succ_{lex_h}$coincident sur les polynômes de même degré.

Exemple D.5. *Soit la précédence des variables $x_1 > x_2 > x_3$:*
– $x_1 x_2^3 x_3 \succ_{lex} x_1 x_2 x_3$, *car la variable x_1 à le même degré dans les deux monômes et $x_2^3 \succ_{lex} x_2$.*
– $x_2^4 x_3^2 \succ_{lex_h} x_1^2 x_2 x_3^2$, *car* $\deg\left(x_2^4 x_3^2\right) (= 6) > \deg\left(x_1^2 x_2 x_3^2\right) (= 5)$.
– $x_1^3 x_2 x_3^2 \succ_{lex_h} x_1^2 x_2 x_3^3$, *car* $\deg\left(x_1^3 x_2 x_3^2\right) = \deg\left(x_1^2 x_2 x_3^2\right)$ *et* $x_1^3 x_2 x_3^2 \succ_{lex} x_1^2 x_2 x_3^3$.

Pour toute la suite de cette annexe, nous fixons l'ordre lexicographique sur les monômes et donc sur les polynômes. Nous le noterons simplement $\succ$.

D.1.3 Division d'un polynôme par une suite de polynômes.

Soient f, un polynôme de $\mathbb{K}[x_1, \ldots, x_n]$ et $\{f_1, \ldots, f_t\}$, une suite de polynômes quelconques de $\mathbb{K}[x_1, \ldots, x_n]$. Soit la précédence des variables $x_1 > \cdots > x_n$, nous admettons le résultat suivant :

Théorème D.6 (Division euclidienne)**.** *Il existe $h, h_1, \ldots, h_t \in \mathbb{K}[x_1, \ldots, x_n]$, uniques, tels que*

$$f = h_1 f_1 + \cdots + h_t f_t + h$$

où $\forall i$, les monômes des $h_i Mm(f_i)$ ne sont divisibles par aucun des $Mm(f_j)$ pour $j < i$ et $\forall i$, les monômes de h ne sont pas divisibles par $Mm(f_i)$.

Définition D.7 (Reste de la division euclidienne)**.** Considérons la division euclidienne de f par la suite de polynôme $f_1, \ldots, f_t$ (voir théorème D.6),

$$f = h_1 f_1 + \cdots + h_t f_t + h.$$

Le polynôme h est appelé reste de la division de f par la suite $f_1, \ldots, f_s$ et est noté $\mathcal{R}_{(f_1,\ldots,f_t)}(f)$.

Il existe un algorithme de décomposition permettant d'obtenir les polynômes $h, h_1, \ldots, h_t$ connaissant les polynômes $f, f_1, \ldots, f_t$. Ce dernier étant relativement complexe, nous ne le présentons pas dans cette annexe mais nous l'utiliserons par la suite. Pour davantage de détails, le lecteur pourra consulter [33].

Exemple D.8. *Soit $f = x_1^3 x_2 \in \mathbb{K}[x_1, x_2]$. Considérons l'ensemble $\{f_1, f_2\} \in \mathbb{K}[x_1, x_2]$ avec $f_1 = x_1^3$ et $f_2 = x_1^2 x_2 - x_2^3$.*
L'algorithme de décomposition nous donne $f = x_2 f_1$. Par unicité de l'écriture (théorème D.6), on en déduit que $\mathcal{R}_{(f_1,f_2)}(f) = 0$.

Exemple D.9. *Considérons maintenant l'ensemble $\{f_1, f_2\} \in \mathbb{K}[x_1, x_2]$ avec $f_1 = x_1^2 x_2 - x_2^3$ et $f_2 = x_1^3$. Nous avons seulement échangé les fonctions f_1 et f_2 de l'exemple précédent.*
La décomposition de f (au sens du théorème D.6) est $f = x_1 f_1 + x_1 x_2^3$. Il s'en suit que $\mathcal{R}_{(f_1,f_2)}(f) = x_1 x_2^3$.

La notion de reste introduite dans ce paragraphe nous sera utile par la suite lors de la construction d'une base de Gröbner.

D.2 Bases de Gröbner

Une base de Gröbner d'un idéal est une famille génératrice particulière de cet idéal possédant de bonnes propriétés concernant la réduction et la division. Plus précisément, on a la définition suivante :

Définition D.10 (Base de Gröbner d'un idéal)**.** Soient $\{f_1, \ldots, f_t\}$, un ensemble de polynômes de $\mathbb{K}[x_1, \ldots, x_n]$ et $\mathfrak{I} = \langle f_1, \ldots, f_t \rangle$ l'idéal de $\mathbb{K}[x_1, \ldots, x_n]$ engendré par cet ensemble. Une base de Gröbner de l'idéal $\mathfrak{I}$ est un système de générateurs $\{g_1, \ldots, g_s\}$ de l'idéal $\mathfrak{I}$ construit de sorte que

$$\left\langle \{M_m(g_i)\}_{i=1,\ldots,s} \right\rangle = \langle \{M_m(f), f \in \mathfrak{I}\} \rangle \tag{D.1}$$

c'est-à-dire que les monômes maximaux des polynômes g_i forment un système de générateurs de l'idéal engendré par les monômes maximaux de tous les éléments de l'idéal $\mathfrak{I}$.

Remarquons tout de suite, à l'aide d'un contre-exemple, qu'en général $\{f_1, \ldots, f_t\}$ n'est pas une base de Gröbner de l'idéal $\langle f_1, \ldots, f_t \rangle$.

Exemple D.11. *Soit $\mathfrak{I} = \langle f_1, f_2 \rangle$ l'idéal de $\mathbb{K}[x_1, x_2]$ engendré par les polynômes $f_1 = x_1^3$ et $f_2 = x_1^2 x_2 - x_2^3$. On a*

$$Mm(f_1) = x_1^3 \text{ et } Mm(f_2) = x_1^2 x_2.$$

Nous allons voir que $\{f_1, f_2\}$ n'est pas une base de Gröbner de l'idéal $\mathfrak{I}$. Pour cela, considérons le polynôme $f \in \mathfrak{I}$ défini par

$$f = x_2 f_1 - x_1 f_2 = x_1 x_2^3.$$

On vérifie aisément que

$$M_m(f) = x_1 x_2^3 \notin \langle M_m(f_1), M_m(f_2) \rangle = \left\langle x_1^3, x_1^2 x_2 \right\rangle.$$

La condition (D.1) n'est pas vérifiée. Dès lors, $\{f_1, f_2\}$ n'est pas une base de Gröbner de l'idéal $\mathfrak{I} = \langle f_1, f_2 \rangle$.

Nous n'entrerons pas dans les détails concernant les propriétés d'une base de Gröbner mais présenterons simplement les grandes lignes de sa construction. Par ailleurs, nous admettrons que tout idéal possède une base de Gröbner.

Avant de présenter l'algorithme de Buchberger, qui permet, à partir d'un système de générateurs d'un idéal de $\mathbb{K}[x_1, \ldots, x_n]$, de construire une base de Gröbner de cet idéal, nous citons le théorème sur lequel repose cet algorithme et précisons un peu les notations qui seront utilisées par la suite.

D.2.1 Préliminaires

Dans ce paragraphe, les résultats énoncés ne seront pas démontrés mais simplement illustrés à l'aide d'exemples concrets, notre objectif étant de donner une idée de la démarche permettant d'arriver à l'algorithme de construction d'une base de Gröbner. Pour davantage de détails, se référer au cours [69].

Définition D.12 (S-polynôme)**.** Pour deux polynômes f_1 et f_2, non nuls de $\mathbb{K}[x_1, \ldots, x_n]$, on définit le S-polynôme de f_1 et f_2, noté $S(f_1, f_2)$ par :

$$S(f_1, f_2) = \frac{ppcm\left(Mm\left(f_1\right), Mm\left(f_2\right)\right)}{Mm\left(f_1\right)} f_1 - \frac{ppcm\left(Mm\left(f_1\right), Mm\left(f_2\right)\right)}{Mm\left(f_2\right)} f_2.$$

Remarque D.13. *Le S-polynôme $S(f_1, f_2)$ est construit avec pour objectif la suppression des monômes maximaux des polynômes f_1 et f_2. De plus, par définition, $S(f_1, f_2) \in \langle f_1, f_2 \rangle$.*

Illustrons ceci par un exemple.

Exemple D.14. *Soient deux polynômes $f_1 = x_1^3$ et $f_2 = x_1^2 x_2 - x_2^3$ de $\mathbb{K}[x_1, x_2]$. On a*

$$Mm(f_1) = x_1^3 \text{ et } Mm(f_2) = x_1^2 x_2,$$

au sens de l'ordre lexicographique avec la précédence $x_1 > x_2$. On en déduit

$$ppcm\left(Mm\left(f_1\right), Mm\left(f_2\right)\right) = x_1^3 x_2.$$

Nous pouvons calculer le S-polynôme de f_1 et f_2

$$\begin{aligned} S(f_1, f_2) &= \frac{x_1^3 x_2}{x_1^3} \times x_1^3 - \frac{x_1^3 x_2}{x_1^2 x_2} \times (x_1^2 x_2 - x_2^3) \\ &= x_2 x_1^3 - x_1 \times (x_1^2 x_2 - x_2^3) \\ &= x_2 x_1^3 - x_2 x_1^3 + x_1 x_2^3 \\ &= x_1 x_2^3. \end{aligned}$$

On a $S(f_1, f_2) = x_2 f_1 - x_1 f_2$ donc $S(f_1, f_2)$ appartient bien à l'idéal $\langle f_1, f_2 \rangle$ de $\mathbb{K}[x_1, x_2]$.

Théorème D.15. *Soit $\mathfrak{I}$ un idéal de $\mathbb{K}[x_1, \ldots, x_n]$ et soit $\{f_1, \ldots, f_t\}$ un ensemble de polynômes non nuls de $\mathfrak{I}$, qui forment un système générateurs de $\mathfrak{I}$. Alors, l'ensemble $\{f_1, \ldots, f_t\}$ est une base de Gröbner de $\mathfrak{I}$ si et seulement si*

$$\forall i, j \in \{1, \ldots, t\},\ \mathcal{R}_{(f_1, \ldots, f_t)}(S(f_i, f_j)) = 0.$$

Illustrons tout de suite ce théorème par un exemple.

Exemple D.16. *Considérons l'ensemble $\{f_1, f_2, f_3, f_4\}$ de polynôme de $\mathbb{K}[x_1, x_2]$ défini par*

$$f_1 = x_1^3, \quad f_2 = x_1^2 x_2 - x_2^3, \quad f_3 = x_1 x_2^3, \quad f_4 = x_2^5.$$

Soit $\mathfrak{I} = \langle f_1, f_2, f_3, f_4 \rangle$ l'idéal engendré par ces 4 polynômes. Montrons que $\{f_1, f_2, f_3, f_4\}$ est une base de Gröbner de $\mathfrak{I}$, c'est-à-dire, d'après le théorème (D.15), que les restes des divisions des $S(f_i, f_j)$ par la suite de polynômes (f_1, f_2, f_3, f_4) sont nuls. On a :

$$Mm(f_1) = x_1^3, \quad Mm(f_2) = x_1^2 x_2, \quad Mm(f_3) = x_1 x_2^3, \quad Mm(f_4) = x_2^5.$$

On en déduit les $S(f_i, f_j)$ suivants,

$$\begin{aligned} &S(f_1, f_2) = x_1 x_2^3 = f_3 \\ &S(f_2, f_3) = -x_2^5 = -f_4 \\ &S(f_2, f_4) = -x_2^7 = -x_2 f_4 \\ &S(f_1, f_3) = S(f_1, f_4) = S(f_3, f_4) = 0. \end{aligned}$$

En conséquence, les restes des divisions des $S(f_i, f_j)$ par les f_1, f_2, f_3, f_4 sont nuls et donc, d'après le théorème D.15, $\{f_1, f_2, f_3, f_4\}$ est une base de Gröbner de $\mathfrak{I}$.

Attention, on rappelle qu'en général, $\{f_1, \ldots, f_t\}$ n'est pas une base de Gröbner de l'idéal $\langle f_1, \ldots, f_t \rangle$ (voir exemple D.11).

Maintenant, nous allons nous intéresser à la construction proprement dite d'une base de Gröbner d'un idéal $\mathfrak{I}$.

D.2.2 Algorithme de Buchberger

Considérons un ensemble $\{f_1, \ldots, f_t\}$ de polynômes de $\mathbb{K}[x_1, \ldots, x_n]$ et l'idéal $\mathfrak{I} = \langle f_1, \ldots, f_t \rangle$ engendré par cet ensemble. Le principe de l'algorithme de Buchberger est d'ajouter à chaque itération un polynôme supplémentaire à l'ensemble $\{f_1, \ldots, f_t\}$ jusqu'à ce que ce dernier devienne une base de Gröbner de l'idéal $\mathfrak{I}$.

Plus précisément, la première étape consiste à vérifier si l'ensemble $\{f_1, \ldots, f_t\}$ est une base de Gröbner de $\mathfrak{I}$ (voir théorème D.15). Dans le cas où $\{f_1, \ldots, f_t\}$, n'est pas une base de Gröbner de $\mathfrak{I}$, on peut trouver deux indices i et j tels que

$$\mathcal{R}_{(f_1,\ldots,f_t)}\left(S\left(f_i, f_j\right)\right) \neq 0.$$

Choisissons i et j de sorte que $i+j$ soit minimal et posons $f_{t+1} = \mathcal{R}_{(f_1,\ldots,f_t)}\left(S\left(f_i, f_j\right)\right)$. Ainsi définie, f_{t+1} est bien un élément de l'idéal $\mathfrak{I}$.
Si $\{f_1, \ldots, f_t, f_{t+1}\}$ n'est toujours pas une base de Gröbner de $\mathfrak{I}$, nous pouvons itérer le procédé et trouver deux nouveaux indices i et j tels que

$$\mathcal{R}_{(f_1,\ldots,f_{t+1})}\left(S\left(f_i, f_j\right)\right) \neq 0.$$

Et ainsi de suite jusqu'à ce que le processus s'arrête.

