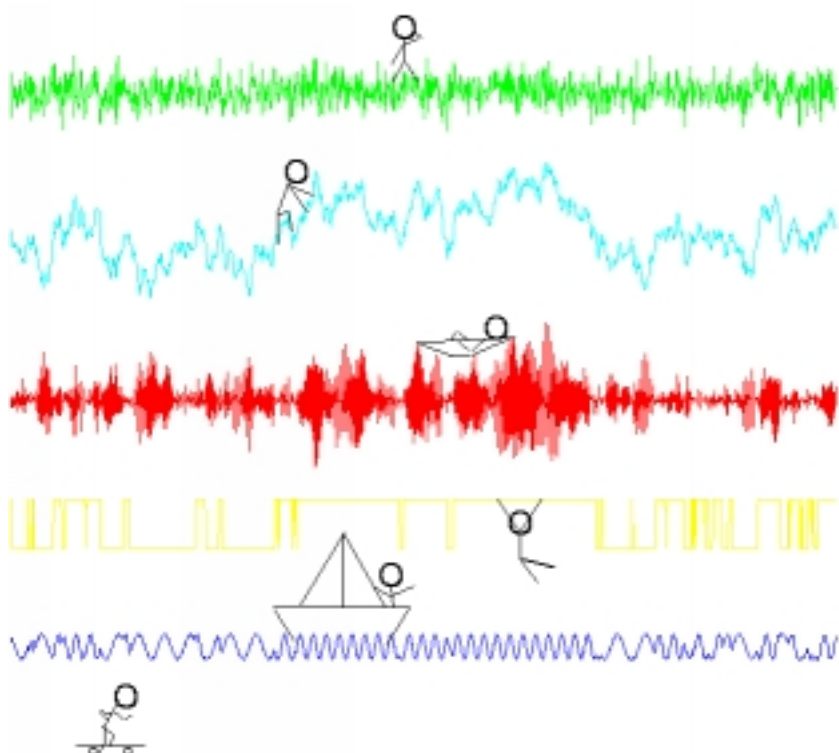


Processus alatoires Traitement des Signaux Mineure TSI

Thierry CHONAVEL

ENST Bretagne



thierry.chonavel@enst-bretagne.fr

avril 2007

Table des matières

1	Introduction	6
1.1	Avertissement	6
1.2	Motivation d'un cours de signal aléatoire	7
1.3	Quelques problèmes classiques de traitement du signal aléatoire	8
1.4	Ce qu'il faut retenir	9
2	Processus aléatoires	10
2.1	Généralités	10
2.1.1	Définition	10
2.1.2	Loi d'un processus aléatoire	11
2.2	Processus du second ordre	12
2.3	Opérations classiques dans $L^2(\Omega, \mathcal{A}, P)$	13
2.3.1	Convergence en moyenne quadratique	13
2.3.2	Continuité en moyenne quadratique	14
2.3.3	Dérivabilité en moyenne quadratique	14
2.3.4	Intégration	15
2.4	Stationnarité et ergodicité	17
2.4.1	Stationnarité	17
2.4.2	Ergodicité	18

	2
Exercices	19
3 Spectre de puissance	21
3.1 Spectres à densité : Densité Spectrale de Puissance (DSP)	21
3.2 Mesure spectrale	23
Exercices	25
4 Représentation spectrale	27
4.1 Transformée de Fourier d'un processus et mesure spectrale	27
4.2 Intégrale stochastique	28
4.3 Echantillonnage	29
Exercices	30
5 Filtrage.	32
5.1 Filtrage des processus stationnaires	32
5.2 Exemples	34
5.2.1 Filtrage passe bande et échantillonnage	34
5.2.2 Filtre dérivateur	35
Exercices	35
6 Processus particuliers importants	37
6.1 Processus gaussien	37
6.2 Processus de Poisson	38
6.3 Bruit blanc	39
6.4 Mouvement brownien	40
6.5 Processus cyclostationnaires	40
6.6 Processus circulaires	41

	3
6.7	Processus multivariés 42
6.8	Chaînes de Markov 43
6.8.1	Définition 43
6.8.2	Espace d'état fini 44
	Exercices 45
7	Transformations non linéaires des processus. 48
7.1	Détecteur quadratique et limiteur saturé. 48
7.1.1	Détecteur quadratique. 48
7.1.2	Limiteur saturé. 49
7.1.3	Théorème de Bussgang. 49
7.2	Modulation en amplitude des processus. 50
7.2.1	Modulation en phase et en quadrature. 51
7.2.2	Représentation analytique et modulation BLU. 51
7.2.3	Représentation de Rice. 52
7.2.4	Démodulation en présence de bruit. 53
	Exercices 54
8	Prédiction linéaire. 57
8.1	Espérance conditionnelle et prédiction linéaire. 57
8.2	Prédiction linéaire à horizon fini. 58
	Exercices 59
9	Filtrage adapté 61
9.1	Problème 61
9.2	Expression du filtre adapté 61
	Exercices 63

10 Processus à spectres rationnels	65
10.1 Equations aux différences et processus à spectres rationnels	65
10.2 Représentation d'état des modèles ARMA.	67
10.3 Identification des processus ARMA.	68
10.3.1 Identification de la partie AR.	68
10.3.2 Identification de la partie MA.	68
Exercices	68
11 Notions d'estimation statistique	70
11.1 Généralités	70
11.2 Estimation empirique	72
11.2.1 Séquence d'observations indépendantes.	72
11.2.2 Processus ergodique.	72
11.3 Notions d'estimation paramétrique	73
11.3.1 Borne de Cramer-Rao	73
11.3.2 Estimateur du maximum de vraisemblance	74
11.3.3 Estimateur linéaire sans biais de variance minimale	74
11.3.4 Méthodes de moindres carrés	74
11.4 Estimation bayésienne	75
11.4.1 Généralités	75
11.4.2 Coût bayésien et estimateurs bayésiens	76
Exercices	77
12 Estimation spectrale non paramétrique	78
12.1 Estimateurs empiriques de la moyenne et des covariances	78
12.1.1 Processus linéaires, et absolue sommabilité des covariances	78

	5
12.1.2 Moyenne empirique	79
12.1.3 Coefficients de covariance empiriques	80
12.2 Estimation empirique de la DSP : le périodogramme	81
12.2.1 Cas des processus linéaires	82
12.2.2 Cas des processus harmoniques	83
12.3 Lissage du périodogramme	83
12.3.1 Fonctions intégrales du périodogramme	83
12.3.2 Périodogramme lissé	84
12.3.3 Périodogramme moyenné	85
Exercices	85
13 Estimation spectrale paramétrique	87
13.1 Introduction	87
13.2 Estimation spectrale des modèles AR	88
13.3 Estimation spectrale des modèles ARMA	88
13.3.1 Estimation de la partie AR	89
13.3.2 Estimation de la partie MA	89
13.4 Synthèse de filtres rationnels	90
Exercices	91
A Annexe - Critère de Dirichlet	93
Classement bibliographique.	96
Bibliographie.	96
Notations et abréviations.	98

Chapitre 1

Introduction

1.1 Avertissement

Ce document expose des notions relatives aux signaux alatoires, encore appels processus alatoires, et au traitement des signal alatoire. Il se développe suivant trois directions

1. La notion de processus alatoire qui permet de décrire des expériences alatoires dont le résultat n'est pas simplement décrit au moyen d'une variable alatoire à valeur scalaire ou vectorielle mais par un signal alatoire. Dans ce contexte, le signal alatoire peut être vu à chaque instant comme une variable alatoire. Mais ici, on est également amené à s'intéresser à la suite des variables alatoires qui constituent les signaux et à leurs relations de dépendance.
2. En particulier, on s'intéressera fortement dans ce cours à la corrélation entre les variables qui constituent un processus à travers la notion de fonction de covariance. De même que la transformée de Fourier joue un rôle central pour l'étude des signaux déterministes, on verra qu'il en est de même pour certaines catégories de processus alatoires. La transformée de Fourier des fonctions de covariance et des processus eux même sera envisagée et exploitée tout au long de ce document. Ainsi, on trouvera principalement des notions relatives à la modélisation spectrale des processus stationnaires au second ordre, ainsi que leur filtrage.
3. Enfin, dans le but d'extraire de signaux alatoires des grandeurs d'intérêt pour l'utilisateur, on introduira quelques éléments de la théorie des statistiques dont le champ d'application dépasse largement le seul cadre du traitement du signal.

Les prérequis mathématiques nécessaires pour aborder ce cours sont essentiellement les bases du cours de probabilités statistiques, ainsi que quelques notions sur la théorie de la mesure.

On trouvera une justification plus approfondie des résultats présentés dans ces notes de cours dans [8, 9].

1.2 Motivation d'un cours de signal aléatoire

Le traitement du signal occupe une place dont les frontières avec des sciences fondamentales telles que les mathématiques ou la physique, et les sciences de l'ingénieur sont assez floues. En sciences de l'ingénieur, le traitement du signal intervient en particulier dans l'électronique, l'informatique, l'automatique, ou les mathématiques appliquées à des domaines aussi divers que la finance ou la biologie. Il en résulte plusieurs approches possibles pour la présentation d'un cours de signal aléatoire, suivant le point de vue de l'auteur. Cette différenciation des approches s'estompe cependant dans la mesure où le traitement du signal devient une discipline clairement identifiée.

Bien souvent, les résultats de la théorie du traitement des signaux déterministes n'est pas suffisante pour aborder certains problèmes. En effet, dans de nombreuses applications les signaux considérés sont aléatoires, en ce sens que la reproduction de la procédure expérimentale ne conduit pas à la même observation du fait de la présence de perturbations que l'opérateur ne peut pas maîtriser totalement. Ces perturbations peuvent être de nature variée : phénomènes de bruit additif ou de distorsion agissant sur les signaux transmis, erreurs liées aux appareils de mesure, ... L'observation comporte donc certains aléas. La théorie des probabilités et des processus aléatoires offre un cadre satisfaisant pour modéliser ces aléas. Dans la pratique du traitement du signal, on emploie d'ailleurs souvent indifféremment le terme de signal pour désigner un processus aléatoire ou une de ses réalisations.

Lorsque l'on cherche à extraire une information recherchée à partir des observations, il faut généralement effectuer un traitement, par exemple une moyenne sur plusieurs expériences successives, afin d'obtenir une estimation de la grandeur étudiée qui soit acceptable pour l'opérateur.

La modélisation probabiliste des signaux trouve son intérêt dans la possibilité que l'on a, pour effectuer cette estimation, d'employer les résultats de la théorie des statistiques. Cette théorie permet également d'envisager la pertinence de la méthode d'estimation retenue. Ainsi, la moyenne évoquée ci-dessus a pour but d'approcher, en vertu de la loi des grands nombres, l'espérance mathématique de la grandeur étudiée, cette grandeur étant décrite par un modèle aléatoire.

Dans certains cas, on ne dispose que d'une seule réalisation de l'expérience. Cela peut être le cas, par exemple, pour un signal enregistré lors d'une expérience de sismique acoustique. Dans un tel contexte, la propriété ergodique, lorsqu'elle est vérifiée, est très utile. Cette propriété permet en effet d'estimer de façon raisonnable des espérances mathématiques par des moyennes temporelles calculées sur la base de la connaissance d'une unique trajectoire du processus observé.

L'étude des signaux déterministes met en évidence l'importance de la transformée de Fourier. La transformée de Fourier permet d'évaluer la répartition de l'énergie ou de la puissance de signaux déterministes en fonction de la fréquence. De plus, de nombreuses transformations des signaux utilisées dans les dispositifs techniques et pour la modélisation de phénomènes naturels sont des opérations de filtrage. Ces filtres, qui sont des opérations de convolution des signaux, apparaissent comme de simples opérations de multiplication

dès lors qu'on travaille dans le domaine fréquentiel. On verra qu'une classe importante et très usitée de processus aléatoires, les processus stationnaires au sens large (on dit aussi au second ordre), admet une représentation dans le domaine fréquentiel. Cette représentation permet de traiter simplement le problème du filtrage. Elle permet également d'évaluer, à travers la notion de spectre d'un processus aléatoire, la façon dont en moyenne la puissance d'un processus se répartit en fonction de la fréquence.

1.3 Quelques problèmes classiques de traitement du signal aléatoire

Indiquons maintenant quelques problèmes très classiques qui s'inscrivent dans le contexte du traitement des signaux aléatoires.

1) On considère un oscillateur électronique que l'on excite au moyen d'une diode à bruit, et dont on cherche à étudier le comportement. Le signal observé en sortie de cet oscillateur ne ressemble que grossièrement à une sinusoïde et n'est pas tout à fait périodique. Dans le domaine spectral on peut envisager la façon dont, en moyenne, la puissance du signal fourni par l'oscillateur se répartit en fonction de la fréquence. Ceci permet en particulier d'évaluer la fréquence centrale de l'oscillateur et de déterminer des indices de performance tels que la largeur de bande de l'oscillateur. Cette répartition moyenne de la puissance en fonction de la fréquence, appelée spectre (chapitre 4), et la façon dont on peut l'estimer ne sont pas complètement intuitives lorsque, comme ici, on décrit le signal observé comme la réalisation d'un processus aléatoire. On verra également que les propriétés de convergence d'estimateurs spectraux classiques tels que le périodogramme peuvent s'avérer décevantes et qu'il faut prendre quelques précautions pour estimer correctement le spectre d'un processus aléatoire.

2) L'utilisation d'un modèle physique pour caractériser les interactions qui conduisent à la formation d'un signal et à celle de son spectre est souvent difficile. On peut alors souvent utiliser des modèles spectraux paramétriques tels que les modèles de spectres rationnels qui conduisent à des descriptions souvent plus simples à estimer mais aussi plus compactes que des modélisations fondées sur l'analyse des phénomènes physiques qui ont engendré le signal. En particulier, pour réaliser des systèmes de synthèse vocale, on cherche à générer des sons qui possèdent des propriétés spectrales analogues à celles de la parole, au moyen de filtres de fonctions de transfert rationnelles. En excitant de tels filtres au moyen de suites de variables aléatoires décorrélatées ou de séquences périodiques d'impulsions, suivant le type de sons que l'on cherche à générer, on peut produire un message oral compréhensible.

3) Lors d'une émission de radio, un signal transmis sur le canal de propagation hertzien y subit l'action d'un bruit additif avant de parvenir au récepteur. Si on considère que le spectre du signal utile n'occupe qu'une bande de fréquences limitée et que l'observation est perturbée par un bruit additif occupant une bande de fréquences beaucoup plus large, on peut chercher à placer un filtre au niveau du récepteur afin de limiter la puissance de bruit reçue. Si le filtre laisse passer une bande de fréquences trop large, la puissance de bruit reçue risque d'être importante. Au contraire, si le filtre ne retient qu'une bande de fréquences très étroite, le signal utile peut se trouver fortement distordu par l'action du

filtre. Il faut alors construire des filtres susceptibles de réaliser un compromis.

4) Pour coder ou transmettre numériquement un signal de parole, on procède classiquement à son échantillonnage puis à la quantification du signal échantillonné. Afin de réduire la quantité d'informations à transmettre, on commence généralement par effectuer avant la quantification un filtrage qui vise à le décorrélérer le signal (filtre d'erreur de prédiction). C'est le signal présent en sortie de ce filtre, appelé signal d'erreur de prédiction, qui est alors quantifié et échantillonné. Ce signal ayant une variance plus faible que celle du signal initial, il est alors possible, pour un niveau fixé de dégradation introduit par la quantification (rapport signal à bruit de quantification), de diminuer le nombre de bits d'information requis pour effectuer cette quantification.

1.4 Ce qu'il faut retenir

Rappelons ici que la connaissance des bases de probabilités est indispensable pour comprendre ce cours et que les notions rappelées en cours lors des premières séances sont supposées assimilées. Chaque semestre, en fonction des thèmes plus spécifiquement étudiés, vous trouverez en encadré dans le document les thèmes qu'il est plus particulièrement important d'avoir compris et retenus.

Chapitre 2

Processus aléatoires

Objectif On présente ici quelques notions relatives aux processus aléatoires en s'intéressant plus particulièrement aux propriétés du second ordre des processus. Cette démarche est motivée par le fait que les statistiques du second ordre sont assez simples à estimer à partir de l'observation d'un processus et qu'elles permettent de mettre en oeuvre de nombreuses méthodes de traitement des signaux.

2.1 Généralités

Sauf indications contraires on suppose que les variables aléatoires mises en jeu sont centrées, à valeurs scalaires et complexes. Soit (Ω, \mathcal{A}, P) un espace probabilisé. Rappelons que l'espace vectoriel $L^2(\Omega, \mathcal{A}, P)$ des variables aléatoires définies sur (Ω, \mathcal{A}, P) et de variance finie est un espace de Hilbert : c'est un espace vectoriel normé complet, muni du produit scalaire défini par $\langle X, Y \rangle = \mathbb{E}[XY^*]$. On notera $\|X\| = \langle X, X \rangle^{1/2}$ la norme correspondante. On trouvera en annexe une démonstration du caractère complet de $L^2(\Omega, \mathcal{A}, P)$.

Les variables aléatoires envisagées dans la suite seront généralement prises dans $L^2(\Omega, \mathcal{A}, P)$. Dès lors, de nombreux problèmes tels que ceux du calcul d'une espérance conditionnelle ou d'une régression linéaire pourront être vus comme des problèmes géométriques.

2.1.1 Définition

On se donne

- un espace probabilisé (Ω, \mathcal{A}, P) , où Ω est l'ensemble des résultats de l'expérience, \mathcal{A} une tribu sur Ω et P une mesure de probabilité définie sur \mathcal{A} ,
- un ensemble d'indices T (typiquement $\mathbb{N}, \mathbb{Z}, \mathbb{R} \dots$) appelé espace des temps,
- un espace mesurable (E, \mathcal{E}) ¹ appelé espace des états. On considérera généralement que

1. C'est à dire que \mathcal{E} est une tribu de E .

$(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ou $(E, \mathcal{E}) = (\mathbb{C}, \mathcal{B}(\mathbb{C}))$, $\mathcal{B}(E)$, désignant la tribu borlienne de E
 – une fonction $X = (X_t)_{t \in T}$ de Ω dans E , telle que pour tout $t \in T$, X_t soit une application mesurable de (Ω, \mathcal{A}) dans (E, \mathcal{E}) . X est appelée fonction aléatoire.

La donnée des grandeurs précédentes définit un processus aléatoire. Quand il n'y a pas d'ambiguïté sur les différents ensembles et sur la mesure de probabilité P , on désigne simplement un processus aléatoire par la fonction aléatoire $X = (X_t)_{t \in T}$.

Notons que X peut être vu comme une fonction de deux variables :

$$X : T \times \Omega \rightarrow E; \quad (t, \omega) \rightarrow X(t, \omega) = X_t(\omega). \quad (2.1)$$

2.1.2 Loi d'un processus aléatoire

Remarquons que pour un processus aléatoire $X = (X_t)_{t \in T}$, X peut être vu comme une variable aléatoire à valeur dans E^T . Notons que E^T peut être vu comme l'ensemble des fonctions de T dans E .

Dans ce qui suit Γ_t représente un élément quelconque de \mathcal{E} . On a vu en cours de probabilités que la notion même de variable aléatoire suppose l'existence d'une tribu sur l'ensemble d'arrivée. Si \mathcal{E} représente la tribu de E , la tribu correspondante sur E^T est définie comme suit :

Définition Etant donné l'espace mesurable (E, \mathcal{E}) , on définit la tribu $\mathcal{E}^{\otimes T}$ de E^T comme la tribu engendrée par les pavés $\prod_{t \in T} \Gamma_t$ de E^T , où $\Gamma_t = E$, sauf en un nombre fini de valeurs de t pour lesquelles $\Gamma_t \in \mathcal{E}$.

On peut en déduire le résultat important suivant :

Proposition 1 *La loi d'un processus X est entièrement définie par la loi de tous les vecteurs aléatoires de taille finie $((X_{t_1}, \dots, X_{t_n}))$ que l'on peut définir à partir de X .*

En effet, une mesure est entièrement définie par les valeurs qu'elle prend sur une famille génératrice de la tribu de l'espace mesurable sur lequel elle est définie. La loi P_X de X , définie par

$$\forall B \in \mathcal{E}^{\otimes T}, P_X(B) = P(X^{-1}(B)), \quad (2.2)$$

est donc entièrement caractérisée par la connaissance de l'ensemble des probabilités

$$\{P((X_{t_1}, \dots, X_{t_n}) \in \Gamma_1 \times \dots \times \Gamma_n); n \in \mathbb{N}, t_1, \dots, t_n \in T, \Gamma_1, \dots, \Gamma_n \in \mathcal{E}\} \quad (2.3)$$

où

$$((X_{t_1}, \dots, X_{t_n}) \in \Gamma_1 \times \dots \times \Gamma_n) = \{\omega \in \Omega; (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in \Gamma_1 \times \dots \times \Gamma_n\}. \quad (2.4)$$

2.2 Processus du second ordre

On supposera que les processus considérés sont à valeurs dans \mathbb{C} , indicés par un ensemble T ($T = \mathbb{N}, \mathbb{Z}, \mathbb{R}, \dots$), et du second ordre, c'est-à-dire vérifient

$$\forall t \in T, \|X_t\|^2 = \mathbb{E}[|X_t|^2] < \infty. \quad (2.5)$$

On définit alors la fonction de moyenne du processus X par

$$m_X(t) = \mathbb{E}[X_t] = \int_{\mathbb{R}} x P_{X_t}(dx), \quad (2.6)$$

où P_{X_t} représente la mesure de probabilité de X_t . La **fonction de covariance** de X est donnée par

$$\begin{aligned} R_X(t_1, t_2) &= \mathbb{E}[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))^*] \\ &= \mathbb{E}[X(t_1)X(t_2)^*] - m_X(t_1)m_X^*(t_2), \end{aligned} \quad (2.7)$$

et sa **fonction de corrélation** est définie par

$$\rho_X(t_1, t_2) = \frac{R_X(t_1, t_2)}{\sqrt{R_X(t_1, t_1)}\sqrt{R_X(t_2, t_2)}}. \quad (2.8)$$

Notons que l'inégalité (2.5) entraîne l'existence des fonctions de moyenne et de covariance du processus considéré.

Une fonction $r(t_i, t_j)$ est dite **de type positif** si

$$\forall M \in \mathbb{N}, \forall t_1, \dots, t_M \in T, \forall a_1, \dots, a_M \in \mathbb{R}, \sum_{i,j=1}^M a_i a_j^* r(t_i, t_j) \geq 0. \quad (2.9)$$

Il est clair que la fonction de covariance d'un processus est de type positif. En effet,

$$\sum_{i,j=1}^M a_i a_j^* R_X(t_i, t_j) = \left\| \sum_{i=1}^M a_i X_{t_i} \right\|^2 \geq 0. \quad (2.10)$$

Cette propriété a des conséquences importantes, comme on le verra au prochain chapitre.

Pour deux processus X et Y , on définit la **fonction d'intercovariance** de X et de Y par

$$\begin{aligned} R_{XY}(t_1, t_2) &= \mathbb{E}[(X(t_1) - m_X(t_1))(Y(t_2) - m_Y(t_2))^*] \\ &= \mathbb{E}[X(t_1)Y(t_2)^*] - m_X(t_1)m_Y(t_2)^*. \end{aligned} \quad (2.11)$$

Cette fonction n'est pas en général de type positif.

2.3 Opérations classiques dans $L^2(\Omega, \mathcal{A}, P)$

Rappelons que $L^2(\Omega, \mathcal{A}, P)$ représente l'ensemble des variables aléatoires X sur (Ω, \mathcal{A}, P) telles que $\mathbb{E}[|X|^2] < \infty$. Les variables X_t d'un processus du second ordre sont donc dans $L^2(\Omega, \mathcal{A}, P)$. Dans ce qui suit $(X_t)_{t \in \mathbb{R}}$ représente un processus du second ordre. Nous allons maintenant rappeler quelques notions et opérations de base dans $L^2(\Omega, \mathcal{A}, P)$, qui sont couramment utilisées en traitement du signal aléatoire. Plus précisément, en pratique, on est par exemple souvent amené à dériver ou intégrer un processus aléatoire par rapport à la variable t comme s'il s'agissait d'une fonction classique. On va voir comment on peut facilement définir ces opérations.

2.3.1 Convergence en moyenne quadratique

Définition On dit que X_t converge en moyenne quadratique vers la variable aléatoire X quand t tend vers t_0 , avec $t_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$, et on note $\lim_{t \rightarrow t_0} X_t \stackrel{m.q.}{=} X$, si

$$\lim_{t \rightarrow t_0} \|X_t - X\|^2 = 0. \quad (2.12)$$

Le théorème suivant permet de vérifier simplement la convergence en moyenne quadratique.

Théorème 1 (Loeve). *La limite en moyenne quadratique au point t_0 existe si et seulement si $\mathbb{E}[X_t X_{t'}^*]$ admet une limite finie lorsque t et t' tendent vers t_0 , indépendamment l'un de l'autre.*

Démonstration Supposons que $\lim_{t \rightarrow t_0} X_t \stackrel{m.q.}{=} X$.

$$\begin{aligned} |\mathbb{E}[X_t X_{t'}^* - X X^*]| &= |\mathbb{E}[X_t (X_{t'} - X)^* + (X_t - X) X^*]| \\ &\leq \|X_t\| \|X_{t'} - X\| + \|X\| \|X - X_t\| \\ &\leq (\|X_t - X\| + \|X\|) \|X_{t'} - X\| + \|X\| \|X - X_t\|, \end{aligned} \quad (2.13)$$

et les termes de droite tendent vers 0 lorsque t et t' tendent vers t_0 . Donc

$$\lim_{t, t' \rightarrow t_0} \mathbb{E}[X_t X_{t'}^*] = \|X\|^2. \quad (2.14)$$

Réciproquement, supposons que $\lim_{t, t' \rightarrow t_0} \mathbb{E}[X_t X_{t'}^*] = \alpha$, avec $\alpha < \infty$.

$$\|X_t - X_{t'}\|^2 = \|X_t\|^2 + \|X_{t'}\|^2 - \mathbb{E}[X_t X_{t'}^*] - \mathbb{E}[X_{t'} X_t^*]. \quad (2.15)$$

Comme chacun des termes de droite tend vers α quand $t, t' \rightarrow t_0$, $\lim_{t, t' \rightarrow t_0} \|X_t - X_{t'}\|^2 = 0$. Donc, pour toute suite $(t_n)_{n \in \mathbb{N}}$ qui tend vers t_0 , $(X_{t_n})_{n \in \mathbb{N}}$ est une suite de Cauchy et converge donc dans $L^2(\Omega, \mathcal{A}, P)$ (puisque c'est un espace complet). Soit X la limite d'une telle suite. Il est aisé de vérifier que cette limite est indépendante de la suite choisie et que $\lim_{t \rightarrow t_0} X_t \stackrel{m.q.}{=} X$, ce qui établit la réciproque du théorème. \square

2. Remarquons que la convergence de $\mathbb{E}[X_t X_{t'}]$ vers $\|X\|^2$ dès lors que $\lim_{t \rightarrow t_0} X_t \stackrel{m.q.}{=} X$ est en fait une conséquence immédiate de la continuité du produit scalaire.

2.3.2 Continuité en moyenne quadratique

Définition On dit que le processus X est continu en moyenne quadratique au point t si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall t', (|t - t'| < \delta \Rightarrow \|X_t - X_{t'}\|^2 < \varepsilon). \quad (2.16)$$

Il apparait que la continuité en moyenne quadratique au point t signifie en fait que $\lim_{t' \rightarrow t} X_{t'} \stackrel{m.q.}{=} X_t$. On peut caractériser la continuité en moyenne quadratique grce au résultat suivant :

Proposition 2 *Pour un processus X centré³, $\lim_{t' \rightarrow t} X_{t'} \stackrel{m.q.}{=} X_t$ si et seulement si $R_X(t_1, t_2)$ est continue au point $(t_1, t_2) = (t, t)$.*

Démonstration

$$\begin{aligned} |R_X(t_1, t_2) - R_X(t, t)| &= \mathbb{E}[X_{t_1} X_{t_2}^* - X_t X_t^*] \\ &= \mathbb{E}[(X_{t_1} - X_t)(X_{t_2}^* - X_t^*)] - \mathbb{E}[X_t(X_t - X_{t_2})^*] \\ &\quad + \mathbb{E}[(X_{t_1} - X_t)X_t^*]. \end{aligned} \quad (2.17)$$

Donc, si X est continu en moyenne quadratique au point t , les trois termes de droite tendent vers 0 lorsque $(t_1, t_2) \rightarrow (t, t)$ (d'après l'inégalité de Cauchy-Schwarz), et $R_X(t_1, t_2)$ est continue au point $(t_1, t_2) = (t, t)$.

La réciproque est immédiate si on considère l'égalité

$$\|X_t - X_{t'}\|^2 = R_X(t, t) - R_X(t, t') - R_X(t', t) + R_X(t', t'). \quad (2.18)$$

□

On voit donc que la continuité en moyenne quadratique de X peut être caractérisée directement à partir de la connaissance de la fonction de covariance de X . On va voir qu'il en est de même pour la dérivabilité ou l'intégrabilité du processus X .

2.3.3 Dérivabilité en moyenne quadratique

Définition On dit que le processus X est dérivable en moyenne quadratique en $t \in T$ si le taux d'accroissement $h^{-1}(X_{t+h} - X_t)$ converge en moyenne quadratique lorsque $h \rightarrow 0$. Bien entendu, la dérivabilité en moyenne quadratique entrane la continuité en moyenne quadratique. Notons X'_t le processus dérivé de X au point t . On a le résultat suivant :

3. Si le processus n'est pas centré, il faut remplacer dans l'énoncé $R_X(t_1, t_2)$ par $\mathbb{E}[X_{t_1} X_{t_2}^*]$.

Proposition 3 *Le processus X est dérivable en moyenne quadratique au point t si et seulement si $\frac{\partial^2}{\partial t_1 \partial t_2} R_X(t_1, t_2)$ existe et est finie au point $(t_1, t_2) = (t, t)$. De plus,*

$$\begin{aligned} R_{X'}(t_1, t_2) &= \frac{\partial^2 R_X(t_1, t_2)}{\partial t_1 \partial t_2} \\ R_{X'X}(t_1, t_2) &= \frac{\partial R_X(t_1, t_2)}{\partial t_1} \\ R_{XX'}(t_1, t_2) &= \frac{\partial R_X(t_1, t_2)}{\partial t_2}. \end{aligned} \quad (2.19)$$

Démonstration On utilise le théorème de Loeve (théorème 1). La dérivabilité en moyenne quadratique de X est équivalente à la convergence de

$$\mathbb{E}\left[\left(\frac{X_{t+h} - X_t}{h}\right)\left(\frac{X_{t+h'} - X_t}{h'}\right)^*\right] \quad (2.20)$$

lorsque h et h' tendent vers 0.

$$\mathbb{E}\left[\left(\frac{X_{t+h} - X_t}{h}\right)\left(\frac{X_{t+h'} - X_t}{h'}\right)^*\right] = \frac{1}{h'} \left[\frac{R_X(t+h, t+h') - R_X(t, t+h')}{h} - \frac{R_X(t+h, t) - R_X(t, t)}{h} \right], \quad (2.21)$$

et la limite, si elle existe, est égale à la valeur de $\frac{\partial^2}{\partial t_1 \partial t_2} R_X(t_1, t_2)$ au point $(t_1, t_2) = (t, t)$. Le reste de la proposition se vérifie directement. \square

2.3.4 Intégration

On procède comme pour l'intégrale de Riemann, en considérant des suites de fonctions en escalier. Le passage à la limite sera pris en moyenne quadratique.