Exemple D.17. *Considérons les polynômes $f_1, f_2, f_3 \in \mathbb{K}[x_1, x_2, x_3]$ définis par*

$$f_1 = x_1^3, \quad f_2 = x_1x_2, \quad f_3 = x_1x_3 + x_3^2.$$

On a

$$Mm(f_1) = x_1^3, \quad Mm(f_2) = x_1x_2, \quad Mm(f_3) = x_1x_3.$$

Déterminons une base de Gröbner de l'idéal $\mathfrak{I} = \langle f_1, f_2, f_3 \rangle$.
Pour commencer, vérifions si $\{f_1, f_2, f_3\}$ est une base de Gröbner de $\mathfrak{I}$. On a :

$$\begin{array}{lll} S(f_1, f_2) = 0 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3)}\left(S(f_1, f_2)\right) = 0, \\ S(f_1, f_3) = -x_1^2x_3^2 = -x_3f_3 + x_3^3 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3)}\left(S(f_1, f_3)\right) = x_3^3, \\ S(f_2, f_3) = -x_2x_3^2 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3)}\left(S(f_2, f_3)\right) = -x_2x_3^2. \end{array}$$

Donc $\{f_1, f_2, f_3\}$ n'est pas une base de Gröbner de $\mathfrak{I}$.
Posons $f_4 = \mathcal{R}_{(f_1,f_2,f_3)}\left(S(f_1, f_3)\right) = x_3^3$ et considérons l'ensemble $\{f_1, f_2, f_3, f_4\}$. On a

$$\begin{array}{lll} S(f_1, f_2) = S(f_1, f_4) = S(f_2, f_4) = 0, & & \\ S(f_1, f_3) = -x_1^2x_3^2 = -x_3f_3 + f_4 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4)}\left(S(f_1, f_3)\right) = 0, \\ S(f_2, f_3) = -x_2x_3^2 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4)}\left(S(f_2, f_3)\right) = -x_2x_3^2, \\ S(f_3, f_4) = x_3^4 = x_3f_4 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4)}\left(S(f_3, f_4)\right) = 0. \end{array}$$

Donc $\{f_1, f_2, f_3, f_4\}$ n'est pas une base de Gröbner de $\mathfrak{I}$.
Posons $f_5 = \mathcal{R}_{(f_1,f_2,f_3,f_4)}S(f_2, f_3)) = -x_2x_3^2$ et considérons l'ensemble $\{f_1, f_2, f_3, f_{4,}f_5\}$. On a

$$\begin{array}{lll} \multicolumn{3}{l}{S(f_1, f_2) = S(f_1, f_4) = S(f_1, f_5) = S(f_2, f_4) = S(f_2, f_5) = S(f_4, f_5) = 0,} \\ S(f_1, f_3) = -x_1^2x_3^2 = -x_3f_3 + f_4 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4,f_5)}\left(S(f_1, f_3)\right) = 0, \\ S(f_2, f_3) = -x_2x_3^2 = f_5 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4,f_5)}\left(S(f_2, f_3)\right) = 0, \\ S(f_3, f_4) = x_3^4 = x_3f_4 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4,f_5)}\left(S(f_3, f_4)\right) = 0, \\ S(f_3, f_5) = -x_2X_3^3 = -x_2f_4 & \Longrightarrow & \mathcal{R}_{(f_1,f_2,f_3,f_4,f_5)}\left(S(f_3, f_5)\right) = 0. \end{array}$$

$\forall i, j \in \{1, \ldots, 5\}, \mathcal{R}_{(f_1,f_2,f_3,f_4,f_5)}\left(S(f_i, f_j)\right) = 0$ *donc $\{f_1, f_2, f_3, f_{4,}f_5\}$ est une base de Gröbner de l'idéal $\mathfrak{I} = \langle f_1, f_2, f_3 \rangle$.*

D.3 Application à la résolution de systèmes d'équations

Dans ce paragraphe, nous allons voir comment les bases de Gröbner peuvent être utilisées afin de caractériser l'ensemble des solutions d'un système d'équations polynomiales.

Le problème est de résoudre un ensemble de t équations $\{f_1 = 0, \dots, f_t = 0\}$ avec $f_1, \dots, f_t \in \mathbb{K}[x_1, \dots, x_n]$. On considère la précédence des variables $x_1 > \cdots > x_n$. Le rôle des base de Gröbner est similaire à celui de l'algorithme d'élimination de Gauss pour les systèmes linéaires.

Nous distinguons deux étapes dans la résolution.

- Tout d'abord, on construit une base de Gröbner de l'idéal engendré par les polynômes $f_1, \dots, f_t$. On réordonne ensuite les éléments de cette base en commençant par le polynôme en x_n, puis le(s) polynôme(s) en x_{n-1} et x_n, etc... Nous obtenons ainsi un nouveau système, triangulaire, équivalent au premier.

$$\begin{cases} f_1(x_n) = 0 \\ f_2(x_n, x_{n-1}) = 0 \\ \vdots \\ f_s(x_n, x_{n-1}, \dots, x_1) = 0 \end{cases}$$

- Ensuite, c'est une phase de calcul. On commence par calculer les racines réelles du polynôme en x_n, puis, chacune de ces racines donne un nouvel ensemble d'équation (on remplace x_n par la valeur obtenue), on peut calculer x_{n-1} et, etc...

Exemple D.18. *Considérons le système suivant, de deux équations à deux inconnues*

$$\begin{cases} x_1^2 + x_1x_2 + 2x_1 + x_2 - 1 = 0 \\ x_1^2 - x_2^2 + 3x_1 + 2x_2 - 1 = 0. \end{cases} \tag{D.2}$$

Posons $f_1 = x_1^2 + x_1x_2 + 2x_1 + x_2 - 1$ *et* $f_2 = x_1^2 - x_2^2 + 3x_1 + 2x_2 - 1$.
Commençons par vérifier si $\{f_1, f_2\}$ *est une base de Gröbner de l'idéal engendré par* f_1 *et* f_2.

$$S(f_1, f_2) = x_1x_2 - x_1 + x_2^2 - x_2 \implies \mathcal{R}_{(f_1,f_2)}(S(f_1, f_2)) = x_1x_2 - x_1 + x_2^2 - x_2.$$

Donc $\{f_1, f_2\}$ *n'est pas une base de Gröbner. Posons* $f_3 = x_1x_2 - x_1 + x_2^2 - x_2$*, on a*

$$\begin{array}{ll} S(f_1, f_2) = x_1x_2 - x_1 + x_2^2 - x_2 = f_3 & \implies \mathcal{R}_{(f_1,f_2,f_3)}(S(f_1, f_2)) = 0 \\ S(f_1, f_3) = x_1^2 + 3x_1x_2 + x_2^2 - x_2 = f_1 + 2f_3 - x_2^2 + 1 & \implies \mathcal{R}_{(f_1,f_2,f_3)}(S(f_1, f_3)) = -x_2^2 + 1 \end{array}$$

Donc $\{f_1, f_2, f_3\}$ *n'est pas une base de Gröbner. Posons* $f_4 = -X_2^2 + 1$*, on obtient*

$$\begin{array}{ll} S(f_1, f_2) = x_1x_2 - x_1 + x_2^2 - x_2 = f_3 & \implies \mathcal{R}_{(f_1,f_2,f_3,f_4)}(S(f_1, f_2)) = 0 \\ S(f_1, f_3) = x_1^2 + 3x_1x_2 + x_2^2 - x_2 = f_1 + 2f_3 + f_4 & \implies \mathcal{R}_{(f_1,f_2,f_3,f_4)}(S(f_1, f_3)) = 0 \\ S(f_2, f_3) = x_1^2 - x_1x_2^2 + 4x_1x_2 - x_2^3 + 2x_2^2 = f_1 - (x_2 - 2)f_3 + f_4 & \implies \mathcal{R}_{(f_1,f_2,f_3,f_4)}(S(f_2, f_3)) = 0 \\ S(f_3, f_4) = x_1x_2 - x_1 - x_2^3 + x_2^2 = f_3 + x_2f_4 & \implies \mathcal{R}_{(f_1,f_2,f_3,f_4)}(S(f_3, f_4)) = 0 \end{array}$$

De plus, on peut vérifier que

$$\mathcal{R}_{(f_1,f_2,f_3,f_4)}(S(f_1, f_4)) = \mathcal{R}_{(f_1,f_2,f_3,f_4)}(S(f_2, f_4)) = 0,$$

donc, $\{f_1, f_2, f_3, f_4\}$ *est une base de Gröbner de* $\langle f_1, f_2\rangle$.
Nous obtenons un nouveau système équivalent au système (D.2) :

$$\begin{cases} -x_2^2 + 1 = 0 \\ x_1x_2 - x_1 + x_2^2 - x_2 = 0 \\ x_1^2 - x_2^2 + 3x_1 + 2x_2 - 1 = 0 \\ x_1^2 + x_1x_2 + 2x_1 + x_2 - 1 = 0. \end{cases}$$

La première équation possède deux solutions $x_2 = 1$ *et* $x_2 = -1$. *En remplaçant* x_2 *par ces valeurs, on obtient deux nouveaux systèmes* (S_1) *et* (S_2), *que l'on résoud aisément.*

$$(S_1)\begin{cases} x_2 = 1 \\ x_1^2 + 3x_1 = 0 \\ x_1^2 + x_1 + 2x_1 = 0 \end{cases} \Longrightarrow \begin{cases} x_2 = 1 \\ x_1 = 0 \textit{ ou } x_1 = -3 \end{cases}$$

$$(S_2)\begin{cases} x_2 = -1 \\ -2x_1 + 2 = 0 \\ x_1^2 + 3x_1 - 4 = 0 \\ x_1^2 - x_1 + 2x_1 - 2 = 0 \end{cases} \Longrightarrow \begin{cases} x_2 = -1 \\ x_1 = 1 \end{cases}$$

Ainsi, les couples solutions du système (D.2) sont (0,1), (-3,1) et (1,-1).

ANNEXE E

Estimation des dérivées : différentiateur

Cette annexe propose différentes méthodes de différentiation. Le lecteur intéressé par cette problématique trouvera dans [24] une liste plus complète et détaillée des méthodes existantes ainsi que des comparaisons entre chacune d'elle.

E.1 Problématique

On souhaite mesurer une grandeur physique $y(t)$ à travers un capteur C. La mesure $y_C(t)$ de $y(t)$ à travers ce capteur peut s'écrire :

$$y_C(t) = G(y(t)) + b_i(t), \tag{E.1}$$

où G est une fonction qui modélise les distorsions linéaires ou non linéaires subies par y lors de la mesure et où le terme $b_i(t)$ modélise la présence d'un bruit de mesure propre au capteur.
En pratique, c'est un peu différent : le signal $y(t)$ à mesurer est toujours noyé dans un bruit de fond, extérieur au capteur, caractéristique de l'environnement de mesure. En tenant compte de ce bruit extérieur, noté $b_e(t)$, on obtient une nouvelle expression de $y_C(t)$:

$$y_C(t) = G(y(t) + b_e(t)) + b_i(t). \tag{E.2}$$

La figure E.1 illustre cette modélisation.

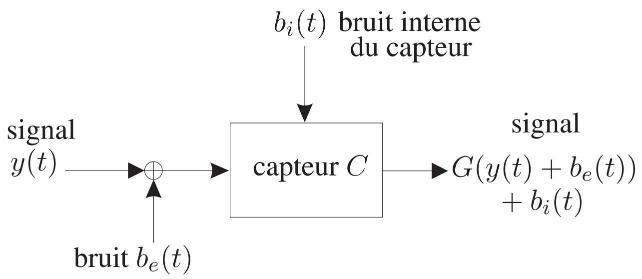


Figure E.1 – Système de mesure.

Dans le cas particulier où l'on se limite à la zone de fonctionnement linéaire du capteur, l'expression (E.2) à la sortie du capteur peut s'écrire

$$y_C(t) = Ky(t) + Kb_e(t) + b_i(t) \tag{E.3}$$

où K est une grandeur constante.
Pour simplifier par la suite, nous nous limiterons au cas du fonctionnement linéaire des capteurs. De plus, nous supposerons possible l'additivité des bruits $Kb_e(t)$ et $b_i(t)$. Nous pouvons donc à nouveau simplifier et l'expression (E.3) devient

$$y_C(t) = Ky(t) + \delta(t) \tag{E.4}$$

où $\delta(t)$ correspond à un bruit parasite regroupant les effets du bruit interne au capteur et du bruit extérieur.

Le différentiateur, dans la chaîne de récupération et de traitement du signal est positionné à la sortie du capteur (voir figure E.1). L'expression (E.2) correspond donc à la forme du signal d'entrée dans le différentiateur. Avec les hypothèses de linéarité du fonctionnement du capteur et d'additivité des bruits, nous considons donc que tout signal bruité $y^\delta(t)$ à l'entrée d'un différentiateur peut s'écrire de la forme

$$y^\delta(t) = y(t) + \delta(t) \tag{E.5}$$

où $y(t)$ correspond au signal utile à dériver et, $\delta(t)$ au bruit parasite décrit précédemment.

Nous allons maintenant nous consacrer exclusivement au rôle du différentiateur et présenter différentes méthodes de différentiation d'un signal. Dans un premier temps, nous évoquerons des méthodes exclusivement numériques puis nous parlerons de la méthode de algébrique proposée par Fliess dans [82].

E.2 Méthodes numériques

Nous nous plaçons dans un environnement idéal, c'est-à-dire non bruité. L'expression du signal à l'entrée du capteur est donc tout simplement $y(t)$. Dans un tel contexte, la méthode la plus intuitive pour obtenir une estimation des n premières dérivées d'un signal est celles des différences finies, basée sur le théorème de Taylor-Laplace.

E.2.1 La méthode des différences finies

Commençons par énoncer le théorème de Taylor-Laplace.

Théorème E.1. *On considère un intervalle $I \subset \mathbb{R}$ non réduit à un point. Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$ sur I et soit $a \in I$, alors :*

$$\forall x \in I, f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)} dt.$$

La première partie de l'expression, $\sum_{k=0}^{k=n} \frac{f^{(k)}(a)}{k!}(x-a)^k$, correspond au développement limité de f à l'ordre n, aussi, la seconde partie, $\frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)} dt$, correspond-elle à la différence entre $f(x)$ et son développement limité à l'ordre n. Cette différence est encore appelée « reste de Laplace ».