Construction de $Y_t = \int_I g(t, \tau) X_\tau d\tau$

Soit $\mathcal{P}_n = \{A_k^n; k = 1, N_n\}$ une suite de partitions d'un intervalle I . On note Δ_k^n la longueur de A_k^n , et on suppose que

$$\lim_{n \rightarrow \infty} (\max_{k=1, N_n} \Delta_k^n) = 0, \quad (2.22)$$

c'est à dire que la taille des intervalles A_k^n décroît vers 0. On définit l'intégrale en moyenne quadratique, notée $Y_t = \int_I g(t, \tau) X_\tau d\tau$ comme la limite en moyenne quadratique, si elle existe, de la suite de variables aléatoires

$$Y_t^n = \sum_{k=1, N_n} g(t, \tau_k^n) X_{\tau_k^n} \Delta_k^n, \quad (2.23)$$

où τ_k^n est un point quelconque de A_k^n , la limite devant être indépendante du choix de la partition. On notera l'analogie de cette construction avec celle de l'intégrale de Riemann. D'après le théorème de Loeve (théorème 1), il est clair que l'intégrale $Y_t = \int_I g(t, \tau) X_\tau d\tau$ est définie en moyenne quadratique si et seulement si $\mathbb{E}[Y_t^n Y_t^{m*}]$ admet une limite lorsque $(m, n) \rightarrow \infty$, c'est à dire si l'intégrale

$$\int_{I \times I} g(t, \tau_1) g(t, \tau_2)^* R_X(\tau_1, \tau_2) d\tau_1 d\tau_2 \quad (2.24)$$

est définie au sens de Riemann. De plus, si l'intégrale Y_t est définie sur tout intervalle de valeurs de t , on obtient sur cet intervalle un processus dont la fonction de covariance est donnée par

$$R_Y(t_1, t_2) = \int_{I \times I} g(t_1, \tau_1) g(t_2, \tau_2)^* R_X(\tau_1, \tau_2) d\tau_1 d\tau_2. \quad (2.25)$$

Intégration en moyenne quadratique avec borne variable

Considérons maintenant la variable aléatoire

$$Y_t = \int_a^t \phi(\tau) X_\tau d\tau, \quad (2.26)$$

où $\phi(t)$ est continue et X_t est continue en moyenne quadratique.

Proposition 4 *Si $\phi(t)$ est continue, et X_t est continue en moyenne quadratique, $Y_t = \int_a^t \phi(\tau) X_\tau d\tau$ est dérivable en moyenne quadratique, de dérivée égale à $\phi(t)X(t)$.*

Démonstration D'après le théorème de Fubini et l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \left\| \frac{Y_{t+h} - Y_t}{h} - \phi(t)X_t \right\|^2 &= \mathbb{E} \left[\frac{1}{h^2} \int_{\tau, \tau' \in [t, t+h]} (\phi(\tau)X_\tau - \phi(t)X_t)(\phi(\tau')X_{\tau'} - \phi(t)X_t)^* d\tau d\tau' \right] \\ &= \frac{1}{h^2} \int_{\tau, \tau' \in [t, t+h]} \mathbb{E} [(\phi(\tau)X_\tau - \phi(t)X_t)(\phi(\tau')X_{\tau'} - \phi(t)X_t)^*] d\tau d\tau' \\ &\leq \frac{1}{h^2} \left(\int_{\tau \in [t, t+h]} \|\phi(\tau)X_\tau - \phi(t)X_t\| d\tau \right)^2. \end{aligned} \quad (2.27)$$

Notons de plus que

$$\|\phi(\tau)X_\tau - \phi(t)X_t\| \leq |\phi(\tau)| \|X_\tau - X_t\| + |\phi(\tau) - \phi(t)| \|X_t\|. \quad (2.28)$$

Donc, d'après la continuité de ϕ et la continuité en moyenne quadratique de X , $\phi(t)X_t$ est continue en moyenne quadratique. Par suite, le terme de droite de (2.28) tend vers 0 lorsque h tend vers 0, ce qui permet de conclure. \square

Intégration par parties

Soient $\phi(t)$ une fonction de classe \mathcal{C}^1 et X_t un processus continu, dérivable et de dérivée continue, en moyenne quadratique. Alors, on peut montrer que

$$\int_{t_0}^t \phi'(\tau) X_\tau d\tau = [\phi(t)X_t - \phi(t_0)X_{t_0}] - \int_{t_0}^t \phi(\tau) X'_\tau d\tau, \quad (2.29)$$

et

$$(\phi(t)X_t)' = \phi'(t)X_t + \phi(t)X'_t. \quad (2.30)$$

Remarque : propriétés des trajectoires Rappelons que les trajectoires d'un processus X sont les fonctions $t \rightarrow X_t(\omega)$, où $\omega \in \Omega$ est fixé. On peut également définir des propriétés

de continuité, de dérivabilité, ou d'intégration sur les trajectoires du processus, lorsque presque sûrement ces trajectoires sont continues, dérivables ou intégrables. Dans la suite, on ne considérera que les définitions en moyenne quadratique de ces notions, sauf dans la définition de l'ergodicité stricte présentée dans le paragraphe suivant.

2.4 Stationnarité et ergodicité

2.4.1 Stationnarité

Définitions Un processus est dit **stationnaire au sens strict** si $\forall n \in \mathbb{N}, \forall t_1, \dots, t_n \in T$, la loi du vecteur $\mathbf{X}_t = (X_{t+t_1}, \dots, X_{t+t_n})$ ne dépend pas de $t : \forall t \in \mathbb{R}, \forall \Gamma_1, \dots, \Gamma_n \in \mathcal{E}$,

$$P((X_{t_1}, \dots, X_{t_n}) \in \Gamma_1 \times \dots \times \Gamma_n) = P((X_{t+t_1}, \dots, X_{t+t_n}) \in \Gamma_1 \times \dots \times \Gamma_n). \quad (2.31)$$

Un processus X est dit **Stationnaire au Sens Large (SSL)**, ou encore stationnaire au second ordre, si les fonctions de moyenne et de covariance de X ne dépendent pas de t :

$$\begin{aligned} \forall t, \tau \in T, \quad m_X(t + \tau) &= m_X(t) \\ \forall t_1, t_2, \tau \in T, \quad R_X(t_1 + \tau, t_2 + \tau) &= R_X(t_1, t_2). \end{aligned} \quad (2.32)$$

Remarquons qu'un processus stationnaire au sens strict l'est également au sens large. En fait, et de façon générale, les propriétés statistiques définies au sens strict sont plus fortes que les propriétés au sens large. L'équivalence entre ces deux types de propriétés correspond au cas où le processus X considéré est gaussien.

Rappelons la distinction qui doit être faite entre les notions de processus stationnaire à l'ordre n qui désigne l'invariance vis à vis de t de la loi des n -uplets $\mathbf{X}_t = (X_{t+t_1}, \dots, X_{t+t_n})$, et celle de processus stationnaire au n -ème ordre, qui indique l'indépendance vis à vis de t des moments du processus d'ordres inférieurs ou égaux à n .

Pour un processus stationnaire au second ordre, $R_X(t_1, t_2) = R_X(t_1 - t_2, 0)$, et on notera alors simplement $R_X(t_1 - t_2)$ cette grandeur.

Notons que pour un processus stationnaire au second ordre les fonctions de covariance et de corrélation coïncident à un facteur près. En effet, dans ce cas, la fonction de covariance s'écrit simplement $\rho_X(t) = R_X^{-1}(0)R_X(t)$. Ceci explique que les deux notions soient souvent confondues dans les ouvrages où les processus considérés sont stationnaires au second ordre.

2.4.2 Ergodicité

L'ergodicité est une propriété importante, qui pour un processus stationnaire X permet d'estimer de façon satisfaisante des grandeurs de la forme $\mathbb{E}[\phi(X_{t_1}, \dots, X_{t_n})]$ à partir de l'observation d'une trajectoire unique du processus.

Définitions Un processus $X = (X_t)_{t \in \mathbb{R}}$ est strictement ergodique si pour toute grandeur aléatoire de la forme $Y_t = \phi(X_{t_1+t}, \dots, X_{t_n+t})$, la valeur de $\lim_{\tau \rightarrow \infty} (2\tau)^{-1} \int_{-\tau}^{\tau} Y_t dt$, où l'intégrale est définie par rapport aux trajectoires de X , est lorsqu'elle existe presque sûrement indépendante de la trajectoire de X considérée.

Cette définition suppose, bien entendu, que presque toutes les trajectoires de Y_t soient des fonctions mesurables de la variable t .

Pour un processus stationnaire continu strictement ergodique, il apparat alors que

$$\lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} \phi(X_{t_1+t}, \dots, X_{t_n+t}) dt \stackrel{\text{p.s.}}{=} \mathbb{E}[\phi(X_{t_1+t}, \dots, X_{t_n+t})], \quad (2.33)$$

c'est à dire que

$$\begin{aligned} \exists N, \quad P(N) = 0, \quad \forall \omega \in \Omega - N, \\ \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} \phi(X_{t_1+t}(\omega), \dots, X_{t_n+t}(\omega)) dt = \mathbb{E}[\phi(X_{t_1+t}, \dots, X_{t_n+t})]. \end{aligned} \quad (2.34)$$

En d'autres termes, l'ensemble des trajectoires du processus pour lesquelles la convergence de l'intégrale vers $\mathbb{E}[\phi(X_{t_1+t}, \dots, X_{t_n+t})]$ n'est pas assurée est de probabilité nulle.

Cette propriété, connue sous le nom de théorème de Birkhoff-Von Neumann, se traduit dans le cas d'un processus stationnaire discret par des relations du type

$$\lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{l=-N, N} \phi(X_{l_1+l}, \dots, X_{l_n+l}) \stackrel{\text{p.s.}}{=} \mathbb{E}[\phi(X_{l_1}, \dots, X_{l_n})]. \quad (2.35)$$

L'ergodicité au second ordre, plus faible que l'ergodicité au sens strict, est caractérisée par le fait que des expression du type

$$\frac{1}{2\tau} \int_{-\tau}^{\tau} X_{t+t_1} dt, \text{ et } \frac{1}{2\tau} \int_{-\tau}^{\tau} X_{t+t_1} X_{t+t_2}^* dt \quad (2.36)$$

dans le cas de processus indicés par \mathbb{R} , ou

$$\frac{1}{2N+1} \sum_{-N}^N X_{n+n_1}, \text{ et } \frac{1}{2N+1} \sum_{-N}^N X_{n+n_1} X_{n+n_2} \quad (2.37)$$

dans le cas de processus indicés par \mathbb{Z} convergent dans $L^2(\Omega, \mathcal{A}, P)$ vers des valeurs constantes. Notons ici la condition de convergence en moyenne quadratique, plus faible que la condition de convergence presque sûre imposée pour l'ergodicité stricte.

Ainsi, un processus SSL valeurs réelles $X = (X_n)_{n \in \mathbb{Z}}$ est ergodique au second ordre si

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{-N}^N X_{n+n_1} &= \mathbb{E}[X_{n_1}], \\ \text{et } \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{-N}^N X_{n+n_1} X_{n+n_2} &= \mathbb{E}[X_{n_1} X_{n_2}]. \end{aligned} \quad (2.38)$$

Pour mettre en évidence l'ergodicité au second ordre d'un processus stationnaire jusqu'au quatrième ordre, on pourra utiliser le théorème de Slutsky, énoncé ici pour un processus à temps discret (on obtient un résultat équivalent dans le cas continu). On ne démontrera pas ce théorème ici.

Théorème 2 (Slutsky). *Un processus $X = (X_n)_{n \in \mathbb{Z}}$, stationnaire jusqu'au quatrième ordre, est ergodique au second ordre si et seulement si*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1, N} R_X(n) = 0, \quad (2.39)$$

$$\forall l \in \mathbb{N}, \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1, N} \mathbb{E}[X_{n+l}^c X_n^{c*} X_l^{c*} X_0^c] - R_X^2(l) = 0, \quad (2.40)$$

et $X_n^c = X_n - m_X$.

Exercices

2.1 Montrez que la somme et le produit de deux fonctions d'autocovariance définissent encore des fonctions d'autocovariance.

2.2 (Processus gaussien) Soit $X = (X_t)_{t \in \mathbb{R}}$ un ensemble de variables aléatoires réelles telles que $\forall n \in \mathbb{N}$ et $\forall t_1, \dots, t_n \in \mathbb{R}$, la loi de $\mathbf{X} = [X_{t_1}, \dots, X_{t_n}]^T$ possède une densité de la forme

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-N/2} |\Gamma_{\mathbf{X}}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - m_{\mathbf{X}})^T \Gamma_{\mathbf{X}}^{-1}(\mathbf{x} - m_{\mathbf{X}})\right], \quad (2.41)$$

où on supposera ici que pour tout vecteur \mathbf{X} , $|\Gamma_{\mathbf{X}}| \neq 0$. X définit un processus gaussien.

a) Vérifiez que $m_{\mathbf{X}} = \mathbb{E}[\mathbf{X}]$ et $\Gamma_{\mathbf{X}} = \text{cov}(\mathbf{X})$.

b) On suppose de plus que X est stationnaire au second ordre. Montrez que si $\lim_{n \rightarrow \infty} R_X(n) = 0$, alors X est ergodique au sens strict. (Indication : rappelons, que pour le vecteur gaussien centré $[X_1, X_2, X_3, X_4]$,

$$\mathbb{E}[X_1 X_2 X_3 X_4] = \mathbb{E}[X_1 X_2] \mathbb{E}[X_3 X_4] + \mathbb{E}[X_1 X_3] \mathbb{E}[X_2 X_4] + \mathbb{E}[X_1 X_4] \mathbb{E}[X_2 X_3]. \quad (2.42)$$

2.3 On considère un processus SSL (Stationnaire au Sens Large) $X = (X_t)_{t \in \mathbb{R}}$. Montrez que $R_X(0) \geq |R_X(t)|$ pour tout t de \mathbb{R} , et que si $R_X(T) = R_X(0)$, alors X est périodique en moyenne quadratique, de période T , c'est à dire que $\|X_{t+T} - X_t\| = 0$.

2.4 Soient X et Y , deux processus SSL centrés tels que X_{t_1} et Y_{t_2} soient indépendants pour tout t_1 et t_2 de \mathbb{R} . Montrez que $Z = XY$ est un processus SSL.

2.5 Montrez que si une fonction de covariance $(R(t))_{t \in \mathbb{R}}$ est continue en $t = 0$, alors, elle est continue en tout point t de \mathbb{R} .

2.6 (Bruit blanc) Montrez que les séquences $(X_n)_{n \in \mathbb{Z}}$ de variables aléatoires indépendantes de distribution identique définissent des processus stationnaires au sens strict et que les séquences de variables aléatoires décorréelées de même moyenne et de même variance définissent des processus SSL.

2.7 (Processus harmoniques) Soit $X = (X_t)_{t \in \mathbb{R}}$, avec $X_t = \sum_{k=1, p} \xi_k e^{2i\pi t f_k}$. Les ξ_k sont des variables aléatoires et les f_k sont des fréquences distinctes.

a) Montrez que X est un processus SSL si et seulement si les ξ_k sont des variables aléatoires centrés et décorréelées.

b) Montrez que X n'est pas ergodique au sens large.

b) Calculez $(2T)^{-1} \int_{-T}^T X_t e^{2i\pi f t} dt$.

2.8 a) Montrez que les séquences indépendantes et de même loi $(X_n)_{n \in \mathbb{Z}}$ telles que $\mathbb{E}[|X_1|] < \infty$ définissent des processus ergodiques au sens strict.

b) Montrez que les processus gaussiens X tels que $\lim_{n \rightarrow \infty} R_X(n) = 0$ sont ergodiques au sens strict.

2.9 Soit $Y_t = X_t - X_{t-\tau}$, où $X = (X_t)_{t \in \mathbb{R}}$ est un processus SSL. Vérifiez que Y est SSL et calculez $R_Y(t)$ en fonction de $R_X(t)$.

2.10 Vérifiez que lorsque une matrice de covariance est inversible, son inverse est encore une matrice de covariance.

2.11 (matrices de Toeplitz) Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus SSL. On note $T_{X,N}$ la matrice de covariance de $\mathbf{X}_N = [X_0, \dots, X_N]^T$.

a) Vérifiez que $T_{X,N}$ est une matrice de Toeplitz, c'est à dire que $[T_{X,N}]_{a+1,b+1} = [T_{X,N}]_{a,b}$ pour tout a et b dans $\{1, \dots, N\}$.

b) On suppose que $T_{X,N}$ est une matrice définie positive. Montrez que $T_{X,N+1}$ est une matrice positive singulière si et seulement si $R_X(N+1)$ appartient à un cercle dont on calculera le centre et le rayon.

2.12 (intégration) Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL et $Y_t = T^{-1} \int_{t-T}^t X_u du$.

a) Calculer la fonction de covariance de Y en fonction de celle de X . Y est-il SSL ?

b) Calculer l'intercovariance de Y et de X . Les deux processus sont-ils stationnairement corrélés ?

Chapitre 3

Spectre de puissance des processus stationnaires au second ordre

Objectif On souhaite évaluer la répartition en moyenne de la puissance d'un processus stationnaire au second ordre en fonction de la fréquence. Cette connaissance permet de caractériser certaines propriétés particulières du processus telles que la présence de composantes harmoniques dans le signal ou permet d'envisager la mise en oeuvre de techniques classiques telles que l'échantillonnage sans perte.

3.1 Spectres à densité : Densité Spectrale de Puissance (DSP)

Dans ce qui suit, on suppose que X est un processus stationnaire au second ordre **centré**.

Si la fonction de covariance de X est intégrable, ($\int_{\mathbb{R}} |R_X(t)| dt < \infty$), on définit la densité spectrale de puissance $S_X(f)$ de X comme la transformée de Fourier de la fonction de covariance :

$$S_X(f) = \int_{\mathbb{R}} R_X(t) e^{-2i\pi ft} dt. (3.1)$$

Dans le cas discret, si $(R_X(n))_{n \in \mathbb{Z}}$ est absolument sommable cette définition devient

$$S_X(f) = \sum_{n \in \mathbb{Z}} R_X(n) e^{-2i\pi nf} (3.2)$$

D'après le théorème de Fubini, il est également clair qu'alors

$$R_X(n) = \int_{\mathcal{I}} e^{2i\pi nf} S_X(f) df, \quad (3.3)$$

avec $\mathcal{I} =] - 1/2, 1/2]$. En effet,

$$\begin{aligned} \int_{\mathcal{I}} e^{2i\pi n f} S_X(f) df &= \sum_{m \in \mathbb{Z}} R_X(m) \int_{\mathcal{I}} e^{2i\pi(n-m)f} df \\ &= R_X(n). \end{aligned} \quad (3.4)$$

Le lien de cette définition avec la répartition moyenne de la puissance du processus en fonction de la fréquence ne semble pas immédiat. Pour en donner une justification intuitive, on pourra par exemple remarquer que dans le cas discret si $\sum_{n=0, \infty} |R_X(n)| < \infty$, alors

$$S_X(f) = \lim_{n \rightarrow \infty} \left\| \frac{1}{\sqrt{n}} \sum_{k=1}^n X_k e^{-2i\pi k f} \right\|^2, \quad (3.5)$$

où, rappelons le, $\|Y\| = \mathbb{E}[|Y|^2]$. En effet, un calcul rapide permet de montrer que

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^n X_k e^{-2i\pi k f} \right\|^2 = \sum_{k=-n, n} \left(1 - \frac{|k|}{n}\right) R_X(k) e^{-2i\pi k f}. \quad (3.6)$$

et tend vers $S_X(f)$ lorsque n tend vers $+\infty$, d'après le théorème de convergence dominée de Lebesgue. $S_X(f)$ apparaît donc comme la variance de la transformée de Fourier du processus à la fréquence f , lorsque la taille du vecteur transformé tend vers $+\infty$.

La fonction $\hat{S}_{X,n}(f) = |n^{-1/2} \sum_{k=1}^n X_k e^{-2i\pi k f}|^2$ est appelée périodogramme d'ordre n de X et sera étudiée dans le cadre de l'estimation spectrale des processus stationnaires.

Il est clair que pour de nombreux processus, tels que $X_t = \xi e^{2i\pi t f_0}$, où $\xi \in L^2(\Omega, \mathcal{A}, P)$, la séquence des covariances n'est pas absolument sommable (ici $R_X(t) = \|\xi\|^2 e^{2i\pi t f_0}$), et les intégrales du type (3.1) ne sont pas définies. Dans le cours de traitement des signaux déterministes, on introduit généralement la théorie des distributions pour pouvoir définir la transformée de Fourier de fonctions qui, en tant que fonctions, ne possèdent pas de transformée de Fourier. Bien que l'on puisse faire la même chose ici, on se contentera d'introduire la notion de mesure spectrale qui constitue un formalisme suffisant pour la plupart des problèmes de traitement du signal aléatoire.

3.2 Mesure spectrale

Dans le cas général, pour un processus $X = (X_t)_{t \in \mathbb{R}}$ stationnaire au second ordre, on peut montrer que sa fonction de covariance R_X peut s'écrire comme la somme de deux termes $R_{X,1}(t)$ et $R_{X,2}(t)$; $R_{X,1}(t)$ admet une transformée de Fourier qui est une fonction positive, notée $S_X(f)$, et $R_{X,2}$ est de la forme $R_{X,2}(t) = \sum_k \rho_k e^{2i\pi t f_k}$, où les coefficients ρ_k sont positifs et le nombre de composantes de $R_{X,2}(t)$ est fini ou dénombrable. Il en résulte que

$$R_X(t) = R_{X,1}(t) + R_{X,2}(t) = \int_{\mathbb{R}} e^{-2i\pi f t} S_X(f) df + \int_{\mathbb{R}} e^{-2i\pi f t} \left(\sum \rho_k \delta_{f_k}(f) \right). \quad (3.7)$$

On pose alors $\mu_X(df) = S_X(f)df + \sum \rho_k \delta_{f_k}(f)$. μ_X est appelée mesure spectrale de X et vérifie

$$R_X(t) = \int_{\mathbb{R}} e^{-2i\pi f t} \mu_X(df). \quad (3.8)$$

De façon analogue, pour un processus à temps discret $X = (X_n)_{n \in \mathbb{Z}}$, on aura

$$R_X(n) = \int_{\mathcal{I}} e^{-2i\pi f n} \mu_X(df), \quad (3.9)$$

avec $\mathcal{I} =]-1/2, 1/2]$ et où la mesure μ_X est de la forme $\mu_X(df) = S_X(f)df + \sum \rho_k \delta_{f_k}(f)$, avec $S_X \geq 0$ et $\rho_k > 0$ ($-1/2 < f \leq 1/2$).

Notons que si la partie discrète du spectre est nulle, c'est à dire si $\mu_X(df) = S_X(f)df$ ou en d'autres termes si $R_{X,2}(t) = 0$, on se retrouve dans les conditions du paragraphe précédant et on directement la DSP $S_X(f)$ par transformée de Fourier de R_X : $S_X(f) = \int e^{2i\pi f t} R_X(t) dt$.

Remarque Les résultats que l'on vient d'indiquer font l'objet d'un théorème, appelé théorème de Bochner¹.

1. On peut définir la notion de mesure spectrale au moyen du résultat suivant :

Théorème 3 (Bochner). Pour une fonction $R(t)$ définie sur \mathbb{R} ,

$$((R(t))_{t \in \mathbb{R}} \text{ est de type positif}) \Leftrightarrow \exists! \mu, \text{ mesure positive, } \forall t \in \mathbb{R}, R(t) = \int_{\mathbb{R}} e^{2i\pi f t} \mu(df), \quad (3.10)$$

et pour une suite indicée par \mathbb{Z} ,

$$(R(n))_{n \in \mathbb{Z}} \text{ est de type positif} \Leftrightarrow \exists! \mu, \text{ mesure positive, } \forall n \in \mathbb{Z}, R(n) = \int_{\mathcal{I}} e^{2i\pi n f} \mu(df). \quad (3.11)$$

Pour un processus X stationnaire au second ordre, sa fonction de covariance $R_X(t)$ est de type positif et en prenant $R(t) = R_X(t)$ dans le théorème de Bochner, l'existence de la mesure spectrale, son unicité et sa positivité apparaissent clairement.

Exemples

a) La mesure spectrale d'un processus X discret défini par une suite Indépendantes et Identiquement Distribuées (IID) de variables aléatoires est constante et égale à la variance σ_X de X_n . En effet, il est clair que les coefficients de covariance de X sont de la forme $R_X(n) = \sigma_X^2 \delta_{0,n}$, et que ces coefficients coïncident avec les coefficients de Fourier de la mesure de \mathcal{I} de densité constante égale à σ_X^2 .

b) Pour un processus harmonique défini par $X_t = \sum_{k=1,p} \xi_k e^{2i\pi f_k t}$, il est clair (le vérifier à titre d'exercice) que la stationnarité au second ordre de X n'est vérifiée que si les variables aléatoires ξ_k sont décorréllées. Si de plus ces variables sont centrées,

$$\begin{aligned} R_X(t) &= \sum_{k=1,p} \|\xi_k\|^2 e^{2i\pi f_k t} \\ &= \int_{\mathbb{R}} e^{2i\pi f t} \left(\sum_{k=1,p} \|\xi_k\|^2 \delta_{f_k} \right), \end{aligned} \quad (3.12)$$

où δ_u représente la mesure de Dirac en u . Par suite, la mesure spectrale de X est $\mu_X(df) = \sum_{k=1,p} \|\xi_k\|^2 \delta_{f_k}$. Il s'agit donc d'une mesure discrète portée par les points f_1, \dots, f_p .

Remarque importante Dans tout ce qui précède, on a supposé que le processus X était centré. Lorsque X n'est pas centré, X peut se réécrire sous la forme $X_t = X_t^c + m_X$, où $X_t^c = X_t - m_X$. Comme le spectre de puissance du signal constant de valeur m_X vaut $m_X^2 \delta_0$, le spectre de X est donné par le spectre de X_t^c (c'est à dire la transformée de Fourier de R_X) plus $m_X^2 \delta_0$:

$$\mu_X(df) = \mu_{X^c}(df) + m_X^2 \delta_0. \quad (3.13)$$

Mesure interspectrale Les processus stationnaires centrés X et Y sont dits stationnairement corrélés si $\mathbb{E}[X_{t+\tau} Y_\tau]$ est indépendant de τ . Dans ce cas, on définit la mesure interspectrale μ_{XY} du couple (X, Y) par

$$\begin{aligned} R_{XY}(t) &= \mathbb{E}[X_{t+\tau} Y_\tau^*] \\ &= \int_{\mathbb{R}} e^{2i\pi f t} \mu_{XY}(df). \end{aligned} \quad (3.14)$$

La mesure μ_{XY} n'est pas positive en général. Notons que μ_{XY} est définie de façon unique

Ainsi, si X et Y sont stationnairement corrélés et possèdent des spectres à densité, le spectre de $aX + bY$, où $a, b \in \mathbb{C}$, est de la forme

$$S_{aX+bY}(f) = |a|^2 S_X(f) + |b|^2 S_Y(f) + ab^* S_{XY}(f) + a^* b S_{YX}(f). \quad (3.15)$$

Notons également que puisque $R_{XY}(t) = R_{YX}^*(-t)$, $S_{XY}(f) = S_{YX}^*(f)$.

Exercices

3.1 Soit $R(t) = e^{-\alpha|t|}$, $t \in \mathbb{R}$ et $\alpha > 0$.

- Vérifiez que R est une fonction de type positif.
- Etudiez le spectre correspondant lorsque α tend vers 0.

3.2 (Estimateur spectral du périodogramme) Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL centré de DSP $S_X(f)$. On suppose que X est de carré intégrable et on définit

$$\hat{S}_{X,T}(f) = \frac{1}{2T} \left| \int_{[-T,T]} X_t e^{-2i\pi ft} dt \right|^2. \quad (3.16)$$

a) Montrez que

$$\mathbb{E}[\hat{S}_{X,T}(f)] = \int_{[-2T,2T]} \left(1 - \frac{|t|}{2T}\right) R_X(t) e^{-2i\pi ft} dt. \quad (3.17)$$

b) Si $\int_{\mathbb{R}} |R_X(t)| dt < \infty$, montrez que $\hat{S}_{X,T}(f)$ converge en moyenne vers $S_X(f)$, c'est à dire que

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\left| \hat{S}_{X,T}(f) - S_X(f) \right| \right] = 0. \quad (3.18)$$

$\hat{S}_{X,T}(f)$ est l'estimateur du périodogramme de $S_X(f)$.

3.3 Soit $V = (V_n)_{n \in \mathbb{Z}}$ une séquence de variables aléatoires de même loi. On suppose que $\mathbb{E}[V_n] = 0$ et que $\|V_n\|^2 = \sigma^2 < \infty$.

- Notons $X_n = \sum_{l=1,q} b_l V_{n-l}$, pour tout n de \mathbb{Z} . Calculez la fonction de covariance et le spectre de $X = (X_n)_{n \in \mathbb{Z}}$.
- On définit $Y = (Y_n)_{n \in \mathbb{Z}}$ par l'équation aux différences $Y_n + aY_{n-1} = V_n$, où $|a| < 1$. Exprimez Y_n en fonction de $(V_{n-k})_{k \geq 0}$. Calculez la fonction de covariance et le spectre de Y .
- Reprenez la question b) mais en prenant pour a une variable aléatoire indépendante de $(V_n)_{n \in \mathbb{Z}}$ et telle que $P(|a| < 1) = 1$.

3.4 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL et $Y_t = X_{t-\tau}$, où τ est une constante.

- Exprimez R_{XY} et μ_{XY} en fonction de R_X et μ_X respectivement.
- Comment peut on retrouver τ à partir de R_{XY} ?

3.5 Soit $X_t = A \cos(2\pi f_0 t + \varphi)$, où A est une variable aléatoire centrée et $\varphi \sim \mathcal{U}_{[0,2\pi]}$. A et φ sont supposées indépendantes.

- Montrez que $X = (X_t)_{t \in \mathbb{R}}$ est SSL. Que ce passerait-il si φ était une grandeur déterministe connue ?

b) Calculez la fonction de covariance et le spectre de X .

c) On suppose maintenant que $Y_t = A \cos(2\pi f_0 t + \varphi + \alpha Z_t)$, où Z est un processus gaussien centré, SSL et décorréolé avec φ , de fonction de covariance $R_Z(t)$. calculez la fonction de covariance de Z .

3.6 Soit $X = (X_n)_{n \in \mathbb{Z}}$, où $X_n = \xi e^{2i\pi\nu n}$. ξ et ν sont des variables aléatoires supposées indépendantes. On suppose que ξ est centrée.

a) Montrez que X est SSL.

b) Notons $\phi_\nu(u) = \mathbb{E}[e^{i\nu u}]$ la fonction caractéristique de ν et p_ν la densité de probabilité de ν . Calculez $R_X(n)$ et $S_X(f)$ en fonction de ces paramètres.

Chapitre 4

Représentation spectrale des processus stationnaires au second ordre

Objectif Au chapitre précédent, on a vu comment représenter la puissance moyenne d'un processus SSL en fonction de la fréquence. Ici, c'est le processus lui-même que l'on va chercher à représenter dans le domaine fréquentielle. Cette représentation permettra d'étudier de façon simple le filtrage des processus stationnaires au chapitre suivant.

4.1 Transformée de Fourier d'un processus et mesure spectrale

On supposera, au besoin en soustrayant la moyenne, que les processus considérés sont centrés.