Considérons maintenant un signal continu $y : t \to y(t)$ et supposons que l'on a obtenu $n+1$ mesures de ce signal, uniformément espacées d'un pas de temps Δt. La première mesure est prise à l'instant t_0, la seconde à l'instant $t_1 = t_0 + \Delta t$, etc...($\forall i \leq n$, $t_i = t_0 + i\Delta t$). Nous obtenons un vecteur de mesure

$$\begin{pmatrix} y(t_0) & y(t_1) & \dots & y(t_n) \end{pmatrix}^T.$$

En appliquant $n+1$ fois le théorème (E.1) aux $n+1$ composantes de ce vecteur, nous obtenons un système de la forme :

$$\begin{pmatrix} y(t_0) \\ y(t_1) \\ \vdots \\ y(t_n) \end{pmatrix} = \mathbf{A} \begin{pmatrix} y(t_0) \\ y^{(1)}(t_0) \\ \vdots \\ y^{(n)}(t_0) \end{pmatrix} + \mathbf{b}, \tag{E.6}$$

où la matrice $\mathbf{A}$ et le vecteur $\mathbf{b}$ sont définis comme suit :

$$\mathbf{A} = \left(\frac{(k\Delta t)^j}{j!}\right)_{0\leq k,j\leq n} \quad \text{et} \quad \mathbf{b} = \left(\frac{1}{n!}\int_{t_0}^{t_0+k\Delta t} (t_k - t)^n \, y^{(n+1)}(t)\, dt\right)_{0\leq k\leq n}$$

Notons que les colonnes de la matrice $\mathbf{A}$ sont indépendantes, elle est donc inversible. Nous obtenons un nouveau système.

$$\begin{pmatrix} y(t_0) \\ y^{(1)}(t_0) \\ \vdots \\ y^{(n)}(t_0) \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} y(t_0) \\ y(t_1) \\ \vdots \\ y(t_n) \end{pmatrix} - \mathbf{A}^{-1}\mathbf{b}.$$

Les valeurs du vecteur $\mathbf{b}$ correspondent aux restes de Laplace des développements limités à l'ordre n du signal aux $n+1$ instants $t_0, t_1, \ldots, t_n$. Ces valeurs et par conséquent le vecteur $\mathbf{A}^{-1}\mathbf{b}$ peuvent être négligés. Nous obtenons alors une estimation des dérivées du signal $y(t)$ à l'instant t_0 :

$$\begin{pmatrix} y(t_0) \\ y^{(1)}(t_0) \\ \vdots \\ y^{(n)}(t_0) \end{pmatrix} \approx \mathbf{A}^{-1} \begin{pmatrix} y(t_0) \\ y(t_1) \\ \vdots \\ y(t_n) \end{pmatrix}. \tag{E.7}$$

Notons pour finir qu'en présence d'un bruit $\delta(t)$, cette méthode n'aurait pû être appliquée car il aurait fallu dériver également $\delta(t)$ et cela aurait perturbé de façon conséquente le résultat de l'approximation.

Une autre approche plus originale pour résoudre ce problème de différentiation est d'utiliser une interpolation polynômiale.

E.2.2 Interpolation polynômiale

On considère ici un signal continu $y : t \to y(t)$ et $n+1$ mesures discrètes de ce signal. La méthode consiste à interpoler ces mesures par un polynôme, appelé polynôme interpolateur. Nous énonçons pour commencer le théorème d'interpolation de Lagrange :

Théorème E.2. *On considère un intervalle $I \subset \mathbb{R}$ non réduit à un point. Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$ sur I. On suppose connues $n+1$ mesures de f, uniformément espacées d'un temps Δt, la première mesure étant prise à l'instant t_0, la seconde à l'instant t_1, etc...($\forall i \leq n$, $t_i = t_0 + i\Delta t$). Alors, il existe un seul et unique polynôme continu sur I de degré inférieur ou égal à n qui interpole f. Ce dernier, appelé polynôme de Lagrange et noté L_n, est défini par :*

$$\forall t \in I, L_n(t) = \sum_{i=0}^{n} f(t_i)\, l_i(t),$$

où

$$l_i(t) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{t - t_j}{(i-j)\,\Delta t}.$$

Pour obtenir une estimation des dérivées du signal $y(t)$, il suffit de considérer les dérivées successives de son polynôme interpolateur, qui s'obtiennent aisément à partir des coefficients de ce dernier. Toutefois, cette méthode possède un inconvénient majeur, le phénomène dit "de Runge". Un changement même infime d'une des mesures $y\left(t_i\right)$ provoque un changement sur l'ensemble des coefficients du polynôme $L_n\left(t\right)$. Cela engendre une modification des dérivées du polynôme $L_n\left(t\right)$, particulièrement au voisinage des points t_0 et t_n. En présence de bruit, cette méthode est peu recommandée dans le sens où un bruit, même très faible provoquerait ce phénomène de Runge et perturberait ainsi fortement les estimations de dérivées obtenues.

En conclusion, dans un environnement bruité, l'interpolation polynomiale tout comme la méthode des différences finies, s'avèrent peu efficaces. C'est d'ailleurs pour cette raison que nous nous sommes placés dans un contexte idéal non bruité. La méthode que nous allons maintenant illustrer est plus robuste dans le cas d'un environnement bruité. Il s'agit d'une approche totalement différente, qui nous amène à un problème d'optimisation sous contrainte.

E.2.3 La régularisation

Pour cette méthode, nous présentons uniquement l'idée générale sans entrer dans les détails calculatoires. Pour davantage de détails, consulter [53].

Problème d'optimisation sous contrainte

Nous nous plaçons, pour ce paragraphe, dans un environnement bruité. On considère un bruit blanc gaussien, noté δ de moyenne nulle et de variance σ^2. Soit $y : \mathbb{R}^+ \to \mathbb{R}$, un signal continu dont on souhaite estimer les q premières dérivées ($q > 0$). Le signal à l'entrée du différentiateur peut donc s'écrire

$$y^\delta(t) = y(t) + \delta(t)$$

On considère $n+1$ mesures bruités $\left(y^\delta\left(t_0\right), y^\delta\left(t_1\right), \ldots, y^\delta\left(t_n\right)\right) \in \mathbb{R}^{n+1}$ de ce signal, prises aux instants $t_0, t_1, \ldots, t_n$ uniformément espacées d'un temps Δt. La méthode, dîte de *régularisation*, cherche à lisser ces mesures bruitées. Pour cela, on se ramène à un problème d'optimisation sous contrainte, qui consiste à minimiser un critère, appelé *critère de régularisation.*

On note $\hat{y} : \left[t_0, t_n\right] \to \mathbb{R} \in \mathcal{C}^q\left[t_0, t_n\right]$ l'estimation du signal bruité y^δ, solution du problème d'optimisation contrainte.

Critère de régularisation

Généralement, le critère de régularisation pour ce type de problème est composé de deux termes :

- Un *terme de lissage,* équivalent à l'intégrale au carré de la $q$ième dérivée, qui minimise l'énergie de la $q$ième dérivée.
- Un *terme de précision*, qui représente l'erreur quadratique moyenne entre la solution du problème et les mesures obtenues.

Pour notre problème, le critère de régularisation s'écrit

$$\min \sum_{i=q-1}^{n-1} \left(\hat{y}^{(q)}(t_i)\left(\Delta t\right)^q\right)^2 \text{ (terme de lissage),} \tag{E.8}$$

sous la contrainte :

$$\frac{1}{n+1} \sum_{i=0}^{n} \left(\hat{y}\left(t_i\right) - y^\delta\left(t_i\right)\right)^2 \leq \sigma^2 \text{ (terme de précision).}$$

Dans (E.8), le terme $\hat{y}^{(q)}(t_i)$ correspond à une estimation de la $q$ième dérivée de $\hat{y}$ à l'instant $t_{i.}$, obtenue par la méthode des différences finies (voir théorème (E.1)). On peut ainsi écrire

$$\hat{y}^{(q)}(t_i) = \frac{\hat{y}^{(q-1)}(t_{i+1}) - \hat{y}^{(q-1)}(t_i)}{\Delta t}.$$

Cette expression montre que le terme de lissage a été défini de manière à lisser la $q$ième dérivée de $\hat{y}$. Après des calculs élémentaires, nous obtenons la formule suivante :

$$\hat{y}^{(q)}(t_i) = \frac{1}{(\Delta t)^q}\sum_{j=0}^{q}(-1)^{q+j}C_q^j\hat{y}(t_{i+j}),$$

où C_q^j est le nombre de combinaisons de j éléments d'un ensemble de q éléments $\left(C_q^j = \frac{q!}{j!(q-j)!}\right)$.

Exemple E.3. *Dans le cas où q=1, le terme de lissage s'écrit*

$$\min\sum_{i=0}^{n-1}\left(\hat{y}(t_{i+1}) - \hat{y}(t_i)\right)^2$$

et correspond à minimiser la somme des écarts quadratiques entre deux valeurs successive de $\hat{y}$.

Estimation

Pour obtenir les dérivées du signal y^δ, la méthode consiste à régulariser ce signal. Nous obtenons la formule suivante pour l'estimée $\hat{y}$ du signal bruité :

$$\hat{y} = \arg\min_{\tilde{y}\in C^q\left[t_0,t_{k+n}\right]}\left[\left(\sum_{i=q-1}^{n-1}\left(\hat{y}^{(q)}(t_i)\,(\Delta t)^q\right)^2\right) + \lambda\left(\sum_{i=k}^{k+n}\left(\tilde{y}\,(t_i) - y^\delta\,(t_i)\right)^2 - (n+1)\,\sigma^2\right)\right] \tag{E.9}$$

où λ est le paramètre de Lagrange. Par construction, $\hat{y}$ a des dérivées lisses.

E.3 Une méthode algébrique

La méthode présentée maintenant a été proposée par M.Fliess et H.Sira-Ramirez dans [82]. Pour obtenir les dérivées successives d'un signal y(t), elle utilise des résultats de l'algèbre différentielle. Les personnes intéressées par davantage de détails sur les techniques de l'algèbre différentielle pourront se référer à [80].

E.3.1 Principe

Considérons un signal arbitraire lisse. Au voisinage d'un instant t_0 fixé, le signal $y(t)$ peut être approximé, d'après le théorème (E.1) par son développement de Taylor à l'ordre $N-1$. Soit $\hat{y}$ l'estimation obtenue, on obtient

$$\hat{y}\,(t) = \sum_{k=0}^{N-1}\frac{y^{(k)}\,(t_0)}{k!}\,(t-t_0)^k\,. \tag{E.10}$$

Remarque E.4. *Plus l'entier N sera grand, meilleure sera l'approximation $\hat{y}$.*

Il suffit donc pour obtenir l'estimée $\hat{y}$ de y, d'obtenir une estimation des dérivées sucessives $y^{(k)}$, $k < N$, du signal initial. Pour cela, commençons par remarquer dans (E.10), que l'estimée $\hat{y}(t)$ peut être vue comme solution du système différentiel :

$$\left\{ \begin{array}{l} \hat{y}^{(N)}(t) = 0 \\ \hat{y}^{(k)}(t_0) = y^{(k)}(t_0) \quad \text{pour } k < N. \end{array} \right. \tag{E.11}$$

Exemple E.5. *Pour $N = 3$, l'égalité (E.10) s'écrit*

$$\hat{y}(t) = y(t_0) + y'(t_0)(t - t_0) + \frac{y^{(2)}(t_0)}{2}(t - t_0)^2.$$

et le système correspondant est le suivant

$$\left\{ \begin{array}{l} \hat{y}^{(3)}(t) = 0 \\ \hat{y}^{(2)}(t_0) = y^{(2)}(t_0) \\ \hat{y}'(t_0) = y'(t_0) \\ \hat{y}(t_0) = y(t_0) \end{array} \right.$$

Après transformation de Laplace du système (E.11), on obtient l'équation suivante :

$$s^N \hat{y}(s) - \sum_{j=1}^{N} s^{N-j} \hat{y}^{(j-1)}(s_0) = 0,$$

que l'on dérive N fois par rapport à la variable de Laplace s pour obtenir une nouvelle équation pour $\hat{y}(s)$ qui ne dépend plus des conditions initiales :

$$\frac{d^N}{ds^N}\left(s^N \hat{y}(s)\right) = 0.$$

On divise ensuite cette dernière équation, respectivement par $s^{N-1}, s^{N-2}, \ldots, s$ et on obtient le système triangulaire suivant :

$$(S) \left\{ \begin{array}{l} \frac{1}{s^{N-1}} \frac{d^N}{ds^N}\left(s^N \hat{y}(s)\right) = 0 \\ \vdots \\ \frac{1}{s} \frac{d^N}{ds^N}\left(s^N \hat{y}(s)\right) = 0 \end{array} \right.$$

Ce système va nous permettre d'obtenir, moyennant une transformée de Laplace inverse, une estimation des dérivées de $y(t)$. La précision de l'approximation dépend du choix de l'entier N. Dans le cas où l'on souhaite un ordre de dérivation égal à q, M. Fliess et H. Sira-Ramirez proposent de prendre $N = 4$ afin d'obtenir une estimation convenable. Nous allons reprendre maintenant l'exemple présenté dans [82] pour illustrer leur méthode.

E.3.2 Exemple

Considérons un signal arbitraire y, lisse dont on souhaite approximer les dérivées jusqu'à l'ordre deux. On prendra donc $N = 4$. Le système (S) à résoudre est le suivant :

$$(S) \left\{ \begin{array}{l} \frac{1}{s^3} \frac{d^4}{ds^4}\left(s^4 \hat{y}(s)\right) = 0 \\ \frac{1}{s^2} \frac{d^4}{ds^4}\left(s^4 \hat{y}(s)\right) = 0 \\ \frac{1}{s} \frac{d^4}{ds^4}\left(s^4 \hat{y}(s)\right) = 0 \end{array} \right. .$$

Un calcul élémentaire nous permet d'obtenir l'expression

$$\frac{d^4}{ds^4}\left(s^4\hat{y}\left(s\right)\right) = 24\hat{y}\left(s\right) + 96s\frac{d\hat{y}\left(s\right)}{ds} + 72s^2\frac{d^2\hat{y}\left(s\right)}{ds^2} + 16s^3\frac{d^3\hat{y}\left(s\right)}{ds^3} + s^4\frac{d^4\hat{y}\left(s\right)}{ds^4}.$$

Le système (S) peut donc s'écrire

$$(S)\left\{\begin{array}{l} \frac{24}{s^3}\hat{y}\left(s\right) + \frac{96}{s^2}\frac{d\hat{y}(s)}{ds} + \frac{72}{s}\frac{d^2\hat{y}(s)}{ds^2} + 16\frac{d^3\hat{y}(s)}{ds^3} + s\frac{d^4\hat{y}(s)}{ds^4} = 0 \\ \frac{24}{s^2}\hat{y}\left(s\right) + \frac{96}{s}\frac{d\hat{y}(s)}{ds} + 72\frac{d^2\hat{y}(s)}{ds^2} + 16s\frac{d^3\hat{y}(s)}{ds^3} + s^2\frac{d^4\hat{y}(s)}{ds^4} = 0 \\ \frac{24}{s}\hat{y}\left(s\right) + 96\frac{d\hat{y}(s)}{ds} + 72s\frac{d^2\hat{y}(s)}{ds^2} + 16s^2\frac{d^3\hat{y}(s)}{ds^3} + s^3\frac{d^4\hat{y}(s)}{ds^4} = 0. \end{array}\right.$$