Rappelons tout d'abord qu'un processus stationnaire au second ordre est d'énergie infinie. Il en résulte que l'intégrale $\int_{[-T,T]} X_t e^{2i\pi f t} dt$ n'est pas définie en moyenne quadratique lorsque $T \rightarrow \infty$. On voit donc ici que la construction de la représentation d'un processus X indic par \mathbb{R} dans le domaine de Fourier n'est pas une opération immédiate.

Cependant, notons qu'il est aisé d'obtenir une représentation d'un processus dans le domaine de Fourier dans le cas où il est harmonique : en effet, si X est un processus harmonique, on pourra exprimer sa valeur courante sous la forme

$$X_t = \sum_k \xi_k e^{2i\pi f_k t} = \int_{\mathbb{R}} e^{2i\pi t f} \left(\sum_k \xi_k \delta_{f_k} \right). \quad (4.1)$$

$\sum_k \xi_k \delta_{f_k}$ apparaît donc comme la représentation du processus dans le domaine fréquentiel. On voit que cette représentation

1. est une mesure, du fait de la présence des mesures de Dirac δ_{f_k} ;
2. est aléatoire, du fait de la présence des variables aléatoires ξ_k .

$\sum_k \xi_k \delta_{f_k}$ est alors appelée la mesure aléatoire de X

Plus généralement, on montre que

tout processus SSL indicé par \mathbb{R} et continu en moyenne quadratique pourra se décomposer sous la forme

$$X_t = \int_{\mathbb{R}} e^{2i\pi t f} \hat{X}(df), \quad (4.2)$$

où \hat{X} sera appelé mesure alatoire de X .

De même, tout processus SSL indicé par \mathbb{Z} , $X = (X_n)_{n \in \mathbb{Z}}$, pourra s'écrire sous la forme

$$X_n = \int_{\mathcal{I}} e^{2i\pi n f} \hat{X}(df). \quad (4.3)$$

4.2 Intégrale stochastique

Considérons un processus SSL X qui admet une mesure alatoire \hat{X} . On cherche ici à donner un sens à des expressions de la forme $\int \phi(f) \hat{X}(df)$, appelées intégrales stochastiques. Ce type d'intégrale sera utile pour introduire le filtrage des processus SSL.

Soit une fonction $\phi(f)$ quelconque telle que $\int |\phi(f)|^2 \mu_X(df) < \infty$. $\phi(f)$ peut s'écrire comme la limite d'une suite de fonctions ϕ_n de la forme $\phi_n(f) = \sum_{k=1, N_n} a_{k,n} e^{2i\pi k f}$. En utilisant la linéarité de l'intégrale il vient que

$$\int \phi_n(f) \hat{X}(df) = \sum_{k=1, N_n} a_{k,n} X_k. \quad (4.4)$$

Par passage à la limite on définira $\int \phi(f) \hat{X}(df)$ par

$$\int \phi(f) \hat{X}(df) = \lim_{n \rightarrow \infty} \int \phi_n(f) \hat{X}(df). \quad (4.5)$$

Indiquons enfin sans démonstration une propriété importante de la représentation spectrale des processus

Théorème 4 Si $\hat{X}(f)$ représente la mesure alatoire du processus X , et $(\phi_k)_{k=1,2}$ sont des fonctions telles que $\int |\phi_k(f)|^2 \mu_X(df) < \infty$ alors

$$\mathbb{E} \left[\int \phi_1(f) \hat{X}(df) \left(\int \phi_2(f) \hat{X}(df) \right)^* \right] = \int_{\mathbb{R}} \phi_1(f) (\phi_2(f))^* \mu_X(df), \quad (4.6)$$

où les intégrales sont à prendre sur \mathbb{R} ou \mathcal{I} selon que les processus sont indiqués par \mathbb{R} ou \mathbb{Z} .

Pour voir tout l'intérêt de la formule (4.6), considérons le cas où $\phi_k(f) = e^{2i\pi t_k f}$ ($k = 1, 2$). On retrouve alors simplement le fait que la fonction de covariance est la transformée de Fourier du spectre :

$$\mathbb{E}[X_{t_1} X_{t_2}^*] = \mathbb{E} \left[\int e^{2i\pi t_1 f} \hat{X}(df) \left(\int e^{2i\pi t_2 f} \hat{X}(df) \right)^* \right] = \int_{\mathbb{R}} e^{2i\pi(t_1 - t_2)f} \mu_X(df). \quad (4.7)$$

Par ailleurs, en prenant $\phi_k(f) = \mathbb{1}_{[f, f+df[}$ ($k = 1, 2$), on obtient les relations différentielles

$$\mathbb{E} \left[\left| \int \mathbb{1}_{[f, f+df[}(f) \hat{X}(df) \right|^2 \right] = \mathbb{E}[|\hat{X}(df)|^2] = \mu_X(df). \quad (4.8)$$

Une autre propriété, qui découle de la relation (4.6), réside dans le fait que les variables aléatoires $\hat{X}(\Delta)$, où Δ est un intervalle, définies par

$$\hat{X}(\Delta) = \int \mathbb{1}_{\Delta}(f) \hat{X}(df) \quad (4.9)$$

sont orthogonales pour des intervalles disjoints, c'est à dire que

$$\Delta_1 \cap \Delta_2 = \emptyset \quad \Rightarrow \quad \mathbb{E}[\hat{X}(\Delta_1) \hat{X}(\Delta_2)^*] = \int_{\mathbb{R}} \mathbb{1}_{\Delta_1}(f) \mathbb{1}_{\Delta_2}(f) \mu_X(df) = 0. \quad (4.10)$$

4.3 Echantillonnage

Soit X un processus stationnaire au second ordre indicé par \mathbb{R} . Un processus discret $Y = (Y_n)_{n \in \mathbb{Z}}$ obtenu par échantillonnage régulier de X est encore stationnaire au second ordre. En effet, en posant $Y_n = X_{nT}$,

$$\begin{aligned} \mathbb{E}[Y_{n+k} Y_n^*] &= \mathbb{E}[X_{(n+k)T} X_{nT}^*] \\ &= R_X(kT). \end{aligned} \quad (4.11)$$

Supposons maintenant que X admet la représentation spectrale $X_t = \int_{\mathbb{R}} e^{2i\pi t f} \hat{X}(df)$. Notons \hat{Y} la mesure spectrale de Y , et cherchons son expression en fonction de celle de X . Comme $Y_n = X_{nT}$,

$$\begin{aligned} Y_n &= \int_{\mathbb{R}} e^{2i\pi n T f} \hat{X}(df) \\ &= \sum_{k \in \mathbb{Z}} \int_{](k-1/2)T^{-1}, (k+1/2)T^{-1}[} e^{2i\pi n T f} \hat{X}(df) \\ &= \sum_{k \in \mathbb{Z}} \int_{\mathcal{I}} e^{2i\pi n f} \hat{X}\left(\frac{df+k}{T}\right) \\ &= \int_{\mathcal{I}} e^{2i\pi n f} \left[\sum_{k \in \mathbb{Z}} \hat{X}\left(\frac{df+k}{T}\right) \right]. \end{aligned} \quad (4.12)$$

Ainsi, la mesure aléatoire de Y est

$$\hat{Y}(df) = \sum_{k \in \mathbb{Z}} \hat{X}\left(\frac{df+k}{T}\right). \quad (4.13)$$

De plus, d'après la relation (4.10),

$$\forall \Delta \in \mathcal{B}(\mathcal{I}), \mathbb{E}[\hat{X}(\frac{\Delta + k}{T})\hat{X}^*(\frac{\Delta + l}{T})] = 0, \text{ pour } k \neq l, \quad (4.14)$$

puisque $\Delta \in \mathcal{I}$ et il est clair que

$$\mu_Y(df) = \sum_{k \in \mathbb{Z}} d\mu_X(\frac{df + k}{T}). \quad (4.15)$$

Dès lors que le support de $\mu_X(df)$ est contenu dans $](-2T)^{-1}, (2T)^{-1}[$, on aura simplement $\mu_Y(df) = \mu_X(T^{-1}df)$, pour $f \in \mathcal{I}$. En particulier, si X possède un spectre à densité, il en est de même pour Y , et $S_Y(f) = T^{-1}S_X(T^{-1}f)$, pour $f \in \mathcal{I}$. On retrouve donc pour les processus stationnaires au second ordre les résultats du théorème d'échantillonnage obtenus dans le cadre du cours de traitement des signaux déterministes.

Exercices

4.1 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL continu en moyenne quadratique et $Y_t = X_{t-\tau}$, avec $\tau \in \mathbb{R}$. Exprimez la mesure spectrale et le spectre de Y en fonction de celle de X .

4.2 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL continu en moyenne quadratique et à valeurs réelles. On pose $\hat{X}(df) = \hat{X}_r(df) + i\hat{X}_i(df)$, où $\hat{X}_r(df)$ et $\hat{X}_i(df)$ représentent respectivement la partie réelle et la partie imaginaire de $\hat{X}(df)$. Montrez que $\hat{X}(-df) = \hat{X}^*(df)$, et vérifiez que

$$X_t = \int_0^\infty \cos(2\pi ft)\hat{X}_1(df) + \int_0^\infty \sin(2\pi ft)\hat{X}_2(df), \quad (4.16)$$

où \hat{X}_1 et \hat{X}_2 sont des mesures stochastiques. Exprimez \hat{X}_1 et \hat{X}_2 en fonction de $\hat{X}_r(df)$ et $\hat{X}_i(df)$.

4.3 Soit \hat{Z} une mesure aléatoire définie sur $E = \mathcal{I}$ ou sur $E = \mathbb{R}$. On définit le processus \mathcal{Z} par $\mathcal{Z}(f) = \hat{Z}([-\infty, f] \cap E)$. Montrer que \mathcal{Z} est à accroissements orthogonaux, c'est à dire que pour tous $f_1 \leq f_2 < f_3 \leq f_4$,

$$\mathbb{E}[(\mathcal{Z}(f_4) - \mathcal{Z}(f_3))(\mathcal{Z}(f_2) - \mathcal{Z}(f_1))^*] = 0. \quad (4.17)$$

4.4 (Matrices de covariance circulantes) Soit $X = (X_n)_{n \in \mathbb{Z}}$, avec $R_X(n)$ périodique, de période N .

a) Montrez que la matrice de covariance $T_{X,N-1}$ de taille $N \times N$ et de terme général $[T_{X,N-1}]_{ab} = R_X(a-b)$ est circulante, c'est à dire que les lignes successives de $T_{X,N-1}$ sont obtenues par permutations circulaires.

b) Montrez que les vecteurs propres de $T_{X,N-1}$ sont les vecteurs

$$d(f_k) = (1/\sqrt{N})[1, e^{2i\pi f_k}, \dots, e^{2i\pi(N-1)f_k}]^T, \quad (4.18)$$

où $f_k = kN^{-1}$ et $k = 0, \dots, N-1$.

- c) Calculez la mesure spectrale de X en fonction de X_0, \dots, X_{n-1} .
(Indication : notez que la mesure aléatoire de X est de la forme $\hat{X} = \sum_{k=0, n-1} \xi_k \delta_{f_k}$.)

Chapitre 5

Filtrage des processus stationnaires au second ordre

Objectif On généralise ici les résultats obtenus dans le cadre du filtrage des signaux déterministes au cas du filtrage des processus stationnaires au second ordre.

5.1 Filtrage des processus stationnaires

Considérons un processus centré X , stationnaire au second ordre et continu en moyenne quadratique. Il admet donc une représentation spectrale de la forme $X_t = \int_{\mathbb{R}} e^{2i\pi ft} \hat{X}(df)$. On définit dans $L^2(\Omega, \mathcal{A}, P)$ les opérations de filtrage de X à partir des fonctions $H(f)$ qui appartiennent à $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$ ¹, par les relations

$$Y_t = \int_{\mathbb{R}} e^{2i\pi ft} H(f) \hat{X}(df). \quad (5.1)$$

On a alors $\hat{Y}(df) = H(f) \hat{X}(df)$. Il est important de se limiter à des fonctions $H(f)$ qui appartiennent à $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$, afin que le processus $Y = (Y_t)_{t \in \mathbb{R}}$ soit également un processus du second ordre, car

$$\|Y_t\|^2 = \int_{\mathbb{R}} |H(f)|^2 \mu_X(df). \quad (5.2)$$

Dans le cas discret, $X_n = \int_{\mathcal{I}} e^{2i\pi nf} \hat{X}(df)$ et les opérations de filtrage de X sont définies à partir des fonctions $H(f) \in L^2(\mathcal{I}, \mathcal{B}(\mathcal{I}), \mu_X)$. On a alors

$$Y_n = \int_{\mathcal{I}} e^{2i\pi nf} H(f) \hat{X}(df). \quad (5.3)$$

1. c'est dire que $\int_{\mathbb{R}} |H(f)|^2 \mu_X(df) < \infty$

D'après la conservation du produit scalaire indiquée par la relation (4.6), les covariances s'expriment en fonction des mesures spectrales par les relations

$$\begin{aligned}
 R_Y(t) &= \mathbb{E}[Y_{t+s}Y_s^*] = \int_{\mathbb{R}} e^{2i\pi ft} |H(f)|^2 \mu_X(df), \\
 R_{YX}(t) &= \mathbb{E}[Y_{t+s}X_s^*] = \int_{\mathbb{R}} e^{2i\pi ft} H(f) \mu_X(df), \\
 R_{XY}(t) &= \mathbb{E}[X_{t+s}Y_s^*] = \int_{\mathbb{R}} e^{2i\pi ft} H^*(f) \mu_X(df).
 \end{aligned} \tag{5.4}$$

Par suite, les mesures spectrales sont liées par les relations

$$\begin{aligned}
 \mu_Y(sf) &= |H(f)|^2 \mu_X(df) \\
 \mu_{YX}(sf) &= H(f) \mu_X(df) \\
 \mu_{XY}(sf) &= H^*(f) \mu_X(df),
 \end{aligned} \tag{5.5}$$

ce qui dans le cas de spectres à densité donne

$$\begin{aligned}
 S_Y(f) &= |H(f)|^2 S_X(f) \\
 S_{YX}(f) &= H(f) S_X(f) \\
 S_{XY}(f) &= H^*(f) S_X(f).
 \end{aligned} \tag{5.6}$$

La transforme de Fourier inverse des relations précédentes conduit aux expressions correspondantes pour les fonctions de covariance, connues sous le nom de **formules des interférences** :

$$\begin{aligned}
 R_Y(t) &= h(t) * h^*(-t) * R_X(t), \\
 R_{YX}(t) &= h(t) * R_X(t), \\
 R_{XY}(t) &= h^*(-t) * R_X(t).
 \end{aligned} \tag{5.7}$$

Etablissons par exemple l'expression (5.7) pour $R_Y(t)$. En utilisant la conservation du produit scalaire (équation (4.6)) lors de l'intégration stochastique et le théorème de Fubini,

$$\begin{aligned}
R_Y(t) &= \mathbb{E}[Y_{u+t}Y_u^*] \\
&= \mathbb{E}\left[\left(\int_{\mathbb{R}} e^{2i\pi f(u+t)} H(f)\hat{X}(df)\right)\left(\int_{\mathbb{R}} e^{2i\pi fu} H(f)\hat{X}(df)\right)^*\right] \\
&= \int_{\mathbb{R}} e^{2i\pi ft} H(f)H^*(f)\mu_X(df) \\
&= \int_{\mathbb{R}^3} e^{2i\pi f(t-t_1+t_2)} h(t_1)h^*(t_2)\mu_X(df)dt_1dt_2 \tag{5.8} \\
&= \int_{\mathbb{R}^2} e^{2i\pi f(t-u)} \left[\int_{\mathbb{R}} h(t_1)h^*(t_1-u)dt_1\right]\mu_X(df)du \\
&= \int_{\mathbb{R}^2} e^{2i\pi f(t-u)} [h(u) * h^*(-u)]\mu_X(df)du \\
&= h(t) * h^*(-t) * R_X(t).
\end{aligned}$$

Si X est un processus stationnaire au second ordre indicé par \mathbb{Z} , on a des formules analogues telles que

$$R_{XY}(t) = \int_{\mathcal{I}} e^{2i\pi n f} H(f)^* \mu_X(df). \tag{5.9}$$

5.2 Exemples

5.2.1 Filtrage passe bande et échantillonnage

On considère un processus X , de représentation spectrale $X_t = \int_{\mathbb{R}} e^{2i\pi ft} \hat{X}(sf)$. Si $H(f) = \mathbb{1}_{[f_1, f_2]}(f)$, alors $Y_t = \int_{\mathbb{R}} e^{2i\pi ft} H(f)\hat{X}(sf)$ a pour mesure spectrale

$$\mu_Y(df) = \mathbb{1}_{[f_1, f_2]}(f)\mu_X(df). \tag{5.10}$$

Ce type de filtrage, appelé filtrage passe bande, conduit à un processus dont la répartition en moyenne de la puissance est la même que celle de X sur $[f_1, f_2]$, et est nulle ailleurs.

Considérons en particulier le problème de l'échantillonnage d'un processus X . Le processus échantillonné à la période T peut être défini par $X_{e,t} = \sum_{n \in \mathbb{Z}} X_{nT} \delta_{nT}$. Supposons maintenant que le spectre de X est porté par un sous intervalle de $[-(2T)^{-1}, (2T)^{-1}]$. On va voir qu'on peut retrouver X à partir de X_e par le filtrage de réponse fréquentielle $H(f) = \mathbb{1}_{[-(2T)^{-1}, (2T)^{-1}]}(f)$, dont on notera $h(t)$ la réponse impulsionnelle. Le processus Y

en sortie de ce filtre vérifie

$$\begin{aligned}
Y_t &= (h * X_e)_t \\
&= \sum_{n \in \mathbb{Z}} h(t - nT) X(nT) \\
&= \sum_{n \in \mathbb{Z}} h(t - nT) \int_{\mathbb{R}} e^{2i\pi n f T} \hat{X}(df) \\
&= \int_{\mathbb{R}} \left(\sum_{n \in \mathbb{Z}} h(t - nT) e^{2i\pi n f T} \right) \hat{X}(df).
\end{aligned} \tag{5.11}$$

Or, pour toute valeur de f contenue dans le support de \hat{X} ,

$$\begin{aligned}
\sum_{n \in \mathbb{Z}} h(t - nT) e^{2i\pi n f T} &= e^{2i\pi f t} [h(u) e^{-2i\pi f u} * \sum_{n \in \mathbb{Z}} \delta_{nT}]_{u=t} \\
&= e^{2i\pi f t} \mathcal{F}^{-1}[H(f' - f) T^{-1} \sum_{n \in \mathbb{Z}} \delta_{nT^{-1}}(f')] \\
&= e^{2i\pi f t} T^{-1}.
\end{aligned} \tag{5.12}$$

Donc, $Y_t = T^{-1} X_t$.

5.2.2 Filtre dérivateur

Pour un processus X stationnaire au second ordre, on vérifie aisément que la condition nécessaire et suffisante de dérivabilité en moyenne quadratique, donnée par l'existence de $\frac{\partial^2}{\partial t_1 \partial t_2} R_X(t_1, t_2)$ au point (t, t) s'exprime par $\int_{\mathbb{R}} f^2 \mu_X(df) < \infty$. Dans ce cas, la fonction

$$h \rightarrow \frac{e^{2i\pi f(t+h)} - e^{2i\pi f t}}{h} \tag{5.13}$$

converge dans $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X(df))$, quand h tend vers 0, vers $2i\pi f e^{2i\pi f t}$. Comme

$$\frac{X_{t+h} - X_t}{h} = \int_{\mathbb{R}} \frac{e^{2i\pi f(t+h)} - e^{2i\pi f t}}{h} \hat{X}(df), \tag{5.14}$$

et $\int_{\mathbb{R}} f^2 \mu_X(df) < \infty$, X_t est dérivable en moyenne quadratique, de dérivée

$$X'_t = \int_{\mathbb{R}} 2i\pi f e^{2i\pi f t} \hat{X}(df). \tag{5.15}$$

La dérivation en moyenne quadratique d'un processus correspond donc à une opération de filtrage, de réponse fréquentielle $H(f) = 2i\pi f$.

Exercices

5.1 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL et un filtre de réponse impulsionnelle $h(t) = \sum_{k=1, p} c_k e^{-\lambda_k t} \mathbb{I}_{\mathbb{R}_+}(t)$. Calculez la fonction de covariance et le spectre du processus Y en

sortie de ce filtre lorsque X en représente l'entrée en fonction de R_X et de μ_X respectivement.

5.2 Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus SSL de densité spectrale de puissance $S_X(f)$, et Y défini par $Y_n = K^{-1} \sum_{k=0, K-1} X_{n-k}$.

a) Calculez le spectre de Y .

b) Calculez $R_Y(n)$ quand X est un bruit blanc de variance σ^2 .

5.4 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL de fonction de covariance $R_X(t)$ et de densité spectrale de puissance $S_X(f)$, qui représente un signal transmis sur un canal de propagation. X se propage vers un récepteur en suivant plusieurs chemins qui introduisent des atténuations et retards. Finalement, on reçoit le signal $Y_t = \sum_{k=1, p} a_k X_{t-\tau_k}$ with $a_1, \dots, a_p, \tau_1, \dots, \tau_p \in \mathbb{R}$.

a) Calculez $R_Y(t)$ et $S_Y(f)$. Montrez que le canal de propagation se comporte comme un filtre dont on donnera la réponse fréquentielle et la réponse impulsionnelle.

b) On suppose maintenant que les coefficients $(a_k)_{k=1, p}$ sont des variables aléatoires centrées, indépendantes entre elles et avec X . Calculez $R_Y(t)$.

5.5 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL. On suppose que $\int_{\mathbb{R}} f^4 \mu_X(df) < \infty$ et que X vérifie l'équation différentielle suivante :

$$\frac{d^2 X}{dt^2} + k^2 X = 0 \quad (5.16)$$

Montrez que $X_k = \xi_1 e^{-ikt} + \xi_2 e^{ikt}$, où ξ_1 et ξ_2 sont des variables aléatoires décorréées.

5.6 Soit un filtre de fonction de transfert

$$g_N(z) = \frac{1}{2N+1} \sum_{n=-N, N} e^{2i\pi n f_0} z^{-n}. \quad (5.17)$$

Tracez $|g_N(e^{2i\pi f})|$ et indiquez l'intérêt d'un tel filtre lorsque N est grand.

5.7 Soit $h(z) = \frac{z^{-1} - \alpha}{1 - \alpha^* z^{-1}}$, avec $|\alpha| < 1$.

a) Montrez que si $X = (X_n)_{n \in \mathbb{Z}}$ est un processus du second ordre, il en est de même pour le processus Y défini par $Y_n = [h(z)]X_n$.

b) Montrez que des processus SSL distincts peuvent avoir le même spectre.

5.8 Etant donné un processus SSL $X = (X_n)_{n \in \mathbb{Z}}$, indiquez quelle condition il existe un processus Y SSL du second ordre solution de l'équation aux différences suivante

$$Y_n + \sum_{k=1, p} a_k Y_{n-k} = \sum_{l=0, q} b_l X_{n-l} \quad (5.18)$$

On montera que Y s'écrit comme une version filtrée de X , dont on précisera la fonction de transfert 10).

Chapitre 6

Processus particuliers importants

Objectif On présente ici brièvement quelques familles de processus qui interviennent souvent en traitement du signal aléatoire. Les processus markoviens indics par \mathbb{R} ne sont pas abordés dans ces notes de cours : un cours entier mériterait de leur être consacré.

6.1 Processus gaussien

Les processus réels gaussiens sont ceux pour lesquels les lois conjointes des variables vectorielles de la forme $\mathbf{X} = [X_{t_1}, \dots, X_{t_n}]^T$ sont gaussiennes. Rappelons qu'un vecteur gaussien \mathbf{X} est caractérisé par une fonction caractéristique de la forme

$$\varphi_{\mathbf{X}}(u) = \exp(i.u^T m_{\mathbf{X}} - (1/2)u^T \Gamma_{\mathbf{X}} u). (6.1)$$

On montre que $m_{\mathbf{X}}$ et $\Gamma_{\mathbf{X}}$ représentent alors respectivement le vecteur moyenne et la matrice de covariance de \mathbf{X} . Lorsque $|\Gamma_{\mathbf{X}}| \neq 0$, \mathbf{X} possède une loi à densité de la forme

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-N/2} |\Gamma_{\mathbf{X}}|^{-1/2} \exp[-\frac{1}{2}(\mathbf{x} - m_{\mathbf{X}})^T \Gamma_{\mathbf{X}}^{-1} (\mathbf{x} - m_{\mathbf{X}})]. (6.2)$$

Pour le cas complexe, la densité de probabilité de \mathbf{X} est obtenue en identifiant \mathbf{X} à un vecteur aléatoire de \mathbb{R}^{2n} dont les composantes sont les parties réelles et imaginaires des composantes de \mathbf{X} .

La donnée d'une fonction $m(t)$ et d'une fonction de type positif $R(t_1, t_2)$ caractérise entièrement la loi d'un processus gaussien pour lequel $\mathbb{E}[X_t] = m(t)$, et $cov[X_{t_1} X_{t_2}] = R(t_1, t_2)$.

On peut montrer que la limite la limite d'une suite de variables aléatoires gaussiennes $(X_n)_{n \in \mathbb{N}}$, convergente dans l'espace $L^2(\Omega, \mathcal{A}, P)$, est une variable aléatoire X gaussienne.

De plus, on montre que pour un couple gaussien de vecteurs (X, Y) , $\mathbb{E}[X|Y]$ coïncide avec la projection orthogonale de X au sens de la norme L^2 sur l'espace vectoriel engendré par les composantes de Y . Plus précisément, on a le résultat suivant :

Théorème 5 *Si (X, Y) est un couple gaussien de vecteurs, et $m_X, m_Y, \Gamma_X, \Gamma_Y, \Gamma_{XY}$ les moyennes, covariances et intercovariance de ces vecteurs,*

$$\mathbb{E}[X|Y] = m_X + \Gamma_{XY}\Gamma_Y^{-1}(Y - m_Y), \quad (6.3)$$

et la loi conditionnelle de X sachant $Y = y$ est une loi

$$\mathcal{N}(m_X + \Gamma_{XY}\Gamma_Y^{-1}(Y - m_Y), \Gamma_X - \Gamma_{XY}\Gamma_Y^{-1}\Gamma_{XY}^T). \quad (6.4)$$

Démonstration Notons

$$Z = X - m_X - \Gamma_{XY}\Gamma_Y^{-1}(Y - m_Y). \quad (6.5)$$

Il est facile de voir que

$$\begin{pmatrix} Z \\ Y \end{pmatrix} = \mathcal{N}\left(\begin{bmatrix} 0 \\ m_Y \end{bmatrix}, \begin{bmatrix} \Gamma_X - \Gamma_{XY}\Gamma_Y^{-1}\Gamma_{XY}^T & 0 \\ 0 & \Gamma_Y \end{bmatrix}\right). \quad (6.6)$$

Par suite,

$$\begin{aligned} \mathbb{E}[X|Y] &= \mathbb{E}[Z|Y] + m_X + \Gamma_{XY}\Gamma_Y^{-1}\mathbb{E}[(Y - m_Y)|Y] \\ &= \mathbb{E}[Z] + m_X + \Gamma_{XY}\Gamma_Y^{-1}(Y - m_Y) \\ &= m_X + \Gamma_{XY}\Gamma_Y^{-1}(Y - m_Y). \end{aligned} \quad (6.7)$$

Il est également clair que

$$\begin{aligned} \mathbb{E}[e^{iu^T X}|Y = y] &= \mathbb{E}[e^{iu^T Z}|Y = y] \exp[iu^T(m_X + \Gamma_{XY}\Gamma_Y^{-1}(y - m_Y))] \\ &= \exp[iu^T(m_X + \Gamma_{XY}\Gamma_Y^{-1}(y - m_Y)) - \frac{1}{2}u^T(\Gamma_X - \Gamma_{XY}\Gamma_Y^{-1}\Gamma_{XY}^T)u], \end{aligned} \quad (6.8)$$

ce qui termine la démonstration. \square

6.2 Processus de Poisson

Les processus de Poisson sont très utiles pour décrire des phénomènes aléatoires tels que les arrivées à des instants indépendants de personnes à un guichet, d'appels sur un réseau de télécommunications, ou le nombre de pannes (supposées indépendantes) d'une machine. L'étude des processus de Poisson est importante en vue de dimensionner correctement les machines ou les services chargés de gérer ces phénomènes.

Définition Un processus de Poisson $X = (X_t)_{t \in \mathbb{R}_+}$ d'intensité λ est un processus à accroissements indépendants, c'est à dire que si $t_1 < t_2 < t_3 < t_4$ alors $X_{t_2} - X_{t_1}$ et $X_{t_4} - X_{t_3}$ sont indépendants, et tel que si $t_1 < t_2$ la loi de $X_{t_2} - X_{t_1}$ est une loi de Poisson de paramètre $\lambda(t_2 - t_1)$, c'est à dire que $P(X_{t_2} - X_{t_1} = k) = (1/k!)\lambda(t_2 - t_1)^k e^{-\lambda(t_2 - t_1)}$.

Pour montrer que les processus de Poisson sont susceptibles de modéliser les phénomènes évoqués plus haut, notons T_n l'instant aléatoire auquel se produit le $n^{\text{ème}}$ événement et $\Delta_n = T_{n+1} - T_n$ l'intervalle de temps entre T_n et T_{n+1} . Le caractère sans mémoire de l'apparition des événements se traduit par les relations

$$P(\Delta_n > t + h | \Delta_n > t) = P(\Delta_n > h). \quad (6.9)$$

On en déduit facilement que Δ_n suit une loi exponentielle dont on notera λ l'intensité. En effet, d'après la relation (6.9) $P(\Delta_n > t + h) = P(\Delta_n > t)P(\Delta_n > h)$. En dérivant cette relation par rapport à h puis en fixant $h = 0$, il vient que la densité de probabilité de Δ_n vérifie l'équation différentielle

$$f_{\Delta_n}(t) = f_{\Delta_n}(0) \int_t^\infty f_{\Delta_n}(u) du, \quad (6.10)$$

dont la solution est

$$f_{\Delta_n}(t) = \lambda \exp(-\lambda t), \quad (6.11)$$

avec $\lambda = f_{\Delta_n}(0)$.

Par suite, comme $T_n = \Delta_1 + \dots + \Delta_n$ et que les variables aléatoires $(\Delta_k)_{k \in \mathbb{N}}$ forment une suite IID, la densité de la loi de T_n est donnée par l'expression

$$f_{T_n}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} \quad (6.12)$$

qui s'obtient facilement en passant par les fonctions caractéristiques. Cette loi est appelée loi d'Erlang d'ordre n et de paramètre λ .

Si maintenant on désigne par N_t le nombre d'événements qui se sont produits jusqu'à l'instant t , il est clair que N_t suit une loi de Poisson de paramètre λt :

$$\begin{aligned} P(N_t = n) &= P(T_n \leq t < T_{n+1}) \\ &= P(T_n \leq t) - P(T_{n+1} \leq t) \\ &= \frac{\lambda^n}{(n-1)!} \left[\int_0^t e^{-\lambda u} u^{n-1} du - \int_0^t \lambda e^{-\lambda u} \frac{u^n}{n} du \right] \\ P(N_t = n) &= e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \end{aligned} \quad (6.13)$$

De plus, il est facile de vérifier que N_t est un processus à accroissements indépendants. N est donc un processus de Poisson.

Des résultats complémentaires concernant les processus de Poisson pourront être trouvés par exemple dans [2], ou [12].

6.3 Bruit blanc

Définition Un bruit blanc est un processus stationnaire au second ordre qui possède une mesure spectrale à densité et dont la DSP est une fonction constante.

Dans le cas d'un processus à indice discret, on vérifie aisément que cela signifie que le processus est constitué d'une suite de variables aléatoires décorréées, de même moyenne et de même variance.

Le cas continu est plus délicat à traiter. En effet, la DSP d'un bruit blanc à indice continu est une fonction constante sur \mathbb{R} . Donc, l'intégrale de la densité spectrale n'est pas définie, ce qui signifie que pour un bruit blanc U on a $\|U_t\|^2 = +\infty$; U n'est donc pas un processus du second ordre! Nous ne traiterons pas ici le cas des bruits blancs à temps continu (voir par exemple [8] à ce sujet).

6.4 Mouvement brownien

Définition Un mouvement brownien (encore appelé processus de Wiener) issu de $t = 0$ est un processus B à accroissements orthogonaux tel que les variables aléatoires $B_t - B_{t'}$ ($t' < t$) ont des lois gaussiennes centrées et de variance $\sigma^2(t - t')$.

Cette définition caractérise un modèle simple de la trajectoire d'une particule dans un liquide, pour laquelle on suppose, outre la continuité de la trajectoire, que la particule subit de nombreux chocs et possède une inertie négligeable (le processus est à accroissements orthogonaux) et que les propriétés du milieu, supposé homogène, n'évoluent pas au cours du temps (la loi de $B_{t+\tau} - B_{t'+\tau}$ ne dépend pas de τ).

Comme $\|B_{t+h} - B_t\|^2 = \sigma^2 h$, il est clair que les trajectoires de B sont continues en moyenne quadratique, mais nulle part différentiables en moyenne quadratique.

6.5 Processus cyclostationnaires

Un processus X est dit cyclostationnaire de période T au sens strict si la loi d'un n -uplet X_{t_1}, \dots, X_{t_n} est la même que celle de $X_{t_1+kT}, \dots, X_{t_n+kT}$, $\forall n, k \in \mathbb{N}$ et $\forall t_1, \dots, t_n \in \mathbb{R}$. Il est dit cyclostationnaire au second ordre si on a simplement $\mathbb{E}[X_{t+kT}] = \mathbb{E}[X_t]$ et $R_X(t_1 + kT, t_2 + kT) = R_X(t_1, t_2)$, $\forall k$.

Exemple : Modulations linéaires en transmissions numériques.

Dans de nombreuses situations, les processus observés en transmissions numériques peuvent se représenter sous la forme

$$X_t = \sum_{k \in \mathbb{Z}} A_k \mathbb{I}_{[0, T[}(t - kT) + B_t, \quad (6.14)$$

où les A_k sont des variables aléatoires indépendantes qui constituent les symboles transmis et B_t un bruit blanc, indépendant des A_k , qui perturbe la transmission. Les A_k prennent leurs valeurs dans un ensemble fini $\{\alpha_1, \dots, \alpha_M\}$ avec des probabilités respectives p_1, \dots, p_M ($\sum_{i=1, M} p_i = 1$). Montrons que X_t est cyclostationnaire au second ordre,

de période T :

$$\begin{aligned}\mathbb{E}[X_t] &= \mathbb{E}[A_k] \\ &= \sum_{i=1, M} \alpha_i p_i, \quad \text{avec } kT \leq t < (k+1)T.\end{aligned}\tag{6.15}$$

Si $kT \leq t < (k+1)T$ et $lT \leq t + \tau < (l+1)T$,

$$\begin{aligned}\mathbb{E}[X_{t+\tau} X_t^*] &= \mathbb{E}[A_k] \mathbb{E}[A_l^*] \\ &= (\sum_{i=1, M} \alpha_i p_i)^2 \quad \text{si } k \neq l, \\ \mathbb{E}[X_{t+\tau} X_t^*] &= \mathbb{E}[|A_k|^2] + \|B_t\|^2 \delta_{0, \tau} \\ &= \sum_{i=1, M} \alpha_i^2 p_i + \|B_t\|^2 \delta_{0, \tau} \quad \text{si } k = l.\end{aligned}\tag{6.16}$$

Pour $0 < \alpha < T$, $\mathbb{E}[X_{t+\tau+\alpha} X_{t+\alpha}^*]$ peut donc différer de $\mathbb{E}[X_{t+\tau} X_t^*]$: prendre par exemple $0 < t < T - \alpha$ et $T - \alpha - t < \tau < T - t$. Mais pour $\alpha = T$ ces deux termes sont identiques, ce qui établit le résultat.

Pour pouvoir utiliser les résultats obtenus au cours de l'étude des processus stationnaires dans un contexte cyclostationnaire, on pourra chercher à se ramener au cas stationnaire en faisant appel au résultat suivant :

Théorème 6 *Si X est cyclostationnaire strict de période T , le processus $Y_t = X_{t-\phi}$, où ϕ est une variable aléatoire de loi uniforme sur $[0, T]$ est stationnaire au sens strict. Les fonctions de répartition F_X et F_Y , notées sous la forme*

$$F_X(x_1, \dots, x_n; t_1, \dots, t_n) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n),\tag{6.17}$$

sont alors liées par la relation

$$F_Y(y_1, \dots, y_n; t_1, \dots, t_n) = \frac{1}{T} \int_0^T F_X(y_1, \dots, y_n; t_1 - \tau, \dots, t_n - \tau) d\tau.\tag{6.18}$$

Si X est cyclostationnaire au second ordre, Y est alors stationnaire au second ordre, et

$$\begin{aligned}\mathbb{E}[Y_t] &= \frac{1}{T} \int_0^T \mathbb{E}[X_\tau] d\tau \\ \text{et } R_Y(t) &= \frac{1}{T} \int_0^T R_X(t + \tau, \tau) d\tau.\end{aligned}\tag{6.19}$$

6.6 Processus circulaires

Un processus X est circulaire si le processus Z défini par $Z_n = X_n e^{i\theta}$, $\forall \theta \in \mathbb{R}$, a la même loi que X . On dit que X est circulaire au p -ème ordre si seulement les moments de Z jusqu'à l'ordre p sont identiques aux moments correspondants de X . On montre alors que

Proposition 5 *Un processus aléatoire complexe X est circulaire au p -ème ordre si et seulement si tous ses moments*

$$\mathbb{E}\left[\prod_{\sum a_m=a} X_m^{a_m} \prod_{\sum b_n=b} (X_n^*)^{b_n}\right] \quad (6.20)$$

sont nuls dès que $a \neq b$ et $a + b \leq p$.

Indiquons également le résultat suivant, caractéristique des variables aléatoires circulaires au second ordre :

Proposition 6 *La variable aléatoire complexe $X = X_1 + iX_2$, où X_1 et X_2 sont des variables aléatoires à valeurs réelles, est circulaire au second ordre si et seulement si elle est centrée, $\|X_1\| = \|X_2\|$ et $\mathbb{E}[X_1X_2] = 0$.*

Démonstration D'après la proposition (5), X est circulaire au second ordre si et seulement si $\mathbb{E}[X] = 0$, et $\mathbb{E}[X^2] = 0$. La deuxième relation se traduit par

$$\mathbb{E}[X^2] = (\|X_1\|^2 - \|X_2\|^2) + 2i\mathbb{E}[X_1X_2] = 0, \quad (6.21)$$

soit $\|X_1\| = \|X_2\|$ et $\mathbb{E}[X_1X_2] = 0$. \square

Notons une conséquence directe de ce résultat : si $X = X_1 + iX_2$ est une variable aléatoire circulaire gaussienne de variance σ_X^2 , on vérifie alors facilement que sa densité de probabilité peut s'exprimer sous la forme d'une fonction de la variable scalaire complexe $x = x_1 + ix_2$:

$$f(x) = \frac{1}{\pi\sigma_X^2} \exp\left(\frac{-|x|^2}{\sigma_X^2}\right). \quad (6.22)$$

Pour ce qui concerne les vecteurs circulaires gaussiens $X = [X_1, \dots, X_n]^T = X_r + iX_i$, on vérifie facilement qu'ils sont caractérisés par les relations $\mathbb{E}[X] = 0$, $\mathbb{E}[X_rX_r^T] = \mathbb{E}[X_iX_i^T]$, et $\mathbb{E}[X_rX_i^T] = -\mathbb{E}[X_iX_r^T]$. En considérant la fonction caractéristique du vecteur $[X_r^T, X_i^T]^T$, il est alors possible d'établir que la densité de probabilité de X peut s'exprimer comme une fonction de la variable vectorielle complexe $x = x_r + ix_i$:

$$f_X(x) = \frac{1}{\pi^N |C_X|} \exp(-x^H C_X^{-1} x), \quad (6.23)$$

où n est la taille de X et $C_X = 2\mathbb{E}[X_rX_r^T]$.

6.7 Processus multivariés

Les processus multivariés sont des processus à valeurs dans un espace produit. On se place en général dans l'espace \mathbb{C}^k . Tous les résultats concernant les propriétés spectrales des processus scalaires stationnaires au second ordre présentés dans ce cours se généralisent

alors au cas vectoriel. Notons simplement qu'alors, la fonction de covariance de X est à valeurs matricielles, et la représentation spectrale est exprimée sous la forme

$$R_X(t) = \int_{\mathbb{R}} e^{2i\pi ft} \Omega(df), \quad (6.24)$$

où Ω est une mesure hermitienne positive ; c'est à dire que les termes de la matrice Ω sont des mesures et $\forall \Delta \in \mathcal{B}(\mathbb{R})$, $\Omega(\Delta)$ est une matrice hermitienne positive.

En traitement du signal l'emploi des processus multivariés peut être utile, en particulier pour décrire un signal observé simultanément sur un ensemble de capteurs. De plus, les équations différentielles linéaires d'ordre k , ou les équations aux différences dans le cas de processus discrets, qui mettent en jeu des processus scalaires pourront être étudiées comme des équations différentielles linéaires du premier ordre relativement à des processus multivariés de dimension k .

6.8 Chaînes de Markov

6.8.1 Définition

On supposera que les variables aléatoires considérées sont à valeurs dans E (typiquement, $E = \mathbb{R}$ ou $E = \mathbb{R}^m$).

Les chaînes de Markov jouent un rôle important dans de nombreux domaines. Ils permettent de décrire des phénomènes qui dépendent à chaque instant (ou en chaque point, selon le sens que l'on donne à la variable t) comme une fonction de la valeur prise à l'instant précédent (resp. aux points voisins) et de la valeur prise à l'instant considéré par un bruit blanc. Cela se traduit par la définition suivante :

Définition 1 Une chaîne de Markov à valeurs dans E est un processus aléatoire $X = (X_n)_{n \in \mathbb{N}}$ tel qu'il existe un bruit blanc $U = (U_n)_{n \in \mathbb{N}}$ et une suite de fonctions $(\Phi_n)_{n \in \mathbb{N}}$ qui vérifient pour tout $n \geq 0$

$$X_{n+1} = \Phi_n(X_n, U_n). \quad (6.25)$$

Lorsque Φ_n ne dépend pas de n , la chaîne de Markov est dite homogène.

Proposition 7 Si X est une chaîne de Markov,

$$P(X_{n+1} \in A | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_{n+1} \in A | X_n = x_n). \quad (6.26)$$

Cette propriété est souvent utilisée comme définition des chaînes de Markov. La définition envisagée ci dessus est mieux adaptée à la simulation informatique où les chaînes de Markov jouent un rôle très important.

Exemple (marche aléatoire) La chaîne de Markov définie par $X_{n+1} = X_n + U_n$ peut décrire la suite des gains dans un jeu de pile ou face si $P(U_n = 1) = P(U_n = -1) = 1/2$ ou la suite des valeurs obtenues par échantillonnage régulier d'un mouvement brownien lorsque $U_n \sim \mathcal{N}(0, \sigma^2)$.

Remarque Dans certaines situations, le processus X n'est pas directement markovien mais on a des relations du type $X_{n+1} = \Phi_n(X_n, X_{n-1}, \dots, X_{n-p}, U_n)$. Il apparaît alors, en posant $\mathbf{Y}_n = [X_n, \dots, X_{n-p}]^T$ que l'on a une relation du type

$$\mathbf{Y}_{n+1} = \begin{pmatrix} \Phi_n(\mathbf{Y}_n, U_n) \\ \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & \dots & 1 & 0 \end{bmatrix} \mathbf{Y}_n \end{pmatrix} = \Psi(\mathbf{Y}_n, U_n). \quad (6.27)$$

On se ramène ainsi à une description markovienne du phénomène.

6.8.2 Espace d'état fini

Supposons maintenant que E soit un ensemble fini, en notant par exemple (et sans perte de généralité) $E = \{1, 2, \dots, N\}$. Supposons également que la chaîne de Markov X soit homogène. On note

$$p_{ij} = P(X_{n+1} = j | X_n = i). \quad (6.28)$$

La matrice \mathbf{P} de taille N et de terme général $[\mathbf{P}]_{ij} = p_{ij}$ est appelée matrice de transition de X .

Proposition 8 Les probabilités de transition à n pas, définies par $p_{ij}^{(n)} = P(X_n = j | X_0 = i)$ sont fournies par les éléments de la matrice \mathbf{P}^n :

$$p_{ij}^{(n)} = [\mathbf{P}^n]_{ij}. \quad (6.29)$$

On déduit alors du fait que $\mathbf{P}^n = \mathbf{P}^l \mathbf{P}^{n-l}$ ($1 \leq l \leq n-1$) que

$$p_{ij}^{(n)} = \sum_{k=1, N} p_{ik}^{(l)} p_{kj}^{(n-l)}. \quad (6.30)$$

Cette formule est appelée **équation de Chapman-Kolmogorov**. Elle signifie que la probabilité de passer de l'état i à l'état j en n étapes revient à aller de l'état i à un état k quelconque en l étapes puis de l'état k à l'état j en $n-l$ étapes. Le calcul des $p_{ij}^{(n)}$ s'effectue en prenant en compte toutes les valeurs possibles de l'état intermédiaire k .

Notons maintenant $\mathbf{p}(n)$ le vecteur ligne qui décrit la loi de X_n . Il est clair que

$$[\mathbf{p}(n+1)]_j = \sum_{i=1, N} [\mathbf{p}(n)]_i p_{ij}. \quad (6.31)$$

On en déduit le résultat suivant :

Proposition 9 pour $m \geq 1$,

$$\mathbf{p}(n) = \mathbf{p}(n-1)\mathbf{P} = \mathbf{p}(0)\mathbf{P}^n. \quad (6.32)$$

Définition 2 Un vecteur v de dimension $N \times 1$ représente une mesure de probabilité stationnaire pour une chaîne de Markov de probabilité de transition \mathbf{P} si

$$v\mathbf{P} = v. \quad (6.33)$$

Notons qu'alors, $v\mathbf{P}^n = v$, c'est à dire que si la loi de X_0 est donnée par $\mathbf{p}(0) = v$, alors la loi de X_n est donnée par $\mathbf{p}(n) = v$.

Il est clair que les mesures stationnaires correspondent à des vecteurs propres à droite de \mathbf{P}^T et que la valeur propre correspondante est égale à 1.

De nombreux ouvrages sont consacrés aux processus markoviens et aux chaînes de Markov auxquels on pourra se référer pour plus de détails.

Exercices

6.1 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL centré et

$$\hat{R}_{X,T}(t) = (2T)^{-1} \int_{-T+|t|/2}^{T-|t|/2} X_{u+t/2} X_{u-t/2} du. \quad (6.34)$$

a) Calculez $\mathbb{E}[\hat{R}_{X,T}(t)]$.

b) Si X est gaussien, montrez que la variance de $\hat{R}_{X,T}(t)$ vaut

$$\frac{1}{2T} \int_{-T+|t|/2}^{T-|t|/2} \left(1 - \frac{|t| + |u|}{2T}\right) (R_X^2(u) + R_X(u+t)R_X(u-t)) du. \quad (6.35)$$

6.2 (Processus à accroissements orthogonaux) Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus centré à accroissements orthogonaux, c'est à dire tel que

$$\mathbb{E}[(X_{t_4} - X_{t_3})(X_{t_2} - X_{t_1})] = 0, \quad (6.36)$$

pour $t_1 \leq t_2 < t_3 \leq t_4$. On suppose aussi que $\|X_{t_2} - X_{t_1}\|^2 = \sigma^2|t_2 - t_1|$. Montrez que les variables aléatoires $X_t - X_s$ sont gaussiennes.

(Indication : considérez la fonction caractéristique de

$$X_t - X_s = \sum_{k=0, N-1} (X_{s+(k+1)(t-s)/N} - X_{s+k(t-s)/N}), \quad (6.37)$$

et faites tendre N vers $+\infty$.)

6.3 (Mouvement brownien) Soit $X_t = h^{-1}(W_{t+h} - W_t)$, où $W = (W_t)_{t \in \mathbb{R}}$ est un mouvement Brownien. Montrez que

$$R_X(t) = h^{-1}(1 - h^{-1}|t|)\mathbb{I}_{[-h,h]}(t), \quad (6.38)$$

et que

$$S_X(f) = \left(\frac{\sin(\pi h f)}{\pi h f} \right)^2. \quad (6.39)$$

6.4 (Spectre de certains codes numériques) La DSP d'un processus cyclostationnaire $X = (X_t)_{t \in \mathbb{R}}$ de période T est définie comme celle du processus stationnaire $Y = (X_{t+\phi})_{t \in \mathbb{R}}$, où ϕ est une variable aléatoire de loi uniforme sur $[0, T[$.

Soit $A = (A_n)_{n \in \mathbb{Z}}$ une séquence stationnaire de symboles réels et $h(t)$ un signal connu, avec $h(t) = 0$ pour $t \notin [0, T[$. $X_t = \sum_{k \in \mathbb{Z}} A_k h(t - kT)$ représente le signal utilisé pour transmettre le symbole A_n .

a) Montrez que le spectre de X est

$$\begin{aligned} \mu_X(df) &= [T^{-1} \sum_{k \in \mathbb{Z}} R_A(k) e^{-2\pi k f T}] |H(f)|^2 df \\ &+ T^{-2} |\mathbb{E}[A_n]|^2 \sum_{k=-\infty, \infty} |H(\frac{k}{T})|^2 \delta_{k/T}. \end{aligned} \quad (6.40)$$

(Remarque : comme $\mathbb{E}[A_n]$ peut être non nul, cette particularité doit être prise en compte dans le calcul du spectre.)

b) Calculez la DSP des codes suivants :

- code NRZ (Non-Return to Zero) : les A_k ($k \in \mathbb{Z}$) sont indépendants, $P(A_k = \pm 1) = 1/2$ et $h(t) = \mathbb{I}_{[0, T[}(t)$;
- code M-aire : les A_k sont indépendants,

$$P(A_k = \pm 1) = P(A_k = \pm 3) = \dots = P(A_k = \pm M - 1) = 1/M \quad (6.41)$$

et $h(t) = \mathbb{I}_{[0, T[}(t)$;

- code RZ (Return to Zero) : les A_k sont indépendants, $P(A_k = \pm 1) = 1/2$ et $h(t) = \mathbb{I}_{[0, \lambda T[}(t)$ ($\lambda \in]0, T[$) ;
- code biphasé : les A_k sont indépendants, $P(A_k = \pm 1) = 1/2$ et $h(t) = \mathbb{I}_{[0, T/2[}(t) - \mathbb{I}_{[T/2, T[}(t)$;
- code AMI (Alternate Marked Inversion) : les A_k ne sont pas indépendants. Une séquence supposée indépendante et équiprobable de 0 et de 1 doit être transmise : les 0 sont codés par $A_k = 0$ et les 1 alternativement par $A_k = 1$ et $A_k = -1$. $h(t) = \mathbb{I}_{[0, T/2[}(t) - \mathbb{I}_{[T/2, T[}(t)$.

c) Quel vous semble être l'intérêt particulier de chacun de ces codes ?

6.5 (Covariance cyclique) Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus cyclostationnaire de période T . Montrez que $\forall m, n \in \mathbb{Z}$ $R_X(m+n, m)$ peut encore s'écrire

$$R_X(m+n, m) = \sum_{k=0, T-1} R_k(n) e^{2i\pi k m / T}, \quad (6.42)$$

et exprimez $R_k(n)$ en fonction de R_X . Les T séquences $(R_k(n))_{n \in \mathbb{Z}}$, $k = 0, T-1$, sont appelées *covariances cycliques*. A quoi la séquence $(R_0(n))_{n \in \mathbb{Z}}$ correspond-elle ?

6.6 Pour $E = \{-2, -1, 0, 1, 2\}$, simulez la chaîne de Markov dont les états initiaux sont équiprobables et dont la matrice de transition est

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0.1 & 0.3 & 0.4 & 0.1 \\ 0 & 0.2 & 0.3 & 0 & 0.5 \\ 0.3 & 0.3 & 0 & 0.4 & 0 \\ 0 & 0.3 & 0.3 & 0.3 & 0.1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (6.43)$$

6.7 Soit X une chaîne de Markov de matrice de transition

$$\mathbf{P} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}. \quad (6.44)$$

a) Calculez \mathbf{P}^m .

b) Etudiez la convergence du vecteur de probabilité $\mathbf{p}(n)$ de X_n .

6.8 On considère une chaîne de Markov homogène à 5 états notés A, B, C, D et E. Les transitions suivantes ont des probabilités non nulles : $A \rightarrow A$, $A \rightarrow B$, $B \rightarrow A$, $B \rightarrow C$, $B \rightarrow D$, $B \rightarrow E$, $C \rightarrow B$, $D \rightarrow D$, $D \rightarrow E$, $E \rightarrow C$, $E \rightarrow D$, $E \rightarrow E$.

a) Dessinez le graphe de ces relations.

b) On suppose que depuis un état quelconque fixé toutes les transitions sont équiprobables. Ecrire la matrice de transition de la chaîne.

c) Rechercher la solution stationnaire de la chaîne.

Chapitre 7

Transformations non linéaires des processus.

Objectif Les transformations subies par les fonctions de covariance des processus dans les systèmes non linéaires ne s'expriment pas au moyen de formules aussi simples que celles obtenues dans le cadre du filtrage. En fait l'étude de ces relations doit généralement être faite au cas par cas. On présente ici quelques exemples simples de transformations non linéaires sans mémoire que l'on rencontre couramment dans les domaines de l'électronique et des télécommunications.

7.1 Détecteur quadratique et limiteur saturé.

Dans ce paragraphe, X désigne un processus indicé par \mathbb{R} , stationnaire au second ordre, centré et à valeurs réelles.

7.1.1 Détecteur quadratique.

Un détecteur quadratique est utilisé par exemple pour mesurer une puissance reçue. La sortie d'un détecteur quadratique de signal d'entrée X_t est de la forme $Y_t = X_t^2$. Si $f_{X_t}(x)$ désigne la densité de probabilité de X_t , celle de Y_t s'en déduit simplement et est donnée par

$$f_{Y_t}(y) = \frac{1}{2\sqrt{y}}(f_{X_t}(\sqrt{y}) + f_{X_t}(-\sqrt{y})). \quad (7.1)$$

Si de plus X est gaussien et centré, il apparaît que

$$\begin{aligned} R_Y(t) &= \mathbb{E}[X_{\tau+t}^2 X_\tau^2] - \mathbb{E}[X_{\tau+t}^2] \mathbb{E}[X_\tau^2] \\ &= 2R_X^2(t), \end{aligned} \quad (7.2)$$

car pour des variables aléatoires réelles gaussiennes X_1, X_2, X_3, X_4 ,

$$\mathbb{E}[X_1 X_2 X_3 X_4] = \mathbb{E}[X_1 X_2] \mathbb{E}[X_3 X_4] + \mathbb{E}[X_1 X_3] \mathbb{E}[X_2 X_4] + \mathbb{E}[X_1 X_4] \mathbb{E}[X_2 X_3]. \quad (7.3)$$

Remarquons maintenant que dans le cas d'un processus réel stationnaire gaussien centré, indicé par \mathbb{R} , le théorème de Slutsky s'exprime par

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{[-T, T]} R_X(t) dt = 0, \quad (7.4)$$

et $\forall l \in \mathbb{R}$,

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{[-T, T]} \mathbb{E}[X_t X_{t+l} X_l X_0] dt - R_X(l)^2 \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{[-T, T]} (R_X(t+l)R_X(t-l) + R_X(t)^2) dt \\ &= 0. \end{aligned} \quad (7.5)$$

D'après l'inégalité de Cauchy-Schwarz, cette dernière condition est vérifiée dès lors que $\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{[-T, T]} R_X^2(t) dt = 0$.

Sous ces conditions d'ergodicité, l'estimateur

$$U_T = \frac{1}{2T} \int_{[-T, T]} X_t^2 dt \quad (7.6)$$

de la puissance du processus converge en moyenne quadratique vers $R_X(0)$ quand T tend vers $+\infty$.

7.1.2 Limiteur saturé.

On considère le limiteur saturé dont la sortie est donnée par $Y_t = \text{sign}(X_t)$. C'est typiquement la forme de la sortie d'un amplificateur opérationnel sans contre-réaction.

Il est clair que $P(Y_t = 1) = P(X_t > 0) = 1 - P(Y_t = -1)$ et que si X est centré

$$R_Y(t) = P(X_{t+\tau} X_\tau > 0) - P(X_{t+\tau} X_\tau < 0). \quad (7.7)$$

On pourra vérifier le résultat suivant à titre d'exercice.

Proposition 10 *Si X est un processus gaussien, centré et stationnaire au second ordre, la fonction de covariance de la sortie du limiteur saturé est donnée par*

$$R_Y(t) = \frac{2}{\pi} \arcsin \left(\frac{R_X(t)}{R_X(0)} \right). \quad (7.8)$$

7.1.3 Théorème de Bussgang.

Le résultat suivant, valable pour les transformations sans mémoire, peut être utilisé pour simplifier le calcul pratique de la covariance d'un processus.

Théorème 7 (Bussgang). *Soit X un processus stationnaire, centré et gaussien, et g une transformation sans mémoire. Alors, en notant $Y_t = g(X_t)$, $R_X(t)$ est proportionnel à $R_{XY}(t)$.*

Démonstration

$$\begin{aligned} R_{XY}(t) &= \mathbb{E}[X_{t+\tau}g(X_\tau)] \\ &= \mathbb{E}[\mathbb{E}[X_{t+\tau}g(X_\tau)|X_\tau]] \\ &= \mathbb{E}[g(X_\tau)\mathbb{E}[X_{t+\tau}|X_\tau]]. \end{aligned} \tag{7.9}$$

Mais comme X est gaussien et centré,

$$\begin{aligned} \mathbb{E}[X_{t+\tau}|X_\tau] &= \mathbb{E}[X_{t+\tau}X_\tau](\mathbb{E}[X_\tau X_\tau])^{-1}X_\tau \\ &= R_X(t)R_X(0)^{-1}X_\tau. \end{aligned} \tag{7.10}$$

Donc $R_{XY}(t) = R_X(0)^{-1}R_{g(X),X}(0)R_X(t)$, ce qui établit le résultat. \square

Ainsi, en reprenant par exemple le cas du limiteur saturé, c'est à dire pour $g(x) = \text{sign}(x)$,

$$R_X(t) = \frac{R_X(0)}{\mathbb{E}[|X_t|]} \mathbb{E}[X_{t+\tau} \text{sign}(X_\tau)]. \tag{7.11}$$

Cette formule peut être utilisée en pratique pour estimer $R_X(t)$ avec un coût de calcul faible.

7.2 Modulation en amplitude des processus.

On considère le processus Y obtenu en modulant l'amplitude d'un signal sinusoïdal par un processus stationnaire au second ordre et centré $X : Y_t = X_t \times \cos(2\pi f_0 t)$. Alors,

$$\begin{aligned} Y_t &= \int_{\mathbb{R}} \frac{1}{2} [e^{2i\pi(f-f_0)t} + e^{2i\pi(f+f_0)t}] \hat{X}(df) \\ &= \int_{\mathbb{R}} e^{2i\pi ft} \left[\frac{\hat{X}(df + f_0) + \hat{X}(df - f_0)}{2} \right] \end{aligned} \tag{7.12}$$

et $\hat{Y}(df) = (1/2)[\hat{X}(df + f_0) + \hat{X}(df - f_0)]$. Supposons que le support de μ_X est dans un intervalle $[-f_c, f_c]$ tel que $f_0 > f_c$. Les supports de $\hat{X}(df + f_0)$ et de $\hat{X}(df - f_0)$ sont alors disjoints et

$$\mu_Y(df) = \frac{1}{4} [\mu_X(df + f_0) + \mu_X(df - f_0)]. \tag{7.13}$$

A un facteur près, le spectre de Y est donc obtenu par translation du spectre de X autour des fréquences $-f_0$ et $+f_0$.

7.2.1 Modulation en phase et en quadrature.

Si $X_{1,t}$ et $X_{2,t}$ sont des processus centrés, stationnaires du second ordre, stationnairement corrélés et à valeurs réelles, on peut définir la modulation en phase et en quadrature de ces processus par

$$Y_{1,t} = X_{1,t} \cos(2\pi f_0 t) - X_{2,t} \sin(2\pi f_0 t). \quad (7.14)$$

Y_1 est centré, et

$$\begin{aligned} \mathbb{E}[Y_{t+\tau} Y_\tau] &= \frac{1}{2}[R_{X_1}(t) + R_{X_2}(t)] \cos(2\pi f_0 t) + \frac{1}{2}[R_{X_1}(t) - R_{X_2}(t)] \cos(4\pi f_0 \tau + 2\pi f_0 t) \\ &\quad - \frac{1}{2}[R_{X_1 X_2}(t) + R_{X_2 X_1}(t)] \sin(4\pi f_0 \tau + 2\pi f_0 t) \\ &\quad + \frac{1}{2}[R_{X_1 X_2}(t) - R_{X_2 X_1}(t)] \sin(2\pi f_0 t). \end{aligned} \quad (7.15)$$

Donc Y_1 est stationnaire au second ordre si et seulement si $R_{X_1}(t) = R_{X_2}(t)$ et $R_{X_1 X_2}(t) = 0$, c'est à dire si X_1 et X_2 sont de même puissance et décorrelés. On supposera dans la suite que cette condition est satisfaite.

La modulation en phase et en quadrature permet de transmettre de façon simultanée deux signaux dans une même bande de fréquence. Les deux signaux ainsi modulés peuvent ensuite être séparés de façon simple. Supposons en effet que les processus X_1 et X_2 ont pour support spectral $[-f_c, f_c]$. Dans ce cas, le spectre de Y est porté par

$$[-f_c - f_0, f_c - f_0] \cup [-f_c + f_0, f_c + f_0]. \quad (7.16)$$

En pratique les fréquences f_0 et f_c sont généralement connues et $f_0 \gg f_c$. X_1 et X_2 peuvent alors être retrouvés à partir de Y par les transformations

$$\begin{aligned} X_{1,t} &= 2F_{f_c}[\cos(2\pi f_0 t) Y_t] \\ X_{2,t} &= 2F_{f_c}[-\sin(2\pi f_0 t) Y_t], \end{aligned} \quad (7.17)$$

où F_{f_c} représente l'opération de filtrage de réponse fréquentielle $\mathbb{I}_{[-f_c, f_c]}(f)$. L'opération (7.17) est connue sous le nom de démodulation en phase et en quadrature.