On effectue une transformation de Laplace inverse. La première équation du système (S) devient

$$24\left(\iiint\hat{y}\right) - 96\left(\iint t\hat{y}\right) + 72\left(\int t^2\hat{y}\right) - 16t^3\hat{y} + \frac{d}{dt}(t^4\hat{y}) = 0,$$

c'est-à-dire

$$24\left(\iiint\hat{y}\right) - 96\left(\iint t\hat{y}\right) + 72\left(\int t^2\hat{y}\right) - 16t^3\hat{y} + (t^4\frac{d\hat{y}}{dt} + 4t^3\hat{y}) = 0,$$

soit encore, après simplification

$$24\left(\iiint\hat{y}\right) - 96\left(\iint t\hat{y}\right) + 72\left(\int t^2\hat{y}\right) - 12t^3\hat{y} + t^4\frac{d\hat{y}}{dt} = 0.$$

En isolant $\frac{d\hat{y}}{dt}$, on obtient :

$$\frac{d\hat{y}}{dt} = -\frac{24}{t^4}\left(\iiint\hat{y}\right) + \frac{96}{t^4}\left(\iint t\hat{y}\right) - \frac{72}{t^4}\left(\int t^2\hat{y}\right) + \frac{12}{t}\hat{y} + t^4 = 0. \tag{E.12}$$

La seconde équation du système (S) devient

$$24\left(\iint\hat{y}\right) - 96\left(\int t\hat{y}\right) + 72t^2\hat{y} - 16(t^3\frac{d\hat{y}}{dt} + 3t^2\hat{y}) + (t^4\frac{d^2\hat{y}}{dt^2} + 8t^3\frac{d\hat{y}}{dt} + 12t^2\hat{y}) = 0,$$

et, après simplification,

$$24\left(\iint\hat{y}\right) - 96\left(\int t\hat{y}\right) + 36t^2\hat{y} - 8t^3\frac{d\hat{y}}{dt} + t^4\frac{d^2\hat{y}}{dt^2} = 0.$$

En isolant $\frac{d^2\hat{y}}{dt}$, on obtient :

$$\frac{d^2\hat{y}}{dt} = -\frac{24}{t^4}\left(\iint\hat{y}\right) + \frac{96}{t^4}\left(\int t\hat{y}\right) - \frac{36}{t^2}\hat{y} + \frac{8}{t}\frac{d\hat{y}}{dt} \tag{E.13}$$

Notons que la troisième équation nous sera inutile ici étant donné que les équations (E.12) et (E.13) obtenues nous permettent déjà d'obtenir une estimation des dérivées, d'ordre un et deux du signal y. Remarquons que (E.12) et (E.13) peuvent s'écrire sous forme du système linéaire à coefficients non stationnaire (c'est-à-dire dont les coefficients varient dans le temps) suivant :

$$\begin{pmatrix} t^4 & 0 \\ -3t^2 & t^4 \end{pmatrix}\begin{pmatrix} \frac{d\hat{y}}{dt} \\ \frac{d^2\hat{y}}{dt} \end{pmatrix} = \begin{pmatrix} 12t^3\hat{y} \\ -36t^2\hat{y} \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

où

$$\left\{\begin{array}{rcl} \dot{z}_1 & = & z_2 - 72t^2\hat{y} \\ \dot{z}_2 & = & z_3 + 96t\hat{y} \\ \dot{z}_3 & = & -24\hat{y} \end{array}\right. .$$

ANNEXE F

Diagonalisation conjointe

Cette annexe est consacrée à la diagonalisation conjointe (ou encore simultanée) de matrices à coefficients réels. Les matrices que nous sommes amenés à diagonaliser dans le cadre de cette thèse sont des matrices d'autocorrélation. Ces dernières étant symétriques, elles sont toujours diagonalisables. L'objectif de cette annexe est de comprendre le principe diagonalisation conjointe.
Nous allons donc présenter, pour commencer, une méthode numérique de diagonalisation d'une matrice symétrique, appelée *méthode de Jacobi* [25].
Ensuite, nous présenterons une extension de cette méthode afin d'obtenir un algorithme de diagonalisation approchée d'une matrice quelconque. Nous chercherons alors à obtenir une matrice la plus diagonale possible, au sens d'un critère qui sera précisé. Enfin, forts de ces techniques, nous aborderons le principe de diagonalisation simultanée d'un ensemble de matrices [18]. Notons que les techniques que nous présentons dans cette annexe s'étendent naturellement aux matrices à coefficients complexes.

Il existe d'autres algorithmes de diagonalisation que celui proposé dans cette annexe. Notre présentation se veut avant tout pédagogique, son but étant de proposer un algorithme simple, complet et détaillé, même si ce dernier n'est pas directement (sous cette forme) applicable dans cette thèse.

F.1 Préliminaires

Dans ce paragraphe, nous énoncons quelques rappels élémentaires d'algèbre matricielle utilisés par la suite.

F.1.1 Matrices orthogonales

Définition F.1 (Matrice orthogonale élémentaire)**.** Une matrice orthogonale élémentaire (ou matrice de rotation) est une matrice, notée $\boldsymbol{\Omega}(p,q,c,s)$ (où $p<q$), de la forme

$$\boldsymbol{\Omega}(p,q,c,s) = \begin{array}{c} \begin{array}{ccccccc} & & p & & q & & \\ & & \downarrow & & \downarrow & & \end{array} \\ \begin{pmatrix} 1 & & 0 & & 0 & & 0 \\ & 1 & \vdots & & \vdots & & \\ 0 & \cdots & c & & s & \cdots & 0 \\ & & & 1 & & & \\ 0 & \cdots & -s & & c & \cdots & 0 \\ & & \vdots & & \vdots & 1 & \\ 0 & & 0 & & 0 & & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \longleftarrow \quad p\text{-ième ligne} \\ \\ \longleftarrow \quad q\text{-ième ligne} \\ \\ \\ \end{array}$$

où $c = \cos\theta$ et $s = \sin\theta$, avec θ un nombre réel.
Ainsi définie, $\mathbf{\Omega}(p,q,c,s)$ correspond à une matrice de rotation d'angle θ dans le plan des p-ièmes et q-ièmes vecteurs de base. Elle est bien orthogonale dans la mesure où

$$\mathbf{\Omega}(p,q,c,s)^T\mathbf{\Omega}(p,q,c,s) = \mathbf{Id}_n.$$

Maintenant, intéressons-nous au produit d'une matrice carré $\mathbf{A} = (a_{ij})_{i,j\in\mathbb{N}}$ quelconque par une matrice orthogonale élémentaire $\mathbf{\Omega}\,(p,q,c,s)$. Soit $\mathbf{B} = (b_{ij})_{i,j\in\mathbb{N}}$ définie par

$$\mathbf{B} = \mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega},$$

où $\mathbf{\Omega} = \mathbf{\Omega}(p,q,c,s)$. On a les relations suivantes entre les coefficients des matrices $\mathbf{A}$ et $\mathbf{B}$:

$$\begin{aligned}
&b_{ij} = a_{ij} \text{ pour } i,j \neq p \text{ et } i,j \neq q,\\
&b_{pj} = ca_{pi} - sa_{qi} \text{ pour } j \neq p,q,\\
&b_{qj} = sa_{pi} + ca_{qi} \text{ pour } j \neq p,q,\\
&b_{pp} = c^2a_{pp} - sca_{pq} - sca_{qp} + s^2a_{qq},\\
&b_{pq} = c^2a_{pq} - s^2a_{qp} + cs(a_{pp} - a_{qq}),\\
&b_{qp} = c^2a_{qp} - s^2a_{pq} + cs(a_{pp} - a_{qq}),\\
&b_{qq} = c^2a_{qq} + sca_{pq} + sca_{qp} + s^2a_{pp}.
\end{aligned} \tag{F.1}$$

Seules les p-ièmes et q-ièmes lignes et colonnes de la matrice $\mathbf{A}$ sont modifiées par la transformation

$$\mathbf{A} \to \mathbf{B} = \mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega}. \tag{F.2}$$

Nous allons voir que les relations (F.1) se simplifient dans le cas où la matrice $\mathbf{A}$ est symétrique.

F.1.2 Cas particulier : Matrice symétrique

Si la matrice $\mathbf{A}$ est symétrique, nous remarquons, d'après les expressions (F.1) et par symétrie de $\mathbf{A}$ (*i.e.*, $\forall i,j, a_{ij} = a_{ji}$), que les coefficients de la matrice $\mathbf{B} = (b_{ij})_{i,j\in\mathbb{N}}$ définie par $\mathbf{B} = \mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega}$ (où $\mathbf{\Omega} = \mathbf{\Omega}(p,q,c,s)$) vérifient

$$b_{pq} = b_{qp},$$

pour p et q fixés.

Proposition F.2. *Soient* $\mathbf{A}$ *une matrice symétrique,* $\mathbf{\Omega} = \mathbf{\Omega}(p,q,c,s)$ *une matrice de rotation élémentaire et* $\mathbf{B} = \mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega}$*. Si l'angle* θ *vérifie la relation*

$$\frac{1}{\tan 2\theta} = \frac{a_{qq} - a_{pp}}{2a_{pq}}, \tag{F.3}$$

alors, on a

$$b_{pq} = b_{qp} = 0. \tag{F.4}$$

Autrement dit, pour un angle θ *convenablement choisi, les coefficients* b_{pq} *et* b_{qp} *de la matrice* $\mathbf{B} = (b_{ij})_{i,j\in\mathbb{N}}$ *sont nuls.*

Démonstration. On recherche une condition sur l'angle θ de sorte que les coefficients b_{pq} et b_{qp} soient nuls. On veut

$$b_{pq} = b_{qp} = 0,$$

c'est-à-dire, d'après (F.1),

$$\left(c^2 - s^2\right) a_{pq} + cs(a_{pp} - a_{qq}) = 0,$$

avec $c = \cos\theta$ et $s = \sin\theta$. En utilisant les relations trigonométriques :

$$\cos^2\theta - \sin^2\theta = \cos 2\theta \text{ et } 2\cos\theta\sin\theta = \sin 2\theta,$$

la condition $b_{pq} = b_{qp} = 0$ devient

$$\cos 2\theta a_{pq} + \frac{\sin 2\theta}{2}(a_{pp} - a_{qq}) = 0,$$

soit encore

$$\frac{1}{\tan 2\theta} = \frac{a_{qq} - a_{pp}}{2a_{pq}}.$$

□

Par conséquent, lorsque θ vérifie (F.3), la transformation $\mathbf{A} \to \mathbf{B} = \mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega}$ permet d'annuler deux termes symétriques non diagonaux de la matrice $\mathbf{A}$. Cette transformation est à la base de l'algorithme de Jacobi pour diagonaliser une matrice symétrique. Toutefois, avant de présenter cet algorithme, nous énonçons un résultat fort utile par la suite.

F.1.3 Norme de Frobénius

Proposition F.3. *Soient* $\mathbf{A} = (a_{ij})_{i,j\in\mathbb{N}}$ *et* $\mathbf{B} = (b_{ij})_{i,j\in\mathbb{N}}$ *deux matrices carrées vérifiant la relation*

$$\mathbf{B} = \mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega},$$

où $\mathbf{\Omega} = \mathbf{\Omega}(p,q,c,s)$ *est une matrice de rotation élémentaire. On a*

$$\|\mathbf{A}\|_F = \|\mathbf{B}\|_F\,,$$

où $\|.\|_F$ *est la norme de Frobénius définie par*

$$\|\cdot\|_F : \left\{ \begin{array}{lcl} \mathbb{R}^{n\times n} & \longrightarrow & \mathbb{R} \\ \mathbf{M} = (m_{ij})_{i,j\in\mathbb{N}} & \longrightarrow & \displaystyle\sum_{i,j\in\mathbb{N}} m_{ij}^2 \end{array} \right. .$$

Démonstration. Ce résultat vient du fait que la norme $\|.\|_F$ de Frobénius est invariante par transformation orthogonale (conséquence de l'inégalité de Cauchy-Schwartz). Pour davantage de détails, le lecteur peut se référer à [25].
On a donc

$$\|\mathbf{B}\|_F = \left\|\mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega}\right\|_F = \|\mathbf{A}\|_F\,.$$

□

Après ces quelques rappels, nous nous intéressons à la méthode de Jacobi pour diagonaliser des matrices symétriques.

F.2 Diagonalisation d'une matrice symétrique

Soit $\mathbf{A} = (a_{ij})_{i,j\in\mathbb{N}}$ une matrice symétrique (donc diagonalisable) à coefficients réels. On souhaite diagonaliser cette matrice.

F.2.1 Principe

L'idée générale de la méthode de Jacobi repose sur la construction d'une suite de matrices orthogonales $(\mathbf{P}_n)_{n\in\mathbb{N}}$, telle que la suite $\left(\mathbf{P}_n^T\mathbf{A}\mathbf{P}_n\right)_{n\in\mathbb{N}}$ converge vers une matrice diagonale $\mathbf{D}$ dont les éléments diagonaux correspondent aux valeurs propres de $\mathbf{A}$ et telle que la suite $(\mathbf{P}_n)_{n\in\mathbb{N}}$ converge vers une matrice dont les colonnes correspondent aux vecteurs propres de la matrice $\mathbf{A}$.

On va chercher à annuler successivement tous les éléments hors diagonaux de la matrice $\mathbf{A}$ en la multipliant par des matrices orthogonales élémentaires judicieusement choisies (c'est-à-dire en utilisant une transformation du type (F.2)).

Itération 1 : Si $\mathbf{A}$ n'est pas diagonale alors il existe p et q, $p \neq q$ tels que $a_{pq} \neq 0$.
Dans le cas où plusieurs éléments non diagonaux a_{pq}, $p \neq q$, sont non nuls, on choisit le couple (p, q) de sorte que le coefficient a_{pq} ait la plus grande valeur absolue parmi les coefficients non diagonaux de $\mathbf{A}$, c'est-à-dire tel que

$$|a_{pq}| = \max_{i\neq j} |a_{ij}| \,. \tag{F.5}$$

Puis, on considère la transformation

$$\mathbf{A} \to \mathbf{B}_1 = \mathbf{\Omega}_1^T\mathbf{A}\mathbf{\Omega}_1,$$

où $\mathbf{\Omega}_1 = \mathbf{\Omega}(p, q, c, s)$, avec $c = \cos\theta$, $s = \sin\theta$ et θ qui vérifie la relation (F.3). La matrice $\mathbf{\Omega}_1$ ainsi choisie permet d'obtenir une nouvelle matrice $\mathbf{B}_1$ telle que $b_{pq} = b_{qp} = 0$. On a ainsi annulé deux éléments non diagonaux de la matrice $\mathbf{A}$.

Itération 2 : Si la matrice $\mathbf{B}_1$ obtenue n'est pas diagonale , on réitère le même procédé afin d'obtenir une nouvelle matrice

$$\mathbf{B}_2 = \mathbf{\Omega}_2^T\mathbf{B}_1\mathbf{\Omega}_2,$$

où $\mathbf{\Omega}_2$ est une matrice orthogonale élémentaire choisie pour annuler le coefficient maximal (au sens de (F.5)) de $\mathbf{B}_1$.