Indiquons par ailleurs que l'on peut introduire une représentation complexe des signaux modulés en phase et en quadrature en notant

$$\begin{aligned} Y_2(t) &= X_{1,t} \sin(2\pi f_0 t) + X_{2,t} \cos(2\pi f_0 t) \\ Z_t &= Y_{1,t} + iY_{2,t} \\ &= (X_{1,t} + iX_{2,t})e^{2i\pi f_0 t}. \end{aligned} \quad (7.18)$$

7.2.2 Représentation analytique et modulation BLU.

Introduisons maintenant la notion de processus analytique associé à un processus. Le filtre de Hilbert est défini par la réponse fréquentielle $H(f) = -i.sign(f)$. La transformée de

Hilbert du processus X est donc donnée par

$$\tilde{X}_t = \int_{\mathbb{R}} -i \cdot \text{sign}(f) e^{2i\pi ft} \hat{X}(df). \quad (7.19)$$

La mesure aléatoire de \tilde{X} est donnée par $-i \cdot \text{sign}(f) \hat{X}(df)$. Comme X est à valeurs réelles, il est clair que $\hat{X}(df) = \hat{X}(-df)^*$. Cette propriété est encore vérifiée par la mesure aléatoire de \tilde{X} ; \tilde{X} est donc également un processus à valeurs réelles, car

$$\tilde{X}_t = 2\mathcal{R}e\left[\int_{\mathbb{R}_+} e^{2i\pi ft} \hat{X}(df)\right]. \quad (7.20)$$

Le processus défini par $Z_t = X_t + i\tilde{X}_t$ possède alors la propriété remarquable d'avoir un support spectral contenu dans \mathbb{R}_+ :

$$\mu_Z(df) = 4 \times \mathbb{1}_{\mathbb{R}_+}(f) \mu_X(df), \quad (7.21)$$

car $\hat{Z}(df) = \hat{X}(df) + i\hat{\tilde{X}}(df) = [1 + \text{sign}(f)]\hat{X}(df)$.

Par ailleurs, X peut être obtenu facilement à partir de Z : $Z_t = X_t + i\tilde{X}_t$ et X_t et \tilde{X}_t sont réels, donc $X_t = (Z_t + Z_t^*)/2$.

Si le support de $S_X(f)$ est borné, il apparaît que $S_Z(f)$ occupe un support de taille 2 fois moindre. Cette propriété est exploitée en transmissions pour limiter la bande passante de signaux transmis sans perte d'information. En effet, on transmet parfois le signal $Z_t \cos(2\pi f_0 t)$ plutôt que $X_t \cos(2\pi f_0 t)$. Cette technique est connue sous le nom de modulation BLU (Bande Latérale Unique).

7.2.3 Représentation de Rice.

On peut se demander si pour un processus stationnaire réel Y_t donné on peut obtenir une représentation du type (7.14). Pour f_0 fixé, on peut poser $W_t = Y_t e^{-2i\pi f_0 t} = X_{1,t} + iX_{2,t}$. On a alors

$$\begin{aligned} Y_t &= \mathcal{R}e[W_t e^{2i\pi f_0 t}] \\ &= X_{1,t} \cos(2\pi f_0 t) - X_{2,t} \sin(2\pi f_0 t). \end{aligned} \quad (7.22)$$

La représentation cherchée existe donc, mais elle n'est pas unique. Ainsi, pour toute valeur de ϕ fixée, on peut par exemple définir $X_{1,t} + iX_{2,t} = Y_t e^{-2i\pi f_0 t + i\phi}$. Considérons la représentation particulière définie par

$$\begin{aligned} Z_t &= X_{1,t} + iX_{2,t} \\ &= (Y_t + i\tilde{Y}_t) e^{-2i\pi f_0 t}, \end{aligned} \quad (7.23)$$

où \tilde{Y} est la transformée de Hilbert de Y . La représentation (7.14) correspondante est appelée représentation de Rice. Le processus Z_t est appelé enveloppe complexe de Y . Sa mesure aléatoire et sa mesure spectrale sont données par

$$\begin{aligned} \hat{Z}(df) &= 2 \times \mathbb{1}_{\mathbb{R}_+}(f + f_0) \hat{Y}(df + f_0) \\ \text{et } \mu_Z(df) &= 4 \times \mathbb{1}_{\mathbb{R}_+}(f + f_0) \mu_Y(df + f_0). \end{aligned} \quad (7.24)$$

Lorsque f_0 ne représente pas une fréquence connue, on considère souvent la représentation obtenue pour la valeur de f_0 qui minimise $\int_{\mathbb{R}} (f - f_0)^2 \mu_Y(df)$, c'est à dire

$$f_0 = \frac{1}{\int_{\mathbb{R}} \mu_Y(df)} \int_{\mathbb{R}} f \mu_Y(df). \quad (7.25)$$

Revenons maintenant sur l'opération de démodulation des signaux modulés en phase et en quadrature. La démodulation du processus Y , dont le spectre est porté par $[-f_c - f_0, f_c - f_0] \cup [-f_c + f_0, f_c + f_0]$, fournit la représentation de Rice de Y . En effet, la mesure aléatoire de $X_1 + iX_2$, où X_1 et X_2 sont fournis par (7.17) est donnée par

$$\begin{aligned} \hat{X}_1(df) + i\hat{X}_2(df) &= 2 \times \mathbb{I}_{[-f_c, f_c]}(f) \hat{Y}(df + f_0) \\ &= 2 \times \mathbb{I}_{\mathbb{R}^+}(f + f_0) \hat{Y}(df + f_0), \end{aligned} \quad (7.26)$$

donc, d'après (7.23) et (7.24), $X_{1,t} + iX_{2,t} = e^{-2i\pi f_0 t} (Y_t + i\tilde{Y}_t)$. On aura intérêt à représenter graphiquement les supports spectraux des processus et des filtres mis en jeu afin de bien visualiser les relations étudiées dans ce paragraphe.

7.2.4 Démodulation en présence de bruit.

Dans un contexte pratique, il est souvent nécessaire d'étudier comment se comporte l'opération de démodulation vis à vis de la présence d'un bruit blanc additif B , de densité spectrale de puissance constante égale à σ^2 . On notera $B_{1,t} = 2F_{f_c} [\cos(2\pi f_0 t) B_t]$ et $B_{2,t} = 2F_{f_c} [-\sin(2\pi f_0 t) B_t]$ le bruit démodulé en phase et en quadrature. La démodulation du processus $Y + B$ conduit alors à $X_1 + B_1$ et $X_2 + B_2$, dont les mesures spectrales sont $\mu_{X_1}(df) + 2 \times \mathbb{I}_{[-f_c, f_c]}(f) \sigma^2 df$ et $\mu_{X_2}(df) + 2 \times \mathbb{I}_{[-f_c, f_c]}(f) \sigma^2 df$. On notera le facteur 2 qui apparaît dans l'expression de la DSP du bruit démodulé.

Ce facteur 2 pourrait laisser croire que l'opération de démodulation entraîne une dégradation du rapport signal à bruit, c'est à dire du rapport de la puissance des signaux utiles et de celle du bruit. Il n'en est rien. En effet, un calcul rapide montre que pour les signaux transmis le rapport signal à bruit total pour le processus $Y + B$ dans la bande utile $[-f_c - f_0, f_c - f_0] \cup [-f_c + f_0, f_c + f_0]$ vaut

$$RSB_{\text{mod}} = \frac{\frac{1}{2} [\int_{\mathbb{R}} \mu_{X_1}(df) + \int_{\mathbb{R}} \mu_{X_2}(df)]}{(4f_c) \sigma^2}, \quad (7.27)$$

et qu'après démodulation, il devient, pour la bande utile $[-f_c, f_c]$,

$$RSB_{\text{démod}} = \frac{\int_{\mathbb{R}} \mu_{X_1}(df) + \int_{\mathbb{R}} \mu_{X_2}(df)}{(2f_c) \times 2\sigma^2 + (2f_c) \times 2\sigma^2}. \quad (7.28)$$

Le rapport signal à bruit n'est donc pas modifié par l'opération de démodulation.

Remarquons de plus que B_1 et B_2 sont décorréliées car

$$\begin{aligned}\mu_{B_1, B_2}(df) &= 4 \times \mathbb{I}_{[-f_c, f_c]}(f) = \times \mathbb{E}\left[\left(\frac{\hat{B}(df - f_0) + \hat{B}(df + f_0)}{2}\right)\left(\frac{\hat{B}(df - f_0) - \hat{B}(df + f_0)}{2i}\right)^*\right] \\ &= -i \times \mathbb{I}_{[-f_c, f_c]}(f)(\sigma^2 df - \sigma^2 df) \\ &= 0.\end{aligned}\tag{7.29}$$

Exercices

7.1 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL et $Y = (Y_t)_{t \in \mathbb{R}}$ avec $Y_t = X_{h(t)}$. A quelle condition sur $h(t)$ Y est-il un processus SSL ?

7.2 Soit $V = (V_n)_{n \in \mathbb{Z}}$ une suite IID. Calculez le spectre de $Y = (Y_n)_{n \in \mathbb{Z}}$ lorsque $Y_n = \text{sign}(bV_n + V_{n-1})$.

7.3 Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus circulaire SSL gaussien de fonction de covariance $R_X(n)$ et $Y = (Y_n)_{n \in \mathbb{Z}}$ défini par

$$Y_n = \sum_{m=0, p} \left(\sum_{k=1, r} a_{m, k} X_{n-m}^k \right).\tag{7.30}$$

Calculez la R_Y .

7.4 Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus gaussien à valeurs réelles et $Y = (Y_t)_{t \in \mathbb{R}}$ avec $Y_t = \sin(X_t)$. Calculez les fonctions de moyenne et de covariance de Y .

7.5 (Détecteur quadratique) Le signal $\mathbb{I}_{[0, T]}(t) \cos(2\pi ft)$ est transmis vers un récepteur sur un canal de propagation de réponse impulsionnelle $h(t) = \sum_{k=1, p} A_k e^{i\phi_k} \delta_{\tau_k}$.

a) Donnez l'expression du signal reçu, noté Y_t .

b) Afin d'estimer A_k et les retards τ_k , Y_t est passé dans un détecteur quadratique qui réalise une démodulation en phase et en quadrature, suivie par un filtrage de réponse impulsionnelle $\mathbb{I}_{[0, T]}(t)$ puis une somme quadratique des deux voies. On obtient ainsi la sortie

$$Z_t = (\mathbb{I}_{[0, T]}(t) * (Y_t \times \cos(2\pi ft)))^2 + (\mathbb{I}_{[0, T]}(t) * (Y_t \times \sin(2\pi ft)))^2.\tag{7.31}$$

On suppose que $f \gg T^{-1}$ et $T < (1/2)|\tau_k - \tau_l|$ pour $k \neq l$. Q

a) Quelle est l'expression de Z_t en fonction des paramètres du canal

b) Comment peut on retrouver les paramètres $(A_k, \tau_k)_{k=1, p}$ à partir de Z_t ?

7.6 (Modulation de phase) Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus ergodique au sens stricte

et and $Y = (Y_t)_{t \in \mathbb{R}}$ défini par

$$Y_t = \exp(2i\pi f_0 t + i\alpha X_t + i\phi), \quad (7.32)$$

où α est une constante et ϕ une variable aléatoire indépendante de X et de loi uniforme sur $[0, 2\pi]$. Calculez $R_Y(t)$ en supposant que X est gaussien, de fonction de covariance $R_X(t)$.

7.7 (Détection de maximum) Soit $X = (X_n)_{n \in \mathbb{Z}}$ une séquence de variables aléatoires centrées et $Y^{(N)} = (Y_n^{(N)})_{n \in \mathbb{Z}}$ défini par $Y_n^{(N)} = \max\{X_n, X_{n-1}, \dots, X_{n-N+1}\}$.

a) Calculez la loi de $Y_n^{(N)}$ en fonction de X_n .

b) Montrez que $Y^{(N)}$ est stationnaire.

7.8 (Limiteur saturé) Soit $X = (X_t)_{t \in \mathbb{R}}$ un processus SSL et $Z = (Z_t)_{t \in \mathbb{R}}$ défini par

$$Z_t = \max\{-1, X_t\} \mathbb{1}_{\mathbb{R}_-}(X_t) + \min\{1, X_t\} \mathbb{1}_{\mathbb{R}_+}(X_t). \quad (7.33)$$

a) Calculez la loi de Z_t en fonction de celle de X_t .

b) Calculez la fonction de covariance de Z dans les cas suivants

– lorsque X est un processus harmonique de la forme $X_t = \xi e^{2i\pi f t}$

– lorsque X est gaussien, de fonction de covariance $R_X(t)$.

7.9 (Effet Doppler) Soit une source mobile émettant un signal $X_t = e^{2i\pi f t}$ depuis un point r et animée d'un vecteur vitesse v . Le signal se propage à la vitesse c jusqu'à un récepteur placé en 0. On note $u = \|r\|^{-1} r$.

a) Montrez qu'à un facteur d'amplitude et à un retard près le récepteur observe le signal

$$y(t) = e^{2i\pi f(1+u^T v/c)^{-1}t}. \quad (7.34)$$

b) Plus généralement, si X est un processus SSL continu en moyenne quadratique transmis par la source, montrez que le signal reçu Y est un processus de la forme

$$Y_t = \int_{\mathbb{R}} e^{2i\pi f(1+u^T v/c)^{-1}t} \hat{X}(df). \quad (7.35)$$

c) Calculez $R_Y(t)$ et $\mu_Y(df)$ en fonction respectivement de $R_X(t)$ et de $\mu_X(df)$.

7.10 (Quantification) Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus stationnaire. Afin d'effectuer un traitement numérique sur X , on le quantifie de la façon suivante : on définit $Y = (Y_n)_{n \in \mathbb{Z}}$ par

$$\begin{cases} Y_n = m^{-1}(k + 1/2) & \text{if } k/m \leq X_n \leq (k + 1)/m \\ & \text{and } k = -N, -N + 1, \dots, N - 1, \\ Y_n = m^{-1}(N - 1/2) & \text{if } X_n \geq N/m \\ Y_n = m^{-1}(-N + 1/2) & \text{if } X_n \leq -N/m. \end{cases} \quad (7.36)$$

a) Calculez la variance de l'erreur de quantification définie par $\|X_n - Y_n\|^2$ quand $X_n \sim \mathcal{U}_{[-N/m, N/m]}$.

b) Plus généralement, on définit K valeurs, notées $(X_{q,k})_{k=1,K}$ et $K-1$ seuils $(S_k)_{k=1,K-1}$ tels que $Y_n = X_{q,1}$ si $X_n < S_1$, $Y_n = X_{q,k}$ si $S_k \leq X_n < S_{k+1}$ pour $1 < k < K-2$, et $Y_n = X_{q,K}$ si $X_n > S_{K-1}$. Exprimez $\|X_n - Y_n\|^2$ en fonction des $X_{q,K}$, of S_k , et de la loi de X .

c) Expliquez comment il est possible de rechercher un minimum local de $\|X_n - Y_n\|^2$ au moyen d'une technique d'optimisation alternée consistant à optimiser le choix des seuils S_k pour des valeurs des $X_{q,k}$ fixées puis à optimiser le choix des $X_{q,K}$ pour des valeurs des seuils fixées. Cette technique de réalisation d'un quantificateur appelé *algorithme de la K-moyenne*.

7.11(Communications AMRC)¹ Soit $X = (X_t)_{t \in \mathbb{R}}$, avec

$$X_t = \sum_{k \in \mathbb{Z}} A_k \mathbb{I}_{[0, T_s]}(t - kT_s + \phi). \quad (7.37)$$

Les coefficients A_k sont des variables aléatoires indépendantes, avec $P(A_k = 1) = P(A_k = -1) = 1/2$ et ϕ est une variable aléatoire indépendante des A_k ($k \in \mathbb{Z}$), de loi uniforme sur $[0, T_s]$.

a) Donnez l'expression du spectre de X .

b) On considère maintenant une séquence de variables aléatoires indépendantes C_m ($m = 0, \dots, N-1$), avec $P(C_m = 1) = P(C_m = -1) = 1/2$, et on note

$$s(t) = \frac{1}{\sqrt{N}} \sum_{m=0, N-1} C_m \mathbb{I}_{[0, T_c]}(t - mT_c), \quad (7.38)$$

où $T_c = T_s/N$. Calculez le spectre de $Y = (Y_t)_{t \in \mathbb{R}}$ défini par

$$Y_t = \sum_{k \in \mathbb{Z}} A_k \mathbb{I}_{[0, T_s]}(t - kT_s + \phi) s(t - kT_s + \phi), \quad (7.39)$$

et comparez le à X .

c) Si le récepteur connaît les séquences $(C_m)_{m=1,N}$ utilisées pour l'étalement de spectre du côté de l'émetteur, montrez que X peut être retrouvé à partir de Y . (Indications : considérez les grandeurs $\int Y_t \times (\sum_{k \in \mathbb{Z}} s(t - kT_s + \phi)) dt$.)

d) Soit $(C_m^{(1)})_{m=0, N-1}$ et $(C_m^{(2)})_{m=0, N-1}$ deux séquences et $s^{(1)}(t)$ and $s^{(2)}(t)$ les signaux correspondants calculés comme dans l'équation (7.38). On suppose que $\sum_{m=0, N-1} C_m^{(1)} C_m^{(2)} = 0$. Montrez qu'alors il est possible de transmettre simultanément des signaux de la forme $X_t^{(1)} = \sum_{k \in \mathbb{Z}} A_k^{(1)} \mathbb{I}_{[0, T_s]}(t - kT_s + \phi)$ et $X_t^{(2)} = \sum_{k \in \mathbb{Z}} A_k^{(2)} \mathbb{I}_{[0, T_s]}(t - kT_s + \phi)$ dans la même bande de fréquence et de les estimer sans faire apparaître de termes d'interférence entre les deux séquences.

1. AMRC : Accès Multiple à Répartition par Codes (en anglais, CDMA : Code Division Multiple Access)

Chapitre 8

Prédiction linéaire horizon fini des processus SSL

Objectif On cherche à approcher, au sens du critère du minimum de variance de l'erreur de prédiction, la variable X_n d'un processus discret X stationnaire au second ordre par une combinaison linéaire des variables aléatoires $(X_{n-k})_{k \geq 1}$. Certaines applications importantes telles que la modélisation des systèmes, la synthèse des filtres, ou encore la compression des signaux, utilisent les résultats de la théorie de la prédiction linéaire.

8.1 Espérance conditionnelle et prédiction linéaire.

A ce point de l'exposé, il est peut être utile de faire quelques rappels concernant l'espérance conditionnelle.

Soient respectivement X et Y une variable aléatoire et un vecteur aléatoire de taille n de $L^2(\Omega, \mathcal{A}, P)$. On est parfois amené à chercher la valeur prise par X à partir de la connaissance de la valeur prise par le vecteur Y . Pour cela, on peut chercher à approcher X par une variable aléatoire de la forme $\phi(Y)$. Un choix possible et assez naturel consiste à choisir cette fonction de telle sorte que la variable aléatoire $\phi(Y)$ soit la plus proche possible de X dans $L^2(\Omega, \mathcal{A}, P)$.

Soit $\hat{\phi}$ la fonction qui réalise cet optimum :

$$\hat{\phi}(Y) = \arg \min_{\phi(Y) \in L^2(\Omega, \mathcal{A}, P)} \| X - \phi(Y) \|^2 \quad (8.1)$$

La solution de ce problème est notée $\mathbb{E}[X|Y]$ et appelée espérance conditionnelle de X sachant Y .

Malheureusement, si la connaissance de la loi de (X, Y) est limitée celle de ses fonctions de moyenne et de covariance, le calcul de $\mathbb{E}[X|Y]$ n'est pas possible, sauf dans le cas où X est un processus gaussien, auquel cas les formules des espérances conditionnelles correspondent

celles de la régression linéaire. De plus, la connaissance de la loi de (X, Y) ne garantit pas que l'on puisse exprimer de façon simple $\mathbb{E}[X|Y]$.

Dans le cadre de la prédiction d'un processus X , on cherche plus particulièrement à évaluer des grandeurs du type $\mathbb{E}[X_n | (X_{n-k})_{k=1,p}]$. La prédiction linéaire à horizon fini consiste plus simplement à calculer la projection de X_n sur l'espace linéaire engendré par un nombre fini X_{n-1}, \dots, X_{n-p} de variables aléatoires du passé de X , et noté $H_{X,n-1,n-p}$. Le calcul de cette projection se ramène à la résolution d'un système d'équations linéaires et est d'un grand intérêt pratique, en particulier lorsqu'on ne dispose que de la connaissance d'un nombre fini de coefficients de covariance de X .

8.2 Prédiction linéaire à horizon fini.

Dans la suite, on notera $H_{X,n_1,n_2} = \text{vect}\{X_{n_1}, X_{n_1+1}, \dots, X_{n_2}\}$ ($n_1 \leq n_2$).

Souvent en pratique, les variables aléatoires X_{n-k} ne sont observées que pour un ensemble limité de valeurs de k : $k = 1, p$. Dans ces conditions, le problème de la prédiction linéaire à horizon fini consiste à chercher la projection de X_n sur $H_{X,n-p,n-1}$, note

$$X_n | H_{X,n-p,n-1} = \sum_{k=1,p} a_k X_{n-k}, \quad (8.2)$$

c'est à dire les coefficients $\alpha_k = a_k$ qui minimisent le critère $\| X_n - \sum_{k=1,p} \alpha_k X_{n-k} \|^2$. $X_n - \sum_{k=1,p} a_k X_{n-k}$ représente donc l'erreur de prédiction.

On notera σ_p^2 le minimum de ce critère qui représente la variance de l'erreur de prédiction.

Il est alors aisé de vérifier que les valeurs a_k des coefficients α_k qui réalisent le minimum sont données par

$$a = T_{p-1}^{-1} r_p, \quad (8.3)$$

où $a = [a_1, \dots, a_p]^T$, $r_p = [R_X(1), \dots, R_X(p)]^T$ et T_{p-1} est la matrice hermitienne de Toeplitz de taille p et de terme général $[T_{p-1}]_{i,j} = R_X(i-j)$, $(i, j) = 1, p$. Rappelons qu'une matrice de Toeplitz T est caractérisée par le fait que les termes d'une même parallèle à la diagonale sont identiques ; c'est à dire que $[T]_{i,j}$ ne dépend que de $i - j$.

Pour justifier la formule (8.3), remarquons que d'après le théorème de projection qui indique que l'erreur de prédiction $X_n - \sum_{k=1,p} a_k X_{n-k}$ est orthogonal à chacune des variables X_{n-l} pour $l = 1, \dots, p$ (eg [?] p.65), les coefficients a_k sont caractérisés par les relations

$$\begin{aligned} \|(X_n - \sum_{k=1,p} a_k X_{n-k}) X_{n-l}^*\|^2 &= R_X(l) - \sum_{k=1,p} a_k R_X(l-k) \\ &= 0, \quad \text{pour } l = 1, p, \end{aligned} \quad (8.4)$$

et les équations de ce système s'écrivent sous forme matricielle, $r_p - T_{p-1} a = 0$, soit $a = T_{p-1}^{-1} r_p$. Les relations

$$R_X(l) - \sum_{k=1,p} a_k R_X(l-k) = 0, \quad l = 1, p, \quad (8.5)$$

sont appelées équations de Yule-Walker, et le filtre de fonction de transfert $a^{(p)}(z) = 1 - \sum_{k=1,p} a_k z^{-k}$ est souvent appelé filtre d'erreur de prédiction d'ordre p .

Si la matrice T_{p-1} n'est pas inversible, on montre que le minimum du critère vaut 0. On peut montrer que lorsque le spectre de X possède une partie à densité non identiquement nulle, T_{p-1} est inversible. On se placera ici dans cette situation.

Proposition 11 *La variance de l'erreur de prédiction à horizon fini est donnée par*

$$\begin{aligned}\sigma_p^2 &= \| X_n - \sum_{k=1,p} a_k X_{n-k} \|^2 \\ &= R_X(0) - r_p^H T_{p-1}^{-1} r_p.\end{aligned}\tag{8.6}$$

Démonstration Comme $a = T_{p-1}^{-1} r_p$, il vient d'après les relations (8.4)

$$\begin{aligned}\sigma_p^2 &= \mathbb{E}[(X_n - \sum_{k=1,p} a_k X_{n-k})(X_n - \sum_{k=1,p} a_k X_{n-k})^*] \\ &= \mathbb{E}[(X_n - \sum_{k=1,p} a_k X_{n-k})X_n^*] \\ &= R_X(0) - \sum_{k=1,p} a_k R_X(-k) \\ &= R_X(0) - r_p^H a \\ &= R_X(0) - r_p^H T_{p-1}^{-1} r_p.\end{aligned}\tag{8.7}$$

□

Exercices

8.1 On cherche à résoudre le problème de prédiction d'un processus à valeurs complexes suivant

$$\min_{\{c_k, d_k\}_{k=1,p}} \| X_n - \sum_{k=1,p} c_k X_{n-k} - \sum_{l=1,p} d_l X_{n-l}^* \|^2.\tag{8.8}$$

a) Calculez les valeurs des coefficients $\{c_k, d_k\}_{k=1,p}$ qui réalisent le minimum de (8.8).

b) Si X est circulaire, montrez que $d_k = 0$ for $k = 1, \dots, p$.

8.2 Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus circulaire et SSL.

a) Calculez les coefficients du polynôme $B_p(z) = \sum_{k=0,p} b_{k,p} z^{-k}$ pour lequel le critère $\| [B_p(z)]X_n \|^2 - 2\mathcal{R}e[b_{0,p}]$ est minimum.

b) Quelle est la relation entre des coefficients de $B_p(z)$ et le filtre d'erreur de prédiction ?

8.3 Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus SSL de mesure spectrale $\mu_X(df) = \sigma^2 df + \delta_{f_0}$. Montrez que X satisfait une équation aux différences de la forme

$$X_n - \alpha X_{n-1} = \beta_0 V_n - \beta_1 V_{n-1}, \quad (8.9)$$

où $V = (V_n)_{n \in \mathbb{Z}}$ est un bruit blanc de variance unitaire : $\|V_n\| = 1$. Calculez α , β_0 et β_1 .

8.4 (interpolation linéaire) Soit $X = (X_n)_{n \in \mathbb{Z}}$ un processus SSL. Résolvez le problème d'interpolation suivant :

$$\min_{(a_k, b_k)_{k=1, p}} \left\| X_n - \sum_{k=1, p} a_k X_{n-k} - \sum_{l=1, p} b_l X_{n+l} \right\|. \quad (8.10)$$

Chapitre 9

Filtrage adapté

Objectif Dans de nombreux problèmes, on cherche à évaluer un processus X à partir de la connaissance partielle d'un processus Y . On s'intéresse ici au filtre adapté qui permet de détecter la présence éventuelle d'un signal déterministe connu, bruité par un processus aléatoire.

9.1 Problème

Dans un certain nombre d'applications telles que le radar ou les transmissions numériques on veut pouvoir détecter la présence d'un signal connu $g(t)$ en entrée d'un récepteur. En fait, $g(t)$ est en général connu seulement à un retard près ou à un coefficient d'amplitude près. C'est ce retard (dans le cas du radar) ou cette amplitude (dans le cas des transmissions) qui constituera l'information recherchée.

On suppose qu'on est en présence d'un bruit additif B qui vient perturber l'observation. Cette observation est donc un processus de la forme

$$X_t = g(t) + B_t \quad (9.1)$$

Pour détecter la présence du signal $g(t)$, on observe la sortie d'un filtre de réponse fréquentielle $H(f)$ excité par X_t . On choisit $H(f)$ de telle sorte que le rapport, noté ρ , de la puissance en sortie du signal utile et de celle du bruit, soit maximal à un instant t_0 fixé.

9.2 Expression du filtre adapté

La sortie du filtre est donnée par

$$\begin{aligned} Y_t &= \int_{\mathbb{R}} e^{2i\pi ft} H(f) G(f) df + \int_{\mathbb{R}} e^{2i\pi ft} H(f) \hat{B}(df) \\ &= Y_{1,t} + Y_{2,t}, \end{aligned} \quad (9.2)$$

où $G(f)$ est la transformée de Fourier de la fonction $g(t)$. Le rapport ρ_t est alors donné par

$$\begin{aligned}\rho_t &= \frac{\|Y_{1,t}\|^2}{\|Y_{2,t}\|^2} \\ &= \frac{|\int_{\mathbb{R}} e^{2i\pi ft} H(f)G(f)df|^2}{\int_{\mathbb{R}} |H(f)|^2 \mu_B(df)}.\end{aligned}\tag{9.3}$$

On supposera que B possède un spectre à densité, de densité $S_B(f)$. On a alors le résultat suivant

Proposition 12 *La maximisation, à l'instant $t = t_0$, du rapport signal à bruit ρ_t du processus $g(t) + B_t$ est obtenue pour le filtre de réponse fréquentielle*

$$H(f) = K e^{-2i\pi ft_0} G^*(f) S_B^{-1}(f), \tag{9.4}$$

où K est une constante. A l'instant $t = t_0$, le rapport signal à bruit est alors donné par

$$\rho_{t_0} = \int_{\mathbb{R}} \frac{|G(f)|^2}{S_B(f)} df. \tag{9.5}$$

Démonstration D'après l'inégalité de Cauchy-Schwarz,

$$\begin{aligned}|\int_{\mathbb{R}} e^{2i\pi ft} H(f)G(f)df|^2 &= |\int_{\mathbb{R}} [e^{2i\pi ft} H(f) S_B^{1/2}(f)] \times [G(f) S_B^{-1/2}(f)] df|^2 \\ &\leq \int_{\mathbb{R}} |H(f)|^2 S_B(f) df \int_{\mathbb{R}} |G(f)|^2 S_B^{-1}(f) df,\end{aligned}\tag{9.6}$$

l'égalité ayant lieu pour $t = t_0$ si et seulement si

$$e^{2i\pi ft_0} H(f) S_B^{1/2}(f) = K G^*(f) S_B^{-1/2}(f), \tag{9.7}$$

où K est une constante, soit

$$H(f) = K e^{-2i\pi ft_0} G^*(f) S_B^{-1}(f). \tag{9.8}$$

La valeur correspondante de ρ_{t_0} s'en déduit immédiatement. \square

En pratique, la position d'un pic en sortie du filtre adapté permet d'estimer un retard et l'amplitude de ce pic permet d'estimer l'amplitude du signal reçu.

Pour un bruit blanc et un signal $g(t)$ de support limité à $[0, T]$, on pourra vérifier à titre d'exercice que la réponse impulsionnelle d'un filtre adapté causal est donnée par

$$h(t) \propto g(t_0 - t), \text{ avec } t_0 \geq t. \tag{9.9}$$

Exercices

9.1 (Transmission d'un symbole en présence de bruit) Le signal $X_t = A\mathbb{1}_{[0,T]}(t)$ est transmis à un récepteur. A est une variable aléatoire de Bernoulli avec $P(A = 1) = p$ et $P(A = -1) = 1 - p$. Le signal $Y_t = X_t + V_t$, où $V = (V_t)_{t \in \mathbb{R}}$ est un bruit blanc gaussien indépendant de A , est observé au niveau du récepteur.

a) A l'instant T , on veut décider si un symbole $+1$ ou -1 a été transmis. Pour cela, on applique un filtre au niveau du récepteur qui fournit un Rapport Signal à Bruit (RSB) maximum à l'instant T . Donnez l'expression de la réponse impulsionnelle du filtre et l'expression du signal à sa sortie.

b) Soit Z_T la sortie du filtre à l'instant T . Un seuil de valeur S permet de prendre une décision sur la valeur transmise comme suit : si $Z_T > S$, on décide que $A = 1$, sinon, on décide que $A = -1$. Calculez la probabilité P_e qu'une mauvaise décision soit prise. Quelle est la valeur de S qui réalise le minimum de P_e ? Considérez les cas particuliers où $p = 1/2$ et $p = 0$.

9.2 (Estimation d'un signal déterministe inconnu en présence de bruit) Soit $Y_t = g(t) + V_t$, où $g(t)$ est une fonction déterministe inconnue et $V = (V_t)_{t \in \mathbb{R}}$ un bruit blanc de variance σ_V^2 . Y_t est observé sur l'intervalle de temps $t \in [0, T]$. Sur cet intervalle, $g(t)$ est approximée par son développement en série de Fourier tronqué $g_N(t) = \sum_{k=-N, N} \hat{g}_k e^{2i\pi kt/T}$.

a) Montrez qu'en utilisant un banc de filtres adaptés pour estimer les coefficients $(\hat{g}_k)_{k=-N, N}$ conduit à l'estimateur suivant de $g_N(t)$:

$$\tilde{g}_N(t) = \sum_{k=-N, N} \left(T^{-1} \int_{[0, T]} e^{-2i\pi kt/T} Y_t dt \right) e^{2i\pi kt/T}. \quad (9.10)$$

b) Montrez que l'erreur quadratique moyenne sur $[0, T]$, définie par $T^{-1} \int_{[0, T]} \|g(t) - \tilde{g}_N(t)\|^2 dt$, est donnée par

$$T^{-1} \int_{[0, T]} \|g(t) - \tilde{g}_N(t)\|^2 dt = \left(\sum_{k > N} |\hat{g}_k|^2 \right) + \frac{2N+1}{T} \sigma_V^2. \quad (9.11)$$

Quel serait un bon choix pour N ?

9.3 Soit $Y_n = X_n + V_n$ un processus SSL, avec V un bruit blanc décorrélé avec X , et $B(z) = \sum_{k=0, p} b_k z^{-k}$ une fonction de transfert. on cherche $b = [b_0, \dots, b_q]^T$ tel que le rapport signal à bruit (RSB) de $[b(z)]Y_n$ défini par

$$\rho_b = \| [b(z)]X_n \|^2 \| [b(z)]V_n \|^2 \quad (9.12)$$

soit maximum.

a) Montrez que pour cela il faut choisir pour b le plus grand vecteur propre de la matrice $T_{X,q}$, de taille $(q+1) \times (q+1)$ et de terme général $[T_{X,q}] = R_X(i-j)$. Donnez la valeur correspondante de ρ_b .

b) Résolvez le problème lorsque V est un bruit corrélé de matrice de covariance $T_{V,q}$.

Chapitre 10

Processus à spectres rationnels

Objectif Les processus à spectre rationnel jouent un rôle très importants pour la modélisation des propriétés du second ordre des processus. En particulier, les DSP continues par morceaux peuvent être approchées avec une précision arbitraire par des DSP rationnelles (et même polynômiales). De plus, les modèles de DSP rationnelles sont parcimonieux, car dans de nombreuses situations l'emploi de fractions rationnelles de degré faible est suffisant. Enfin, l'identification ou l'estimation des paramètres d'une DSP rationnelle est relativement simple, comme on le verra dans les chapitres suivants. On va donc indiquer quelques résultats significatifs concernant ces processus.

10.1 Equations aux différences et processus à spectres rationnels

Considérons l'équation aux différences

$$X_n + \sum_{k=1,p} a_k X_{n-k} = \sum_{l=0,q} b_l U_{n-l}, \quad (10.1)$$

où $U = (U_n)_{n \in \mathbb{Z}}$ est un processus stationnaire au second ordre fixé, de spectre absolument continu. Supposons de plus que la DSP $S_U(f)$ de U vérifie $0 < m \leq S_U(f) \leq M < \infty$. S'il existe un processus X stationnaire solution de l'équation (10.1), alors on a

$$a(e^{2i\pi f}) \hat{X}(df) = b(e^{2i\pi f}) \hat{U}(df), \quad (10.2)$$

avec

$$a(e^{2i\pi f}) = 1 + \sum_{k=1,p} a_k e^{-2i\pi k f}, \quad (10.3)$$

$$\text{et } b(e^{2i\pi f}) = \sum_{l=0,q} b_l e^{-2i\pi l f}.$$

Remarquons que si $a(z)$ s'annule sur le cercle unité, la réponse fréquentielle $b(e^{2i\pi f})a^{-1}(e^{2i\pi f})$ n'appartient pas à $L^2(\mathcal{I}, B(\mathcal{I}), S_U(f)df)$ et le problème n'a pas de solution. On supposera donc que $a(e^{2i\pi f}) \neq 0, \forall f$. Alors,

$$X_n = \int_{\mathcal{I}} e^{2i\pi n f} \frac{b(e^{2i\pi f})}{a(e^{2i\pi f})} \hat{U}(df) \quad (10.4)$$

est l'unique solution du problème.

On pourra encore écrire $X_n = \sum_{k \in \mathbb{Z}} h_k U_{n-k}$, où les coefficients h_k sont ceux du développement en série de Laurent de $b(z)a^{-1}(z)$ au voisinage du cercle unité. Remarquons que la série $\sum_{k \in \mathbb{Z}} h_k$ est absolument sommable.

Si de plus $b(z)a^{-1}(z)$ est la fonction de transfert d'un filtre causal, c'est à dire si $h_k = 0$ pour $k < 0$, alors la série $\sum_{k \in \mathbb{N}} h_k z^{-k}$ est convergente dans un domaine $|z| > 1 - \varepsilon$ ($\varepsilon > 0$). $b(z)a^{-1}(z)$ est donc holomorphe dans ce domaine et $a(z) \neq 0$ pour $|z| \geq 1$. La réciproque est immédiate et le théorème suivant résume ces résultats.

Théorème 8 *Si U est un processus stationnaire qui a un spectre à densité et dont la DSP vérifie*

$$0 < m \leq S_U(f) \leq M < \infty, \quad (10.5)$$

l'équation aux différences

$$X_n + \sum_{k=1,p} a_k X_{n-k} = \sum_{l=0,q} b_l U_{n-l}. \quad (10.6)$$

admet une solution si et seulement si $a(z) = 1 + \sum_{k=1,p} a_k z^{-k} \neq 0$ pour $|z| = 1$. On a alors

$$X_n = \int_{\mathcal{I}} e^{2i\pi n f} \frac{b(e^{2i\pi f})}{a(e^{2i\pi f})} \hat{U}(df), \quad (10.7)$$

avec $b(z) = \sum_{k=0,q} b_k z^{-k}$. De plus, le filtre de fonction de transfert $b(z)a^{-1}(z)$ est causal si et seulement si $a(z) \neq 0$ pour $|z| \geq 1$.

La DSP de X est donnée par

$$S_X(f) = \left| \frac{b(e^{2i\pi f})}{a(e^{2i\pi f})} \right|^2 S_U(f). \quad (10.8)$$

Si X possède une DSP qui est une fraction rationnelle de $e^{-2i\pi f}$, on dit que X est un processus à spectre rationnel, ou encore un processus ARMA¹. C'est en particulier le cas lorsque U est un bruit blanc puisque $S_U(f)$ est alors constant. On supposera désormais que U est un bruit blanc.

Indiquons maintenant une propriété intéressante des covariances des processus ARMA

Proposition 13 *Les coefficients de covariance des processus à spectre rationnel décroissent de façon exponentielle.*

Un processus X à spectre rationnel peut donc être représenté comme la sortie d'un filtre excité par un bruit blanc et de fonction de transfert rationnelle $b(z)a^{-1}(z)$. On parle alors de représentation ARMA (de l'anglais *Auto Regressive Moving Average*) de X .

1. ARMA : AutoRegressive Moving Average

10.2 Représentation d'état des modèles ARMA.

Les modèles ARMA peuvent s'écrire sous la forme d'un modèle d'état linéaire et stationnaire. Considérons en effet un modèle ARMA défini par l'équation

$$Y_n + \sum_{k=1,p} a_k Y_{n-k} = \sum_{l=0,q} b_l U_{n-k} \quad (10.9)$$

et commençons par supposer que $p = q + 1$. Considérons maintenant le modèle d'état

$$\begin{cases} \mathbf{X}_{n+1} &= A\mathbf{X}_n + BU_{n+1} \\ Z_n &= C\mathbf{X}_n \end{cases} \quad (10.10)$$

avec $\mathbf{X}_n = [X_n, \dots, X_{n-p+1}]^T$, $B = [1, 0, \dots, 0]^T$, $C = [b_0, \dots, b_q]$ et

$$A = \begin{bmatrix} -a_1 & \cdot & \cdot & \cdot & -a_p \\ 1 & 0 & \cdot & \cdot & 0 \\ 0 & 1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 1 & 0 \end{bmatrix}. \quad (10.11)$$

Montrons que $Z = Y$.

$$\begin{aligned} Z_n &= \sum_{l=0,q} b_l X_{n-l} \\ \sum_{k=0,p} a_k Z_{n-k} &= \sum_{l=0,q} b_l \left(\sum_{k=0,p} a_k X_{n-l-k} \right) \\ &= \sum_{l=0,q} b_l U_{n-l}. \end{aligned} \quad (10.12)$$

Donc $Z = Y$.

Si $p \neq q + 1$, il suffit de procéder de la même manière en complétant la plus courte des séquences $(a_k)_{k=1,p}$ et $(b_l)_{l=0,q}$ par des coefficients égaux à 0, jusqu'à obtenir deux séquences de même longueur.

Considérons réciproquement un modèle d'état linéaire et stationnaire de la forme

$$\begin{cases} \mathbf{X}_{n+1} &= A\mathbf{X}_n + BU_{n+1} \\ Y_n &= C\mathbf{X}_n \end{cases} \quad (10.13)$$

Pour un tel modèle, $[I_p - Az^{-1}]\mathbf{X}_n = BU_n$, donc

$$Y_n = [C(I - Az^{-1})^{-1}B]U_n. \quad (10.14)$$

Il est clair que $C(I - Az^{-1})^{-1}B$ est une fraction rationnelle en z^{-1} , car les coefficients de la matrice $(I - Az^{-1})^{-1}$ s'écrivent comme le rapport des cofacteurs de $I - Az^{-1}$ et du déterminant de $I - Az^{-1}$. Donc, si U est un bruit blanc, Y est un processus ARMA si les zéros de $|I - Az^{-1}|$, c'est à dire les valeurs propres de A , sont dans le disque unité. On vérifie que $|I - Az^{-1}| = a(z)$, avec $a(z) = \sum_{k=0,p} a_k z^{-k}$, ce qui montre au passage que le calcul des racines d'un polynôme $a(z)$ se ramène à celui des valeurs propres de la matrice A qui lui correspond par la relation (10.11), et qui est appelée matrice compagnon de $a(z)$.

10.3 Identification des processus ARMA.

Etant donné un processus X dont le spectre s'écrit comme une fraction rationnelle de $e^{2i\pi f}$, on cherche à identifier un filtre de fonction de transfert rationnelle tel que X puisse être vu comme la sortie d'un tel filtre excité par un bruit blanc :

$$X_n + \sum_{k=1,p} a_k X_{n-k} = \sum_{l=0,q} b_l U_{n-l}, \quad (10.15)$$

On cherche à identifier ce filtre à partir de la connaissance des coefficients de covariance $(R_X(k))_{k=0,N}$.

10.3.1 Identification de la partie AR.

Pour identifier les $(a_k)_{k=1,p}$, remarquons que

$$\begin{aligned} X_n &= -\sum_{k=1,p} a_k X_{n-k} + \sum_{l=0,q} b_l U_{n-l} \\ &= \sum_{k=0,\infty} h_k U_{n-k}. \end{aligned} \quad (10.16)$$

Comme U est un bruit blanc, $X_{n-l} \in H_{U,n-l}$ et $X_n + \sum_{k=1,p} a_k X_{n-k} \in H_{U,n-q,n}$. Donc,

$$\begin{aligned} \mathbb{E}[(X_n + \sum_{k=1,p} a_k X_{n-k})X_{n-l}^*] &= R_X(l) + \sum_{k=1,p} a_k R_X(l-k) \\ &= 0, \quad \text{pour } l > q. \end{aligned} \quad (10.17)$$

La résolution du système linéaire ainsi obtenu fournit les coefficients $(a_k)_{k=1,p}$.

10.3.2 Identification de la partie MA.

Pour identifier les $(b_l)_{l=0,q}$, remarquons que $S_X(f)|a(e^{2i\pi f})|^2 = |b(e^{2i\pi f})|^2$. Il est alors clair qu'on peut calculer les coefficients $(b_l)_{l=0,q}$ en identifiant les coefficients de l'égalité polynômiale

$$\left(\sum_{n=-(p+q)}^{p+q} R_X(n)e^{-2i\pi n f} \right) |a(e^{2i\pi f})|^2 = |b(e^{2i\pi f})|^2. \quad (10.18)$$

On obtient ainsi des équations non linéaires dont la solution peut être trouvée par des techniques numériques.

Exercices

Dans ce qui suit $X = (X_n)_{n \in \mathbb{Z}}$ est un processus ARMA(p, q) qui vérifie l'équation aux différences

$$X_n + \sum_{k=1,p} a_k X_{n-k} = \sum_{l=0,q} b_l V_{n-l}, \quad (10.19)$$

où $V = (V_n)_{n \in \mathbb{Z}}$ est un bruit blanc de variance σ_V^2 .

10.1 Si X est un processus AR(1), vérifiez que $\sigma_X^2 = \sigma_V^2(1 - |a|)^2$.

10.2 (filtre AR causal) Soit la fonction de transfert $h(z) = (1 + a_1 z^{-1} + a_2 z^{-2})^{-1}$. Pour quelles valeurs de (a_1, a_2) dans le plan \mathbb{R}^2 le filtre est-il causal ?

10.3 (Equation aux différences avec des conditions initiales) Soit $Y = (Y_n)_{n \in \mathbb{N}}$ un processus tel que $Y_0 = 0$ et $Y_n = aY_{n-1} + V_n$ pour $n > 0$, avec $|a| < 1$. Les variables aléatoires $(V_n)_{n \in \mathbb{N}^*}$ sont décorrélées, de variance σ_V^2 . Calculez $\text{cov}(Y_{m+n}, Y_m)$ et étudiez son comportement lorsque m tend vers $+\infty$.

Chapitre 11

Notions d'estimation statistique

Objectif Les statistiques permettent d'évaluer des paramètres relatifs à la loi de phénomènes aléatoires lorsque ces lois ne sont pas connues mais que l'on dispose d'un certain nombre de réalisations du phénomène étudié.

11.1 Généralités

On comprend bien que des paramètres qui s'écrivent comme l'espérance mathématique d'une certaine fonction des variables aléatoires mises en jeu pourront souvent être estimés de façon satisfaisante par des moyennes, en vertu de la loi des grands nombres et du théorème de la limite centrale. La thorie des statistiques offre un cadre gnral au dveloppement de ce type de dmarche.

Indiquons pour commencer quelques notions élémentaires relatives à l'estimation statistique. On se place dans un espace probabilisé (Ω, \mathcal{A}, P) où P est inconnue. Dans le cas de l'estimation non paramétrique, l'information sur P est peu précise (eg P possède une densité). Dans le cas de l'estimation paramétrique, on suppose que P appartient à un ensemble $\{P_\theta; \theta \in \Theta\}$, où θ est un vecteur de paramètres qui caractérise entièrement P_θ . On s'arrange généralement pour que Θ et $\{P_\theta; \theta \in \Theta\}$ soient en correspondance bijective.

Modèle !statistique On définit un modèle statistique comme un triplet $(\Omega, \mathcal{A}, \mathcal{P})$, où (Ω, \mathcal{A}) est un espace probabilisable et \mathcal{P} une famille de lois de probabilités sur (Ω, \mathcal{A}) . Dans le cas de l'estimation paramétrique, on aura $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ et un modèle de la forme $(\Omega, \mathcal{A}, \{P_\theta; \theta \in \Theta\})$. C'est par exemple le cas si le processus étudié est un bruit blanc gaussien de moyenne m et de variance σ^2 inconnues : ici $\theta = (m, \sigma^2)$.

Dans le cadre de l'estimation statistique, on est amené à introduire les notions suivantes

Observation L'observation x est constituée par la réalisation d'une variable aléatoire X , ou d'une séquence (X_1, \dots, X_n) de variables aléatoires. Si les variables aléatoires $(X_i)_{i=1,n}$ constituent une suite IID, on parle d'échantillon de taille n .

Estimateur L'estimation ponctuelle vise à évaluer à partir de l'observation x une grandeur caractérisée par la probabilité P , et notée $g(P)$ ou plus simplement g . Dans le cas de l'estimation paramétrique la mesure de probabilité dépend de la valeur θ du paramètre. g est alors une fonction de θ . On estime g par une fonction $f(x)$ de l'observation telle que $f(X)$ soit une variable aléatoire. $f(X)$ est alors appelée estimateur de g .

Qualité d'un estimateur Pour évaluer la qualité d'un estimateur, on définit des fonctions de coût $C(g, f(X))$ qui quantifient la dissemblance entre g et son estimateur $f(X)$. Les fonctions de coût sont encore appelées fonction de perte.

En estimation, on cherche souvent un estimateur qui est maximum pour la relation d'ordre définie par $f_1 \geq f_2$ si

$$\forall P \in \mathcal{P}, \mathbb{E}_P[C(g, f_1(X))] \leq \mathbb{E}_P[C(g, f_2(X))], \quad (11.1)$$

où \mathbb{E}_P représente l'espérance mathématique sous la loi de probabilité P , c'est à dire que $\mathbb{E}_P[C(g, f(X))] = \int C(g, f(x))P(dx)$. $\mathbb{E}[C(g, f(X))]$ est appelé coût moyen, ou risque, de l'estimateur $f(X)$. Malheureusement, il s'agit d'une relation d'ordre partiel, et deux estimateurs ne peuvent pas toujours être comparés. On est alors souvent amené à restreindre l'ensemble des estimateurs à une sous famille de la forme $\{f(X); f \in F\}$ au sein de laquelle on pourra définir un meilleur estimateur, c'est à dire trouver un estimateur $\hat{f}(X)$ tel que

$$\forall P \in \mathcal{P}, \hat{f} = \arg \min_{f \in F} \mathbb{E}_P[C(g, f(X))]. \quad (11.2)$$

Le critère d'erreur quadratique moyenne défini par

$$\|f(X) - g\|^2 = \text{var}[f(X)] + |\mathbb{E}[f(X)] - g|^2, \quad (11.3)$$

La grandeur $\mathbb{E}[f(X)] - g$ est appelée **biais** de l'estimateur $f(X)$. $\|f(X) - g\|^2$ est une fonction de risque souvent utilisée en traitement du signal. On restreint souvent la classe des estimateurs envisagés en se limitant aux **estimateurs sans biais**, auquel cas l'existence d'un meilleur estimateur \hat{f} est assurée ([18] p.96). Le critère (11.3) présente alors l'avantage de souvent conduire à une solution f facile à calculer, et de pouvoir être interprété comme la puissance moyenne de l'erreur entre g et son estimateur.

Convergence Une suite d'estimateurs $f_n(X_1, \dots, X_n)$ qui converge en probabilité vers g quand la taille n de l'observation tend vers $+\infty$ est dit faiblement consistante, et lorsque la convergence est presque sûre, on parle d'estimateur fortement consistant. Lorsque la moyenne d'ordre p de $f_n(X_1, \dots, X_n) - g$ tend vers 0, on parle de consistance en moyenne d'ordre p . De plus, on dira qu'une suite d'estimateurs de g est asymptotiquement sans biais si la suite de leurs moyennes converge vers g .

11.2 Estimation empirique

11.2.1 Séquence d'observations indépendantes.

Pour un échantillon de taille n , $\mathbf{X}_n = [X_1, \dots, X_n]^T$, on définit la mesure de probabilité empirique et la fonction de répartition empirique par

$$\begin{aligned}\hat{F}_n &= \frac{1}{n} \sum_{i=1, n} \delta_{X_i} \\ \text{et } \hat{F}_n(y) &= \frac{1}{n} \sum_{i=1, n} \mathbb{I}_{X_i \leq y}.\end{aligned}\tag{11.4}$$

L'estimateur empirique d'une grandeur $g = f(F)$, où F est la fonction de répartition associée à P est défini par $f(\hat{F}_n)$. L'utilisation de cet estimateur est justifiée, pour de nombreuses fonctions f , par les résultats suivants (voir par exemple [16], chapitre 1)

Théorème 9 *En tout point y , $\lim_{n \rightarrow \infty} \hat{F}_n(y) \stackrel{\text{p.s.}}{=} F_X(y)$, où $F_X(y)$ est la fonction de répartition de X .*

Démonstration $\mathbb{E}[\mathbb{I}_{X \leq y}] = P(X \leq y) = F_X(y)$, et la loi forte des grands nombres permet de conclure que $\lim_{n \rightarrow \infty} \hat{F}_n(y) \stackrel{\text{p.s.}}{=} F_X(y)$. \square

Théorème 10 *Soit une fonction G de la forme*

$$G(F) = h\left(\int_{\mathbb{R}} g(x) dF(x)\right),\tag{11.5}$$

où h est continue au point $\int_{\mathbb{R}} g(x) dF_X(x)$. Alors, $\lim_{n \rightarrow \infty} G(\hat{F}_n) \stackrel{\text{p.s.}}{=} G(F_X)$.

Ce résultat assure en particulier la convergence presque sûre des moments empiriques : en notant $\hat{m}_{X^k, n} = n^{-1} \sum_{i=1, n} X_i^k$,

$$\begin{aligned}\lim_{n \rightarrow \infty} \hat{m}_{X^k, n} &\stackrel{\text{p.s.}}{=} \mathbb{E}[X^k] \\ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1, n} (X_i - \hat{m}_{X, n})^k &\stackrel{\text{p.s.}}{=} \mathbb{E}[(X - \mathbb{E}[X])^k].\end{aligned}\tag{11.6}$$

Malheureusement, lorsque les X_i sont les valeurs d'un processus aléatoire, on ne peut plus appliquer de façon directe les résultats précédents car, sauf cas particuliers, les X_i ne sont pas indépendants.

11.2.2 Processus ergodique.

Rappelons que pour un processus X , strictement stationnaire et strictement ergodique,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1, n} f(X_{i+1}, \dots, X_{i+k}) \stackrel{\text{p.s.}}{=} \mathbb{E}[f(X_1, \dots, X_k)].\tag{11.7}$$

En particulier, on a

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{m}_{X,n} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1,n} X_i \\ &\stackrel{\text{p.s.}}{=} m_X \\ \text{et } \lim_{n \rightarrow \infty} \hat{R}_{X,n}(k) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1,n} X_i X_{i+k} - \hat{m}_{X,n}^2 \\ &\stackrel{\text{p.s.}}{=} R_X(k). \end{aligned} \tag{11.8}$$

Lorsque le processus n'est stationnaire et ergodique qu'au second ordre, les propriétés (11.8) sont encore vérifiées si on remplace la convergence presque sûre par la convergence en moyenne quadratique.

Dans le chapitre suivant, on verra des conditions qui permettent de garantir la convergence presque sûre des fonctions de moyenne et de covariance, pour une classe importante de processus, celle des processus linéaires.

11.3 Notions d'estimation paramétrique

11.3.1 Borne de Cramer-Rao

Supposons que les lois P_θ du modèle statistique $(\Omega, \mathcal{A}, \{P_\theta; \theta \in \Theta\})$ étudié soient définies par des densités de probabilité, notées p_θ , telles que $p_\theta(x)$ soit deux fois dérivable par rapport à θ , presque partout en x . Supposons de plus que

1. Θ est un ouvert de \mathbb{R}^p ,
2. pour tout élément A de \mathcal{A} , $P_\theta(A) = \int_A p_\theta(x) dx$ est deux fois dérivable en θ ,
3. les opérations de dérivation et d'intégration peuvent être permutées.

Alors, pour tout estimateur sans biais $T(\mathbf{X})$ ($\mathbf{X} = [X_1, \dots, X_n]^T$) d'une grandeur $g(\theta)$, la matrice de covariance Σ_T de $T(\mathbf{X})$ est inférieurement bornée par la borne dite de Cramer-Rao (voir par exemple [18] p.98) :

$$\Sigma_T \geq \frac{\partial g(\theta)}{\partial \theta} I_\theta^{-1} \frac{\partial g(\theta)^H}{\partial \theta}, \tag{11.9}$$

où I_θ représente la matrice d'information de Fisher. Le terme général de I_θ est donné par

$$\begin{aligned} [I_\theta]_{i,j} &= \mathbb{E}\left[\left(\frac{\partial}{\partial \theta_i} \log p_\theta\right)\left(\frac{\partial}{\partial \theta_j} \log p_\theta\right)\right] \\ &= -\mathbb{E}\left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log p_\theta\right], \end{aligned} \tag{11.10}$$

et le terme général de la matrice $\frac{\partial}{\partial \theta} g(\theta)$ est donné par

$$\left[\frac{\partial g(\theta)}{\partial \theta}\right]_{i,j} = \frac{\partial g_i(\theta)}{\partial \theta_j}. \tag{11.11}$$

Un estimateur sans biais dont la matrice de covariance atteint la borne de Cramer-Rao est dit **efficace**. Cette borne est un majorant de la variance pour un estimateur sans biais quelconque, mais elle ne peut pas toujours être atteinte.

11.3.2 Estimateur du maximum de vraisemblance

Pour le modèle statistique $(\Omega, \mathcal{A}, \{P_\theta; \theta \in \Theta\})$, l'estimateur du maximum de vraisemblance de θ est l'estimateur de θ qui maximise $P_\theta(\mathbf{X})$.

L'estimation au sens du maximum de vraisemblance est une technique très appréciée en statistiques, quand elle peut être mise en oeuvre. Ceci tient au fait que si pour le modèle statistique $(\Omega, \mathcal{A}, \{P_\theta; \theta \in \Theta\})$ il existe un estimateur efficace de θ , alors il concide avec l'estimateur du maximum de vraisemblance de θ ([18] p.110). Par ailleurs, on montre que l'estimateur du maximum de vraisemblance d'une grandeur $g(\theta)$ est $g(\hat{\theta})$, où $\hat{\theta}$ est l'estimateur du maximum de vraisemblance de θ ([18] p.105).

11.3.3 Estimateur linéaire sans biais de variance minimale

En l'absence de modèle statistique précis, on suppose parfois que le vecteur d'observation \mathbf{X} suit un modèle linéaire de la forme $\mathbf{X} = H\theta + W$, où W est un vecteur aléatoire de moyenne nulle et de matrice de covariance C supposée connue. On cherche un estimateur linéaire sans biais de θ sous la forme $\hat{\theta} = M\mathbf{X}$. Plaçons nous ici dans le cas réel. L'estimateur linéaire sans biais de variance minimale de θ , dont on peut vérifier qu'il est obtenu lorsque M est la solution du problème contraint

$$\begin{cases} \min_M \text{Tr}(MCM^T) \\ MH = I, \end{cases} \quad (11.12)$$

est donné par

$$\hat{\theta} = (H^T C^{-1} H)^{-1} H^T C^{-1} \mathbf{X}. \quad (11.13)$$

La matrice de covariance de $\hat{\theta}$ est alors égale à $(H^T C^{-1} H)^{-1}$. Lorsque W est gaussien, il est clair que $\hat{\theta}$ coïncide avec l'estimateur du maximum de vraisemblance de θ pour le modèle $\mathbf{X} = H\theta + W$.

11.3.4 Méthodes de moindres carrés

Pour un modèle linéaire de la forme $\mathbf{X} = H\theta + W$, où W est un vecteur aléatoire de matrice de covariance inconnue, l'estimation au sens des moindres carrés de θ est donnée par

$$\begin{aligned} \hat{\theta} &= \arg \min_{\theta} \| \mathbf{X} - H\theta \|_K^2 \\ &= (H^T K H)^{-1} H^T K \mathbf{X}, \end{aligned} \quad (11.14)$$

où la norme $\| \cdot \|_K$ est définie par $\| u \|_K^2 = u^T K u$. Lorsque K est égal à l'inverse de la matrice de covariance C de W , on retrouve l'estimateur linéaire sans biais de variance minimale.

Si le modèle n'est pas linéaire : $\mathbf{X} = g(\theta) + W$, le critère $J(\theta) = \| \mathbf{X} - g(\theta) \|_K^2$ à minimiser n'est plus une fonction quadratique du paramètre, et on doit généralement mettre en oeuvre des algorithmes itératifs d'optimisation pour trouver l'optimum.

11.4 Estimation bayésienne

11.4.1 Généralités

De nombreuses techniques d'estimation paramétrique, et en particulier les méthodes d'estimation au sens du maximum de vraisemblance, peuvent présenter certains inconvénients.

En particulier, elles n'intègrent pas d'autres informations que les données pour effectuer l'estimation, alors qu'il pourrait parfois être intéressant, voir indispensable, d'exploiter certaines informations ou contraintes relatives au vecteur des paramètres à estimer, noté θ . La connaissance a priori de l'ordre de grandeur que doit avoir un paramètre, ou celle de contraintes qu'il doit vérifier, peut être prise en compte par la donnée d'une loi de probabilité relative au vecteur de paramètres θ , appelée loi a priori.

Par ailleurs, les méthodes de maximum de vraisemblance tendent à surparamétrer l'ordre des modèles. En effet, on conçoit bien qu'en augmentant le nombre de paramètres d'un modèle, on arrive généralement à un nouveau modèle, pour lequel la valeur du maximum de la vraisemblance est plus grande. Comme la méthode du maximum de vraisemblance ne permet pas de définir une taille raisonnable pour le modèle, on utilise alors parfois des termes de pénalisation de la vraisemblance. L'introduction d'un a priori sur θ peut constituer une façon élégante de résoudre ce problème.

L'approche bayésienne consiste donc à voir les paramètres recherchés comme des variables aléatoires pour lesquelles on se donne des lois de probabilité dites lois a priori. L'utilisation de telles lois permet d'incorporer les informations a priori relatives aux paramètres et de gérer efficacement les problèmes de taille des modèles. Cette approche est relativement souple dans la mesure où il est toujours possible d'utiliser des lois a priori peu informatives (telles que des lois uniformes), ou paramétrées, lorsque les informations disponibles sur le vecteur de paramètres θ sont trop limitées. Les paramètres qui interviennent éventuellement dans la description des lois a priori sont appelés hyperparamètres.

Afin de simplifier les notations, θ désignera indifféremment un paramètre aléatoire vectoriel ou une réalisation de cette variable aléatoire.