Ainsi de suite, on itére le procédé jusqu'à obtention d'une matrice diagonale.

Nous admettons ici que ce processus converge, c'est-à-dire qu'il existe $N \in \mathbb{N}$ tel que $\mathbf{B}_N$ soit diagonale (pour davantage de détails, consulter [25]). On définit alors la suite $(\mathbf{P}_n)_{n\in\mathbb{N}}$ par $\mathbf{P}_n = \mathbf{\Omega}_n \ldots \mathbf{\Omega}_2\mathbf{\Omega}$. Ainsi, la suite $\left(\mathbf{P}_n^T\mathbf{A}\mathbf{P}_n\right)_{n\in\mathbb{N}}$ converge vers une matrice diagonale $\mathbf{D} = \mathbf{P}_N^T\mathbf{A}\mathbf{P}_N$. Et, dans le cas où les valeurs propres de $\mathbf{A}$ sont distinctes, les vecteurs colonnes de la matrice $\mathbf{P}$ correspondent aux vecteurs propres associés à ces valeurs propres.

F.2.2 Algorithme de Jacobi

Entrées : $\mathbf{A} = (a_{ij})_{i,j\in\mathbb{N}}$: matrice symétrique quelconque

Sorties : $\mathbf{A}$: matrice diagonale

Algorithme :

Tant qu'il existe un couple $p \neq q$ tel que $a_{pq} \neq 0$
- Choisir le couple (p, q) tel que $|a_{pq}| = \max_{i \neq j} |a_{ij}|$,
- Calculer θ de sorte que $\frac{1}{\tan 2\theta} = \frac{a_{qq} - a_{pp}}{2a_{pq}}$,
- En déduire la matrice $\mathbf{\Omega}(p, q, c, s)$ correspondante,
- $\mathbf{A} := \mathbf{\Omega}(p, q, c, s)^T \mathbf{A} \mathbf{\Omega}(p, q, c, s)$.

Fin Tant que

F.2.3 Interprétation géométrique

Illustrons graphiquement la méthode de Jacobi dans le cas particulier de la diagonalisation d'une matrice symétrique carrée d'ordre 2.

Soit $\mathbf{A}$, la matrice définie par

$$\mathbf{A} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}.$$

La matrice $\mathbf{A}$ peut être caractérisée par une ellipse (voir figure F.1) notée $(\mathcal{E}_1)$, qui correspond à la représentation graphique de la courbe d'équation

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = cste,$$

c'est-à-dire d'équation

$$x_1^2 + x_1 x_2 + x_2^2 = cste.$$

.

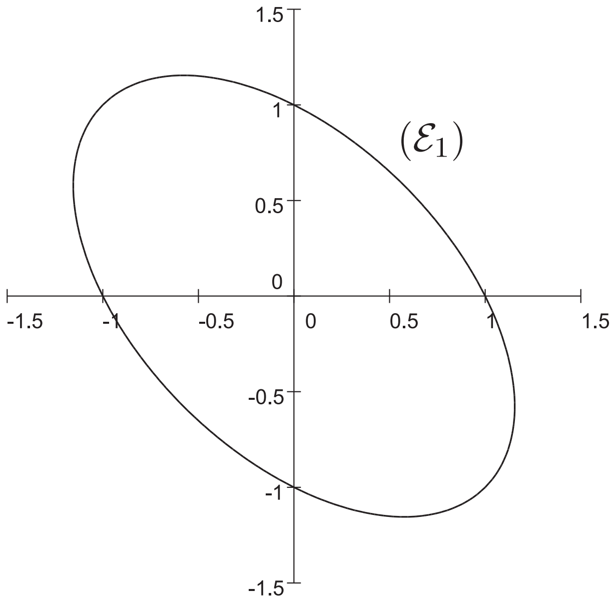


Figure F.1 – Ellipse $(\mathcal{E}_1)$ associée à la matrice $\mathbf{A}$.

Diagonalisons la matrice $\mathbf{A}$. Considérons la matrice de rotation $\mathbf{\Omega}$ suivante,

$$\mathbf{\Omega} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

où θ vérifie la relation (F.3), c'est-à-dire

$$\frac{1}{\tan 2\theta} = 0 \Longleftrightarrow \theta = \frac{\pi}{4}.$$

On a

$$\mathbf{\Omega} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix},$$

et

$$\mathbf{D} = \mathbf{\Omega}^T \mathbf{A} \mathbf{\Omega} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

On obtient bien une matrice diagonale.

Rappelons que la diagonalisation peut être vue comme un changement de base

$$\mathbf{x}' = \mathbf{\Omega}\mathbf{x}.$$

Notons que les vecteurs colonnes de la matrice $\mathbf{\Omega}$ correspondent bien aux vecteurs propres orthonormés de la matrice $\mathbf{A}$, associés respectivement aux valeurs propres $\frac{1}{2}$ et $\frac{3}{2}$.

La figure F.2 correspond à la représentation graphique de la courbe d'équation

$$\mathbf{x}^T \mathbf{D} \mathbf{x} = cste,$$

c'est-à-dire d'équation

$$x_1^2 + \left(\sqrt{3}x_2\right)^2 = cste.$$

Il s'agit d'une ellipse, notée $(\mathcal{E}_2)$. Cette ellipse est l'image par une rotation d'angle $\theta = \frac{\pi}{4}$ de l'ellipse $(\mathcal{E}_1)$ de la figure F.1.

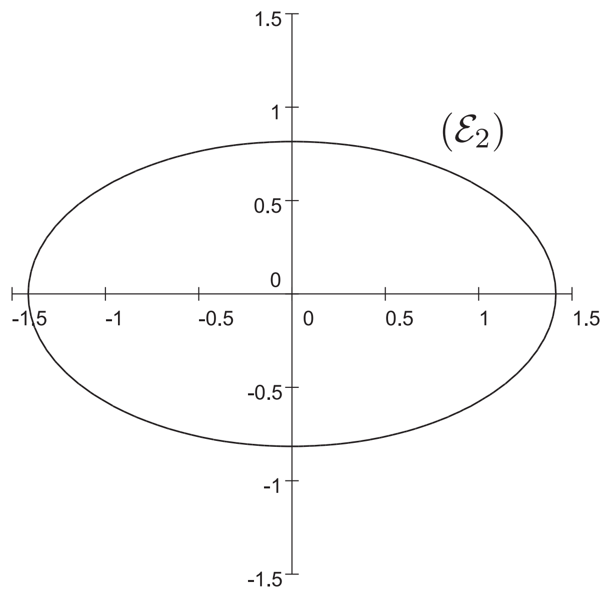


Figure F.2 – Ellipse $(\mathcal{E}_2)$ associée à la matrice diagonale $\mathbf{D}$.

F.2.4 Conclusion

L'algorithme proposé ici dans le cadre de la diagonalisation d'une matrice symétrique est relativement simple. Nous nous intéressons maintenant à sa généralisation pour une matrice quelconque (diagonalisable ou non) afin d'obtenir une matrice aussi diagonale que possible, au sens d'un certain critère. Nous présentons une méthode inspirée de l'algorithme de Jacobi pour la diagonalisation approchée d'une matrice à coefficients réels.

Les algorithmes que nous présentons dans les paragraphes suivants sont inspirés de [18].

F.3 Diagonalisation approchée

Dans ce paragraphe, nous présentons l'extension de la méthode de Jacobi pour la diagonalisation approchée d'une matrice réelle. Considérons une matrice $\mathbf{A}$ quelconque à coefficients réels.

F.3.1 Introduction

Le principe de la diagonalisation approchée est le même que précédemment, il s'agit de rechercher une suite de matrices orthogonales $(\mathbf{P}_n)_{n\in\mathbb{N}}$, telle que la suite $\left(\mathbf{P}_n^T\mathbf{A}\mathbf{P}_n\right)_{n\in\mathbb{N}}$ converge vers une matrice $\mathbf{M}$ la plus diagonale possible au sens du critère suivant :

$$\underset{\mathbf{M}=(m_{ij})}{\text{Minimiser}} \sum_{i\neq j} |m_{ij}|^2 \,. \tag{F.6}$$

Autrement dit, nous cherchons la matrice $\mathbf{M} = (m_{ij})_{i,j\in\mathbb{N}}$ dont la somme quadratique des éléments non diagonaux est minimale.

On va chercher à minimiser les éléments non diagonaux de la matrice $\mathbf{A}$ en utilisant une transformation du type (F.2), c'est-à-dire en considérant les produits de $\mathbf{A}$ par des matrices orthogonales élémentaires.

Avant de préciser le principe de l'algorithme, étudions plus en détails le critère (F.6).

F.3.2 Etude du critère à minimiser

Nous cherchons à remplacer le critère (F.6) par un critère équivalent plus simple [18].

Considérons le problème suivant

$$\underset{c,s}{\text{Minimiser}} \sum_{i\neq j} |b_{ij}|^2 \,, \tag{F.7}$$

où les b_{ij} sont les coefficients de la matrice $\mathbf{B}_1 = \mathbf{\Omega}_1^T\mathbf{A}\mathbf{\Omega}_1$. On rappelle que l'expression des coefficients b_{ij} en fonction des coefficients a_{ij} de la matrice $\mathbf{A}$ sont donnés par les formules (F.1). On a

$$\sum_{i\neq j} b_{ij}^2 = \sum_{i,j} b_{ij}^2 - \sum_{i} b_{ii}^2,$$

donc, d'après la proposition F.3,

$$\sum_{i\neq j} b_{ij}^2 = \sum_{i,j} a_{ij}^2 - \sum_{i} b_{ii}^2.$$

Or, les coefficients a_{ij} de la matrice $\mathbf{A}$ sont indépendants de c et s, donc le problème (F.7) devient

$$\underset{c,s}{\text{Maximiser}} \sum_i b_{ii}^2. \tag{F.8}$$

Etant donné que b_{pp} et b_{qq} sont les seuls éléments diagonaux de $\mathbf{B}_1$ dépendant de s et c, le problème (F.8) est équivalent à

$$\underset{c,s}{\text{Maximiser}}\, b_{pp}^2 + b_{qq}^2. \tag{F.9}$$

En appliquant l'égalité triangulaire[1] à b_{pp} et b_{qq}, on trouve

$$\begin{aligned} b_{pp}^2 + b_{qq}^2 &= \frac{1}{2}\left[(b_{pp}+b_{qq})^2 + (b_{pp}-b_{qq})^2\right] \\ &= \frac{1}{2}\left[\left(c^2 a_{pp} + s^2 a_{qq} + c^2 a_{qq} + s^2 a_{pp}\right)^2 + (b_{pp}-b_{qq})^2\right] \\ &= \frac{1}{2}\left[(a_{pp}+a_{qq})^2 + (b_{pp}-b_{qq})^2\right]. \end{aligned}$$

De même que précedemment, les coefficients a_{ij} étant indépendants de c et s, le problème (F.9) devient

$$\underset{c,s}{\text{Maximiser}}\; (b_{pp}-b_{qq})^2. \tag{F.10}$$

Or, en remplaçant b_{pp} et b_{qq} par leurs valeurs données par les expressions (F.1), on a

$$\begin{aligned} b_{pp} - b_{qq} &= \left(c^2 a_{pp} - sc a_{pq} - sc a_{qp} + s^2 a_{qq}\right) - \left(c^2 a_{qq} + sc a_{pq} + sc a_{qp} + s^2 a_{pp}\right) \\ &= \left(c^2 - s^2\right)(a_{pp} - a_{qq}) - 2sc\,(a_{pq} + a_{qp}). \end{aligned}$$

Posons

$$b_{pp} - b_{qq} = \mathbf{x}(c,s)^T \mathbf{y}_A,$$

où $\mathbf{x}(c,s)$ et $\mathbf{y}_A$ sont des vecteurs définis par

$$\mathbf{x}(c,s) = \begin{pmatrix} c^2 - s^2 \\ -2sc \end{pmatrix} \text{ et } \mathbf{y}_A = \begin{pmatrix} a_{pp} - a_{qq} \\ a_{pq} + a_{qp} \end{pmatrix}.$$

Ainsi, le problème (F.10) devient

$$\underset{c,s}{\text{Maximiser}}\, \mathbf{x}(c,s)^T \mathbf{y}_A \mathbf{y}_A^T \mathbf{x}(c,s), \tag{F.11}$$

et finalement, en posant $\mathbf{Q} = \mathbf{y}_A \mathbf{y}_A^T$, ceci équivaut à

$$\underset{c,s}{\text{Maximiser}}\, \mathbf{x}(c,s)^T \mathbf{Q} \mathbf{x}(c,s). \tag{F.12}$$

Autrement dit, il s'agit de maximiser la forme quadratique $q(\mathbf{x}) = \mathbf{x}^T \mathbf{Q} \mathbf{x}$ sur $\left\{\mathbf{x}(c,s), c^2+s^2=1\right\}$. Nous énonçons le résultat suivant :

Proposition F.4.

$$\left\{\mathbf{x}(c,s), c^2+s^2=1\right\} = \left\{\begin{pmatrix} x_1 & x_2 \end{pmatrix}^T, x_1^2 + x_2^2 = 1\right\}.$$

[1] $\forall x, y \in \mathbb{R}, x^2 + y^2 = \frac{1}{2}\left((x+y)^2 - (x-y)^2\right).$

Démonstration. Soit $\mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^{\mathrm{T}}$ tel que $x_1^2 + x_2^2 = 1$. Posons

$$c = \begin{cases} 1 & \text{si } x_1 = -1 \text{ et } x_2 = 0, \\ \sqrt{\frac{x_1+1}{2}} & \text{sinon.} \end{cases}$$

$$s = \begin{cases} 0 & \text{si } x_1 = -1 \text{ et } x_2 = 0, \\ \frac{x_2}{\sqrt{2(x_1+1)}} & \text{sinon.} \end{cases}$$

Ainsi définis, c et s vérifient

$$c^2 + s^2 = 1,$$

donc $\left\{ \begin{pmatrix} x_1 & x_2 \end{pmatrix}^{\mathrm{T}}, x_1^2 + x_2^2 = 1 \right\} \subset \left\{ \mathbf{x}(c,s), c^2 + s^2 = 1 \right\}$.