11.4.2 Coût bayésien et estimateurs bayésiens

Dans ce qui suit, on supposera, pour simplifier les notations, et sans perte de généralité, que les lois de probabilité mises en oeuvre sont absolument continues (par rapport à la mesure de Lebesgue). La densité de la loi a priori de θ sera notée $\pi(\theta)$, celle de la loi des observations paramétrée par θ $f(x|\theta)$, et celle de la loi des variables observées pour la loi a priori $\pi(\theta)$, $f_\pi(x)$. On notera $f(\theta|x)$ la densité de la loi de θ conditionnellement à la valeur x prise par le vecteur X des variables aléatoires observées, appelée loi a posteriori. E et Θ désigneront les espaces dans lesquels X et θ prennent leurs valeurs. Notons que

$$f_\pi(x) = \int_{\Theta} f(x|\theta)\pi(\theta)d\theta, \quad (11.15)$$

et que d'après la règle de Bayes,

$$\begin{aligned} f(\theta|x) &= \frac{f(x|\theta)\pi(\theta)}{f_\pi(x)} \\ &= \frac{f(x|\theta)\pi(\theta)}{\int_{\Theta} f(x|\theta)\pi(\theta)d\theta}. \end{aligned} \quad (11.16)$$

En estimation non bayésienne, on estime généralement θ en minimisant une certaine fonction de coût moyen de la forme $\mathbb{E}[C(\theta, T(X))]$, où θ est un vecteur déterministe de paramètres inconnus. En estimation bayésienne, θ est un vecteur aléatoire, et la fonction de coût prend la forme

$$\begin{aligned} \mathbb{E}[C(\theta, T(X))] &= \mathbb{E}[\mathbb{E}[C(\theta, T(X))|\theta]] \\ &= \int_{\Theta} \mathbb{E}[C(\theta, T(X))|\theta]\pi(\theta)d\theta. \end{aligned} \quad (11.17)$$

Pour une fonction de coût quadratique, on obtient l'estimateur $\mathbb{E}[\theta|X]$. Sa valeur $\mathbb{E}[\theta|X = x]$ prise pour $X = x$, sera notée simplement $\mathbb{E}[\theta|x]$. Notons le cas important suivant

Théorème 11 *Pour des modèles exponentiels, c'est à dire des lois dont la densité $f(x|\theta)$ est de la forme $h(x)\exp(\theta^T x - \phi(\theta))$,*

$$\mathbb{E}[\theta|x] = \nabla_x \log f_\pi(x) - \nabla_x \log h(x). \quad (11.18)$$

Démonstration

$$\begin{aligned} \mathbb{E}[\theta_i|x] &= \frac{1}{f_\pi(x)} \int_{\Theta} \theta_i h(x) e^{\theta^T x - \phi(\theta)} \pi(\theta) d\theta \\ &= \frac{1}{f_\pi(x)} \frac{\partial}{\partial x_i} \int_{\Theta} h(x) e^{\theta^T x - \phi(\theta)} \pi(\theta) d\theta - \frac{1}{h(x)} \frac{\partial h(x)}{\partial x_i} \\ &= \frac{\partial}{\partial x_i} [\log f_\pi(x) - \log h(x)]. \end{aligned} \quad (11.19)$$

□

L'estimateur du maximum a posteriori, est obtenu en considérant les fonctions de coût égales à un, partout sauf sur un petit voisinage des valeurs de θ où $f(\theta|x)$ est maximale et où la fonction de coût vaut zéro. On obtient l'estimateur du maximum a posteriori en faisant tendre vers 0 la taille de ce voisinage.

Pour ce qui concerne le choix de la loi a priori, une technique souvent envisagée consiste à utiliser des familles de lois dites conjuguées, c'est à dire des familles de lois paramétrées, telles que si la loi a priori appartient à cette famille, alors il en est de même pour la loi a posteriori. Une autre technique consiste à prendre en compte certaines propriétés d'invariance qu'il peut être naturel d'envisager pour la loi de θ , telles que l'invariance par translation ou par changement d'échelle de la densité a priori de θ .

Les choix précédents de lois a priori ne sont pas toujours possibles ou pas toujours justifiés. Même lorsqu'ils le sont, il apparaît bien souvent que le calcul d'expressions intégrales de la forme $\int_{\Theta} \mathbb{E}[C(\theta, T(X))|\theta] \pi(\theta) d\theta$, ou la maximisation de la vraisemblance a posteriori $f(\theta|x)$ ne peuvent pas être effectuées de façon analytique. Les méthodes de simulation de Monte Carlo permettent alors de résoudre numériquement les problèmes envisagés.

Exercices

11.1 Vérifiez les formules (11.13) et (11.14).

11.2 On observe un échantillon $\mathbf{x} = [x_1, \dots, x_n]$ de réalisations indépendantes d'une variable aléatoire X de loi exponentielle de densité $p_{\theta}(x) = \theta \exp(-\theta x)$.

a) Donnez l'estimateur du maximum de vraisemblance de θ . Calculez le biais et la variance de cet estimateur

b) On suppose maintenant que le paramètre θ est aléatoire de loi a priori de densité $\pi(\theta) = \lambda \exp(-\lambda\theta)$, où λ est un paramètre connu. Calculez l'estimateur du maximum de vraisemblance a posteriori de λ défini par $\arg \max_{\lambda} p(\lambda|\mathbf{x})$.

11.3 On observe un échantillon $[x_1, \dots, x_n]$ de réalisations indépendantes d'une variable aléatoire X de loi gaussienne de densité $p(x) = (\sqrt{2\pi}\sigma)^{-1} \exp(-(x-m)^2/(2\sigma^2))$, où m et σ^2 sont des paramètres inconnus. Calculez l'estimateur du maximum de vraisemblance de ces paramètres.

Chapitre 12

Estimation spectrale non paramétrique

Objectif En pratique, les coefficients de covariance d'un processus ne sont pas connus de façon exacte mais seulement estimés. On envisage ici le problème de l'estimation non paramétrique de ces coefficients, ainsi que celui de l'estimation spectrale non paramétrique par la méthode du périodogramme.

12.1 Estimateurs empiriques de la moyenne et des covariances

Comme en pratique on est en général amené à travailler à partir d'une version échantillonnée d'une réalisation du processus étudié, on se placera dans le cas discret. Il est clair que l'échantillonnage à intervalles de temps réguliers d'un processus stationnaire (respectivement ergodique) fournit un processus discret qui est encore stationnaire (respectivement ergodique). On va maintenant indiquer quelques résultats concernant les estimateurs empiriques des fonctions de moyenne et de covariance des processus discrets stationnaires au second ordre.

Bien entendu, on ne peut pas ici appliquer de façon directe la loi des grands nombres ou le théorème de la limite centrale du fait que les variables mises en jeu ne sont pas indépendantes. Aussi, on va introduire une classe assez générale de processus, celle des processus linéaires. Pour les processus linéaires, on verra que l'estimation empirique de la moyenne et des covariances vérifient encore la loi des grands nombres ou le théorème de la limite centrale.

12.1.1 Processus linéaires, et absolue sommabilité des covariances

Dans ce qui suit, on sera souvent amené à considérer des processus linéaires pour lesquels il est assez simple d'établir un certain nombre de résultats.

Définition Un processus X défini par

$$X_n = m_X + \sum_{k \in \mathbb{Z}} h_k V_{n-k}, \quad (12.1)$$

où V est un bruit blanc centré de variance σ^2 finie et $\sum_{k \in \mathbb{Z}} |h_k| < \infty$, est appelé processus linéaire.

La classe des processus linéaires contient notamment les processus ARMA.

Pour un processus linéaire X , le théorème de convergence dominée de Lebesgue conduit aux relations

$$\begin{aligned} R_X(k) &= \mathbb{E}[X_{n+k} X_n^*] \\ &= \sigma^2 \sum_{l \in \mathbb{Z}} h_{k+l} h_l^* \\ \sum_{k \in \mathbb{Z}} |R_X(k)| &\leq \sigma^2 (\sum_{k \in \mathbb{Z}} |h_k|)^2 \\ \sum_{k \in \mathbb{Z}} R_X(k) &= \sigma^2 |\sum_{k \in \mathbb{Z}} h_k|^2. \end{aligned} \quad (12.2)$$

Notons que les sommes précédentes sont finies. De façon générale, l'absolue sommabilité des covariances entraîne la continuité de la DSP :

Proposition 14 *Pour un processus X stationnaire au second ordre l'hypothèse*

$$\sum_{k \in \mathbb{Z}} |R_X(k)| < \infty \quad (12.3)$$

entraîne la continuité de $S_X(f)$.

Notons que réciproquement l'hypothèse $\sum_{k \in \mathbb{Z}} |R_X(k)| < \infty$ souvent faite dans ce document et vérifiée pour les processus linéaires l'est également pour les processus dont la mesure spectrale est à densité et dont la DSP est continue, dérivable par morceaux, et de dérivée bornée.

Rappelons enfin que pour un processus X dont la DSP est de classe \mathcal{C}^1 au voisinage de f ,

$$\sum_{k \in \mathbb{Z}} e^{-2i\pi kf} R_X(k) = S_X(f), \quad (12.4)$$

d'après le critère de Dirichlet (voir [?] p.539 et l'annexe).

12.1.2 Moyenne empirique

Soit $X = (X_n)_{n \in \mathbb{Z}}$, un processus stationnaire au second ordre. Sa moyenne empirique définie par $\hat{m}_{X,n} = n^{-1} \sum_{k=1,n} X_k$ est un estimateur sans biais :

$$\begin{aligned} \mathbb{E}[\hat{m}_{X,n}] &= \mathbb{E}[X_n] \\ &= m_X. \end{aligned} \quad (12.5)$$

On vérifie que la variance de $\hat{m}_{X,n}$ est donnée par

$$\| \hat{m}_{X,n} - m_X \|^2 = \frac{1}{n} \sum_{|j| < n} \left(1 - \frac{|j|}{n}\right) R_X(j). \quad (12.6)$$

Donc,

$$\| \hat{m}_{X,n} - m_X \|^2 \leq \frac{1}{n} \sum_{|j| < n} |R_X(j)| \quad (12.7)$$

et si $\lim_{n \rightarrow \infty} R_X(n) = 0$, $\hat{m}_{X,n}$ converge en moyenne quadratique vers m_X . Cette condition est en fait également nécessaire.

Si de plus $\sum_{n \in \mathbb{Z}} |R_X(j)| < \infty$, il vient, d'après (12.6) et le théorème de la convergence dominée de Lebesgue que

$$\lim_{n \rightarrow \infty} n \| \hat{m}_{X,n} - m_X \|^2 = \sum_{j \in \mathbb{Z}} R_X(j). \quad (12.8)$$

En outre, si $S_X(f)$ est de classe \mathcal{C}^1 au voisinage de 0, $\sum_{n \in \mathbb{Z}} R_X(j) = S_X(0)$ d'après le critère de Dirichlet. Le fait qu'asymptotiquement la variance de $\hat{m}_{X,n}$ soit équivalente à $n^{-1} S_X(0)$ suggère alors la possible existence d'un théorème central limite pour $\hat{m}_{X,n}$. Le résultat suivant formalise cette idée

Théorème 12 *Si X est un processus linéaire de variance σ^2 , avec*

$$X_n = m_X + \sum_{k \in \mathbb{Z}} h_k V_{n-k}, \quad (12.9)$$

et $\sum_{k \in \mathbb{Z}} h_k \neq 0$, alors

$$\lim_{n \rightarrow \infty} \sqrt{n} [\hat{m}_{X,n} - m_X] \sim \mathcal{N}(0, \sigma^2 | \sum_{k \in \mathbb{Z}} h_k|^2). \quad (12.10)$$

Notons que le théorème central limite assure la convergence presque sûre de la suite des variables vers la valeur moyenne de la limite.

12.1.3 Coefficients de covariance empiriques

On utilise généralement les estimateurs empiriques suivants des covariances, définis pour $k \geq 0$ par

$$\tilde{R}_{X,n}(k) = \frac{1}{n-k} \sum_{j=1, n-k} (X_{j+k} - \hat{m}_{X,n})(X_j - \hat{m}_{X,n})^*, \quad (12.11)$$

$$\text{et } \hat{R}_{X,n}(k) = \frac{1}{n} \sum_{j=1, n-k} (X_{j+k} - \hat{m}_{X,n})(X_j - \hat{m}_{X,n})^*,$$

et $\tilde{R}_{X,n}(k) = \tilde{R}_{X,n}^*(-k)$, et $\hat{R}_{X,n}(k) = \hat{R}_{X,n}^*(-k)$, pour $k < 0$.

Ces estimateurs sont biaisés, et

$$\begin{aligned}
\mathbb{E}[\hat{R}_{X,n}(k)] &= \frac{n-k}{n} \mathbb{E}[\tilde{R}_{X,n}(k)] \\
&= \frac{1}{n} \sum_{j=1,n} \mathbb{E}[(X_{j+k} - \hat{m}_{X,n})(X_j - \hat{m}_{X,n})^*] + O\left(\frac{k}{n}\right) \\
&= \frac{1}{n} \sum_{j=1,n} \mathbb{E}[(X_{j+k} - m_X)(X_j - m_X)^*] - \|\hat{m}_{X,n} - m_X\|^2 + O\left(\frac{k}{n}\right) \\
&= R_X(k) - \frac{1}{n} \sum_{j=-n,n} \left(1 - \frac{|j|}{n}\right) R_X(j) + O\left(\frac{k}{n}\right).
\end{aligned} \tag{12.12}$$

Il est clair que si $\lim_{n \rightarrow \infty} R_X(n) = 0$, les estimateurs $\tilde{R}_{X,n}(k)$ et $\hat{R}_{X,n}(k)$ sont asymptotiquement sans biais.

Remarquons que si le processus X est de moyenne m_X connue, on peut remplacer $\hat{m}_{X,n}$ par m_X dans les définitions de $\tilde{R}_{X,n}(k)$ et de $\hat{R}_{X,n}(k)$. Dans ce cas, $\tilde{R}_{X,n}(k)$ devient un estimateur sans biais. Cependant, on préférera souvent utiliser l'estimateur biaisé $\hat{R}_{X,n}(k)$ car à la différence de $(\tilde{R}_{X,n}(k))_{k=0,n-1}$ la séquence $(\hat{R}_{X,n}(k))_{k=0,n-1}$ est de type positif, et donc il existe une mesure spectrale dont les coefficients $(\hat{R}_{X,n}(k))_{k=0,n-1}$ sont les premiers coefficients de Fourier. Aussi, dans la suite, on travaillera toujours avec l'estimateur $\hat{R}_{X,n}(k)$. Afin de simplifier l'exposé, on supposera que le processus X considéré est de moyenne nulle.

12.2 Estimation empirique de la DSP : le périodogramme

L'estimateur du périodogramme (encore appelé spectrogramme) de la DSP d'un processus X supposé centré et stationnaire au second ordre est défini par

$$\hat{S}_{X,n}(f) = \frac{1}{n} \left| \sum_{k=1,n} X_k e^{-2i\pi k f} \right|^2. \tag{12.13}$$

On vérifie aisément que $\hat{S}_{X,n}(f)$ est défini de façon équivalente par

$$\hat{S}_{X,n}(f) = \sum_{k=-n+1}^{n-1} \hat{R}_{X,n}(k) e^{-2i\pi k f}. \tag{12.14}$$

Cette expression est à rapprocher de la relation $S_X(f) = \sum_{k \in \mathbb{Z}} R_X(k) e^{-2i\pi k f}$, qui donne l'expression de la DSP lorsque $\sum_{k \in \mathbb{Z}} |R_X(k)| < \infty$, et qui suggère que le périodogramme est un estimateur pertinent de la DSP.

Malheureusement, on verra un peu plus loin que le périodogramme n'est pas un estimateur consistant ! On montrera cependant qu'il est possible en lissant le périodogramme, c'est à dire en lui faisant subir un filtrage, d'obtenir un estimateur consistant.

De (12.14), il vient que

$$\mathbb{E}[\hat{S}_{X,n}(f)] = \sum_{k=-n+1}^{n-1} \left(1 - \frac{|k|}{n}\right) R_X(k) e^{-2i\pi k f}, \quad (12.15)$$

et lorsque $\sum_{k \in \mathbb{Z}} |R_X(k)| < \infty$ le théorème de la convergence dominée conduit à

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}[\hat{S}_{X,n}(f)] &= \sum_{k \in \mathbb{Z}} R_X(k) e^{-2i\pi k f} \\ &= S_X(f). \end{aligned} \quad (12.16)$$

Remarquons qu'alors la convergence en moyenne est uniforme en f car

$$|\mathbb{E}[\hat{S}_{X,n}(f)] - S_X(f)| \leq \frac{1}{\sqrt{n}} \sum_{|k| \leq \sqrt{n}} |R_X(k)| + \sum_{|k| > \sqrt{n}} |R_X(k)|, \quad (12.17)$$

et les deux termes de droite tendent vers 0 quand n tend vers $+\infty$.

Nous allons maintenant indiquer les propriétés asymptotiques du périodogramme en terme de variance. Comme le cas général n'est pas simple, nous allons nous limiter à la classe importante des processus linéaires et valeurs reelles ainsi qu'aux processus harmoniques.

12.2.1 Cas des processus linéaires

On suppose maintenant que X est un processus linéaire défini par $X_n = \sum_{k \in \mathbb{Z}} h_k V_{n-k}$, où V est un bruit blanc de variance égale à un et $\sum_{k \in \mathbb{Z}} |h_k| < \infty$. La DSP de X est donc

$$\begin{aligned} S_X(f) &= \left| \sum_{k \in \mathbb{Z}} h_k e^{-2i\pi k f} \right|^2 S_V(f) \\ &= |h(e^{2i\pi f})|^2. \end{aligned} \quad (12.18)$$

Comme $\sum_{k \in \mathbb{Z}} |R_X(k)| \leq (\sum_{k \in \mathbb{Z}} |h_k|)^2 < \infty$, la relation (12.17) montre que $\mathbb{E}[S_{X,n}(f)]$ converge uniformément en f vers $S_X(f)$.

Moyennant quelques calculs supplémentaires préciser le comportement asymptotique de $\hat{S}_{X,n}(f)$:

Théorème 13 *Pour un processus linéaire X dont la densité spectrale de puissance $S_X(f)$ ne s'annule pas, $\hat{S}_{X,n}(f)$ converge en loi vers une variable aléatoire, de loi exponentielle si $0 < f < 1/2$, et du $\chi^2(1)$ pour $f \in \{0, 1/2\}$, dont la moyenne vaut $S_X(f)$.*

De plus, si $\mathbb{E}[|V_n|^4] = \nu < \infty$ ($\sigma^2 = 1$) et $\sum_{k \in \mathbb{Z}} \sqrt{|k|} |h_k| < \infty$, on obtient les variances asymptotiques suivantes :

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{var}[\hat{S}_{X,n}(f)] &= 2S_X^2(f) & \text{si } f = 0, 1/2 \\ \lim_{n \rightarrow \infty} \text{var}[\hat{S}_{X,n}(f)] &= S_X^2(f) & \text{si } f \neq 0, 1/2 \\ \lim_{n \rightarrow \infty} \text{cov}[\hat{S}_{X,n}(f_1), \hat{S}_{X,n}(f_2)] &= 0 & \text{si } 0 < f_1 < f_2 < 1/2. \end{aligned} \quad (12.19)$$

12.2.2 Cas des processus harmoniques

Pour compléter l'étude du périodogramme, envisageons le cas d'un processus X défini par $X_n = \sum_{k=1,p} \xi_k e^{2i\pi n f_k}$. Pour cela, commençons par rappeler le résultat suivant

Proposition 15 *Le noyau de Fejer d'ordre n défini par*

$$K_n(f) = \frac{1}{n} \left(\frac{\sin(\pi n f)}{\sin \pi f} \right)^2, \quad (12.20)$$

vérifie

$$\int_{\mathcal{I}} K_n(f) df = 1, \quad (12.21)$$

$$\text{et } \forall \varepsilon \in]0, 1/2[\quad \lim_{n \rightarrow \infty} \int_{-\varepsilon}^{\varepsilon} K_n(f) df = 1.$$

En notant $[x]$ la partie entière de x , remarquons maintenant que

$$\begin{aligned} \hat{S}_{X,n}(f) &= \frac{1}{n} \left| \sum_{k=1,n} (\sum_{l=1,p} \xi_l e^{2i\pi k f_l}) e^{-2i\pi k f} \right|^2 \\ &= \sum_{l=1,p} |\xi_l|^2 K_n(f_l - f) + \\ &\quad \sum_{l \neq m} \xi_l \xi_m^* e^{i\pi [n/2](f_l - f_m)} K_{[n/2]}^{1/2}(f_l - f) K_{[n/2]}^{1/2}(f_m - f) + O(1/\sqrt{n}). \end{aligned} \quad (12.22)$$

D'après la proposition 15, il est clair que pour toute fonction g continue,

$$\lim_{n \rightarrow \infty} \int_{\mathcal{I}} g(f) \hat{S}_{X,n}(f) df = \sum_{l=1,p} |\xi_l|^2 g(f_l). \quad (12.23)$$

Il apparaît donc que $\hat{S}_{X,n}(f)$ converge (au sens des distributions) vers la mesure aléatoire $\sum_{l=1,p} |\xi_l|^2 \delta_{f_l}$. Donc,

12.3 Lissage du périodogramme

L'estimateur spectral du périodogramme n'est pas très satisfaisant. En effet, la convergence du périodogramme a lieu en moyenne mais $\hat{S}_{X,n}(f)$ ne converge pas en moyenne quadratique. De plus, la décorrélation asymptotique à des fréquences différentes entraîne une allure irrégulière du périodogramme, indépendamment de la régularité de la DSP $S_X(f)$.

12.3.1 Fonctions intégrales du périodogramme

Cependant, lorsque le spectre de X intervient au travers de relations intégrales de la forme

$$\int_{\mathcal{I}} g(f) S_X(f) df, \quad (12.24)$$

où g est une fonction continue, l'utilisation de l'estimateur du périodogramme peut souvent être envisagée de façon satisfaisante, comme le montre le résultat suivant, valable pour un processus linéaire X défini par $X_n = \sum_{k \in \mathbb{Z}} h_k V_{n-k}$ qui satisfait

$$\mathbb{E}[|V_n|^4] < \infty, \quad (12.25)$$

$$\text{et } \sum_{k \in \mathbb{Z}} \sqrt{|k|} |h_k| < \infty.$$

Proposition 16 *Pour un processus X défini par $X_n = \sum_{k \in \mathbb{Z}} h_k V_{n-k}$, où V est un bruit blanc constitué d'une suite IID et qui vérifie (12.25), et une fonction g continue sur \mathcal{I} ,*

$$\lim_{n \rightarrow \infty} \int_{\mathcal{I}} g(f) \hat{S}_{X,n}(f) df \stackrel{m.q.}{=} \int_{\mathcal{I}} g(f) S_X(f) df. \quad (12.26)$$

12.3.2 Périodogramme lissé

La proposition précédente justifie qu'en pratique on remplace souvent l'estimateur du périodogramme par des versions lissées de celui-ci, pour lesquelles on estime le spectre à la fréquence f en moyennant les valeurs du périodogramme sur une fenêtre de fréquences centrée sur f . Notons $W(f)$ la fonction de pondération utilisée. En prolongeant $W(f)$ hors de \mathcal{I} de façon périodique, le nouvel estimateur spectral est donné par

$$\hat{S}_{X,n}^W(f) = \int_{\mathcal{I}} W(f-u) \hat{S}_{X,n}(u) du. \quad (12.27)$$

Dans le cas où X est un bruit blanc, pour que $\hat{S}_{X,n}^W(f)$ soit un estimateur asymptotiquement sans biais de $S_X(f)$ il faut clairement que $\int_{\mathcal{I}} W(u) du = 1$. De plus, pour assurer la positivité de $\hat{S}_{X,n}^W(f)$, et le fait que ce soit une fonction paire lorsque X est un processus à valeurs réelles, il est naturel de prendre pour W une fonction paire positive. Lorsque n tend vers l'infini, on pourra diminuer la taille du support de la fonction W , alors notée W_n , afin d'obtenir une suite d'estimateur $\hat{S}_{X,n}^{W_n}(f)$ de $S_X(f)$ asymptotiquement sans biais. On vient donc de formuler les hypothèses suivantes

$$\begin{aligned} W_n(f) &\geq 0 \\ W_n(-f) &= W_n(f) \\ \int_{\mathcal{I}} W_n(f) df &= 1 \end{aligned} \quad (12.28)$$

le support $[-f_n, f_n]$ de W_n vérifie $\lim_{n \rightarrow \infty} f_n = 0$.

On peut montrer que si $S_X(f)$ est continue alors $\hat{S}_{X,n}^{W_n}(f)$ est asymptotiquement sans biais.

En pratique, pour construire une approximation asymptotiquement sans biais et de variance asymptotique nulle du périodogramme, on peut commencer par effectuer la FFT de (X_1, \dots, X_n) . En prenant le module au carré de la FFT on obtient l'estimateur spectral suivant :

$$\tilde{S}_{X,n}(f) = \hat{S}_{X,n}(f_{n,k}), \quad (12.29)$$

pour $f \in \mathcal{I}_k$, avec $f_{n,k} = k/n$ et

$$\begin{cases} \mathcal{I}_k = [f_{n,k}, f_{n,k} + 1/n[, & \text{et } k = -n/2, n/2 - 1 & \text{si } n \text{ est pair} \\ \mathcal{I}_k = [f_{n,k} - 1/(2n), f_{n,k} + 1/(2n)[, & \text{et } k = -(n-1)/2, (n-1)/2 & \text{si } n \text{ est impair.} \end{cases} \quad (12.30)$$

En considérant le lissage défini par

$$W_n = \sum_{|k| \leq m_n} W_{n,k} \delta_{f_{n,k}}, \quad (12.31)$$

la version lissée de $\tilde{S}_{X,n}(f)$ est donnée par

$$\begin{aligned} \tilde{S}_{X,n}^{W_n}(f) &= \sum_{|l| \leq m_n} \tilde{S}_{X,n}(f - f_{n,l}) W_{n,l} \\ &= \sum_{|l| \leq m_n} \hat{S}_{X,n}(f_{n,k} - f_{n,l}) W_{n,l}, \end{aligned} \quad (12.32)$$

pour $f \in \mathcal{I}_k$. Les conditions (12.28) se traduisent ici par

$$\begin{aligned} W_{n,k} &\geq 0 \\ W_{n,-k} &= W_{n,k} \\ \sum_{|k| \leq m_n} W_{n,k} &= 1 \\ \lim_{n \rightarrow \infty} m_n/n &= 0. \end{aligned} \quad (12.33)$$

En imposant de plus que $\sum_{|k| \leq m_n} W_{n,k}^2 = 0$, on peut montrer que $\tilde{S}_{X,n}$ est asymptotiquement sans biais et de variance asymptotique nulle.

Il existe de nombreux choix de $W_{n,k}$ qui permettent d'assurer simultanément la décroissance vers 0 du biais et de la variance quand $n \rightarrow \infty$. Par exemple, on peut prendre $W_{n,k} = m_{n,k}^{-1}$, avec $m_{n,k} = \sqrt{n}$.

12.3.3 Périodogramme moyenné

Une autre approche, souvent utilisée de façon conjointe à la précédente, consiste à d'obtenir une réduction de variance du périodogramme en moyennant plusieurs réalisations de celui-ci. En particulier, dans le cas de réalisations indépendantes de l'estimateur du périodogramme la variance est réduite d'un facteur K^{-1} , où K est le nombre de réalisations du périodogramme qui sont moyennées.

Exercices

Ici, $X = (X_n)_{n \in \mathbb{Z}}$ représente un processus SSL.

12.1 Si $X = (X_n)_{n \in \mathbb{Z}}$ est un processus AR(1), avec $X_n = aX_{n-1} + V_n$ et $|a| < 1$, calculez la variance de l'estimateur empirique de la moyenne $N^{-1} \sum_{n=1, N} X_n$ et comparez le à la variance de l'estimateur empirique de la moyenne d'un bruit blanc.

12.2 Reprenez l'exercice précédent en prenant pour X un processus MA(1).

a) Calculez $\mathbb{E}[\hat{R}_{Y,T}(t)]$.

b) Si X est gaussien, montrez que

$$\begin{aligned} \lim_{T \rightarrow \infty} T \cdot \text{cov}(\hat{R}_{Y,T}(t), \hat{R}_{Y,T}(t')) \\ = \frac{1}{T} \int_{[0, T-t]} (e^{2i\pi f(t-t')} + e^{2i\pi f(t+t')}) S_Y(f) df. \end{aligned} \quad (12.34)$$

c) Comparez ce résultat avec le cas discret.

12.3 (Données manquantes) On suppose que $X = (X_n)_{n \in \mathbb{Z}}$ est un processus stationnaire au second ordre, linéaire, de fonction de covariance $R_X(n)$. On suppose que lors de l'observation de X certaines valeurs sont manquantes ou aberrantes et, dans ce cas, éliminées. On est donc conduit à un processus d'observation Y de la forme $Y_n = g_n X_n$, avec $g_n = 1$ si X_n est enregistrée et 0 sinon. On définit

$$R_g(n) = \lim_{N \rightarrow \infty} N^{-1} \sum_{k=1, N-n} g_{k+n} g_k, \quad (12.35)$$

et on supposera ici que cette limite existe effectivement. Montrer que $R_g(n)^{-1} \hat{R}_{Y,N}(n)$ est un estimateur asymptotiquement sans biais de $R_X(n)$.

Chapitre 13

Estimation spectrale paramétrique

Objectif L'emploi de l'estimateur empirique du périodogramme ou du périodogramme lissé est mal adapté à certaines situations, en particulier lorsque seul un faible nombre de coefficients de covariance du processus étudié est accessible à partir de l'observation. Dans de tels cas, on envisage plutôt l'utilisation de techniques d'estimation spectrale paramétriques, notamment en employant des modèles de spectres rationnels.

13.1 Introduction

Ainsi qu'on l'a vu, dans le cadre de l'estimation spectrale non paramétrique, l'estimation spectrale s'effectue plus ou moins directement à partir de la connaissance d'un estimateur des premiers coefficients de covariance associés au spectre.

Dans certaines applications, les données disponibles ne permettent d'estimer qu'un nombre très réduit de coefficients de covariance du processus étudié. Dans ce cas, l'estimateur du périodogramme aura une résolution très médiocre (on peut définir la résolution comme l'écart de fréquence minimal à partir duquel il est possible de discriminer deux composantes harmoniques, au sens d'un certain critère à préciser).

Pour illustrer ce problème, considérons l'exemple suivant : en traitement du signal multi-capteurs, on dispose souvent d'un nombre important de réalisations d'une séquence $\mathbf{X} = [X_1, \dots, X_n]^T$ d'un processus $X = (X_z)_{z \in \mathbb{R}}$ indicé par une variable de position spatiale z et échantillonné aux points $1, \dots, n$ au moyen de n capteurs. Les différentes réalisations sont obtenues en considérant les valeurs prises par \mathbf{X} à différents instants. Comme le nombre de capteurs est généralement faible, on cherche d'abord à estimer les n premiers coefficients de covariance de X à partir desquels on estime ensuite la mesure spectrale du processus par des techniques paramétriques qui offrent une meilleure résolution qu'une transformée de Fourier réalisée sur n points.

De plus, à la différence de l'estimation non paramétrique, l'approche paramétrique permet d'intégrer des contraintes sur les coefficients de covariance et sur le spectre, relatives à d'éventuelles informations a priori concernant le processus. Ainsi, pour certains processus,

on possède des informations spécifiques sur la partie absolument continue et sur la partie singulière du spectre, telles que la partie absolument continue est constante, la partie singulière est portée par p points, le support de la mesure spectrale est connu,...

On a déjà évoqué dans le chapitre consacré aux processus à spectre rationnel l'intérêt de la modélisation ARMA. On pourra estimer les paramètres d'un modèle ARMA en optimisant des critères statistiques classiques tels que le critère du maximum du vraisemblance.

Dans ce chapitre, on se limitera à des approches de l'estimation des modèles ARMA par la minimisation de critères d'erreur quadratique moyenne.

13.2 Estimation spectrale des modèles AR

Pour estimer les coefficients d'un modèle AR, d'ordre p

$$X_n + \sum_{k=1,p} a_k X_{n-k} = V_n \quad (13.1)$$

on peut utiliser les équations de Yule-Walker, obtenues dans le cadre de la prédiction linéaire à horizon fini (paragraphe 8.2)

$$R_X(l) + \sum_{k=1,p} a_k R_X(l-k) = 0, \quad l = 1, p. \quad (13.2)$$

Rappelons que ces équations sont obtenues en minimisant le critère d'erreur quadratique moyenne $\| X_n + \sum_{k=1,p} a_k X_{n-k} \|^2$.

Pour estimer les coefficients $(a_k)_{k=1,p}$ du modèle AR, on peut alors chercher à résoudre les équations de Yule-Walker en remplaçant les coefficients de covariance $R_X(k)$ par leurs estimateurs $\hat{R}_{X,N}(k)$. On peut démontrer que la solution alors obtenue pour les coefficients $(a_k)_{k=1,p}$ tend à correspondre à l'estimateur du maximum de vraisemblance lorsque N devient grand (voir par exemple [9] p. 178). On verra en cours de traitement de la parole qu'il existe un algorithme rapide pour résoudre les équations de Yule-Walker : l'algorithme de Levinson.