Réciproquement, soit $\mathbf{x}(c,s) = \begin{pmatrix} c^2 - s^2 \\ -2sc \end{pmatrix}$ tel que $c^2 + s^2 = 1$. On a

$$\left(c^2 - s^2\right)^2 + (-2sc)^2 = c^4 + s^4 + 2s^2c^2 = (c^2 + s^2)^2 = 1.$$

Donc, $\left\{ \mathbf{x}(c,s), c^2 + s^2 = 1 \right\} \subset \left\{ \begin{pmatrix} x_1 & x_2 \end{pmatrix}^{\mathrm{T}}, x_1^2 + x_2^2 = 1 \right\}$. □

Ainsi, grâce à ce résultat, maximiser la forme quadratique $q(\mathbf{x}) = \mathbf{x}^T \mathbf{Q} \mathbf{x}$ sur $\left\{ \mathbf{x}(c,s), c^2 + s^2 = 1 \right\}$ est équivalent à la maximiser sur la sphère unité $\left\{ \begin{pmatrix} x_1 & x_2 \end{pmatrix}^{\mathrm{T}}, x_1^2 + x_2^2 = 1 \right\}$. On s'est donc ramené au problème suivant

$$\underset{\{(x_1 \;\; x_2)^{\mathrm{T}}, x_1^2 + x_2^2 = 1\}}{\text{Maximiser}} \; q(\mathbf{x}). \tag{F.13}$$

Cette dernière condition revient à trouver le vecteur propre associé à la plus grande valeur propre de la matrice $\mathbf{Q}$.

F.3.3 Principe

Détaillons maintenant l'algorithme de diagonalisation. Considérons le problème suivant :

Itération 1 : Si $\mathbf{A}$ n'est pas diagonale alors il existe p et q, $p \neq q$ tels que $a_{pq} \neq 0$. On choisit le couple (p,q) de sorte que le coefficient a_{pq} soit maximal au sens du critère (F.5).
Puis, on considère la transformation

$$\mathbf{A} \to \mathbf{B}_1 = \mathbf{\Omega}_1^T \mathbf{A} \mathbf{\Omega}_1,$$

où $\mathbf{\Omega}_1 = \mathbf{\Omega}(p,q,c,s)$. On choisit les paramètres c et s de la matrice $\mathbf{\Omega}_1$ de sorte que la somme quadratique des éléments non diagonaux de $\mathbf{\Omega}_1^T \mathbf{A} \mathbf{\Omega}_1$ soit minimale. Pour cela, on cherche le vecteur propre associé à la plus grande valeur propre de la matrice $\mathbf{Q} = \mathbf{y}_A \mathbf{y}_A^T$ (voir paragraphe F.3.2). On obtient ainsi une matrice $\mathbf{B}_1 = \mathbf{\Omega}_1^T \mathbf{A} \mathbf{\Omega}_1$ dont la somme quadratique des éléments non diagonaux est inférieure à la somme quadratique des éléments non diagonaux de la matrice $\mathbf{A}$, autrement dit

$$\sum_{i \neq j} |b_{ij}|^2 \leq \sum_{i \neq j} |a_{ij}|^2,$$

où $\mathbf{A} = (a_{ij})_{i,j}$ et $\mathbf{B}_1 = (b_{ij})_{i,j}$.

Itération 2 : On réitère le même procédé avec la matrice $\mathbf{B}_1$ afin d'obtenir une nouvelle matrice $\mathbf{B}_2 = \mathbf{\Omega}_2^T \mathbf{B}_1 \mathbf{\Omega}_2$, où $\mathbf{\Omega}_2$ est une matrice orthogonale élémentaire choisie pour diminuer la somme quadratique des éléments non diagonaux de $\mathbf{B}_1$.

Ainsi de suite, on itére le procédé jusqu'à obtention d'une matrice la plus diagonale possible au sens du critère (F.6).

Convergence : La convergence de ce processus est assurée étant donné que toute suite décroissante minorée de $\mathbb{R}$ converge. Toutefois, il n'a pas été démontré qu'il converge bien vers le minimum au sens de (F.6). Avant d'observer sur un exemple concret la convergence du processus, présentons l'algorithme de diagonalisation approchée d'une matrice réelle.

F.3.4 Algorithme

Entrées : $\mathbf{A} = (a_{ij})_{i,j\in\mathbb{N}}$: matrice quelconque à coefficients réels.
$\mathbf{M} = (m_{ij})_{i,j\in\mathbb{N}}$: matrice initialement égale à la matrice identité.

Sorties : $\mathbf{A}$: matrice la plus diagonale possible au sens du critère (F.6)

Algorithme :

Tant que $\sum_{i\neq j} |a_{ij}|^2 < \sum_{i\neq j} |m_{ij}|^2$,

- Choisir le couple (p, q) tel que $|a_{pq}| = \max_{i\neq j} |a_{ij}|$,
- Calculer les paramètres c et s qui minimisent la somme quadratique des éléments non diagonaux de la matrice $\mathbf{\Omega}(p, q, c, s)^T \mathbf{A}\mathbf{\Omega}(p, q, c, s)$.
- $\mathbf{M} := \mathbf{A}$.
- $\mathbf{A} := \mathbf{\Omega}(p, q, c, s)^T \mathbf{A}\mathbf{\Omega}(p, q, c, s)$.

Fin Tant que

F.3.5 Exemple

Nous illustrons cet algorithme dans le cas particulier d'une matrice $\mathbf{A}$ de dimension 4×4, définie par

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & 1 & 0 \\ 0 & 1 & 0 & 4 \\ -1 & 0 & -1 & 3 \\ 0 & 0 & 0 & 7 \end{pmatrix}.$$

Pour cette matrice, la somme quadratique s_1 des éléments non diagonaux vaut

$$s_1 = \sum_{i\neq j} |a_{ij}|^2 = 52.$$

Nous cherchons à rendre la matrice $\mathbf{A}$ la "plus diagonale possible" au sens de (F.6). Nous allons détailler les deux premières itérations de l'algorithme afin de mettre en évidence la convergence du processus.

Itération 1 : Le coefficient de la matrice $\mathbf{A}$ ayant la plus grande valeur absolue est $a_{12} = 5$ $(p = 1, q = 2)$. Considérons donc la matrice $\mathbf{\Omega}_1$ définie par

$$\mathbf{\Omega}_1 = \begin{pmatrix} c & s & 0 & 0 \\ -s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

où $c^2 + s^2 = 1$. On doit choisir c et s de sorte que la somme quadratique des éléments non diagonaux de $\mathbf{\Omega}^T\mathbf{A}\mathbf{\Omega}$ soit minimale. Nous avons vu que ceci revient à résoudre le problème (F.12), autrement dit, à trouver le vecteur propre associé à la plus grande valeur propre de

$$\mathbf{Q}_1 = \mathbf{y}_A\mathbf{y}_A^T = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}.$$

La plus grande valeur propre de $\mathbf{Q}$ est 25. Le vecteur propre associé est $\mathbf{x} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. On obtient $c = \frac{\sqrt{2}}{2}$ et $s = \frac{\sqrt{2}}{2}$ donc

$$\mathbf{\Omega}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

et on en déduit

$$\mathbf{A}_2 = \mathbf{\Omega}_1^T\mathbf{A}\mathbf{\Omega}_1 = \begin{pmatrix} -\frac{3}{2} & \frac{5}{2} & \frac{\sqrt{2}}{2} & -2\sqrt{2} \\ -\frac{5}{2} & \frac{7}{2} & \frac{\sqrt{2}}{2} & 2\sqrt{2} \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & -1 & 3 \\ 0 & 0 & 0 & 7 \end{pmatrix}.$$

La somme quadratique s_2 des éléments non diagonaux de $\mathbf{A}_2$ vaut

$$s_2 = \sum_{i \neq j} |a_{ij}|^2 = 39,5.$$

$\mathbf{A}_2$ est donc plus diagonale que $\mathbf{A}$ au sens du critère (F.6).

Itération 2 : Le coefficient de la matrice $\mathbf{A}_2$ ayant la plus grande valeur absolue est $a_{34} = 3$. Considérons la matrice $\mathbf{\Omega}_2$ définie par

$$\mathbf{\Omega}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & s \\ 0 & 0 & -s & c \end{pmatrix}$$

où $c^2 + s^2 = 1$. On doit choisir c et s de sorte que la somme quadratique des éléments non diagonaux de $\mathbf{\Omega}_2^T\mathbf{A}\mathbf{\Omega}_2$ soit minimale. Ceci revient à trouver le vecteur propre associé à la plus grande valeur propre de

$$\mathbf{Q}_2 = \mathbf{y}_{A_2}\mathbf{y}_{A_2}^T = \begin{pmatrix} 64 & -24 \\ -24 & 9 \end{pmatrix}.$$

La plus grande valeur propre de $\mathbf{Q}_2$ est 73. Le vecteur propre associé est $\mathbf{x} = \begin{pmatrix} -\frac{8}{\sqrt{73}} \\ \frac{3}{\sqrt{73}} \end{pmatrix}$. On obtient $c = \sqrt{\frac{-\frac{8}{\sqrt{73}}+1}{2}}$ et $s = \frac{\frac{3}{\sqrt{73}}}{\sqrt{2\left(-\frac{8}{\sqrt{73}}+1\right)}}$ donc

$$\mathbf{\Omega}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \sqrt{\frac{-\frac{8}{\sqrt{73}}+1}{2}} & \frac{\frac{3}{\sqrt{73}}}{\sqrt{2\left(-\frac{8}{\sqrt{73}}+1\right)}} \\ 0 & 0 & -\frac{\frac{3}{\sqrt{73}}}{\sqrt{2\left(-\frac{8}{\sqrt{73}}+1\right)}} & \sqrt{\frac{-\frac{8}{\sqrt{73}}+1}{2}} \end{pmatrix}.$$

et on en déduit

$$A_3 = \mathbf{\Omega}_2^T \mathbf{A}_2 \mathbf{\Omega}_2 = \begin{pmatrix} -1.5 & 2.5 & 2.909\,2 & 0.191\,10 \\ -2.5 & 3.5 & -2.656\,9 & 1.200\,4 \\ -0.126\,17 & -0.126\,17 & 6.218\,6 & -1.309\,0 \\ -0.695\,76 & -0.695\,76 & -4.309\,0 & -0.218\,63 \end{pmatrix}.$$

La somme quadratique s_3 des éléments non diagonaux de $\mathbf{A}_3$ vaut

$$s_3 = \sum_{i \neq j} |a_{ij}|^2 = 21.304.$$

$\mathbf{A}_3$ est donc plus diagonale que $\mathbf{A}_2$ au sens du critère (F.6).

Ainsi de suite, on itère le processus afin d'obtenir une matrice la plus diagonale possible.

F.4 Diagonalisation simultanée approchée de matrices

Dans cette dernière partie, nous allons nous intéresser à la diagonalisation conjointe approchée de matrices réelles.

Le problème est le suivant : on considère un ensemble

$$E = \{\mathbf{A}_k, k = 1, \ldots, K\}$$

de K matrices que l'on souhaite diagonaliser simultanément. C'est-à-dire que nous cherchons une matrice orthogonale $\mathbf{P}$ telle que pour tout k,

$$\mathbf{P}^T \mathbf{A}_k \mathbf{P}$$

soit diagonale ou la plus diagonale possible (au sens du critère (F.6) précédemment défini).

Dans un premier temps, nous présentons l'algorithme de type Jacobi, détaillé dans [18], permettant de résoudre ce problème. Puis, nous illustrons cet algorithme dans le cas de deux matrices symétriques de dimension 2×2.

F.4.1 Principe

Le principe est similaire à celui présenté dans le cadre de la diagonalisation approchée d'une matrice réelle. Nous allons construire une suite de matrices orthogonales $(\mathbf{P}_n)_{n\in\mathbb{N}}$, telle que pour tout k, la suite $\left(\mathbf{P}_n^T \mathbf{A}_k \mathbf{P}_n\right)_{n\in\mathbb{N}}$ converge vers une matrice $\mathbf{M}_k$ la plus diagonale possible au sens du critère (F.6). On va chercher à minimiser les éléments non diagonaux des matrices $\mathbf{A}_k = \left(a_{k_{ij}}\right)_{i,j}$ en utilisant des transformations du type (F.2).

Itération 1 : On choisit le couple (p, q) de sorte que

$$\sum_{k=1}^{K} \left|a_{k_{pq}}\right| = \max_{i \neq j} \sum_{k=1}^{K} \left|a_{k_{ij}}\right|,$$

c'est-à-dire celui dont la somme des valeurs absolues des coefficients des matrices $\mathbf{A}_k$ aux indices de ligne et colonne respectivement p et q est maximale.

On choisit une matrice orthogonale élémentaire $\mathbf{\Omega}_1 = \mathbf{\Omega}_1(p, q, c, s)$ telle que, pour tout k, la matrice $\mathbf{\Omega}_1^{\mathrm{T}} \mathbf{A}_k \mathbf{\Omega}_1$ soit plus la plus diagonale possible au sens de (F.6) qui équivaut à (F.12). Le problème est donc le suivant :

$$\underset{c,s}{\text{Minimiser}} \sum_{k=1}^{K} \sum_{i \neq j} \left|a_{k_{ij}}\right|^2, \tag{F.14}$$

où les $a_{k_{ij}}$ sont les coefficients des matrices $\mathbf{A}_k$. Par analogie avec l'algorithme présenté au paragraphe F.3, la matrice $\mathbf{\Omega}_1$ doit donc répondre au problème suivant :

$$\underset{\{(x_1,x_2)\}, x_1^2 + x_2^2 = 1\}}{\text{Maximiser}} \sum_{k=1}^{K} q_k(\mathbf{x}), \tag{F.15}$$

où $q_k(\mathbf{x}) = \mathbf{x} \mathbf{y}_{A_k} \mathbf{y}_{A_k}^T \mathbf{x}^T$.

Cette dernière condition revient à trouver le vecteur propre associé à la plus grande valeur propre de la matrice

$$\mathbf{Q} = \sum_{k=1}^{K} \mathbf{y}_{A_k} \mathbf{y}_{A_k}^T.$$

et ainsi trouver les paramètres c et s de la matrice $\mathbf{\Omega}_1$ (tels que pour tout k, les matrices $\mathbf{\Omega}_1^T \mathbf{A}_k \mathbf{\Omega}_1$ soient plus diagonales que les matrices $\mathbf{A}_k$, au sens du critère (F.14)).

Itération 2 : On réitère le même procédé avec les matrices $\mathbf{\Omega}_1^T \mathbf{A}_k \mathbf{\Omega}_1$ $(\forall k)$ obtenues. On recherche ainsi une nouvelle matrice orthogonale élémentaire $\mathbf{\Omega}_2$ afin de diminuer la somme des sommes quadratiques des éléments non diagonaux des matrices $\mathbf{\Omega}_1^T \mathbf{A}_k \mathbf{\Omega}_1$.

Ainsi de suite, on itére le procédé jusqu'à obtention de matrices les plus diagonales possible au sens du critère (F.15).

La convergence de ce processus est assuré étant donné que toute suite décroissante minorée de $\mathbb{R}$ converge mais comme dans le cas précédent, il n'est pas certain qu'il converge vers le minimum.

F.4.2 Algorithme

Entrées : $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_K$: matrices quelconques à coefficients réels.
$\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_K$: matrices initialement matrices identités.

Sorties : $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_K$: matrices les plus diagonales possibles.

Algorithme :

Tant que $\sum_{k=1}^{K}\sum_{i\neq j}\left|a_{k_{ij}}\right|^2 < \sum_{k=1}^{K}\sum_{i\neq j}\left|m_{k_{ij}}\right|^2$,

- Choisir le couple (p,q) tel que $\sum_{k=1}^{K}\left|a_{k_{pq}}\right| = \max_{i\neq j}\sum_{k=1}^{K}\left|a_{k_{ij}}\right|$,
- Calculer les paramètres c et s qui minimisent la somme des sommes quadratiques des éléments non diagonaux des matrices $\boldsymbol{\Omega}(p,q,c,s)^T\mathbf{A}_k\boldsymbol{\Omega}(p,q,c,s)$ au sens de (F.14).
- $\mathbf{M}_k := \mathbf{A}_k$.
- $\mathbf{A}_k := \boldsymbol{\Omega}(p,q,c,s)^T\mathbf{A}_k\boldsymbol{\Omega}(p,q,c,s)$.

Fin Tant que

F.4.3 Exemple

Nous illustrons la méthode de diagonalisation simultanée dans le cas particulier d'un couple de matrices symétriques de dimension 2×2. Dans un premier temps, nous proposons un résultat général issu de [75] puis nous l'illustrons graphiquement.

F.4.3.1 Diagonalisation conjointe de deux matrices symétriques de dimension 2×2

Dans [75], Pham propose une solution analytique pour la diagonalisation conjointe de deux matrices symétriques réelles de dimension 2×2. Nous présentons ici ses résultats.

Proposition F.5. *Soient $\mathbf{A}_1$ et $\mathbf{A}_2$, deux matrices symétriques réelles non proportionnelles définies par*

$$\mathbf{A}_1 = \begin{pmatrix} a & c \\ c & b \end{pmatrix},$$
$$\mathbf{A}_2 = \begin{pmatrix} d & f \\ f & e \end{pmatrix}.$$

On suppose que $\mathbf{A}_1$ est définie positive. Alors, la matrice $\mathbf{P}$ définie par

$$\mathbf{P} = \begin{pmatrix} \delta & 2\gamma \\ 2\alpha & \delta \end{pmatrix},$$

où

$$\begin{aligned}
&\alpha = bf - ce\\
&\beta = ae - bd\\
&\gamma = cd - af\\
&\delta = \beta + sign\,(\beta)\,\sqrt{\beta^2 - 4\alpha\gamma}.
\end{aligned}$$

diagonalise conjointement $\mathbf{A}_1$ et $\mathbf{A}_2$, autrement dit, les matrices $\mathbf{P}^T\mathbf{A}_1\mathbf{P}$ et $\mathbf{P}^T\mathbf{A}_2\mathbf{P}$ sont diagonales.

Remarque F.6. *Bien entendu, toutes les matrices de la forme $\mathbf{DP}$, où $\mathbf{D}\in\mathfrak{D}$, diagonalisent également les matrices $\mathbf{A}_1$ et $\mathbf{A}_2$.*

Pour clore cette annexe, nous avons choisi d'illustrer graphiquement la diagonalisation conjointe de deux matrices symétriques réelles (dont l'une est définie positive), afin de mettre en évidence le rôle des matrices orthogonales élémentaires $\boldsymbol{\Omega}$ (matrices de passage).

F.4.3.2 Illustration graphique

Considérons les matrices $\mathbf{A}_1$ et $\mathbf{A}_2$ symétriques suivantes

$$\mathbf{A}_1 = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$$
$$\mathbf{A}_2 = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Les ellipses $(\mathcal{E}_1)$ et $(\mathcal{E}_2)$, représentées sur la figure F.3 correspondent aux représentations graphiques des formes quadratiques associées aux matrices $\mathbf{A}_1$ et $\mathbf{A}_2$.

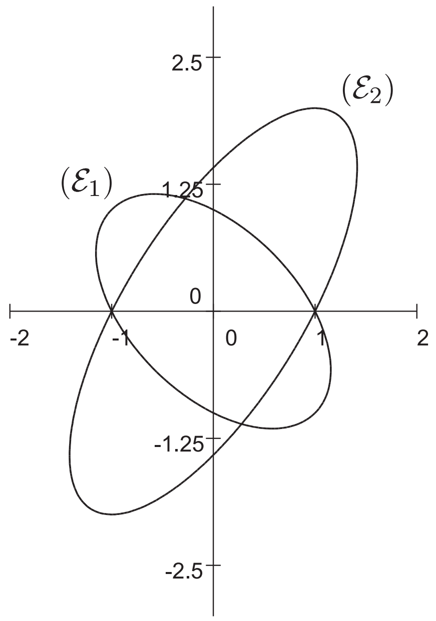


Figure F.3 – Ellipses $(\mathcal{E}_1)$ et $(\mathcal{E}_2)$ associées respectivement aux matrices $\mathbf{A}_1$ et $\mathbf{A}_2$.

La matrice $\mathbf{A}_1$ étant définie positive, ces deux matrices sont conjointement diagonalisables. Nous allons détailler chaque transformation subie par les matrices $\mathbf{A}_1$ et $\mathbf{A}_2$ jusqu'à obtention de deux matrices diagonales.

Etape 1 : On considère la matrice orthogonale élémentaire

$$\mathbf{\Omega}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

On obtient

$$\mathbf{\Omega}_1^T \mathbf{A}_1 \mathbf{\Omega}_1 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix},$$
$$\mathbf{\Omega}_1^T \mathbf{A}_2 \mathbf{\Omega}_1 = \begin{pmatrix} \frac{5}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Les ellipses $(\mathcal{E}_1')$ et $(\mathcal{E}_2')$, représentées sur la figure F.4, correspondent aux représentations graphiques des formes quadratiques associées aux matrices $\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1$ et $\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1$.

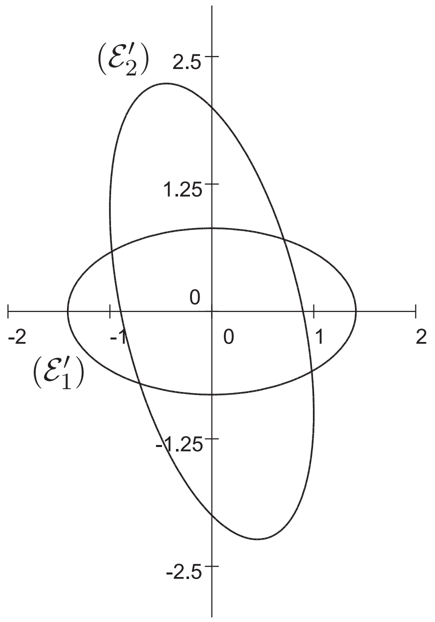


Figure F.4 – Ellipses $(\mathcal{E}_1')$ et $(\mathcal{E}_2')$ associées respectivement aux matrices $\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1$ et $\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1$.

Etape 2 : On considère la matrice diagonale

$$\mathbf{D}_1 = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{\frac{2}{3}} \end{pmatrix}.$$

On obtient

$$\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1\mathbf{D}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1\mathbf{D}_1 = \begin{pmatrix} \frac{5}{2} & \frac{1}{6}\sqrt{3} \\ \frac{1}{6}\sqrt{3} & \frac{1}{6} \end{pmatrix}.$$

Le cercle $(\mathcal{E}_1'')$ et l'ellipse $(\mathcal{E}_2'')$, représentés sur la figure F.5, correspondent aux représentations graphiques des formes quadratiques associées aux matrices $\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1\mathbf{D}_1$ et $\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1\mathbf{D}_1$.

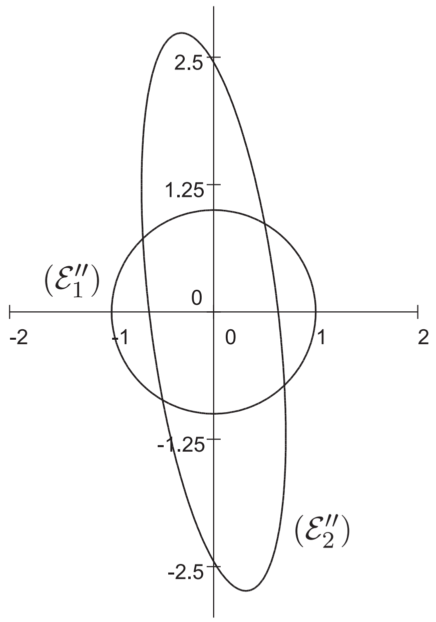


Figure F.5 – Ellipses $(\mathcal{E}_1'')$ et $(\mathcal{E}_2'')$ associées respectivement aux matrices $\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1\mathbf{D}_1$ et $\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1\mathbf{D}_1$.

Etape 3 : On considère la matrice orthogonale élémentaire

$$\begin{aligned}
\mathbf{\Omega}_2 &= \begin{pmatrix} \frac{\frac{7}{3}\sqrt{3}-\frac{2}{3}\sqrt{3}\sqrt{13}}{\sqrt{\left(\frac{7}{3}\sqrt{3}-\frac{2}{3}\sqrt{3}\sqrt{13}\right)^2+1}} & \frac{\frac{7}{3}\sqrt{3}+\frac{2}{3}\sqrt{3}\sqrt{13}}{\sqrt{\left(\frac{7}{3}\sqrt{3}+\frac{2}{3}\sqrt{3}\sqrt{13}\right)^2+1}} \\ \frac{1}{\sqrt{\left(\frac{7}{3}\sqrt{3}-\frac{2}{3}\sqrt{3}\sqrt{13}\right)^2+1}} & \frac{1}{\sqrt{\left(\frac{7}{3}\sqrt{3}+\frac{2}{3}\sqrt{3}\sqrt{13}\right)^2+1}} \end{pmatrix}. \\
&\approx \begin{pmatrix} -0.120\,984\,826\,7 & 0.992\,654\,356\,6 \\ 0.992\,654\,356\,6 & 0.120\,984\,826\,7 \end{pmatrix}
\end{aligned}$$

On obtient

$$\begin{aligned}
\mathbf{\Omega}_2^T\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\mathbf{\Omega}_2^T\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2 &\approx \begin{pmatrix} 0.131\,48 & 0 \\ 0 & 2.\,535\,18 \end{pmatrix}.
\end{aligned}$$

Ces deux matrices sont bien diagonales. Ainsi, la matrice $\mathbf{P} = \mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2$ diagonalise conjointement $\mathbf{A}_1$ et $\mathbf{A}_2$.

Le cercle $(\mathcal{E}_1''')$ et l'ellipse $(\mathcal{E}_2''')$, représentés sur la figure F.6 correspondent aux représentations graphiques des formes quadratiques associées aux matrices diagonales $\mathbf{\Omega}_2^T\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2$ et $\mathbf{\Omega}_2^T\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2$.

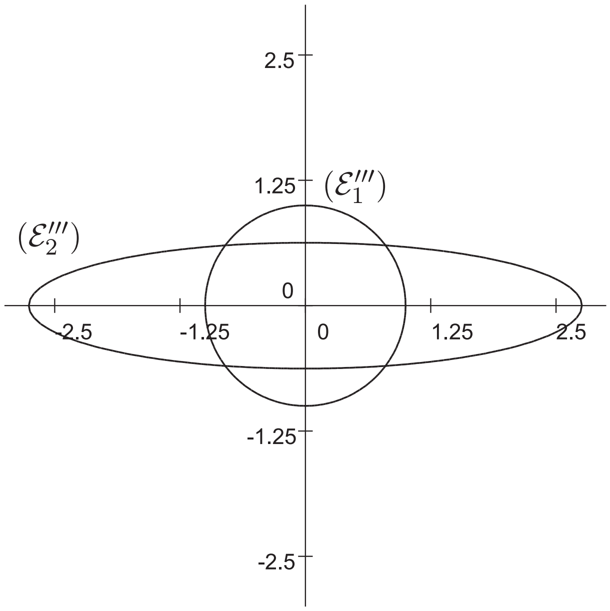


Figure F.6 – Ellipses $(\mathcal{E}_1''')$ et $(\mathcal{E}_2''')$ associées respectivement aux matrices $\mathbf{\Omega}_2^T\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_1\mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2$ et $\mathbf{\Omega}_2^T\mathbf{D}_1\mathbf{\Omega}_1^T\mathbf{A}_2\mathbf{\Omega}_1\mathbf{D}_1\mathbf{\Omega}_2$.

Pour obtenir les courbes $(\mathcal{E}_1''')$ et $(\mathcal{E}_2''')$ à partir des courbes $(\mathcal{E}_1)$ et $(\mathcal{E}_2)$, nous avons successivement effectué une rotation d'angle $\frac{\pi}{4}$ puis une homothétie et à nouveau une rotation d'angle $\alpha \approx 1.69$.

La diagonalisation conjointe revient à effectuer une première rotation afin de rendre la matrice $\mathbf{A}_1$ diagonale puis une homothétie pour que $\mathbf{A}_1$ soit représentée par un cercle et enfin une dernière rotation pour rendre $\mathbf{A}_2$ diagonale (à noter que cette dernière rotation n'affecte pas le caractère diagonale de $\mathbf{A}_1$ puisque que la matrice $\mathbf{A}_1$ est représentée par un cercle).

Bibliographie

[1] A. Akhenak, M. Cahldi, D. Maquin, and J. Ragot. Sliding mode multiple observer for fault detection and isolation. In *42nd IEEE Conf. on Decision and Control*, volume 1, pages 953–958, Hawaii, USA, december 2003.

[2] S. Amari, A. Cichocki, and H. H. Yang. A new learning algorithm for blind signal separation. *Advances in neural information processing systems*, 8, 1996.

[3] M. Babaie-Zadeh, A. Mansour, C. Jutten, and F. Marvasti. A geometric approach for separating speech signals. In *In proceedings of fifth International Conference on Independant Component Analysis and Blind Signal Separation*, pages 798–806, Granada, Spain, September 2004. Springer.

[4] PH. Barbe and M. Ledoux. *Probabilité*. Belin, 1998.

[5] T. Bell and T. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. In *Neural Computation*, volume 7, pages 1004–1034, 1995.

[6] A. Belouchrani, K. Abed Meraim, J.F. Cardoso, and E. Moulines. A blind source separation technique using second order statistics. *IEEE Transactions on Signal Processing*, 45:210–220, 1997.

[7] A. Blanc-Lapierre and B. Picinbono. *Fonctions Aléatoires*. Masson, Paris, 1981.