13.3 Estimation spectrale des modèles ARMA

La modélisation ARMA du spectre d'un processus consiste à identifier ce processus à la sortie d'un filtre de fonction de transfert est connu,....

$$h(z) = \frac{b(z)}{a(z)} = \frac{\sum_{l=0,q} b_l z^{-l}}{1 + \sum_{k=1,p} a_k z^{-k}}, \quad (13.3)$$

excité par un bruit blanc de variance unité. Le spectre est alors estimé par

$$\tilde{S}_X(f) = \frac{|b(e^{2i\pi f})|^2}{|a(e^{2i\pi f})|^2}. \quad (13.4)$$

13.3.1 Estimation de la partie AR

On notera ici $(X_k)_{k=1,N}$ les variables observées. On peut montrer que

$$R_X(n) + \sum_{k=1,p} a_k R_X(n-k) = 0, \quad n > q. \quad (13.5)$$

Pour $k = 0, N-1$, on estime les $R_X(k)$ par

$$\begin{aligned} \hat{R}_{X,N}(k) &= \frac{1}{N} \sum_{n=1, N-|k|} X_{n+|k|} X_n^* \\ &= \hat{R}_{X,N}^*(-k). \end{aligned} \quad (13.6)$$

En général, on n'utilise pas toutes les valeurs $\hat{R}_{X,N}(k)$ disponibles, car les variances des estimateurs de $R_X(k)$ obtenus pour k proche de N sont élevées. Notons $(\hat{R}_{X,N}(k))_{k=-M,M}$, avec $M < N$, la séquence utilisée.

Les coefficients de covariance estimés vérifient les relations

$$\hat{R}_{X,N}(n) + \sum_{k=1,p} a_k \hat{R}_{X,N}(n-k) = \varepsilon_n \quad n > q, \quad (13.7)$$

où ε_n représente une erreur liée à l'estimation des coefficients de covariance. L'estimation de $a = [a_1, \dots, a_p]^T$ au sens des moindres carrés à partir de $(\hat{R}_{X,N}(k))_{k=-M,M}$ est donc donnée par

$$\hat{a} = \arg \min_a \sum_{l=q+1,M} |\hat{R}_{X,N}(l) + \sum_{k=1,p} a_k \hat{R}_{X,N}(l-k)|^2, \quad (13.8)$$

soit, en supposant que $M - q \geq p$,

$$\hat{a} = -(\hat{R}^H \hat{R})^{-1} \hat{R}^H [\hat{R}_{X,N}(q+1), \dots, \hat{R}_{X,N}(M)]^T, \quad (13.9)$$

avec

$$\hat{R} = \begin{bmatrix} \hat{R}_{X,N}(q) & \dots & \hat{R}_{X,N}(q+1-p) \\ \vdots & \ddots & \vdots \\ \hat{R}_{X,N}(M-1) & \dots & \hat{R}_{X,N}(M-p) \end{bmatrix}. \quad (13.10)$$

13.3.2 Estimation de la partie MA

L'estimation de la partie MA est plus délicate et ne sera pas envisagée ici. Notons simplement que si l'on ne cherche pas les coefficients b_l mais seulement la DSP de X , il suffit alors de poser

$$\left| \sum_{l=0,q} b_l e^{-2i\pi l f} \right|^2 = \sum_{k=-q,q} \beta_k e^{-2i\pi k f}, \quad (13.11)$$

et de remarquer que les coefficients de Fourier de Y défini par

$$Y_n = X_n + \sum_{k=1,p} \hat{a}_k X_{n-k}. \quad (13.12)$$

vérifient $R_Y(k) = \beta_k$ ($k = -q, q$). On identifie alors les estimateurs des coefficients β_k par les relations

$$\hat{\beta}_k = \hat{R}_{Y,N}(k) = [1 \quad \hat{a}^H] \begin{bmatrix} \hat{R}_{X,N}(k) & \cdot & \hat{R}_{X,N}(k-p) \\ \cdot & \cdot & \cdot \\ \hat{R}_{X,N}(k+p) & \cdot & \hat{R}_{X,N}(k) \end{bmatrix} \begin{bmatrix} 1 \\ \hat{a} \end{bmatrix}. \quad (13.13)$$

13.4 Synthèse de filtres rationnels

Dans le même ordre d'idée que l'estimation spectrale ARMA discutée ci-dessus, où il s'agissait de trouver les coefficients d'une DSP rationnelle à partir de ses coefficients de covariance estimés, on peut chercher à estimer une fonction de transfert rationnelle qui corresponde au mieux, au sens d'un critère quadratique, à une réponse impulsionnelle fixée. Si ce problème n'entre pas directement dans le cadre de l'estimation des spectres rationnels, la démarche proposée présente cependant des analogies avec l'étude précédente.

Pour réaliser un filtre causal stable dont la réponse impulsionnelle est $(h_{d,n})_{n \in \mathbb{N}}$, on cherche à approcher ce filtre par un filtre causal stable de fonction de transfert rationnelle

$$h(z) = \frac{b(z)}{a(z)} = \frac{\sum_{l=0,q} b_l z^{-l}}{1 + \sum_{k=1,p} a_k z^{-k}}, \quad (13.14)$$

que l'on note également $h(z) = \sum_{n \geq 0} h_n z^{-n}$. On peut alors chercher les coefficients $(a_k)_{k=1,p}$ et $(b_l)_{l=0,q}$ qui minimisent le critère

$$J = \sum_{n=0,N} |h_{d,n} - h_n|^2, \quad (13.15)$$

qui prend en compte les $N+1$ premiers coefficients de la réponse impulsionnelle à identifier.

En considérant par exemple le cas simple où

$$\begin{aligned} h(z) &= b(1 + az^{-1})^{-1} \\ &= \sum_{k \in \mathbb{N}} b(-a)^k z^{-k}, \end{aligned} \quad (13.16)$$

avec $|a| < 1$, il apparaît que J est une fonction non linéaire des paramètres à estimer. Aussi, pour simplifier le traitement du problème, on utilise généralement la méthode de Prony qui consiste à minimiser la distance entre les coefficients des réponses impulsionnelles des filtres de fonctions de transfert $h(z)a(z)$ et $b(z)$.

Ainsi, en notant $a_0 = 1$, et $b_n = 0$ pour $n > q$, on est conduit à minimiser la fonctionnelle

$$J(a, b) = \sum_{n=0,N} \left| \sum_{k=0,p} a_k h_{d,n-k} - b_n \right|^2. \quad (13.17)$$

En considérant $J(a, b)$ comme une fonction de la variable b paramétrée par a , on calcule classiquement l'optimum $b = g(a)$ puis on cherche l'optimum en a de $J(a, g(a))$. Ainsi, comme

$$J(a, b) = \| H_0 [1 \ a^T]^T - b \|^2 + \| v_h + Ha \|^2, \quad (13.18)$$

avec

$$H_0 = \begin{bmatrix} h_{d,0} & 0 & \cdot & \cdot & 0 \\ h_{d,1} & h_{d,0} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{d,q} & h_{d,q-1} & \cdot & \cdot & h_{d,q-p} \end{bmatrix},$$

$$H = \begin{bmatrix} h_{d,q} & \cdot & \cdot & h_{d,q-p+1} \\ \cdot & \cdot & \cdot & \cdot \\ h_{d,N-1} & \cdot & \cdot & h_{d,N-p} \end{bmatrix},$$
(13.19)

$$\text{et } v_h = [h_{d,q+1}, \dots, h_{d,N}]^T,$$

il vient, en supposant que $N - q > p$,

$$\hat{a} = -(H^H H)^{-1} H^H v_h,$$
(13.20)

$$\text{et } \hat{b} = H_0 [1 \ \hat{a}^T]^T.$$

Exercices

Dans cette section, $X = (X_n)_{n \in \mathbb{Z}}$ représente un processus SSL.

13.1 Montrer les équations (13.5).

13.1 (Estimation de l'amplitude d'une sinusoïde bruitée) Soit X un processus SSL de la forme

$$X_n = \xi e^{2i\pi n f_0} + V_n, \quad (13.21)$$

où ξ est une variable aléatoire centrée et $V = (V_n)_{n \in \mathbb{Z}}$ un bruit blanc circulaire, décorrélé avec ξ . f_0 est supposée connue de même que les premiers coefficients de covariance de X , notés $(R_X(k))_{k=0,N}$. On cherche ici la valeur de ξ .

a) ξ est estimée comme une fonction linéaire X_n, \dots, X_{n+N} :

$$\hat{\xi} = \sum_{k=0,N} h_k^* X_{n-k}. \quad (13.22)$$

On cherche le vecteur $h = [h_1, \dots, h_N]^T$ solution du problème d'optimisation suivant :

$$\begin{cases} \min_h \|h^H \mathbf{X}_n\|^2 \\ h^H d(f_0) = 1, \end{cases} \quad (13.23)$$

où $\mathbf{X}_n = [X_n, \dots, X_{n+N}]^T$ et $d(f) = e^{2i\pi n f} [1, e^{2i\pi f}, \dots, e^{2i\pi N f}]^T$. Donnez une interprétation de ce critère. Soit $T_{X,N}$ la matrice de covariance de \mathbf{X}_n . Montrez que l'optimum de (13.23) est donné par

$$\hat{h} = \frac{T_{X,N}^{-1} d(f_0)}{d(f_0)^H T_{X,N}^{-1} d(f_0)}. \quad (13.24)$$

b) Soit $T_{V,N}$ la matrice de covariance de $\mathbf{V}_n = [V_n, \dots, V_{n+N}]^T$ et $T_{S,N} = T_{X,N} - T_{V,N}$. Lorsque h est choisi comme la solution du problème suivant

$$\max_h \frac{h^H T_{S,N} h}{h^H T_{V,N} h}, \quad (13.25)$$

que représente t'il? Montrez que l'optimum de ce problème est défini à un facteur près et donnez l'expression de sa solution. En introduisant la contrainte de normalisation supplémentaire $h^H d(f_0) = 1$, montrez que l'on obtient la solution du problème (13.23).

c) On suppose maintenant que V est gaussien. Calculez l'estimateur du maximum de vraisemblance de ξ obtenu en minimisant la densité de probabilité de \mathbf{X}_n conditionnellement à ξ . Quelle remarque peut-on faire?

13.2 (Estimateur spectral de Capon) Soit X un processus SSL. Pour trouver un estimateur du spectre de X pour la fréquence f , on se donne le modèle suivant : $X_n = \xi_f e^{2i\pi n f} + V_n$, où $V = (V_n)_{n \in \mathbb{Z}}$ est un bruit blanc décorrélé avec ξ_f .

a) On estime ξ_f comme dans l'exercice précédant par $h_f^H \mathbf{X}_n$, avec $\mathbf{X}_n = [X_n, \dots, X_{n+N}]^T$. Montrez que l'estimateur du spectre de X , noté $S_C(f)$, appelé *estimateur spectral de Capon*, et donné par $S_C(f) = (N+1) \|h_f^H \mathbf{X}_n\|^2$, peut se réécrire

$$S_{MV}(f) = \frac{N+1}{d(f)^H \hat{T}_{X,N}^{-1} d(f)}, \quad (13.26)$$

où $[\hat{T}_{X,N}]_{ab} = \hat{R}_{X,N}(a-b)$ et $d(f) = [1, e^{2i\pi f}, \dots, e^{2i\pi(N-1)f}]^T$.

b) Calculez $S_C(f)$ lorsque X est un bruit blanc et lorsque X est de la forme $X_n = \xi e^{2i\pi n f_0} + V_n$, où V est un bruit blanc.

(13.3 (Sinusoïde bruitée)) Soit X tel que

$$X_n = A e^{2i\pi n f_0} + V_n, \quad (13.27)$$

où A et f_0 sont inconnues ($A \in \mathbb{C}$) et $V = (V_n)_{n \in \mathbb{Z}}$ est un bruit blanc gaussien circulaire.

a) Montrez que l'estimateur du maximum de vraisemblance de (A, f_0) , noté (\hat{A}, \hat{f}_0) , est donné par

$$\begin{cases} \hat{f}_0 = \arg \max_f d(f)^H \hat{T}_{X,N} d(f) \\ \hat{A} = \|d(\hat{f}_0)\|^{-2} d(\hat{f}_0)^H \mathbf{X}_N, \end{cases} \quad (13.28)$$

où $[\hat{T}_{X,N}]_{ab} = \hat{R}_{X,N}(a-b)$ et $d(f) = [1, e^{2i\pi f t}, \dots, e^{2i\pi(N-1)ft}]^T$.

b) Interprétez (\hat{A}, \hat{f}_0) en termes de transformée de Fourier et de périodogramme de X .

13.4 On suppose que X est un processus MA(q). Exprimez la log-vraisemblance de X_1, \dots, X_N en fonction des coefficients du modèle.

Annexe A

Annexe - Critère de Dirichlet

Lemme 1 (de Riemman-Lebesgue). Si $S(f)$ est une fonction continue, périodique de période 1,

$$\lim_{t \rightarrow \infty} \int_{\mathcal{I}} e^{2i\pi t f} S(f) df = 0. \quad (\text{A.1})$$

Théorème 14 Si une fonction S , périodique de période 1, admet des limites $S(f^-)$ et $S(f^+)$ à gauche et à droite au point f , et si la fonction

$$u \rightarrow u^{-1}[S(f+u) + S(f-u) - S(f^+) - S(f^-)] \quad (\text{A.2})$$

est bornée au voisinage de 0, la série de Fourier associée à S converge vers $\frac{1}{2}[S(f^+) + S(f^-)]$ au point f .

Corollaire 1 Si S , périodique de période 1, est continue, dérivable par morceaux, et de fonction dérivée bornée, la série de Fourier associée à S converge vers $S(f)$ en tout point f .

Démonstration On vérifie facilement que le lemme est vérifié pour les fonctions en escalier. Comme toute fonction continue $S(f)$ est limite uniforme d'une suite de fonctions en escalier $(\phi_n(f))_{f \geq 0}$,

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n > N, \exists T_n \in \mathbb{R}, \forall t > T_n,$$

$$\begin{aligned} \left| \int_{\mathcal{I}} e^{2i\pi t f} S(f) df \right| &\leq \left| \int_{\mathcal{I}} e^{2i\pi t f} \phi_n(f) df \right| + \int_{\mathcal{I}} |\phi_n(f) - S(f)| df \\ &\leq \varepsilon. \end{aligned} \quad (\text{A.3})$$

Pour montrer le théorème, remarquons par

$$\begin{aligned}
\sum_{k=-n,n} R(k)e^{-2i\pi kf} &= \int_{\mathcal{I}} \sum_{k=-n,n} e^{2i\pi k(u-f)} S(u) du \\
&= \int_{\mathcal{I}} \frac{\sin(\pi(2n+1)(u-f))}{\sin(\pi(u-f))} S(u) du \\
&= \int_{\mathcal{I}} \frac{\sin(\pi(2n+1)u)}{\sin(\pi u)} S(u+f) du \\
&= \frac{1}{2} \int_{\mathcal{I}} \frac{\sin(\pi(2n+1)u)}{\sin(\pi u)} [S(f+u) + S(f-u)] du.
\end{aligned} \tag{A.4}$$

Pour $S(f) = 1$ sur \mathcal{I} , $R(k) = \delta_{0,k}$, donc $\int_{\mathcal{I}} \frac{\sin(\pi(2n+1)u)}{\sin(\pi u)} du = 1$. Ainsi,

$$\begin{aligned}
&|\sum_{k=-n,n} R(k)e^{-2i\pi kf} - \frac{1}{2}[S(f^+) + S(f^-)]| \\
&\leq \frac{1}{2} \left| \int_{\mathcal{I}} \left(u \frac{\sin(\pi(2n+1)u)}{\sin(\pi u)} \right) \left(\frac{1}{u} [S(f+u) + S(f-u) - S(f^+) - S(f^-)] \right) du \right| \\
&\leq C \left| \int_{\mathcal{I}} \frac{u}{\sin(\pi u)} \sin(\pi(2n+1)u) du \right|,
\end{aligned} \tag{A.5}$$

où C est une constante. Comme $\frac{u}{\sin(\pi u)}$ est continue sur $\mathcal{I} - \{0\}$, et se prolonge par continuité sur \mathcal{I} , le lemme de Riemman-Lebesgue permet de conclure.

Le corrolaire est une conséquence directe du théorème. \square

Remarque L'hypothèse $\sum_{k \in \mathbb{Z}} |R_X(k)| < \infty$ souvent faite dans ce document se trouve également vérifiée pour les processus dont la mesure spectrale est absolument continue (par rapport à la mesure de Lebesgue) et dont la DSP est continue et dérivable par morceaux, et de dérivée bornée.

En effet, pour un tel processus X , de DSP $S_X(f)$, il vient, en effectuant une intégration par partie

$$\begin{aligned}
R_X(n) &= \int_{\mathcal{I}} e^{2i\pi n f} S_X(f) df = [S_X(f) \frac{e^{2i\pi n f}}{2i\pi n}]_{-1/2}^{1/2} - \int_{\mathcal{I}} \frac{e^{2i\pi n f}}{2i\pi n} S'_X(f) df \\
&= - \int_{\mathcal{I}} \frac{e^{2i\pi n f}}{2i\pi n} S'_X(f) df,
\end{aligned} \tag{A.6}$$

avec $S'_X(f) = \frac{d}{df} S_X(f)$. En notant c_n les coefficients de Fourier de $S'_X(f)$, il apparaît que $R_X(n) = -(2i\pi n)^{-1} c_n$. Donc

$$\begin{aligned}
(|c_n| - n^{-1})^2 &= |c_n|^2 - 4\pi |R_X(n)| + n^{-2} \\
&\geq 0,
\end{aligned} \tag{A.7}$$

et d'après l'inégalité de Parseval,

$$\begin{aligned} \sum_{n \in \mathbb{Z}} |R_X(n)| &\leq R_X(0) + \frac{1}{4\pi} \sum_{n \in \mathbb{Z}^*} (|c_n|^2 + n^{-2}) \\ &\leq R_X(0) + \frac{1}{4\pi} \sum_{n \in \mathbb{N}^*} n^{-2} + \frac{1}{4\pi} \int_{\mathcal{I}} |S'_X(f)|^2 df, \end{aligned} \tag{A.8}$$

ce qui justifie la remarque.

Classement bibliographique.

- Probabilités et processus : [1]-[5]
- Processus stationnaires : [6]-[15]
- Statistiques : [16]-[18]
- Analyse de Fourier : [19]-[20]

Bibliographie

- [1] P.B. BILLINGSLEY, *Probability and Measure, Wiley Series in Probability and Mathematical Statistics*, 1979.
- [2] N. BOULEAU, *Processus Stochastiques et Applications*, Hermann, 1988.
- [3] P. DOUKHAN, *Mixing, Properties and Examples*, Lecture Notes in Statistics, Springer-Verlag, 1994.
- [4] M. METIVIER, *Notions Fondamentales de la Théorie des Probabilités*, Dunod, Paris, 1972.
- [5] A. PAPOULIS, *Probability, Random Variables, and Stochastic Processes*, Mac Graw Hill, 1991.
- [6] R.AZENCOTT, D. DACUNHA-CASTELLE, *Séries d'Observations Irrégulières, Modélisation et Prévission*, Masson, 1984.
- [7] P.J. BROCKWELL, R.A. DAVIS, *Time Series : Theory and Methods*, Springer Verlag, New York, 1991.
- [8] T. CHONAVEL, *Traitement du Signal Aléatoire*, Collection Scientifique et Technique des Télécommunications, Springer, 2000.
- [9] T. CHONAVEL, *Statistical Signal Processing. Modelling and estimation*, Springer, 2002.
- [10] D. DACUNHA-CASTELLE, M.DUFLO, *Probabilités et Statistiques - Tome 1 : problèmes à temps fixe.*, Masson, 1982.
- [11] D. DACUNHA-CASTELLE, M.DUFLO, *Probabilités et Statistiques - Tome 2 : problèmes à temps mobile.*, Masson, 1983.
- [12] D. DACUNHA-CASTELLE, M.DUFLO, *Exercices de Probabilités et de Statistiques, 1- Problèmes à temps fixe.* Masson, 1990.
- [13] S.M. KAY, *Modern Spectral Estimation, Theory and Applications*, Prentice Hall, Englewood Cliffs, N.J., 1988.
- [14] B. PICINBONO, *Signaux Aléatoires - Tome 2 : fonctions aléatoires et modèles avec problèmes résolus*, Dunod, 1993.
- [15] A. PAPOULIS, *Signal Analysis*, Mac Graw Hill, 1984.
- [16] A. BOROVKOV, *Statistiques Mathématiques*, Mir, Moscou, 1987.
- [17] M.KENDALL, A.STUART, *The Advanced Theory of Statistics*, C.Griffin, 1977.
- [18] A. MONFORT, *Cours de Statistiques Mathématiques*, Economica, Paris, 1982.
- [19] H. DYM, H.P. Mc KEAN, *Fourier Series and Integrals*, Academic Press, Londres, 1972.
- [20] A.PAPOULIS, *The Fourier Integral and its Applications*, Mac Graw hill, 1962.

Notations et abréviations

$ x , M $	module d'un nombre complexe, déterminant d'une matrice
v^T, M^T	transposé d'un vecteur, ou d'une matrice
x^*, v^*, M^*	valeur conjuguée d'un scalaire, d'un vecteur, ou d'une matrice
v^H, M^H	valeur transposée et conjugué d'un vecteur ou d'une matrice
$[v]_i, [M]_{ij}$	élément d'indice i d'un vecteur, ou (i, j) d'une matrice
$\langle x, y \rangle$	produit scalaire de x et de y
$Tr(M)$	trace d'une matrice
$sign(x)$	$sign(x) = +1, -1, 0$, si x est positif, négatif, nul
$[x]$	partie entière du nombre x
$\delta_{a,b}$	symbole de Kronecker : $\delta_{a,b} = 1$ si $a = b$, et 0 sinon
δ_a	mesure de Dirac portée par a
$\nabla_x f$	gradient de la fonction f par rapport à la variable x
$diag(\lambda_1, \dots, \lambda_n)$	matrice diagonale de termes diagonaux $\lambda_1, \dots, \lambda_n$
I, I_n	matrice identité, matrice identité de taille n
$\ X\ $	norme de X (la norme choisie est définie par le contexte)
$\Re[z], \Im[z]$	partie réelle, imaginaire, de z
$a = b \bmod [p]$	a est le reste de la division de b par p
$\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$	ensembles des nombres entiers, entiers relatifs, réels, et complexes

\mathbb{D}	disque unité ouvert
\mathcal{I}	intervalle $]-1/2, 1/2]$
\mathcal{C}^k	ensemble des fonctions k fois dérivables, de dérivées continues
$\mathcal{C}^\infty(K)$	ensemble des fonctions infiniment dérivables, de support compact et contenu dans K
$\mathcal{B}(\mathbb{R}^n)$	tribu borelienne de \mathbb{R}^n
$\text{vect}\{(X_i)_{i \in I}\}$	espace vectoriel engendré par les combinaisons linéaires finies des X_i
$\overline{\text{vect}\{(X_i)_{i \in I}\}}$	prolongement de $\text{vect}\{(X_i)_{i \in I}\}$ en un espace complet
$\mathbb{1}_A$	fonction indicatrice de l'ensemble A
$\mathbb{E}[X]$	espérance mathématique de X
$\mathbb{E}[X Y]$	espérance de X conditionnellement à Y
$\text{var}[X]$	variance de X
$\text{cov}[X, Y]$	covariance de X et de Y
$X H, X Y$	projection orthogonale de X sur l'espace H , sur l'espace $\text{vect}\{Y\}$
$X \perp Y$	les variables X et Y sont orthogonales, ie $\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^*] = 0$
$[h(z)]X_n$	valeur à l'instant n de la sortie du filtre de fonction de transfert $h(z)$ excité par le processus $(X_n)_{n \in \mathbb{Z}}$
\oplus	somme orthogonale d'espaces vectoriels
$m.q., p.s.$	en moyenne quadratique, presque sûrement
$\xrightarrow{p.s.}, \xrightarrow{m.q.}, \xrightarrow{P}, \xrightarrow{L}$	convergence presque sûre, en moyenne quadratique, en probabilité, en loi
DSP	densité spectrale de puissance
suite IID	suite de variables aléatoires indépendantes et identiquement distribuées

Index

- échantillonnage, 28
- équation
 - aux différences, 62
 - d'état, 64
 - de Chapman-Kolmogorov, 42
 - de Yule-Walker, 57
- algorithme
 - de Levinson, 85
 - K-moyenne, 54
- borne de Cramer-Rao, 70
- bruit
 - blanc, 19
- bruit blanc, 37
- chaîne de Markov
 - homogène, 41
- chaînes de Markov, 41
- chaîne de Markov, 41
- communications
 - à spectre étalé, 54
- convergence
 - d'un estimateur, 68
 - en moyenne quadratique, 13
- covariances cycliques, 44
- critère
 - d'erreur quadratique moyenne, 68
 - de Dirichlet, 90
 - des moindres carrés, 71
 - du maximum de vraisemblance, 71
- démodulation, 51
- Détecteur quadratique, 52
- détecteur quadratique, 46
- Détection de maximum, 53
- DSP, 21
- effet Doppler, 53
- erreur de prédiction
 - à horizon fini, 57
- espérance conditionnelle, 55
- estimateur
 - de Capon, 89
 - biais, 68
 - consistant, 68
 - des moindres carrés, 71
 - du maximum de vraisemblance, 71
 - efficace, 71
 - linéaire sans biais, 71
- estimation
 - des DSP rationnelles, 85
 - des filtres ARMA, 85
 - des filtres MA, 87
 - empirique de la DSP, 78
 - empirique de la moyenne, 76
 - empirique des covariances, 77
 - empirique des moments, 70
 - paramétrique de la DSP, 84
- estimator
 - periodogram, 24
- filtrage
 - des processus stationnaires, 30
 - relations
 - covariances et spectres, 31
- filtre
 - adapté, 59
 - dérivateur, 33
 - de Hilbert, 49
 - passé bande, 32
- fonction
 - de coût, 68
- formule
 - des interférences, 31
- harmonic process, 19
- interpolation linéaire, 58
- lemme
 - de Riemman-Lebesgue, 90
- Limiteur saturé, 53
- limiteur saturé, 47
- méthode

- de Prony, 87
- matrice
 - companion, 64
 - de Toeplitz, 56
 - de transition, 42
- matrices de covariance
 - circulantes, 29
- matrices de covariance circulantes, 29
- matrix
 - Toeplitz matrix, 20
- mesure
 - spectrale, 22
- modèle
 - statistique, 67
- modulation
 - BLU, 49
 - en amplitude, 48
 - en phase et en quadrature, 49
- modulation de phase, 52
- mouvement brownien, 38
- moyenne quadratique
 - continuité, 14
 - convergence, 13
 - dérivation, 14
 - intégration, 15
- périodogramme, 78
 - intégration, 80
 - lissage, 81
 - moyennage, 82
 - variance, 79
- periodogram, 24
- prédiction linéaire
 - à horizon fini, 56
- probabilité
 - stationnaire, 43
- process
 - harmonic process, 19
 - with orthogonal increments, 29
- processus
 - à spectre rationnel, 62
 - de poisson, 36
 - ARMA, 63
 - bruit blanc, 37
 - circulaire, 39
 - cyclostationnaires, 38
 - du second ordre, 12
 - ergodique, 17, 69
 - gaussien, 35
 - harmonique, 80
 - linéaire, 75, 79
 - markovien, 41
 - mouvement brownien, 42
 - multivarié, 40
 - stationnaire, 17
- quantification, 53
- représentation
 - analytique, 49
 - d'état, 64
 - de Rice, 50
- spectre
 - de raies, 80
- théorème
 - de Birkhoff-Von Neumann, 18
 - de Slutsky, 19
 - de Bochner, 23
 - de Bussgang, 48
- Toeplitz matrix, 20
- variance
 - périodogramme, 79

Index

- échantillonnage, 29
- équation
 - aux différences, 65
 - d'état, 67
 - de Chapman-Kolmogorov, 44
 - de Yule-Walker, 59
- algorithme
 - de Levinson, 88
 - K-moyenne, 56
- borne de Cramer-Rao, 73
- bruit
 - blanc, 20
- bruit blanc, 39
- chaîne de Markov
 - homogène, 43
- chaînes de Markov, 43
- chaîne de Markov, 43
- communications
 - à spectre étalé, 56
- convergence
 - d'un estimateur, 71
 - en moyenne quadratique, 13
- covariances cycliques, 46
- critère
 - d'erreur quadratique moyenne, 71
 - de Dirichlet, 93
 - des moindres carrés, 74
 - du maximum de vraisemblance, 74
- démodulation, 53
- Détecteur quadratique, 54
- détecteur quadratique, 48
- Détection de maximum, 55
- DSP, 21
- effet Doppler, 55
- erreur de prédiction
 - à horizon fini, 59
- espérance conditionnelle, 57
- estimateur
 - de Capon, 92
 - biais, 71
 - consistant, 71
 - des moindres carrés, 74
 - du maximum de vraisemblance, 74
 - efficace, 74
 - linéaire sans biais, 74
- estimation
 - des DSP rationnelles, 88
 - des filtres ARMA, 88
 - des filtres MA, 90
 - empirique de la DSP, 81
 - empirique de la moyenne, 79
 - empirique des covariances, 80
 - empirique des moments, 73
 - paramétrique de la DSP, 87
- estimator
 - periodogram, 25
- filtrage
 - des processus stationnaires, 32
 - relations
 - covariances et spectres, 33
- filtre
 - adapté, 61
 - dérivateur, 35
 - de Hilbert, 51
 - passé bande, 34
- fonction
 - de coût, 71
- formule
 - des interférences, 33
- harmonic process, 20
- interpolation linéaire, 60
- lemme
 - de Riemman-Lebesgue, 93
- Limiteur saturé, 55
- limiteur saturé, 49
- méthode

- de Prony, 90
- matrice
 - companion, 67
 - de Toeplitz, 58
 - de transition, 44
- matrices de covariance
 - circulantes, 30
- matrices de covariance circulantes, 30
- matrix
 - Toeplitz matrix, 20
- mesure
 - spectrale, 23
- modèle
 - statistique, 70
- modulation
 - BLU, 51
 - en amplitude, 50
 - en phase et en quadrature, 51
- modulation de phase, 54
- mouvement brownien, 40
- moyenne quadratique
 - continuité, 14
 - convergence, 13
 - dérivation, 14
 - intégration, 15
- périodogramme, 81
 - intégration, 83
 - lissage, 84
 - moyennage, 85
 - variance, 82
- periodogram, 25
- prédiction linéaire
 - à horizon fini, 58
- probabilité
 - stationnaire, 45
- process
 - harmonic process, 20
 - with orthogonal increments, 30
- processus
 - à spectre rationnel, 65
 - de poisson, 38
 - ARMA, 66
 - bruit blanc, 39
 - circulaire, 41
 - cyclostationnaires, 40
 - du second ordre, 12
 - ergodique, 18, 72
 - gaussien, 37
 - harmonique, 83
 - linéaire, 78, 82
 - markovien, 43
 - mouvement brownien, 44
 - multivarié, 42
 - stationnaire, 17
- quantification, 55
- représentation
 - analytique, 51
 - d'état, 67
 - de Rice, 52
- spectre
 - de raies, 83
- théorème
 - de Birkhoff-Von Neumann, 18
 - de Slutsky, 19
 - de Bochner, 23
 - de Bussgang, 50
- Toeplitz matrix, 20
- variance
 - périodogramme, 82