[8] P. Borne, G. Dauphin-Tanguy, J. P. Richard, F. Rotella, and I. Zambettakis. *Modélisation et Identification des processus*. Technip, 1992.

[9] H. Bousbiah-Salah, A. Belouchrani, and K. Abed-Meraim. Jacobi-like algorithm for blind signal separation of convolutive mixtures. *Electronics Letters*, 37(16):1049–1050, Auguste 2001.

[10] M. Briane and G. Pagès. *Théorie de l'intégration*. Vuibert, 2004.

[11] B. Buchberger. *An Algorithm for finding a basis for the residue class ring of a zero-dimensional polynomial ideal (in German)*. PhD thesis, University of Innsbruck, Autriche, 1965.

[12] B. Buchberger. Grobner bases: An algorithmic method in polynomial ideal theory. *Multidimensional Systems Theory*, pages 184–232, 1985.

[13] N. Budin and S. Abbound. Real time multichannel abdominal fetal ECG monitor using a digital signal coprocessor. In *Comp Biol. Med.*, volume 24, pages 451–462, 1994.

[14] D. Callaerts. *Signal separation methods based on singular value decomposition and their application to the real time extraction of the fetal ECG from cutaneaous recordings*. PhD thesis, Katolieke Universiteit of Leuven, 1989.

[15] J. F. Cardoso. Infomax and maximum likelihood for blind source separation. *IEEE Signal Processing Letters*, 4(4):112–114, April 1997.

[16] J. F. Cardoso. Independent component analysis of the cosmic microwave background. In *In proceedings of fourth International Conference on Independant Component Analysis and Blind Signal Separation*, pages 657–661, Nara, Japan, April 2003.

[17] J. F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, 44(12):3017–3030, December 1996.

[18] J.-F. Cardoso and A. Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1):161–164, January 1996.

[19] J.F. Cardoso. Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 2655–2658, Albuquerque, NM, USA, 3-6 April 1990.

[20] J.F. Cardoso. Estimating equations for source separation. In *IEEE International Conference on Acoustics Speech and Signal Processing*, volume 5, page 3449, 1997.

[21] J.F. Cardoso. Blind signal separation : statistical principles. *Procedings of the IEEE. Specials issues on on blind identifiacation and estimation Transactions on Signal Processing*, 86(10):2009–2025, October 1998.

[22] J.-F. Cavassilas, B. Xerri, and G. Chabriel. Séparation autodidacte de sources temporellement corrélées. In *Seizième colloque GRETSI*, pages 107–110, Grenoble, Septembre 1997.

[23] E. Chaumette, P. Comon, and D. Muller. Applications of ICA to airport surveillance. In *HOS 93*, pages 210–214, South Lake Tahoe, California, 7-9 June 1993.

[24] T. Cheviron. Différentiateurs pour la commande et l'observation de systèmes non linéaires : Application à la reconstruction de perturbation et à leur rejet. Rapport de DEA, IRCCYN-ENSIETA, 2005.

[25] P. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Dunod, Paris, 1998.

[26] A. Cichocki and S. Amari. *Adaptative Blind Signal and Image Processing : Learning algoritms and applications*. John Wiley, 2002.

[27] P. Comon. Independent component analysis. In *International Signal Proceedings Workshop on High-Order Statistics*, pages 29–38, 1991.

[28] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, April 1994.

[29] P. Comon. Blind identification and source separation in 2×3 under determined mixtures. *IEEE Transactions on Signal Processing*, 52(1):287–314, January 2004.

[30] P. Comon, C. Jutten, and J. Hérault. Blind separation of sources, part ii : Problems statement. *signal Processing*, (24):11–20, 1991.

[31] P. Comon and L. Rota. Blind separation of independent sources from convolutive mixtures. In *IEICE Transaction on Fundamentals of Elect. Com. Comput. Sciences*, volume E86-A, March 2003.

[32] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, 1991.

[33] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms : An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer-Verlag, 1992.

[34] G. Darmois. Analyse des liaisons de probabilit´e. In *In Proceedings Intern. Stat. conference 1947*, volume III A, page 231, Washington (D.C.), 1951.

[35] D. Declerq and A. Quinquis. *Le signal aléatoire*. Hermes, 1996.

[36] J. P. Demailly. *Analyse numérique et équations différentielles*. Presses Universitaires de Grenoble, 1996.

[37] D. Donoho. On minimun entropy deconvolution. pages 565–609, 1981.

[38] J. L. Ferrier and M. Rivoire. *Cours d'automatique*. Eyrolles, Paris, 1990.

[39] M. Fliess, C. Join, M. Mboup, and A. Sedoglavic. Estimation des derivées d'un signal multidimensionnel avec applications aux images et aux vidéos. In *Actes 20e coll. GRETSI*, september 2005.

[40] M. Fliess, J. Lévine, Ph. Martin, and P. Rouchon. Flatness and defect of non linear systems : Introductory theory and examples. In *Int J. Control*, number 61, pages 1327–1361, 1995.

[41] D. Foata and A. Fuchs. *Calcul des probabilités*. Dunod, 1998.

[42] A.J. Fossard and D. Normand-Cyrot. *Systèmes non linéaires*, volume 1.modélisation-estimation. Masson, 1993.

[43] M. Gaeta and J.L. Lacoume. Sources separation without a priori knowledge: The maximum likelihood solution. In *European Signal Processing Conference*, volume 1, pages 621–624, Barcelona, Spain, 18-21 September 1990.

[44] I. Guikhman and A. Skorokhod. *Introduction à la théorie des processus aléatoires*. Editions Mir, 1980.

[45] E. R. Hansen. A generalised interval arithmetic. In *Interval Mathematics, Lecture notes in Computer Science*, pages 7–18. Springer Verlag, 1975.

[46] E. R. Hansen. *Global Optimization Using Interval Analysis*. Marcel Dekker, New-York, 1992.

[47] H. Hironaka. Resolution of singularities of an algebric variety over a field of characteristic zero. In *Annals of Mathematics*, number 79, pages 109–326, 1964.

[48] J. Hérault and B. Ans. Réseaux de neurones à synapses modifiables : décodage de messages sensoriels composites par une apprentissage non supervisé et permanent. In *C.R. Académie des Sciences*, volume Série III, Paris, 1984.

[49] J. Hérault and C. Jutten. Space or time adaptive signal processing by neural network models. In *AIP Conference Proceedings*, volume 151, pages 206–211, August 1986.

[50] J. Hérault, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *Actes du Xème colloque GRETSI*, pages 1017–1022, Nice, Mai 1985.

[51] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley, 2001.

[52] A. Hyvärinen and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492, 1997.

[53] S. Ibrir. *Méthodes numériques pour l'observation et la commande des systèmes non linéaires.* PhD thesis, Université Paris Sud Orsay, 2000.

[54] J. Jaulin, M. Kieffer, D. Didrit, and E. Walter. *Applied Interval Analysis*. Springer, 2001.

[55] L. Jaulin. *Solution globale et garantie de problèmes ensemblistes ; applications à l'estimation non linéaire et à la commande robuste*. PhD thesis, Université Paris-Sud, Orsay, 1994.

[56] L. Jaulin and E. Walter. Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis. *Math. Comput. Simulation*, 35:123–127, 1993.

[57] L. Jaulin and E. Walter. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica*, 29(4):1053–1064, 1993.

[58] P.O. Amblard J.L. Lacoume and P. Comon. *Statistiques d'ordre supérieur pour le traitement du signal*. Masson edition, 1997.

[59] C. Jutten. *Calcul neuromimétique et traitement du signal, Analyse en composantes indépendantes*. PhD thesis, INPG-USMG, Grenoble, 1987.

[60] C. Jutten and J. Hérault. Independent component analysis versus PCA. In *Proceedings EUSIPCO,*, pages 643–646, Grenoble, France, 1988.

[61] C. Jutten and J. Karhunen. Advances in nonlinear blind source separation. In *In proceedings of fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, pages 245–256, Nara, Japon, April 2003.

[62] C. Jutten and A. Taleb. Source separation from dusk till dawn. In *Independent Component Analysis*, pages 15–26, Helsinki, Juin 2000.

[63] M.G. Kendall and A. Stuart. *The Advanced Theory of Statistics.* C. Griffin, London, 1963.

[64] M. Kieffer. *Estimation ensembliste par analyse par intervalles, applications à la localisation d'un véhicule*. PhD thesis, Université Paris-Sud, Orsay, 1999.

[65] S. Lagrange, L. Jaulin, V. Vigneron, and C. Jutten. Analytical solution of the blind source separation problem using derivatives. In *In proceedings of fifth International Conference on Independant Component Analysis and Blind Signal Separation*, pages 81–88, Granada, Spain, September 2004. Springer.

[66] F. Lamnabhi-Lagarrigue and P. Rouchon. *Commandes non linéaires*. Systèmes automatisés. Lavoisier, Janvier 2003.

[67] L. De Lathauwer, P. Comon, B. De Moor, and J. Vandewalle. ICA algorithms for 3 sources and 2 sensors. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 116–120, 14-16 June 1999.

[68] Ph. Martin, R.M. Murray, and P. Rouchon. Flat system, equivalence and trajectory generation. Technical report, ensmp, April 2003. *http:/www.cds.caltech.edu/∼murray/papers/*2003*d_mmr03-cds.html*.

[69] M. Martin-Deschamps. Méthodes algébriques effectives. Technical report, Université de Versailles - Saint Quentin en Yvelines.

[70] K. Matsuoka, M. Ohya, and M. Kawamoto. A neural net for blind separation of nonstationary signals. In *Neural Networks*, volume 8, pages 411–419, 1995.

[71] S. Méléard. *Probabilités, Concepts fondamentaux*. Techniques de l'ingénieur, 2001.

[72] R. Moore. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, 1966.

[73] J. Neveu. *Bases mathématiques du calcul des probabilités*. MASSON, 1964.

[74] D. T. Pham. Fast algorithm for estimating mutual information, entropies and score functions. In *In proceedings of fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japon, April 2003.

[75] D. T. Pham and J. F. Cardoso. Blind separation of instantaneous mixtures of non stationary sources. *IEEE Transaction on Signal Processing*, 9(49):1837–1848, September 2001.

[76] D. T. Pham, P. Garat, and C. Jutten. Separation of a mixture of independent sources through a maximum likelihood approach. In M. Moonen J. Vandewalle, R. Boite and A. Oosterlinck, editors, *In Signal Processing VI, Theories and Applications*, pages 771–774, Brussels, Belgium, August 1992.

[77] B. Picinbono. *Théorie des signaux et des systèmes*. Dunod, 1989.

[78] C. G. Puntonet, A. Mansour, and C. Jutten. Geometrical algorithm for blind separation of sources. In *In Actes du XVeme colloque GRETSI*, pages 273 – 276, Juan-Les-Pins, France, september 1995.

[79] C. G. Puntonet, A. Prieto, C. Jutten, M. Rodriguez-Alvarez, and J. Ortega. Separation of sources: A geometry-based procedure for reconstruction of n-valued signals. *IEEE Transaction on Signal Processing*, 46(3):267–284, 1995.

[80] J. F. Ritt. *Differential Algebra*. Dover Publications, New-York, 1966.

[81] P. Ruiz and J.L. Lacoume. Extraction of independent sources from correlated inputs : A solution based on cumulants. In *Workshop on Higher-Order Spectral Analysis*, pages 146–151, Vail, USA, 28-30 June 1989.

[82] H. Sira-Ramirez and M. Fliess. On the output feedback control of a synchronous generator. In *Proceedings of the 43 rd IEEE Conf. Decision Control*, Bahamas, 2004.

[83] S. Skelboe. Computation of rational interval functions. *BIT*, 14:87–95, 1974.

[84] E. Sorouchyari. Blind separation of sources, part iii : Stability analysis. *signal Processing*, (24):21–29, 1991.

[85] T. Sunaga. Theory of interval algebra and its applications to numerical analysis. *RAAG Memoirs*, 1958.

[86] A. Taleb and C. Jutten. On underdetermined source separation. In *In Proceedings ICASSP 99*, volume 3, pages 1445–1448, Arizona, Phoenix, March 1999.

[87] A. Taleb and C. Jutten. Source separation in post non linear mixture. *IEEE Transactions on signal Processing*, 47(10):2807–2820, October 1999.

[88] L. Tong, R. Liu, V. Soon, and Y.F. Huang. Indeterminacy and identifiability of blind identification. *IEEE Transactions on Circuits and Systems*, 21(5):499–509, May 1991.

[89] V. Vigneron, S. Lagrange, and C. Jutten. Calcul ensembliste pour la séparation de sources et caractérisation des solutions stables. In *Proceedings of the XIth European Symposium on Artificial Neural Networks*, pages 143–148, Bruges, 23-25 Avri 2003.

[90] E. Walter and L. Pronzato. *Identification of parametric models from experimental data*. Springer, 1997.

Contributions aux méthodes d'estimation en aveugle.

Sébastien LAGRANGE

Résumé

Le travail présenté dans cette thèse concerne d'une part, le problème de séparation aveugle de sources et d'autre part, celui d'estimation de paramètres en aveugle. Pour ces deux problèmes, une nouvelle classe de modèles de mélanges est étudiée, celle des mélanges inversibles, décrits par des équations différentielles.

Pour le problème de séparation aveugle de source, nous définissons la notion de séparabilité selon un modèle d'entrées fixé, puis proposons des méthodes de séparation basées sur des statistiques des signaux et de leurs dérivées.

Pour le problème d'estimation aveugle de paramètres, nous définissons la notion d'identifiabilité en aveugle selon un modèle d'entrées fixé et présentons une méthode d'estimation exploitant également les résultats sur les dérivées de signaux.

Les techniques d'analyse par intervalles sont exploitées afin d'obtenir des solutions garanties.

Cette étude, principalement théorique, est illustrée par de nombreux exemples simples.

Mots-clés : Séparation aveugle de sources, Mélange inversible, Estimation aveugle de paramètres, Analyse par intervalles, Dérivées de signaux aléatoires, Séparabilité, Identifiabilité en aveugle

Abstract

This dissertation presents the problem of blind source separation and the problem of blind parameter estimation. For these problems, a new class of mixture models is considered : invertible mixtures, described by differential equations. For the blind source separation problem, respectively blind parameter estimation problem, we define the concept of separability, resp. of blind identifiability, according to an input model. Then, we propose a separation, resp. an estimation, method based on new results concerning random signal derivative. Interval analysis methods are exploited in order to obtain guaranteed solutions. Although our study is mainly theoretical, many illustrative examples are proposed.

Keywords: Blind source separation, Invertible mixture, Blind parameter estimation, Interval analysis, Derivative of random signal, Separability, Blind identifiability